The Annals of Probability
2001, Vol. 29, No. 4, 1547–1562

A CLASSIFICATION OF COALESCENT PROCESSES FOR HAPLOID EXCHANGEABLE POPULATION MODELS

BY MARTIN MÖHLE AND SERIK SAGITOV[1]

Johannes Gutenberg-Universität, Mainz and Chalmers and Göteborgs Universities, Göteborg

We consider a class of haploid population models with nonoverlapping generations and fixed population size N assuming that the family sizes within a generation are exchangeable random variables. A weak convergence criterion is established for a properly scaled ancestral process as $N \to \infty$. It results in a full classification of the coalescent generators in the case of exchangeable reproduction. In general the coalescent process allows for simultaneous multiple mergers of ancestral lines.

1. Introduction and motivation. Cannings (1974, 1975) introduced a certain class of haploid population models with fixed population size $N \in \mathbb{N} := \{1, 2, \ldots\}$ and nonoverlapping generations $r \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}$. As our interest is in ancestral population genetics, the generations are labelled backwards in time; that is, $r = 0$ is the current generation, $r = 1$ brings us one generation backward in time and so on. Each model in this class is characterized by a family of random variables $\{\nu_i^{(r)}\}$, $i \in \{1, \ldots, N\}$, $r \in \mathbb{N}$, where $\nu_i^{(r)}$ denotes the number of offspring of the ith individual alive in the rth generation. As the population size is fixed the condition

$$\nu_1^{(r)} + \cdots + \nu_N^{(r)} = N, \qquad r \in \mathbb{N} \tag{1}$$

has to be satisfied. Cannings assumed that the reproduction law does not change and is independent from generation to generation; that is:

(I) The offspring vectors $(\nu_1^{(r)}, \ldots, \nu_N^{(r)})$, $r \in \mathbb{N}$, are independent and identically distributed

and that the individuals present at a given generation have the same propensity to reproduce; that is:

(II) For each fixed $r \in \mathbb{N}$ the family sizes $\nu_1^{(r)}, \ldots, \nu_N^{(r)}$ are exchangeable.

For convenience the notation $\nu_i := \nu_i^{(1)}$, $i \in \{1, \ldots, N\}$ is used. Recall that (II) ensures that the distribution of the random vector $(\nu_{i_1}, \ldots, \nu_{i_k})$ with pairwise distinct indices depends only upon k and not upon the particular set of indices.

Received February 1999; revised November 2000.
[1]Supported by the Bank of Sweden Tercentenary Foundation project Dependence and Interaction in Stochastic Population Dynamics.
AMS 2000 *subject classifications*. Primary 92D25, 60J70; secondary 92D15, 60F17.
Key words and phrases. Ancestral processes, coalescent, exchangeability, generator, neutrality, population genetics, weak convergence.

We are interested in the asymptotics of the genealogical structure in such a population in the spirit of Kingman (1982a, b, c). Fix $n \le N$ and sample n individuals at random from the 0th generation. Let $\mathscr{R}_r$ denote the equivalence relation which contains the pair (i, j) if and only if the ith and the jth individual of this sample have a common ancestor in the rth generation backwards in time. The assumption (I) ensures that the so-called *ancestral process* $(\mathscr{R}_r)_{r\in\mathbb{N}_0}$ is a time homogeneous Markov chain with the state space

$$\mathscr{E}_n = \text{the set of all equivalence relations on } \{1, \dots, n\}$$

and the initial value $\mathscr{R}_0 = \xi_0$,

$$\xi_0 = \{\text{all equivalence classes are singletons } \{i\},\ i = 1, \dots, n\}. \tag{2}$$

Since the transition probability $p_{\xi\eta} := P(\mathscr{R}_r = \eta \,|\, \mathscr{R}_{r-1} = \xi)$ is equal to zero for $\xi \not\subseteq \eta$, the focus will be on such pairs $\xi, \eta \in \mathscr{E}_n$ that $\xi \subseteq \eta$. The relation $\xi \subseteq \eta$ implies that every equivalence class of η is either a union of several equivalence classes of ξ or coincides with an equivalence class of ξ. Reflecting this observation, write a for the number of η-classes and $b = b_1 + \cdots + b_a$ for the number of ξ-classes, where $b_1 \ge \cdots \ge b_g \ge 2$ are ordered group sizes of merging ξ-classes and $b_{g+1} = \cdots = b_a = 1$. Notice that $g = 0$ if $\xi = \eta$ and $g \ge 1$ if $\xi \subset \eta$. With this notation it follows from the assumption (II) via a combinatorial "putting balls into boxes" argument that the transition probability is given by

$$p_{\xi\eta} = \frac{1}{(N)_b} \sum_{\substack{i_1,\dots,i_a=1 \\ \text{all distinct}}}^{N} \mathrm{E}((\nu_{i_1})_{b_1} \cdots (\nu_{i_a})_{b_a}) = \frac{(N)_a}{(N)_b} \mathrm{E}((\nu_1)_{b_1} \cdots (\nu_a)_{b_a}), \tag{3}$$

where $(N)_b := N(N-1)\cdots(N-b+1)$.

Let c_N denote the probability that two individuals, chosen randomly without replacement from some generation, have a common ancestor one generation backwards in time, that is,

$$c_N := \frac{1}{(N)_2} \sum_{i=1}^{N} \mathrm{E}((\nu_i)_2) = \frac{\mathrm{E}((\nu_1)_2)}{N-1} = \frac{\mathrm{Var}(\nu_1)}{N-1} = 1 - \mathrm{E}(\nu_1\nu_2), \tag{4}$$

where the formula $\mathrm{E}(\nu_1) = 1$ has been used. This probability, called the *coalescence probability* is of fundamental interest in the coalescent theory as c_N^{-1} is the proper time scale to get convergence to the coalescent [it is only natural to assume that $c_N > 0$ for sufficiently large N because the case $c_N = 0$ corresponds to the trivial reproduction law $P(\nu_1 = 1, \dots, \nu_N = 1) = 1$]. The coalescence probability is also important as it is directly connected via $c_N = 1 - \lambda_2$ to the eigenvalue $\lambda_2 := \mathrm{E}(\nu_1\nu_2)$ of the transition matrix of the descendant process, that is, the genealogical process looking forwards in time [see Cannings (1974)].

