J. Appl. Prob. **53,** 502–517 (2016)
*doi:*10.1017/jpr.2016.16

CONVERGENCE TO THE STRUCTURED COALESCENT PROCESS

RYOUTA KOZAKAI,* **
AKINOBU SHIMIZU * AND
MORIHIRO NOTOHARA,* *Nagoya City University*

Abstract

The coalescent was introduced by Kingman (1982a), (1982b) and Tajima (1983) as a continuous-time Markov chain model describing the genealogical relationship among sampled genes from a panmictic population of a species. The random mating in a population is a strict condition and the genealogical structure of the population has a strong influence on the genetic variability and the evolution of the species. In this paper, starting from a discrete-time Markov chain model, we show the weak convergence to a continuous-time Markov chain, called the structured coalescent model, describing the genealogy of the sampled genes from whole population by means of passing the limit of the population size. Herbots (1997) proved the weak convergence to the structured coalescent on the condition of conservative migration and Wright–Fisher-type reproduction. We will give the proof on the condition of general migration rates and exchangeable reproduction.

Keywords: Structured coalescent; nonconservative migration; Cannings' reproduction; weak convergence

2010 Mathematics Subject Classification: Primary 92D25; 92D10
Secondary 60F17

1. Introduction

The standard coalescent process describes the genealogical relationship among randomly sampled genes from a panmictic population. The number of ancestors of sampled genes is reduced one by one through coalescent events if we trace the ancestral lineage. In the limit as the population size goes to ∞, Kingman (1982a), (1982b) proved the convergence to the ancestral limit process for a broad class of reproduction model called the exchangeable model, including the Wright–Fisher model and the Moran model, on some condition of the number of children produced by each individual. The natural population is hardly panmictic, and the geographic structure of the population affects the evolution and the genetic diversity of the population. The coalescent process with geographical structure is called the structured coalescent process. The structured coalescent model has been investigated by Takahata (1988), Notohara (1990), (1993), (1997), (2001), (2010), Nath and Griffiths (1993), Herbots (1994), (1997), Wilkinson-Herbots (1998), Bahlo and Griffiths (2000), Sampson (2006), and Wakeley (1998), (2009). Takahata (1988) introduced the structured coalescent consisting of two subpopulations, and Notohara (1990) formulated it in the general form. A rigorous mathematical derivation was given by Herbots (1994), (1997) on the condition of conservative migration and Wright–Fisher-type reproduction. Another mathematical proof of the weak convergence of the discrete-time model

Received 6 January 2015; revision received 10 June 2015.
* Postal address: Graduate School of Natural Sciences, Nagoya City University, Mizuho, Nagoya, 467-8501, Japan.
** Email address: r.kosakai@nsc.nagoya-cu.ac.jp

was given by Sampson (2006) on the condition of general migration rates, where he studied the case with rapidly changing population sizes, finite subpopulations and the Wright–Fisher reproduction. The proof is based on the convergence theorem proved by Möhle (1998) and similar applications of the convergence theorem to structured population models can be found in the literature dealing with the weak convergence of discrete-time Markov chains; see, for example, Hössjer (2011), Kaj *et al.* (2001), Nordborg and Krone (2002), Pollak (2011), and Sagitov and Jagers (2005). In these studies, the number of ancestors in the limiting process follows Kingman's coalescent.

In this paper we will show that the sequence of discrete-time Markov chain models with countable subpopulations, whose migration steps are based on the condition including nonconservative migration and whose backward reproduction steps are deduced from Cannings' model (Cannings (1974)), converges weakly to the structured coalescent, which is the continuous-time Markov chain determined by the matrix $\boldsymbol{Q}$ stated below. The structured coalescent model is described as follows. Let S be the countable set of subpopulation labels. We denote by $N_i(= 2c_iN)$ the number of individuals after reproduction step in subpopulation $i \in S$, where c_i is a positive integer constant and the number of individuals after migration in subpopulation i is denoted by N_i^*. In general, N_i^* does not equal N_i, which includes the case of nonconservative migration. We focus on the genealogical relationship of sampled genes from an entire population. We randomly sample α_k genes from colony $k \in S$ and denote the set of all sampled genes by $\alpha = (\alpha_k : k \in S)$. We assume that the total number of sampled genes $|\alpha| = \sum_{k\in S}\alpha_k < \infty$. Let $E = \{\alpha = (\alpha_k : k \in S) \in \mathbb{Z}_+{}^S : |\alpha| \le n\}$ for a fixed natural number n, where $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ denotes the set of nonnegative integers. We define binary operations $\alpha \pm \beta$ for $\alpha, \beta \in E$ by $(\alpha \pm \beta)_i = \alpha_i \pm \beta_i$ provided $\alpha \pm \beta \in E$. Let $\boldsymbol{e}^i$ be a unit vector such that $(\boldsymbol{e}^i)_j = \delta_{i,j}$(1 for $i = j$ and 0 otherwise) for each $i \in S$. The structured coalescent is the continuous-time Markov chain $\{\alpha(t) : t \ge 0\}$ on the state space E generated by the matrix $\boldsymbol{Q}$ whose entries are

$$\boldsymbol{Q}_{\alpha,\beta} = \begin{cases} -\displaystyle\sum_{i\in S}\left\{\alpha_i \frac{M_i}{2} + \frac{\sigma^2\alpha_i(\alpha_i - 1)}{2c_i}\right\} & \text{if } \beta = \alpha, \\ \alpha_i \dfrac{M_{i,j}}{2} & \text{if } \beta = \alpha - \boldsymbol{e}^i + \boldsymbol{e}^j \ (i \ne j), \\ \dfrac{\sigma^2\alpha_i(\alpha_i - 1)}{2c_i} & \text{if } \beta = \alpha - \boldsymbol{e}^i, \\ 0 & \text{otherwise,} \end{cases} \tag{1.1}$$

where $M_{i,j}$ is the scaled migration rate (backward in time) from subpopulation $i \in S$ to $j \in S$, $M_i = \sum_{j\ne i} M_{i,j}$, and σ is a positive constant. The constants $M_{i,j}$, M_i, and σ^2 are discussed in the second section. The ancestral process $\alpha(t) = (\alpha_i(t) : i \in S)$ with $\alpha(0) = \alpha \in E$ represents the geographical configuration of ancestors of a sample α at time t backward, $\alpha_i(t)$ is the number of distinct ancestors locating at colony $i \in S$. In Shiga (1980a), (1980b) the author investigated the structure of stationary states and the ergodic property of the stepping stone model described by the infinite-dimensional diffusion process. The duality relation between Shiga's stepping stone model and the structured coalescent was discussed in Notohara (1990). In the next section we will explain our model in detail. Section 3 is devoted to the proof of the convergence of the transition probability, and this implies the convergence of finite-dimensional distributions to the continuous-time Markov chain. In Section 4 we will prove the weak convergence to the structured coalescent on the condition of general migration rates and exchangeable reproduction.

2. The discrete-time model

We consider a discrete-time Markov chain model, where each generation is made up of a migration step and a reproduction step. At each generation, a fixed proportion $q_{i,j}$ of individuals born in subpopulation i are assumed to migrate to subpopulation j. Here, it is assumed that $q_{i,j}$ is constant for each $(i, j) \in S \times S$, and

$$q_{i,j} \geq 0, \qquad \sum_{j \in S,\, j \neq k} q_{k,j} \leq 1.$$

First, we will discuss migration rates and backward migration rates. We assume the following conditions on the constants c_i, $i \in S$, and the migration rates $\{q_{k,j}\}$.

