ON THE EXTENDED MORAN MODEL AND ITS RELATION TO COALESCENTS WITH MULTIPLE COLLISIONS

Thierry Huillet, Martin Moehle

ON THE EXTENDED MORAN MODEL AND ITS RELATION TO COALESCENTS WITH MULTIPLE COLLISIONS

T. Huillet[1] and M. Möhle[2]

April 28, 2011

Abstract

We study the asymptotics of the extended Moran model as the total population size N tends to infinity. Two convergence results are provided, the first result leading to discrete-time limiting coalescent processes and the second result leading to continuous-time limiting coalescent processes. The limiting coalescent processes allow for multiple mergers of ancestral lineages (Λ-coalescent). It is furthermore verified that any continuous time Λ-coalescent (with Λ any probability distribution) can arise in the limit. Typical examples of extended Moran models are discussed, with an emphasis on models being in the domain of attraction of beta coalescents or Λ-coalescents with Λ being log infinitely divisible.

Keywords: beta coalescent; Cannings model; Λ-coalescent; exchangeability; log infinitely divisible distributions; Moran model; multiple collisions

2010 Mathematics Subject Classification: Primary 92D25; 60J70 Secondary 92D15; 60F17

1 Introduction

Cannings models [4, 5] are population models with non-overlapping generations and fixed population size $N \in \mathbb{N} := \{1, 2, \ldots\}$. These models are characterized by a family of random variables $\nu_i(t), t \in \mathbb{Z} := \{\ldots, -1, 0, 1, \ldots\}, i \in \{1, \ldots, N\}$, where $\nu_i(t)$ denotes the number of offspring of individual i of generation t. It is assumed that for each generation $t \in \mathbb{Z}$ the offspring variables $\nu_1(t), \ldots, \nu_N(t)$ are exchangeable.

We consider a particular subclass of Cannings population models in which in each generation $t \in \mathbb{Z}$ only one of the N individuals is allowed to have more than 1 offspring. More precisely, our model is defined in terms of a family of random variables $U_N(t)$, $t \in \mathbb{Z}$, each taking values in $\{0, \ldots, N\}$, as follows. For $t \in \mathbb{Z}$ and $i \in \{1, \ldots, N\}$ define

$$\mu_i(t) := \begin{cases} 1 & \text{if } i \in \{1, \ldots, N - U_N(t)\}, \\ U_N(t) & \text{if } i = N - U_N(t) + 1, \\ 0 & \text{if } i \in \{N - U_N(t) + 2, \ldots, N\}. \end{cases}$$

[1]Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 2 Avenue Adophe Chauvin, 95302 Cergy-Pontoise, France, E-mail address: thierry.huillet@u-cergy.fr

[2]Mathematisches Institut, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany, E-mail address: martin.moehle@uni-tuebingen.de

Now let $\nu_1(t), \ldots, \nu_N(t)$ be a random permutation of $\mu_1(t), \ldots, \mu_N(t)$. For each fixed $t \in \mathbb{Z}$ the random variables $\nu_1(t), \ldots, \nu_N(t)$ are then exchangeable and we interpret $\nu_i(t)$ as the number of offspring of individual i of generation t of a corresponding exchangeable Cannings model. It is assumed that, for each fixed $N \in \mathbb{N}$, the random variables $U_N(t)$, $t \in \mathbb{Z}$, are independent and identically distributed. We write $\nu_i := \nu_i(0)$ and $U_N := U_N(0)$ for convenience. The most celebrated example is the standard Moran model (see [16] or [7, p. 636]) corresponding to $U_N \equiv 2$, in which one randomly selected individual has two offspring, one other randomly selected individual has no offspring and all the other $N-2$ individuals have exactly one offspring. For $U_N \equiv 1$ (and, by definition, as well for $U_N \equiv 0$) we obtain the trivial model in which every individual has exactly one offspring. For $U_N \equiv N$ one randomly chosen individual is the parent of all the N children of the next generation. Eldon and Wakeley [6] focus on a model where U_N takes the value 2 (standard Moran model) with high probability and a value of order N (corresponding to a large reproduction event) with complementary probability. Extended Moran models are briefly mentioned by Birkner and Blath [2] and Lessard [13], and are a bit further analyzed in [9]. The main goal of this paper is to provide further insight in this subclass of Cannings models, in particular in the behavior of the model as the total population size N tends to infinity. Typically, for large population size, the ancestry of a sample of size n taken from some generation is well approximated by an n-coalescent with multiple collisions (Proposition 3.1 and Theorem 3.3), and essentially any such coalescent with multiple collisions can arise in the limit (Proposition 3.4). Similar results are known for related but different Cannings population models, for example models based on particular sampling schemes [22] and models based on certain conditioning procedures [9]. Our results heavily gain from the theory on Cannings models and from coalescent theory [15].

The paper is organized as follows. In Section 2 results are presented which hold for arbitrary but fixed population size N. The main Section 3 presents asymptotic results as the total population size N tends to infinity. Further examples are discussed in Section 4. As a side effect we provide in the appendix (Section 5) exact and asymptotic formulas for the total rates of arbitrary beta coalescents.

We use the notation $(x)_k := x(x-1)\cdots(x-k+1)$, $x \in \mathbb{R}$, $k \in \mathbb{N}_0$, for the descending factorials, with the convention that $(x)_0 := 1$. Throughout the article $B(.,.)$ denotes the beta function and $\beta(a,b)$ the beta distribution with parameters $a, b \in (0,\infty)$. We furthermore write $X \stackrel{d}{=} Q$ if the random variable X has distribution Q.

2 Offspring, descendants, and ancestors

We start with expressing the joint offspring distribution of $(\nu_1, \ldots, \nu_N)$ in terms of U_N. Obviously, $\mathbb{P}(\nu_1 = 1, \ldots, \nu_N = 1) = \mathbb{P}(U_N \in \{0,1\})$. It is furthermore

readily checked that

$$\mathbb{P}(\nu_1 = j_1, \ldots, \nu_N = j_N) \;=\; \frac{(N-j_k)!(j_k-1)!}{N!}\mathbb{P}(U_N = j_k)$$

whenever there exists one index $k \in \{1,\ldots,N\}$ such that $j_k \in \{2,\ldots,N\}$, $j_l \in \{0,1\}$ for all $l \in \{1,\ldots,N\} \setminus \{k\}$ and $j_1 + \cdots + j_N = N$. The marginal distribution of ν_i, $i \in \{1,\ldots,N\}$, is therefore given by

$$\mathbb{P}(\nu_i = j) \;=\; \begin{cases} \displaystyle\sum_{k=2}^{N} \frac{k-1}{N}\mathbb{P}(U_N = k) & \text{for } j = 0, \\ \displaystyle 1 - \sum_{k=2}^{N} \frac{k}{N}\mathbb{P}(U_N = k) & \text{for } j = 1, \\ \displaystyle \frac{\mathbb{P}(U_N = j)}{N} & \text{for } j \in \{2,\ldots,N\}. \end{cases}$$

Let us briefly consider the model forwards in time. Take a sample of $n \in \{0,\ldots,N\}$ individuals from generation 0 and, for $t \in \mathbb{N}_0$, let X_t denote the number of descendants of these n individuals t generations forwards in time. It is well known ([9, Example 2.2]) that the process $X := (X_t)_{t\in\mathbb{N}_0}$, called the forward process, is a time-homogeneous Markov chain with state space $\{0,\ldots,N\}$, initial state $X_0 = n$. If $U_N \geq 1$ almost surely, then X has transition probabilities

$$P_{i,j} \;:=\; \mathbb{P}(X_{t+1} = j \,|\, X_t = i) \;=\; \begin{cases} \mathbb{E}\big(\binom{N-U_N}{j}\binom{U_N-1}{i-j}\big)\big/\binom{N}{i} & \text{if } j < i, \\ \mathbb{E}\big(\binom{N-U_N}{j} + \binom{N-U_N}{N-j}\big)\big/\binom{N}{i} & \text{if } j = i, \\ \mathbb{E}\big(\binom{N-U_N}{N-j}\binom{U_N-1}{j-i}\big)\big/\binom{N}{i} & \text{if } j > i. \end{cases} \tag{1}$$

If $\mathbb{P}(U_N = 0) > 0$, then on the right hand side in (1) each U_N has to be replaced by $\max(U_N, 1)$.

For our purposes it is more important to analyze the model backwards in time. Take a sample of $n \in \{1,\ldots,N\}$ individuals from generation 0, and for each $t \in \mathbb{N}_0$ define a random equivalence relation $\mathcal{R}_t = {}_{(N)}\mathcal{R}_t^{(n)}$ such that, by definition, $(i,j) \in \mathcal{R}_t$ if and only if the individuals i and j of the sample have a common parent t generations backwards in time. Since it is assumed that the random variables $U_N(t)$, $t \in \mathbb{Z}$, are independent and identically distributed, the so-called ancestral process $(\mathcal{R}_t^{(n)})_{t\in\mathbb{N}_0}$ is a time-homogeneous Markov chain with state space $\mathcal{E}_n$, the set of equivalence relations on $\{1,\ldots,n\}$. If the ancestral process is in a state $\xi \in \mathcal{E}_n$ with $i := |\xi|$ blocks, then the process moves to any equivalence relation $\eta \in \mathcal{E}_n$, which is obtained from ξ by merging exactly $i - j + 1 \geq 2$ blocks of ξ ($\Rightarrow |\eta| = j$), with probability

$$p_{\xi\eta} \;:=\; \mathbb{P}(\mathcal{R}_{t+1}^{(n)} = \eta \,|\, \mathcal{R}_t^{(n)} = \xi) \;=\; \frac{\mathbb{E}((U_N)_{i-j+1}(N-U_N)_{j-1})}{(N)_i}, \tag{2}$$

$|\eta| = j \in \{1,\ldots,i-1\}$, and the process stays in the state ξ with probability

$$p_{\xi\xi} \;:=\; \mathbb{P}(\mathcal{R}_{t+1}^{(n)} = \xi \,|\, \mathcal{R}_t^{(n)} = \xi)$$

$$
\begin{aligned}
&= 1 - \sum_{j=1}^{i-1} \binom{i}{j-1} \frac{\mathbb{E}((U_N)_{i-j+1}(N-U_N)_{j-1})}{(N)_i} \\
&= 1 - \sum_{j=1}^{i-1} \frac{\mathbb{E}\big(\binom{U_N}{i-j+1}\binom{N-U_N}{j-1}\big)}{\binom{N}{i}} \\
&= \sum_{j=i}^{i+1} \frac{\mathbb{E}\big(\binom{U_N}{i-j+1}\binom{N-U_N}{j-1}\big)}{\binom{N}{i}} \;=\; \frac{\mathbb{E}\big(U_N\binom{N-U_N}{i-1} + \binom{N-U_N}{i}\big)}{\binom{N}{i}}. \qquad (3)
\end{aligned}
$$

From (2) and (3) it is readily checked that the backward process $(Y_t^{(n)})_{t\in\mathbb{N}_0} := (|\mathcal{R}_t^{(n)}|)_{t\in\mathbb{N}_0}$, which counts the number of ancestors, is a Markov chain with state space $\{1,\dots,n\}$ and transition probabilities

$$
\widehat{P}_{i,j} \;:=\; \mathbb{P}(Y_{t+1}^{(n)} = j \,|\, Y_t^{(n)} = i) \;=\; \begin{cases} \mathbb{E}\big(\binom{N-U_N}{j-1}\binom{U_N}{i-j+1}\big)/\binom{N}{i} & \text{if } j < i, \\ \mathbb{E}\big(\binom{N-U_N}{i} + U_N\binom{N-U_N}{i-1}\big)/\binom{N}{i} & \text{if } j = i, \\ 0 & \text{if } j > i, \end{cases}
$$

in agreement with [9, Section 2]. In particular, the backward process moves from a state $i \geq 2$ to the state 1 with transition probability $\widehat{P}_{i,1} = \mathbb{E}((U_N)_i)/(N)_i$. The extended Moran model has therefore coalescence probability $c_N := \widehat{P}_{2,1} = \mathbb{E}((U_N)_2)/(N)_2$, in agreement with [6, Eq. (2)]. Note that c_N is the probability that two individuals, chosen at random from some generation, have a common parent. Depending on the choice of U_N, the effective population size $N_e := 1/c_N$ can be of very different nature. For example, if $U_N \equiv N$, then $N_e = 1$ for all N. On the other hand, since $c_N \leq \mathbb{P}(U_N \geq 2)$, N_e can grow arbitrarily fast with N, namely when the probability $\mathbb{P}(U_N \geq 2)$ tends to 0 extremely fast as $N \to \infty$.