Kingman (1982b) has shown that given $\sup_N \mathrm{E}(\nu_1^k) < \infty$, $k \ge 2$ (this holds, e.g., for the Moran model and the Wright–Fisher model) the convergence of finite-dimensional distributions

$$(\mathscr{R}_{[t/c_N]})_{t\geq 0} \to (R_t)_{t\geq 0}, \qquad N \to \infty \tag{5}$$

takes place. The limit process $(R_t)_{t\geq 0}$, the so-called (standard) n-coalescent process, is a continuous time Markov chain with state space $\mathscr{E}_n$, initial state (2) and infinitesimal generator $Q = (q_{\xi\eta})_{\xi,\eta\in\mathscr{E}_n}$ given by

$$q_{\xi\eta} := \begin{cases} -|\xi|(|\xi|-1)/2, & \text{if } \xi = \eta, \\ 1, & \text{if } \xi \prec \eta, \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

where $\xi \prec \eta$ means that $\xi \subset \eta$, $g = 1$ and $b_1 = 2$; that is, during the transition exactly two ancestral lines merge together.

The convergence (5) is based on the asymptotic formula

$$p_{\xi\eta} = \delta_{\xi\eta} + c_N\, q_{\xi\eta} + o(c_N), \qquad \xi, \eta \in \mathscr{E}_n, \tag{7}$$

which is often written in matrix notation

$$P_N = I + c_N\, Q + o(c_N), \tag{8}$$

where $P_N := (p_{\xi\eta})_{\xi,\eta\in\mathscr{E}_n}$ denotes the transition matrix of the ancestral process. In Möhle (1998, 1999) Kingman's result was extended beyond the framework of exchangeable population models and it was shown that (5) holds even in the sense of the weak convergence of stochastic processes.

Recently a richer class of the coalescent generators Q allowing for multiple mergers with $g = 1$ and $b_1 \geq 2$ was found independently by Pitman (1999) and Sagitov (1999). A member Q of this class is characterized by a probability measure F on the unit interval $[0, 1]$ via the formula

$$q_{\xi\eta} = \begin{cases} -\displaystyle\int_{[0,1]} \frac{1-(1-x)^{b-1}(1-x+bx)}{x^2}\, F(dx), & \text{if } \xi = \eta, \\ \displaystyle\int_{[0,1]} x^{b_1-2}(1-x)^{b-b_1}\, F(dx), & \text{if } \xi \subset \eta,\ g = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

For example if F is uniformly distributed on $[0, 1]$ then the generator Q corresponds to the clustering process recently constructed by Bolthausen and Sznitman (1998) in the context of Ruelle's probability cascades. The generator (9) is equal to the generator (6) of the standard n-coalescent if and only if the probability measure $F = \delta_0$ is concentrated in zero.

The present paper is based upon an instrumental development of Möhle and Sagitov (1998) of the method of Sagitov (1999). We establish a general coalescent structure allowing for simultaneous mergers of ancestral lines ($g \geq 1$). Due to the main result of this paper, Theorem 2.1, in general, a coalescent generator $Q = (q_{\xi\eta})_{\xi,\eta\in\mathscr{E}_n}$ is characterized by a sequence of symmetric measures F_r, $r \in \mathbb{N}$, where each F_r is concentrated on the simplex

$$\Delta_r := \{(y_1, \ldots, y_r) \in [0, 1]^r \mid y_1 + \cdots + y_r \leq 1\}$$

with

$$1 = F_1(\Delta_1) \geq F_2(\Delta_2) \geq \cdots. \tag{10}$$

If $\xi \subset \eta$, then the corresponding entry has the form

$$q_{\xi\eta} = \sum_{r=g}^{[(a+g)/2]} \int_{\Delta_r} x_1^{b_1-2} \cdots x_g^{b_g-2} T^{(r)}_{g,\,a-g}(x_1, \dots, x_r)\, F_r(dx_1, \dots, dx_r). \tag{11}$$

Here the set of polynomials

$$T^{(r)}_{j,s}(x_1, \dots, x_r), \qquad 1 \leq j \leq r,\ \ r \in \mathbb{N},\ \ s \in \mathbb{N}_0 \tag{12}$$

is defined explicitly by the formula

$$T^{(r)}_{r,\,s}(x_1, \dots, x_r) = (1 - x_1 - \cdots - x_r)^s \tag{13}$$

and

$$\begin{aligned} &T^{(r)}_{r-j,\,s}(x_1, \dots, x_r) \\ &\quad = (-1)^{j+1} \sum_{i_j=2j-1}^{i_{j+1}-2} \cdots \sum_{i_1=1}^{i_2-2} \prod_{k=0}^{j} i_k \left(1 - \sum_{i=1}^{r-k} x_i\right)^{i_{k+1}-i_k-2}, \qquad j = 1, \dots, r, \end{aligned} \tag{14}$$

where $i_0 = -1$ and $i_{j+1} = s + 1$. Note that this implies $T^{(r)}_{r-j,\,s} \equiv 0$ for $s < 2j$. The diagonal entries of Q are given by

$$q_{\xi\xi} = -\sum_{j=1}^{b-1} j \sum_{r=1}^{[(j+1)/2]} \int_{\Delta_r} T^{(r)}_{1,\,j-1}(x_1, \dots, x_r) F_r(dx_1, \dots, dx_r). \tag{15}$$

Another representation for the diagonal entries is given at the end of Section 3. Observe that in the case when $F_2(\Delta_2) = 0$ the formulas (11), (15) and $T^{(1)}_{1,\,s}(x) = (1-x)^s$ bring us back to (9) with $F = F_1$.

The article is organized as follows. Section 2 presents the results and gives some intuitive explanations, while the rigorous proofs are given in Sections 3 and 4. A concrete example connected to the Wright–Fisher model is presented in Section 5. The last section gives some historical perspectives and puts the coalescent with simultaneous multiple collisions in the context of recent research and developments in coalescent theory.