(A.1) There exists a constant K such that $1 \leq c_i < K$, $i \in S$.

(A.2) The migration rate $q_{i,j}$, $i \neq j$, can be written as

$$q_{i,j} = \frac{q^*_{i,j}}{4N} \quad (i \neq j),$$

where

$$\sup_{k \in S} \sum_{j \in S,\, j \neq k} q^*_{k,j} < \infty \quad \text{and} \quad \sup_{k \in S} \sum_{j \in S,\, j \neq k} q^*_{j,k} < \infty \tag{2.1}$$

are satisfied.

The backward migration rate $m_{i,j}$ from subpopulation i to subpopulation j, defined as the proportion of individuals in subpopulation i immediately after the migration step who were born in subpopulation j, is given by $m_{i,j} = N_j q_{j,i} / N^*_i$, and we define

$$m_i = \sum_{j \neq i} m_{i,j} = \sum_{j \neq i} \frac{N_j q_{j,i}}{N^*_i}.$$

The population size after migration step N^*_i is given by

$$\begin{aligned}
N^*_i &= \sum_{k \neq i} N_k q_{k,i} + N_i \biggl(1 - \sum_{j \neq i} q_{i,j}\biggr) \\
&= \sum_{k \neq i} 2c_k N \frac{q^*_{k,j}}{4N} + 2c_i N \biggl(1 - \sum_{j \neq i} \frac{q^*_{i,j}}{4N}\biggr) \\
&= 2c_i N + \frac{1}{2}\biggl(\sum_{k \neq i} c_k q^*_{k,i} - c_i \sum_{j \neq i} q^*_{i,j}\biggr) \\
&= 2c_i N + \frac{Q_i}{2},
\end{aligned}$$

where

$$Q_i = \sum_{k \neq i} c_k q^*_{k,i} - c_i \sum_{j \neq i} q^*_{i,j}.$$

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

If $Q_i = 0$ holds for all $i \in S$ then it is said to be conservative. The absolute value of Q_i can be evaluated as

$$\begin{aligned} |Q_i| &= \left| \sum_{k\neq i} c_k q^*_{k,i} - c_i \sum_{j\neq i} q^*_{i,j} \right| \\ &\le K\left(\sum_{k\neq i} q^*_{k,i} + \sum_{j\neq i} q^*_{i,j} \right) \\ &\le K\left(\sup_k \sum_{j\neq k} q^*_{k,j} + \sup_k \sum_{j\neq k} q^*_{j,k} \right) \\ &\le C, \end{aligned} \tag{2.2}$$

where C is a positive constant, and we have

$$m_{i,j} = \frac{N_j q_{j,i}}{N^*_i} = \frac{2c_j N q_{j,i}}{2c_i N + Q_i/2} \le \frac{K}{2} \frac{q^*_{j,i}}{2N - C/2}, \qquad \lim_{N\to\infty} 4Nm_{i,j} = \frac{c_j}{c_i} q^*_{j,i},$$
$$\frac{c_j q^*_{j,i}}{c_i + Q_i/4N} \le \frac{c_j q^*_{j,i}}{1 - C/4N} \le \frac{K q^*_{j,i}}{1 - C/4N} < \infty,$$

where the inequality $c_i \ge 1, i \in S$, in (A.1) is used. Thus, using Lebesgue's convergence theorem and (2.1), we have

$$\lim_{N\to\infty} 4Nm_i = \lim_{N\to\infty} \sum_{j\neq i} 4Nm_{i,j} = \lim_{N\to\infty} \sum_{j\neq i} \frac{c_j q^*_{j,i}}{c_i + Q_i/4N} = \sum_{j\neq i} \frac{c_j}{c_i} q^*_{j,i}.$$

We denote the scaled migration rate $M_{i,j}$ by

$$M_{i,j} = \frac{c_j}{c_i} q^*_{j,i} \quad \text{and} \quad M_i = \sum_{j\neq i} \frac{c_j}{c_i} q^*_{j,i}.$$

If N is sufficiently large, we have

$$4Nm_i \le \sum_{j\neq i} \frac{c_j q^*_{j,i}}{c_i - C/4N} \le 2K \sum_{j\neq i} q^*_{j,i} = 2K\left(\sup_i \sum_{j\neq i} q^*_{j,i} \right) := M,$$

where M is a positive constant. Thus, we have

$$\sup_{i\in S} m_i \le \frac{M}{4N} \quad \text{and} \quad \sup_{i\in S} M_i \le K \sup_{i\in S} \sum_{j\neq i} q^*_{j,i} \le \frac{M}{2}. \tag{2.3}$$

2.1. Cannings' exchangeable reproduction

In what follows the generations are labeled backwards in time. The first generation means the parent's generation, the second generation is the grandparent's generation, and so on. Let $\nu_i^{(l,r)}$ be the number of offspring of the ith individual alive in the rth generation in subpopulation $l \in S$. We assume that the following statements hold.

(A.3) $(\nu_1^{(l,r)}, \nu_2^{(l,r)}, \dots, \nu_{N^*_l}^{(l,r)})$ is exchangeable for fixed $l \in S$ and $r \in \mathbb{Z}_+$, and

$$\sum_{i=1,2,\dots,N^*_l} \nu_i^{(l,r)} = N_l.$$

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

(A.4) $(\nu_1^{(l,r)}, \nu_2^{(l,r)}, \dots, \nu_{N_l^*}^{(l,r)})$, $r \in \mathbb{Z}_+$, is independent and identically distributed for fixed l.

(A.5) $\{(\nu_1^{(l,r)}, \nu_2^{(l,r)}, \dots, \nu_{N_l^*}^{(l,r)}),\ r \in \mathbb{Z}_+\}$, $l \in S$, are independent.

Also

(A.6) $\lim_{N\to\infty} \sup_{l\in S} |\mathbb{E}[\{\nu_1^{(l,r)}\}^2] - (\sigma^2+1)| = 0$ for a constant $\sigma^2 > 0$, and that

(A.7) $K_p^* = \sup_{l\in S, N} \mathbb{E}[\{\nu_1^{(l,r)}\}^p] < \infty$ for any positive integer p.

Now, c_N^l is the probability that two individuals, chosen randomly without replacement in subpopulation l at a generation, share a parent. This probability is given by

$$c_N^l = \frac{\sum_{i=1}^{N_l^*} \mathbb{E}[\nu_i^{(l,r)}(\nu_i^{(l,r)} - 1)]}{N_l(N_l - 1)} = \frac{1}{2N}\left\{\left(1 + \frac{Q_l}{4c_l N}\right)\frac{\mathbb{E}[\{\nu_1^{(l,r)}\}^2]}{c_l - 1/2N} - \frac{1}{c_l - 1/2N}\right\}.$$

From the above, we have $\lim_{N\to\infty} 2Nc_N^l = \sigma^2/c_l$.