3 Asymptotic results

In this section we are interested in the behavior of the model if the total population size N tends to infinity. It will turn out that, under mild conditions, the ancestral process of the extended Moran model can be approximated by a coalescent process with multiple collisions (Λ-coalescent). For fundamental information on Λ-coalescents we refer the reader to the papers of Pitman [18], Sagitov [19], and Schweinsberg [21].

For $n \in \mathbb{N}$ let ϱ_n denote the restriction from $\mathcal{E}$, the set of all equivalence relations on $\mathbb{N}$, to $\mathcal{E}_n$. As in Definition 2.1 of [9], we say that the considered population model is in the domain of attraction of a continuous-time coalescent process $\Pi = (\Pi_t)_{t\in[0,\infty)}$, if, for each sample size $n \in \mathbb{N}$, the time-scaled ancestral process $(\mathcal{R}^{(n)}_{\lfloor t/c_N \rfloor})_{t\in[0,\infty)}$ weakly converges to $(\varrho_n \circ \Pi_t)_{t\in[0,\infty)}$ as $N \to \infty$. Analogously, we say that the considered population model is in the domain of attraction of a discrete-time coalescent process $(\Pi_t)_{t\in\mathbb{N}_0}$, if, for each sample size $n \in \mathbb{N}$, the ancestral process $(\mathcal{R}_t^{(n)})_{t\in\mathbb{N}_0}$ weakly converges to $(\varrho_n \circ \Pi_t)_{t\in\mathbb{N}_0}$ as $N \to \infty$.

Proposition 3.1 *Suppose that $U_N/N \to U$ in distribution as $N \to \infty$ for some random variable U satisfying $\mathbb{E}(U) > 0$. Then the extended Moran model is in the domain of attraction of a discrete-time Λ-coalescent with $\Lambda(dx) := x^2\nu(dx)$, where ν denotes the distribution of U.*

Remark. The measure Λ in Proposition 3.1 satisfies $\Lambda(\{0\}) = 0$ and $\int_{(0,1]} x^{-2}\,\Lambda(dx) = \nu((0,1]) \leq 1$.

Proof. The convergence $U_N/N \to U$ in distribution as $N \to \infty$ is equivalent to the convergence of all the moments $\mathbb{E}((U_N/N)^k) \to \mathbb{E}(U^k)$, $k \in \mathbb{N}_0$. For all $k \in \{2, 3, \ldots\}$ it follows that

$$\widehat{P}_{k,1} \;=\; \frac{\mathrm{E}((U_N)_k)}{(N)_k} \;=\; \frac{1}{(N)_k}\sum_{l=0}^{k} s(k,l)\,\mathbb{E}(U_N^l) \;\to\; \mathbb{E}(U^k) \;=:\; \phi_1(k) \;>\; 0,$$

where the $s(k,l)$ denote the Stirling numbers of the first kind. By definition of the extended Moran model, simultaneous multiple collisions of ancestral lineages are impossible. This property carries over to a potential limiting process, so if a limiting coalescent process exists this limiting coalescent cannot have simultaneous multiple collisions and hence must be a Λ-coalescent allowing only for multiple collisions. One may deduce this property also more formally using Theorem 2.1 and Corollary 2.2 of Sagitov [20]. By [15, Theorem 2.1], the convergence $\widehat{P}_{k,1} \to \phi_1(k)$ for all $k \in \{2, 3, \ldots\}$ implies that the model is in the domain of attraction of a discrete-time Λ-coalescent. The measure Λ is uniquely determined via its moments $\phi_1(k) = \int_{[0,1]} x^{k-2}\,\Lambda(dx)$, $k \geq 2$. Comparing with $\phi_1(k) = \mathbb{E}(U^k)$ shows that $\Lambda(dx) = x^2\nu(dx)$, where ν is the distribution of U.□

Examples. Let U be a random variable taking values in $[0,1]$.

1. Suppose that $U_N = \lfloor NU \rfloor$ for all $N \in \mathbb{N}$. From $NU - 1 \leq U_N \leq NU$ it follows that $U - 1/N \leq U_N/N \leq U$. Thus, $U_N/N \to U$ almost surely and Proposition 3.1 is applicable whenever $\mathbb{E}(U) > 0$.

2. Suppose that for each $N \in \mathbb{N}$, conditional on U, U_N has a binomial distribution with parameters N and U. Then, $\mathbb{P}(U_N = j) = \binom{N}{j}\mathbb{E}(U^j(1-U)^{N-j})$, $j \in \{0, \ldots, N\}$, and $\mathbb{E}((U_N)_k) = (N)_k\mathbb{E}(U^k)$, $k \in \mathbb{N}_0$. In particular, for arbitrary but fixed $k \in \mathbb{N}_0$ and all $N \geq k$, $\mathbb{E}((U_N)_k)/(N)_k = \mathbb{E}(U^k)$ does not depend on N. For all $k \in \mathbb{N}_0$ it follows that $\mathbb{E}((U_N/N)^k) = N^{-k}\sum_{l=0}^{k} S(k,l)\mathbb{E}((U_N)_l) \to \mathbb{E}(U^k)$ as $N \to \infty$, where the $S(k,l)$ are the Stirling numbers of the second kind. Since $0 \leq U_N/N \leq 1$, this convergence of moments is equivalent to the convergence $U_N/N \to U$ in distribution as $N \to \infty$. Again, Proposition 3.1 is applicable whenever $\mathbb{E}(U) > 0$.

   A canonical such sequence $U_1, U_2, \ldots$ of random variables can be constructed as follows. Let $W_1, W_2, \ldots$ be independent random variables (and independent of U) and suppose that each W_n is uniformly distributed on

$[0,1]$. For $n \in \mathbb{N}$ define the Bernoulli variables $B_n := 1_{\{W_n \leq U\}}$ and let $U_N := U \odot N := \sum_{n=1}^{N} B_n$ denote the number of successes in the first N trials of the infinite exchangeable sequence $(B_n)_{n\in\mathbb{N}}$. Here $\odot$ denotes the binomial thinning operator of Steutel and van Harn [23]. In this case, by de Finetti's theorem (see, for example, [8, Theorem 1.3]) U_N/N converges almost surely to a random variable having the same distribution as U.

Remark. The previous examples show that any discrete-time Λ-coalescent (with Λ being any measure satisfying $\Lambda(\{0\}) = 0$ and $\int_{(0,1]} x^{-2}\,\Lambda(dx) \leq 1$) can arise in Proposition 3.1 as a limiting process.

Obviously, there exist extended Moran models (take for example $U_N \equiv 2$, the standard Moran model) satisfying $U_N/N \to 0$ in distribution as $N \to \infty$. Proposition 3.1 is then not applicable. It turns out that in this situation, one has to compare the factorial moments $\mathbb{E}((U_N)_k)$ for $k \geq 2$ with $\mathbb{E}((U_N)_2)$. We start with the following basic but fundamental observation.

Lemma 3.2 *The following four conditions are equivalent.*

(i) $U_N/N \to 0$ *in distribution as* $N \to \infty$.

(ii) $\lim_{N\to\infty} \mathbb{E}((U_N/N)^p) = 0$ *for all* $p > 0$.

(iii) $\lim_{N\to\infty} \mathbb{E}((U_N/N)^p) = 0$ *for some* $p > 0$.

(iv) $\lim_{N\to\infty} c_N = 0$.

Proof. Obviously, (ii) follows from (i), since for each $p > 0$ the function $x \mapsto x^p$ is continuous and bounded on $[0,1]$. That (ii) implies (iii) is trivial. If (iii) holds, then, for all $\varepsilon > 0$, the Tschebyscheff–Markov inequality yields $\mathbb{P}(U_N/N > \varepsilon) \leq \varepsilon^{-p}\mathbb{E}((U_N/N)^p) \to 0$ as $N \to \infty$. Thus, $U_N/N \to 0$ in probability which implies (i). Therefore, the first three conditions are equivalent. Thus, it remains to verify that (iv) is equivalent to $\lim_{N\to\infty} \mathbb{E}((U_N/N)^2) = 0$. But this is clear, since the coalescence probability c_N satisfies

$$c_N \;=\; \frac{\mathbb{E}((U_N)_2)}{(N)_2} \;\leq\; \frac{\mathbb{E}(U_N^2)}{(N)_2} \;\sim\; \frac{\mathbb{E}(U_N^2)}{N^2}$$

and

$$c_N \;\geq\; \frac{\mathbb{E}(U_N^2) - \mathbb{E}(U_N)}{N^2} \;\geq\; \frac{\mathbb{E}(U_N^2)}{N^2} - \frac{1}{N},$$

since $U_N \leq N$. □

We are now able to formulate the following theorem.

Theorem 3.3 *Suppose that* $U_N/N \to 0$ *in distribution as* $N \to \infty$. *If all the limits*

$$\phi_1(k) \;:=\; \lim_{N\to\infty} \frac{\mathbb{E}((U_N)_k)}{N^{k-2}\mathbb{E}((U_N)_2)}, \qquad k \in \{2,3,\ldots\} \tag{4}$$

exist, then the extended Moran model is in the domain of attraction of a continuous-time Λ*-coalescent, where* Λ *is a probability measure on* $[0,1]$ *uniquely determined via its moments* $\int_{[0,1]} x^{k-2}\,\Lambda(dx) = \phi_1(k)$, $k \in \{2,3,\ldots\}$.