2. A weak convergence criterion.

This section presents the main result of the paper, Theorem 2.1, which shows that the formulas (11) and (15) fully describe the class of coalescent generators for the population models with exchangeable reproduction. The central condition of Theorem 2.1 requires the existence of the limits

$$\lim_{N\to\infty} \frac{\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j})}{N^{k_1+\cdots+k_j-j} c_N} = \phi_j(k_1, \dots, k_j) \tag{16}$$

for all $j \in \mathbb{N}$ and $k_1 \geq \cdots \geq k_j \geq 2$. To justify the denominator in the l.h.s. of (16), turn to the chain of inequalities

$$
\begin{aligned}
&\sum_{\substack{i_1,\dots,i_j=1\\ \text{all distinct}}}^{N} (\nu_{i_1})_{k_1} \cdots (\nu_{i_j})_{k_j} \\
&\quad \leq \sum_{\substack{i_1,\dots,i_l=1\\ \text{all distinct}}}^{N} (\nu_{i_1})_{m_1} \nu_{i_1}^{k_1-m_1} \cdots (\nu_{i_l})_{m_l} \nu_{i_l}^{k_l-m_l} \sum_{i_{l+1},\dots,i_j=1}^{N} \nu_{i_{l+1}}^{k_{l+1}} \cdots \nu_{i_j}^{k_j} \\
(17) \qquad &\quad \leq \sum_{\substack{i_1,\dots,i_l=1\\ \text{all distinct}}}^{N} (\nu_{i_1})_{m_1} N^{k_1-m_1} \cdots (\nu_{i_l})_{m_l} N^{k_l-m_l} (\nu_1 + \cdots + \nu_N)^{k_{l+1}+\cdots+k_j} \\
&\quad = \frac{N^{k_1+\cdots+k_j}}{N^{m_1+\cdots+m_l}} \sum_{\substack{i_1,\dots,i_l=1\\ \text{all distinct}}}^{N} (\nu_{i_1})_{m_1} \cdots (\nu_{i_l})_{m_l},
\end{aligned}
$$

where $l \leq j$, $k_1 \geq m_1 \geq 1, \dots, k_l \geq m_l \geq 1$. With $l = 1$ and $m_1 = 2$ it entails that in general

$$\limsup_{N\to\infty} \frac{\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j})}{N^{k_1+\cdots+k_j-j} c_N} \leq 1, \qquad k_1 \geq \cdots \geq k_j \geq 2.$$

Relation (17) implies also that the set of the limits (16) is monotone:

$$(18) \quad \phi_j(k_1,\dots,k_j) \leq \phi_l(m_1,\dots,m_l) \quad \text{whenever } j \geq l,\ k_1 \geq m_1,\dots,k_l \geq m_l.$$

THEOREM 2.1. *If the limits* (16) *exist for all* $j \in \mathbb{N}$ *and* $k_1 \geq \cdots \geq k_j \geq 2$, *then for each sample size* $n \in \mathbb{N}$ *the asymptotic formula* (8) *holds with* $Q = (q_{\xi\eta})_{\xi,\,\eta\in\mathscr{E}_n}$ *defined by* (11) *and* (15). *The corresponding symmetric measures* F_r, $r \in \mathbb{N}$ *are uniquely determined via their moments*

$$(19) \quad \int_{\Delta_r} x_1^{k_1-2} \cdots x_r^{k_r-2} F_r(dx_1,\dots,dx_r) = \phi_r(k_1,\dots,k_r), \qquad k_1 \geq \cdots \geq k_r \geq 2.$$

Conversely, if (8) *holds, then all the limits* (16), $j \in \mathbb{N}$, $k_1 \geq \cdots \geq k_j \geq 2$, *exist.*

Given (8) *and*

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

the ancestral process $(\mathscr{R}_r)_{r\in\mathbb{N}_0}$ *converges weakly to a discrete time Markov chain* $(R_r)_{r\in\mathbb{N}_0}$ *with the initial state* (2) *and the transition matrix* $I + cQ$.

In the case

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

the asymptotic formula (8) *implies the convergence* (5) *in the Skorohod sense, where the limit coalescent process* $(R_t)_{t\geq 0}$ *is a continuous time Markov chain with the initial state* (2) *and the transition matrix* e^{tQ}.

Condition (16) for an arbitrary fixed $j \in \mathbb{N}$ has the following two equivalent versions (cf. Section 4 for the proof). One version of (16) is just a restatement in terms of the central moments for the joint distribution of the family sizes

$$\lim_{N\to\infty} \frac{\mathrm{E}((\nu_1-1)^{k_1}\cdots(\nu_j-1)^{k_j})}{N^{k_1+\cdots+k_j-j}c_N} = \phi_j(k_1,\dots,k_j), \qquad k_1 \ge \cdots \ge k_j \ge 2. \tag{20}$$

The other equivalent version of (16) requires the existence of a symmetric measure F_j defined on the simplex Δ_j such that

$$\lim_{N\to\infty} \frac{\mathrm{E}((\nu_1)_2\cdots(\nu_j)_2)}{N^j c_N} = F_j(\Delta_j) \tag{21}$$

(which is obviously true for $j = 1$), and

$$\lim_{N\to\infty} \frac{N^j}{c_N} P(\nu_1 > Nx_1, \dots, \nu_j > Nx_j) = \int_{x_1}^1 \cdots \int_{x_j}^1 \frac{F_j(dy_1,\dots,dy_j)}{y_1^2\cdots y_j^2}, \tag{22}$$

holding for all points $(x_1,\dots,x_j)$ of continuity for the measure F_j.

Versions (21) and (22) bring the following picture of the asymptotic coalescent structure. Call *large* every family whose size is of order N. Obviously, every large family with a positive probability embraces two or more sampled ancestral lines (in other words, begets a multiple merger). Due to the condition

$$\lim_{N\to\infty} Nc_N^{-1} P(\nu_1 > Nx_1) = \int_{x_1}^1 y^{-2} F_1(dy),$$

a finite number of large families is encountered with a positive probability while scanning N generations in the population.