2.2. Backward migration matrix

We denote by $R_N^{\mathrm{m}}(\alpha)$ the probability that two or more individuals in α are migrant, and it satisfies

$$\begin{aligned} R_N^{\mathrm{m}}(\alpha) &\le \sum_{k\in S} \frac{\binom{\alpha_k}{2}\binom{N_k^*-2}{m_kN_k^*-2}}{\binom{N_k^*}{m_kN_k^*}} + \sum_{k\in S} \frac{\binom{\alpha_k}{1}\binom{N_k^*-1}{m_kN_k^*-1}}{\binom{N_k^*}{m_kN_k^*}} \sum_{l\neq k} \frac{\binom{\alpha_l}{1}\binom{N_l^*-1}{m_lN_l^*-1}}{\binom{N_l^*}{m_lN_l^*}} \\ &\le \sum_{k\in S}(\alpha_k m_k)^2 + \sum_{k\in S}\alpha_k m_k \sum_{l\neq k}\alpha_l m_l \\ &= \left(\sum_{k\in S}\alpha_k m_k\right)^2. \end{aligned}$$

From (2.3), it follows that

$$R_N^{\mathrm{m}}(\alpha) \le \frac{M^2|\alpha|^2}{16N^2}.$$

The probability that exactly one individual in α migrates backward in time from subpopulation i to subpopulation $j(\neq i)$ is

$$\begin{aligned} &\frac{\binom{\alpha_i}{1}\binom{N_i^*-\alpha_i}{m_{i,j}N_i^*-1}}{\binom{N_i^*}{m_{i,j}N_i^*}} \frac{\binom{N_i^*-m_{i,j}N_i^*-\alpha_i+1}{m_iN_i^*-m_{i,j}N_i^*}}{\binom{N_i^*-m_{i,j}N_i^*}{m_iN_i^*-m_{i,j}N_i^*}} \prod_{k\neq i} \frac{\binom{N_k^*-\alpha_k}{m_kN_k^*}}{\binom{N_k^*}{m_kN_k^*}} \\ &\quad = \alpha_i m_{i,j} \frac{N_i^*}{N_i^* - m_iN_i^* - \alpha_i + 1} \prod_{k\in S} \prod_{a=0,\dots,\alpha_k-1} \frac{N_k^* - m_kN_k^* - a}{N_k^* - a}. \end{aligned} \tag{2.4}$$

Let $R_N^{\mathrm{m}}(\alpha,\beta)$ be the probability that the backward migration step changes the value of the ancestral process from α to β and two or more individuals in α are migrants. By (2.4), we see

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

that the transition probability in one backward migration step is given by

$$P_N^{\mathrm{m}}(\beta \mid \alpha) = \begin{cases} 1 - \sum_{i \in S} \alpha_i m_i \frac{N_i^*}{N_i^* - m_i N_i^* - \alpha_i + 1} \\ \quad \times \prod_{k \in S} \prod_{a=0,\dots,\alpha_k - 1} \frac{N_k^* - m_k N_k^* - a}{N_k^* - a} \\ \quad - \sum_{\gamma \neq \alpha} R_N^{\mathrm{m}}(\alpha, \gamma) & \text{if } \beta = \alpha, \\ \alpha_i m_{i,j} \frac{N_i^*}{N_i^* - m_i N_i^* - \alpha_i + 1} \\ \quad \times \prod_{k \in S} \prod_{a=0,\dots,\alpha_k - 1} \frac{N_k^* - m_k N_k^* - a}{N_k^* - a} \\ \quad + R_N^{\mathrm{m}}(\alpha, \alpha - \boldsymbol{e}^i + \boldsymbol{e}^j) & \text{if } \beta = \alpha - \boldsymbol{e}^i + \boldsymbol{e}^j (j \neq i), \\ R_N^{\mathrm{m}}(\alpha, \beta) & \text{otherwise,} \end{cases} \tag{2.5}$$

where

$$\sum_{\beta \neq \alpha} R_N^{\mathrm{m}}(\alpha, \beta) \leq R_N^{\mathrm{m}}(\alpha) \leq \left(\sum_{k \in S} \alpha_k m_k \right)^2.$$

2.3. Backward reproduction matrix

We denote by $R_N^{\mathrm{r}}(\alpha)$ the probability that two or more pairs of individuals belonging to α each share a parent. This probability is given by

$$R_N^{\mathrm{r}}(\alpha) = \sum_{v=2}^{\infty} P_N^{\mathrm{r}}\{\text{exactly } v \text{ pairs of individuals in } \alpha \text{ each share a parent}\}. \tag{2.6}$$

We will discuss the probability $R_N^{\mathrm{r}}(\alpha)$. Suppose that $a_l = N_l^*$ and $b_l = N_l$. We have

$$\frac{a_l}{b_l} = 1 + \frac{1}{2N} \frac{Q_l}{2c_l} \leq 2, \qquad \frac{a_l - 1}{b_l - 1} \leq 2\frac{a_l}{b_l} \leq 4, \qquad \frac{1}{b_l - 1} \leq \frac{1}{b_l - 2} \leq \frac{1}{b_l - 3} \leq \frac{1}{N},$$

and

$$|c_N^l| \leq \frac{a_l K_2^*}{b_l(b_l - 1)} \leq \frac{2K_2^*}{N} \quad \text{for } N \geq \max\left(\frac{C}{4}, 3\right).$$

The event that two or more pairs of individuals belonging to α each share a parent is included in the union of the event $C(1)$ and $C(2)$. Here, $C(1)$ denotes the event that two or more pairs of individuals which are drawn from at least two distinct subpopulations each share a parent and $C(2)$ stands for the event that two or more pairs of individuals which are drawn from one subpopulation each share a parent. Thus, we see that

$$R_N^{\mathrm{r}}(\alpha) \leq P(C(1)) + P(C(2)). \tag{2.7}$$

First, we will discuss $P(C(1))$. Since

$$\sum_{i=1}^{a_{l_1}} \frac{\nu_i^{(l_1,r)}(\nu_i^{(l_1,r)} - 1)}{b_{l_1}(b_{l_1} - 1)} \quad \text{and} \quad \sum_{i=1}^{a_{l_2}} \frac{\nu_i^{(l_2,r)}(\nu_i^{(l_2,r)} - 1)}{b_{l_2}(b_{l_2} - 1)}$$

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

are mutually independent for $l_1 \neq l_2$, we have

$$P(C(1)) \le \binom{V}{2}\frac{(2K_2^*)^2}{N^2} \le \frac{n^4(2K_2^*)^2}{8N^2} \le \frac{C_1}{N^2}, \tag{2.8}$$

where $V = \sum_{i\in S}\binom{\alpha_i}{2}$ and $C_1 = n^4(K_2^*)^2/2$.

