J. Appl. Prob. **52,** 538–557 (2015)
Printed in England
© *Applied Probability Trust* 2015

THE COALESCENT IN PERIPATRIC METAPOPULATIONS

AMAURY LAMBERT,* *Université Pierre et Marie Curie and Collège de France*
CHUNHUA MA,** *Université Pierre et Marie Curie, Collège de France and Nankai University*

Abstract

We consider a dynamic metapopulation involving one large population of size N surrounded by colonies of size $\varepsilon_N N$, usually called peripheral isolates in ecology, where $N \to \infty$ and $\varepsilon_N \to 0$ in such a way that $\varepsilon_N N \to \infty$. The main population, as well as the colonies, independently send propagules to found new colonies (emigration), and each colony independently, eventually merges with the main population (fusion). Our aim is to study the genealogical history of a finite number of lineages sampled at stationarity in such a metapopulation. We make assumptions on model parameters ensuring that the total outer population has size of the order of N and that each colony has a lifetime of the same order. We prove that under these assumptions, the scaling limit of the genealogical process of a finite sample is a censored coalescent where each lineage can be in one of two states: an inner lineage (belonging to the main population) or an outer lineage (belonging to some peripheral isolate). Lineages change state at constant rate and (only) inner lineages coalesce at constant rate per pair. This two-state censored coalescent is also shown to converge weakly, as the landscape dynamics accelerate, to a time-changed Kingman coalescent.

Keywords: Censored coalescent; metapopulation; weak convergence; peripheral isolate; population genetics; peripatric speciation

2010 Mathematics Subject Classification: Primary 60K35
Secondary 60J05; 60G10; 92D10; 92D15

1. Introduction

Many plant and animal populations in nature are highly fragmented, and this fragmentation plays a prominent role in the context of adaptation and speciation. Indeed, the emergence of new species is usually thought to be driven by geographical processes [5]. First, *allopatric speciation* occurs when various subpopulations belonging to the same initial species are separated by a geographical barrier that prevents hybridization between them (gene flow) and allows them to diverge (genetical differentiation) by local adaptation. Second, *parapatric speciation* is a version of allopatric speciation where local adaptation is mediated by the existence of an environmental gradient (resource availability, environmental conditions). Third, when a species is present in one large, panmictic population surrounded by small colonies, usually called peripheral isolates, it is believed that the combination of founder events and of local adaptation to borderline environmental conditions leads to the formation of new species within the isolates. This phenomenon is called *peripatric speciation*. We aim to study the genealogy of populations

Received 20 March 2013; revision received 23 June 2014.
* Postal address: Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, CNRS UMR 7599, Paris, France. Email address: amaury.lambert@upmc.fr
** Postal address: School of Mathematical Sciences and LPMC, Nankai University, Tianjin, P. R. China. Email address: mach@nankai.edu.cn

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

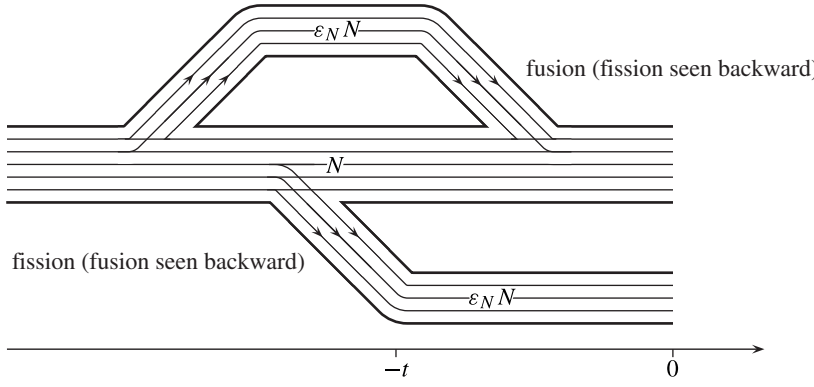

FIGURE 1: Dynamic metapopulation.

embedded in such a spatial context. The present study would give a first step for our future work in the field of speciation modeling.

Population dynamic models specifying explicitly the spatial context are called *metapopulation models* (see Hanski and Gilpin [8]). Typical models include: island model, isolation by distance, stepping stone models, and extinction–recolonization models. From the point of view of speciation, all these models suffer from the same defect: they assume a given, constant number of subpopulations in the metapopulation, with fixed migration rates between them. As one of the authors of the present paper suggested (see Lambert [13]), an alternative method would consist of considering a species as 'spread out on a randomly evolving number of locations, allowing for repeated fragmentations of colonies, colonizations of new locations, as well as secondary contacts between subpopulations'. This author and others have designed such dynamic landscape models (see [2], [3], and [11]), but usually in a detailed ecological context whose study is only possible through numerical simulations (with the exception of [1]).

Here, we propose a mathematical study of a dynamic landscape of the peripatric type. More specifically, we consider a dynamic population subdivision which involves one large main population surrounded by a random number of small peripheral isolates, that we will call colonies for simplicity. The size of the main population is constant equal to N, the size of each colony is constant equal to $\varepsilon_N N$ and the reproduction mechanism in each population is given by the Moran model. The number of colonies at time t is denoted by $\xi_N(t)$. The landscape dynamics is as follows (see Figure 1):

- each individual sends independently $\varepsilon_N N$ offspring to found a new colony at constant rate θ_N if dwelling in the mainland, or at constant rate β_N if dwelling in a colony;
- each colony independently merges again with the main population at rate $\gamma_N \xi_N^{\alpha-1}$, where $\alpha \geq 1$ and γ_N is the fusion rate; at such a so-called fusion time, $\varepsilon_N N$ individuals among the new $(1+\varepsilon_N)N$ individuals of the main population are chosen uniformly and simultaneously killed in order to keep its size constant.

Note that $(\xi_N(t); t \geq 0)$ is a density-dependent birth–death process with immigration. The parameter α is meant to model the competition for space, since the fusion rate per colony grows with the number of colonies. This density-dependence disappears if α is chosen to equal 1.

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

The main purpose of this paper is to investigate the genealogy of a finite sample of lineages in the above peripatric metapopulation model. We will show that the history of such a sample, viewed backward in time, can be approximated, as $N \to \infty$ under certain assumptions, by a two-state censored coalescent, where the state of a lineage can be inner (lying in the main population) or outer (lying in a colony). Lineages change state at a constant rate per lineage, but only inner lineages can coalesce, at a constant rate per pair of lineages, as in Kingman's coalescent [12].

A two-state censored coalescent can be viewed as a new type of structured coalescent. The structured coalescent describes the ancestral genealogical process of a sample of lineages in a subdivided population connected by migration. The coalescent on two subpopulations was considered by Takahata [17]; for a finite number of subpopulations by Notohara [15], and placed in a rigorous framework by Herbots [9]. To date, there have been a number of works dealing with the structured coalescent arising in various special types of metapopulations; see Nordborg and Krone [14] and Eldon [6] and the references therein. Our results show that new types of structured coalescents can arise in some specific dynamic metapopulations.

We now provide the heuristic methods used to obtain our result. We assume that $N \to \infty$ and $\varepsilon_N \to 0$ in such a way that $\varepsilon_N N \to \infty$, so that the size of colonies is large but negligible compared to the main population (assumption A). It is known that in a Moran model, inner lineages coalesce at constant rate per pair when time is rescaled by N (Kingman's coalescent [12]). We make assumptions on the parameters ensuring that all events changing the configuration of ancestral lineages occur on this time- scale. This can only be done to the exception of coalescences in colonies, which happen instantaneously in the new time-scale, leading to outer lineages which always all lie in different colonies. Also, in order to have a total outer population size of the order of N, we need to have a number of colonies of the order of ε_N^{-1}. This can be achieved by the following choice of parameters (assumption B). The *per capita* emigration rate θ_N in the mainland is taken equal to

$$\theta_N = \frac{\theta}{\varepsilon_N N^2}.$$

The *per capita* emigration rate β_N in one colony is taken equal to

$$\beta_N = \frac{\beta}{\varepsilon_N N^2}$$

and the fusion rate γ_N is taken equal to

$$\gamma_N = \gamma \frac{\varepsilon_N^{\alpha-1}}{N},$$

where $\beta < \gamma$ if $\alpha = 1$. Under these assumptions, the number of colonies is asymptotically deterministic, equal to $\varepsilon_N^{-1}\kappa$, where κ is the unique solution of $\beta z + \theta = \gamma z^\alpha$.

Looking backward in time, the rate at which a single inner lineage changes state is the rate at which a single lineage is taken in a fusion event, which happens at rate

$$\frac{\varepsilon_N}{1+\varepsilon_N}(\gamma_N \xi_N^{\alpha-1})\xi_N \approx \frac{\varepsilon_N}{1+\varepsilon_N}\gamma_N \left(\frac{\kappa}{\varepsilon_N}\right)^\alpha,$$

which is equivalent to $\gamma(\kappa^\alpha/N)$ as $N \to \infty$. As a consequence, in the new time-scale, inner lineages become outer lineages at rate $\gamma\kappa^\alpha$.

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Note that the lifetime of a colony is exponential with parameter

$$\gamma_N \xi_N^{\alpha-1} \approx \gamma_N \varepsilon_N^{1-\alpha} \kappa^{\alpha-1}.$$

Then the rate at which each colony coalesces with the mainland is

$$\gamma_N \xi_N^{\alpha-1} \frac{N}{\xi_N \varepsilon_N N + N} \approx \frac{\gamma \kappa^{\alpha-1}}{N(\kappa+1)} = \frac{\beta\kappa + \theta}{\kappa} \frac{1}{\kappa+1}$$

and the rate at which each colony coalesces with another given colony is

$$\gamma_N \xi_N^{\alpha-1} \frac{\varepsilon_N N}{\xi_N \varepsilon_N N + N} = o\left(\frac{1}{N}\right).$$

Thus, the following probability vanishes: the probability that two given colonies will coalesce into a colony before one of them coalesces with the mainland. Also note that the probability that two lineages are taken in the same fusion with the mainland vanishes, so that no two lineages can lie within the same colony. As a consequence, in the new time-scale, outer lineages are not allowed to coalesce and they become inner lineages at rate $(\beta\kappa + \theta)/\kappa(\kappa+1)$. By making these heuristics rigorous we obtain the results stated in Theorem 3.1. Namely, the genealogical history of a finite sample of lineages, seen as a process backward in time, converges weakly (except at time 0, where instantaneous coalescences within colonies makes the limiting process not right-continuous) to the following two-state censored coalescent. Inner lineages coalesce at constant rate 1 per pair, and lineages change type at constant rate per lineage: inner lineages become outer lineages at rate $\gamma\kappa^\alpha$ and outer lineages become inner lineages at rate $(\beta\kappa + \theta)/\kappa(\kappa+1)$.