A large family caused by a multiple merger might, in a sense, trigger a chain reaction of mergers within the same generation. To see this, observe that the total number of families in a generation is equal to N and the relation

$$\lim_{N\to\infty} NP(\nu_2 > Nx_2 \,|\, \nu_1 > Nx_1) = \frac{\int_{x_1}^1 \int_{x_2}^1 y_1^{-2} y_2^{-2} F_2(dy_1, dy_2)}{\int_{x_1}^1 y^{-2} F_1(dy)}$$

indicates to the possibility that we might encounter another large family outside the initial one provided $F_2(\Delta_2) > 0$. Furthermore, if $F_3(\Delta_3) > 0$, the second large family leaves room for the third one:

$$\lim_{N\to\infty} NP(\nu_3 > Nx_3 \,|\, \nu_1 > Nx_1, \nu_2 > Nx_2)$$
$$= \frac{\int_{x_1}^1 \int_{x_2}^1 \int_{x_3}^1 y_1^{-2} y_2^{-2} y_3^{-2} F_3(dy_1, dy_2, dy_3)}{\int_{x_1}^1 \int_{x_2}^1 y_1^{-2} y_2^{-2} F_2(dy_1, dy_2)}$$

and so on. This imaginary chain reaction of mergers (there is no real order for the mergers happening within one generation) is bound to stop after a random number of rounds because the population of size N might host only a finite number of large families [given $F_{l+1}(\Delta_{l+1}) = 0$ this number of rounds never exceeds l].

REMARK. According to Theorem 2.1 we have $F_r(\Delta_r) = \phi_r(2,\dots,2)$ so that (10) follows from (18). Note that $F_{l+1}(\Delta_{l+1}) = 0$ in particular when the random variables $(\nu_1)_2\cdots(\nu_l)_2$ and $(\nu_{l+1})_2$ are not positively correlated, provided $\lim_{N\to\infty} c_N = 0$. Indeed in this case,

$$\mathrm{E}((\nu_1)_2\cdots(\nu_{l+1})_2) \le \mathrm{E}((\nu_1)_2\cdots(\nu_l)_2)\cdot\mathrm{E}((\nu_1)_2) \le \frac{N^{2l}c_N}{(N)_l}\cdot Nc_N$$

and hence

$$F_{l+1}(\Delta_{l+1}) = \phi_{l+1}(2,\dots,2) = \lim_{N\to\infty}\frac{\mathrm{E}((\nu_1)_2\cdots(\nu_{l+1})_2)}{N^{l+1}c_N} \le \lim_{N\to\infty} c_N = 0.$$

3. The proof of the criterion.

LEMMA 3.1. *If the limits* (16) *exist for some* $j\in\mathbb{N}$, *then there exists a measure* F_j *uniquely determined on the simplex* Δ_j *by its moments* (19).

PROOF. If $\phi_j(2,\dots,2) = 0$ then (19) implies $F_j(\Delta_j) = 0$. In the case $\phi_j(2,\dots,2) > 0$ we have $\mathrm{E}((\nu_1)_2\cdots(\nu_j)_2) > 0$ for sufficiently large N. Let $Y_{1,j},\dots,Y_{j,j}$ be the random variables with the joint distribution

$$P(Y_{1,j}=i_1,\dots,Y_{j,j}=i_j) := \frac{(i_1)_2\cdots(i_j)_2}{\mathrm{E}((\nu_1)_2\cdots(\nu_j)_2)}P(\nu_1=i_1,\dots,\nu_j=i_j), \tag{23}$$

where $i_1,\dots,i_j\in\{2,\dots,N\}$. The representation

$$\begin{aligned}\mathrm{E}\Big(Y_{1,j}^{k_1}\cdots Y_{j,j}^{k_j}\Big) &= \sum_{i_1,\dots,i_j}\frac{(i_1)^{k_1}\cdots(i_j)^{k_j}(i_1)_2\cdots(i_j)_2}{\mathrm{E}((\nu_1)_2\cdots(\nu_j)_2)}P(\nu_1=i_1,\dots,\nu_j=i_j)\\ &= \frac{\mathrm{E}((\nu_1^{k_1+2}-\nu_1^{k_1+1})\cdots(\nu_j^{k_j+2}-\nu_j^{k_j+1}))}{\mathrm{E}((\nu_1)_2\cdots(\nu_j)_2)}\end{aligned}$$

in view of the equation $t^k = \sum_{l=1}^k (t)_l S_{kl}$, $t\in\mathbb{R}$, $k\ge 1$ (S_{kl} are the Stirling numbers of the second kind) leads to

$$\lim_{N\to\infty}\mathrm{E}\bigg(\Big(\frac{Y_{1,j}}{N}\Big)^{k_1}\cdots\Big(\frac{Y_{j,j}}{N}\Big)^{k_j}\bigg) \overset{(16)}{=} \frac{\phi_j(k_1+2,\dots,k_j+2)}{\phi_j(2,\dots,2)},\quad k_1,\dots,k_j\in\mathbb{N}_0. \tag{24}$$

This convergence of moments implies [see Feller (1971), Chapter 8, Section 1] the weak convergence of the probability distributions on Δ_j:

$$P\bigg(\frac{Y_{1,j}}{N}\in dy_1,\dots,\frac{Y_{j,j}}{N}\in dy_j\bigg)\to P_j(dy_1,\dots,dy_j),\qquad N\to\infty. \tag{25}$$

Comparison between (24) with (25) shows that (19) holds with

$$F_j(dx_1,\dots,dx_j) = \phi_j(2,\dots,2)\cdot P_j(dx_1,\dots,dx_j).$$

The uniqueness of F_j is because the limit moments (24) fully characterize the probability measure P_j. □

DEFINITION 3.2. For $j \in \mathbb{N}$, $k_1, \dots, k_j \geq 2$ and $s \in \mathbb{N}_0$ define

$$\psi_{j,s}(k_1, \dots, k_j) := \lim_{N\to\infty} \frac{\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s})}{N^{k_1+\cdots+k_j-j} c_N}$$

as long as this limit exists.

LEMMA 3.3. *The following recursion over s holds*:

$$\begin{aligned}
&\psi_{j,s+1}(k_1, \dots, k_j) \\
&\quad = \psi_{j,s}(k_1, \dots, k_j) - \sum_{i=1}^{j} \psi_{j,s}(k_1, \dots, k_{i-1}, k_i + 1, k_{i+1}, \dots, k_j) \\
&\qquad - s\psi_{j+1,s-1}(k_1, \dots, k_j, 2)
\end{aligned}$$

for all $j \in \mathbb{N}$, $k_1, \dots, k_j \geq 2$ *and all* $s \in \mathbb{N}_0$.

