Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE DIRECT®
ELSEVIER
Stochastic Processes and their Applications 106 (2003) 107–139
stochastic processes and their applications
www.elsevier.com/locate/spa

Coalescent processes obtained from supercritical Galton–Watson processes

Jason Schweinsberg*,[1]

Department of Mathematics, Malott Hall, Cornell University, Ithaca, NY 14853-4201, USA

Received 5 October 2002; received in revised form 1 February 2003; accepted 21 February 2003

Abstract

Consider a population model in which there are N individuals in each generation. One can obtain a coalescent tree by sampling n individuals from the current generation and following their ancestral lines backwards in time. It is well-known that under certain conditions on the joint distribution of the family sizes, one gets a limiting coalescent process as $N \to \infty$ after a suitable rescaling. Here we consider a model in which the numbers of offspring for the individuals are independent, but in each generation only N of the offspring are chosen at random for survival. We assume further that if X is the number of offspring of an individual, then $P(X \geqslant k) \sim Ck^{-a}$ for some $a > 0$ and $C > 0$. We show that, depending on the value of a, the limit may be Kingman's coalescent, in which each pair of ancestral lines merges at rate one, a coalescent with multiple collisions, or a coalescent with simultaneous multiple collisions.

MSC: primary 60J80; secondary 60J10; 60J27

Keywords: Coalescence; Galton–Watson processes; Ancestral processes; Poisson–Dirichlet distribution

1. Introduction

We consider a population model in which the number of offspring of each individual is chosen independently from some distribution on $\{0, 1, \ldots\}$ with mean greater than 1, as with supercritical Galton–Watson processes. However, we assume that only N of the offspring survive to form the next generation, so that the population size remains fixed. For a concrete example, consider annual plants. Each plant produces many seeds, but

* Tel.: +1-607-2555503.
E-mail address: jasonsch@math.cornell.edu (J. Schweinsberg).
URL: http://www.math.cornell.edu/~jasonsch/webpage.html
[1] Supported by an NSF Postdoctoral Fellowship.

0304-4149/03/$ - see front matter
doi:10.1016/S0304-4149(03)00028-0

the total population size remains roughly constant since the environment can support a limited number of plants per acre. If we sample n individuals from one generation, we can follow their ancestral lines backwards in time to obtain a coalescent tree. The goal of this paper is to determine what coalescent processes arise in the limit, after suitable rescaling, as $N \to \infty$.

We focus on the case in which the probability that an individual has k or more offspring decays like k^{-a} for some $a > 0$. We show that when $a \geqslant 2$, the limit is Kingman's coalescent, a coalescent process in which each pair of blocks merges at rate 1. When $1 \leqslant a < 2$, the limit is a coalescent with multiple collisions, in which many blocks can merge at once but no two such mergers occur at the same time. When $0 < a < 1$, the limit is a discrete-time coalescent with simultaneous multiple collisions, in which many blocks can merge at one time, and many such mergers can occur simultaneously.

The rest of this paper is organized as follows. In Section 1.1, we introduce the relevant families of coalescent processes. In Section 1.2, we describe how these coalescent processes can be obtained by taking limits of ancestral processes. In Section 1.3, we formulate our model more precisely and state the convergence results. Then in Section 2, we prove convergence to Kingman's coalescent when $a \geqslant 2$. In Section 3, we consider the case $1 \leqslant a < 2$. In Section 4, we prove the convergence result for $0 < a < 1$ and establish a formula for the transition probabilities of the limiting coalescent process.

1.1. Coalescent processes

We first review some facts about coalescent processes that we will need. For a more thorough survey of coalescence, see Aldous (1999). Let $\mathscr{P}_n$ be the set of partitions of $\{1, \ldots, n\}$, and let $\mathscr{P}_\infty$ be the set of partitions of $\mathbb{N}$. Kingman (1982a) introduced the n-coalescent, which is a $\mathscr{P}_n$-valued continuous-time Markov process $(\Pi_n(t),\ t \geqslant 0)$ such that $\Pi_n(0)$ is the partition of $\{1, \ldots, n\}$ into singletons, and then each pair of blocks merges at rate one. More precisely, for $\xi, \eta \in \mathscr{P}_n$, write $\xi \prec \eta$ if η can be obtained by merging two blocks of ξ. If ξ and η are distinct partitions of $\{1, \ldots, n\}$, then

$$\lim_{t \downarrow 0} t^{-1} P(\Pi_n(s+t) = \eta | \Pi_n(s) = \xi) = \begin{cases} 1 & \text{if } \xi \prec \eta, \\ 0 & \text{otherwise.} \end{cases}$$

Kingman (1982a) also showed that there exists a $\mathscr{P}_\infty$-valued Markov process $(\Pi_\infty(t),\ t \geqslant 0)$, which we call Kingman's coalescent, whose restriction to the first n positive integers is the n-coalescent. That is, we have

$$(R_n \Pi_\infty(t),\ t \geqslant 0) =_d (\Pi_n(t),\ t \geqslant 0),$$

where for all $\pi \in \mathscr{P}_\infty$, the partition $R_n \pi \in \mathscr{P}_n$ is defined such that two integers i and j are in the same block of $R_n \pi$ if and only if they are in the same block of π.

Pitman (1999) introduced coalescents with multiple collisions. A coalescent with multiple collisions is a $\mathscr{P}_\infty$-valued Markov process $(\Pi_\infty(t),\ t \geqslant 0)$ such that for each

n, the restriction $(R_n\Pi_\infty(t),\ t \geqslant 0)$ is a $\mathscr{P}_n$-valued Markov process with the property that whenever there are b blocks, each k-tuple of blocks is merging to form a single block at some rate $\lambda_{b,k}$, and no other transitions are possible. The rates $\lambda_{b,k}$ do not depend on n or the sizes of the blocks. Pitman showed that the transition rates satisfy

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

for some finite measure Λ on $[0,1]$. Eq. (1) sets up a one-to-one correspondence between coalescents with multiple collisions and finite measures Λ on $[0,1]$, and the process whose transition rates are given by (1) is called the Λ-coalescent. Kingman's coalescent is the special case in which Λ is the unit mass at zero. See Sagitov (1999) and Schweinsberg (2000a) for further results on coalescents with multiple collisions.

One can generalize these processes further to obtain coalescents with simultaneous multiple collisions, which were studied by Möhle and Sagitov (2001) and Schweinsberg (2000b). Suppose $b, k_1, \ldots, k_r, s$ are nonnegative integers such that $k_j \geqslant 2$ for $j = 1, \ldots, r$ and $b = s + \sum_{j=1}^r k_j$. 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 after the collision are unions of $k_1, \ldots, k_r$ blocks before the collision. A coalescent with simultaneous multiple collisions is a $\mathscr{P}_\infty$-valued Markov process $(\Pi_\infty(t),\ t \geqslant 0)$ such that for each n, the process $(R_n\Pi_\infty(t),\ t \geqslant 0)$ is a $\mathscr{P}_n$-valued Markov process with the property that when there are b blocks, each $(b; k_1, \ldots, k_r; s)$-collision is occurring at some fixed rate $\lambda_{b;k_1,\ldots,k_r;s}$. Schweinsberg (2000b) showed that there must be a finite measure Ξ on the space $\Delta = \{(x_1, x_2, \ldots)\colon x_1 \geqslant x_2 \geqslant \cdots \geqslant 0,\ \sum_{i=1}^\infty x_i \leqslant 1\}$ such that if $\Xi = \Xi_0 + a\delta_0$, where Ξ_0 has no atom at zero and δ_0 is a unit mass at zero, then $\lambda_{b;k_1,\ldots,k_r;s}$ equals

$$\int_\Delta \left(\sum_{l=0}^{s} \sum_{\substack{i_1,\ldots,i_{r+l}=1 \\ \text{all distinct}}}^{\infty} \binom{s}{l} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r} x_{i_{r+1}} \ldots x_{i_{r+l}} \left(1 - \sum_{j=1}^\infty x_j\right)^{s-l} \right) \Bigg/ \sum_{j=1}^\infty x_j^2 \,\Xi_0(\mathrm{d}x) + a 1_{\{r=1, k_1=2\}}. \tag{2}$$

The coalescent process whose transition rates are given by (2) is called the Ξ-coalescent. As noted in Lemma 3.3 of Möhle and Sagitov (2001) and Lemma 18 of Schweinsberg (2000b), these transition rates satisfy the recursion

$$\lambda_{b+1;\,k_1,\ldots,k_r;\,s+1} = \lambda_{b;\,k_1,\ldots,k_r;\,s} - \sum_{j=1}^r \lambda_{b+1;\,k_1,\ldots,k_{j-1},k_j+1,k_{j+1},\ldots,k_r;\,s} - s\lambda_{b+1;\,k_1,\ldots,k_r,2;\,s-1}. \tag{3}$$

Using (3), we can compute all of the transition rates from those in which $s = 0$.

If Ξ has no atom at zero and

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

then regardless of the number of blocks, the total rate of all transitions is at most 1. Consequently, the Ξ-coalescent is a jump-hold Markov process in which the expected holding time in every state is at least 1. Schweinsberg (2000b) showed that in this case, one can define a discrete-time Ξ-coalescent such that the formula (2) gives transition probabilities rather than transition rates. More precisely, the discrete-time Ξ-coalescent is a $\mathscr{P}_\infty$-valued Markov chain $(Y_m)_{m=0}^{\infty}$ such that for each n, the restricted process $(R_n Y_m)_{m=0}^{\infty}$ is also a Markov chain. If η and ξ are partitions of $\{1,\ldots,n\}$ such that ξ contains b blocks, η contains $r+s$ blocks, s blocks of η consist of a single block of ξ, and the remaining r blocks of η are unions of $k_1,\ldots,k_r$ blocks of ξ, then the transition probability $P(R_n Y_{m+1} = \eta | R_n Y_m = \xi)$, which we denote by $p_{b;k_1,\ldots,k_r;s}$, is given by (2). These transition probabilities also satisfy the recursion (3).

1.2. Coalescents as limits of ancestral processes

It is known that the coalescent processes defined above can arise as limits of ancestral processes in a population model that was introduced by Cannings (1974, 1975). Here we describe the model and the convergence results. Assume that the population has N individuals in each generation, and that there are infinitely many generations backwards in time. Generations do not overlap. For all nonnegative integers m and all $i \in \{1,\ldots,N\}$, let $v_{i,N}^{(m)}$ denote the number of children of the ith individual in generation $-(m+1)$. Note that $v_{1,N}^{(m)} + \cdots + v_{N,N}^{(m)} = N$ because the population size is fixed. We assume that the family sizes in different generations are independent. We also assume that the vector of family sizes $(v_{1,N}^{(m)},\ldots,v_{N,N}^{(m)})$ has the same distribution for all m, so we will drop the superscript in the notation when we are concerned only with the distribution of the family sizes. See Möhle (2002) for some results when the family size distribution may depend on m. We also assume here that $(v_{1,N},\ldots,v_{N,N})$ is exchangeable; see Möhle (1998) and Griffiths and Tavaré (1994) for some results that do not require exchangeability. As we will explain in more detail in the next subsection, the model described at the beginning of the introduction satisfies these conditions.

Sample $n \leqslant N$ distinct individuals at random from generation 0. Define a discrete-time $\mathscr{P}_n$-valued Markov chain $(\Psi_{n,N}(m))_{m=0}^{\infty}$, where $\Psi_{n,N}(m)$ is the 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 the same ancestor in generation $-m$. Let

$$c_N = \frac{E[(v_{1,N})_2]}{N-1},$$

where $(r)_0=1$ and $(r)_k = r(r-1)\ldots(r-k+1)$ for positive integers k. Note that c_N is the probability that two individuals chosen at random from one generation have the same ancestor in the previous generation. We review here some conditions under which the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$, converge as $N \to \infty$ to a limiting coalescent process. (Since these Markov processes have a finite state space, we can equivalently consider either convergence of finite-dimensional distributions or convergence in the Skorohod topology.) The time scaling by a factor of c_N ensures that regardless of the population size, the expected time for two given ancestral lines to merge equals 1.