Next, we will discuss $P(C(2))$. We can evaluate $P(C(2))$ as

$$P(C(2)) \le \sum_{l\in S}(P_l^{(1)} + P_l^{(2)}),$$

where

$$P_l^{(1)} = \binom{\alpha_l}{3}\sum_{i=1}^{a_l}\frac{\mathbb{E}[\nu_i^{(l,r)}(\nu_i^{(l,r)}-1)(\nu_i^{(l,r)}-2)]}{b_l(b_l-1)(b_l-2)}$$

and

$$P_l^{(2)} = \left(\binom{\binom{\alpha_l}{2}}{2}\right)\sum_{i\neq j,\ 1\le i,j\le a_l}\frac{\mathbb{E}[\nu_i^{(l,r)}(\nu_i^{(l,r)}-1)\nu_j^{(l,r)}(\nu_j^{(l,r)}-1)]}{b_l(b_l-1)(b_l-2)(b_l-3)}.$$

It follows that $P_l^{(1)}$ satisfies

$$P_l^{(1)} \le \frac{{\alpha_l}^3}{6}\frac{a_l\mathbb{E}[{\nu_1^{(l,r)}}^3]}{b_l(b_l-1)(b_l-2)} \le \frac{{\alpha_l}^3}{6}\frac{2K_3^*}{N^2}.$$

From the above, we have

$$\sum_{l\in S}P_l^{(1)} \le \frac{n^3K_3^*}{3N^2}.$$

Then $P_l^{(2)}$ satisfies

$$P_l^{(2)} \le \left(\binom{\binom{\alpha_l}{2}}{2}\right)\frac{a_l(a_l-1)\mathbb{E}[{\nu_1^{(l,r)}}^2{\nu_2^{(l,r)}}^2]}{b_l(b_l-1)(b_l-2)(b_l-3)} \le \frac{{\alpha_l}^4K_4^*}{(b_l-2)(b_l-3)} \le \frac{{\alpha_l}^4K_4^*}{N^2}.$$

From the above, we have

$$\sum_{l\in S}P_l^{(2)} \le \frac{n^4K_4^*}{N^2}.$$

By the above arguments, we obtain

$$P(C(2)) \le \left(\frac{n^3}{3}K_3^* + n^4K_4^*\right)\frac{1}{N^2} = \frac{C_2}{N^2}, \tag{2.9}$$

where $C_2 = (n^3/3)K_3^* + n^4K_4^*$. By (2.7)–(2.9), we have

$$R_N^{\mathrm{r}}(\alpha) \le \frac{C_1 + C_2}{N^2}. \tag{2.10}$$

We denote by $R_N^{\mathrm{r}}(\alpha,\beta)$ the probability that the backward reproduction step changes the value of the ancestral process from α to $\beta \notin \{\alpha\}\cup\{\alpha - \boldsymbol{e}^i;\ i\in S\}$. Now, we have

$$R_N^{\mathrm{r}}(\alpha) = \sum_{\beta\notin\{\alpha\}\cup\{\alpha-\boldsymbol{e}^i,\ i\in S\}} R_N^{\mathrm{r}}(\alpha,\beta).$$

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

It follows that $R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i)$ is defined by

$$P_N^{\mathrm{r}}(\alpha - \boldsymbol{e}^i \mid \alpha) = \binom{\alpha_i}{2} c_N^i - R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i),$$

where $P_N^{\mathrm{r}}(\alpha - \boldsymbol{e}^i \mid \alpha)$ is the probability that the backward reproduction step changes the value of the ancestral process from α to $\alpha - \boldsymbol{e}^i$. We have

$$\sum_{i \in S} P_N^{\mathrm{r}}(\alpha - \boldsymbol{e}^i \mid \alpha) = P_N^{\mathrm{r}}\{\text{exactly one pair of individuals in } \alpha \text{ each share a parent}\} \tag{2.11}$$

and

$$\sum_{i \in S} \binom{\alpha_i}{2} c_N^i = \sum_{v=1}^{\infty} v P_N^{\mathrm{r}}\{\text{exactly } v \text{ pairs of individuals in } \alpha \text{ each share a parent}\}. \tag{2.12}$$

From (2.11) and (2.12), it follows that

$$\sum_{i \in S} R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i) = \sum_{v=2}^{\infty} v P_N^{\mathrm{r}}\{\text{exactly } v \text{ pairs of individuals in } \alpha \text{ each share a parent}\}. \tag{2.13}$$

By (2.6) and (2.13), we have

$$2R_N^{\mathrm{r}}(\alpha) \le \sum_{i \in S} R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i) \le \binom{n}{2} R_N^{\mathrm{r}}(\alpha). \tag{2.14}$$

From the above, we have

$$\sum_{\beta \neq \alpha} R_N^{\mathrm{r}}(\alpha, \beta) \le \left\{\binom{n}{2} + 1\right\} R_N^{\mathrm{r}}(\alpha). \tag{2.15}$$

Combining the above, the transition probability of the ancestral process in one backward reproduction step is given by

$$P_N^{\mathrm{r}}(\beta \mid \alpha) = \begin{cases} 1 - \sum_{i \in S} \binom{\alpha_i}{2} c_N^i + \sum_{i \in S} R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i) \\ \quad - \sum_{\gamma \notin \{\alpha\} \cup \{\alpha - \boldsymbol{e}^i,\, i \in S\}} R_N^{\mathrm{r}}(\alpha, \gamma) & \text{if } \beta = \alpha, \\ \binom{\alpha_i}{2} c_N^i - R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i) & \text{if } \beta = \alpha - \boldsymbol{e}^i, \\ R_N^{\mathrm{r}}(\alpha, \beta) & \text{otherwise.} \end{cases} \tag{2.16}$$

3. Convergence of the finite-dimensional distribution

As migration and reproduction operate independently, the one-generation transition probabilities of the ancestral process are found from the transition probabilities in one backward migration step and one backward reproduction step as

$$P_N(\beta \mid \alpha) = \sum_{\gamma} P_N^{\mathrm{r}}(\gamma \mid \alpha) P_N^{\mathrm{m}}(\beta \mid \gamma).$$

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

In matrix notation, the transition matrix $\boldsymbol{P}_N$ in one generation is given by

$$\boldsymbol{P}_N = \boldsymbol{P}_N^{\mathrm{r}} \boldsymbol{P}_N^{\mathrm{m}}, \tag{3.1}$$

where $\boldsymbol{P}_N^{\mathrm{m}}$ and $\boldsymbol{P}_N^{\mathrm{r}}$ are the transition matrices in one backward migration step and one backward reproduction step given by (2.5) and (2.16), respectively. Let $\{\alpha^{(N)}(\tau)\}_{\tau\in\mathbb{Z}_+} = \{(\alpha_i^{(N)}(\tau))_{i\in S}\}_{\tau\in\mathbb{Z}_+}$ be the discrete-time Markov chain on E whose transition probability is P_N with the initial state $\alpha^{(N)}(0) = \alpha (\in E)$. In this section we will show that the sequence of the finite-dimensional distributions of $\{\alpha_i^{(N)}([2Nt])\}$ converges to that of the structured coalescent process $\{\alpha(t)\}_{t\ge 0}$ with $\alpha(0) = \alpha$. Let $\boldsymbol{I}$ denote the identity matrix. Namely,

$$\boldsymbol{I}_{\alpha,\beta} = \delta_{\alpha,\beta} = \begin{cases} 1 & \text{if } \beta = \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\boldsymbol{P}_N^{\mathrm{m}} = \boldsymbol{I} + \frac{\boldsymbol{Q}_N^{\mathrm{m}}}{2N} + \boldsymbol{R}_N^{\mathrm{m}}, \qquad \boldsymbol{P}_N^{\mathrm{r}} = \boldsymbol{I} + \frac{\boldsymbol{Q}_N^{\mathrm{r}}}{2N} + \boldsymbol{R}_N^{\mathrm{r}},$$

where

$$(\boldsymbol{R}_N^{\mathrm{m}})_{\alpha,\beta} = \begin{cases} -\sum_{\gamma\ne\alpha} R_N^{\mathrm{m}}(\alpha,\gamma) & \text{if } \beta = \alpha, \\ R_N^{\mathrm{m}}(\alpha,\beta) & \text{otherwise,} \end{cases}$$