PROOF. Take the l.h.s. and the r.h.s. in the following chain of equalities:

$$\begin{aligned}
&(N - j - s)\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s+1}) \\
&\quad \overset{\text{(II)}}{=} \mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s}(\nu_{j+s+1} + \cdots + \nu_N)) \\
&\quad \overset{(1)}{=} \mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s}(N - \nu_1 - \cdots - \nu_{j+s})) \\
&\quad = \mathrm{E}\Biggl((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s}\Biggl(N - (k_1 + \cdots + k_j) - s \\
&\qquad\qquad - \sum_{i=1}^{j}(\nu_i - k_i) - \sum_{i=j+1}^{j+s}(\nu_i - 1)\Biggr)\Biggr) \\
&\quad = (N - (k_1 + \cdots + k_j) - s)\, \mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s}) \\
&\qquad - \sum_{i=1}^{j} \mathrm{E}((\nu_1)_{k_1} \cdots (\nu_i)_{k_i+1} \cdots (\nu_j)_{k_j} \nu_{j+1} \cdots \nu_{j+s}) \\
&\qquad - s\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} (\nu_{j+1})_2 \nu_{j+2} \cdots \nu_{j+s}),
\end{aligned}$$

and divide them by $N^{k_1+\cdots+k_j+1-j} c_N$. After letting $N \to \infty$ we get the asserted recursion equation. □

LEMMA 3.4. *The polynomials* (12) *defined by relations* (13) *and* (14) *satisfy*

$$T^{(r)}_{r,s+1}(x_1, \dots, x_r) = \left(1 - \sum_{i=1}^{r} x_i\right) T^{(r)}_{r,s}(x_1, \dots, x_r) \tag{26}$$

and for $j = 1, \dots, r - 1$,

$$T^{(r)}_{j,s+1}(x_1, \dots, x_r) = \left(1 - \sum_{i=1}^{j} x_i\right) T^{(r)}_{j,s}(x_1, \dots, x_r) - sT^{(r)}_{j+1,s-1}(x_1, \dots, x_r). \tag{27}$$

PROOF. Formula (26) is obvious in view of (13). To verify (27) rewrite it as

$$T^{(r)}_{r-j,s+1}(x_1,\dots,x_r)=\left(1-\sum_{i=1}^{r-j}x_i\right)T^{(r)}_{r-j,s}(x_1,\dots,x_r)-sT^{(r)}_{r-j+1,s-1}(x_1,\dots,x_r)$$

and apply (14). □

LEMMA 3.5. *If the limits* (16) *exist for all* $j\in\mathbb{N}$, *then*

$$\psi_{j,s}(k_1,\dots,k_j)=\sum_{r\ge j}\int_{\Delta_r}x_1^{k_1-2}\cdots x_j^{k_j-2}T^{(r)}_{j,s}(x_1,\dots,x_r)\,F_r(dx_1,\dots,dx_r)$$

for all $j\in\mathbb{N}$, $k_1,\dots,k_j\ge 2$ *and all* $s\in\mathbb{N}_0$.

PROOF. We use induction over s. The case $s=0$ follows from the equality

$$\psi_{j,0}(k_1,\dots,k_j)=\phi_j(k_1,\dots,k_j)\overset{(19)}{=}\int_{\Delta_j}x_1^{k_1-2}\cdots x_j^{k_j-2}F_j(dx_1,\dots,dx_j).$$

Lemmas 3.3 and 3.4 ensure that the induction assumption implies the asserted formula

$$\begin{aligned}
&\psi_{j,s+1}(k_1,\dots,k_j)\\
&\quad\overset{\text{L.3.3}}{=}\psi_{j,s}(k_1,\dots,k_j)-\sum_{i=1}^{j}\psi_{j,s}(k_1,\dots,k_{i-1},k_i+1,k_{i+1},\dots,k_j)\\
&\qquad-s\psi_{j+1,s-1}(k_1,\dots,k_j,2)\\
&\quad\overset{\text{ind}}{=}\sum_{r\ge j}\int_{\Delta_r}x_1^{k_1-2}\cdots x_j^{k_j-2}\left(1-\sum_{i=1}^{j}x_i\right)T^{(r)}_{j,s}(x_1,\dots,x_r)\,F_r(dx_1,\dots,dx_r)\\
&\qquad-s\sum_{r\ge j+1}\int_{\Delta_r}x_1^{k_1-2}\cdots x_j^{k_j-2}T^{(r)}_{j+1,s-1}(x_1,\dots,x_r)\,F_r(dx_1,\dots,dx_r)\\
&\quad\overset{\text{L.3.4}}{=}\sum_{r\ge j}\int_{\Delta_r}x_1^{k_1-2}\cdots x_j^{k_j-2}T^{(r)}_{j,s+1}(x_1,\dots,x_r)\,F_r(dx_1,\dots,dx_r).
\end{aligned}$$
□

To finish the proof of Theorem 2.1, turn to the equality

$$\begin{aligned}
q_{\xi\eta}&=\lim_{N\to\infty}\frac{p_{\xi\eta}}{c_N}\\
&\overset{(3)}{=}\lim_{N\to\infty}\frac{(N)_a}{(N)_b c_N}\mathrm{E}((\nu_1)_{b_1}\cdots(\nu_g)_{b_g}\nu_{g+1}\cdots\nu_a)\\
&=\lim_{N\to\infty}\frac{\mathrm{E}((\nu_1)_{b_1}\cdots(\nu_g)_{b_g}\nu_{g+1}\cdots\nu_a)}{N^{b-a}c_N}=\psi_{g,a-g}(b_1,\dots,b_g),
\end{aligned}\tag{28}$$

saying that for any pair $\xi\subset\eta$ the l.h.s. and the r.h.s. exist or do not exist simultaneously and coincide when they exist.