We first state the most general convergence result, due to Möhle and Sagitov (2001), which gives conditions under which the processes converge to a coalescent with simultaneous multiple collisions. See Sagitov (2002) for another approach to proving convergence.

Proposition 1. *Suppose that for all $r \geqslant 1$ and $k_1, \ldots, k_r \geqslant 2$, the limits*

$$\lim_{N \to \infty} \frac{E[(v_{1,N})_{k_1}(v_{2,N})_{k_2} \ldots (v_{r,N})_{k_r}]}{N^{k_1+\cdots+k_r-r} c_N} \tag{4}$$

exist. If

$$\lim_{N \to \infty} c_N = 0,$$

then the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ converge as $N \to \infty$ to a process $(\Psi_{n,\infty}(t),\ t \geqslant 0)$ which has the same law as the restriction to $\{1, \ldots, n\}$ of a coalescent with simultaneous multiple collisions. Furthermore, if $k_1, \ldots, k_r \geqslant 2$ and $b = k_1 + \cdots + k_r$, then the transition rate $\lambda_{b;\, k_1,\ldots,k_r;\, 0}$ is given by the limit in (4), and these transition rates uniquely determine all the transition rates. If

$$\lim_{N \to \infty} c_N = c > 0, \tag{5}$$

then the processes $(\Psi_{n,N}(m))_{m=0}^{\infty}$ converge as $N \to \infty$ to a process $(Y_m)_{m=0}^{\infty}$, which has the same law as the restriction to $\{1, \ldots, n\}$ of a discrete-time coalescent with simultaneous multiple collisions. When $k_1, \ldots, k_r \geqslant 2$ and $b = k_1 + \cdots + k_r$, the transition probabilities $p_{b;\, k_1,\ldots,k_r;\, 0}$ satisfy

$$p_{b;\, k_1,\ldots,k_r;\, 0} = \lim_{N \to \infty} \frac{E[(v_{1,N})_{k_1}(v_{2,N})_{k_2} \ldots (v_{r,N})_{k_r}]}{N^{k_1+\cdots+k_r-r}}.$$

Furthermore, these transition probabilities uniquely determine all of the transition probabilities.

The convergence results in Proposition 1 are part of Theorem 2.1 of Möhle and Sagitov (2001). The formulas for the transition rates and transition probabilities can be seen from Theorem 2.1 and Eq. (28) of Möhle and Sagitov (2001). The fact that the transition rates $\lambda_{b;\, k_1,\ldots,k_r;\, 0}$ and transition probabilities $p_{b;\, k_1,\ldots,k_r;\, 0}$ uniquely determine the remaining transition rates and transition probabilities is a consequence of recursion (3).

A consequence of Proposition 1 is that to prove that the processes $(\Psi_{n,N}(m))_{m=0}^{\infty}$ converge to the restriction to $\{1, \ldots, n\}$ of the discrete-time Ξ-coalescent, where Ξ has no atom at zero, it suffices to show (5) and to show that for all $r \geqslant 1$ and $k_1, \ldots, k_r \geqslant 2$, we have

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_{k_1}\ldots(\nu_{r,N})_{k_r}]}{N^{k_1+\cdots+k_r-r}} = \int_{\Delta} \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{\infty} x_{i_1}^{k_1}\ldots x_{i_r}^{k_r} \Big/ \sum_{j=1}^{\infty} x_j^2 \,\Xi(\mathrm{d}x). \tag{6}$$

Note that the right-hand side of (6) is the expression in (2) when $s=0$ and Ξ has no atom at zero.

Although all coalescents with simultaneous multiple collisions, up to a time-scaling constant, can arise as limits in these models (see Schweinsberg, 2000b), one gets Kingman's coalescent in the limit as long as the family sizes are not too large. Kingman (1982b) showed that the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ converge to the n-coalescent as long as $\mathrm{Var}(\nu_{1,N})$ converges to a finite limit as $N \to \infty$ and the higher moments $E[\nu_{1,N}^k]$ are bounded as $N \to \infty$. Möhle (1998) gave some other conditions that guarantee convergence to the n-coalescent. To prove convergence to the n-coalescent, we will use the following result from Section 4 of Möhle (2000).

Proposition 2. *Suppose*

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_3]}{N^2 c_N} = 0. \tag{7}$$

Then, as $N \to \infty$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ converge to the n-coalescent.

Möhle (2000) showed that (7) implies that $\lim_{N\to\infty} c_N = 0$, which is why (7) ensures that we get a continuous-time process in the limit.

Coalescents with multiple collisions arise in the limit when there are occasionally large families whose size is order N, but typically at most one large family per generation. Since coalescents with multiple collisions are special cases of coalescents with simultaneous multiple collisions, one could establish convergence to the Λ-coalescent using Proposition 1. We will find it more convenient, however, to use an equivalent condition that can be expressed in terms of the tail probabilities of the family sizes. Proposition 3 below can be deduced from Theorem 2.1 and Remark 1 in Möhle and Sagitov (1998). Alternatively, the result follows from Theorem 3.1 of Sagitov (1999), the equivalence of the limits in (16) and (20) of Möhle and Sagitov (2001), and the monotonicity condition given in Eq. (17) of Möhle and Sagitov (2001).

Proposition 3. *Suppose*

$$\lim_{N\to\infty} c_N = 0 \tag{8}$$

and

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_2(\nu_{2,N})_2]}{N^2 c_N} = 0. \tag{9}$$

Also, assume that for some probability measure Λ on $[0,1]$, we have

$$\lim_{N\to\infty} \frac{N}{c_N} P(\nu_{1,N} > Nx) = \int_x^1 y^{-2}\Lambda(\mathrm{d}y)$$

for all $x \in (0,1)$ *at which the limit function is continuous. Then, as* $N \to \infty$, *the processes* $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ *converge to a process* $(\Psi_{n,\infty}(t),\ t \geqslant 0)$ *that has the same law as the restriction to* $\{1,\dots,n\}$ *of the* Λ*-coalescent.*

1.3. A model involving supercritical Galton–Watson processes

We now describe more precisely the model mentioned at the beginning of this section. The model is of the form considered in Section 1.2, with a fixed population size N and infinitely many generations backwards in time. Note that the model is determined by the distribution of the vectors $(v_{1,N}^{(m)},\dots,v_{N,N}^{(m)})$ of family sizes. We will obtain these family sizes by sampling from the offspring of a supercritical Galton–Watson process. To obtain the family sizes in generation $-m$, start with i.i.d. random variables $X_1,\dots,X_N$. Here X_i represents the number of offspring of the ith individual in generation $-(m+1)$, but only some will survive to form the next generation. If $X_1+\cdots+X_N \geqslant N$, then we obtain the next generation by sampling N of these offspring at random without replacement. We define $v_{i,N}^{(m)}$ to be the number of offspring of the ith individual in generation $-(m+1)$ that are among the N chosen for survival. Note that $v_{1,N}^{(m)}+\cdots+v_{N,N}^{(m)}=N$.

We make two assumptions on the distribution of the number of offspring. First, we assume that

$$E[X_1] > 1. \tag{10}$$

As we will see, this assumption ensures that $X_1+\cdots+X_N \geqslant N$ with sufficiently high probability that we may define the family sizes $(v_{1,N}^{(m)},\dots,v_{N,N}^{(m)})$ arbitrarily when $X_1+\cdots+X_N < N$ without affecting the results. Secondly, for most of our results, we assume that there exist constants $C>0$ and $a>0$ such that

$$P(X_1 \geqslant k) \sim Ck^{-a}, \tag{11}$$

where $\sim$ means that the ratio of the two sides approaches 1 as $k \to \infty$.

Before we state our main result, recall that the Poisson–Dirichlet distributions, which were studied extensively by Pitman and Yor (1997), are a two-parameter family of probability distributions on Δ. The parameters are denoted by (α,θ), and the distribution is defined when $0 \leqslant \alpha < 1$ and $\theta > -\alpha$. When $\theta=0$ and $0<\alpha<1$, which is the case that we will consider in this paper, the Poisson–Dirichlet distribution can be constructed as follows (see Perman et al. (1992) and Proposition 6 of Pitman and Yor (1997)). Let $Z_1 \geqslant Z_2 \geqslant \cdots$ be the ranked points of a Poisson point process on $(0,\infty)$ with characteristic measure Λ_α, where $\Lambda_\alpha((x,\infty))=Cx^{-\alpha}$ for all $x>0$. Note that $(Z_1,Z_2,\dots)$ has the same law as the ranked jump sizes up to time 1 of a pure-jump subordinator $(\tau_t,\ t \geqslant 0)$ whose Lévy measure is Λ_α. This process is a stable subordinator of index α and satisfies $E[\mathrm{e}^{-\lambda\tau_t}]=\mathrm{e}^{-tC\Gamma(1-\alpha)\lambda^\alpha}$. For all j, let $W_j=Z_j/\sum_{i=1}^\infty Z_i$. Then, the sequence $(W_1,W_2,\dots)$ has the Poisson–Dirichlet distribution with parameters $(\alpha,0)$.

Theorem 4. *Assume* (10) *is satisfied.*

(a) *If* $E[X_1^2]<\infty$ (*in particular, if* (11) *holds and* $a>2$), *then the processes* $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ *converge as* $N \to \infty$ *to the n-coalescent.*

(b) *If* (11) *holds and* $a = 2$, *then the processes* $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ *converge as* $N \to \infty$ *to the n-coalescent.*

(c) *When* (11) *holds with* $1 \leqslant a < 2$, *the processes* $(\Psi_{n,N}(\lfloor t/c_N \rfloor),\ t \geqslant 0)$ *converge as* $N \to \infty$ *to a continuous-time process* $(\Psi_{n,\infty}(t),\ t \geqslant 0)$ *that has the same law as the restriction to* $\{1,\ldots,n\}$ *of the* Λ*-coalescent, where* Λ *is the probability measure associated with the Beta*$(2-a,a)$ *distribution. The transition rates are given by*

$$\lambda_{b,k} = \frac{B(k-a, b-k+a)}{B(2-a,a)}, \tag{12}$$

where $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ *is the beta function.*

(d) *For* $0 < a < 1$, *let* Θ_a *be the probability measure on* Δ *associated with the Poisson–Dirichlet distribution with parameters* $(a,0)$. *Let* Ξ_a *be the measure on* Δ *defined by*

$$\Xi_a(\mathrm{d}x) = \left(\sum_{j=1}^{\infty} x_j^2 \right) \Theta_a(\mathrm{d}x).$$

When (11) *holds with* $0 < a < 1$, *the processes* $(\Psi_{n,N}(m))_{m=0}^{\infty}$ *converge as* $N \to \infty$ *to a discrete-time Markov chain* $(Y_m)_{m=0}^{\infty}$ *that has the same law as the restriction to* $\{1,\ldots,n\}$ *of a discrete-time* Ξ_a*-coalescent. The transition probabilities are given by*

$$p_{b;\,k_1,\ldots,k_r;\,s} = \frac{a^{r+s-1}(r+s-1)!}{(b-1)!} \prod_{i=1}^{r} (k_i - a)_{k_i}. \tag{13}$$

Note that large family sizes are more likely for small values of a. The above theorem says that when $a \geqslant 2$, large families are sufficiently rare that we get Kingman's coalescent in the limit. When $1 \leqslant a < 2$, there are enough large families to produce multiple collisions, but it is rare to have two large families in any one generation. When $0 < a < 1$, each generation has many large families, which is why we get a discrete-time coalescent with simultaneous multiple collisions.