The paper is organized as follows. In Section 2 we provide a detailed description of our dynamic metapopulation model in forward and backward time. The main result, Theorem 3.1, is stated in Section 3. In addition, we also prove that under fast landscape dynamics, the censored coalescent converges weakly to a time-changed version of the Kingman coalescent [12]. Finally, Section 3.3 is dedicated to the formal proofs of the above results.

2. Metapopulation model

2.1. Forward dynamics

Let $N \in \mathbb{N}$ with $\mathbb{N} := \{0, 1, 2, \ldots\}$ and let ε_N be any positive number such that $\varepsilon_N N \in \mathbb{N}$. Let θ_N, γ_N and α be positive constants. Consider a dynamic metapopulation model involving one large population of size N, called main population, and a random number of small populations, called colonies, of size $\varepsilon_N N$. The main population and the colonies periodically send propagules (or emigrants) that found new colonies and ultimately each colony merges again with the main population. A further assumption is as follows. See Figure 1 for an illustration.

Assumption 2.1. (a) *The number of colonies, denoted by* $\{\xi_N(t) : t \geq 0\}$, *evolves as a density-dependent birth–death process with immigration and the transition rates are given by*

$$\begin{aligned} j &\to j+1 \quad \text{at rate } N\theta_N + \varepsilon_N N \beta_N j, \\ j &\to j-1 \quad \text{at rate } \gamma_N j^\alpha. \end{aligned} \tag{2.1}$$

When $\alpha = 1$, *we require that* $\beta < \gamma$. *In this case the process* $\{\xi_N(t)\}$ *is reduced to a subcritical birth–death branching process with immigration. It follows from Kelly [10] that* $\{\xi_N(t)\}$ *with*

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

any initial value has the stationary distribution π_N given by

$$\pi_N(0) = \left(1 + \sum_{j=1}^{\infty} w_{N,j}\right)^{-1} \quad \text{and} \quad \pi_N(k) = w_{N,j}\pi_N(0) \quad \text{for } k \geq 1, \tag{2.2}$$

where

$$w_{N,j} = \frac{N\theta_N}{\gamma_N j^\alpha} \prod_{i=1}^{j-1} \frac{N\theta_N + \varepsilon_N N\beta_N i}{\gamma_N i^\alpha}.$$

We assume that $\xi_N(0)$ is distributed as π_N. Then $\{\xi_N(t)\}$ is a stationary Markov chain. Let $(P_t^N)_{t\geq 0}$ be its semigroup. For any finite set $\{t_1 < t_2 < \cdots < t_n\} \subset \mathbb{R}$ define the probability measure on $\mathbb{N}$ by

$$\eta^N_{t_1,t_2,\ldots,t_n}(j_1, j_2, \ldots, j_n) = \pi_N(j_1)P^N_{t_2-t_1}(j_1, j_2)\cdots P^N_{t_n-t_{n-1}}(j_{n-1}, j_n). \tag{2.3}$$

Then $\{\eta^N_{t_1,t_2,\ldots,t_n} : t_1 < t_2 < \cdots < t_n \in \mathbb{R}\}$ is a consistent family. By Kolmogorov's theorem, there is a stochastic process $\{\xi_N(t) : t \in \mathbb{R}\}$ with finite-dimensional distributions given by (2.3). Clearly, $\{\xi_N(t) : t \in \mathbb{R}\}$ is a stationary Markov chain with one-dimensional marginal distribution π_N and transition semigroup $(P_t^N)_{t\geq 0}$.

(b) *At the jump times of $\xi_N(t)$ from j to $j+1$, one individual, chosen uniformly at random from the population (including the mainland and all colonies), gives birth to $\varepsilon_N N$ emigrant offspring individuals which found a new colony. We refer to such an event as 'emigration' (of new colonies) or 'fission'. Note that our assumptions amount to saying that each individual independently founds a new colony at rate θ_N if dwelling in the mainland or at rate β_N if dwelling in a colony.*

(c) *At the jump times of $\xi_N(t)$ from j to $j-1$, one colony is chosen at random from the j current colonies and all the $\varepsilon_N N$ individuals within this colony immediately migrate back into the main population. We refer to such an event as a 'fusion' (of colonies with the main population). Instead of keeping all those $(1+\varepsilon_N)N$ individuals in the main population alive, only N of them survive this fusion event, which are chosen uniformly at random among the $(1+\varepsilon_N)N$ previously existing individuals.*

(d) *Between the jump times of $\xi_N(t)$, the large population and the colonies independently evolve as Moran models, that is, at rate* 1 *each individual independently gives birth to a single offspring, and simultaneously a uniformly chosen individual is killed.*

2.2. Backward dynamics

Now we start with a sample of n lineages at time 0 and proceed backward in time. Let $\boldsymbol{X}_N(t) = (X_N^0(t), X_N^1(t), \ldots, X_N^n(t))$ be the ancestral process of this sample defined for $t \geq 0$ by

$X_N^0(t) =$ the number of lineages in the main population at time $-t$,

$X_N^i(t) =$ the number of colonies containing i lineages at time $-t$ $(1 \leq i \leq n)$.

We set $\boldsymbol{X}_N(0) = \boldsymbol{x}$, where $\boldsymbol{x} = (x_0, x_1, \ldots, x_n) \in \mathbb{N}^{n+1}$ with $x_0 + \sum_{j=1}^n jx_j = n$. The process $\{\boldsymbol{X}_N(t) : t \geq 0\}$ has state-space

$$E := \left\{(x_0, x_1, \ldots, x_n) \in \mathbb{N}^{n+1} : 1 \leq x_0 + \sum_{j=1}^{n} jx_j \leq n\right\}.$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Define the subspace Π of E by

$$\Pi := \{(x_0, x_1, 0, \dots, 0) \in \mathbb{N}^{n+1} \colon 1 \le x_0 + x_1 \le n\}.$$

Consider the projection $\Gamma \colon (x_0, x_1, 0, \dots, 0) \mapsto (x_0, x_1)$ from Π to $\mathbb{N}^2$:

$$\Gamma(\Pi) = \{(x_0, x_1) \in \mathbb{N}^2 \colon 1 \le x_0 + x_1 \le n\}.$$

By the action of the homeomorphism Γ, $\Gamma(\Pi)$ can be regarded as a subspace of E, and we thus still denote it by Π for simplicity. For $\boldsymbol{x} \in E$, let

$$\bar{\boldsymbol{x}} := \left(x_0, \sum_{j=1}^{n} x_j\right).$$

We will use this notation for the following reason: in the new time-scale lineages lying in the same colony will immediately coalesce, the configuration $\boldsymbol{x}$ immediately turns into $(x_0, \sum_{j=1}^{n} x_j, 0, \dots, 0)$ where all outer lineages are now alone in their respective colonies. Note that $\boldsymbol{x} \mapsto \bar{\boldsymbol{x}}$ is an injection from E to Π. We also write $\boldsymbol{e}_j = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{N}^{n+1}$ whose $(j+1)$th component is 1 for $j = 0, \dots, n$.

Let $\eta_N(t) = \xi_N(-t)$ for $t \ge 0$. It follows from [10, Lemma 1.5, p. 9] that $\{\eta_N(t) \colon t \ge 0\}$ is still a stationary Markov process with the same transition rates as (2.1). Thus, conditioned on $\eta_N(t)$, the fission events (fusions seen backward in time) happen at rate $\gamma_N \eta_N^{\alpha}(t)$ and the fusion events (fissions seen backward in time) happen at rate $N(\beta_N \varepsilon_N \eta_N + \theta_N)$. At any fission event seen backward, every lineage independently exits from the main population with probability $\varepsilon_N/(1+\varepsilon_N)$. At any fusion event seen backward, one colony is chosen at random from the existing colonies and the (say) i lineages in this colony enter one population (the mainland or another colony), and simultaneously coalesce together (if $i \ge 2$), and coalesce with their ancestor in this population (if it is also in the sample; but asymptotically, with high probability $i = 1$, the ancestor is not in the sample and the lineages enter the mainland). Between fission and fusion times, coalescences within the main population or within colonies may happen. We again refer to Figure 1 for an illustration.

