Λ-coalescents arising in a population with dormancy[*]

Fernando Cordero[†]   Adrián González Casanova[‡]
Jason Schweinsberg[§]   Maite Wilke-Berenguer[¶]

Abstract

Consider a population evolving from year to year through three seasons: spring, summer and winter. Every spring starts with N dormant individuals waking up independently of each other according to a given distribution. Once an individual is awake, it starts reproducing at a constant rate. By the end of spring, all individuals are awake and continue reproducing independently as Yule processes during the whole summer. In the winter, N individuals chosen uniformly at random go to sleep until the next spring, and the other individuals die. We show that because an individual that wakes up unusually early can have a large number of surviving descendants, for some choices of model parameters the genealogy of the population will be described by a Λ-coalescent. In particular, the beta coalescent can describe the genealogy when the rate at which individuals wake up increases exponentially over time. We also characterize the set of all Λ-coalescents that can arise in this framework.

1 Introduction

Dormancy is a widespread evolutionary strategy. A significant proportion of all living microorganisms are in some form of latent state of life [22]. The persistence of a latent subset of a population creates a buffer against selection [15, 16] and contributes to the maintenance of biodiversity [18]. Recently, probabilistic models have been successful in explaining some aspects of this facet of evolution and opening interesting questions in

[†]Technische Fakultät, Universität Bielefeld, Germany. E-mail: fcordero@techfak.uni-bielefeld.de
[‡]Instituto de Matemáticas, Universidad Nacional Autónoma de México, Mexico.
 E-mail: adrian.gonzalez@im.unam.mx
[§]Department of Mathematics, University of California San Diego, USA. E-mail: jschweinsberg@ucsd.edu
[¶]Institut für Mathematik, Humboldt-Universität zu Berlin, Germany. E-mail: mwb@math.hu-berlin.de

the fields of population genetics and beyond. The impact of dormancy at different scales on the coalescent processes describing the genealogies of populations were investigated in [19, 3, 4, 5]. Branching processes in random environment explain how dormancy can be selectively advantageous under fluctuating environmental conditions [6] while models from adaptive dynamics uncover that dormancy can arise from competition [7].

While some microbial organisms can last dormant for millions of years [27], most deactivation periods are much shorter. For example, mosquitos survive hostile environmental conditions by producing eggs which resist low temperatures and dry conditions [14]. Water triggers the eclosion of the eggs, so typically the newborns have favorable weather conditions and avoid the dry season. The earlier an individual reaches the reproductive state the higher its chances of having a large number of descendants in the next generation. This induces a selective pressure favoring mosquitos that are fortunate to be in contact with water soon after the rainy season starts.

The mechanisms for leaving a dormant state have been observed to be under selective pressure not only in the case of mosquitos [32] but also in different contexts such as experimental evolution; see, for example, figure 2 of [23] and [24]. In such experiments, individuals reproduce until the resources are depleted, and after that some of them are sampled and propagated to fresh identical media. The process of taking some bacteria from a depleted to a fully replenished environment induces a form of latency. Similar to the mosquitos, the bacteria undergoing the experiment are subjected to selective pressures resulting from the randomness of the activation time, and waking up rapidly from the dormant state provides an important advantage.

Wright and Vetsigian [33] postulated that the randomness in the times when individuals emerge from a dormant state could cause the distribution of the numbers of offspring produced by different individuals to become highly skewed. Indeed, they demonstrated in their bacterial experiments that "the heavy-tailed nature of the distribution of descendants can, in our case, be largely explained by phenotypic variability in lag time before exponential growth." It is well established in the probability literature [30, 31] that heavy-tailed offspring distributions can affect the genealogical structure of the population. Whereas the genealogy of populations under a wide range of conditions can be modeled by Kingman's coalescent [21], in which only two lineages can merge at a time, the genealogies of populations with heavy-tailed offspring distributions are sometimes best described by coalescents with multiple mergers, also known as Λ-coalescents, which allow many ancestral lines to merge at once. The primary aim of this paper is to describe how the randomness in the times when individuals emerge from dormancy affects the genealogy of the population and, in particular, to understand the conditions under which the genealogy is best described by a Λ-coalescent. Wright and Vetsigian [33] report that "it is unlikely that the variance diverges with population size for the particular species and conditions we examined," indicating that Λ-coalescents probably do not arise in this instance. Nevertheless, we find that Λ-coalescent genealogies can appear if some rare individuals emerge from dormancy sufficiently early.

1.1 A model involving dormancy

We now describe a population model involving dormancy, which is very similar to the one introduced by Wright and Vetsigian in Section 2.7 of [33]. Note that we refer to the time period in the model as a day, which might be most appropriate when considering an evolutionary experiment involving bacteria, but in other contexts it will be more natural to think of this time period as lasting a year, and we will refer to different seasons. The model evolves as follows. We begin every day (or year) with a population of N dormant individuals. Each day (or year) has length T_N and consists of three phases:

- **Activation phase (Spring)**: This phase has length t_N. Each individual wakes up at some random time before t_N. Once an individual is awake, it reproduces at rate λ_N, meaning that, as in a Yule process, it gives birth to a new individual after an exponentially distributed time with rate λ_N, and then the process repeats.

- **Active phase (Summer)**: This phase has length $T_N - t_N$, and during this phase all individuals are awake and reproducing at rate λ_N.

- **Sampling phase (Winter)**: At time T_N, we choose N individuals uniformly at random from the population to go to sleep until the start of the next day (or year), and all other individuals die.

Note that we have compressed the winter into a single time point, whereas it may be more realistic to think of the deaths as occurring gradually over a longer winter period. However, this will not substantially affect the results that follow.