Note that the transition rates in (12) follow immediately from (1) and the fact that $\Lambda(\mathrm{d}x) = B(2-a,a)^{-1}x^{1-a}(1-x)^{a-1}\,\mathrm{d}x$. When $a = 1$, the measure Λ is the uniform distribution on $(0,1)$, and (12) becomes

$$\lambda_{b,k} = \frac{\Gamma(k-1)\Gamma(b-k+1)}{\Gamma(b)} = \frac{(k-2)!(b-k)!}{(b-1)!},$$

as noted in Pitman (1999). The limiting coalescent process in this case is the Bolthausen–Sznitman coalescent, which was introduced by Bolthausen and Sznitman (1998). Bertoin and Le Gall (2000) showed that the Bolthausen–Sznitman coalescent is closely related to a continuous-state branching process that was first studied by Neveu (1992). See Bertoin and Pitman (2000) and Pitman (1999) for additional work related to the Bolthausen–Sznitman coalescent.

Schweinsberg (2000a) showed that if Λ is the Beta(α, β) distribution, then the Λ-coalescent "comes down from infinity," meaning that there are only finitely many blocks at all positive times even when there are infinitely many blocks at time zero, if and only if $\alpha < 1$. If T_n denotes the first time at which the integers $\{1, \ldots, n\}$ are in the same block (or, equivalently, the first time at which the process restricted to $\{1, \ldots, n\}$ contains only a single block), then the condition that the coalescent comes down from infinity was shown in Schweinsberg (2000a) to be equivalent to

$$\lim_{n \to \infty} E[T_n] < \infty. \tag{14}$$

Thus, in part (c) of Theorem 4, the limiting coalescent process satisfies (14) when $1 < a < 2$, but not when $a = 1$. The fact that the Bolthausen–Sznitman coalescent does not come down from infinity had previously been shown by Bolthausen and Sznitman (1998), and Sagitov (1999) considered some related examples.

2. Convergence to Kingman's coalescent when $a \geqslant 2$

Throughout this section, as well as the next two sections, we will assume that we are working with the model defined in Section 1.3, and that Eq. (10) holds. We first introduce some notation. Let

$$S_N = X_1 + \cdots + X_N.$$

Let

$$\mu = E[X_1].$$

When (11) holds, we can define positive constants C' and C'' such that

$$C' k^{-a} \leqslant P(X_1 \geqslant k) \leqslant C'' k^{-a} \tag{15}$$

for all positive integers k.

Our goal in this section is to show that when $E[X_1^2] < \infty$ or when (11) holds and $a = 2$, the limit of the ancestral processes defined in the introduction is the n-coalescent. We will need to check the condition (7) in Proposition 2. We begin with two lemmas that can be applied whenever (10) holds.

Lemma 5. *There exists a constant $A < 1$ such that $P(S_N \leqslant N) \leqslant A^N$ for all N.*

Proof. For all $r \in [0, 1]$, let $\rho(r) = E[r^{X_1}]$. Then $\mu = \rho'(1)$, regardless of whether μ is finite or infinite (see, for example, Theorem 13 in chapter 5 of Fristedt and Gray (1997)). Note that $E[r^{S_N}] \geqslant r^N P(S_N \leqslant N)$, which means

$$P(S_N \leqslant N) \leqslant r^{-N} E[r^{S_N}] = (r^{-1}\rho(r))^N$$

for all $r \in [0, 1]$. Since $\rho(1) = 1$ and $\rho'(1) = \mu > 1$, there exists a number $r \in (0, 1)$ such that $\rho(r) < r$ and therefore $r^{-1}\rho(r) < 1$. It follows that $P(S_N \leqslant N) \leqslant A^N$, where $A = r^{-1}\rho(r)$. $\square$

Lemma 6. *For all $r \geqslant 1$ and $k_1, \ldots, k_r \geqslant 2$, we have*

$$\lim_{N\to\infty} \frac{E[(v_{1,N})_{k_1} \ldots (v_{r,N})_{k_r}]}{N^{k_1+\cdots+k_r-r} c_N} = \lim_{N\to\infty} \frac{N^r}{c_N} E\left[\frac{(X_1)_{k_1} \ldots (X_r)_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right], \tag{16}$$

in the sense that if either limit exists, then so does the other, and the limits are equal. Also,

$$c_N \sim NE\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right], \tag{17}$$

where $\sim$ means that the ratio of the sides approaches zero as $N \to \infty$. Moreover, there exists a constant $A_1 > 0$ such that

$$c_N \geqslant \frac{A_1}{N} \tag{18}$$

for all N.

Proof. Place the individuals in the current generation in random order. Independently, place the individuals in the previous generation in random order. Let $B_{k_1,\ldots,k_r}$ be the event that the first k_1 individuals in the current generation are descended from the first individual in the previous generation, the next k_2 individuals in the current generation are descended from the second individual in the previous generation, and so on. We have

$$P(B_{k_1,\ldots,k_r}) = E[P(B_{k_1,\ldots,k_r} | v_{1,N}, \ldots, v_{N,N})] = E\left[\frac{(v_{1,N})_{k_1} \ldots (v_{r,N})_{k_r}}{(N)_{k_1+\cdots+k_r}}\right]. \tag{19}$$

Thus,

$$\frac{N^r}{c_N} P(B_{k_1,\ldots,k_r}) \sim \frac{E[(v_{1,N})_{k_1} \ldots (v_{r,N})_{k_r}]}{N^{k_1+\cdots+k_r-r} c_N}. \tag{20}$$

We have

$$\begin{aligned} P(B_{k_1,\ldots,k_r}) &= P(B_{k_1,\ldots,k_r} \cap \{S_N \geqslant N\}) + P(B_{k_1,\ldots,k_r} \cap \{S_N < N\}) \\ &= E[P(B_{k_1,\ldots,k_r} \cap \{S_N \geqslant N\}) | X_1, \ldots, X_N] + P(B_{k_1,\ldots,k_r} \cap \{S_N < N\}) \\ &= E\left[\frac{(X_1)_{k_1} \ldots (X_r)_{k_r}}{(S_N)_{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right] + P(B_{k_1,\ldots,k_r} \cap \{S_N < N\}). \end{aligned} \tag{21}$$

Eq. (19) gives $c_N = E[(v_{1,N})_2]/(N-1) = NP(B_2)$. Therefore, using (21) for the first inequality, Jensen's inequality for the fourth, and Lemma 5 for the last, we have

$$\begin{aligned} c_N &= NP(B_2) \geqslant NE\left[\frac{(X_1)_2}{(S_N)_2} 1_{\{S_N \geqslant N\}}\right] \geqslant NE\left[\frac{X_1(X_1-1)}{S_N^2} 1_{\{S_N \geqslant N\}}\right] \\ &\geqslant \frac{N}{2} E\left[\left(\frac{X_1}{S_N}\right)^2 1_{\{X_1 \geqslant 2, S_N \geqslant N\}}\right] \geqslant \frac{N}{2}\left(E\left[\frac{X_1}{S_N} 1_{\{X_1 \geqslant 2, S_N \geqslant N\}}\right]\right)^2 \end{aligned}$$

$$= \frac{N}{2}\left(E\left[\frac{X_1}{S_N}\middle| X_1 \geqslant 2,\ S_N \geqslant N\right] P(X_1 \geqslant 2,\ S_N \geqslant N)\right)^2$$

$$\geqslant \frac{N}{2}\left(E\left[\frac{X_1}{S_N}\middle| X_1 \geqslant 2,\ S_N \geqslant N\right] \left(P(X_1 \geqslant 2) - A^N\right)\right)^2.$$

For all $k \geqslant 1$, we have $E[X_1/S_N|S_N = k] = 1/N$. Therefore $E[X_1/S_N|X_1 \geqslant 2,\ S_N = k] \geqslant 1/N$. Thus,

$$c_N \geqslant \frac{N}{2}\left(\frac{P(X_1 \geqslant 2) - A^N}{N}\right)^2 = \frac{(P(X_1 \geqslant 2) - A^N)^2}{2N}.$$

Since $\mu > 1$, we have $P(X_1 \geqslant 2) > 0$. Since $A < 1$, for sufficiently large N we have $P(X_1 \geqslant 2) - A^N > 0$. Thus, there exists a positive constant A_1 such that (18) holds.

To prove (16), note that $0 \leqslant P(B_{k_1,\ldots,k_r} \cap \{S_N < N\}) \leqslant A^N$ by Lemma 5. Combining this fact with (18), we get

$$\lim_{N\to\infty} \frac{N^r}{c_N} P(B_{k_1,\ldots,k_r} \cap \{S_N < N\}) = 0. \tag{22}$$

It follows from (21) and (22) that

$$\lim_{N\to\infty} \frac{N^r}{c_N} P(B_{k_1,\ldots,k_r}) = \lim_{N\to\infty} \frac{N^r}{c_N} E\left[\frac{(X_1)_{k_1}\ldots(X_r)_{k_r}}{(S_N)_{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right]$$

$$= \lim_{N\to\infty} \frac{N^r}{c_N} E\left[\frac{(X_1)_{k_1}\ldots(X_r)_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right].$$

This fact, combined with (20), implies (16).

Finally, (17) follows from (16) when $r=1$ and $k_1=2$, as the left-hand side of (16) equals 1 in that case. □

The next proposition, combined with Proposition 2, proves part (a) of Theorem 4.

Proposition 7. *If* $E[X_1^2] < \infty$, *then*

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_3]}{N^2 c_N} = 0.$$

Proof. Using (16) and (18) from Lemma 6, we get

$$\limsup_{N\to\infty} \frac{E[(\nu_{1,N})_3]}{N^2 c_N} = \limsup_{N\to\infty} \frac{N}{c_N} E\left[\frac{(X_1)_3}{S_N^3} 1_{\{S_N \geqslant N\}}\right]$$

$$\leqslant \limsup_{N\to\infty} \frac{N^2}{A_1} E\left[\frac{(X_1)_3}{S_N^3} 1_{\{S_N \geqslant N\}}\right].$$

Thus, to prove the proposition, it suffices to show that

$$\lim_{N\to\infty} N^2 E\left[\frac{(X_1)_3}{S_N^3} 1_{\{S_N \geqslant N\}}\right] = 0. \tag{23}$$

We have

$$\begin{aligned} N^2 E\left[\frac{(X_1)_3}{S_N^3} 1_{\{S_N \geqslant N\}}\right] &\leqslant N^2 E\left[\frac{X_1^3}{\max\{X_1^3, N^3\}}\right] \\ &= N^2 \left(\sum_{k=0}^{N-1} \frac{k^3}{N^3} P(X_1 = k) + \sum_{k=N}^{\infty} P(X_1 = k)\right) \\ &= \frac{1}{N} \sum_{k=0}^{N-1} k^3 P(X_1 = k) + N^2 P(X_1 \geqslant N). \end{aligned} \tag{24}$$

Since $E[X_1^2] < \infty$, we have

$$\limsup_{N\to\infty} N^2 P(X_1 \geqslant N) \leqslant \limsup_{N\to\infty} E[X_1^2 1_{\{X_1 \geqslant N\}}] = 0. \tag{25}$$

Let $\varepsilon > 0$. Choose a positive integer L such that $E[X_1^2 1_{\{X_1 \geqslant L\}}] < \varepsilon/2$. Suppose N is large enough that $LE[X_1^2]/N < \varepsilon/2$. Then,

$$\begin{aligned} \frac{1}{N}\sum_{k=0}^{N-1} k^3 P(X_1 = k) &= \frac{1}{N}\sum_{k=0}^{L-1} k^3 P(X_1 = k) + \frac{1}{N}\sum_{k=L}^{N-1} k^3 P(X_1 = k) \\ &\leqslant \frac{L}{N}\sum_{k=0}^{L-1} k^2 P(X_1 = k) + \sum_{k=L}^{N-1} k^2 P(X_1 = k) \\ &\leqslant \frac{LE[X_1^2]}{N} + E[X_1^2 1_{\{X_1 \geqslant L\}}] < \varepsilon. \end{aligned}$$

Therefore,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} k^3 P(X_1 = k) = 0. \tag{26}$$

Equations (24)–(26) imply (23). □

We now establish the results necessary to handle the case in which (11) holds and $a = 2$.