Remarks. 1. If $\phi_1(3) = 0$, then $\Lambda = \delta_0$ is the Dirac measure at 0, so in this case the limiting coalescent is the Kingman coalescent [12].
2. Condition (4) has the following probabilistic interpretation (see also the proof of Lemma 3.1 in [15]): Let X_N be a sort of size biased random variable with distribution $\mathbb{P}(X_N = j) := (j)_2\mathbb{P}(U_N = j)/\mathbb{E}((U_N)_2)$, $j \in \{2, \dots, N\}$. From $\mathbb{E}(X_N^k) = \mathbb{E}(U_N^{k+2} - U_N^{k+1})/\mathbb{E}((U_N)_2)$, $k \in \mathbb{N}_0$, it is readily seen that (4) is equivalent to $\mathbb{E}((X_N/N)^k) \to \phi_1(k+2)$ for all $k \in \mathbb{N}$, and, hence, equivalent to the convergence $X_N/N \to X$ in distribution for some random variable X taking values in $[0, 1]$ with moments $\mathbb{E}(X^k) = \phi_1(k+2)$, $k \in \mathbb{N}$. The probability measure Λ occurring in Theorem 3.3 is the distribution of X.
3. Note that $\mathbb{E}(U_N^2) - \sqrt{\mathbb{E}(U_N^2)} \leq \mathbb{E}(U_N^2) - \mathbb{E}(U_N) = \mathbb{E}((U_N)_2) \leq \mathbb{E}(U_N^2)$ or, equivalently,

$$1 - \frac{1}{\sqrt{\mathbb{E}(U_N^2)}} \;\leq\; \frac{\mathbb{E}((U_N)_2)}{\mathbb{E}(U_N^2)} \;\leq\; 1.$$

Thus, if $\mathbb{E}(U_N^2) \to \infty$ as $N \to \infty$, then in the denominator in (4), $\mathbb{E}((U_N)_2)$ can be replaced by $\mathbb{E}(U_N^2)$. Since $(\mathbb{E}(U_N))^2 \leq \mathbb{E}(U_N^2)$, this replacement is in particular allowed if $\mathbb{E}(U_N) \to \infty$ as $N \to \infty$. A typical model which does not satisfy $\mathbb{E}(U_N) \to \infty$ is given as follows. For each $N \in \mathbb{N}$ let V_N be a random variable having a beta distribution with parameters 1 and N. Define $U_N := V_N \odot N$, $N \in \mathbb{N}$. Then, for all $k \in \mathbb{N}$, $\mathbb{E}((U_N)_k) = (N)_k\mathbb{E}(V_N^k) = (N)_k B(k+1, N)/B(1, N) \to k!$ as $N \to \infty$.

Proof. (of Theorem 3.3) By Lemma 3.2, the assumption $U_N/N \to 0$ in distribution is equivalent to $\lim_{N\to\infty} c_N = 0$. In particular, under this condition, the extended Moran model cannot be in the domain of attraction of a discrete-time limiting coalescent. For $k \in \{2, 3 \dots\}$ we have

$$\frac{\widehat{P}_{k,1}}{c_N} \;=\; \frac{(N)_2\mathbb{E}((U_N)_k)}{(N)_k\mathbb{E}((U_N)_2)} \;\sim\; \frac{\mathbb{E}((U_N)_k)}{N^{k-2}\mathbb{E}((U_N)_2)} \;\to\; \phi_1(k).$$

Moreover, by definition, simultaneous multiple collisions do not occur in the extended Moran model, and, therefore, as also explained in the proof of Proposition 3.1, a potential limiting process can only allow for multiple collisions of ancestral lineages. It remains to apply well known convergence theorems from coalescent theory (see, for example, [15, Theorem 2.1]). □

It is natural to ask if any Λ-coalescent (with Λ being any probability measure on $[0, 1]$) can occur as a limiting process in Theorem 3.3. The following proposition answers this question positively.

Proposition 3.4 *Let Λ be a probability distribution on $[0, 1]$. Then, for each $N \in \mathbb{N}$, there exists a random variable U_N taking values in $\{1, \dots, N\}$ such that $(U_N)_{N\in\mathbb{N}}$ satisfies the conditions of Theorem 3.3 and such that the continuous-time limiting coalescent arising in Theorem 3.3 has characterizing measure Λ.*

Proof. We essentially exploit the construction of a Cannings population model provided on p. 40 in [14]. Fix the probability measure Λ on $[0, 1]$, let X be

a random variable with distribution Λ, and let $\Pi = (\Pi_t)_{t\geq 0}$ be a continuous-time Λ-coalescent. The basic idea (see Case 1 below) is to define U_N to be the number of particles involved in the first collision event of the restricted coalescent $(\varrho_N \circ \Pi_t)_{t\geq 0}$. For $N \in \mathbb{N}$ and $j \in \{1, \ldots, N-1\}$ we need to introduce $\lambda_{Nj} := \int_{[0,1]} x^{j-2}(1-x)^{N-j}\, \Lambda(dx) = \mathbb{E}(X^{j-2}(1-X)^{N-j})$, $j \in \{2, \ldots, N\}$, and the total rates

$$\begin{aligned}
\lambda_N &:= \sum_{j=2}^{N} \binom{N}{j} \lambda_{Nj} \\
&= \binom{N}{2} \Lambda(\{0\}) + \int_{(0,1]} \frac{1-(1-x)^N - Nx(1-x)^{N-1})}{x^2}\, \Lambda(dx) \\
&= \binom{N}{2} \mathbb{P}(X=0) + \mathbb{E}\bigg(\frac{1-(1-X)^N - NX(1-X)^{N-1}}{X^2} 1_{\{X>0\}} \bigg)
\end{aligned}$$

of the Λ-coalescent (see, for example, [18, Eqs. (4) and (6)]). Note that $\lambda_2 = \lambda_{22} = 1$. It is readily checked that for arbitrary but fixed $x \in (0,1]$ the expression $(1-(1-x)^N - Nx(1-x)^{N-1})/x^2$ below the above integral is non-decreasing in N. Thus, the sequence $(\lambda_N)_{N\in\mathbb{N}}$ is non-decreasing. Moreover, $\lambda_N \to \infty$ if $\Lambda(\{0\}) > 0$ and, by the monotone convergence theorem, $\lambda_N \to \int_{(0,1]} x^{-2}\Lambda(dx) \in [0,\infty]$, if $\Lambda(\{0\}) = 0$. In the following it is assumed without loss of generality that $N \geq 2$. Let $\xi \in \mathcal{E}_N$ be an equivalence relation having $j := |\xi| < N$ equivalence classes (blocks) and having exactly one block of size greater than 1. The restricted coalescent $(\varrho_N \circ \Pi_t)_{t\geq 0}$ jumps at its first jump time $T_N := \inf\{t > 0 : |\varrho_N \circ \Pi_t| < N\}$ to the state ξ with probability $\mathbb{P}(\varrho_N \circ \Pi_{T_N} = \xi) = \lambda_{N,N-j+1}/\lambda_N$, A backward induction on k yields

$$\mathbb{P}(\varrho_k \circ \Pi_{T_N} = \xi) \;=\; \frac{\lambda_{k,k-j+1}}{\lambda_N} \tag{5}$$

for all $k \in \{2, \ldots, N\}$ and all $\xi \in \mathcal{E}_k$ with $j := |\xi| < k$ and such that ξ has exactly one block of size greater than 1.

There exists the following alternative representation of (5). Let $B := |\varrho_N \circ \Pi_{T_N}|$ denote the (random) number of blocks of $\varrho_N \circ \Pi_{T_N}$ and let $\mu_1 \geq \cdots \geq \mu_B$ be the sizes of these B blocks. Define $\mu_i := 0$ for $i \in \{B+1, \ldots, N\}$. Since the Λ-coalescent Π allows only for multiple collisions, we have $\mu_2 = \cdots = \mu_B = 1$. Only $\mu_1 = \mu_1(N) = N - B + 1$ may depend on N and takes values in the set $\{2, \ldots, N\}$. Since Π_{T_N} is an exchangeable random partition, it follows from Kingman's representation [12, Eq. (3.13)] that

$$\begin{aligned}
\mathbb{P}(\varrho_k \circ \Pi_{T_N} = \xi) &= \sum_{\substack{r_1,\ldots,r_j=1 \\ \text{all distinct}}}^{N} \frac{\mathbb{E}((\mu_{r_1})_{k-j+1}\mu_{r_2}\cdots\mu_{r_j})}{(N)_k} \\
&= \frac{\mathbb{E}((\mu_1)_{k-j+1}(N-\mu_1)_{j-1})}{(N)_k}.
\end{aligned} \tag{6}$$

The right-hand sides in (5) and (6) coincide, in particular for $j = 1$, which yields

$$\frac{\mathbb{E}((\mu_1)_k)}{(N)_k} \;=\; \frac{\lambda_{kk}}{\lambda_N}, \qquad k \in \{2, \dots, N\}. \tag{7}$$

In particular, for $k \in \{2, 3, \dots\}$ and all $N \geq k$, the quantity

$$\frac{(N)_2 \mathbb{E}((\mu_1)_k)}{(N)_k \mathbb{E}((\mu_1)_2)} \;=\; \frac{\lambda_{kk}}{\lambda_{22}} \;=\; \int_{[0,1]} x^{k-2}\, \Lambda(dx) \tag{8}$$

does not depend on N. In order to finish the proof, two cases are distinguished.

Case 1. Suppose that $\Lambda(\{0\}) > 0$ or that $\int_{(0,1]} x^{-2} \Lambda(dx) = \infty$. Note that this situation is equivalent to $\lim_{N\to\infty} \lambda_N = \infty$. Then, by (7), $\mathbb{E}(\mu_1^2)/N^2 \to 0$. Thus (see also Lemma 3.2), the conditions of Theorem 3.3 are satisfied with $U_N := \mu_1(N)$, $N \in \mathbb{N}$. In other words, we define an extended Moran model for which the number of offspring U_N coincides pathwise with the size $\mu_1(N)$ of the largest block of the state of the restricted Λ-coalescent $(\varrho_N \circ \Pi_t)_{t\geq 0}$ after its first jump. By (8), the limiting continuous-time coalescent arising in Theorem 3.3 has characterizing measure Λ. Note that U_N has distribution

$$\begin{aligned} \mathbb{P}(U_N = j) \;&=\; \mathbb{P}(|\varrho_N \circ \Pi_{T_N}| = N - j + 1) \;=\; \binom{N}{j} \frac{\lambda_{Nj}}{\lambda_N} \\ &=\; \frac{1}{\lambda_N} \binom{N}{j} \mathbb{E}(X^{j-2}(1-X)^{N-j}), \quad j \in \{2, \dots, N\}. \end{aligned} \tag{9}$$

For further properties of U_N we refer the reader to the following remark and the examples thereafter.