The formalization is strongly inspired by the model introduced in [31]. For $d \in \mathbb{Z}$ and $i \in [N] := \{1, \ldots, N\}$, we denote by $\tau_{i,N}^{(d)}$ the random time in $[0, t_N]$ when the ith individual at the start of day d becomes active. We assume that the random variables $\tau_{1,N}^{(d)}, \ldots, \tau_{N,N}^{(d)}$ are independent and identically distributed. We denote by $X_{i,N}^{(d)}$ and $\nu_{i,N}^{(d)}$ the number of descendants of the ith individual starting day d at time T_N before and after sampling respectively. The total number of individuals at the end of day d before sampling is then

$$S_N^{(d)} := \sum_{i \in [N]} X_{i,N}^{(d)}.$$

Note also that $\sum_{i \in [N]} \nu_{i,N}^{(d)} = N$. Since the distributions of the $\tau_{i,N}^{(d)}$, $X_{i,N}^{(d)}$, and $\nu_{i,N}^{(d)}$ do not depend on d, we will frequently drop the superscripts when we are concerned only with the distributions of these quantities. By a well-known fact about Yule processes, conditional on the activation time $\tau_{i,N}$, the random variable $X_{i,N}$ is geometrically distributed with parameter $\exp(-\lambda_N(T_N - \tau_{i,N}))$. Moreover, conditional on $(X_{1,N}, \ldots, X_{N,N})$, we see that $\nu_{i,N}$ has a hypergeometric distribution with parameters S_N, $X_{i,N}$ and N.

We now explain how we will represent the genealogy of this population. Let us assume that we sample $n \in [N]$ individuals at random on day 0. We define a discrete-time Markov chain $(\Psi_{n,N}(d))_{d=0}^{\infty}$ taking its values on the set of partitions $\mathcal{P}_n$ of $[n]$, by letting $\Psi_{n,N}(d)$ be the partition of $[n]$ such that i and j are in the same block if and only if the ith and jth individuals in the sample have the same ancestor on day $-d$. We will be interested in the asymptotic behaviour as $N \to \infty$ of this ancestral process. The quantity

$$c_N := \frac{E[(\nu_{1,N})_2]}{N-1}, \tag{1.1}$$

where $(\cdot)_n$ denotes the falling factorial, will play a crucial role. Note that c_N is the probability that two individuals chosen uniformly at random from one generation have the same ancestor in the previous generation. It therefore establishes the appropriate time scale on which to study the process because after scaling time by $1/c_N$, the expected time for two randomly chosen individuals to trace their lineages back to a common ancestor will equal 1.

Tools for studying the limits as $N \to \infty$ in models such as this one were developed in [30, 26, 31]. It is well-known that when the distribution of the family sizes $\nu_{1,N}$ is highly skewed, the genealogy can sometimes be described by a Λ-coalescent. The Λ-coalescents were introduced independently by Pitman [29] and Sagitov [30] and also appeared around the same time in work of Donnelly and Kurtz [12]. Whenever Λ is a finite measure on $[0, 1]$, the Λ-coalescent is a $\mathcal{P}_n$-valued Markov process having the

property that whenever there are b blocks, each possible transition that involves k of the blocks merging into one happens at rate

$$\lambda_{b,k} := \int_0^1 y^{k-2}(1-y)^{b-k}\,\Lambda(\mathrm{d}y).$$

When Λ is a unit mass at 0, we have $\lambda_{b,2} = 1$ and $\lambda_{b,k} = 0$ whenever $k \geq 3$, so we obtain Kingman's coalescent in which each pair of blocks merges at rate 1. When $\Lambda(\{0\}) = 0$, the Λ-coalescent can be constructed from a Poisson point process on $[0,\infty) \times (0,\infty)$ with intensity $\mathrm{d}t \times y^{-2}\,\Lambda(\mathrm{d}y)$, in such a way that if (t,y) is a point of this Poisson point process, then at time t we have a merger event in which each lineage independently participates in the merger with probability y. This viewpoint is useful for understanding the results that follow. In particular, when Λ is a unit mass at one, we obtain a process known as the star-shaped coalescent in which all blocks merge after a waiting time whose distribution is exponential(1).

In Section 1.2, we consider a simple case in which there is no summer, and individuals can only wake up at the very beginning and at the very end of the spring. The genealogy of the population in this case is described in Theorem 1.1. In Section 1.3, we consider the case in which there is also no summer, but the rate at which individuals wake up from dormancy increases exponentially over time. The genealogy of the population in this case is stated in Theorem 1.3. In Section 1.4, we state Proposition 1.5, which gives conditions under which the summer period does not affect the genealogy of the population. In Section 1.5, we state Theorem 1.7, which characterizes the possible Λ-coalescents that can arise as limits in this model. After establishing some preliminary results in Section 2, we prove Theorem 1.1 in Section 3, Theorem 1.3 in Section 4, Proposition 1.5 in Section 5, and Theorem 1.7 in Section 6.

1.2 A two-point distribution for the exit time from dormancy

Here we describe the limiting genealogy in a simple instance of the model introduced above in which there is no summer, meaning that $T_N = t_N$, and the random variables $\tau_{i,N}$ can take only the two values 0 and T_N. We write

$$P(\tau_{i,N} = 0) = \omega_N = 1 - P(\tau_{i,N} = T_N),$$

where ω_N is assumed to satisfy

$$\lim_{N \to \infty} N\omega_N = 0. \tag{1.2}$$

Note that the progeny of individual i at time T_N is given by

$$X_{i,N} = b_{i,N}G_{i,N} + (1 - b_{i,N}),$$

where $(b_{i,N})_{i=1}^N$ is an i.i.d. family of Bernoulli random variables with parameter ω_N and $(G_{i,N})_{i=1}^N$ is an i.i.d. family of geometric random variables with parameters $e^{-\lambda_N T_N}$ independent of $(b_{i,N})_{i=1}^N$. Let us assume in addition that

$$\lambda_N T_N = \beta \log(\kappa N), \qquad \text{for some } \kappa, \beta > 0. \tag{1.3}$$

We denote by $\mathbb{D}_S[0,\infty)$ the set of càdlàg functions from $[0,\infty)$ to S, equipped with the usual Skorohod J_1 topology. We then have the following result. Throughout the paper, convergence of ancestral processes as $N \to \infty$ refers to weak convergence of stochastic processes in $\mathbb{D}_{\mathcal{P}_n}[0,\infty)$.