Based on the above description, it is not hard to see that $\{(\boldsymbol{X}_N(t), \eta_N(t)) \colon t \ge 0\}$ is a time-homogeneous Markov chain taking values in $E \times \mathbb{R}_+$. The corresponding generator is given by

$$\bar{A}_N g(\boldsymbol{x}, k) = \bar{\psi}_N g(\boldsymbol{x}, k) + \bar{\phi}_N g(\boldsymbol{x}, k) + \bar{\Gamma}_N g(\boldsymbol{x}, k) \tag{2.4}$$

for any bounded function g on $E \times \mathbb{N}$. Here,

$$\bar{\psi}_N g(\boldsymbol{x}, k) = \sum_{j=2}^{n} x_j \binom{j}{2} \frac{2}{\varepsilon_N N - 1} (g(\boldsymbol{x} - \boldsymbol{e}_j + \boldsymbol{e}_{j-1}, k) - g(\boldsymbol{x}, k)),$$

which corresponds to coalescence of lineages in each colony. Since each colony evolves independently as a Moran model with size $\varepsilon_N N$ until fusing with the mainland, the coalescent rate for any two given lineages in the same colony is of the order $1/(\varepsilon_N N)$. Note that the

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

generator $\psi_N g(\boldsymbol{x}, u) \equiv 0$ if $\boldsymbol{x} \in \Pi$. In this case no two lineages lie in the same colony. Then

$$
\begin{aligned}
&\bar{\phi}_N g(\boldsymbol{x}, k) \\
&\quad = \binom{x_0}{2} \frac{2}{N-1} (g(\boldsymbol{x} - \boldsymbol{e}_1, k) - g(\boldsymbol{x}, k)) \\
&\qquad + N(\beta_N k \varepsilon_N + \theta_N) \sum_{r=1}^{x_0} \binom{x_0}{r} \left(\frac{\varepsilon_N}{1+\varepsilon_N}\right)^r \left(\frac{1}{1+\varepsilon_N}\right)^{x_0 - r} (g(\boldsymbol{x} - r\boldsymbol{e}_0 + \boldsymbol{e}_r, k+1) \\
&\qquad\qquad - g(\boldsymbol{x}, k)) \\
&\qquad + \gamma_N k^\alpha \frac{N}{k\varepsilon_N N + N} \left(\frac{x_1}{k}\right)\left(1 - \frac{x_0}{N}\right) (g(\boldsymbol{x} - \boldsymbol{e}_1 + \boldsymbol{e}_0, k-1) - g(\boldsymbol{x}, k)) \mathbf{1}_{\{k>0\}} \\
&\qquad + \gamma_N k^\alpha \frac{N}{k\varepsilon_N N + N} \sum_{j=2}^{n} \left(\frac{x_j}{k}\right)\left(1 - \frac{x_0}{N}\right) (g(\boldsymbol{x} - \boldsymbol{e}_j + \boldsymbol{e}_0, k-1) - g(\boldsymbol{x}, k)) \mathbf{1}_{\{k>0\}} \\
&\qquad + \gamma_N k^\alpha \frac{N}{k\varepsilon_N N + N} \sum_{j=1}^{n} \left(\frac{x_j}{k}\right)\left(\frac{x_0}{N}\right) (g(\boldsymbol{x} - \boldsymbol{e}_j, k-1) - g(\boldsymbol{x}, k)) \mathbf{1}_{\{k>0\}} \\
&\qquad + \gamma_N k^\alpha \frac{k\varepsilon_N N}{k\varepsilon_N N + N} \sum_{i \neq j, i, j \geq 1}^{n} \frac{(x_i/k)(x_j/k) j}{\varepsilon_N N} (g(\boldsymbol{x} - \boldsymbol{e}_i + \boldsymbol{e}_j, k-1) - g(\boldsymbol{x}, k)) \mathbf{1}_{\{k>0\}} \\
&\qquad + \gamma_N k^\alpha \frac{k\varepsilon_N N}{k\varepsilon_N N + N} \sum_{i \neq j, i, j \geq 1}^{n} \left(\frac{x_i}{k}\right)\left(\frac{x_j}{k}\right)\left(\frac{1-j}{\varepsilon_N N}\right) \\
&\qquad\qquad \times (g(\boldsymbol{x} - \boldsymbol{e}_i - \boldsymbol{e}_j + \boldsymbol{e}_{j+1}, k-1) - g(\boldsymbol{x}, k)) \mathbf{1}_{\{k>0\}},
\end{aligned}
$$

where the first term corresponds to coalescence of lineages within the mainland, the second term corresponds to migration of lineages from the mainland to colonies, the third term corresponds to migration of lineages from each colony containing one lineage, and the fourth term from each colony containing more than one lineage, which is identically equal to 0 if $\boldsymbol{x} \in \Pi$. We denote by $\mathbf{1}_{\{\cdot\}}$ the indicator function. The fifth term corresponds to coalescence of lineages between the mainland and colonies. The last two terms correspond to coalescence or migration of lineages among colonies. Finally,

$$
\begin{aligned}
\bar{\Gamma}_N g(\boldsymbol{x}, k) &= N(\beta_N \varepsilon_N k + \theta_N) \left(\frac{1}{1+\varepsilon_N}\right)^{x_0} (g(\boldsymbol{x}, k+1) - g(\boldsymbol{x}, k)) \\
&\quad + \gamma_N k^\alpha \left(1 - \sum_{j=1}^{n} \frac{x_j}{k}\right) (g(\boldsymbol{x}, k-1) - g(\boldsymbol{x}, k)) \mathbf{1}_{\{k>0\}},
\end{aligned}
$$

which corresponds to the event that the number of colonies increases or decreases but the ancestral process does not change.

3. Convergence to the two-state censored coalescent

3.1. Main results

Let $D([0, \infty), S)$ be the space of all càdlàg functions $x\colon [0, \infty) \to S$ endowed with the Skorokhod topology for any separable and complete metric space S; see Ethier and Kurtz [7, p. 116] for details. For $N \in \mathbb{N}$, we consider the sequence of processes $\{(\boldsymbol{X}_N(\cdot), \eta_N(\cdot))\}$. Define

$$
\boldsymbol{Y}_N(t) = \boldsymbol{X}_N(Nt) \quad \text{and} \quad \tilde{\eta}_N(t) = \varepsilon_N \eta_N(Nt).
$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Let $\beta > 0$, $\theta > 0$, and $\gamma > 0$ be constants satisfying $\beta < \gamma$ if $\alpha = 1$. We further assume the following conditions.

Condition 3.1. *It holds that $\varepsilon = \varepsilon_N$ satisfying $\varepsilon_N \to 0$ and $\varepsilon_N N \to \infty$ as $N \to \infty$;*

Condition 3.2. *It holds that $\beta_N = \beta/(\varepsilon_N N^2)$, $\theta_N = \theta/(\varepsilon_N N^2)$, and $\gamma_N = \gamma \varepsilon_N^{\alpha-1}/N$.*

Recall that $\boldsymbol{y} \in E$ and the corresponding $\bar{\boldsymbol{y}} \in \Pi$. The main result of the paper follows.

Theorem 3.1. *Under Conditions 3.1 and 3.2, the finite-dimensional distributions of the ancestral process $\{\boldsymbol{Y}_N(t), t \geq 0\}$ starting at $\boldsymbol{y}$ converge to those of a Π-valued continuous time Markov chain $\{\boldsymbol{Y}(t), t \geq 0\}$ starting at $\bar{\boldsymbol{y}}$, except at time 0. The corresponding infinitesimal generator $\boldsymbol{Q} = (q_{\boldsymbol{r},\boldsymbol{r}'})_{\boldsymbol{r},\boldsymbol{r}' \in \Pi}$ is given by*

$$q_{\boldsymbol{r},\boldsymbol{r}'} = \begin{cases} -((\beta\kappa + \theta)r_0 + \dfrac{\gamma\kappa^{\alpha-1}}{\kappa+1} r_1 + r_0(r_0 - 1)) & \text{if } \boldsymbol{r}' = \boldsymbol{r}, \\ \gamma\kappa^{\alpha} r_0 & \text{if } r_0 \neq 0 \text{ and } \boldsymbol{r}' = \boldsymbol{r} + (-1, 1), \\ \dfrac{\beta\kappa + \theta}{\kappa(\kappa+1)} r_1 & \text{if } r_1 \neq 0 \text{ and } \boldsymbol{r}' = \boldsymbol{r} + (1, -1), \\ r_0(r_0 - 1) & \text{if } \boldsymbol{r}' = \boldsymbol{r} + (-1, 0), \\ 0 & \text{otherwise}, \end{cases} \tag{3.1}$$

where $\boldsymbol{r} = (r_0, r_1) \in \Pi$ and κ is the unique solution of the equation $\beta z + \theta = \gamma z^{\alpha}$. Furthermore, if the initial value $\boldsymbol{y} \in \Pi$, weak convergence on $D([0, \infty), \Pi)$ to $\{\boldsymbol{Y}(t)\}$ holds.

Remark 3.1. For $\boldsymbol{y} \in E$, let $\boldsymbol{Y}(\cdot)$ be the above process (3.1) with the initial value $\boldsymbol{Y}(0) = \bar{\boldsymbol{y}}$. Define

$$\boldsymbol{Y}^*(t) = \begin{cases} \boldsymbol{Y}(t) & \text{if } t > 0, \\ \boldsymbol{y} & \text{if } t = 0. \end{cases} \tag{3.2}$$

Then $\boldsymbol{Y}^*(\cdot)$ is an E-valued continuous time Markov chain. The proof of Theorem 3.1 actually shows that the finite-dimensional distributions of $\{\boldsymbol{Y}_N(t), t \geq 0\}$ converge to those of $\{\boldsymbol{Y}^*(t), t \geq 0\}$. For the limiting process $\boldsymbol{Y}^*(\cdot)$, we have $\boldsymbol{Y}^*(0) = \boldsymbol{y}$ while $\boldsymbol{Y}^*(0+) = \bar{\boldsymbol{y}}$, which implies that lineages lying in the same colony immediately coalesce.

The previous statement describes the asymptotic genealogical history of a finite sample of lineages, seen as a process backward in time. Except at time 0, where instantaneous coalescences within colonies makes the limiting process not right-continuous, this process converges weakly to a two-state censored coalescent, where type 0 corresponds to inner lineages (lying in the main population) and type 1 to outer lineages (lying in pairwise distinct colonies). Inner lineages coalesce at constant rate 1 per (ordered) pair, and lineages change type at constant rate per lineage: inner lineages become outer lineages at rate $\gamma\kappa^{\alpha}$ and outer lineages become inner lineages at rate $(\beta\kappa + \theta)/\kappa(\kappa + 1)$.

Nordborg and Krone [14] studied the behavior of a geographically structured population with strong migration; see also Notohara [16]. If migration is strong to some extent, the total population behaves like a panmictic population with an effective population size depending on the population structure and the strength of migration. In other words, if migration occurs on a time-scale that is much faster than the coalescent time-scale then the limiting ancestral process viewed on the coalescent time-scale is the Kingman coalescent with a so-called 'effective'

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

coalescence rate. Inspired by this idea, we consider a sequence of censored coalescent processes $\{\boldsymbol{Y}_k(t)\}$ defined by

$$q_{\boldsymbol{r},\boldsymbol{r}'} = \begin{cases} -(a_k r_0 + b_k r_1 + r_0(r_0-1)) & \text{if } \boldsymbol{r}' = \boldsymbol{r}, \\ a_k r_0 & \text{if } r_0 \neq 0 \text{ and } \boldsymbol{r}' = \boldsymbol{r} + (-1, 1), \\ b_k r_1 & \text{if } r_1 \neq 0 \text{ and } \boldsymbol{r}' = \boldsymbol{r} + (1, -1), \\ r_0(r_0-1) & \text{if } \boldsymbol{r}' = \boldsymbol{r} + (-1, 0), \\ 0 & \text{otherwise,} \end{cases}$$

where the parameters a_k and b_k satisfy the following condition.