$$(\boldsymbol{R}_N^{\mathrm{r}})_{\alpha,\beta} = \begin{cases} \sum_{i\in S} R_N^{\mathrm{r}}(\alpha,\alpha-\boldsymbol{e}^i) \quad - \sum_{\beta\notin\{\alpha\}\cup\{\alpha-\boldsymbol{e}^i\}} R_N^{\mathrm{r}}(\alpha,\beta) & \text{if } \beta = \alpha, \\ \quad -R_N^{\mathrm{r}}(\alpha,\alpha-\boldsymbol{e}^i) & \text{if } \beta = \alpha-\boldsymbol{e}^i, \\ R_N^{\mathrm{r}}(\alpha,\beta) & \text{otherwise,} \end{cases}$$

$$(\boldsymbol{Q}_N^{\mathrm{m}})_{\alpha,\beta} = \begin{cases} -\sum_{i\in S} \alpha_i (2Nm_i) \dfrac{N_i^*}{N_i^* - m_i N_i^* - \alpha_i + 1} \\ \quad \times \prod_{k\in S} \prod_{a=0,\dots,\alpha_k-1} \dfrac{N_k^* - m_k N_k^* - a}{N_k^* - a} & \text{if } \beta = \alpha, \\ \alpha_i (2Nm_{i,j}) \dfrac{N_i^*}{N_i^* - m_i N_i^* - \alpha_i + 1} \\ \quad \times \prod_{k\in S} \prod_{a=0,\dots,\alpha_k-1} \dfrac{N_k^* - m_k N_k^* - a}{N_k^* - a} & \text{if } \beta = \alpha - \boldsymbol{e}^i + \boldsymbol{e}^j,\ j \ne i, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2}$$

and $\boldsymbol{Q}_N^{\mathrm{r}}$ is given by

$$(\boldsymbol{Q}_N^{\mathrm{r}})_{\alpha,\beta} = \begin{cases} -\sum_{i\in S} \dbinom{\alpha_i}{2} \left\{ \left(1 + \dfrac{Q_i}{4c_i N}\right) \dfrac{\mathbb{E}[\{\nu_1{}^{(i,r)}\}^2]}{c_i - 1/2N} \right. \\ \qquad\qquad \left. - \dfrac{1}{c_i - 1/2N} \right\} & \text{if } \beta = \alpha, \\ \dbinom{\alpha_i}{2} \left\{ \left(1 + \dfrac{Q_i}{4c_i N}\right) \dfrac{\mathbb{E}[\{\nu_1^{(i,r)}\}^2]}{c_i - 1/2N} - \dfrac{1}{c_i - 1/2N} \right\} & \text{if } \beta = \alpha - \boldsymbol{e}^i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Note that (3.2) and (3.3) follow from (2.5) and (2.16), respectively.

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

Recall that the number of subpopulations is infinite and countable. We consider the following norm of any matrix $\boldsymbol{A}$:

$$||\boldsymbol{A}|| = \sup_{\alpha} \sum_{\beta} |\boldsymbol{A}_{\alpha,\beta}|.$$

By noting that

$$\begin{aligned}
&\sum_{i\in S} \alpha_i m_i \frac{N_i^*}{N_i^* - m_i N_i^* - \alpha_i + 1} \prod_{k\in S} \prod_{a=0,\dots,\alpha_k - 1} \frac{N_k^* - m_k N_k^* - a}{N_k^* - a} \\
&\quad = \sum_{i\in S} \alpha_i m_i \prod_{l=0}^{\alpha_i - 2} \frac{N_i^* - m_i N_i^* - 1}{N_i^* - 1 - l} \prod_{k\in S,\, k\neq i} \prod_{\alpha=0}^{\alpha_k - 1} \frac{N_k^* - m_k N_k^* - a}{N_k^* - a} \\
&\quad \leq \frac{nM}{4N},
\end{aligned}$$

we can see that

$$||\boldsymbol{Q}_N^{\mathrm{m}}|| \leq nM. \tag{3.4}$$

By (2.2) and (3.3), we have

$$\begin{aligned}
||\boldsymbol{Q}_N^{\mathrm{r}}|| &\leq 2 \sum_{i\in S} \binom{\alpha_i}{2} \left\{ \left(1 + \frac{Q_i}{4c_i N}\right) \frac{\mathbb{E}[\{\nu_1^{(i,r)}\}^2]}{c_i - 1/2N} - \frac{1}{c_i - 1/2N} \right\} \\
&\leq 2 \sum_{i\in S} \binom{\alpha_i}{2} \left\{ \left(1 + \frac{C}{4N}\right) \frac{\mathbb{E}[\{\nu_1^{(i,r)}\}^2]}{1 - 1/2N} \right\} \\
&\leq 4 \binom{n}{2} (1 + C) K_2^* \\
&< \infty,
\end{aligned} \tag{3.5}$$

$$||\boldsymbol{R}_N^{\mathrm{m}}|| \leq \frac{M^2 n^2}{8N^2} \quad \text{and} \quad ||\boldsymbol{R}_N^{\mathrm{r}}|| \leq \frac{n^2 (C_1 + C_2)}{N^2}, \tag{3.6}$$

where the following two statements deduced from (2.10), (2.14), and (2.15) are used:

$$\sum_{i\in S} R_N^{\mathrm{r}}(\alpha, \alpha - \boldsymbol{e}^i) \leq \binom{n}{2} R_N^{\mathrm{r}}(\alpha) \leq \binom{n}{2} \frac{(C_1 + C_2)}{N^2},$$

$$\sum_{\beta\neq\alpha} R_N^{\mathrm{r}}(\alpha, \beta) \leq \left\{ \binom{n}{2} + 1 \right\} R_N^{\mathrm{r}}(\alpha) \leq \left\{ \binom{n}{2} + 1 \right\} \frac{(C_1 + C_2)}{N^2}.$$

Below we will show the convergence of the finite-dimensional distributions. Next, we consider

$$\boldsymbol{P}_N = \boldsymbol{I} + \frac{\boldsymbol{Q}_N + \boldsymbol{\pi}_N}{2N}, \tag{3.7}$$

where $\boldsymbol{Q}_N = \boldsymbol{Q}_N^{\mathrm{m}} + \boldsymbol{Q}_N^{\mathrm{r}}$ and

$$\boldsymbol{\pi}_N = 2N \left(\boldsymbol{R}_N^{\mathrm{r}} + \boldsymbol{R}_N^{\mathrm{m}} + \boldsymbol{R}_N^{\mathrm{r}} \boldsymbol{R}_N^{\mathrm{m}} + \frac{\boldsymbol{Q}_N^{\mathrm{r}} \boldsymbol{R}_N^{\mathrm{m}}}{2N} + \frac{\boldsymbol{R}_N^{\mathrm{r}} \boldsymbol{Q}_N^{\mathrm{m}}}{2N} + \frac{\boldsymbol{Q}_N^{\mathrm{r}} \boldsymbol{Q}_N^{\mathrm{m}}}{4N^2} \right). \tag{3.8}$$

Clearly, $\lim_{N\to\infty} \boldsymbol{Q}_N = \boldsymbol{Q}$; namely, $\lim_{N\to\infty} (\boldsymbol{Q}_N)_{\alpha,\beta} = \boldsymbol{Q}_{\alpha,\beta}$ holds for any $\alpha, \beta \in E$, where the matrix $\boldsymbol{Q}$ is defined by (1.1).

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

Employing the above results, we will discuss the rescaled process $\{\alpha^{(N)}([2Nt])\colon t \geq 0\}$ in what follows.