Theorem 1.1. *Assume that conditions* (1.2) *and* (1.3) *hold. Then:*

1. *If $\beta > 1$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ to the star-shaped coalescent.*

2. *If $\beta = 1$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ to the Λ-coalescent, with Λ being the probability measure characterized by $y^{-2} \Lambda_\kappa(\mathrm{d}y) := \frac{1}{E[Y_\kappa^2]} P(Y_\kappa \in \mathrm{d}y)$, where Y_κ is a $(0,1)$-valued random variable whose distribution is determined by*

$$P(Y_\kappa > x) := e^{-\frac{x}{\kappa(1-x)}}, \qquad x \in (0,1). \tag{1.4}$$

3. *If $\beta < 1$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ to Kingman's coalescent.*

This result can be understood as follows. Suppose an individual wakes up from dormancy unusually early, at time 0 rather than at time T_N. Then this individual spawns a branching process in which individuals give birth at rate λ_N, meaning that the number of descendants alive at time T_N will be geometrically distributed with mean $e^{\lambda_N T_N}$. Therefore, condition (1.3) with $\beta = 1$ is precisely what is needed for the number of descendants of this individual to be a constant multiple of N. Indeed, suppose (1.3) holds with $\beta = 1$. Then the number of descendants alive at time T_N can be approximately expressed as $\kappa N W$, where W has an exponential(1) distribution. Because there will also be $N - 1$ individuals who wake up at time T_N, the fraction of the population at time T_N that is descended from the individual who woke up early will be approximately $\kappa W/(\kappa W + 1)$, which means approximately a fraction $\kappa W/(\kappa W + 1)$ of the ancestral lines will merge at this time. Indeed it is straightforward to check using (1.4) that we have the equality in distribution

$$Y_\kappa =_d \frac{\kappa W}{\kappa W + 1}. \tag{1.5}$$

If instead (1.3) holds with $\beta < 1$, then the number of descendants at time T_N of the individual who woke up early will be much smaller than N, making multiple mergers unlikely and giving rise to a Kingman's coalescent genealogy. On the other hand, if $\beta > 1$, then the number of descendants of this individual will be much larger than N, meaning that with high probability all of the sampled ancestral lineages merge, leading to a star-shaped genealogy.

Remark 1.2. The precise asymptotic behavior of $(c_N)_{N=1}^\infty$ is characterized in Lemma 3.2.

1.3 Exponentially increasing rates of exit from dormancy

We now consider a possibly more realistic scenario in which the rate at which individuals exit from dormancy increases approximately exponentially over time. As in the model in Section 1.2, we assume there is no summer, so $T_N = t_N$. We assume that $\lambda_N = \lambda > 0$ for all N. We also assume that $T_N - \tau_{i,N} = \zeta_i \wedge T_N$, where $(\zeta_i)_{i=1}^N$ is a sequence of i.i.d. random variables whose distribution does not depend on N. We assume there exist constants $\gamma > 0$ and $c > 0$ such that, using $\sim$ to denote that the ratio of the two sides tends to one, we have

$$P(\zeta_1 > y) \sim ce^{-\gamma y}, \quad \text{as } y \to \infty, \tag{1.6}$$

Finally, we assume that

$$\lim_{N \to \infty} \frac{\log N}{T_N} = 0. \tag{1.7}$$

The condition (1.7) ensures that the spring period is long enough to allow the individuals that emerge unusually early from dormancy to have a number of descendants which is a constant multiple of N, making multiple mergers of ancestral lines possible. We obtain the following result.

Theorem 1.3. *Assume that conditions (1.6) and (1.7) hold, and let $a = \gamma/\lambda$. Then:*

1. *If $a \geq 2$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ to Kingman's coalescent.*

2. *If $1 \leq a < 2$, the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ to the Λ-coalescent in which Λ is the Beta$(2-a, a)$ distribution.*

3. *If $0 < a < 1$, the processes $(\Psi_{n,N}(\lfloor d \rfloor))_{d=0}^{\infty}$ converge as $N \to \infty$ to the discrete-time Ξ_a-coalescent described in Theorem 4(d) of [31].*

To understand this result, consider the distribution of the random variables $X_{i,N}$. If we disregard the truncation at T_N, which will turn out to have minimal effect, then conditional on $\zeta_i = u$, the distribution of $X_{i,N}$ is geometric with parameter $e^{-\lambda u}$. We assume for simplicity that the distribution of ζ_i is exactly exponential with rate γ. Then, making the change of variables $s = e^{-u\lambda}$ and using Stirling's approximation as $k \to \infty$ in the last step, we obtain

$$P(X_{i,N} > k) = \int_0^{\infty} (1 - e^{-\lambda u})^k \gamma e^{-\gamma u} \mathrm{d}u$$

$$= a \int_0^1 (1-s)^k s^{a-1} \mathrm{d}s = \frac{a\Gamma(k+1)\Gamma(a)}{\Gamma(k+1+a)} \sim \Gamma(a+1)k^{-a}.$$

Therefore, this result fits into the framework of Theorem 4 of [31], where it was established that beta coalescents describe the limiting genealogies in populations with these heavy-tailed offspring distributions. Note also that when $0 < a < 1$, there will be multiple individuals in each day whose descendants comprise a substantial fraction of the population, which is why the limit of the ancestral processes is not a Λ-coalescent but rather a discrete-time process in which multiple groups of lineages merge in each time step.