Condition 3.3. *As $k \to \infty$, $a_k \to \infty$, $b_k \to \infty$ and $b_k/a_k \to p$ for some constant $p > 0$.*

The above condition corresponds to the acceleration of the landscape dynamics (strong migration between mainland and colonies). Furthermore, we suppose that the initial value $\boldsymbol{Y}_k(0) = \boldsymbol{y} \in \Pi$ with $y_0 + y_1 = n$. Let $Y_k(t) = Y_k^0(t) + Y_k^1(t)$ and let $I_n = \{0, 1, 2, \dots, n\}$. The following theorem states that such a strong migration limit gives rise to a single state coalescent process, where coalescence rates are obtained by averaging over the probability of presence in the main population.

Theorem 3.2. *Under Condition 3.3, the process $\{Y_k(t), t \geq 0\}$ starting at n converges weakly to the time-changed n-Kingman coalescent process $\{K(t), t \geq 0\}$ on $D([0, \infty), I_n)$. When $n = l$, the coalescence rate is given by*

$$c_l = \sum_{j=1}^{l} j(j-1)\binom{l}{j}\Big(\frac{p}{1+p}\Big)^j\Big(\frac{1}{1+p}\Big)^{l-j}.$$

Remark 3.2. Condition 3.3 has the intuitive biological meaning that migrations occur at a much faster rate than coalescences.

Remark 3.3. Note that $b_k/a_k \to p$ as $k \to \infty$. It is easy to see that if $p = 0$, which corresponds to predominant emigrations, $\{Y_k(t), t \geq 0\}$ converges weakly to the constant process $\{K(t) \equiv n, t \geq 0\}$; if $p = \infty$ which corresponds to predominant fusions, $\{Y_k(t), t \geq 0\}$ converges weakly to the standard Kingman coalescent $\{K(t), t \geq 0\}$ (i.e. $c_l = l(l-1)$).

3.2. Further discussions

Density dependence. We can allow for more general metapopulation dynamics by changing the dynamics to include both types of density dependence (in birth rate or in death rate). To do so, we would have to parametrize the fission rate by a nonnegative function $b_N(x)$ on $\mathbb{R}_+$ and the fusion rate by a nonnegative function $x d_N(x)$ on $\mathbb{R}_+$, in such a way that we have convergence of the number of colonies, denoted by $\xi_N(t)$, after scaling, to deterministic dynamics with a single stable equilibrium. More precisely,

- at constant rate ρ_N, the mainland sends $\varepsilon_N N$ offspring to found a new colony;
- at rate $b_N(\xi_N)$, each colony independently sends $\varepsilon_N N$ offspring to found a new colony;
- each colony independently merges again with the mainland at rate $d_N(\xi_N)$.

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Then the number of colonies, still denoted by $\xi_N(t)$, is a birth–death density dependent process with immigration and the transition rate is given by

$$j \to \begin{cases} j+1 & \text{at rate } jb_N(j) + \rho_N, \\ j-1 & \text{at rate } jd_N(j). \end{cases}$$

We assume the following condition.

Condition 3.4. (a) *For each N, $xb_N(x)$ is nondecreasing and concave, and $xd_N(x)$ is nondecreasing and convex.*

(b) *There exist functions $b(\cdot)$, $d(\cdot) \in C^1(\mathbb{R}_+)$, and some constant $\rho > 0$ such that*

$$\rho_N N\varepsilon_N \to \rho, \qquad Nb_N\left(\frac{x}{\varepsilon_N}\right) \to b(x), \qquad Nd_N\left(\frac{x}{\varepsilon_N}\right) \to d(x)$$

as $N \to \infty$ for all $x \in \mathbb{R}_+$.

(c) *It holds that* $\limsup_{x\to\infty}(b(x) - d(x)) < 0$.

Then, in a similar manner to the proof of Lemma 3.1, we obtain

$$\varepsilon_N \xi_N(Nt) \xrightarrow{\mathbb{P}} \kappa$$

in $D([0,\infty), \mathbb{R}_+)$, where κ is the unique positive solution of $xb(x) + \rho = xd(x)$ and '$\xrightarrow{\mathbb{P}}$' denotes convergence in probability. Thus, by virtually the same proof as in Section 3.2, we still have that the ancestral process converges weakly to a two-state censored coalescent defined by (3.1), where inner lineages become outer lineages at rate $d(\kappa)$ and outer lineages become inner lineages at rate $(\kappa b(\kappa) + \rho)/\kappa(\kappa + 1)$.

3.2.1. *Extinction of colonies.* Conditioned on $\xi_N(t)$, the number of colonies at time t, the lifetime of a colony is exponential with parameter $\gamma_N \xi_N^{\alpha-1}$. Let us suppose that at the end of lifetime of the colony, it will merge with the mainland with probability p or be extinct with probability $1-p$. Then it is not hard to see that ξ_N is still asymptotically deterministic, equal to $\varepsilon_N^{-1}\kappa$, and we obtain the same result as Theorem 3.1 except for the inner to outer rate, where γ becomes γp.

3.2.2. *Random size for each colony.* We might allow for colonies with random size (either constant but random, or ergodic), provided we can guarantee the total population size is of the order of N, and provided we are able to estimate the probability that, looking back in time, at a given fusion event, a given inner lineage originates from the merging colony.

3.2.3. *Reproduction mechanism.* We could have assumed that all individuals (in the mainland) give birth to $N\varepsilon_N$ offspring regardless of whether there is a founding event or not. More precisely, in addition to founding new colonies, each individual independently at rate 1 gives birth to $\varepsilon_N N$ offsprings, and simultaneously $\varepsilon_N N$ uniformly chosen individuals are killed. In this case, one would still obtain a two-state censored coalescent as defined in Theorem 3.1.

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

3.3. Proofs

To prove Theorem 3.1, we start by proving the following lemmas.

Lemma 3.1. *Under Conditions 3.4(a) and 3.4(b), as $N \to \infty$,*

$$\tilde{\eta}_N(\cdot) \xrightarrow{\mathbb{P}} \kappa$$

in $D([0,\infty),\mathbb{R}_+)$, where κ is the unique solution of $\beta z + \theta = \gamma z^\alpha$.

Proof. Recall that the number $\xi_N(Nt)$ of colonies of size $\varepsilon_N N$ is a birth–death density dependent process with immigration with transition rates given by

$$\begin{aligned} j &\to j+1 \quad \text{at rate} \quad N^2\theta_N + \varepsilon_N N^2 \beta_N j, \\ j &\to j-1 \quad \text{at rate} \quad N\gamma_N j^\alpha. \end{aligned}$$

It has the stationary distribution π_N given by (2.2). Note that

$$\frac{N^2\theta_N + \varepsilon_N N^2\beta_N i}{N\gamma_N i^\alpha} = \frac{1}{\gamma}\left(\frac{\theta}{\varepsilon_N^\alpha i^\alpha} + \frac{\beta}{\varepsilon_N^{\alpha-1} i^{\alpha-1}}\right),$$

which is a decreasing function of i. It is easy to find some constant $0 < \rho < 1$ and $c > 0$ such that for any $i > M_N$, where $M_N := [c/\varepsilon_N]$ and $[x]$ denotes the integer part of real number x,

$$\frac{N^2\theta_N + \varepsilon_N N^2\beta_N i}{N\gamma_N i^\alpha} < \rho < 1.$$

Then, when $j > M_N$,

$$\begin{aligned} \pi_N(k) &= w_{N,M_N} \frac{N\theta_N + \varepsilon_N N\beta_N M_N}{\gamma_N j^\alpha} \prod_{i=M_N+1}^{j-1} \frac{N\theta_N + \varepsilon_N N\beta_N i}{\gamma_N i^\alpha} \left[1 + \sum_{j=1}^{\infty} w_{N,j}\right]^{-1} \\ &\le \frac{N\theta_N + \varepsilon_N N\beta_N M_N}{\gamma_N j^\alpha} \prod_{i=M_N+1}^{j-1} \frac{N\theta_N + \varepsilon_N N\beta_N i}{\gamma_N i^\alpha} \\ &\le \rho^{j-M_N}. \end{aligned}$$

Thus, we have

$$\sum_{j=2M_N}^{\infty} \pi_N(j) \le \sum_{j=2M_N}^{\infty} \rho^{j-M_N} = \frac{\rho^{M_N}}{1-\rho},$$

which implies that $\pi_N([2M_N,\infty)) = O(\rho^{M_N})$. Suppose that $\xi_N(0)$, the initial value of $\xi_N(\cdot)$, is distributed as π_N. Then the sequence $\{\varepsilon_N \xi_N(0)\}$ is tight. On the other hand, $\{\varepsilon_N \xi_N(Nt)\}$ takes values in $\{i\varepsilon_N : i \in \mathbb{N}\}$ and its generator is given by

$$L_N f(z) = N^2\left(\theta_N + \frac{\varepsilon_N \beta_N z}{\varepsilon_N}\right)(f(z+\varepsilon_N) - f(z)) + N\gamma_N \left(\frac{z}{\varepsilon_N}\right)^\alpha (f(z-\varepsilon_N) - f(z))$$

for any continuous bounded function f on $\mathbb{R}_+$. Let $C_c^2(\mathbb{R}_+)$ be the set of twice differentiable functions with compact support on $\mathbb{R}_+$. It is not hard to see that as $N \to \infty$ for $f \in C_c^2(\mathbb{R}_+)$,

$$\|L_N f - Lf\| \to 0 \quad \text{and} \quad Lf(z) = (\beta z + \theta - \gamma z^\alpha) f'(z), \tag{3.3}$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

where $\|f\| = \sup_{x\in\mathbb{R}_+} |f(x)|$. The Markov process ξ with generator L is actually deterministic and satisfies the ordinary differential equation:

$$\xi'(t) = \beta\xi(t) + \theta - \gamma\xi^{\alpha}(t).$$

Note that there exists a unique positive solution κ such that $\beta\kappa + \theta - \gamma\kappa^{\alpha} = 0$. Furthermore, it follows that $(\beta x + \theta - \gamma x^{\alpha})'|_{x=\kappa} < 0$, and for any $\delta > 0$, $\inf_{\{|z-\kappa|\geq\delta\}\cap\mathbb{R}_+} |\beta z + \theta - \gamma z^{\alpha}| > 0$. Thus, the above ordinary differential equation has the unique equilibrium point κ. It follows from (3.3), [7, Theorem 6.1, p. 28], and [7, Theorem 9.10, p. 244] that $\varepsilon_N\xi_N(0) \xrightarrow{\mathrm{w}} \kappa$ as $N \to \infty$, where '$\xrightarrow{\mathrm{w}}$' denotes weak convergence. Again by (3.3), [7, Corollary 8.7, p. 231] shows that $\{\xi_N(t) : t \geq 0\} \xrightarrow{\mathrm{w}} \{\xi(t) \equiv \kappa, t \geq 0\}$ on $D([0,\infty), \mathbb{R}_+)$. The lemma is proved.