Case 2. Suppose that $\Lambda(\{0\}) = 0$ and that $\int_{(0,1]} x^{-2} \Lambda(dx) < \infty$, or, equivalently, that $\lim_{N\to\infty} \lambda_N < \infty$. Then, instead of $\mu_1(N)$, consider the modified random variables $U_N := \mu_1(N)\, 1_{A_N} + 1_{A_N^c}$, $N \in \mathbb{N}$, where $A_1, A_2, \dots$ are events independent of $(\mu_1(N))_{N\in\mathbb{N}}$ with $\mathbb{P}(A_N) \to 0$ as $N \to \infty$. Then, $\mathbb{E}((U_N)_2)/(N)_2 = \mathbb{E}((\mu_1)_2)/(N)_2 \mathbb{P}(A_N) \leq \mathbb{P}(A_N) \to 0$. Moreover, from $\mathbb{E}((U_N)_k) = \mathbb{E}((\mu_1)_k)\mathbb{P}(A_N)$ for all $k \in \{2, 3, \dots\}$ it follows that

$$\frac{(N)_2 \mathbb{E}((U_N)_k)}{(N)_k \mathbb{E}((U_N)_2)} \;=\; \frac{(N)_2 \mathbb{E}((\mu_1)_k)}{(N)_k \mathbb{E}((\mu_1)_2)}.$$

Again, by (8), the limiting continuous-time coalescent arising in Theorem 3.3 has characterizing measure Λ. □

Remark. (Further properties of U_N) By (9), the random variable $U_N := \mu_1(N)$ constructed in Case 1 of the previous proof has mean

$$\begin{aligned} \mathbb{E}(U_N) \;&=\; \sum_{j=2}^{N} j\, \mathbb{P}(U_N = j) \;=\; \frac{1}{\lambda_N} \mathbb{E}\bigg(\sum_{j=2}^{N} j \binom{N}{j} X^{j-2}(1-X)^{N-j} \bigg) \\ &=\; \frac{(N)_2}{\lambda_N} \mathbb{P}(X = 0) + \frac{N}{\lambda_N} \mathbb{E}\bigg(\frac{1 - X - (1-X)^N}{X(1-X)}\, 1_{\{0<X<1\}} \bigg) + \frac{N}{\lambda_N} \mathbb{P}(X = 1). \end{aligned}$$

By (7), the higher factorial moments of U_N are

$$\mathbb{E}((U_N)_k) \;=\; \frac{(N)_k}{\lambda_N}\mathbb{E}(X^{k-2}), \qquad k \in \{2,\dots,N\}. \tag{10}$$

Alternatively, (10) can be also deduced directly from (9). From (10) it follows in particular that the model has coalescence probability $c_N = \mathbb{E}((U_N)_2)/(N)_2 = 1/\lambda_N$, so the effective population size $N_e = 1/c_N = \lambda_N$ coincides with the total rate λ_N. Since U_N is the number of particles involved in the first collision event of the restricted coalescent $(\varrho_N \circ \Pi_t)_{t\geq 0}$, many properties of Π can be expressed in terms of U_N. For example, the rate at which the number of blocks is decreasing (see [21, Eq. (2)]) is $\gamma_N = \lambda_N \mathbb{E}(U_N - 1)$. Provided that Λ has no atom at 1, by [21, Theorem 1], Π comes down from infinity if and only if $\sum_{N=2}^{\infty} 1/(\lambda_N \mathbb{E}(U_N - 1)) < \infty$.

Examples. 1. (Kingman coalescent) Obviously, if $\Lambda = \delta_0$ is the Dirac measure at 0, then the construction provided in the proof of Proposition 3.4 leads to $U_N \equiv 2$, the standard Moran model, which is well known to be in the domain of attraction of the Kingman coalescent.

2. (Bolthausen–Sznitman coalescent) If Λ is the uniform distribution on $[0,1]$, then $\lambda_{Nj} = B(j-1, N-j+1)$ (beta function) and $\lambda_N = N-1$, leading to $\mathbb{P}(U_N = j) = N/((N-1)j(j-1))$, $j \in \{2,\dots,N\}$, $\mathbb{E}(U_N) = Nh_{N-1}/(N-1) = \log N + \gamma + O((\log N)/N)$, where $h_N := \sum_{i=1}^{N} 1/i$ denotes the Nth harmonic number and γ denotes the Euler constant, and $\mathbb{E}((U_N)_k) = (N)_k/((N-1)(k-1))$ for $k \geq 2$. Note that $U_N \to U$ in distribution as $N \to \infty$, where U has distribution $\mathbb{P}(U = j) = 1/(j(j-1))$, $j \in \{2,3,\dots\}$. The associated extended Moran model has coalescence probability $c_N = \mathbb{E}((U_N)_2)/(N)_2 = 1/(N-1)$ and, by construction, this model is in the domain of attraction of the Bolthausen–Sznitman coalescent [3].

3. (beta coalescent) Fix $a, b \in (0,\infty)$ and suppose that Λ has beta density $x \mapsto (B(a,b))^{-1} x^{a-1}(1-x)^{b-1}$, $x \in (0,1)$, with respect to the Lebesgue measure on $(0,1)$. Obviously, $\lambda_{Nj} = (B(a,b))^{-1} \int_0^1 x^{j+a-3}(1-x)^{N-j+b-1}\,dx = B(j+a-2, N-j+b)/B(a,b)$, $2 \leq j \leq N$. The distribution (9) takes the form

$$\mathbb{P}(U_N = j) \;=\; \frac{1}{\lambda_N}\binom{N}{j}\frac{B(a+j-2, b+N-j)}{B(a,b)}, \qquad j \in \{2,\dots,N\}. \tag{11}$$

Formulas for the total rates λ_N, $N \in \mathbb{N}$, are provided in the appendix (Section 5). In order to determine the behavior of the distribution (11) for large N three cases need to be distinguished.

Case (i): If $a < 2$, then it follows from (11) and Corollary 5.1 that $U_N \to U$ in distribution, where U is a random variable taking the value $j \in \{2,3,\dots\}$ with probability

$$\mathbb{P}(U = j) \;=\; \frac{(2-a)\Gamma(a+j-2)}{\Gamma(a)\Gamma(j+1)}.$$

The distribution of U (shifted by 1) occurs in a similar context in [10, p. 225]. Using (17), it is easily seen that U is heavy tailed with $\mathbb{P}(U > j) = \Gamma(a+j-1)/(\Gamma(a)\Gamma(j+1))$, $j \in \mathbb{N}$, and $\mathbb{P}(U > j) \sim 1/(j^{2-a}\Gamma(a))$ as $j \to \infty$. Note that $\mathbb{E}(U^\eta) = \infty$ for all $\eta \geq 2-a$. It is remarkable that the distribution of U does not depend on the parameter b.

Case (ii): If $a > 2$, then, by Corollary 5.1, $\lambda_N \to (a+b-2)(a+b-1)/((a-2)(a-1))$. From (11) it follows that $U_N/N \to U$ in distribution as $N \to \infty$, where U is beta distributed with parameters $a-2$ and b.

Case (iii): If $a = 2$, then, by the remark after the proof of Proposition 3.4,

$$\begin{aligned}
\mathbb{E}(U_N) &= \frac{N}{\lambda_N}\mathbb{E}\Big(\frac{1-X-(1-X)^N}{X(1-X)}\Big) \\
&= \frac{N}{\lambda_N}\frac{1}{B(2,b)}\int_0^1 \frac{1-x-(1-x)^N}{x(1-x)} x(1-x)^{b-1}\,dx \\
&= \frac{N}{\lambda_N}\frac{1}{B(2,b)}\int_0^1 ((1-x)^{b-1}-(1-x)^{N+b-2})\,dx \\
&= \frac{N}{\lambda_N}\frac{1}{B(2,b)}\Big(\frac{1}{b}-\frac{1}{N+b-1}\Big) = \frac{N}{\lambda_N}\frac{(b+1)(N-1)}{N+b-1}
\end{aligned}$$

and, for $k \geq 2$,

$$\mathbb{E}((U_N)_k) = \frac{(N)_k}{\lambda_N}\mathbb{E}(X^{k-2}) = \frac{(N)_k}{\lambda_N}\frac{B(k,b)}{B(2,b)}.$$

By Corollary 5.1, $\lambda_N \sim b(b+1)\log N$ as $N \to \infty$. Thus, $\mathbb{E}((U_N)_k) \sim B(k,b)N^k/\log N$ for all $k \in \mathbb{N}$. In particular, $\mathbb{E}(U_N) \sim N/(b\log N)$ and $\mathrm{Var}(U_N) \sim \mathbb{E}((U_N)_2) = (N)_2/\lambda_N \sim N^2/(b(b+1)\log N)$ as $N \to \infty$. This asymptotics of the mean and the variance of U_N however does not provide direct information on the limiting behavior of U_N as $N \to \infty$. Clearly, $\mathbb{E}((U_N)_2)/(N)_2 = 1/\lambda_N \to 0$ as $N \to \infty$, or, equivalently (see Lemma 3.2), $U_N/N \to 0$ in distribution as $N \to \infty$. Based on the formula

$$\begin{aligned}
\mathbb{P}(U_N \leq k) &= \sum_{j=2}^k \mathbb{P}(U_N = j) = \frac{1}{\lambda_N B(2,b)}\sum_{j=2}^k \binom{N}{j} B(j, N-j+b) \\
&= \frac{1}{\lambda_N B(2,b)}\sum_{j=2}^k \frac{1}{j}\frac{\Gamma(N+1)\Gamma(N-j+b)}{\Gamma(N+b)\Gamma(N-j+1)} \qquad k \in \{2,\ldots,N\},
\end{aligned}$$

for the distribution function of U_N, it follows that $\mathbb{P}(\log U_N/\log N \leq x) = \mathbb{P}(U_N \leq N^x) \to x$ as $N \to \infty$ for all $x \in [0,1]$. Thus,

$$\frac{\log U_N}{\log N} \to U$$

in distribution as $N \to \infty$, where U is uniformly distributed on the unit interval. Note that the latter convergence does not depend on the parameter $b > 0$. Two

additional convergence results for U_N are now provided for the situation when $a = 2$.

Corollary 3.5 *If both N and b tend to infinity such that $b/N \to \varrho$ for some constant $\varrho \in (0, \infty)$, then $U_N \to L$ in distribution as $N \to \infty$, where L has distribution $\mathbb{P}(L = j) = c_\varrho^{-1}(1+\varrho)^{-j}/j$, $j \in \{2, 3, \ldots\}$ with normalizing constant $c_\varrho := \log(1 + 1/\varrho) - 1/(1 + \varrho)$, so L has a logarithmic distribution conditioned on being larger than* 1.

Proof. Using Stirling's formula $\Gamma(x-1) \sim (x/e)^x\sqrt{2\pi x}$ as $x \to \infty$, it is readily checked that

$$\mathbb{P}(U_N = j) \;=\; \frac{1}{B(2,b)\lambda_N}\binom{N}{j}B(j, b+N-j) \;\to\; c_\varrho^{-1}\,\frac{(1+\varrho)^{-j}}{j}$$

for all $j \in \{2, 3, \ldots\}$, with $c_\varrho := \lim_{N\to\infty} B(2,b)\lambda_N = \log(1 + 1/\varrho) - 1/(1 + \varrho)$, where the last equality follows from (20) in the appendix. Thus, $U_N \to L$ in distribution. □

For fixed parameter b we may also look at the thinned random variable $V_N := \alpha_N \odot U_N$ (discrete scaling of U_N by α_N) with, for example, $\alpha_N \sim \alpha/N$ for some constant $\alpha \in (0, \infty)$. Note that $\mathbb{E}(V_N) = \alpha_N\mathbb{E}(U_N) \sim \alpha/(b \log N) \to 0$ as $N \to \infty$. Therefore, $V_N \to 0$ in distribution as $N \to \infty$, which is as well intuitively clear since $\alpha_N \sim \alpha/N$ is small for large N, so the binomial thinning $V_N = \alpha_N \odot U_N$ causes V_N to have high mass at 0. The following lemma shows that conditioning V_N to be not equal to 0 leads to a proper discrete limiting distribution.