Remark 1.4. The asymptotic behavior of $(c_N)_{N=1}^{\infty}$ is considered in more detail in [31]. Lower bounds are given in Lemma 6 and Lemma 10 for the cases of $a > 2$ and $a = 2$ respectively. Lemma 13 yields $\lim_{N \to \infty} N^{a-1} c_N = C$ for a suitable constant C when $1 < a < 2$, and finally Lemma 16 states that $\lim_{N \to \infty} c_N \log(N) = 1$ for $a = 1$.

1.4 The effect of the summer on the genealogy

In Sections 1.2 and 1.3, we assumed there was no summer. We now consider how the inclusion of a summer period impacts the genealogy of the population. For this, we compare two populations whose spring has length t_N and whose reproduction rate is λ_N, which are subject to the same activation times. One population has a summer of length $T_N - t_N > 0$, and the other one has no summer, so that years have length t_N. Let $(\hat{\Psi}_{n,N}(d))_{d=0}^{\infty}$ denote the ancestral process associated with a sample of n individuals at time 0 from the population without a summer, and let $(\Psi_{n,N}(d))_{d=0}^{\infty}$ denote the corresponding ancestral process for the model with a summer. We obtain the following result.

Proposition 1.5. *Assume that there is a sequence $(\rho_N)_{N=1}^{\infty}$ of positive numbers such that $(\hat{\Psi}_{n,N}(\lfloor \rho_N t \rfloor))_{t \geq 0} \Rightarrow (\Psi_n(t))_{t \geq 0}$ in $\mathbb{D}_{\mathcal{P}_n}[0, \infty)$, where Ψ_n is a continuous-time Markov chain with values in $\mathcal{P}_n$. For the model with a summer, let $Y_{i,N}$ denote the number of descendants of the ith individual at the beginning of a given year who are alive at time t_N, and assume that*

$$\lim_{N \to \infty} \rho_N E\left[\frac{1}{Y_{1,N} + \cdots + Y_{N,N}}\right] = 0. \tag{1.8}$$

Then

$$(\Psi_{n,N}(\lfloor \rho_N t \rfloor))_{t \geq 0} \Rightarrow (\Psi_n(t))_{t \geq 0}.$$

When condition (1.8) holds, coalescence of lineages during the summer is sufficiently rare that the inclusion of the summer period does not affect the genealogy of the population in the limit. To understand why (1.8) is the correct condition, note that under the usual scaling $\rho_N = 1/c_N$, the probability that two randomly chosen lineages coalesce is $c_N = 1/\rho_N$. On the other hand, conditional on there being M individuals alive at the end of the spring, the probability that two randomly chosen lineages merge during the summer, regardless of the length of the summer, is bounded above by $2/(M+1)$, as a consequence of Lemma 5.1. Therefore, condition (1.8) implies that the probability that two lineages coalesce in the summer is of smaller order than the probability that two lineages coalesce in the spring.

On the other hand, if (1.8) fails, then as long as the summer has a length that does not tend to zero as $N \to \infty$, the summer period will cause additional pairwise mergers of ancestral lines.

Remark 1.6. Because $Y_{i,N} \geq 1$ for all i, the expectation in (1.8) is bounded above by $1/N$. Therefore, assuming we take $\rho_N = 1/c_N$ as in Theorems 1.1 and 1.3, condition (1.8) will hold as long as

$$\lim_{N\to\infty} Nc_N = \infty.$$

In the setting of Theorem 1.1, we therefore see from Lemma 3.2 that when $\beta \geq 1$, condition (1.8) holds whenever $\lim_{N\to\infty} N^2\omega_N = \infty$, which ensures that early emergence from dormancy is not too rare. When $\beta < 1$, condition (1.8) holds when $\lim_{N\to\infty} N^{2\beta}\omega_N = \infty$, which in view of (1.2) is only possible if $\beta > 1/2$. In the setting of Theorem 1.3, we can see from Remark 1.4 that condition (1.8) holds whenever $a \geq 2$. In particular, when the genealogy is given by a Λ-coalescent, the multiple mergers happen rapidly enough that the inclusion of a summer period would not affect the genealogy of the population in the limit.

1.5 A characterization of the possible Λ-coalescent limits

Theorem 1.7 below characterizes the Λ-coalescents that can arise as limits of ancestral processes in the model introduced in Section 1.1. Essentially, the possible measures Λ are those that can be expressed as mixtures of the measures Λ that appear as limits in Theorem 1.1. Note that the model in Section 1.1 is characterized by the birth rates $(\lambda_N)_{N=1}^\infty$, the times $(t_N)_{N=1}^\infty$ and $(T_N)_{N=1}^\infty$, and the distributions of the random variables $(\tau_{1,N})_{N=1}^\infty$. Therefore, to say that the Λ-coalescent can arise as a limit in this model means that there are choices of these parameters for which the rescaled ancestral processes converge to the Λ-coalescent.

Theorem 1.7. *It is possible for the Λ-coalescent to arise as the limit of the rescaled ancestral processes* $(\Psi_{n,N}(\lfloor \rho_N t \rfloor))_{t \geq 0}$ *in the population model defined in Section 1.1 if and only if we can write*

$$\Lambda = b_0\delta_0 + b_1\delta_1 + \Lambda', \tag{1.9}$$

where b_0 and b_1 are nonnegative real numbers and Λ' is a measure on $(0,1)$ with density h with respect to Lebesgue measure, where

$$h(y) = \int_0^\infty \frac{1}{\kappa}\left(\frac{y}{1-y}\right)^2 e^{-\frac{y}{\kappa(1-y)}} \, \eta(\mathrm{d}\kappa) \tag{1.10}$$

for all $y \in (0,1)$ and $\int_0^\infty (1 \wedge \kappa^2) \, \eta(\mathrm{d}\kappa) < \infty$.