Lemma 8. *If $\mu < \infty$, then there exists a constant $A_2 > 0$ such that*

$$c_N \geqslant A_2 N E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right].$$

Proof. Using Lemma 6 for the asymptotic equivalence in the first line, Lemma 5 in the third line, and Markov's inequality in the fourth line, we obtain

$$
\begin{aligned}
c_N &\sim NE\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] \geqslant NE\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}} 1_{\{X_2+\cdots+X_N \leqslant 2(N-1)\mu\}}\right] \\
&\geqslant NE\left[\frac{(X_1)_2}{(X_1+2\mu(N-1))^2} 1_{\{X_2+\cdots+X_N \leqslant 2(N-1)\mu\}}\right] - NP(S_N < N) \\
&\geqslant NE\left[\frac{(X_1)_2}{(X_1+2\mu(N-1))^2}\right] P(X_2+\cdots+X_N \leqslant 2(N-1)\mu) - NA^N \\
&\geqslant \frac{N}{2} E\left[\frac{(X_1)_2}{(X_1+2\mu(N-1))^2}\right] - NA^N \geqslant \frac{N}{8\mu^2} E\left[\frac{(X_1)_2}{(X_1+N)^2}\right] - NA^N \\
&\geqslant \frac{N}{32\mu^2} E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right] - NA^N. \qquad (27)
\end{aligned}
$$

The lemma follows from (27) and the fact that $c_N \geqslant A_1/N$ by Lemma 6. □

Lemma 9. *Let g be a real-valued function defined on $\{0,1,\ldots\}$, and let X be a $\{0,1,\ldots\}$-valued random variable. Then*

$$
\begin{aligned}
&\sum_{k=0}^{N} g(k)P(X=k) \\
&\quad = g(0) - g(N)P(X \geqslant N+1) + \sum_{k=1}^{N} [g(k)-g(k-1)]P(X \geqslant k). \qquad (28)
\end{aligned}
$$

If $\lim_{N\to\infty} g(N)P(X \geqslant N-1) = 0$, *then*

$$
E[g(X)] = \sum_{k=0}^{\infty} g(k)P(X=k) = g(0) + \sum_{k=1}^{\infty} [g(k)-g(k-1)]P(X \geqslant k). \qquad (29)
$$

Proof. This result is just summation by parts. We have

$$
\begin{aligned}
\sum_{k=0}^{N} g(k)P(X=k) &= \sum_{k=0}^{N} g(k)[P(X \geqslant k) - P(X \geqslant k+1)] \\
&= g(0)P(X \geqslant 0) - g(N)P(X \geqslant N+1) \\
&\quad + \sum_{k=1}^{N} [g(k)-g(k-1)]P(X \geqslant k),
\end{aligned}
$$

which implies (28) because $P(X_1 \geqslant 0) = 1$. Equation (29) follows by taking the limit as $N \to \infty$ in (28). □

Lemma 10. *If* (11) *holds with* $a=2$, *then there exists a constant* A_3 *such that*

$$c_N \geqslant \frac{A_3 \log N}{N}$$

for all N.

Proof. Note that $(k)_2 - (k-1)_2 = 2(k-1)$. Using this fact, Lemmas 8 and 9, and Eq. (15), we get

$$\begin{aligned} c_N &\geqslant A_2 N E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right] \geqslant A_2 N \sum_{k=0}^{N} \frac{(k)_2}{N^2} P(X_1 = k) \\ &= \frac{A_2}{N}\left(-N(N-1)P(X_1 \geqslant N+1) + \sum_{k=1}^{N} 2(k-1)P(X_1 \geqslant k)\right) \\ &\geqslant \frac{A_2}{N}\left(-\frac{C''N(N-1)}{(N+1)^2} + \sum_{k=1}^{N}\frac{2C'(k-1)}{k^2}\right). \end{aligned} \tag{30}$$

The first term inside the parentheses on the right-hand side of (30) stays bounded as $N \to \infty$, while the second term is asymptotically equivalent to $2C' \log N$. The lemma follows from these observations. $\square$

The next proposition implies part (b) of Theorem 4.

Proposition 11. *If* (11) *holds with* $a=2$, *then*

$$\lim_{N\to\infty} \frac{E[(v_{1,N})_3]}{N^2 c_N} = 0.$$

Proof. By Lemma 6, it suffices to show that

$$\lim_{N\to\infty} \frac{N}{c_N} E\left[\frac{(X_1)_3}{S_N^3} 1_{\{S_N \geqslant N\}}\right] = 0. \tag{31}$$

It follows from (24) that

$$\frac{N}{c_N} E\left[\frac{(X_1)_3}{S_N^3} 1_{\{S_N \geqslant N\}}\right] \leqslant \frac{1}{N^2 c_N}\sum_{k=0}^{N-1} k^3 P(X_1 = k) + \frac{N}{c_N} P(X_1 \geqslant N). \tag{32}$$

By Lemmas 9 and 10 and the upper bound from (15),

$$\begin{aligned} \frac{1}{N^2 c_N}\sum_{k=0}^{N-1} k^3 P(X_1 = k) &\leqslant \frac{1}{A_3 N \log N}\sum_{k=1}^{N-1} (k^3 - (k-1)^3) P(X \geqslant k) \\ &\leqslant \frac{1}{A_3 N \log N}\sum_{k=1}^{N} (3k^2)(C''k^{-2}) = \frac{3C''}{A_3 \log N}. \end{aligned} \tag{33}$$

Likewise,

$$\frac{N}{c_N} P(X_1 \geqslant N) \leqslant \frac{N^2}{A_3 \log N}(C'' N^{-2}) = \frac{C''}{A_3 \log N}. \tag{34}$$

Eqs. (32), (33), and (34) imply (31). □

3. Convergence to the Λ-coalescent when $1 \leqslant a < 2$

Our goal in this section is to prove that when $1 \leqslant a < 2$, the limiting coalescent process is the Λ-coalescent, where Λ has the beta density given by $\Lambda(\mathrm{d}x) = B(2-a, a)^{-1} x^{1-a}(1-x)^{a-1}\,\mathrm{d}x$. When $a = 1$, we have $E[X_1] = \infty$, so this case requires a different argument from the case $1 < a < 2$.

Lemma 12. *If* (11) *holds with* $1 \leqslant a < 2$, *then*

$$\lim_{M\to\infty} M^a E\left[\frac{(X_1)_2}{(X_1+M)^2}\right] = CaB(2-a, a).$$

Proof. By Lemma 9,

$$\begin{aligned} M^a E\left[\frac{(X_1)_2}{(X_1+M)^2}\right] &= M^a \sum_{k=1}^{\infty} \left(\frac{k(k-1)}{(k+M)^2} - \frac{(k-1)(k-2)}{(k-1+M)^2}\right) P(X_1 \geqslant k) \\ &= M^a \sum_{k=1}^{\infty} \left(\frac{(k-1)(2M(k+M)+k)}{(k+M)^2(k-1+M)^2}\right) P(X_1 \geqslant k). \end{aligned}$$

Let $\varepsilon > 0$. Choose L large enough that if $k \geqslant L$, then $(1-\varepsilon)Ck^{-a} \leqslant P(X_1 \geqslant k) \leqslant (1+\varepsilon)Ck^{-a}$, and if $k \geqslant L$ and $M \geqslant L$, then

$$\begin{aligned} (1-\varepsilon)\int_k^{k+1} \frac{2Mx^{1-a}}{(x+M)^3}\,\mathrm{d}x &\leqslant \frac{(k-1)[2M(k+M)+k]k^{-a}}{(k+M)^2(k-1+M)^2} \\ &\leqslant (1+\varepsilon)\int_k^{k+1} \frac{2Mx^{1-a}}{(x+M)^3}\,\mathrm{d}x. \end{aligned}$$

Since $a < 2$, we have

$$\lim_{M\to\infty} M^a \sum_{k=1}^{L-1} \left(\frac{(k-1)(2M(k+M)+k)}{(k+M)^2(k-1+M)^2}\right) P(X_1 \geqslant k) = 0.$$

Therefore,

$$\limsup_{M\to\infty} M^a E\left[\frac{(X_1)_2}{(X_1+M)^2}\right] \leqslant (1+\varepsilon)^2 C M^a \int_L^{\infty} \frac{2Mx^{1-a}}{(x+M)^3}\,\mathrm{d}x \tag{35}$$

and

$$\liminf_{M\to\infty} M^a E\left[\frac{(X_1)_2}{(X_1+M)^2}\right] \geqslant (1-\varepsilon)^2 C M^a \int_L^\infty \frac{2Mx^{1-a}}{(x+M)^3}\,dx. \tag{36}$$

Making the substitution $y = M/(M+x)$, so that $x = My^{-1}(1-y)$ and $dx/dy = -My^{-2}$, we get

$$\int_L^\infty \frac{x^{1-a}}{(x+M)^3}\,dx = \int_0^{M/(M+L)} \left(\frac{M(1-y)}{y}\right)^{1-a} \left(\frac{y}{M}\right)^3 My^{-2}\,dy$$
$$= M^{-1-a} \int_0^{M/(M+L)} y^a (1-y)^{1-a}\,dy.$$

Therefore,

$$\lim_{M\to\infty} M^a \int_L^\infty \frac{2Mx^{1-a}}{(x+M)^3}\,dx = 2\int_0^1 y^a(1-y)^{1-a}\,dy = \frac{2\Gamma(a+1)\Gamma(2-a)}{\Gamma(3)}$$
$$= \frac{a\Gamma(a)\Gamma(2-a)}{\Gamma(2)} = aB(2-a,a). \tag{37}$$

By letting $\varepsilon \to 0$, we obtain the conclusion of the lemma from (35)–(37). □

Lemma 13. *If* (11) *holds with* $1 < a < 2$, *then*

$$\lim_{N\to\infty} N^{a-1} c_N = Ca\mu^{-a} B(2-a,a).$$

Proof. By Lemma 6, it suffices to show that

$$\lim_{N\to\infty} N^a E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] = Ca\mu^{-a}B(2-a,a). \tag{38}$$

Let $\varepsilon > 0$, and choose $\delta > 0$ to be small enough that $(1-\delta)\mu > 1$. By the law of large numbers,

$$P((1-\delta)N\mu \leqslant X_2 + \cdots + X_N \leqslant (1+\delta)N\mu) > 1-\varepsilon \tag{39}$$

for sufficiently large N. For N large enough that (39) holds,

$$E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N\geqslant N\}}\right] = E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N\geqslant N\}} 1_{\{X_2+\cdots+X_N<(1-\delta)N\mu\}}\right]$$
$$+E\left[\frac{(X_1)_2}{S_N^2} 1_{\{X_2+\cdots+X_N\geqslant(1-\delta)N\mu\}}\right]$$

$$\leqslant \varepsilon E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right] + E\left[\frac{(X_1)_2}{(X_1+(1-\delta)N\mu)^2}\right]$$

$$\leqslant 4\varepsilon E\left[\frac{(X_1)_2}{(X_1+N)^2}\right] + E\left[\frac{(X_1)_2}{(X_1+(1-\delta)N\mu)^2}\right] \tag{40}$$

and

$$E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] \geqslant E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}} 1_{\{(1-\delta)N\mu \leqslant X_2+\cdots+X_N \leqslant (1+\delta)N\mu\}}\right]$$

$$\geqslant (1-\varepsilon)E\left[\frac{(X_1)_2}{(X_1+(1+\delta)N\mu)^2}\right]. \tag{41}$$

It follows from (40) and Lemma 12 that

$$\limsup_{N\to\infty} N^a E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right]$$

$$\leqslant 4\varepsilon CaB(2-a,a) + [(1-\delta)\mu]^{-a} CaB(2-a,a). \tag{42}$$