Corollary 3.6 *Let $a = 2$, fix $\alpha \in (0, \infty)$ and let $\alpha_1, \alpha_2, \ldots \in (0, 1)$ such that $N\alpha_N \to \alpha$ as $N \to \infty$. Then, the random variable $V_N := \alpha_N \odot U_N$, conditional on $V_N > 0$, converges in distribution as $N \to \infty$ to $q := (q_1, q_2, \ldots)$, where $q_j := r_j/r$ with $r_j := (\alpha^j/j!)\sum_{k=0}^{\infty}(-\alpha)^k B(j + k, b)/k!$ for $j \in \mathbb{N}$ and $r := \sum_{j=1}^{\infty} r_j = -\sum_{k=1}^{\infty}(-\alpha)^k B(k, b)/k!$.*

Proof. Conditional on U_N, V_N has a binomial distribution with parameters U_N and α_N. Therefore, for all $j \in \mathbb{N}_0$,

$$\begin{aligned}
\mathbb{P}(V_N = j) &= \mathbb{E}\left(\binom{U_N}{j}\alpha_N^j(1-\alpha_N)^{U_N-j}\right)\\
&= \mathbb{E}\left(\binom{U_N}{j}\alpha_N^j\sum_{k=0}^{U_N-j}\binom{U_N-j}{k}(-\alpha_N)^k\right)\\
&= \frac{\alpha_N^j}{j!}\sum_{k=0}^{\infty}\frac{(-\alpha_N)^k}{k!}\mathbb{E}((U_N)_{j+k}). \qquad (12)
\end{aligned}$$

Note that the last sum is finite since $\mathbb{E}((U_N)_{j+k}) = 0$ for $k > N - j$. Since $\alpha_N \sim \alpha/N$ and $\mathbb{E}((U_N)_k) \sim B(k, b)N^k/\log N$ as $N \to \infty$ for all $k \in \mathbb{N}$, it

follows from (12) that

$$(\log N)(1-\mathbb{P}(V_N=0)) \;\to\; -\sum_{k=1}^{\infty}\frac{(-\alpha)^k}{k!}B(k,b) \;=\; r$$

and that

$$(\log N)\mathbb{P}(V_N=j) \;\to\; \frac{\alpha^j}{j!}\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!}B(j+k,b) \;=\; r_j, \qquad j\in\mathbb{N}.$$

Note that

$$\begin{aligned} r &= -\sum_{k=1}^{\infty}\frac{(-\alpha)^k}{k!}\int_0^1 x^{k-1}(1-x)^{b-1}\,dx \\ &= \int_0^1 \frac{1-e^{-\alpha x}}{x}(1-x)^{b-1}\,dx \;\in(0,\infty) \end{aligned}$$

and that

$$\begin{aligned} r_j &= \frac{\alpha^j}{j!}\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!}B(j+k,b) \;=\; \frac{\alpha^j}{j!}\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!}\int_0^1 x^{j+k-1}(1-x)^{b-1}\,dx \\ &= \frac{\alpha^j}{j!}\int_0^1 x^{j-1}(1-x)^{b-1}e^{-\alpha x}\,dx \;\in\; (0,\infty), \qquad j\in\mathbb{N}. \end{aligned}$$

Moreover,

$$\begin{aligned} \sum_{j=1}^{\infty} r_j &= \sum_{j=1}^{\infty}\frac{\alpha^j}{j!}\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!}B(j+k,b) \;=\; \sum_{j=1}^{\infty}\frac{\alpha^j}{j!}\sum_{l=j}^{\infty}\frac{(-\alpha)^{l-j}}{(l-j)!}B(l,b) \\ &= \sum_{l=1}^{\infty}\frac{\alpha^l}{l!}B(l,b)\sum_{j=1}^{l}\binom{l}{j}(-1)^{l-j} \;=\; \sum_{l=1}^{\infty}\frac{\alpha^l}{l!}B(l,b)(-1)^{l+1} \;=\; r. \end{aligned}$$

In particular, $\mathbb{P}(V_N=j\,|\,V_N>0)=\mathbb{P}(V_N=j)/\mathbb{P}(V_N>0)\to q_j:=r_j/r$ for all $j\in\mathbb{N}$. Therefore, V_N, conditional on $V_N>0$, converges in distribution to $q:=(q_1,q_2,\ldots)$. □

Remark. An alternative proof of Proposition 3.4 is presented which shows in particular that the sequence $(U_N)_{N\in\mathbb{N}}$ in Proposition 3.4 is far from being unique. Fix the probability measure Λ and suppose first that $a:=\Lambda(\{0\})=0$. Define $\nu(dx):=x^{-2}\Lambda(dx)$ and choose a sequence $(p_N)_{N\in\mathbb{N}}$ of positive real numbers satisfying $1\geq p_1\geq p_2\geq\cdots$ and $\lim_{N\to\infty}p_N=0$. From $\nu((p_N,1])\to\nu((0,1])>0$ it follows that there exists a constant $N_0\in\mathbb{N}$ such that $\nu((p_N,1])>0$ for all $N>N_0$. Without loss of generality we will construct the random variable U_N only for all $N>N_0$. For $N\leq N_0$ simply choose some arbitrary random variable taking values in $\{1,\ldots,N\}$. Define the sequence $(\nu_N)_{N>N_0}$ of probability measures via

$$\nu_N(B) \;:=\; \frac{\nu([p_N,1]\cap B)}{\nu([p_N,1])} \tag{13}$$

for any Borel set $B \subseteq [p_N, 1]$. Let V_N be a random variable with distribution ν_N. Note that V_N takes values in $[p_N, 1]$ almost surely. The binomially thinned random variable $S_N := V_N \odot N$ (compare with Example 2 after Proposition 3.1) has distribution $\mathbb{P}(S_N = j) = \binom{N}{j}\mathbb{E}(V_N^j(1-V_N)^{N-j})$, $j \in \{0, \ldots, N\}$, and factorial moments $\mathbb{E}((S_N)_k) = (N)_k\mathbb{E}(V_N^k)$, $k \in \mathbb{N}_0$. Note that

$$\begin{aligned} q_N &:= \mathbb{P}(S_N \geq 2) = 1 - \mathbb{P}(S_N = 0) - \mathbb{P}(S_N = 1) \\ &= 1 - \mathbb{E}((1-V_N)^N) - N\mathbb{E}(V_N(1-V_N)^{N-1}) \\ &= \frac{1}{\nu([p_N,1])}\int_{[p_N,1]} (1-(1-x)^N - Nx(1-x)^{N-1})\,\nu(dx) \;>\; 0 \end{aligned}$$

for all $N > N_0$. Now let U_N be S_N conditioned on the event that $S_N \geq 2$. Clearly, U_N has distribution $\mathbb{P}(U_N = j) = \mathbb{P}(S_N = j)/q_N$, $j \in \{2, \ldots, N\}$, mean $\mathbb{E}(U_N) = (\mathbb{E}(S_N) - \mathbb{P}(S_N = 1))/q_N = N\mathbb{E}(V_N(1-(1-V_N)^{N-1}))/q_N$, and higher factorial moments $\mathbb{E}((U_N)_k) = \mathbb{E}((S_N)_k)/q_N = (N)_k\mathbb{E}(V_N^k)/q_N$, $k \in \{2, \ldots, N\}$. For all $k \in \{2, 3, \ldots\}$ and all $N \geq \max(N_0 + 1, k)$ it follows that

$$\frac{(N)_2\mathbb{E}((U_N)_k)}{(N)_k\mathbb{E}((U_N)_2)} \;=\; \frac{\mathbb{E}(V_N^k)}{\mathbb{E}(V_N^2)} \;=\; \frac{\int_{[p_N,1]} x^k\,\nu(dx)}{\int_{[p_N,1]} x^2\,\nu(dx)} \;\to\; \int_{[0,1]} x^{k-2}\,\Lambda(dx)$$

as $N \to \infty$. Moreover, the coalescence probability is given by

$$\begin{aligned} c_N &= \frac{\mathbb{E}((U_N)_2)}{(N)_2} \;=\; \frac{\mathbb{E}(V_N^2)}{q_N} \\ &= \frac{1}{q_N\nu([p_N,1])}\int_{[p_N,1]} x^2\,\nu(dx) \;=\; \frac{1}{\theta_N}\int_{[p_N,1]} \Lambda(dx), \end{aligned} \tag{14}$$

where

$$\begin{aligned} \theta_N &:= \int_{[p_N,1]} (1-(1-x)^N - Nx(1-x)^{N-1})\,\nu(dx) \\ &= \int_{(0,1]} 1_{[p_N,1]}(x)\frac{1-(1-x)^N - Nx(1-x)^{N-1}}{x^2}\,\Lambda(dx). \end{aligned}$$

For arbitrary but fixed $x \in (0, 1]$ the expression below the last integral is non-decreasing in N with limit $1/x^2$. Thus, by the monotone convergence theorem, $\theta_N \to \int_{(0,1]} x^{-2}\Lambda(dx)$, which can be finite or infinite.
If $\int_{(0,1]} x^{-2}\Lambda(dx) = \infty$, then, $\theta_N \to \infty$, and, by (14), $c_N \to 0$. Thus, the sequence $(U_N)_{N\in\mathbb{N}}$ satisfies the conditions of Theorem 3.3.
If Λ has positive mass $a := \Lambda(\{0\}) \in (0, 1]$ at zero, then decompose $\Lambda = a\delta_0 + (1-a)\Lambda_0$, where Λ_0 is a probability measure having no mass at zero, and replace ν by $\nu(dx) := x^{-2}\Lambda_0(dx)$ and the measure ν_N in (13) by

$$\nu_N(B) \;:=\; \frac{(1-a)\nu([p_N,1]\cap B) + ap_N^{-2}\delta_{p_N}(B)}{(1-a)\nu([p_N,1]) + ap_N^{-2}}$$

for any Borel set $B \subseteq [p_N, 1]$. Then everything works in the same manner leading to

$$\begin{aligned}\theta_N &= (1-a)\int_{[p_N,1]} (1-(1-x)^N - Nx(1-x)^{N-1})\,\nu(dx) \\ &\quad + a\frac{1-(1-p_N)^N - Np_N(1-p_N)^{N-1}}{p_N^2},\end{aligned}$$

which diverges due to the fraction occurring after the constant a, for example whenever $p_N = O(1/N)$. Thus, the construction of a sequence $(U_N)_{N\in\mathbb{N}}$ is established whenever $\Lambda(\{0\}) > 0$ or $\int_{(0,1]} x^{-2}\Lambda(dx) = \infty$. Otherwise proceed as in Case 2 of the proof of Proposition 3.4.