We now make a few remarks concerning this result:

1. To relate this result to Theorem 1.1, note that the density of the random variable

Y_κ described in (1.4) is given by

$$f_\kappa(y) = \frac{1}{\kappa(1-y)^2} e^{-\frac{y}{\kappa(1-y)}}, \qquad y \in [0,1].$$

Therefore, if Λ' has density h with respect to Lebesgue measure, where h is given by (1.10), then $y^{-2}\Lambda'(\mathrm{d}y)/\mathrm{d}y = \int_0^\infty f_\kappa(y)\,\eta(\mathrm{d}\kappa)$. Note also that

$$\Lambda'(\mathrm{d}y) = h(y)\mathrm{d}y = \left(\int_0^\infty y^2 f_\kappa(y)\eta(\mathrm{d}\kappa)\right)\mathrm{d}y = \int_0^\infty \Lambda_\kappa(\mathrm{d}y)E[Y_\kappa^2]\eta(\mathrm{d}\kappa). \qquad (1.11)$$

2. Suppose $\eta(\mathrm{d}\kappa) = \kappa^{-1-a}\,\mathrm{d}\kappa$, where $0 < a < 2$. Then $\int_0^\infty (1 \wedge \kappa^2)\,\eta(\mathrm{d}\kappa) < \infty$. In this case, making the substitution $x = y/(\kappa(1-y))$, we have

$$h(y) = \left(\frac{y}{1-y}\right)^2 \int_0^\infty e^{-\frac{y}{\kappa(1-y)}} \kappa^{-2-a}\,\mathrm{d}\kappa = \left(\frac{y}{1-y}\right)^{1-a} \int_0^\infty e^{-x}x^a\,\mathrm{d}x,$$

which is a constant multiple of the $\mathrm{Beta}(2-a,a)$ density appearing in Theorem 1.3.

3. Typically, when rescaling time in the ancestral processes, we take $\rho_N = 1/c_N$, so that the expected time required for two randomly chosen lineages to merge equals 1. When $\rho_N = 1/c_N$, the measure Λ that appears in the limit must be a probability measure. Allowing arbitrary scaling constants $(\rho_N)_{N=1}^\infty$ simply allows finite measures Λ that are not probability measure to arise in the limit. See Remark 6.1 for details.

4. For the purposes of this result, we may assume without loss of generality that $t_N = T_N$, because regardless of the choice of t_N, it is always possible that the distribution of the times $\tau_{i,N}$ may be concentrated on a smaller subinterval of $[0, T_N]$. That is, every Λ-coalescent that can arise in this model can arise in an instance of the model in which there is no summer period.

5. The decomposition (1.10) in Theorem 1.7 is unique. Indeed, if two measures η and $\hat\eta$ define the same function h via (1.10), then the Laplace transforms of the measures $z\eta \circ \phi^{-1}(\mathrm{d}z)$ and $z\hat\eta \circ \phi^{-1}(\mathrm{d}z)$, where $\phi(\kappa) = 1/\kappa$, coincide in $(0, \infty)$. We conclude that $\eta \circ \phi^{-1} = \hat\eta \circ \phi^{-1}$, and hence $\eta = \hat\eta$. In combination with Theorem 1.7 we infer that $\{\delta_0, \delta_1\} \cup \{\Lambda_\kappa\}_{\kappa>0}$ is the set of extremal points of the convex set of probability measures Λ appearing in Theorem 1.7. Note that (1.11) shows how any probability measure Λ appearing in Theorem 1.7 can be written as a mixture of the probability measures Λ_κ.

We end this section with an alternate characterization of the measures Λ' appearing in Theorem 1.7.

Proposition 1.8. *An integrable function $h : (0,1) \to (0,\infty)$ can be expressed as in (1.10) for some measure η on $[0,\infty)$ if and only if there is a completely monotone function $g : (0,\infty) \to (0,\infty)$ such that*

$$h(y) = \frac{y^2}{(1-y)^2}\, g\left(\frac{y}{1-y}\right),$$

for all $y \in (0,1)$ and $\int_0^\infty g(v)(1 \wedge v^2)\,\mathrm{d}v < \infty$.

2 Genealogies in Cannings models

The model introduced in Section 1.1 is an example of a Cannings model. Cannings models, first introduced in [9, 10], have discrete generations and a fixed population size N, and the distribution of the family size vector $(\nu_{1,N}, \ldots, \nu_{N,N})$ is required to be

exchangeable. For such models we always define the coalescence probability c_N as in (1.1). There is by now a standard set of tools for studying the genealogy of such models. To prove convergence of the ancestral process in our model to a Λ-coalescent, we will mainly use the following result, which is essentially Theorem 3.1 of [30]; see also Proposition 3 of [31].