Theorem 3.1. *Under conditions (A.1)–(A.6), the following equality holds:*

$$\lim_{N\to\infty} \boldsymbol{P}_N^{[2Nt]} = \mathrm{e}^{t\boldsymbol{Q}}.$$

Here $\boldsymbol{P}_N$ *and* $\boldsymbol{Q}$ *are defined by (3.1) and (1.1), respectively.*

Proof. Since

$$||\boldsymbol{Q}|| \leq nM + 2\sigma^2 \binom{n}{2} < \infty,$$

we have

$$|(\mathrm{e}^{t\boldsymbol{Q}})_{\alpha,\beta}| \leq \sum_{v=0}^{\infty} \frac{t^v |(\boldsymbol{Q}^v)_{\alpha,\beta}|}{v!} \leq \sum_{v=0}^{\infty} \frac{t^v ||\boldsymbol{Q}||^v}{v!} = \mathrm{e}^{t||\boldsymbol{Q}||} < \infty \quad \text{for any } \alpha, \beta \in E.$$

Thus, we see that $\mathrm{e}^{t\boldsymbol{Q}} = \sum_{v=0}^{\infty} t^v \boldsymbol{Q}^v / v!$ exists. Obviously,

$$\begin{aligned} \boldsymbol{P}_N^{[2Nt]} &= \left\{\boldsymbol{I} + \frac{\boldsymbol{Q}_N + \boldsymbol{\pi}_N}{2N}\right\}^{[2Nt]} \\ &= \sum_{v=0}^{[2Nt]} \binom{[2Nt]}{v} \left(\frac{1}{2N}\right)^v (\boldsymbol{Q}_N + \boldsymbol{\pi}_N)^v \\ &= \sum_{v=0}^{[2Nt]} \frac{[2Nt]([2Nt]-1)\cdots([2Nt]-v+1)}{(2N)^v} \frac{(\boldsymbol{Q}_N + \boldsymbol{\pi}_N)^v}{v!} \end{aligned}$$

holds. Thus, we have

$$(\boldsymbol{P}_N^{[2Nt]})_{\alpha,\beta} = \sum_{v=0}^{\infty} a_{v,N}, \tag{3.9}$$

where

$$a_{v,N} = \mathbf{1}_{\{v\leq[2Nt]\}} \frac{[2Nt]([2Nt]-1)\cdots([2Nt]-v+1)}{(2N)^v} \frac{((\boldsymbol{Q}_N + \boldsymbol{\pi}_N)^v)_{\alpha,\beta}}{v!} \tag{3.10}$$

and

$$\mathbf{1}_{\{v\leq[2Nt]\}} = \begin{cases} 1 & \text{if } v \leq [2Nt], \\ 0 & \text{otherwise.} \end{cases}$$

By (3.4) and (3.5), we have

$$||\boldsymbol{Q}_N|| \leq ||\boldsymbol{Q}_N^{\mathrm{m}}|| + ||\boldsymbol{Q}_N^{\mathrm{r}}|| \leq C^* < \infty, \tag{3.11}$$

where

$$C^* := nM + 4\binom{n}{2}(1+C)K_2^*. \tag{3.12}$$

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

By (3.4)–(3.6) and (3.8), we obtain

$$||\boldsymbol{\pi}_N|| \le \frac{k_1}{N}, \qquad ||\boldsymbol{A}_N|| \le \frac{k_2}{N}, \tag{3.13}$$

where $\boldsymbol{A}_N$ is defined by $(\boldsymbol{Q}_N + \boldsymbol{\pi}_N)^v = \boldsymbol{Q}_N^v + \boldsymbol{A}_N$, and the constants k_1 and k_2 do not depend on N. Now, a matrix $\boldsymbol{V}$ is defined by

$$(\boldsymbol{V})_{\alpha,\beta} = \begin{cases} nK\sum_{j\ne i} q^*_{i,j} + 2\binom{\alpha_i}{2}\{(1+C)K_2^*\} & \text{if } \beta = \alpha, \\ nKq^*_{i,j} & \text{if } \beta = \alpha - \boldsymbol{e}^i + \boldsymbol{e}^j\ (i \ne j), \\ 2\binom{\alpha_i}{2}\{(1+C)K_2^*\} & \text{if } \beta = \alpha - \boldsymbol{e}^i, \\ 0 & \text{otherwise.} \end{cases}$$

Then, it follows that

$$||\boldsymbol{V}|| \le 2nK \sup_{i\in S}\sum_{j\ne i} q^*_{i,j} + 4\binom{n}{2}(1+C)K_2^* < \infty, \qquad |(\boldsymbol{Q}_N)_{\alpha,\beta}| \le (\boldsymbol{V})_{\alpha,\beta}.$$

By the dominated convergence theorem, we have

$$\begin{aligned} \lim_{N\to\infty}(\boldsymbol{Q}_N^v)_{\alpha,\beta} &= \lim_{N\to\infty}\sum_{\gamma_1,\gamma_2,\dots,\gamma_{v-1}} (\boldsymbol{Q}_N)_{\alpha,\gamma_1}(\boldsymbol{Q}_N)_{\gamma_1,\gamma_2}\cdots(\boldsymbol{Q}_N)_{\gamma_{v-1},\beta} \\ &= \sum_{\gamma_1,\gamma_2,\dots,\gamma_{v-1}} (\boldsymbol{Q})_{\alpha,\gamma_1}(\boldsymbol{Q})_{\gamma_1,\gamma_2}\cdots(\boldsymbol{Q})_{\gamma_{v-1},\beta} \\ &= (\boldsymbol{Q}^v)_{\alpha,\beta}. \end{aligned}$$

Thus, we have $\lim_{N\to\infty}(\boldsymbol{Q}_N + \boldsymbol{\pi}_N)^v = \boldsymbol{Q}^v$ and

$$\lim_{N\to\infty} a_{v,N} = \frac{t^v(\boldsymbol{Q}^v)_{\alpha,\beta}}{v!}, \qquad \alpha, \beta \in S \tag{3.14}$$

for any v. By (3.10), (3.11), and (3.13), we have, for sufficiently large N,

$$|a_{v,N}| \le \frac{t^v||\boldsymbol{Q}_N + \boldsymbol{\pi}_N||^v}{v!} \le \frac{t^v(C^*+1)^v}{v!}, \qquad \alpha, \beta \in S \tag{3.15}$$

for any $N \in \mathbb{N}$. By (3.9), (3.10), (3.14), (3.15), and the dominated convergence theorem, we obtain

$$\lim_{N\to\infty}(\boldsymbol{P}_N^{[2Nt]})_{\alpha,\beta} = \sum_{v=0}^{\infty}\frac{t^v(\boldsymbol{Q}^v)_{\alpha,\beta}}{v!} = (\mathrm{e}^{t\boldsymbol{Q}})_{\alpha,\beta} \quad \text{for any } \alpha, \beta \in E.$$

This completes the proof. □

This statement implies that the finite-dimensional distributions of the process $\{\alpha^{(N)}([2Nt])\colon t \ge 0\}$ converge to those of the structured coalescent $\{\alpha(t)\colon t \ge 0\}$ as $N \to \infty$, because E is countable.

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

4. Weak convergence to the structured coalescent process

We regard E as a subspace of $\boldsymbol{R}^S$ (where $\boldsymbol{R}$ is the set of real numbers), endowed with the norm

$$||\boldsymbol{X}|| = \sup_{i \in S} |x_i| \quad (\boldsymbol{X} = (x_i)_{i \in S} \in \boldsymbol{R}^S).$$

Note that E is a separable complete metric space with this norm.