As in Section 2 it is easy to see that $(\boldsymbol{Y}_N(\cdot), \tilde{\eta}_N(\cdot))$ is a continuous-time Markov chain taking values in $E \times \mathbb{R}_+$. Based on (2.4) and Conditions 3.1 and 3.2, a simple calculation shows that the corresponding generator is given by

$$A_N g(\boldsymbol{y}, u) = \psi_N g(\boldsymbol{y}, u) + \phi_N g(\boldsymbol{y}, u) + \Gamma_N g(\boldsymbol{y}, u) \tag{3.4}$$

for any bounded function g on $E \times \mathbb{R}_+$. Here,

$$\psi_N g(\boldsymbol{y}, u) = 2\sum_{j=2}^{n} y_j \binom{j}{2} \frac{1}{\varepsilon_N}\left(1 + \frac{1}{\varepsilon_N N - 1}\right)(g(\boldsymbol{y} - \boldsymbol{e}_j + \boldsymbol{e}_{j-1}, u) - g(\boldsymbol{y}, u)).$$

Note that $1/(\varepsilon_N N - 1) \to 0$ as $N \to \infty$ by Condition 3.1. We also have

$$\begin{aligned}
\phi_N g(\boldsymbol{y}, u) = {} & 2\binom{y_0}{2}(g(\boldsymbol{y} - \boldsymbol{e}_0, u) - g(\boldsymbol{y}, u)) \\
& + (\beta u + \theta) y_0 (g(\boldsymbol{y} - \boldsymbol{e}_0 + \boldsymbol{e}_1, u + \varepsilon_N) - g(\boldsymbol{y}, u)) \\
& + \frac{\gamma u^{\alpha-1}}{u+1} y_1 (g(\boldsymbol{y} - \boldsymbol{e}_1 + \boldsymbol{e}_0, u - \varepsilon_N) - g(\boldsymbol{y}, u))\, \mathbf{1}_{\{u>0\}} \\
& + \frac{\gamma u^{\alpha-1}}{u+1} \sum_{j=2}^{n} y_j (g(\boldsymbol{y} - \boldsymbol{e}_j + \boldsymbol{e}_0, u - \varepsilon_N) - g(\boldsymbol{y}, u))\, \mathbf{1}_{\{u>0\}} \\
& + \left(\varepsilon_N R_{1,N} g(\boldsymbol{y}, u) + \left(\varepsilon_N + \frac{1}{N}\right)\frac{u^{\alpha-1}}{u+1}\, \mathbf{1}_{\{u>0\}}\, R_{2,N} g(\boldsymbol{y}, u)\right).
\end{aligned}$$

Here the fourth term is identically equal to 0 if $\boldsymbol{y} \in \Pi$. In the last term, $R_{1,N}$ and $R_{2,N}$ are bounded linear operators satisfying $\|R_{i,N}\| \leq C$ for some constant C. The last term vanishes as $N \to \infty$ if $c_1 \leq u \leq c_2$ for positive numbers c_1 and c_2. Last, we have

$$\begin{aligned}
\Gamma_N g(\boldsymbol{y}, u) = {} & (\beta u + \theta)\varepsilon_N^{-1}(1 - y_0\varepsilon_N)(g(\boldsymbol{y}, u + \varepsilon_N) - g(\boldsymbol{y}, u)) \\
& + \gamma u^{\alpha}\varepsilon_N^{-1}(1 - \varepsilon_N u^{-1}\sum_{j=1}^{n} y_j)(g(\boldsymbol{y}, u - \varepsilon_N) - g(\boldsymbol{y}, u))\, \mathbf{1}_{\{u>0\}}.
\end{aligned}$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Let us write $c_\psi^N(\boldsymbol{y})$ (respectively $c_\phi^N(\boldsymbol{y},u)$, $c_\Gamma^N(\boldsymbol{y},u)$) the total rate of the events generated by ψ_N (respectively ϕ_N, Γ_N) when $A_N g$ is applied to $(\boldsymbol{y},u)$. Then

$$c_\psi^N(\boldsymbol{y}) = 2\sum_{j=2}^{n} y_j \binom{j}{2} \frac{1}{\varepsilon_N}\left(1+\frac{1}{\varepsilon_N N-1}\right),$$

$$c_\phi^N(\boldsymbol{y},u) = 2\binom{y_0}{2} + (\beta u+\theta)y_0 + \frac{\gamma u^{\alpha-1}}{u+1}\mathbf{1}_{\{u>0\}}\sum_{j=1}^{n} y_j + \varepsilon_N\left(1+\frac{u^{\alpha-1}}{u+1}\mathbf{1}_{\{u>0\}}\right),$$

and

$$c_\Gamma^N(\boldsymbol{y},u) = (\beta u+\theta)\varepsilon_N^{-1}(1-y_0\varepsilon_N) + \gamma u^\alpha \varepsilon_N^{-1}(1-\varepsilon_N u^{-1}\sum_{j=1}^{n} y_j)\,\mathbf{1}_{\{u>0\}}.$$

Let us introduce the following notation,

$$\sigma_0^N = \inf\{t\ge 0\colon \boldsymbol{Y}_N(t)\in\Pi\}$$

and

$$\sigma_1^N = \inf\{t\ge 0\colon \text{a } \phi_N\text{-event occurs at } t\}.$$

Lemma 3.2. *It holds that* $\sigma_0^N \xrightarrow{\mathbb{P}} 0$ *as* $N\to\infty$.

Proof. By Lemma 3.1, we have for any T and $0<\delta<\kappa$,

$$\mathbb{P}\Big\{\sup_{0\le t\le T}|\tilde{\eta}_N(t)-\kappa|>\delta\Big\}\to 0 \quad \text{as } N\to\infty, \tag{3.5}$$

where $\mathbb{P}$ is the probability measure. Fix δ above. Let $c_1=\kappa-\delta$ and $c_2=\kappa+\delta$. Conditioned on $(\boldsymbol{Y}_N(t),\tilde{\eta}_N(t))=(\boldsymbol{y},u)$ with $(\boldsymbol{y},u)\in(E\setminus\Pi)\times[c_1,c_2]$ at the current time t,

$$\mathbb{P}\{\text{the next event is a } \phi_N\text{-event}\} = \frac{c_\phi^N(\boldsymbol{y},u)}{c_\psi^N(\boldsymbol{y})+c_\phi^N(\boldsymbol{y},u)+c_\Gamma^N(\boldsymbol{y},u)} \le C\varepsilon_N$$

for some positive constant C,

$$\mathbb{P}\{\text{the next event is a } \psi_N\text{-event}\} = \frac{c_\psi^N(\boldsymbol{y})}{c_\psi^N(\boldsymbol{y})+c_\phi^N(\boldsymbol{y},u)+c_\Gamma^N(\boldsymbol{y},u)} \le \frac{2n^3}{2n^3+\beta c_1+\gamma c_1^\alpha}$$

for sufficiently large N,

$$\mathbb{P}\{\text{the next event is a } \Gamma_N\text{-event}\} = \frac{c_\Gamma^N(\boldsymbol{y},u)}{c_\psi^N(\boldsymbol{y})+c_\phi^N(\boldsymbol{y},u)+c_\Gamma^N(\boldsymbol{y},y)} \le \frac{(\beta c_2+\theta)+\gamma c_2^\alpha}{2+(\beta c_2+\theta)+\gamma c_2^\alpha}$$

for sufficiently large N. Inspired by Taylor and Véber [18, Lemma 3.1], we fix some $s>0$ and consider

$$\mathbb{P}\{\sigma_0^N>s\} = \mathbb{P}(D)+o(1),$$

where

$$D = \Big\{\sigma_0^N>s,\ \sup_{0\le t\le s}|\tilde{\eta}_N(t)-\kappa|\le\delta\Big\}.$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Then

$$\begin{aligned}\mathbb{P}(D) &= \mathbb{P}\{\{\text{at most } n\psi_N\text{-events occur in } [0,s]\} \cap D\}\\ &= \mathbb{P}\{\{\text{at most } n\psi_N\text{- and at least a } \phi_N\text{-events occur in } [0,s]\} \cap D\}\\ &\quad + \mathbb{P}\{\{\text{at most } n\psi_N\text{- and no } \phi_N\text{-events occur in } [0,s]\} \cap D\}\\ &=: I_1 + I_2.\end{aligned}$$

Note that, we have $Y^N(t) \in E \setminus \Pi$ for $t \in [0,s]$ if $\sigma_0^N > s$. Let

$$p = \frac{2n^3}{2n^3 + \beta c_1 + \gamma c_1^\alpha} \vee \frac{(\beta c_2 + \theta) + \gamma c_2^\alpha}{2 + (\beta c_2 + \theta) + \gamma c_2^\alpha}.$$