Examples. 1. (Kingman coalescent) If $\Lambda = \delta_0$ is the Dirac measure at 0, then we can, for example, choose any constant $c \in (0, 1]$ and work with the sequence $p_N := c/N$. Obviously, $V_N \equiv p_N$, and S_N is binomially distributed with parameters N and p_N. The distribution of U_N is therefore given by $\mathbb{P}(U_N = j) = \binom{N}{j} p_N^j (1-p_N)^{N-j}/q_N$, $j \in \{2, \ldots, N\}$, with $q_N = 1-(1-p_N)^N - Np_N(1-p_N)^{N-1} \sim 1-(c+1)e^{-c} \in (0,1)$. The corresponding extended Moran model has coalescence probability $c_N = \mathbb{E}(V_N^2)/q_N = p_N^2/q_N \sim c^2/(1-(c+1)e^{-c})N^{-2}$, and the model is in the domain of attraction of the Kingman coalescent.
2. (Bolthausen–Sznitman coalescent) If Λ is uniformly distributed on the unit interval $[0, 1]$, then it turns out that we can work with any sequence $(p_N)_{N\in\mathbb{N}}$ of positive real numbers satisfying $1 > p_1 \geq p_2 \geq \cdots$ and $\lim_{N\to\infty} p_N = 0$. It is readily checked that V_N has density $x \mapsto p_N/((1-p_N)x^2)$, $x \in [p_N, 1]$, and, hence, mean $\mathbb{E}(V_N) = p_N(-\log p_N)/(1-p_N)$ and higher moments

$$\mathbb{E}(V_N^k) = \frac{p_N(1-p_N^{k-1})}{(1-p_N)(k-1)}, \qquad k \in \{2, 3, \ldots\}.$$

In particular, $\mathbb{E}(V_N^2) = p_N$. From

$$1-(1-x)^N - Nx(1-x)^{N-1} = \sum_{k=2}^{N} (-1)^k \binom{N}{k} (k-1)x^k$$

it follows that

$$\begin{aligned}\theta_N &= \int_{p_N}^1 \frac{1-(1-x)^N - Nx(1-x)^{N-1}}{x^2}\,dx \\ &= \sum_{k=2}^{N} (-1)^k \binom{N}{k} (k-1)\int_{p_N}^1 x^{k-2}\,dx \\ &= \sum_{k=2}^{N} (-1)^k \binom{N}{k} (1-p_N^{k-1}) = \frac{1-p_N}{p_N}(1-(1-p_N)^{N-1}),\end{aligned}$$

and $q_N = \theta_N/\nu([p_N,1]) = 1-(1-p_N)^{N-1}$, $N \in \mathbb{N}$. For $N \geq 2$, the corresponding extended Moran model has coalescence probability

$$c_N \;=\; \frac{\mathbb{E}(V_N^2)}{q_N} \;=\; \frac{p_N}{1-(1-p_N)^{N-1}} \;\leq\; \frac{N-2}{N-1}p_N + \frac{1}{N-1} \;\to\; 0,$$

since $p_N \to 0$ as $N \to \infty$. Since V_N has density $x \mapsto p_N/((1-p_N)x^2)$, $x \in [p_N,1]$, it follows that U_N has distribution

$$\mathbb{P}(U_N = j) \;=\; \frac{\binom{N}{j}}{q_N}\mathbb{E}(V_N^j(1-V_N)^{N-j}) \;=\; \frac{\binom{N}{j}}{q_N}\frac{p_N}{1-p_N}\int_{p_N}^1 x^{j-2}(1-x)^{N-j}\,dx,$$

$j \in \{2,\ldots,N\}$. The distribution of U_N is therefore related to the incomplete beta function and does not seem to simplify further. By construction, this extended Moran model is in the domain of attraction of the Bolthausen–Sznitman coalescent.

4 Further examples

Examples of extended Moran models are now studied which do not coincide with the constructions provided in Section 3, but nevertheless satisfy the assumptions of Theorem 3.3.

Example 4.1 Let $A_1, A_2, \ldots$ be events satisfying $\lim_{N\to\infty}\mathbb{P}(A_N) = 0$ or, equivalently, $1_{A_N} \to 0$ in distribution as $N \to \infty$. Let furthermore $V_1, V_2, \ldots$ be random variables independent of $(A_N)_{N\in\mathbb{N}}$ taking values in $[0,1]$. Suppose that, for $N \in \mathbb{N}$, $U_N = \lfloor NV_N \rfloor 1_{A_N} + 21_{A_N^c}$, so with probability $\mathbb{P}(A_N)$ extreme reproduction events occur whereas with complementary probability $\mathbb{P}(A_N^c)$, the population evolves according to the standard Moran model. From $U_N \leq N1_{A_N} + 2$ it follows that $U_N/N \leq 1_{A_N} + 2/N \to 0$ in distribution as $N \to \infty$, indicating that this model basically belongs to the situation covered by Theorem 3.3. We have

$$\mathbb{E}\Big(\Big(\frac{U_N}{N}\Big)^k\Big) \;=\; \mathbb{E}\Big(\Big(\frac{\lfloor NV_N \rfloor}{N}\Big)^k\Big)\mathbb{P}(A_N) + \Big(\frac{2}{N}\Big)^k\mathbb{P}(A_N^c), \qquad k \in \mathbb{N}_0.$$

Suppose now that $V_N = V$ does not depend on N. Then $\mathbb{E}((\lfloor NV \rfloor/N)^k) = \mathbb{E}(V^k) + O(N^{-1})$, and it follows that

$$\mathbb{E}\Big(\Big(\frac{U_N}{N}\Big)^k\Big) \;\sim\; \mathbb{E}(V^k)\mathbb{P}(A_N) + \Big(\frac{2}{N}\Big)^k\mathbb{P}(A_N^c), \qquad k \in \mathbb{N}_0. \tag{15}$$

If, for example, $\mathbb{P}(A_N) = N^{-\gamma}$ for some constant $\gamma \in (0,\infty)$, then

$$\mathbb{E}\Big(\Big(\frac{U_N}{N}\Big)^k\Big) \;\sim\; \begin{cases} \mathbb{E}(V^k)N^{-\gamma} & \text{if } \gamma < k, \\ \mathbb{E}(V^k)N^{-\gamma} + (2/N)^k & \text{if } \gamma = k, \\ (2/N)^k & \text{if } \gamma > k. \end{cases} \tag{16}$$

Three cases are distinguished.

Case 1. If $0 < \gamma < 2$ then, $\mathbb{E}(U_N^2) \sim \mathbb{E}(V^2)N^{2-\gamma} \to \infty$ as $N \to \infty$. By (16),

$$\phi_1(k) \;=\; \lim_{N\to\infty} \frac{\mathbb{E}(U_N^k)}{N^{k-2}\mathbb{E}((U_N)_2)} \;=\; \lim_{N\to\infty} \frac{\mathbb{E}((U_N/N)^k)}{\mathbb{E}((U_N/N)^2)} \;=\; \frac{\mathbb{E}(V^k)}{\mathbb{E}(V^2)}$$

for all $k \in \{2, 3, \ldots\}$. By Theorem 3.3, the extended Moran model is in the domain of attraction of a continuous-time Λ-coalescent with $\Lambda(dx) := x^2\nu(dx)$ and $\nu(dx) := P_V(dx)/\mathbb{E}(V^2)$, where P_V denotes the distribution of V.

Case 2. If $\gamma = 2$, then, by (16), $\mathbb{E}(U_N) \to 2$, $\mathbb{E}(U_N^2) \to \mathbb{E}(V^2) + 4$ and, hence, $\mathbb{E}((U_N)_2) \to \mathbb{E}(V^2) + 2$. Moreover, by (16), for $k \in \{3, 4, \ldots\}$, $\mathbb{E}(U_N^k) \sim N^{k-2}\mathbb{E}(V^k)$ and for $k \in \{3, 4, \ldots\}$ it follows that

$$\phi_1(k) \;=\; \lim_{N\to\infty} \frac{\mathbb{E}(U_N^k)}{N^{k-2}\mathbb{E}((U_N)_2)} \;=\; \frac{\mathbb{E}(V^k)}{\mathbb{E}(V^2)+2}.$$

By Theorem 3.3, the model is in the domain of attraction of a continuous-time Λ-coalescent with

$$\Lambda(dx) \;:=\; \frac{x^2 P_V(dx) + 2\delta_0(dx)}{\mathbb{E}(V^2)+2}.$$

Case 3. If $\gamma > 2$, then it follows from (16) that $\mathbb{E}(U_N) \to 2$ and $\mathbb{E}(U_N^2) \to 4$. Thus, $\mathbb{E}((U_N)_2) \to 2$. Moreover, $\mathbb{E}(U_N^3) \sim \mathbb{E}(V^3)N^{3-\gamma}$ if $2 < \gamma < 3$, $\mathbb{E}(U_N^3) \to \mathbb{E}(V^3) + 8$ for $\gamma = 3$, and $\mathbb{E}(U_N^3) \to 8$, if $\gamma > 3$. It follows that

$$\phi_1(3) \;=\; \lim_{N\to\infty} \frac{\mathbb{E}(U_N^3)}{N\mathbb{E}((U_N)_2)} \;=\; \frac{1}{2}\lim_{N\to\infty} \frac{\mathbb{E}(U_N^3)}{N} \;=\; 0.$$

The model is hence in the domain of attraction of the Kingman coalescent.

Remarks. 1. Eldon and Wakeley [6, Eq. (7)] study the particular model in which $\mathbb{P}(A_N) = N^{-\gamma}$ for some constant $\gamma \in (1, \infty)$ and $V \equiv \psi$ is equal to a given constant $\psi \in [0, 1]$.
2. The asymptotics of the model in the previous example depends on how fast the probabilities $\mathbb{P}(A_N)$ tend to zero. For example, if $\mathbb{P}(A_N) = \lambda^{-N}$ for some constant $\lambda \in (0, \infty)$, then, by (15), $\mathbb{E}((U_N/N)^k) \sim (2/N)^k$ for all $k \in \mathbb{N}$. Thus, $\phi_1(3) = 0$ and the model is in the domain of attraction of the Kingman coalescent. Intuitively, since $\mathbb{P}(A_N)$ tends to zero very fast, extreme reproduction events are very rare. For large N the standard Moran model dominates leading to the Kingman coalescent. In contrast, if $\mathbb{P}(A_N)$ tends slowly to 0, say $\mathbb{P}(A_N) = (\log N)^{-\beta}$ for some constant $\beta > 0$, then, by (15), $\mathbb{E}((U_N/N)^k) \sim \mathbb{E}(V^k)(\log N)^{-\beta}$ for all $k \in \mathbb{N}$. We are hence essentially in the same situation as in Case 1 of the previous example. The model is in the domain of attraction of a continuous-time Λ-coalescent with Λ defined as in Case 1 of the previous example.

Example 4.2 Let $V_1, V_2, \ldots$ be random variables taking values in $[0, 1]$ satisfying $V_N \to 0$ in distribution as $N \to \infty$ and suppose that $U_N = V_N \odot N$, $N \in \mathbb{N}$.