Theorem 2.1. *Consider a Cannings model in which the sizes of the N families are denoted by $\nu_{1,N}, \ldots, \nu_{N,N}$. Now sample n individuals at random from the population in generation zero, and let $\Psi_{n,N}(k)$ be the partition of $[n]$ such that i and j are in the same block of $\Psi_{n,N}(k)$ if and only if the ith and jth individuals in the sample have the same ancestor in generation $-k$. Let Λ be a probability measure on $[0,1]$. Then the processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge in $\mathbb{D}_{\mathcal{P}_n}[0,\infty)$ to the Λ-coalescent for all $n \in \mathbb{N}$ if and only if the following three conditions hold:*

1. *We have*
$$\lim_{N \to \infty} c_N = 0.$$

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

3. *For all $x \in (0,1)$ such that $\Lambda(\{x\}) = 0$, we have*
$$\lim_{N \to \infty} \frac{N}{c_N} P(\nu_{1,N} > Nx) = \int_x^1 y^{-2} \, \Lambda(\mathrm{d}y).$$

Remark 2.2. In place of condition 1 in Theorem 2.1, Sagitov [30] has the condition that $N^{-1} E[(\nu_{1,N} - 1)^2] \to 0$, which is equivalent because the fact that $E[\nu_{1,N}] = 1$ implies that $E[(\nu_{1,N})_2] = E[(\nu_{1,N} - 1)^2]$. Also, in place of condition 2, Sagitov has the condition that for all integers $a \geq 2$, we have
$$\lim_{N \to \infty} \frac{E[(\nu_{1,N} - 1)^2 \cdots (\nu_{a,N} - 1)^2]}{N^a c_N} = 0.$$

However, the equivalence of (16) and (20) in [26] implies that we can consider $(\nu_{k,N})_2$ in place of $(\nu_{k,N} - 1)^2$, and equation (17) of [26] implies that the limit is zero for all $a \geq 2$ if and only if the limit is zero when $a = 2$. These observations lead to the formulation of the result given above.

Remark 2.3. In principle, one could enrich the Cannings models with neutral genetic types to obtain a forward model for the evolution of type frequencies. To this end, we assign types to each individual at generation zero and propagate them forwards in time by imposing the rule that each individual inherits the type of its parent. Assume now that we have two types, 0 and 1, and denote by Z_d^N the frequency of type 1 individuals in generation $d \in \mathbb{N}$. Let A_d^N denote the number of blocks in the partition $\Psi_{n,N}(d)$. One can use sampling duality as described in [8, Thm. 3.1] to discover that

$$E\left[(Z_{\lfloor t/c_N \rfloor}^N)^n \mid Z_0^N = z\right] = E\left[z^{A_{\lfloor t/c_N \rfloor}^N} \mid A_0^N = n\right] + o(1), \quad \text{as } N \to \infty. \tag{2.1}$$

Assume now that the ancestral process $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converges to the Λ-coalescent as $N \to \infty$ (this assumption is satisfied for example in the setting of Theorems 1.1 and 1.3). Relation (2.1) allows one to prove uniform convergence on compact time intervals of the semigroup of $(Z_{\lfloor t/c_N \rfloor}^N)_{t \geq 0}$ to that of the two-types Λ-Fleming Viot process, when applied to polynomials. A standard density argument and Theorem [20, Thm. 19.28] yields the convergence of the involved processes. This line of thought was formalized in the diplom thesis [28] and can be generalized to the infinitely many types setting in the sense of [13].

For the rest of this section, we consider a subclass of Cannings models in which the family sizes in each generation are obtained in the following way. We consider a sequence of independent and identically distributed positive integer-valued random variables $X_{1,N}, \ldots, X_{N,N}$, where $X_{k,N}$ denotes the number of offspring produced by the kth individual. Note that we do not allow the random variables $X_{k,N}$ to take the value zero. We let $S_N = X_{1,N} + \cdots + X_{N,N}$ be the total number of offspring. We then sample N of the S_N offspring without replacement to form the next generation, and denote by $\nu_{k,N}$ the number of offspring of the kth individual that are sampled. Note that the model introduced in Section 1.1 fits into this framework.

In this section, we will use the notation $f(N) \ll g(N)$ to mean $\lim_{N \to \infty} f(N)/g(N) = 0$ and $f(N) \lesssim g(N)$ to mean $\sup_N f(N)/g(N) < \infty$.

Lemma 2.4 is useful for translating properties of the family sizes after sampling to properties of the family sizes before sampling, and vice versa.

Lemma 2.4. *For $r \in \mathbb{N}$, $k_1, \ldots, k_r \geq 2$, and $N > k_1 + \cdots + k_r$, we have*

$$\frac{E[(\nu_{1,N})_{k_1} \cdots (\nu_{r,N})_{k_r}]}{(N)_{k_1 + \cdots + k_r}} = E\left[\frac{(X_{1,N})_{k_1} \cdots (X_{r,N})_{k_r}}{(S_N)_{k_1 + \cdots + k_r}}\right]. \tag{2.2}$$

In particular,

$$c_N := \frac{E[(\nu_{1,N})_2]}{N-1} = NE\left[\frac{X_{1,N}(X_{1,N}-1)}{S_N(S_N-1)}\right]. \tag{2.3}$$

Moreover,

$$c_N \geq \frac{P(X_{1,N} \geq 2)^2}{2N}. \tag{2.4}$$

Proof. The identity (2.2) follows from equations (19) and (21) in the proof of Lemma 6 of [31]. Both sides give the probability that, if we sample $k_1 + \cdots + k_r$ individuals, the first k_1 are descended from the first individual in the previous generation, the next k_2 are descended from the second individual in the previous generation, and so on. The result (2.3) is a special case of (2.2). It remains to prove (2.4). Using (2.3) for the first inequality and Jensen's inequality for the third, we have

$$c_N \geq NE\left[\frac{X_{1,N}(X_{1,N}-1)}{S_N^2}\right] \geq \frac{N}{2}E\left[\frac{X_{1,N}^2}{S_N^2}\mathbb{1}_{\{X_{1,N} \geq 2\}}\right]$$

$$\geq \frac{N}{2}\left(E\left[\frac{X_{1,N}}{S_N}\mathbb{1}_{\{X_{1,N} \geq 2\}}\right]\right)^2 \geq \frac{N}{2}\left(\frac{1}{N}P(X_{1,N} \geq 2)\right)^2,$$

which ends the proof. $\qquad \square$

Lemma 2.5 shows that when $c_N \to 0$, the distribution of the total number of offspring is highly concentrated around some value a_N when N is large. This result implies that, when an unusually large family arises that will produce multiple mergers of ancestral lines, the total size of the remaining $N-1$ families in that generation can be treated as being essentially nonrandom, so that the size of the large family will determine the proportion of ancestral lineages that merge in this generation.