Then

$$\begin{aligned}I_1 &\le \sum_{k=0}^{n} \mathbb{P}\{\{\text{exactly } k\psi_N\text{-events before a } \phi_N\text{-event occurs in } [0,s]\} \cap D\}\\ &= \sum_{k=0}^{n}\sum_{l=0}^{\infty} \mathbb{P}\{\{\text{exactly } k\psi_N\text{- and } l\Gamma_N\text{-events before a } \phi_N\text{-event in } [0,s]\} \cap D\}\\ &\le \sum_{k=0}^{n}\sum_{l=0}^{\infty} \binom{k+l}{k} p^{k+l}(C\varepsilon_N), \qquad (3.6)\end{aligned}$$

Since $0 < p < 1$, $\sum_{k=0}^{n}\sum_{l=0}^{\infty}\binom{k+l}{k}p^{k+l} < \infty$. Then $I_1 \to 0$ as $N \to \infty$. Let U_j^N be the arrival time of the jth event occurring to $(\boldsymbol{Y}_N, \tilde{\eta}_N)$. For I_2,

$$\begin{aligned}I_2 &= \sum_{k=0}^{n}\sum_{l=0}^{\infty} \mathbb{P}\{\{\text{exactly } k\psi_N\text{-events, } l\Gamma_N\text{-events and no } \phi_N\text{-events occur in } [0,s]\} \cap D\}\\ &\le \sum_{k=0}^{n}\sum_{l=0}^{\infty} \binom{k+l}{k} p^{k+l}\mathbb{P}\{\{U_{k+l}^N < s, U_{k+l+1}^N > s\} \cap D\}.\end{aligned}$$

Conditioned on $(\boldsymbol{Y}_N(t), \tilde{\eta}_N(t)) = (\boldsymbol{y}, u)$ with $\boldsymbol{y} \in (E \setminus \Pi)$, the rate for the event occurring to $(\boldsymbol{Y}_N, \tilde{\eta}_N)$ at time t is $c_\psi^N(\boldsymbol{y}) + c_\phi^N(\boldsymbol{y}, u) + c_\Gamma^N(\boldsymbol{y}, u)$ and $c_\psi^N(\boldsymbol{y}) \ge 2/\varepsilon_N$. Then U_{k+l+1}^N is stochastically bounded by the sum of $k+l+1$ independent and identically distributed (i.i.d.) exponential variables with parameter $2/\varepsilon_N$ whose distribution becomes concentrated close to 0 as $N \to \infty$. Thus,

$$\mathbb{P}\{\{U_{k+l}^N < s, U_{k+l+1}^N > s\} \cap D\} \to 0 \quad \text{as } N \to \infty,$$

and by the dominated convergence theorem, $I_2 \to 0$.

Lemma 3.3. *There exist positive constants M and K_1 such that for any $s > 0$,*

$$\limsup_{N\to\infty} \mathbb{P}\{\sigma_1^N \le s\} \le M(1 - \mathrm{e}^{-K_1 s}).$$

Proof. By the proof of (3.6), $\mathbb{P}\{\text{at least one } \phi_N\text{-event occurs before } \sigma_0^N\} \to 0$ as $N \to \infty$. Then by (3.5), we have

$$\mathbb{P}\{\sigma_1^N \le s\} = \mathbb{P}(G) + o(1),$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

where

$$G = \Big\{\text{only } \psi_N\text{- or } \Gamma_N\text{- events before } \sigma_0^N,\ \sup_{0\le t\le s} |\tilde{\eta}_N(t) - \kappa| \le \delta \text{ and } \sigma_1^N \le s\Big\}.$$

Recall that $\boldsymbol{Y}_N(0) = \boldsymbol{y}$. If only ψ_N- or Γ_N- events occur before σ_0^N, $\sigma_0^N < \sigma_1^N$, and $\boldsymbol{Y}_N(t) = \bar{\boldsymbol{y}}$ for $t \in [\sigma_0^N, \sigma_1^N]$. Furthermore, $Y_N(t) \in \Pi$ and $c_\psi^N \equiv 0$ for $t \ge \sigma_0^N$. Conditioned on $(\boldsymbol{Y}_N(t), \tilde{\eta}_N(t)) = (\boldsymbol{y}, u)$ with $(\boldsymbol{y}, u) \in \Pi \times [c_1, c_2]$,

$$K_1 \le c_\phi^N(\boldsymbol{y}, u) \le K_2, \qquad \frac{\varepsilon_N^{-1}(\beta c_1 + \gamma c_1^\alpha)}{2} \le c_\Gamma^N(\boldsymbol{y}, u) \le \varepsilon_N^{-1}(\theta + \beta c_2 + \gamma c_2^\alpha)$$

for sufficiently large N, where $K_1 = [\gamma(c_1^{\alpha-1} \wedge c_2^{\alpha-1})] \wedge (\beta c_1)$ and $K_2 = n^2 + n(\beta c_2 + \theta) + \gamma n(c_1^{\alpha-1} \vee c_2^{\alpha-1})$. Then

$$\frac{c_\phi^N(\boldsymbol{y}, u)}{c_\phi^N(\boldsymbol{y}, u) + c_\Gamma^N(\boldsymbol{y}, u)} \le \frac{2K_2\varepsilon_N}{2K_2\varepsilon_N + \beta c_1 + \gamma c_1^\alpha},$$

$$\frac{c_\Gamma^N(\boldsymbol{y}, u)}{c_\phi^N(\boldsymbol{y}, u) + c_\Gamma^N(\boldsymbol{y}, u)} \le \frac{\theta + \beta c_2 + \gamma c_2^\alpha}{K_1\varepsilon_N + \theta + \beta c_2 + \gamma c_2^\alpha}.$$

For $(\boldsymbol{Y}_N(\cdot), \tilde{\eta}_N(\cdot))$ with initial value $(\boldsymbol{y}, u) \in \Pi \times [c_1, c_2]$, recall that U_j^N denotes the arrival time of the jth event occurring to $(\boldsymbol{Y}_N, \tilde{\eta}_N)$ and $U_0^N = 0$. It is not hard to see that U_j^N is stochastically larger than the sum of j i.i.d. exponential variables with parameter $\varepsilon_N^{-1}(\theta + \beta c_2 + \gamma c_2^\alpha) + K_2$. We have

$$\begin{aligned}
\mathbb{P}(G) &= \sum_{k=0}^{\infty} \mathbb{P}\{\{\text{exactly } k\Gamma_N\text{-events occur in } [\sigma_0^N, \sigma_1^N]\} \cap G\} \\
&\le \sum_{k=0}^{\infty} \left(\frac{\theta + \beta c_2 + \gamma c_2^\alpha}{K_1\varepsilon_N + \theta + \beta c_2 + \gamma c_2^\alpha}\right)^k \frac{2K_2\varepsilon_N}{2K_2\varepsilon_N + \beta c_1 + \gamma c_1^\alpha} \mathbb{P}\{\sigma_0^N + U_k^N \le s\} \\
&\le \sum_{k=0}^{\infty} \left(\frac{\theta + \beta c_2 + \gamma c_2^\alpha}{K_1\varepsilon_N + \theta + \beta c_2 + \gamma c_2^\alpha}\right)^k \frac{2K_2\varepsilon_N}{2K_2\varepsilon_N + \beta c_1 + \gamma c_1^\alpha} \mathbb{P}\Big\{\sigma_0^N + \sum_{j=1}^{k} \tilde{V}_j^N \le s\Big\} \\
&\le M\mathbb{P}\Big\{\sigma_0^N + \sum_{j=1}^{T_N} \tilde{V}_j^N \le s\Big\}
\end{aligned}$$

for some positive constant M and sufficiently large N, where $\{\tilde{V}_j^N\}$ are i.i.d. exponential variables with parameter $\varepsilon_N^{-1}(\theta+\beta c_2+\gamma c_2^\alpha)+K_2$, and T_N is a geometric variable with parameter $K_1\varepsilon_N/(K_1\varepsilon_N + (\beta c_2 + \theta) + \gamma c_2^\alpha)$ independent of $\{\tilde{V}_j^N\}$. Since $\sigma_0^N \xrightarrow{\mathbb{P}} 0$, a simple calculation shows that $\sigma_0^N + \sum_{j=0}^{M_N} V_j^N$ converges weakly to an exponential variable with parameter K_1. The lemma is proved.

Lemma 3.4. *Under Conditions 3.1 and 3.2, the ancestral process $\{\boldsymbol{Y}_N(t), t \ge 0\}$ starting at $\boldsymbol{y}$ with $\boldsymbol{y} \in \Pi$ converges weakly on $D([0, \infty), \Pi)$ to $\{\boldsymbol{Y}(t), t \ge 0\}$ given by (3.1) starting at $\boldsymbol{y}$.*

Proof. If the initial value $\boldsymbol{Y}_N(0) = \boldsymbol{y} \in \Pi$ then $(\boldsymbol{Y}_N(t), t \ge 0) \in D([0, \infty), \Pi)$ and, thus, ψ_N and the fourth term in ϕ_N vanishes. In this case for any bounded function g on $\Pi \times \mathbb{R}_+$,

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

the generator of $(\boldsymbol{Y}_N(t), \tilde{\eta}_N(t))$ is given by $B_N g = \tilde{\phi}_N g + \Gamma_N g$, where Γ_N is given in (3.4) and

$$\begin{aligned}\tilde{\phi}_N g(\boldsymbol{y}, u) = {} & 2\binom{y_0}{2}(g(\boldsymbol{y} - \boldsymbol{e}_0, u) - g(\boldsymbol{y}, u)) \\ & + (\beta u + \theta) y_0 (g(\boldsymbol{y} - \boldsymbol{e}_0 + \boldsymbol{e}_1, u + \varepsilon_N) - g(\boldsymbol{y}, u)) \\ & + \frac{\gamma u^{\alpha-1}}{u+1} y_1 (g(\boldsymbol{y} - \boldsymbol{e}_1 + \boldsymbol{e}_0, u - \varepsilon_N) - g(\boldsymbol{y}, u))\, \mathbf{1}_{\{u>0\}} \\ & + \left(\varepsilon_N R_{1,N} g(\boldsymbol{y}, u) + \left(\varepsilon_N + \frac{1}{N}\right) \frac{u^{\alpha-1}}{u+1} \mathbf{1}_{\{u>0\}} R_{2,N} g(\boldsymbol{y}, u)\right).\end{aligned}$$