Likewise, (41) and Lemma 12 give

$$\liminf_{N\to\infty} N^a E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] \geqslant (1-\varepsilon)[(1+\delta)\mu]^{-a} CaB(2-a,a). \tag{43}$$

Finally, (38) follows from (42) and (43) after letting $\delta, \varepsilon \to 0$. □

Lemma 14. *If* (11) *holds with* $1 < a < 2$, *then for all* $x \in (0,1)$,

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) = \frac{1}{B(2-a,a)} \int_x^1 y^{-1-a}(1-y)^{a-1}\,\mathrm{d}y.$$

Proof. Fix $x \in (0,1)$. Let $\varepsilon > 0$, and choose $\delta > 0$ to be small enough that $(1-\delta)\mu > 1$. Assume N is large enough that (39) holds. By considering separately the events $\{X_2 + \cdots + X_N < (1-\delta)N\mu\}$ and $\{X_2 + \cdots + X_N \geqslant (1-\delta)N\mu\}$, we get

$$P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \leqslant \varepsilon P\left(\frac{X_1}{N} \geqslant x\right) + P\left(\frac{X_1}{X_1+(1-\delta)N\mu} \geqslant x\right). \tag{44}$$

By considering only the event $\{(1-\delta)N\mu \leqslant X_2 + \cdots + X_N \leqslant (1+\delta)N\mu\}$, we obtain

$$P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \geqslant (1-\varepsilon)P\left(\frac{X_1}{X_1+(1+\delta)N\mu} \geqslant x\right). \tag{45}$$

Combining (44) with (11) and Lemma 13, we get

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \\
&\quad \leqslant \limsup_{N\to\infty} \frac{N}{c_N} \left(\varepsilon P(X_1 \geqslant xN) + P\left(X_1 \geqslant \frac{x}{1-x}(1-\delta)N\mu\right)\right) \\
&\quad = \limsup_{N\to\infty} \frac{CN}{c_N} \left(\varepsilon x^{-a} N^{-a} + \left(\frac{x}{1-x}\right)^{-a} (1-\delta)^{-a} \mu^{-a} N^{-a}\right) \\
&\quad = \frac{1}{B(2-a,a)} \left(\frac{\varepsilon x^{-a}}{a\mu^{-a}} + \left(\frac{1-x}{x}\right)^{a} \frac{(1-\delta)^{-a}}{a}\right).
\end{aligned} \tag{46}$$

Likewise, from (45) and Lemma 13, we get

$$\liminf_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \geqslant \frac{1-\varepsilon}{B(2-a,a)} \left(\frac{1-x}{x}\right)^{a} \frac{(1+\delta)^{-a}}{a}. \tag{47}$$

Taking the limit as $\delta, \varepsilon \to 0$ in Eqs. (46) and (47), we get

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) = \frac{1}{aB(2-a,a)} \left(\frac{1-x}{x}\right)^{a}.$$

The lemma now follows from the fact that

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

which can be seen by substituting $z = (1-y)/y$. □

Lemma 15. *If $\mu < \infty$, then*

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_2(\nu_{2,N})_2]}{N^2 c_N} = 0.$$

Proof. By Lemma 6, it suffices to show

$$\lim_{N\to\infty} \frac{N^2}{c_N} E\left[\frac{(X_1)_2(X_2)_2}{S_N^4} 1_{\{S_N \geqslant N\}}\right] = 0. \tag{48}$$

Note that

$$\begin{aligned}
E\left[\frac{(X_1)_2(X_2)_2}{S_N^4} 1_{\{S_N \geqslant N\}}\right] &\leqslant E\left[\frac{(X_1)_2(X_2)_2}{(\max\{X_1^2, N^2\})(\max\{X_2^2, N^2\})}\right] \\
&= \left(E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right]\right)^2.
\end{aligned}$$

By Lemma 8,

$$E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right] \leqslant \frac{c_N}{A_2 N}.$$

Since $\lim_{N\to\infty} c_N = 0$ by Lemma 13,

$$\limsup_{N\to\infty} \frac{N^2}{c_N} E\left[\frac{(X_1)_2(X_2)_2}{S_N^4} 1_{\{S_N \geqslant N\}}\right] \leqslant \limsup_{N\to\infty} \frac{N^2}{c_N}\left(\frac{c_N}{A_2 N}\right)^2 = \limsup_{N\to\infty} \frac{c_N}{A_2^2} = 0,$$

which proves the lemma. □

We next prove three lemmas that pertain to the case in which $a = 1$. These lemmas are similar to Lemmas 13–15.

Lemma 16. *If* (11) *holds with* $a = 1$, *then* $\lim_{N\to\infty} (\log N) c_N = 1$.

Proof. By Lemma 6, it suffices to show that

$$\lim_{N\to\infty} (N \log N) E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] = 1.$$

Fix $B > 0$. Define $Y_i = X_i 1_{\{X_i \leqslant BN\}}$. For all positive integers L such that $L \leqslant BN$,

$$\begin{aligned} E[Y_1] &= \sum_{k=1}^{\infty} P(Y_1 \geqslant k) = \sum_{k=1}^{\lfloor BN \rfloor} [P(X_1 \geqslant k) - P(X_1 > BN)] \\ &= \sum_{k=1}^{L-1} P(X_1 \geqslant k) + \sum_{k=L}^{\lfloor BN \rfloor} P(X_1 \geqslant k) - \sum_{k=1}^{\lfloor BN \rfloor} P(X_1 > BN). \end{aligned} \tag{49}$$

Since $P(X_1 \geqslant k) \leqslant 1$ for all k and $P(X_1 > BN) \leqslant C''(BN)^{-1}$ by (15),

$$\lim_{N\to\infty} \frac{1}{\log N}\left(\sum_{k=1}^{L-1} P(X_1 \geqslant k) - \sum_{k=1}^{\lfloor BN \rfloor} P(X_1 > BN)\right) = 0 \tag{50}$$

for all L. Let $\eta > 0$. By (11), we can choose L large enough that

$$C(1-\eta)\int_k^{k+1} x^{-1}\,\mathrm{d}x \leqslant P(X_1 \geqslant k) \leqslant C(1+\eta)\int_k^{k+1} x^{-1}\,\mathrm{d}x \tag{51}$$

for all $k \geqslant L$. Note that

$$\lim_{N\to\infty} \frac{C}{\log N}\int_L^{\lfloor BN \rfloor + 1} x^{-1}\,\mathrm{d}x = \lim_{N\to\infty} \frac{C}{\log N}(\log(\lfloor BN \rfloor + 1) - \log L) = C. \tag{52}$$

It follows by combining (49)–(52) and letting $\eta \to 0$ that

$$\lim_{N\to\infty} E\left[\frac{Y_1}{\log N}\right] = C.$$

Also, using (15) and Lemma 9,

$$\operatorname{Var}(Y_1) \leqslant E[Y_1^2] = \sum_{k=1}^{\infty}[k^2-(k-1)^2]P(Y_1 \geqslant k)$$
$$\leqslant \sum_{k=1}^{\lfloor BN \rfloor} 2kP(X_1 \geqslant k) \leqslant 2C''BN. \tag{53}$$

Furthermore, we have

$$\lim_{N\to\infty} P\left(\max_{1\leqslant i\leqslant N} X_i > BN\right) = \lim_{N\to\infty} 1-(1-P(X_1 > BN))^N = 1-e^{-C/B}.$$

Let $0 < \delta < 1/2$, and let $\varepsilon > 0$. Choose B large enough that $1-e^{-C/B} < \varepsilon/4$. Choose N large enough that the following three conditions hold:

$$\left|1 - E\left[\frac{Y_2+\cdots+Y_N}{CN\log N}\right]\right| < \frac{\delta}{2},$$

$$\left|P\left(\max_{1\leqslant i\leqslant N} X_i > BN\right) - 1 + e^{-C/B}\right| < \frac{\varepsilon}{4},$$

$$\frac{8C''B}{C^2(\log N)^2\delta^2} < \frac{\varepsilon}{2}.$$

By Chebyshev's inequality and (53),

$$P\left(\left|\frac{X_2+\cdots+X_N}{CN\log N} - 1\right| \geqslant \delta\right)$$
$$\leqslant P\left(\left|\frac{Y_2+\cdots+Y_N}{CN\log N} - E\left[\frac{Y_2+\cdots+Y_N}{CN\log N}\right]\right| \geqslant \frac{\delta}{2}\right) + P\left(\max_{1\leqslant i\leqslant N} X_i > BN\right)$$
$$\leqslant \operatorname{Var}\left(\frac{Y_2+\cdots+Y_N}{CN\log N}\right)\left(\frac{\delta}{2}\right)^{-2} + \frac{\varepsilon}{2} \leqslant \frac{8C''BN(N-1)}{C^2N^2(\log N)^2\delta^2} + \frac{\varepsilon}{2} < \varepsilon. \tag{54}$$

Also, since $\delta < 1/2$, Chebyshev's inequality and (53) give

$$P\left(X_2+\cdots+X_N \leqslant \frac{C}{2}N\log N\right)$$
$$\leqslant P\left(Y_2+\cdots+Y_N \leqslant \frac{C}{2}N\log N\right)$$
$$\leqslant P\left(\left|\frac{Y_2+\cdots+Y_N}{CN\log N} - E\left[\frac{Y_2+\cdots+Y_N}{CN\log N}\right]\right| \geqslant \frac{1}{4}\right)$$
$$\leqslant 16\operatorname{Var}\left(\frac{Y_2+\cdots+Y_N}{CN\log N}\right) \leqslant \frac{32C''B}{C^2(\log N)^2}. \tag{55}$$

We now consider separately the events $D_1 = \{X_2 + \cdots + X_N < \frac{C}{2} N \log N\}$, $D_2 = \{\frac{C}{2} N \log N \leqslant X_2 + \cdots + X_N < C(1-\delta) N \log N\}$, and $D_3 = \{X_2 + \cdots + X_N \geqslant C(1-\delta) N \log N\}$. Inequalities (54) and (55) imply

$$\begin{aligned} E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] &\leqslant P(D_1) E\left[\frac{(X_1)_2}{\max\{X_1^2, N^2\}}\right] + P(D_2) E\left[\frac{(X_1)_2}{(X_1 + \frac{C}{2} N \log N)^2}\right] \\ &\quad + E\left[\frac{(X_1)_2}{(X_1 + C(1-\delta) N \log N)^2}\right] \\ &\leqslant \frac{128 C'' B}{C^2 (\log N)^2} E\left[\frac{(X_1)_2}{(X_1 + N)^2}\right] + \varepsilon E\left[\frac{(X_1)_2}{(X_1 + \frac{C}{2} N \log N)^2}\right] \\ &\quad + E\left[\frac{(X_1)_2}{(X_1 + C(1-\delta) N \log N)^2}\right]. \end{aligned}$$

Let $D_4 = \{C(1-\delta) N \log N \leqslant X_2 + \cdots + X_N \leqslant C(1+\delta) N \log N\}$. Then

$$\begin{aligned} E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] &\geqslant P(D_4) E\left[\frac{(X_1)_2}{(X_1 + C(1+\delta) N \log N)^2}\right] \\ &\geqslant (1-\varepsilon) E\left[\frac{(X_1)_2}{(X_1 + C(1+\delta) N \log N)^2}\right]. \end{aligned}$$

Note that $CaB(a, 2-a) = C$ when $a = 1$. Therefore, by applying Lemma 12 with $M = N$, $M = \frac{C}{2} N \log N$, $M = C(1-\delta) N \log N$, and $M = C(1+\delta) N \log N$, we get

$$\begin{aligned} \limsup_{N\to\infty} (N \log N) E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] &\leqslant \limsup_{N\to\infty} \left(\frac{128 C'' B}{C \log N} + 2\varepsilon + \frac{1}{1-\delta}\right) \\ &= 2\varepsilon + \frac{1}{1-\delta} \end{aligned}$$

and

$$\liminf_{N\to\infty} (N \log N) E\left[\frac{(X_1)_2}{S_N^2} 1_{\{S_N \geqslant N\}}\right] \geqslant \frac{1-\varepsilon}{1+\delta}.$$