Then, $\mathbb{E}((U_N)_k) = (N)_k\mathbb{E}(V_N^k)$, $k \in \mathbb{N}_0$. In particular, $\mathbb{E}(U_N/N) = \mathbb{E}(V_N) \to 0$ as $N \to \infty$, or, equivalently, $U_N/N \to 0$ in distribution as $N \to \infty$. We are therefore in principle in the situation of Theorem 3.3. We now discuss some particular choices for V_N.

(i) Let $(p_N)_{N\in\mathbb{N}}$ be a sequence of real numbers satisfying $0 < p_N \leq 1$ for all $N \in \mathbb{N}$ and $p_N \to 0$. If $V_N \equiv p_N$ for all $N \in \mathbb{N}$, then $\mathbb{E}((U_N)_3)/(N\mathbb{E}((U_N)_2)) \sim (N)_3p_N^3/(N(N)_2p_N^2) \sim p_N \to 0$. Thus, $\phi_1(3) = 0$ and, by the remark after Theorem 3.3, the model is in the domain of attraction of the Kingman coalescent.

(ii) Suppose that V_N is beta distributed with parameters $a_N, b_N > 0$. Then

$$\mathbb{E}(V_N^k) \;=\; \frac{\Gamma(a_N + k)\Gamma(a_N + b_N)}{\Gamma(a_N)\Gamma(a_N + b_N + k)}, \qquad k \in \mathbb{N}_0.$$

In particular, $\mathbb{E}(V_N) = a_N/(a_N+b_N) = 1/(1+b_N/a_N) \to 0$ whenever $b_N/a_N \to \infty$. Suppose now in addition that $a_N, b_N \to 0$. Then, for $k \in \{2, 3, \ldots\}$,

$$\frac{\mathbb{E}((U_N)_k)}{N^{k-2}\mathbb{E}((U_N)_2)} \;\sim\; \frac{\mathbb{E}(V_N^k)}{\mathbb{E}(V_N^2)} \;=\; \prod_{i=3}^{k} \frac{a_N + i - 1}{a_N + b_N + i - 1} \;\to\; 1.$$

Thus, $\phi_1(k) = 1$ for all $k \in \{2, 3, \ldots\}$. By Theorem 3.3, the model is in the domain of attraction of the star-shaped coalescent, where $\Lambda = \delta_1$ is the Dirac measure at 1.

(iii) Let Z be a proper random variable satisfying $\mathbb{P}(0 \leq Z < \infty) = 1$ and let $\phi(\eta) := \mathbb{E}(e^{-\eta Z})$, $\eta \geq 0$, denote its Laplace transform. Note that $\phi(0) = 1$, that ϕ is completely monotone on $(0, \infty)$, and that $\phi(\eta) = \mathbb{E}(W^\eta)$ is also the moment function of the random variable $W := e^{-Z}$ satisfying $\mathbb{P}(0 < W \leq 1) = 1$.
Let $W_1, W_2, \ldots$ be $(0, 1]$-valued random variables satisfying $W_N \to W$ in distribution as $N \to \infty$, or, equivalently, $\phi_N(\eta) := \mathbb{E}(W_N^\eta) \to \mathbb{E}(W^\eta) = \phi(\eta)$ as $N \to \infty$ for all $\eta \in [0, \infty)$. Let furthermore $(a_N)_{N\in\mathbb{N}}$ be a $(0, 1)$-valued sequence of real numbers converging to 0. For $N \in \mathbb{N}$ let V_N be a $[0, 1]$-valued random variable with moment function $\mathbb{E}(V_N^\eta) := a_N\phi_N(\eta)$, $\eta \in [0, \infty)$. Thus, V_N has distribution $(1 - a_N)\delta_0 + a_NP_{W_N}$, where P_{W_N} denotes the distribution of W_N. In particular, $V_N \to 0$ in distribution as $N \to \infty$. Let X_N be the second order size-biased of V_N. For some more information on size-biasing and its relation to renewal processes we refer the reader to [1] and [17]. We have

$$\mathbb{E}(X_N^\eta) \;=\; \frac{\mathbb{E}(V_N^{\eta+2})}{\mathbb{E}(V_N^2)} \;=\; \frac{\mathbb{E}(W_N^{\eta+2})}{\mathbb{E}(W_N^2)} \;\to\; \frac{\phi(\eta + 2)}{\phi(2)} \;=\; \mathbb{E}(X^\eta),$$

where X denotes the second order size-biased of W.
Let $U_N := V_N \odot N$, $N \in \mathbb{N}$. Clearly, $U_N/N \to 0$ in distribution as $N \to \infty$. Furthermore, $\mathbb{E}((U_N)_k) = (N)_k\mathbb{E}(V_N^k)$. Thus, for all $k \in \{2, 3, \ldots\}$,

$$\frac{\mathbb{E}((U_N)_k)}{N^{k-2}\mathbb{E}((U_N)_2)} \;=\; \frac{(N)_k\mathbb{E}(V_N^k)}{N^{k-2}(N)_2\mathbb{E}(V_N^2)} \;\sim\; \frac{\mathbb{E}(V_N^k)}{\mathbb{E}(V_N^2)} \;\to\; \mathbb{E}(X^{k-2}).$$

By Theorem 3.3, choosing U_N as just described, we are led to a continuous-time Λ-coalescent, with Λ being the distribution of X. Note that $\Lambda(\{0\}) = 0$.

(iv) Consider the previous case (iii) and assume in addition that Z is infinitely divisible, or, equivalently, that $W = e^{-Z}$ is log infinitely divisible. Then, $\phi(\eta) = e^{-\gamma\psi(\eta)}$ for all $\eta \in [0, \infty)$, where $\gamma > 0$ and ψ is some Laplace exponent. Note that $\psi(0) = 0$ and that ψ' is completely monotone on $(0, \infty)$. Let us distinguish two cases.
a) Suppose first that Z is compound Poisson, so W is log compound Poisson. Then $\psi(\eta) = 1 - h(\eta)$, where h is completely monotone with $h(0) = 1$ and $h(\infty) := \lim_{\eta\to\infty} h(\eta) = 0$. Thus, $\phi(\infty) := \lim_{\eta\to\infty} \phi(\eta) = e^{-\gamma}$ meaning that W has an atom at 1 with $\mathbb{P}(W = 1) = e^{-\gamma}$. In this case, we have

$$\mathbb{E}(X^\eta) \;=\; \frac{\phi(\eta+2)}{\phi(2)} \;=\; e^{-\gamma h(2)(1-g(\eta))} \;=\; \int_{(0,1]} x^\eta \, \Lambda(dx),$$

where $g(\eta) := h(\eta + 2)/h(2)$ is completely monotone with $g(0) = 1$ and $g(\infty) = 0$. We therefore obtain a log compound Poisson Λ-coalescent whose moment function of $X \overset{d}{=} \Lambda$ is compound Poisson with rate $\gamma h(2)$ and jump height Laplace transform g. Note that $\Lambda(\{1\}) = e^{-\gamma h(2)} > 0$.
b) Suppose now that Z is infinitely divisible but not in the compound Poisson class (any such Z is classically obtained as a weak limit of a compound Poisson sequence). Then, $\phi(\eta) = e^{-\gamma\psi(\eta)}$ for all $\eta \in [0, \infty)$, where the Laplace exponent ψ satisfies $\psi(\infty) = \infty$. Then W has no more an atom at 1. In this latter case

$$\mathbb{E}(X^\eta) \;=\; \frac{\phi(\eta+2)}{\phi(2)} \;=\; e^{-\gamma\widetilde{\psi}(\eta)} \;=\; \int_{(0,1)} x^\eta \, \Lambda(dx),$$

where $\widetilde{\psi}(\eta) := \psi(\eta + 2) - \psi(2)$ is a new Laplace exponent. Note that $\widetilde{\psi}(0) = 1$, $\widetilde{\psi}(\infty) = \infty$, and that $\widetilde{\psi}'$ is completely monotone. We get a log infinitely divisible Λ-coalescent whose moment function of $X \overset{d}{=} \Lambda$ is log infinitely divisible as before, but not in the compound Poisson class. Note that Λ has no atom at 1. In order to give a concrete example suppose that $X \overset{d}{=} \beta(a, b)$ is beta distributed with parameters $a, b > 0$. It is known [11] that X is log infinitely divisible. If $a > 2$ and $b > 0$, the $\beta(a, b)$-coalescent can be produced as described above provided we choose $W \overset{d}{=} \beta(a - 2, b)$.
In (iii) and (iv) we exhibited sequences U_N satisfying $U_N/N \to 0$ in distribution as $N \to \infty$ as a result of the law of V_N being a mixture of δ_0 and the law of some $(0, 1]$-valued random variable W_N. In (iv) we were led to Λ-coalescents with Λ being log infinitely divisible. Let us now present a special family of log infinitely divisible coalescents for which $U_N = V_N \odot N$ with the property $V_N \to 0$ in distribution as $N \to \infty$, but without V_N having mass at 0.

(v) Fix $\gamma \in (0, \infty)$ and let $(\alpha_N)_{N\in\mathbb{N}}$ be a sequence of positive real numbers satisfying $\alpha_N \to 0$ as $N \to \infty$. Suppose that $-\log V_N$ is gamma distributed with shape parameter γ and scale parameter α_N, so $-\log V_N$ has density $x \mapsto$

$(1/\Gamma(\gamma))x^{\gamma-1}\alpha_N^\gamma e^{-\alpha_N x}$, $x > 0$. Then, $\mathbb{E}(V_N^k) = (1 + k/\alpha_N)^{-\gamma}$, $k \in \mathbb{N}_0$. In particular $\mathbb{E}(V_N) = (1 + 1/\alpha_N)^{-\gamma} \sim \alpha_N^\gamma \to 0$ as $N \to \infty$. For $k \in \{2, 3, \ldots\}$ it follows that

$$\frac{\mathbb{E}((U_N)_k)}{N^{k-2}\mathbb{E}((U_N)_2)} \sim \frac{\mathbb{E}(V_N^k)}{\mathbb{E}(V_N^2)} = \Big(\frac{\alpha_N + 2}{\alpha_N + k}\Big)^\gamma \to \Big(\frac{2}{k}\Big)^\gamma.$$

Thus $\phi_1(k) = (2/k)^\gamma$, $k \in \{2, 3, \ldots\}$. By Theorem 3.3, the model is in the domain of attraction of a continuous-time Λ-coalescent with Λ being a log-gamma distribution with parameters γ and 2 of the form

$$\Lambda(dx) = \frac{2^\gamma}{\Gamma(\gamma)} x(-\log x)^{\gamma-1}\, dx, \qquad x \in (0, 1).$$