Furthermore, it follows from Lemma 3.1 that

$$(\boldsymbol{Y}_N(0), \tilde{\eta}_N(0)) \xrightarrow{\mathbb{P}} (\boldsymbol{y}, \kappa) \quad \text{as } N \to \infty.$$

We can choose some $\delta > 0$ such that $\kappa - \delta > 0$ and we let $I = [\kappa - \delta, \kappa + \delta]$. Also by Lemma 3.1,

$$\lim_{n\to\infty} \mathbb{P}\{(\boldsymbol{Y}_N(t), \tilde{\eta}_N(t)) \in \Pi \times I,\ 0 \le t \le T\} = 1.$$

Note that Π is a finite set, so the discrete topology on Π makes it a complete and compact metric space. Then the above limit implies the compact containment condition of Ethier and Kurtz [7, Equation (7.9), Chapter 3] for $(\boldsymbol{Y}_N(t), \tilde{\eta}_N(t))$. It is not hard to see that for any continuous and bounded function g on $\Pi \times \mathbb{R}_+$,

$$\lim_{n\to\infty} \sup_{(\boldsymbol{y},u)\in\Pi\times I} |B_N g(\boldsymbol{y}, u) - Bg(\boldsymbol{y}, u)| = 0,$$

where

$$\begin{aligned}Bg(\boldsymbol{y}, u) = {} & 2\binom{y_0}{2}(g(\boldsymbol{y} - \boldsymbol{e}_0, u) - g(\boldsymbol{y}, u)) + (\beta u + \theta) y_0 (g(\boldsymbol{y} - \boldsymbol{e}_0 + \boldsymbol{e}_1, u) - g(\boldsymbol{y}, u)) \\ & + \frac{\gamma u^{\alpha-1}}{u+1} y_1 (g(\boldsymbol{y} - \boldsymbol{e}_1 + \boldsymbol{e}_0, u) - g(\boldsymbol{y}, u)) + (\beta u + \theta - \gamma u^{\alpha}) \frac{\partial}{\partial u} g(\boldsymbol{y}, u).\end{aligned}$$

It follows from [7, Corollary 8.7] that $\{(\boldsymbol{Y}_N(t), \tilde{\eta}_N(t)) : t \ge 0\}$ converges weakly on $D([0, \infty), \Pi \times \mathbb{R}_+)$ to $\{(\boldsymbol{Y}(t), \kappa) : t \ge 0\}$ as $N \to \infty$.

Proof of Theorem 3.1. Let $\mathbb{P}_{\boldsymbol{y}}(\cdot)$ be the distribution of $(\boldsymbol{Y}_N(\cdot), \tilde{\eta}_N(\cdot))$ with initial value $(\boldsymbol{y}, \tilde{\eta}_N(0))$, where $\tilde{\eta}_N(\cdot)$ is distributed as π_N given in Section 2. Let $f_1, \dots, f_k$ be real-valued functions on E. Choose $0 < s < t_1 < \cdots < t_k$. Let $Q_N = \{\sigma_0^N < s < \sigma_1^N\}$. Then

$$\begin{aligned}& \mathbb{E}_{\boldsymbol{y}}\left\{\prod_{i=1}^k f_i(\boldsymbol{Y}_N(t_i))\, \mathbf{1}_{Q_N}\right\} \\ & \quad = \mathbb{E}_{\boldsymbol{y}}\left\{\mathbf{1}_{Q_N}\, \mathbb{E}\left\{\prod_{i=1}^k f_i(\boldsymbol{Y}_N(t_i)) \,\middle|\, \mathcal{F}_s^N\right\}\right\} \\ & \quad = \mathbb{E}_{\boldsymbol{y}}\left\{\mathbf{1}_{Q_N}\, \mathbb{E}_{(\bar{\boldsymbol{y}}, \tilde{\eta}_N(s))}\left\{\prod_{i=1}^k f_i(\boldsymbol{Y}_N(t_i - s))\right\}\right\}\end{aligned}$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

$$= \mathbb{E}_{\boldsymbol{y}}\left\{\mathbb{E}_{(\bar{\boldsymbol{y}},\tilde{\eta}_N(s))}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}_N(t_i - s))\right\}\right\} - \mathbb{E}_{\boldsymbol{y}}\left\{\mathbf{1}_{\bar{Q}_N}\,\mathbb{E}_{((\bar{\boldsymbol{y}},\tilde{\eta}_N(s))}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}_N(t_i - s))\right\}\right\}$$
$$= \mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}_N(t_i - s))\right\} - \mathbb{E}_{\boldsymbol{y}}\left\{\mathbf{1}_{\bar{Q}_N}\,\mathbb{E}_{(\bar{\boldsymbol{y}},\tilde{\eta}_N(s))}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}_N(t_i - s))\right\}\right\},$$

where $\bar{Q}_N$ is the complement of the set Q_N and $\mathbb{E}$ is the expectation value. The last equality follows from the fact that $\tilde{\eta}_N(\cdot)$ is stationary. Then by Lemmas 3.2, 3.3, and 3.4,

$$\begin{aligned}
&\limsup_{N\to\infty}\left|\mathbb{E}_{\boldsymbol{y}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}_N(t_i))\right\} - \mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}(t_i - s))\right\}\right| \\
&\quad\le 2\max_i \|f_i\| \limsup_{N\to\infty}\mathbb{P}(\bar{Q}_N) \\
&\qquad + \limsup_{N\to\infty}\left|\mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}_N(t_i - s))\right\} - \mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}(t_i - s))\right\}\right| \\
&\quad\le 2M\max_i \|f_i\|(1 - \mathrm{e}^{-K_1 s}) \\
&\quad\to 0 \quad \text{as } s \to 0.
\end{aligned}$$

Since $\boldsymbol{Y}(t)$ is stochastically continuous,

$$\mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}(t_i - s))\right\} \xrightarrow{\mathbb{P}} \mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}(t_i))\right\} \quad \text{as } s \to 0.$$

Then, we have

$$\lim_{N\to\infty}\mathbb{E}_{\boldsymbol{y}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}^N(t_i))\right\} = \mathbb{E}_{\bar{\boldsymbol{y}}}\left\{\prod_{i=1}^{k} f_i(\boldsymbol{Y}(t_i))\right\}. \tag{3.7}$$

On the other hand, for the process $\boldsymbol{Y}^*(\cdot)$ defined by (3.2), the transition probability $p_t^*(\boldsymbol{y},\cdot)$ is given by

$$p_t^*(\boldsymbol{y},\cdot) = \begin{cases} p_t(\bar{\boldsymbol{y}},\cdot) & \text{if } t > 0, \\ \delta_{\boldsymbol{y}}(\cdot) & \text{if } t = 0, \end{cases} \tag{3.8}$$

where $p_t(\bar{\boldsymbol{y}},\cdot)$ is the transition probability of $\boldsymbol{Y}(\cdot)$. Let $t_0 = 0$. We also have

$$\lim_{N\to\infty}\mathbb{E}\left\{\prod_{i=0}^{k} f_i(\boldsymbol{Y}^N(t_i))\right\} = \mathbb{E}\left\{\prod_{i=0}^{k} f_i(\boldsymbol{Y}^*(t_i))\right\}.$$

Proof of Theorem 3.2. Step 1. Recall the notation in Section 2.2. Under the homeomorphism, the subspace Π can be regarded as $\Gamma(\Pi)$ for simplicity. It is not hard to see that for any function f on Π,

$$f(\boldsymbol{Y}_k(t)) - f(\boldsymbol{y}) - \int_0^t (B_k f)(\boldsymbol{Y}_k(s))\,\mathrm{d}s$$

is a martingale, where

$$\begin{aligned}
B_k f(\boldsymbol{y})r = {}& 2\binom{y_0}{2}(f(\boldsymbol{y} + (-1,0)) - f(\boldsymbol{y})) + a_k y_0(f(\boldsymbol{y} + (-1,1)) - f(\boldsymbol{y})) \\
&+ b_k y_1(f(\boldsymbol{y} + (1,-1)) - f(\boldsymbol{y})).
\end{aligned}$$

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

Recall that $\boldsymbol{Y}_k(t) = (Y_k^0(t), Y_k^1(t))$ and $Y_k(t) = Y_k^0(t) + Y_k^1(t)$. For any function g on I_n, let $f(\boldsymbol{y}) = g(y_0 + y_1)$ for $\boldsymbol{y} \in \Pi$. Then

$$g(Y_k(t)) - g(n) - \int_0^t (\tilde{B}_k g)(Y_k^0(s), Y_k^1(s))\,\mathrm{d}s$$

is also a martingale, where $\tilde{B}_k g(y) = 2\binom{y_0}{2}(g(y-1) - g(y))$. Note that I_n is a finite set, so the discrete topology on I_n makes it a complete and compact metric space. Then $Y_k(\cdot)$ satisfies the compact containment condition. For each $T > 0$, $\sup_k \int_0^T |\tilde{B}_k g(Y_k^0(s), Y_k(s))|\,\mathrm{d}s \le 2n^2 T\|g\|$, where $\|g\| = \sup_{y\in I_n} |g(y)|$. By Ethier and Kurtz [7, Theorems 9.1 and 9.4, p. 142], $Y_k(\cdot)$ is relatively compact in $D([0,\infty), I_n)$.