Assume that the limits (16) exist for all $j \in \mathbb{N}$. Then according to Lemma 3.5 and (28) the formula (7) with (11) is valid for all $\xi \subset \eta$. The quantities $\gamma_b := \lim_{N\to\infty}(1-\mathrm{E}(\nu_1 \cdots \nu_b))/c_N$ satisfy the recursion $\gamma_{b+1} = \gamma_b + b\psi_{1,b-1}(2)$. Thus (15) follows from

$$q_{\xi\xi} = -\gamma_b = -\sum_{j=1}^{b-1} j\psi_{1,j-1}(2)$$

and Lemma 3.5. Hence the asymptotic formula (8) holds. Conversely, if the asymptotic formula (8) holds, then the existence of all the limits (16), $j \in \mathbb{N}$, $k_1 \geq \cdots, k_j \geq 2$, follows from the equality

$$\phi_j(b_1, \ldots, b_a) = \psi_{g,0}(b_1, \ldots, b_g) \stackrel{(28)}{=} q_{\xi\eta},$$

which holds provided $a = g$. Assume now that (8) holds and that the limit $c := \lim_{N\to\infty} c_N$ exists. If $c > 0$ then the asymptotic formula (8) is equivalent to $\lim_{N\to\infty} P_N = I + cQ$ and the convergence of the finite-dimensional distributions of $(\mathscr{R}_r)_{r\in\mathbb{N}_0}$ follows immediately. The weak convergence is also established as for processes with time-set $\mathbb{N}_0$ the convergence of the finite-dimensional distributions is equivalent to the weak convergence [Billingsley (1968), page 19]. Assume now that $c = 0$. Then

$$\left|\left|P_N^{[t/c_N]} - (I + c_N Q)^{[t/c_N]}\right|\right| \;\leq\; \frac{t}{c_N}||P_N - (I + c_N Q)|| = t\left|\left|\frac{P_N - I}{c_N} - Q\right|\right|$$

converges to zero as N tends to infinity and hence

$$\lim_{N\to\infty} P_N^{[t/c_N]} = \lim_{N\to\infty}(I + c_N Q)^{[t/c_N]} = e^{tQ}.$$

Thus the convergence of the finite-dimensional distributions of the time-scaled ancestral process $(\mathscr{R}_{[t/c_N]})_{t\geq 0}$ follows immediately. The convergence (5) in the Skorohod sense, that is, in $D_{\mathscr{E}_n}([0,\infty))$, is obtained as in Möhle (1999). □

REMARK. In this remark another representation for the diagonal entries $q_{\xi\xi}$ is derived. For $\xi \in \mathscr{E}_n$ with $b := |\xi|$ it follows that

$$\begin{aligned} q_{\xi\xi} &= -\sum_{\substack{\eta\in\mathscr{E}_n \\ \xi\subset\eta}} q_{\xi\eta} \\ &= -\sum_{a=1}^{b-1} \sum_{\substack{b_1\geq\cdots\geq b_a\geq 1 \\ b_1+\cdots+b_a=b}} \frac{b!}{b_1!\cdots b_a! n_1!\cdots n_b!} q_{\xi\eta} \\ &= -\sum_{g=1}^{[b/2]} \sum_{a=g}^{b-g} \sum_{\substack{b_1\geq\cdots\geq b_g\geq 2 \\ b_1+\cdots+b_g=b-(a-g)}} \frac{b!}{b_1!\cdots b_g!(a-g)! n_2!\cdots n_b!} q_{\xi\eta}, \end{aligned}$$

where $n_l := |\{i \in \{1, \dots, a\} \mid b_i = l\}|$ denotes the number of b_i's equal to l, $l \in \{1, \dots, b\}$. Thus (11) entails

$$q_{\xi\xi} = -\sum_{g=1}^{[b/2]} \sum_{a=g}^{b-g} \sum_{r=g}^{[(a+g)/2]} \int_{\Delta_r} S^{(r)}_{g,a-g}(x_1, \dots, x_r) F_r(dx_1, \dots, dx_r),$$

where

$$S^{(r)}_{g,s}(x_1, \dots, x_r) = \sum_{\substack{b_1 \geq \cdots \geq b_g \geq 2 \\ b_1 + \cdots + b_g = b-g}} \frac{b!\, x_1^{b_1-2} \cdots x_g^{b_g-2}\, T^{(r)}_{g,s}(x_1, \dots, x_r)}{b_1! \cdots b_g!\, s!\, n_2! \cdots n_b!}.$$

4. The equivalence of (16), (20), (21) and (22). Fix some $j \in \mathbb{N}$. Here we show the equivalence of the conditions (16), (20), (21) and (22) with the measure F_j and the function ϕ_j being linked by (19).

(16) $\Leftrightarrow$ (20). The proof of this equivalence is based on the decomposition

$$(\nu_1)_{k_1} \cdots (\nu_j)_{k_j} = \sum_{i_1=1}^{k_1} \cdots \sum_{i_j=1}^{k_j} \alpha_{i_1,\dots,i_j} (\nu_1 - 1)^{i_1} \cdots (\nu_j - 1)^{i_j}, \tag{29}$$

where the $\alpha_{i_1,\dots,i_j}$ are some finite coefficients and $\alpha_{k_1,\dots,k_j} = 1$. It suffices to verify that

$$\mathrm{E}((\nu_1 - 1)^{i_1} \cdots (\nu_l - 1)^{i_l}) = o(N^{k_1+\cdots+k_j-j} c_N), \qquad N \to \infty \tag{30}$$

for all $(i_1, \dots, i_l) \in \{1, \dots, k_1\} \times \cdots \times \{1, \dots, k_l\} \setminus \{(k_1, \dots, k_l)\}$, $1 \leq l \leq j$, $k_1 \geq \cdots \geq k_j \geq 2$.

To prove (30) notice first that $\mathrm{E}(\nu_1 - 1) = 0$, $\mathrm{E}((\nu_1 - 1)^2) = \mathrm{E}((\nu_1)_2) = (N-1)c_N$. Turning to a counterpart of (17) for $\mathrm{E}|(\nu_1 - 1)^{i_1} \cdots (\nu_l - 1)^{i_l}|$ we see that (30) is true when at least one i_r is greater than or equal to 2. In the remaining case $i_1 = \cdots = i_l = 1$ the equality chain

$$\begin{aligned}
&(N - l + 1)\mathrm{E}(\nu_1 - 1) \cdots (\nu_l - 1) \\
&\quad \overset{\text{(II)}}{=} \mathrm{E}((\nu_1 - 1) \cdots (\nu_{l-1} - 1)((\nu_l - 1) + \cdots + (\nu_N - 1))) \\
&\quad \overset{(1)}{=} -\mathrm{E}((\nu_1 - 1) \cdots (\nu_{l-1} - 1)((\nu_1 - 1) + \cdots + (\nu_{l-1} - 1))) \\
&\quad = -(l-1)\mathrm{E}((\nu_1 - 1)^2(\nu_2 - 1) \cdots (\nu_{l-1} - 1))
\end{aligned}$$

ends with a term $o(Nc_N)$ in accordance with the previous argument.