The lemma follows by taking $\delta, \varepsilon \to 0$. □

Lemma 17. *If* (11) *holds with* $a = 1$, *then for all* $x \in (0, 1)$,

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) = \int_x^1 y^{-2} \, \mathrm{d}y.$$

Proof. Let $0 < \delta < 1/2$, and let $\varepsilon > 0$. As shown in the proof of Lemma 16, Eqs. (54) and (55) hold for large enough N. Define the events D_1, D_2, D_3, and D_4 as in the proof of Lemma 16. Then, for all $x \in (0,1)$,

$$P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \leqslant P(D_1)P\left(\frac{X_1}{N} \geqslant x\right) + P(D_2)P\left(\frac{X_1}{X_1 + \frac{C}{2} N \log N} \geqslant x\right)$$
$$+ P\left(\frac{X_1}{X_1 + C(1-\delta)N \log N} \geqslant x\right)$$
$$\leqslant \frac{32C''B}{C^2(\log N)^2} P(X_1 \geqslant Nx) + \varepsilon P\left(X_1 \geqslant \frac{xCN \log N}{2(1-x)}\right)$$
$$+ P\left(X_1 \geqslant \frac{xC(1-\delta)N \log N}{1-x}\right).$$

Also,

$$P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \geqslant P(D_4)P\left(\frac{X_1}{X_1 + C(1+\delta)N \log N} \geqslant x\right)$$
$$\geqslant (1-\varepsilon)P\left(X_1 \geqslant \frac{xC(1+\delta)N \log N}{1-x}\right).$$

Using (11) and Lemma 16,

$$\limsup_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right)$$
$$\leqslant \limsup_{N\to\infty} (N \log N)$$
$$\left[\frac{32C''B}{C(\log N)^2}(Nx)^{-1} + \varepsilon C\left(\frac{2(1-x)}{xCN \log N}\right) + C\left(\frac{1-x}{xC(1-\delta)N \log N}\right)\right]$$
$$= \frac{2(1-x)\varepsilon}{x} + \frac{1}{1-\delta}\left(\frac{1-x}{x}\right). \tag{56}$$

Also,

$$\liminf_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) \geqslant (N \log N)(1-\varepsilon)C\left(\frac{1-x}{xC(1+\delta)N \log N}\right)$$
$$= \frac{1-\varepsilon}{1+\delta}\left(\frac{1-x}{x}\right). \tag{57}$$

Combining (56) and (57) and letting $\delta, \varepsilon \to 0$, we get

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x\right) = \frac{1-x}{x} = \int_x^1 y^{-2}\, dy,$$

as claimed. □

Lemma 18. *If* (11) *holds with* $a=1$, *then*

$$\lim_{N\to\infty}\frac{E[(v_{1,N})_2(v_{2,N})_2]}{N^2c_N}=0.$$

Proof. As in the proof of Lemma 15, it suffices to prove (48). By the argument used to prove (55), there exists a constant K such that

$$P\left(X_3+\cdots+X_N\leqslant\frac{C}{2}N\log N\right)\leqslant\frac{K}{(\log N)^2}.$$

Therefore,

$$\begin{aligned}
&E\left[\frac{(X_1)_2(X_2)_2}{S_N^4}1_{\{S_N\geqslant N\}}\right]\\
&\quad\leqslant\frac{K}{(\log N)^2}E\left[\left(\frac{(X_1)_2}{\max\{X_1^2,N^2\}}\right)\left(\frac{(X_2)_2}{\max\{X_2^2,N^2\}}\right)\right]\\
&\qquad+E\left[\left(\frac{(X_1)_2}{\max\{X_1^2,(\frac{C}{2}N\log N)^2\}}\right)\left(\frac{(X_2)_2}{\max\{X_2^2,(\frac{C}{2}N\log N)^2\}}\right)\right]\\
&\quad\leqslant\frac{K}{(\log N)^2}\left(4E\left[\frac{(X_1)_2}{(X_1+N)^2}\right]\right)^2+\left(4E\left[\frac{(X_1)_2}{(X_1+\frac{C}{2}N\log N)^2}\right]\right)^2.
\end{aligned}\tag{58}$$

Using (58) and Lemmas 12 and 16, we get

$$\begin{aligned}
&\limsup_{N\to\infty}\frac{N^2}{c_N}E\left[\frac{(X_1)_2(X_2)_2}{S_N^4}1_{\{S_N\geqslant N\}}\right]\\
&\quad\leqslant\limsup_{N\to\infty}(N^2\log N)\left(\frac{K}{(\log N)^2}\frac{16C^2}{N^2}+\frac{16C^2}{(\frac{C}{2}N\log N)^2}\right)=0,
\end{aligned}$$

and (48) follows. □

We now combine the results of the previous six lemmas to prove the desired convergence. We first state one additional lemma regarding the tail probabilities of the hypergeometric distribution. Recall that the hypergeometric distribution with parameters (N,R,n) is the distribution of the number of red balls drawn, when n balls are chosen without replacement from an urn containing N balls, R of which are red. The following bound appears in Chvátal (1979).

Lemma 19. *Suppose* X *has a hypergeometric distribution with parameters* (N,R,n), *where* $n\leqslant N$. *Let* $\varepsilon>0$. *Then*

$$P\left(X\geqslant\left(\frac{R}{N}+\varepsilon\right)n\right)\leqslant e^{-2\varepsilon^2n}$$

and

$$P\left(X \leqslant \left(\frac{R}{N} - \varepsilon\right) n\right) \leqslant e^{-2\varepsilon^2 n}.$$

Proof of part (c) of Theorem 4. We need to verify the three conditions of Proposition 3. When $1 < a < 2$, condition (8) can be deduced from Lemma 13. When $a = 1$, (8) follows from Lemma 16. Condition (9) is a consequence of Lemma 15 when $1 < a < 2$ and Lemma 18 when $a = 1$. It remains to show that if (11) holds with $1 \leqslant a < 2$, then

$$\lim_{N\to\infty} \frac{N}{c_N} P(v_{1,N} > Nx) = \int_x^1 y^{-2} \Lambda(\mathrm{d}y)$$

$$= \frac{1}{B(2-a,a)} \int_x^1 y^{-1-a}(1-y)^{a-1}\,\mathrm{d}y \tag{59}$$

for all $x \in (0,1)$.

Fix $x \in (0,1)$. Note that

$$\lim_{N\to\infty} \frac{N}{c_N} P(v_{1,N} > Nx) = \lim_{N\to\infty} \frac{N}{c_N} E[P(v_{1,N} > Nx | X_1, \ldots, X_N)]$$

$$= \lim_{N\to\infty} \frac{N}{c_N} E[P(v_{1,N} > Nx | X_1, \ldots, X_N) 1_{\{S_N \geqslant N\}}], \tag{60}$$

where the second equality holds because

$$\lim_{N\to\infty} \frac{N}{c_N} P(S_N < N) = 0$$

by Lemma 5 and Eq. (18) of Lemma 6. Let $0 < \varepsilon < x$. On $\{S_N \geqslant N\}$, the conditional distribution of $v_{1,N}$ given $X_1, \ldots, X_N$ is hypergeometric with parameters (S_N, X_1, N). Therefore, by Lemma 19,

$$\lim_{N\to\infty} \frac{N}{c_N} E[P(v_{1,N} > Nx | X_1, \ldots, X_N) 1_{\{S_N \geqslant N\}} 1_{\{X_1/S_N \leqslant x - \varepsilon\}}] = 0 \tag{61}$$

and

$$\lim_{N\to\infty} \frac{N}{c_N} E[P(v_{1,N} \leqslant Nx | X_1, \ldots, X_N) 1_{\{S_N \geqslant N\}} 1_{\{X_1/S_N \geqslant x + \varepsilon\}}] = 0. \tag{62}$$

By (60) and (61),

$$\limsup_{N\to\infty} \frac{N}{c_N} P(v_{1,N} > Nx)$$

$$\leqslant \limsup_{N\to\infty} \frac{N}{c_N} E[P(v_{1,N} > Nx | X_1, \ldots, X_N) 1_{\{S_N \geqslant N\}} 1_{\{X_1/S_N \geqslant x - \varepsilon\}}]$$

$$\leqslant \limsup_{N\to\infty} \frac{N}{c_N} P\left(\{S_N \geqslant N\} \cap \left\{\frac{X_1}{S_N} \geqslant x - \varepsilon\right\}\right)$$

$$= \limsup_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x - \varepsilon\right). \tag{63}$$

By (60) and (62),

$$\liminf_{N\to\infty} \frac{N}{c_N} P(v_{1,N} > Nx) \geqslant \liminf_{N\to\infty} \frac{N}{c_N} P\left(\{S_N \geqslant N\} \cap \left\{\frac{X_1}{S_N} \geqslant x + \varepsilon\right\}\right)$$

$$= \liminf_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_1}{S_N} 1_{\{S_N \geqslant N\}} \geqslant x + \varepsilon\right). \tag{64}$$

By combining equations (63) and (64) with Lemmas 14 and 17, and then letting $\varepsilon \to 0$, we obtain (59). □

4. Convergence to the Ξ-coalescent when $0 < a < 1$

In this section, we consider the case in which $0 < a < 1$ and prove part (d) of Theorem 4. Define $Y_{1,N} \geqslant Y_{2,N} \geqslant \cdots \geqslant Y_{N,N}$ by ranking in decreasing order the values of $N^{-1/a}X_1, \ldots, N^{-1/a}X_N$. Let $Z_1 \geqslant Z_2 \geqslant \ldots$ be the ranked points of a Poisson point process on $(0, \infty)$ with characteristic measure Λ_a, where $\Lambda_a((x,\infty)) = Cx^{-a}$ for all $x > 0$. For all j, let $W_j = Z_j / \sum_{i=1}^{\infty} Z_i$. Recall from the introduction that the sequence $(W_1, W_2, \ldots)$ has the Poisson–Dirichlet distribution with parameters $(a, 0)$.

Lemma 20. *For all positive integers j, we have*

$$(Y_{1,N}, \ldots, Y_{j,N}) \to_{\mathrm{d}} (Z_1, \ldots, Z_j)$$

as $N \to \infty$.