Theorem 4.1. *Suppose that conditions (A.1)–(A.6) hold. Then, the process* $\{\alpha^{(N)}([2Nt])\colon t \geq 0\}$ *converges weakly in* $D_E[0, \infty)$ *to the structured coalescent* $\{\alpha(t)\colon t \geq 0\}$*; namely,*

$$\{\alpha^{(N)}([2Nt])\colon t \geq 0\} \xrightarrow{\mathrm{w}} \{\alpha(t)\colon t \geq 0\}.$$

Proof. According to Ethier and Kurtz (1986, Chapter 3, Corollary 7.4), the relative compactness of $\{\alpha^{(N)}([2Nt])\}$ is guaranteed if we prove the following two conditions:

(i) for every $\eta > 0$ and $t \geq 0$, there exists a compact set $\Gamma_{\eta,t} \subset E$ such that

$$\liminf_{N \to \infty} P\{\alpha^{(N)}([2Nt]) \in \Gamma_{\eta,t}\} \geq 1 - \eta;$$

(ii) for every $\eta > 0$ and $T \geq 0$, there exists $\delta > 0$ such that

$$\limsup_{N \to \infty} P\{\omega'(\alpha^{(N)}([2Nt]), \delta, T) \geq \eta\} \leq \eta,$$

where $\omega'(\alpha^{(N)}([2Nt]), \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s,t \in [t_{i-1}, t_i)} ||\alpha^{(N)}([2Ns]) - \alpha^{(N)}([2Nt])||$, with the sequences $\{t_i\}$ satisfying $0 = t_0 < t_1 < \cdots < t_{k-1} < T \leq t_k$ and $\min_i (t_i - t_{i-1}) > \delta$.

First, we will verify condition (i). Let t be fixed. We consider a sequence of finite subsets $\{S_m\}_{m \in \mathbb{Z}_+ \setminus \{0\}}$ satisfying $S_1 \subset S_2 \subset \cdots \subset S_m \subset S_{m+1} \subset \cdots$ and $\bigcup_{m=1}^{\infty} S_m = S$. Define $\Gamma_m = \{\alpha \in E, \alpha_i = 0 \text{ if } i \notin S_m\}$ for $m \in \mathbb{Z}_+ \setminus \{0\}$ which is a finite subset of E. This implies that $\Gamma_1 \subset \Gamma_2 \subset \cdots$ and $\bigcup_{m=1}^{\infty} \Gamma_m = E$. Then, we have

$$\lim_{n \to \infty} P\{\alpha(t) \in \Gamma_n\} = P\left\{\alpha(t) \in \bigcup_{n=1}^{\infty} \Gamma_n\right\} = P\{\alpha(t) \in E\} = 1.$$

By the convergence of finite-dimensional distributions, we have

$$P\{\alpha(t) \in \Gamma_n\} = \lim_{N \to \infty} P\{\alpha^{(N)}([2Nt]) \in \Gamma_n\}.$$

Hence, we see that condition (i) is satisfied.

We will prove (ii). We define p_N by

$$p_N = \frac{C^* + 1}{2N}, \qquad N \in \mathbb{Z}_+, \tag{4.1}$$

where C^* is defined by (3.12). If N is large enough, we have $p_N < 1$. For each sufficiently large N, we define $(Z^{(N)}, \xi^{(N)}) = \{(Z^{(N)}(\tau), \xi^{(N)}(\tau)), \tau = 0, 1, 2, \ldots\}$ by the Markov chain

https://doi.org/10.1017/jpr.2016.16 Published online by Cambridge University Press

with state space $(\mathbb{Z}_+ \setminus \{0\}) \times E$ and transition probabilities

$$P\{(Z^{(N)}(\tau+1), \xi^{(N)}(\tau+1)) = (j, \beta) \mid (Z^{(N)}(\tau), \xi^{(N)}(\tau)) = (i, \alpha)\}$$
$$= \begin{cases} 1 - p_N & \text{if } j = i \text{ and } \beta = \alpha, \\ p_N - \displaystyle\sum_{\gamma \in E,\, \gamma \neq \alpha} P_N(\gamma \mid \alpha) & \text{if } j = i+1 \text{ and } \beta = \alpha, \\ P_N(\beta \mid \alpha) & \text{if } j = i+1 \text{ and } \beta \neq \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

where $P_N(\beta \mid \alpha)$ is the transition probability of the ancestral process $\{\alpha^{(N)}(\tau)\}_{\tau \in \mathbb{Z}_+}$ from α to β in one generation. Using (3.7), (3.11), (3.13), (4.1), and the restrictions made on N, we have, for any $\alpha \in E$,

$$\sum_{\gamma \in E,\, \gamma \neq \alpha} P_N(\gamma \mid \alpha) = \frac{\sum_{\gamma \neq \alpha} (\boldsymbol{Q}_N + \boldsymbol{\pi}_N)_{\alpha,\gamma}}{2N} \leq p_N.$$

Since the distribution of $\xi^{(N)}$ is that of the process $\alpha^{(N)}$, we have

$$P\{\omega'(\alpha^{(N)}([2N\cdot]), \delta, T) \geq \eta\} = P\{\omega'(\xi^{(N)}([2N\cdot]), \delta, T) \geq \eta\}$$

for every $\eta > 0$ and $T > 0$. The process $(Z^{(N)}, \xi^{(N)})$ jumps with probability p_N at every generation; at each jump, $Z^{(N)}$ increases by 1. The construction is such that every time $\xi^{(N)}$ jumps, $Z^{(N)}$ jumps as well. We denote the ith jump time of $Z^{(N)}$ by ρ_i, $i = 0, 1, 2, \dots$. Obviously, $0 = \rho_0 < \rho_1 < \cdots$. Let us denote τ_i (interjump times)$= \rho_i - \rho_{i-1}$, $i \in \mathbb{Z}_+$. Let τ_i be mutually independent and each τ_i is geometrically distributed with mean $1/p_N$. Now, fix $\eta > 0$ and $T > 0$. Here, we will prove that

$$\{\rho_J \geq 2NT \text{ and } \tau_i > 2N\delta,\ i = 1, 2, \dots, J\} \subset \{\omega'(\xi^{(N)}([2N\cdot]), \delta, T) < \eta\} \tag{4.2}$$

for any $J \in \mathbb{Z}_+$ and $\delta > 0$. To show (4.2), we will discuss the following argument for any fixed element of the left-hand side of (4.2). Denoting $k_N = \min\{i : \rho_i \geq 2NT\}$, we have $1 \leq k_N \leq J$ and the partition

$$t_i = \frac{\rho_i}{2N} \quad (i = 0, 1, \dots, k_N)$$

satisfies $0 = t_0 < t_1 < \cdots < t_{k_{N-1}} < T \leq t_{k_N}$ and $t_i - t_{i-1} > \delta\ (i = 1, \dots, k_N)$. As the process $(Z^{(N)}(\tau), \xi^{(N)}(\tau))$ is constant for $\rho_{i-1} \leq \tau < \rho_i$, $i = 1, 2, \dots, J$, we have

$$\omega'(\xi^{(N)}([2N\cdot]), \delta, T) = 0.$$

Thus, we obtain (4.2). Hence, for every $J \in \mathbb{Z}_+$ and $\delta > 0$, we obtain

$$P\{\omega'(\xi^{(N)}([2N\cdot]), \delta, T) < \eta\} \geq P\{\rho_J \geq 2NT \text{ and } \tau_i > 2N\delta,\ i = 1, 2, \dots, J\}.$$