Thus (30) holds and we can conclude from (29) that for any fixed set of indices $k_1 \geq \cdots \geq k_j \geq 2$ the two limits (16) and (20) are equal when they exist with the existence of one entailing the existence of the other. This conclusion is slightly stronger than the asserted equivalence. □

(16) $\Leftrightarrow$ (21) and (22). Due to Lemma 3.1 the condition (16) is equivalent to the weak convergence of probability measures (25) so that the problem can

be replaced by (25) $\Leftrightarrow$ (21) and (22). The latter equivalence follows from the equality

$$P(\nu_1 > Nx_1, \dots, \nu_j > Nx_j) = \int_{x_1}^1 \cdots \int_{x_j}^1 P\left(\frac{\nu_1}{N} \in dy_1, \dots, \frac{\nu_j}{N} \in dy_j\right)$$

$$\overset{(23)}{=} N^{-2j}\mathrm{E}((\nu_1)_2 \cdots (\nu_j)_2) \int_{x_1}^1 \cdots \int_{x_j}^1 \frac{P\left(\frac{Y_{1,j}}{N} \in dy_1, \dots, \frac{Y_{j,j}}{N} \in dy_j\right)}{y_1\left(y_1 - \frac{1}{N}\right) \cdots y_j\left(y_j - \frac{1}{N}\right)}. \qquad \square$$

5. The Wright–Fisher model as a limit. Recall that the Wright–Fisher model describes a population of a fixed size (say l), where every individual chooses its parent at random among l individuals constituting the previous generation. Here we discuss a simple exchangeable population model whose time-scaled ancestral process converges to the ancestral process of the Wright–Fisher model.

Take a fixed constant $1 \le l \le N/2$ and consider such an exchangeable population model that in each generation exactly l families are of size $[N/l]$ while other family sizes are zeros and ones. In this case,

$$P(\nu_1 = \cdots = \nu_l = [N/l], \nu_{l+1} = \cdots = \nu_{l+l_1} = 1, \nu_{l+l_1+1} = \cdots = \nu_N = 0)$$
$$= \frac{1}{\binom{N}{l}\binom{N-l}{l_1}} = \frac{l!l_1!}{(N)_{l+l_1}},$$

where $l_1 := N - l[N/l]$. It follows that

$$\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j}) \sim \frac{(l)_j}{(N)_j}\left(\frac{N}{l}\right)_{k_1} \cdots \left(\frac{N}{l}\right)_{k_j}, \qquad N \to \infty$$

for all $k_1, \dots, k_j \ge 2$ and hence

$$c := \lim_{N\to\infty} c_N = 1/l.$$

This entails

$$\phi_j(k_1, \dots, k_j) = \lim_{N\to\infty} \frac{\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j})}{N^{k_1+\cdots+k_j-j} c_N} = (l)_j l^{1-k_1-\cdots-k_j}.$$

Thus for this particular model the limit measure F_j assigns its total mass $\phi_j(2, \dots, 2) = (l)_j l^{1-2j}$ to the single point $(1/l, \dots, 1/l) \in \mathbb{R}^j$ being a zero measure for $j > l$. Now using Lemma 3.3 and induction over s we can show that

$$\psi_{j,s}(k_1, \dots, k_j) = (l)_{j+s} l^{1-s-k_1-\cdots-k_j}. \tag{31}$$

The case $s = 0$ follows from $\psi_{j,0}(k_1, \dots, k_j) = \phi_j(k_1, \dots, k_j) = (l)_j l^{1-k_1-\cdots-k_j}$. The step from s to $s+1$ is given by

$$\begin{aligned}\psi_{j,s+1}(k_1, \dots, k_j) &\overset{\text{L.3.3}}{=} (l)_{j+s} l^{1-s-k} - \sum_{i=1}^{j} (l)_{j+s} l^{-s-k} - s(l)_{j+s} l^{-s-k} \\ &= (l)_{j+s} l^{-s-k}(l - j - s) = (l)_{j+s+1} l^{1-(s+1)-k},\end{aligned}$$

where $k := k_1 + \cdots + k_j$. We conclude that for $\xi \subset \eta$,

$$q_{\xi\eta} \overset{(28)}{=} \psi_{g,a-g}(b_1, \dots, b_g) \overset{(31)}{=} (l)_a l^{1-(a-g)-b_1-\cdots-b_g} = (l)_a l^{1-b}$$

so that the transition matrix $\Pi = I + cQ$ for the limit Markov chain has entries $\pi_{\xi\eta} = (l)_a l^{-b}$ for $\xi \subset \eta$ and the resulting coalescent process coincides with the ancestral process for the Wright–Fisher model with the population size l.

As l tends to infinity, the generator Q converges to the generator of the standard n-coalescent in agreement with the weak convergence of the measure F_1 to the point measure in zero. For $j > 1$ the total mass of F_j converges to zero as l tends to infinity.

To generalize our example take an integer valued random variable L_N with

$$P(1 \le L_N \le N/2) = 1, \qquad N \in \mathbb{N}$$

and conditional on $\{L_N = l\}$, $l \in \mathbb{N}$ define a population model as before. Assuming that L_N converges weakly as N tends to infinity to some random variable L, we deduce

$$\begin{aligned}\phi_j(k_1, \dots, k_j) &= \lim_{N\to\infty} \frac{\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j})}{N^{k_1+\cdots+k_j-j} c_N} \\ &= \lim_{N\to\infty} \sum_{l=1}^{N/2} \frac{\mathrm{E}((\nu_1)_{k_1} \cdots (\nu_j)_{k_j} \mid L_N = l)}{N^{k_1+\cdots+k_j-j} c_N} P(L_N = l) \\ &= \sum_{l=1}^{\infty} (l)_j l^{1-k_1-\cdots-k_j} P(L = l) = \mathrm{E}((L)_j L^{1-k_1-\cdots-k_j})\end{aligned}$$

and

$$c := \lim_{N\to\infty} c_N = \mathrm{E}(1/L).$$

Note that the last expectation is positive even if we allow for the possibility $0 \le P(L = \infty) < 1$. In particular, if $L - 1$ has a Poisson distribution with parameter $\lambda > 0$, then

$$c = \int_0^1 \mathrm{E}(x^{L-1})\, dx = \int_0^1 e^{\lambda(x-1)}\, dx = \frac{1 - e^{-\lambda}}{\lambda}.$$

For the generalized example it follows that the entries of the limit generator Q are given by $q_{\xi\eta} = \mathrm{E}((L)_a L^{1-b})$ and the transition matrix $\Pi = I + cQ$ for the limit Markov chain has entries $\pi_{\xi\eta} = \mathrm{E}(1/L)\mathrm{E}((L)_a L^{1-b})$ for $\xi \subset \eta$.