This example can be extended as follows. Suppose that V_N has moments $\mathbb{E}(V_N^k) = (1 + \psi(k)/\alpha_N)^{-\gamma}$, $k \in \mathbb{N}_0$, where ψ is the Laplace exponent of some infinitely divisible distribution on $(0, \infty)$. Note that $-\log V_N$ is infinitely divisible, since it is distributed as Z_1, where $(Z_t)_{t\geq 0} := (X_{Y_t})_{t\geq 0}$ is the subordinator obtained by subordinating the subordinator $(X_t)_{t\geq 0}$ with Laplace exponent ψ to the gamma subordinator $(Y_t)_{t\geq 0}$ with parameters γ and α_N. Suppose that $\psi(\eta) > 0$ for all $\eta > 0$. We have $\mathbb{E}(V_N^k) \to 0$ for all $k \in \mathbb{N}$ and, therefore, $V_N \to 0$ in distribution as $N \to \infty$. Moreover, if $\psi'(0+) < \infty$, then $\mathbb{E}(V_N^{\alpha_N \eta}) = (1 + \psi(\alpha_N\eta)/\alpha_N)^{-\gamma} \to (1 + \eta\psi'(0))^{-\gamma}$ as $N \to \infty$ for all $\eta \geq 0$, so $V_N^{\alpha_N} \to V_\infty$ in distribution as $N \to \infty$, where $-\log V_\infty$ is gamma distributed with parameters γ and $1/\psi'(0+)$. The same argument as above shows that the model is in the domain of attraction of a continuous-time Λ-coalescent, where the measure Λ is uniquely determined via its moments $\int_{[0,1]} x^{k-2}\, \Lambda(dx) = (\psi(2)/\psi(k))^\gamma$, $k \in \{2, 3, \ldots\}$. If X is a random variable with distribution Λ, then X has moments $\mathbb{E}(X^k) = (\psi(2)/\psi(k+2))^\gamma = (1+\tilde{\psi}(k))^{-\gamma}$, $k \in \mathbb{N}_0$, where $\tilde{\psi}(\eta) := \psi(\eta + 2)/\psi(2) - 1$, $\eta \in [0, \infty)$. In particular, $-\log X$ is infinitely divisible since the derivative of $\tilde{\psi}$ is completely monotone. For example, if $\psi(\eta) = \eta^\alpha$ for some constant $\alpha \in (0, 1]$, then $\int_{[0,1]} x^{k-2}\, \Lambda(dx) = (2/k)^{\alpha\gamma}$, $k \in \{2, 3, \ldots\}$ and Λ is the log-gamma distribution with parameters $\alpha\gamma$ and 2.

5 Appendix

In this appendix we provide exact and asymptotic formulas for the total rates λ_N, $N \in \mathbb{N}$, for the Λ-coalescent with $\Lambda = \beta(a, b)$ being the beta distribution with arbitrary parameters $a, b > 0$. These formulas provide further insight not only in the structure of beta coalescents in general but also in the extended Moran model constructed in the proof of Proposition 3.4, since this model has effective population size $N_e = \lambda_N$.

Recall that all Λ-coalescents satisfy the consistency relation $\lambda_{k+1} - \lambda_k = k\lambda_{k+1,2}$, $k \in \mathbb{N}$. Therefore, for the $\beta(a, b)$-coalescent, we obtain

$$\lambda_N = \sum_{k=1}^{N-1} (\lambda_{k+1} - \lambda_k) = \sum_{k=1}^{N-1} k\lambda_{k+1,2} = \frac{1}{B(a, b)} \sum_{k=1}^{N-1} kB(a, b + k - 1).$$

For $\alpha > 0$ and $\beta > -1$ with $\alpha \neq \beta$ the formula

$$\sum_{k=1}^{n} \frac{\Gamma(k+\alpha-1)}{\Gamma(k+\beta)} \;=\; \frac{1}{\beta-\alpha}\Big(\frac{\beta\Gamma(\alpha)}{\Gamma(\beta+1)} - \frac{\Gamma(n+\alpha)}{\Gamma(n+\beta)}\Big), \quad n \in \mathbb{N} \tag{17}$$

can be easily verified by induction on $n \in \mathbb{N}$. Applying this formula yields

$$\begin{aligned}
\sum_{k=1}^{N-1} kB(a,b+k-1) &= \sum_{k=1}^{N-1} (b+k-1-(b-1))B(a,b+k-1) \\
&= \Gamma(a)\left(\sum_{k=1}^{N-1} \frac{\Gamma(b+k)}{\Gamma(a+b+k-1)} - (b-1)\sum_{k=1}^{N-1} \frac{\Gamma(b+k-1)}{\Gamma(a+b+k-1)}\right) \\
&= \frac{(a+b-2)(a+b-1)}{(a-2)(a-1)}B(a,b) \\
&\qquad - \frac{(a+b+N-2)(aN+b-N-1)}{(a-2)(a-1)}B(a,b+N-1)
\end{aligned}$$

for $a, b > 0$ with $a \notin \{1,2\}$. Division by $B(a,b)$ yields the general formula for the total rates of the beta coalescent, namely

$$\begin{aligned}
\lambda_N \;=\; & \frac{(a+b-2)(a+b-1)}{(a-2)(a-1)} \\
& - \frac{(a+b+N-2)(aN+b-N-1)}{(a-2)(a-1)} \frac{\Gamma(a+b)\Gamma(b+N-1)}{\Gamma(b)\Gamma(a+b+N-1)},
\end{aligned} \tag{18}$$

provided that $a \notin \{1,2\}$. The particular cases $a \in \{1,2\}$ need to be handled separately. For $a = 1$,

$$\begin{aligned}
\lambda_N &= \frac{1}{B(1,b)} \sum_{k=1}^{N-1} kB(1,b+k-1) \\
&= b\sum_{k=1}^{N-1} \frac{k}{b+k-1} \;=\; b\sum_{k=1}^{N-1} \Big(1 - \frac{b-1}{b+k-1}\Big) \\
&= b(N-1) - b(b-1)\sum_{k=1}^{N-1} \frac{1}{b+k-1} \\
&= b(N-1) - b(b-1)(\Psi(N+b-1) - \Psi(b)),
\end{aligned} \tag{19}$$

where $\Psi(x) := \Gamma'(x)/\Gamma(x)$, $x > 0$. For $a = 2$,

$$\begin{aligned}
\lambda_N &= \frac{1}{B(2,b)} \sum_{k=1}^{N-1} kB(2,b+k-1) \\
&= b(b+1)\sum_{k=1}^{N-1} \frac{k}{(b+k)(b+k-1)} \\
&= b(b+1)\left(\frac{1-N}{N+b-1} + \Psi(N+b-1) - \Psi(b)\right).
\end{aligned} \tag{20}$$

Corollary 5.1 *(Asymptotics of the total rate λ_N) Fix $a, b > 0$. For the $\beta(a,b)$-coalescent, as $N \to \infty$, the total rates λ_N, $N \in \mathbb{N}$, satisfy*

$$\lambda_N \sim \begin{cases} \dfrac{\Gamma(a+b)}{(2-a)\Gamma(b)} N^{2-a} & \text{for } 0 < a < 2, \\ b(b+1)\log N & \text{for } a = 2, \\ \dfrac{(a+b-2)(a+b-1)}{(a-2)(a-1)} & \text{for } 2 < a < \infty. \end{cases} \tag{21}$$

In particular $\lim_{N\to\infty} \lambda_N = \infty$ if and only if $a \leq 2$.

Remark. In particular, for the $\beta(2-\alpha,\alpha)$-coalescent with parameter $\alpha \in (0,2)$, which is the subject of particular studies in the literature, the total rates satisfy $\lambda_N \sim N^\alpha/\Gamma(\alpha+1) \to \infty$ as $N \to \infty$.

Proof. For $a \in \{1,2\}$, (21) follows from (19), (20) and $\Psi(N) \sim \log N$, $N \to \infty$. If $a \notin \{1,2\}$, then apply the formula $\Gamma(N+c) \sim N^c\Gamma(N)$, $N \to \infty$, $c \in \mathbb{R}$, to the two gamma functions having N in their argument on the right hand side in (18). □

Acknowledgement. This work benefited from the support of the ANR MANEGE (ANR-09-BLAN-0215) headed by Sylvie Méléard.

References

[1] Bertoin, J., van Harn, K., and Steutel, F. W. (1999) Renewal theory and level passage by subordinators. *Stat. Probab. Lett.* **45**, 65–69. MR1718352

[2] Birkner, M. and Blath, J. (2008) Computing likelihoods for coalescents with multiple collisions in the infinitely many sites model. *J. Math. Biol.* **57**, 435–465. MR2411228

[3] Bolthausen, E. and Sznitman, A.-S. (1998) On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**, 247–276. MR1652734

[4] Cannings, C. (1974) The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Adv. Appl. Probab.* **6**, 260–290. MR0343949

[5] Cannings, C. (1975) The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. *Adv. Appl. Probab.* **7**, 264–282. MR0371430

[6] Eldon, B. and Wakeley, J. (2006) Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* **172**, 2621–2633.

[7] Gladstien, K. (1978) The characteristic values and vectors for a class of stochastic matrices arising in genetics. *SIAM J. Appl. Math.* **34**, 630–642. MR0475977

[8] Gnedin, A. and Pitman, J. (2008) Moments of convex distribution functions and completely alternating sequences. *Probability and statistics: essays in honor of David A. Freedman, Inst. Math. Stat. Collect.* **2**, 30–41. MR2459948

[9] Huillet, T. and Möhle, M. (2011) Population genetics models with skewed fertilities: a forward and backward analysis. *Stochastic Models* **27**, to appear.

[10] Iksanov, A. and Möhle, M. (2008) On the number of jumps of random walks with a barrier. *Adv. Appl. Probab.* **40**, 206–228. MR2411821

[11] Keilson, J. and Sumita, U. (1983) A decomposition of the beta distribution, related order and asymptotic behavior. *Ann. Inst. Statist. Math.* **35**, 243–253. MR0716035

[12] Kingman, J. F. C. (1982) The coalescent. *Stoch. Process. Appl.* **13**, 235–248. MR0671034

[13] Lessard, S. (2010) Recurrence equations for the probability distribution of sample configurations in exact population genetics models. *J. Appl. Probab.* **47**, 732–751. MR2731345

[14] Möhle, M. (2006) On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli* **12**, 35–53. MR2202319

[15] Möhle, M. and Sagitov, S. (2001) A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, 1547–1562. MR1880231

[16] Moran, P. A. P. (1958) Random processes in genetics. *Proc. Camb. Phil. Soc.* **54**, 60–71. MR0127989

[17] Pakes, A. G. (1997) Characterization by invariance under length-biasing and random scaling. *J. Statist. Plann. Inference* **63**, 285–310. MR1491587

[18] Pitman, J. (1999) Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870–1902. MR1742892

[19] Sagitov, S. (1999) The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36**, 1116–1125. MR1742154

[20] Sagitov, S. (2003) Convergence to the coalescent with simultaneous multiple mergers. *J. Appl. Probab.* **40**, 839–854. MR2012671

[21] Schweinsberg, J. (2000) A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.* **5**, 1–11. MR1736720

[22] Schweinsberg, J. (2003) Coalescent processes obtained from supercritical Galton–Watson processes. *Stoch. Process. Appl.* **106**, 107–139. MR1983046

[23] Steutel F. W. and van Harn, K. (1979) Discrete analogues of self-decomposability and stability. *Ann. Probab.* **7**, 893–899. MR0542141