Lemma 2.5. *Suppose $\lim_{N \to \infty} c_N = 0$. Then there exists a sequence of positive numbers $(a_N)_{N=1}^{\infty}$ such that the following hold:*

1. *We have $S_N/a_N \to_p 1$, where $\to_p$ denotes convergence in probability as $N \to \infty$.*
2. *There exists $\delta > 0$ such that*

$$\lim_{N \to \infty} c_N^{-1} P\left(S_N - \max_{1 \leq k \leq N} X_{k,N} < \delta a_N\right) = 0.$$

Proof. Let $\bar{S}_{N-1} = S_N - X_{1,N} = X_{2,N} + \cdots + X_{N,N}$. Let

$$m_N = \inf\{k : P(\bar{S}_{N-1} \leq k) \geq 1/2\}. \tag{2.5}$$

Let $0 < \delta < 1/2$. We have

$$P(X_{1,N} \geq \delta \bar{S}_{N-1}) \geq P(X_{1,N} \geq \delta m_N) P(\bar{S}_{N-1} \leq m_N) \geq \frac{1}{2} P(X_{1,N} \geq \delta m_N).$$

Because $\bar{S}_{N-1} = S_N - X_{1,N}$, it follows that

$$NP(X_{1,N} \geq \delta m_N) \leq 2NP(X_{1,N} \geq \delta \bar{S}_{N-1}) = 2NP\left(X_{1,N} \geq \frac{\delta}{1+\delta} S_N\right). \tag{2.6}$$

Using the fact that $S_N \geq N$ followed by Markov's inequality, we have

$$
\begin{aligned}
P\left(X_{1,N} \geq \frac{\delta}{1+\delta} S_N\right) &= P\left(X_{1,N}(X_{1,N}-1) \geq \left(\frac{\delta}{1+\delta} S_N\right)\left(\frac{\delta}{1+\delta} S_N - 1\right)\right) \\
&= P\left(X_{1,N}(X_{1,N}-1) \geq \frac{\delta^2}{(1+\delta)^2} S_N(S_N-1)\left(\frac{S_N - \frac{1+\delta}{\delta}}{S_N - 1}\right)\right) \\
&\leq P\left(\frac{X_{1,N}(X_{1,N}-1)}{S_N(S_N-1)} \geq \frac{\delta^2}{(1+\delta)^2} \cdot \frac{N - \frac{1+\delta}{\delta}}{N-1}\right) \\
&\leq \left(\frac{(1+\delta)^2}{\delta^2} \cdot \frac{N-1}{N - \frac{1+\delta}{\delta}}\right) E\left[\frac{X_{1,N}(X_{1,N}-1)}{S_N(S_N-1)}\right].
\end{aligned} \tag{2.7}
$$

From (2.3), (2.6), and (2.7) we get

$$NP(X_{1,N} \geq \delta m_N) \leq \left(\frac{2(1+\delta)^2}{\delta^2} \cdot \frac{N-1}{N - \frac{1+\delta}{\delta}}\right) c_N, \tag{2.8}$$

which tends to zero as $N \to \infty$ by assumption.

For $1 \leq k \leq N$, let

$$W_{k,N} = X_{k,N} \mathbb{1}_{\{X_{k,N} \leq \delta m_N\}} + \mathbb{1}_{\{X_{k,N} > \delta m_N\}}.$$

Let $U_N = W_{1,N} + \cdots + W_{N,N}$, and let $\bar{U}_{N-1} = U_N - W_{1,N}$. Equation (2.8) implies

$$P(S_N \neq U_N) \leq NP(X_{1,N} \neq W_{1,N}) \to 0 \quad \text{as } N \to \infty. \tag{2.9}$$

Let $d_N = E[W_{1,N}]$, and let

$$a_N = Nd_N. \tag{2.10}$$

Then $E[U_N] = a_N$. It follows from (2.5) that

$$P(\bar{S}_{N-1} \leq m_N - 1) < 1/2,$$

and therefore by Markov's inequality,

$$
\begin{aligned}
1/2 &< P(\bar{S}_{N-1} \geq m_N) \\
&\leq P(\bar{U}_{N-1} \geq m_N) + (N-1)P(X_{1,N} \neq W_{1,N}) \\
&\leq \frac{E[\bar{U}_{N-1}]}{m_N} + (N-1)P(X_{1,N} \neq W_{1,N}) \\
&= \frac{(N-1)a_N}{Nm_N} + (N-1)P(X_{1,N} \neq W_{1,N}).
\end{aligned}
$$

Because $(N-1)P(X_{1,N} \neq W_{1,N}) \to 0$ by (2.9), it follows that for sufficiently large N, we have

$$\frac{a_N}{m_N} \geq \frac{1}{3}. \tag{2.11}$$

We now establish the result using a second moment argument. Given $r = 1/\ell$ for some positive integer ℓ and an integer k such that $1 \leq k \leq \ell$, define

$$U_{k,r,N} = W_{\lfloor (k-1)rN \rfloor+1,N} + \cdots + W_{\lfloor krN \rfloor,N}.$$