The resulting coalescent process depends on the observed value of the limit random variable L. If $L = l < \infty$, the coalescent is the ancestral process of the Wright–Fisher model with the population size l. When $L = \infty$ the sampled ancestral lines never merge.

6. Discussion, historical perspectives and recent developments. Coalescent theory has its origin in physics and in genetics. In statistical physics models are studied where objects of different masses move in space. When two objects, of masses x and y say, come close they may coalesce into one object of mass $x + y$. A rate kernel $K(x, y)$ specifies the propensity that these objects merge together. Models with a fixed number n of objects are studied as well as models with countable many objects. Evans and Pitman (1998) introduce such Markovian coalescent processes in general. Aldous and Pitman (1998) study the standard additive coalescent ($K(x, y) = x + y$) and present a construction via a Poisson splitting procedure on continuum random trees. Aldous (1997) is in particular interested in the multiplicative coalescent ($K(x, y) = xy$). See also Aldous and Limic (1998) for further details.

In mathematical population genetics the coalescent theory goes back to Kingman (1982a, b, c). The genetics impact of the coalescent theory is well described in the review-like articles of Donnelly and Tavaré (1995), Hudson (1991), Li and Fu (1999), Möhle (2000) and Nordborg (2001) and references therein.

A special (and probably first) example for a coalescent process with multiple collisions goes back to Bolthausen and Sznitman (1998). Their process, motivated by Ruelle's probability cascades, corresponds to the special case when the measure F in (9) is uniformly distributed on $[0, 1]$. Other recent constructions of this process, based on the genealogy of continuous-state branching processes and on subordination schemes, are presented in Bertoin and Le Gall (2000) and Bertoin and Pitman (2000).

Sagitov (1999) relaxes the reproduction moment conditions put by Kingman (1982b). As the population size tends to infinity he derives a class of limit coalescent processes allowing for multiple mergers of ancestral lines. The corresponding generator Q with entries (9) appears when the limits $\phi_1(k)$, $k \geq 3$, in (16) are allowed to be larger than zero. On the other hand the generator Q with entries (9) prohibits simultaneous collisions ($\phi_2(2, 2) = 0$). Pitman (1999) and Schweinsberg (2000a) study directly the class of coalescent processes with generator (9) without having a limit from a discrete population model in the background.

The present paper generalizes the methods of Möhle (1999) and Sagitov (1999) and leads to a full classification of the limit coalescent processes for the case of exchangeable reproduction. The corresponding generator allows for simultaneous multiple collisions; that is, many multiple collisions can appear at the same time. Such simultaneous collisions appear if and only if the limit $\phi_2(2, 2)$ in (16) is larger than zero. In a recent paper Schweinsberg (2000b) goes further in studying the coalescent processes with simultaneous multiple mergers introduced here. In particular he characterizes the corresponding generator in terms of a single measure on an infinite-dimensional simplex.

REFERENCES

ALDOUS, D. J. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** 812–854.

ALDOUS, D. J. and LIMIC, V. (1998). The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.* **3**.

ALDOUS, D. J. and PITMAN, J. (1998). The standard additive coalescent. *Ann. Probab.* **26** 1703–1726.

BERTOIN, J. and 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. and PITMAN, J. (2000). Two coalescents derived from the ranges of stable subordinators. *Electron J. Probab.* **5**.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

BOLTHAUSEN, E. and 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. in 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. in Appl. Probab.* **7** 264–282.

DONNELLY, P. and TAVARÉ, S. (1995). Coalescents and genealogical structure under neutrality. *Ann. Rev. Genet.* **29** 401–421.

EVANS, S. N. and PITMAN, J. (1998). Construction of Markovian coalescents. *Ann. Inst. H. Poincaré Probab. Statist.* **34** 339–383.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, **1**, 2nd ed. Wiley, New York.

HUDSON, R. R. (1991). Gene genealogies and the coalescent process. *Oxford Survey Evol. Biol.* **7** 1–44.

KINGMAN, J. F. C. (1982a). On the genealogy of large populations. *J. Appl. Probab.* **19A** 27–43.

KINGMAN, J. F. C. (1982b). Exchangeability and the evolution of large populations. In *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino, eds.) 97–112. North–Holland, Amsterdam.

KINGMAN, J. F. C. (1982c). The coalescent. *Stochastic Process. Appl.* **13** 235–248.

LI, W. H. and FU, Y. X. (1999). Coalescent theory and its applications in population genetics. In *Statistics in Genetics* (M. E. Halloran and S. Geisser, eds.) Springer, Berlin.

MÖHLE, M. (1998). Robustness results for the coalescent. *J. Appl. Probab.* **35** 438–447.

MÖHLE, M. (1999). Weak convergence to the coalescent in neutral population models. *J. Appl. Probab.* **36** 446–460.

MÖHLE, M. (2000). Ancestral processes in population genetics: the coalescent. *J. Theoret. Biol.* **204** 629–638.

MÖHLE, M. and SAGITOV, S. (1998). A characterisation of ancestral limit processes arising in haploid population genetics models. Preprint. Johannes Gutenberg-Univ., Mainz.

NORDBORG, M. (2001). Coalescent theory. In *Handbook of Statistical Genetics* (D. J. Balding, C. Cannings and M. Bishop, eds.) Wiley, Chichester. To appear.

PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902.

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

SCHWEINSBERG, J. (2000a). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.* **5**.

SCHWEINSBERG, J. (2000b). Coalescents with simultaneous multiple collision. Technical Report *Electron. J. Probab.* **5**.

DEPARTMENT OF MATHEMATICS
JOHANNES GUTENBERG UNIVERSITY MAINZ
SAARSTRAßE 21
55099 MAINZ
GERMANY
E-MAIL: moehle@mathematik.uni-mainz.de

DEPARTMENT OF MATHEMATICAL STATISTICS
CHALMERS AND GÖTEBORGS UNIVERSITIES
41296 GÖTEBORG
SWEDEN
E-MAIL: serik@math.chalmers.se