Proof. Fix positive real numbers $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_j$, and take $x_0 = \infty$. For $1 \leqslant i \leqslant j$, let $L_i^N = \#\{k\colon x_i \leqslant Y_{k,N} < x_{i-1}\}$ and $L_i = \#\{k\colon x_i \leqslant Z_k < x_{i-1}\}$, where $\#S$ denotes the cardinality of the set S. Note that L_i is the number of points of the Poisson process between x_i and x_{i-1}. Thus, $L_1, \ldots, L_j$ are independent Poisson random variables, and

$$E[L_i] = \Lambda_a((x_i, x_{i-1})) = Cx_i^{-a} - Cx_{i-1}^{-a}.$$

Also, $(L_1^N, \ldots, L_j^N, N - L_1^N - \cdots - L_j^N)$ has a multinomial distribution with parameters $(N; p_1, \ldots, p_j, p)$, where $p_i = P(x_i \leqslant Y_{1,N} < x_{i-1})$ and $p = 1 - p_1 - \cdots - p_j = 1 - P(Y_1 \geqslant x_j)$. Using $\sim$ to denote that the ratio of the two sides tends to 1 as $N \to \infty$, we have

$$p_i = P(X_1 \geqslant N^{1/a}x_i) - P(X_1 \geqslant N^{1/a}x_{i-1}) \sim N^{-1}(Cx_i^{-a} - Cx_{i-1}^{-a})$$

for $1 \leqslant i \leqslant j$ and

$$p^{N-n_1-\cdots-n_j} = [1 - P(X_1 \geqslant N^{1/a}x_j)]^{N-n_1-\cdots-n_j}$$

$$\sim (1 - CN^{-1}x_j^{-a})^N \sim \mathrm{e}^{-Cx_j^{-a}}$$

for all nonnegative integers $n_1,\ldots,n_j$. Therefore, for all nonnegative integers $n_1,\ldots,n_j$,

$$\begin{aligned} P(L_1^N=n_1,\ldots,L_j^N=n_j) &= \frac{(N)_{n_1+\cdots+n_j}}{n_1!\ldots n_j!}\,p_1^{n_1}\ldots p_j^{n_j}\,p^{N-n_1-\cdots-n_j} \\ &\sim \frac{N^{n_1+\cdots+n_j}}{n_1!\ldots n_j!}\left(\prod_{i=1}^{j} N^{-n_i}(Cx_i^{-a}-Cx_{i-1}^{-a})^{n_i}\right)\mathrm{e}^{-Cx_j^{-a}} \\ &= \prod_{i=1}^{j}\frac{\mathrm{e}^{-(Cx_i^{-a}-Cx_{i-1}^{-a})}(Cx_i^{-a}-Cx_{i-1}^{-a})^{n_i}}{n_i!} \\ &= P(L_1=n_1,\ldots,L_j=n_j). \end{aligned} \tag{65}$$

Note that $Y_{i,N}\geqslant x_i$ if and only if $L_1^N+\cdots+L_i^N\geqslant i$ and $Z_i\geqslant x_i$ if and only if $L_1+\cdots+L_i\geqslant i$. It thus follows from (65) that

$$\begin{aligned} \lim_{N\to\infty} P(Y_{1,N}\geqslant x_1,\ldots,Y_{j,N}\geqslant x_j) &= \lim_{N\to\infty} P(L_i^N\geqslant i \text{ for } 1\leqslant i\leqslant j) \\ &= P(L_i\geqslant i \text{ for } 1\leqslant i\leqslant j) \\ &= P(Z_1\geqslant x_1,\ldots,Z_j\geqslant x_j). \end{aligned}$$

This is enough to establish that as $N\to\infty$, $(Y_{1,N},\ldots,Y_{j,N})$ converges weakly to $(Z_1,\ldots,Z_j)$ (see Section 2 of Billingsley, 1999). □

Lemma 21. *For all positive integers j,*

$$\left(Y_{1,N},\ldots,Y_{j,N},\sum_{i=j+1}^{N}Y_{i,N}\right)\to_{\mathrm{d}}\left(Z_1,\ldots,Z_j,\sum_{i=j+1}^{\infty}Z_i\right)$$

as $N\to\infty$.

Proof. For all Borel subsets A of $\mathbb{R}^{j+1}$ and all $r>0$, let $A^r=\{x\in\mathbb{R}^{j+1}\colon |x-y|<r$ for some $y\in A\}$. Let d be the Prohorov metric on the set of probability measures on $\mathbb{R}^{j+1}$, defined by $d(P,Q)=\inf\{r>0\colon P(A)\leqslant Q(A^r)+r$ and $Q(A)\leqslant P(A^r)+r$ for all Borel sets $A\}$. Convergence in the Prohorov metric is equivalent to weak convergence (see Chapter 6 of Billingsley, 1999).

Let P_M be the distribution of the random vector $(Z_1,\ldots,Z_j,\sum_{i=j+1}^{M}Z_i)$, and let P be the distribution of $(Z_1,\ldots,Z_j,\sum_{i=j+1}^{\infty}Z_i)$. Also, let $Q_{M,N}$ denote the distribution of $(Y_{1,N},\ldots,Y_{j,N},\sum_{i=j+1}^{M}Y_{i,N})$, and let Q_N be the distribution of $(Y_{1,N},\ldots,Y_{j,N},\sum_{i=j+1}^{N}Y_{i,N})$. Let $\varepsilon>0$. Since $(Z_1,\ldots,Z_j,\sum_{i=j+1}^{M}Z_i)$ converges almost surely to $(Z_1,\ldots,Z_j,\sum_{i=j+1}^{\infty}Z_i)$ as $M\to\infty$, there exists M_1 such that $d(P_M,P)<\varepsilon/3$ for all $M\geqslant M_1$. Let $\delta=[\varepsilon^2(1-a)/(18C'')]^{1/(1-a)}$. Choose $M\geqslant M_1$ such that $P(Z_M\geqslant\delta/2)<\varepsilon/12$. Note that $Y_{M,N}$ converges weakly to Z_M as $N\to\infty$ by Lemma 20, so there exists a positive integer N_1 such that for all $N\geqslant N_1$, we have $P(Y_{M,N}\geqslant\delta)<\varepsilon/6$. Also, we have

$$E\left[\sum_{i=1}^{N} Y_{1,N} 1_{\{Y_{i,N} \leqslant \delta\}}\right] = N^{1-1/a} E[X_1 1_{\{X_1 \leqslant N^{1/a}\delta\}}]$$

$$= N^{1-1/a} \sum_{k=1}^{\infty} P(X_1 1_{\{X_1 \leqslant N^{1/a}\delta\}} \geqslant k)$$

$$\leqslant N^{1-1/a} \sum_{k=1}^{\lfloor N^{1/a}\delta \rfloor} P(X_1 \geqslant k) \leqslant C'' N^{1-1/a} \sum_{k=1}^{\lfloor N^{1/a}\delta \rfloor} k^{-a}$$

$$\leqslant C'' N^{1-1/a} \int_0^{N^{1/a}\delta} x^{-a}\, \mathrm{d}x = \frac{C''}{1-a} \delta^{1-a}. \tag{66}$$

By (66) and Markov's inequality,

$$P\left(\sum_{i=M+1}^{N} Y_{i,N} \geqslant \frac{\varepsilon}{3}\right) \leqslant P(Y_{M,N} \geqslant \delta) + P\left(\sum_{i=1}^{N} Y_{i,N} 1_{\{Y_{i,N} \leqslant \delta\}} \geqslant \frac{\varepsilon}{3}\right)$$

$$\leqslant \frac{\varepsilon}{6} + \frac{3}{\varepsilon}\left(\frac{C''}{1-a}\delta^{1-a}\right) = \frac{\varepsilon}{3}.$$

Therefore, for $N \geqslant N_1$, we have $d(Q_{M,N}, Q_N) \leqslant \varepsilon/3$. It follows from Lemma 20 that $(Y_{1,N}, \ldots, Y_{M,N}) \to_{\mathrm{d}} (Z_1, \ldots, Z_M)$, which means there exists N_2 such that $d(Q_{M,N}, P_M) < \varepsilon/3$ for all $N \geqslant N_2$. Thus, for $N \geqslant \max\{N_1, N_2\}$, we have

$$d(Q_N, P) \leqslant d(Q_N, Q_{M,N}) + d(Q_{M,N}, P_M) + d(P_M, P) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

which proves the lemma. □

Lemma 22. *As $N \to \infty$, we have*

$$\left(\frac{Y_{1,N}}{Y_{1,N} + \cdots + Y_{N,N}}, \ldots, \frac{Y_{N,N}}{Y_{1,N} + \cdots + Y_{N,N}}, 0, 0, \ldots\right) \to_{\mathrm{d}} (W_1, W_2, \ldots).$$

Proof. Fix $j \in \mathbb{N}$. The function

$$(x_1, \ldots, x_{j+1}) \mapsto (x_1/(x_1 + \cdots + x_{j+1}), \ldots, x_j/(x_1 + \cdots + x_{j+1}))$$

is continuous except on $\{x_1 + \cdots + x_{j+1} = 0\}$. Since $P(\sum_{i=1}^{\infty} Z_i = 0) = 0$, it follows from Lemma 21 and the Mapping Theorem (Theorem 2.7 of Billingsley, 1999) that

$$\left(\frac{Y_{1,N}}{Y_{1,N} + \cdots + Y_{N,N}}, \ldots, \frac{Y_{j,N}}{Y_{1,N} + \cdots + Y_{N,N}}\right) \to_{\mathrm{d}} \left(\frac{Z_1}{\sum_{i=1}^{\infty} Z_i}, \ldots, \frac{Z_j}{\sum_{i=1}^{\infty} Z_i}\right)$$
$$= (W_1, \ldots, W_j). \tag{67}$$

For sequences in Δ, weak convergence of a finite number of coordinates implies weak convergence of the sequence (see chapter 4 of Billingsley, 1999), so the lemma follows from (67). $\square$

Lemma 23. *Suppose* (11) *holds with* $0 < a < 1$. *For all* $k_1, \ldots, k_r \geqslant 2$,

$$\lim_{N\to\infty} N^r E\left[\frac{(X_1)_{k_1}\ldots(X_r)_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right] = \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{\infty} E[W_{i_1}^{k_1}\ldots W_{i_r}^{k_r}].$$

Proof. We have

$$N^r E\left[\frac{(X_1)_{k_1}\ldots(X_r)_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right] \sim \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{N} E\left[\frac{(X_{i_1})_{k_1}\ldots(X_{i_r})_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right], \tag{68}$$

where $\sim$ means that the ratio of the two sides approaches 1 as $N \to \infty$. Also,

$$\sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{N} E\left[\frac{(X_{i_1})_{k_1}\ldots(X_{i_r})_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}} 1_{\{X_{i_j} \geqslant N^{1/4} \text{ for } j=1,\ldots,r\}}\right]$$

$$\sim \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{N} E\left[\frac{X_{i_1}^{k_1}\ldots X_{i_r}^{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}} 1_{\{X_{i_j} \geqslant N^{1/4} \text{ for}=1,\ldots,r\}}\right]. \tag{69}$$

Note that

$$\sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{N} E\left[\frac{X_{i_1}^{k_1}\ldots X_{i_r}^{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}} 1_{\{X_{i_1} \leqslant N^{1/4}\}}\right]$$

$$\leqslant \sum_{i_1,\ldots,i_r=1}^{N} E\left[\left(\frac{N^{1/4}}{N}\right)^{k_1}\left(\frac{X_{i_2}}{S_N}\right)^{k_2}\cdots\left(\frac{X_{i_r}}{S_N}\right)^{k_r}\right]$$

$$\leqslant \left(\frac{N^{1/4}}{N}\right)^2 \sum_{i_1=1}^{N} E\left[\sum_{i_2,\ldots,i_r=1}^{N}\left(\frac{X_{i_2}}{S_N}\right)\cdots\left(\frac{X_{i_r}}{S_N}\right)\right]$$

$$\leqslant \frac{1}{N^{3/2}}\sum_{i_1=1}^{N} 1 = \frac{1}{N^{1/2}}, \tag{70}$$

and the same result would hold with $1_{\{X_{i_j} \leqslant N^{1/4}\}}$ in the expectation on the left-hand side for any $j = 2, \ldots, r$ in place of $1_{\{X_{i_1} \leqslant N^{1/4}\}}$. Furthermore, Lemma 5 implies that

$$\lim_{N\to\infty} \sum_{i_1,\dots,i_r=1}^{N} E\left[\frac{X_{i_1}^{k_1}\dots X_{i_r}^{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N<N\}}\right] = 0. \tag{71}$$

Combining (68)–(71), we get

$$\lim_{N\to\infty} N^r E\left[\frac{(X_1)_{k_1}\dots(X_r)_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N\geqslant N\}}\right]$$

$$= \lim_{N\to\infty} \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{N} E\left[\frac{X_{i_1}^{k_1}\dots X_{i_r}^{k_r}}{S_N^{k_1+\cdots+k_r}}\right]$$

$$= \lim_{N\to\infty} \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{N} E\left[\left(\frac{Y_{i_1,N}}{Y_{1,N}+\cdots+Y_{N,N}}\right)^{k_1}\dots\left(\frac{Y_{i_r,N}}{Y_{1,N}+\cdots+Y_{N,N}}\right)^{k_r}\right], \tag{72}$$

in the sense that if one of the limits exists, then so do the other two, and the values are equal. The function $f : \Delta \to \Delta$ defined by

$$f(x_1,x_2,\dots) = \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{\infty} x_{i_1}^{k_1}\dots x_{i_r}^{k_r}$$

is bounded and continuous (see the proof of Lemma 26 of Schweinsberg, 2000b). Thus, by Lemma 22,

$$\lim_{N\to\infty} \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{N} E\left[\left(\frac{Y_{i_1,N}}{Y_{1,N}+\cdots+Y_{N,N}}\right)^{k_1}\dots\left(\frac{Y_{i_r,N}}{Y_{1,N}+\cdots+Y_{N,N}}\right)^{k_r}\right]$$

$$= \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{\infty} E[W_{i_1}^{k_1}\dots W_{i_r}^{k_r}].$$

Combining this result with (72) yields the conclusion of the lemma. □

Before proceeding with the proof of part (d) of Theorem 4, we review some facts about the partition structures associated with Poisson–Dirichlet distributions. These partition structures were studied in Pitman (1995), and connections with excursions of Brownian motion and Bessel processes were explained in Pitman (1997).