Thus, in order to prove condition (ii) it is sufficient to find $J \in \mathbb{Z}_+$ and $\delta > 0$ such that

$$\liminf_{N \to \infty} P\{\rho_J \geq 2NT \text{ and } \tau_i > 2N\delta,\ i = 1, 2, \dots, J\} \geq 1 - \eta.$$

Now, we have

$$\begin{aligned} &P\{\rho_J \geq 2NT \text{ and } \tau_i > 2N\delta,\ i = 1, 2, \dots, J\} \\ &\quad = P\{\rho_J \geq 2NT \mid \tau_i > 2N\delta,\ i = 1, 2, \dots, J\} P\{\tau_i > 2N\delta,\ i = 1, 2, \dots, J\} \\ &\quad = P\{\rho_J \geq 2NT \mid \tau_i > 2N\delta,\ i = 1, 2, \dots, J\} (P\{\tau_1 > 2N\delta\})^J. \end{aligned}$$

Since $\tau_i,\ i = 1, 2, \dots, J$ are independent and identically distributed random variables such that $P\{\tau_i = k\} = p_N(1 - p_N)^{k-1},\ k \geq 1$, and $\rho_J = \sum_{i=1}^J \tau_i$, we have

$$P\{\rho_J \geq 2NT \mid \tau_i > 2N\delta,\ i = 1, 2, \dots, J\} \geq P\{\rho_J \geq 2NT\}.$$

Since we have

$$P\{\rho_J \geq 2NT\} = P\{Z^{(N)}([2NT]) - Z^{(N)}(0) < J\},$$

we obtain

$$P\{\omega'(\alpha^{(N)}([2N\cdot]), \delta, T) \leq \eta\} \geq P\{Z^{(N)}([2NT]) - Z^{(N)}(0) < J\} \left(P\left\{ \frac{\tau_1}{2N} > \delta \right\} \right)^J \tag{4.3}$$

for every $J \in \mathbb{Z}_+$ and $\delta > 0$. Because the distribution of τ_i is geometric and $\tau_1/2N$ converges in distribution as $N \to \infty$ to an exponentially distributed random variable X with mean $1/(C^* + 1)$ and the distribution of $Z^{(N)}([2NT]) - Z^{(N)}(0)$ is binomial with parameters $[2NT]$ and p_N, it follows that as $N \to \infty$, $Z^{(N)}([2NT]) - Z^{(N)}(0)$ converges to a Poisson distributed random variable Z with mean $T(C^* + 1)$. From (4.3), it follows that

$$\liminf_{N \to \infty} P\{\omega'(\alpha^{(N)}([2N\cdot]), \delta, T) \geq \eta\} \geq P\{Z < J\}(P\{X > \delta\})^J \tag{4.4}$$

since the right-hand since of (4.4) can be made arbitrary close to 1 for sufficiently large J and sufficiently small $\delta > 0$ depending on J, we see that the proof of (ii) is complete.

Thus, by Ethier and Kurtz (1986, Chapter 3, Theorem 7.8), the process $\{\alpha^{(N)}([2Nt]),\ t \geq 0\}$ converges weakly in $D_E[0, \infty)$ to the structured coalescent $\{\alpha(t),\ t \geq 0\}$ defined in the introduction. □

Acknowledgement

We are grateful to the anonymous referee for a careful reading and useful suggestions.

References

Bahlo, M. and Griffiths, R. C. (2000). Inference from gene trees in a subdivided population. *Theoret. Pop. Biol.* **57,** 79–95.

Bahlo, M. and Griffiths, R. C. (2001). Coalescence times for two genes from a subdivided population. *J. Math. Biol.* **43,** 397–410.

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

Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* John Wiley, New York.

Herbots, H. M. (1994). Stochastic models in population genetics: genealogy and genetic differentiation in structured populations. Doctoral Thesis, University of London.

Herbots, H. M. (1997). The structured coalescent. In *Progress in Population Genetics and Human Evolution* (IMA Vol. Math. Appl. **87**). Springer, New York, pp. 231–255.

Hössjer, O. (2011). Coalescence theory for a general class of structured populations with fast migration. *Adv. Appl. Prob.* **43,** 1027–1047.

KAJ, I., KRONE, S. M. AND LASCOUX, M. (2001). Coalescent theory for seed bank models. *J. Appl. Prob.* **38,** 285–300.

KINGMAN, J. F. C. (1982a). On the genealogy of large populations. In *Essays in Statistical Science* (J. Appl. Prob. Spec. Vol **19A**), Applied Probability Trust, Sheffield, pp. 27–43.

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

MÖHLE, M. (1998). A convergence theorem for Markov chains arising in population genetics and the coalescent with selfing. *Adv. Appl. Prob.* **30,** 493–512.

NATH, H. B. AND GRIFFITHS, R. C. (1993). The coalescent in two colonies with symmetric migration. *J. Math. Biol.* **31,** 841–851.

NORDBORG, M AND KRONE, S. M. (2002). Separation of timescales and convergence to the coalescent in structured populations. In *Modern Developments In Theoretical Population Genetics,* Oxford University Press, pp. 194–232.

NOTOHARA, M. (1990). The coalescent and the genealogical process in geographically structured population. *J. Math. Biol.* **29,** 59–75.

NOTOHARA, M. (1993). The strong-migration limit for the genealogical process in geographically structured populations. *J. Math. Biol.* **31,** 115–122.

NOTOHARA, M. (1997). The number of segregating sites in a sample of DNA sequences from a geographically structured population. *J. Math. Biol.* **36,** 188–200.

NOTOHARA, M. (2001). The structured coalescent process with weak migration. *J. Appl. Prob.* **38,** 1–17.

NOTOHARA, M. (2010). An application of the central limit theorem to coalescence times in the structured coalescent model with strong migration. *J. Math. Biol.* **61,** 695–714.

POLLAK, E. (2011). Coalescent theory for age-structured random mating populations with two sexes. *Math. Biosci.* **233,** 126–134.

SAGITOV, S. AND JAGERS, P. (2005). The coalescent effective size of age-structured populations. *Ann. Appl. Prob.* **15,** 1778–1797.

SAMPSON, K. Y. (2006). Structured coalescent with nonconservative migration. *J. Appl. Prob.* **43,** 351–362.

SHIGA, T. (1980a). An interacting system in population genetics. *J. Math. Kyoto Univ.* **20,** 213–242.

SHIGA, T. (1980b). An interacting system in population genetics. II *J. Math. Kyoto Univ.* **20,** 723–733.

TAJIMA, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105,** 437–460.

TAKAHATA, N. (1988). The coalescent in two partially isolated diffusion populations. *Genetical Res.* **52,** 213–222.

WAKELEY, J. (1998). Segregating sites in Wright's island model. *Theoret. Pop. Biol.* **53,** 166–174.

WAKELEY, J. (2009). *Coalescent Theory: An Introduction.* Roberts, Greenwood Village, CO.

WILKINSON-HERBOTS, H. M. (1998). Genealogy and subpopulation differentiation under various models of population structure. *J. Math. Biol.* **37,** 535–585.