Step 2. Suppose that $\{\xi_j^k(\cdot)\}_{j=1}^n$ is the sequence of i.i.d. Markov chains taking values in $\{0, 1\}$ and whose transition rate matrix is given by

$$\begin{pmatrix} -1 & 1 \\ \dfrac{b_k}{a_k} & -\dfrac{b_k}{a_k} \end{pmatrix}.$$

Let $P_{ij}^k(t) = \mathbb{P}\{\xi_1^k(t) = j \mid \xi_1^k(0) = i\}$. A simple calculation shows that

$$P_{00}^k(t) = 1 - P_{01}^k(t) = \frac{b_k}{a_k + b_k} + \frac{a_k}{a_k + b_k}\mathrm{e}^{-(1+(b_k/a_k))t},$$
$$P_{10}^k(t) = 1 - P_{11}^k(t) = \frac{b_k}{a_k + b_k} - \frac{b_k}{a_k + b_k}\mathrm{e}^{-(1+(b_k/a_k))t}.$$

Let $\zeta_n^k(t) = \sum_{j=1}^n \mathbf{1}_{\{\xi_j^k(t)=0\}}$. Since $\{\xi_i^k(t)\}_{i=1}^n$ are independent of each other, it is not hard to see that for any g on I_n,

$$\sup_{x,y\in I_n} |\mathbb{E}_x\{g(\zeta_n^k(t))\} - \mathbb{E}_y\{g(\zeta_n^k(t))\}| \le 2n\|g\|\mathrm{e}^{-(1+(b_k/a_k))t}, \qquad t \ge 0.$$

This implies that $\zeta_n^k(t)$ satisfies the ϕ-mixing condition (see [4, p. 111]). By (1.13) of [4, p. 109],

$$\begin{aligned}\sup_{y\in I_n} &|\mathbb{E}_y\{g(\zeta_n^k(t_2))g(\zeta_n^k(t_1))\} - \mathbb{E}_y\{g(\zeta_n^k(t_2))\}\mathbb{E}_y\{g(\zeta_n^k(t_1))\}| \\ &\le 2\sqrt{2n}\|g\|^2\mathrm{e}^{-(1+(b_k/a_k))(t_2-t_1)/2}\end{aligned}$$

for any $t_2 \ge t_1 \ge 0$. Then

$$\begin{aligned}&\mathbb{E}_y\left\{\left(\int_0^t (g(\zeta_n^k(a_k s)) - \mathbb{E}_y[g(\zeta_n^k(a_k s))])\,\mathrm{d}s\right)^2\right\} \\ &\quad= \mathbb{E}_y\left\{\int_0^t\int_0^t \mathrm{d}s_1\,\mathrm{d}s_2(g(\zeta_n^k(a_k s_1)) - \mathbb{E}_y[g(\zeta_n^k(a_k s_1))])(g(\zeta_n^k(a_k s_2)) - \mathbb{E}_y[g(\zeta_n^k(a_k s_2))])\right\} \\ &\quad= \int_0^t\int_0^t \mathrm{d}s_1\,\mathrm{d}s_2(\mathbb{E}_y\{g(\zeta_n^k(a_k s_2))g(\zeta_n^k(a_k s_1))\} - \mathbb{E}_y\{g(\zeta_n^k(a_k s_2))\}\mathbb{E}_y\{g(\zeta_n^k(a_k s_1))\}) \\ &\quad\le C(n)\|g\|^2\int_0^t \mathrm{d}s_2\int_0^t \mathrm{e}^{-(a_k+b_k)|s_2-s_1|/2}\,\mathrm{d}s_1 \\ &\quad\le C(n)\|g\|^2 t/a_k, \qquad (3.9)\end{aligned}$$

where $C(n)$ is a constant depending only on n. Since $P_{00}^k(a_k t) \to p/(1+p)$ and $P_{01}^k(a_k t) \to 1/(1+p)$ as $k \to \infty$, it is easy to see for any $t \ge 0$, $\zeta_n^k(a_k t) \overset{\mathrm{D}}{=} \zeta_n$ as $k \to \infty$, where ζ_n follows

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

the binomial distribution, i.e. $\zeta_n \sim \mathrm{bn}(n, p/(1+p))$. Note that I_n is finite. The dominated convergence theorem shows that

$$\sup_{y\in I_n} \int_0^t |\mathbb{E}_y\{g(\zeta_n^k(a_k s))\} - \mathbb{E}\{g(\zeta_n)\}|\,\mathrm{d}s \to 0 \quad \text{as } k \to \infty.$$

Combined with (3.9), we have

$$\sup_{y\in I_n} \mathbb{E}_y\left\{\left(\int_0^t (g(\zeta_n^k(a_k s)) - \mathbb{E}[g(\zeta_n)])\,\mathrm{d}s\right)^2\right\} \to 0 \quad \text{as } k \to \infty. \tag{3.10}$$

Step 3. It holds that $(Y_k^0(t), Y_k(t))$ is a Markov process as in step 1 and $Y_k^0(0) = y \in I_n$ and $Y_k(0) = n$. Let $\mathcal{F}_t^k = \sigma\{(Y_k^0(s), Y_k(s)) : 0 \le s \le t\}$. Define $T_j^k = \inf\{t \ge 0 : Y_k(t) = n - j\}$ with $T_0^k = 0$ and $\tau_j^k = T_j^k - T_{j-1}^k$. Set $h(y) = y(y-1)$ for $y \in I_n$. By (3.10), we have

$$\begin{aligned}
\mathbb{P}\{\tau_{j+1}^k > t\} &= \mathbb{E}\{\mathbb{P}\{\tau_{j+1}^k > t \mid \mathcal{F}_{T_j^k}^k\}\} \\
&= \mathbb{E}\{\mathbb{P}_{\{Y_k^0(T_j^k), n-j\}}(\tau_{j+1}^k > t)\} \\
&= \mathbb{E}\left\{\mathbb{E}_{Y_k^0(T_j^k)}\left(\exp\left\{-\int_0^t h(\zeta_{n-j}^k(a_k s))\,\mathrm{d}s\right\}\right)\right\} \\
&\to \mathrm{e}^{-\mathbb{E}\{h(\zeta_{n-j})\}t} \quad \text{as } k \to \infty.
\end{aligned}$$

Similarly,

$$\mathbb{P}\{\tau_1^k > t, \tau_2^k > s\} = \mathbb{E}\left\{\mathbf{1}_{\{\tau_1^k > t\}}\,\mathbb{E}_{Y_k^0(T_1^k)}\left(\exp\left\{-\int_0^t h(\zeta_{n-1}^k(a_k s))\,\mathrm{d}s\right\}\right)\right\}.$$

Then

$$\begin{aligned}
&|\mathbb{P}\{\tau_1^k > s, \tau_2^k > t\} - \mathrm{e}^{-\mathbb{E}\{h(\zeta_n)\}s - \mathbb{E}\{h(\zeta_{n-1})\}t}| \\
&\quad \le |\mathbb{P}\{\tau_1^k > s\} - \mathrm{e}^{-\mathbb{E}\{h(\zeta_n)\}s}| + \sup_{y\in I_n} \mathbb{E}_y\left\{\left|\int_0^t (h(\zeta_{n-1}^k(a_k s)) - \mathbb{E}\{h(\zeta_{n-1})\})\,\mathrm{d}s\right|\right\}.
\end{aligned}$$

By (3.10) it follows that the second term in the right-hand side of the above inequality goes to 0 as $k \to \infty$. By induction, $(\tau_1^k, \dots, \tau_{n-1}^k) \overset{\mathrm{D}}{=} (\tau_1, \dots, \tau_{n-1})$, where $\{\tau_j\}_{j=1}^{n-1}$ are independent of each other and τ_j follows the exponential distribution with parameter c_{n-j+1}. It follows that $\{Y_k(t), t \ge 0\}$ converges in the sense of finite-dimensional distributions to the n-Kingman coalescent process $\{K(t), t \ge 0\}$. Since $\{Y_k(t)\}$ is relatively compact, the theorem is proved.

Acknowledgements

AL and CM were financially supported by grant MANEGE 'Modèles Aléatoires en Écologie, Génétique et Évolution' 09-BLAN-0215 of ANR (French national research agency). AL also thanks the Center for Interdisciplinary Research in Biology (Collège de France) for funding. CM also thanks the National Natural Science Foundation of China (NSFC) (grant number 11001137 and grant number 11271204) and the China Scholarship Council (CSC).

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press

References

[1] Aguilée, R., Claessen, D. and Lambert, A. (2009). Allele fixation in a dynamic metapopulation: founder effects vs refuge effects. *Theoret. Popul. Biol.* **76,** 105–117.

[2] Aguilée, R., Claessen, D. and Lambert, A. (2013). Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. *Evolution* **67,** 1291–1306.

[3] Aguilée, A., Lambert, A. and Claessen, D. (2011). Ecological speciation in dynamic landscapes. *J. Evolutionary Biol.* **24,** 2663–2677.

[4] Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Prob. Surveys* **2,** 107–144.

[5] Coyne, J. A. and Orr, H. A. (2004). *Speciation*. Sinauer Associates, Sunderland, MA.

[6] Eldon, B. (2009). Structured coalescent processes from a modified Moran model with large offspring numbers. *Theoret. Popul. Biol.* **76,** 92–104.

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

[8] Hanski, I. and Gilpin, M. E. (eds) (1997). *Metapopulation Biology: Ecology, Genetics, and Evolution.* Academic Press, San Diego, CA.

[9] Herbots, H. M. (1997). The structured coalescent. In *Progress in Population Genetics and Human Evolution*, Springer, New York, pp. 231–255.

[10] Kelly, F. P. (1979). *Reversibility and Stochastic Networks.* John Wiley, Chichester.

[11] Keymer, J. E., Marquet, P. A., Velasco-Hernández, J. X. and Levin, S. A. (2000). Extinction thresholds and metapopulation persistence in dynamic landscapes. *Amer. Naturalist* **156,** 478–494.

[12] Kingman, J. F. C. (1982). The coalescent. *Stoch. Process. Appl.* **13,** 235–248.

[13] Lambert, A. (2010). Population genetics, ecology and the size of populations. *J. Math. Biol.* **60,** 469–472.

[14] Nordborg, M. and Krone, S. M. (2002). Separation of time scales and convergence to the coalescent in structured populations. In *Modern Developments in Theoretical Population Genetics*, Oxford University Press, pp. 194–232.

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

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

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

[18] Taylor, J. E. and Véber, A. (2009). Coalescent processes in subdivided populations subject to recurrent mass extinctions. *Electron. J. Prob.* **14,** 242–288.

https://doi.org/10.1239/jap/1437658614 Published online by Cambridge University Press