Suppose $(V_j)_{j=1}^{\infty}$ has a Poisson–Dirichlet distribution with parameters $(\alpha, 0)$, where $0<\alpha<1$. Define random variables $U_1, U_2, \dots$ to be conditionally i.i.d. given $(V_j)_{j=1}^{\infty}$, with $P(U_j = N|V_1, V_2, \dots) = V_N$. Define a random partition Π_n of $\{1,\dots,n\}$ such that

i and j are in the same block of Π_n if and only if $U_i = U_j$. Let π be a partition of $\{1,\ldots,n\}$ containing k blocks of sizes $n_1,\ldots,n_k$. For real numbers x and a and nonnegative integers m, define $[x]_{0;a} = 1$ and $[x]_{m;a} = x(x+a)\ldots(x+(m-1)a)$ for $m \geqslant 1$. It was shown in Proposition 9 of Pitman (1995) that

$$P(\Pi_n = \pi) = p(n_1,\ldots,n_k) = \frac{\alpha^{k-1}(k-1)!}{(n-1)!}\prod_{i=1}^{k}[1-\alpha]_{n_i-1;1}. \tag{73}$$

The function $p(n_1,\ldots,n_k)$, which is called the exchangeable probability function, is a symmetric function of $n_1,\ldots,n_k$. As noted in Proposition 10 of Pitman (1995), we have the recursion

$$p(n_1,\ldots,n_k) = \sum_{j=1}^{k} p(n_1,\ldots,n_{j-1},n_j+1,n_{j+1},\ldots,n_k) + p(n_1,\ldots,n_k,1). \tag{74}$$

Fix $k_1,\ldots,k_r \geqslant 2$. Let π be the partition of $\{1,\ldots,n\}$ with the property that the first k_1 integers are in one block, the next k_2 are in another block, and so on. Conditional on $(V_j)_{j=1}^{\infty}$, the probability that $U_1 = \cdots = U_{k_1} = i_1$, $U_{k_1+1} = \cdots = U_{k_1+k_2} = i_2$, and so on is $V_{i_1}^{k_1}\ldots V_{i_r}^{k_r}$. By summing over the possible values of $i_1,\ldots,i_r$ and taking expectations, we see that

$$P(\Pi_n = \pi) = \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{\infty} E[V_{i_1}^{k_1}\ldots V_{i_r}^{k_r}]. \tag{75}$$

Proof of part (d) of Theorem 4. Since $(W_1, W_2, \ldots)$ has the Poisson–Dirichlet distribution with parameters $(a, 0)$, if we define the measures Θ_a and Ξ_a as in Theorem 4, then

$$E[W_{i_1}^{k_1}\ldots W_{i_r}^{k_r}] = \int_{\Delta} x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}\,\Theta_a(\mathrm{d}x)$$

for all $i_1,\ldots,i_r$. Combining this result with Lemma 23, we get

$$\begin{aligned}
&\lim_{N\to\infty} N^r E\left[\frac{(X_1)_{k_1}\ldots(X_r)_{k_r}}{S_N^{k_1+\cdots+k_r}} 1_{\{S_N \geqslant N\}}\right] \\
&\quad = \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{\infty} \int_{\Delta} x_{i_1}^{k_1}\ldots x_{i_r}^{k_r}\,\Theta_a(\mathrm{d}x) \\
&\quad = \int_{\Delta} \sum_{\substack{i_1,\ldots,i_r=1 \\ \text{all distinct}}}^{\infty} x_{i_1}^{k_1}\ldots x_{i_r}^{k_r} \Bigg/ \sum_{j=1}^{\infty} x_j^2\,\Xi_a(\mathrm{d}x).
\end{aligned} \tag{76}$$

By combining (76) with Lemma 6 and the remarks following Proposition 1, we see that to prove that the processes $(\Psi_{n,N}(m))_{m=0}^{\infty}$ defined in Theorem 4 converge as $N \to \infty$

to the restriction to $\{1,\dots,n\}$ of a discrete-time Ξ_a-coalescent, it remains only to verify that

$$\lim_{N\to\infty} c_N = c \tag{77}$$

for some $c > 0$.

Denote by $p(n_1,\dots,n_k)$ the exchangeable partition function defined above when $(V_j)_{j=1}^{\infty}$ has the Poisson–Dirichlet distribution with parameters $(a,0)$. Since $(W_j)_{j=1}^{\infty}$ has the Poisson–Dirichlet distribution with parameters $(a,0)$, it follows from (73) and (75) that

$$\begin{aligned} p(k_1,\dots,k_r) &= \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{\infty} E[W_{i_1}^{k_1}\dots W_{i_r}^{k_r}] \\ &= \frac{a^{r-1}(r-1)!}{(k_1+\cdots+k_r-1)!}\prod_{i=1}^{r}[1-a]_{k_i-1;1}. \end{aligned} \tag{78}$$

Combining this result with Lemmas 6 and 23, we get

$$\lim_{N\to\infty} c_N = \lim_{N\to\infty} NE\left[\frac{(X_1)_2}{S_N^2}1_{\{S_N\geqslant N\}}\right] = \sum_{i=1}^{\infty}E[W_i^2] = p(2) = 1-a > 0,$$

which proves (77).

We now verify the formula (13) for the transition probabilities. It follows from Proposition 1, Lemmas 6 and 23, and Eq. (78) that if $k_1,\dots,k_r\geqslant 2$ and $b=k_1+\cdots+k_r$, then

$$\begin{aligned} p_{b;k_1,\dots,k_r;0} &= \sum_{\substack{i_1,\dots,i_r=1\\ \text{all distinct}}}^{\infty} E[W_{i_1}^{k_1}\dots W_{i_r}^{k_r}] \\ &= p(k_1,\dots,k_r) = \frac{a^{r-1}(r-1)!}{(b-1)!}\prod_{i=1}^{r}[1-a]_{k_i-1;1}. \end{aligned}$$

Recall that to obtain a formula for $p_{b;k_1,\dots,k_r;s}$ when $s\neq 0$, we can use the recursion (3). Likewise, let $p_s(k_1,\dots,k_r)=p(k_1,\dots,k_r,1,\dots,1)$, where $k_1,\dots,k_r\geqslant 2$ and there are s ones on the right-hand side. Then, using the fact that $p(n_1,\dots,n_k)$ is a symmetric function of $n_1,\dots,n_k$, we can write (74) as

$$\begin{aligned} p_{s+1}(k_1,\dots,k_r) = {} & p_s(k_1,\dots,k_r) - \sum_{j=1}^{r} p_s(k_1,\dots,k_{j-1},k_j+1,k_{j+1},\dots,k_r) \\ & - s\,p_{s-1}(k_1,\dots,k_r,2). \end{aligned}$$

Thus, the transition probabilities and the exchangeable probability function satisfy the same recursion. We conclude that for all $r\geqslant 1$, $k_1,\dots,k_r\geqslant 2$, and $s\geqslant 0$ such that

$b = k_1 + \cdots + k_r + s$, we have

$$p_{b;\,k_1,\dots,k_r;\,s} = p_s(k_1,\dots,k_r) = \frac{a^{r+s-1}(r+s-1)!}{(b-1)!} \prod_{i=1}^{r} [1-a]_{k_i;1}. \tag{79}$$

The right-hand side of (79) is equivalent to the right-hand side of (13), so the proof of part (d) of Theorem 4 is now complete. □

Acknowledgements

The author thanks Rick Durrett for suggesting this problem, for helpful discussions, and for comments on earlier drafts of this work. He also thanks a referee for some suggestions.

References

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

Bertoin, J., Le Gall, J.-F., 2000. The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. Probab. Theory Related Fields 117, 249–266.

Bertoin, J., Pitman, J., 2000. Two coalescents derived from the ranges of stable subordinators. Electron. J. Probab. 5, 1–17.

Billingsley, P., 1999. Convergence of Probability Measures, 2nd Edition. Wiley, New York.

Bolthausen, E., Sznitman, A.-S., 1998. On Ruelle's probability cascades and an abstract cavity method. Comm. Math. Phys. 197, 247–276.

Cannings, C., 1974. The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. Adv. Appl. Probab. 6, 260–290.

Cannings, C., 1975. The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. Adv. Appl. Probab. 7, 264–282.

Chvátal, V., 1979. The tail of the hypergeometric distribution. Discrete Math. 25, 285–287.

Fristedt, B., Gray, L., 1997. A Modern Approach to Probability Theory. Birkhäuser, Boston.

Griffiths, R.C., Tavaré, S., 1994. Sampling theory for neutral alleles in a varying environment. Phil. Trans. Roy. Soc. London B. 344, 403–410.

Kingman, J.F.C., 1982a. The coalescent. Stochast. Process. Appl. 13, 235–248.

Kingman, J.F.C., 1982b. On the genealogy of large populations. In: Gani, J., Hannan, E.J. (Eds.), Essays in Statistical Science, (J. Appl. Probab. 19A, 27–43.) Applied Probability Trust, Sheffield.

Möhle, M., 1998. Robustness results for the coalescent. J. Appl. Probab. 35, 438–447.

Möhle, M., 2000. Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. Adv. Appl. Probab. 32, 983–993.

Möhle, M., 2002. The coalescent in population models with time-inhomogeneous environment. Stochast. Process. Appl. 97, 199–227.

Möhle, M., Sagitov, S., 1998. A characterization of ancestral limit processes arising in haploid population genetics models. Preprint. Johannes Gutenberg-Univ., Mainz.

Möhle, M., Sagitov, S., 2001. A classification of coalescent processes for haploid exchangeable population models. Ann. Probab. 29, 1547–1562.

Neveu, J., 1992. A continuous-state branching process in relation with the GREM model of spin glass theory. Rapport interne no. 267, Ecole Polytechnique.

Perman, M., Pitman, J., Yor, M., 1992. Size-biased sampling of Poisson point processes and excursions. Probab. Theory Related Fields 92, 21–39.

Pitman, J., 1995. Exchangeable and partially exchangeable random partitions. Probab. Theory. Related Fields 102, 145–158.

Pitman, J., 1997. Partition structures derived from Brownian motion and stable subordinators. Bernoulli 3, 79–96.

Pitman, J., 1999. Coalescents with multiple collisions. Ann. Probab. 27, 1870–1902.

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

Sagitov, S., 1999. The general coalescent with asynchronous mergers of ancestral lines. J. Appl. Probab. 36, 1116–1125.

Sagitov, S., 2002. Convergence to the coalescent with simultaneous multiple mergers. Preprint. Available via http://www.math.chalmers.se/∼serik/coal.html.

Schweinsberg, J., 2000a. A necessary and sufficient condition for the Λ-coalescent to come down from infinity. Electron. Comm. Probab. 5, 1–11.

Schweinsberg, J., 2000b. Coalescents with simultaneous multiple collisions. Electron. J. Probab. 5, 1–50.