Note that $U_N = U_{1,1,N}$, and that $U_{k,r,N}$ is the sum of between $rN-1$ and $rN+1$ random variables. Therefore,

$$E[U_{k,rN}] \geq \Big(r - \frac{1}{N}\Big)a_N, \tag{2.12}$$

and using that $W_{1,N} \geq 1$, we have

$$\begin{aligned}
\mathrm{Var}(U_{k,r,N}) &\leq (rN+1)\mathrm{Var}(W_{1,N}) \\
&= (rN+1)\big(E[W_{1,N}^2] - (E[W_{1,N}])^2\big) \\
&\leq (rN+1)E[W_{1,N}(W_{1,N}-1)]. \tag{2.13}
\end{aligned}$$

Using that $X_{1,N} \leq \delta m_N$ on the event $\{W_{1,N}-1 \neq 0\}$, and that $W_{1,N}$ is independent of $\bar{S}_{N-1}$, we have

$$\begin{aligned}
c_N &\geq N E\left[\frac{W_{1,N}(W_{1,N}-1)}{S_N^2}\right] \geq N E\left[\frac{W_{1,N}(W_{1,N}-1)}{(\delta m_N + \bar{S}_{N-1})^2}\right] \\
&= N E[W_{1,N}(W_{1,N}-1)]E\left[\frac{1}{(\delta m_N + \bar{S}_{N-1})^2}\right]. \tag{2.14}
\end{aligned}$$

Now

$$E\left[\frac{1}{(\delta m_N + \bar{S}_{N-1})^2}\right] \geq P(\bar{S}_{N-1} = \bar{U}_{N-1})E\left[\frac{1}{(\delta m_N + \bar{S}_{N-1})^2}\Big|\bar{S}_{N-1} = \bar{U}_{N-1}\right]. \tag{2.15}$$

Let $Z_{1,N}, \ldots, Z_{N,N}$ be independent random variables whose distribution is the conditional distribution of $X_{1,N}$ given $X_{1,N} \leq \delta m_N$. Note that

$$E[Z_{1,N}] = E[W_{1,N}|X_{1,N} \leq \delta m_N] \leq \frac{E[W_{1,N}]}{P(X_{1,N} \leq \delta m_N)} = \frac{a_N}{NP(X_{1,N} \leq \delta m_N)}. \tag{2.16}$$

Let $\bar{V}_{N-1} = Z_{2,N} + \cdots + Z_{N,N}$. Then, using Jensen's inequality, followed by (2.16) and then (2.11), for N large enough that $\delta m_N > 1$ we have

$$\begin{aligned}
E\left[\frac{1}{(\delta m_N + \bar{S}_{N-1})^2}\Big|\bar{S}_{N-1} = \bar{U}_{N-1}\right] &= E\left[\frac{1}{(\delta m_N + \bar{V}_{N-1})^2}\right] \\
&\geq \frac{1}{(\delta m_N + E[\bar{V}_{N-1}])^2} \\
&\geq \Big(\delta m_N + \frac{(N-1)a_N}{NP(X_{1,N} \leq \delta m_N)}\Big)^{-2} \\
&\geq \frac{1}{a_N^2}\Big(3\delta + \frac{N-1}{NP(X_{1,N} \leq \delta m_N)}\Big)^{-2}.
\end{aligned}$$

It follows from (2.9) that $\lim_{N\to\infty}(N-1)/(NP(X_{1,N} \leq \delta m_N)) = 1$, and therefore for sufficiently large N, we have

$$E\left[\frac{1}{(\delta m_N + \bar{S}_{N-1})^2}\Big|\bar{S}_{N-1} = \bar{U}_{N-1}\right] \geq \frac{1}{a_N^2(1+4\delta)^2}.$$

Combining this result with (2.14) and (2.15), we get

$$c_N \geq NE[W_{1,N}(W_{1,N} - 1)]P(\bar{S}_{N-1} = \bar{U}_{N-1}) \cdot \frac{1}{a_N^2(1 + 4\delta)^2},$$

and therefore

$$E[W_{1,N}(W_{1,N} - 1)] \leq \frac{c_N a_N^2 (1 + 4\delta)^2}{NP(\bar{S}_{N-1} = \bar{U}_{N-1})}.$$

By (2.13),

$$\text{Var}(U_{k,r,N}) \leq \frac{(rN + 1)c_N a_N^2 (1 + 4\delta)^2}{NP(\bar{S}_{N-1} = \bar{U}_{N-1})}. \tag{2.17}$$

Let $\varepsilon > 0$. Taking $r = 1$ and using Chebyshev's inequality along with (2.9) and the assumption that $\lim_{N \to \infty} c_N = 0$, we get

$$P\left(\left|\frac{S_N}{a_N} - 1\right| > \varepsilon\right) \leq P(S_N \neq U_N) + \frac{\text{Var}(U_N)}{a_N^2 \varepsilon^2} \leq P(S_N \neq U_N) + \frac{(N + 1)c_N(1 + 4\delta)^2}{NP(\bar{S}_{N-1} = \bar{U}_{N-1})\varepsilon^2} \to 0.$$

That is, we have $S_N/a_N \to_p 1$, which is part 1 of the result. To prove part 2, we take $r = 1/3$ and note that we can have

$$S_N - \max_{1 \leq k \leq N} X_{k,N} < \delta a_N$$

only if at least two of the random variables $U_{k,r,N}$, for $k \in \{1, 2, 3\}$, are less than δa_N. Let $0 < \delta < 1/3$. By (2.12), (2.17), and Chebyshev's inequality,

$$\begin{aligned}
P(U_{k,r,N} < \delta a_N) &\leq P\left(\left|U_{k,r,N} - E[U_{k,r,N}]\right| > \left(r - \delta - \frac{1}{N}\right)a_N\right) \\
&\leq \left(r - \delta - \frac{1}{N}\right)^{-2} \frac{\text{Var}(U_{k,r,N})}{a_N^2} \\
&\leq \left(r - \delta - \frac{1}{N}\right)^{-2} \frac{(rN + 1)c_N(1 + 4\delta)^2}{NP(\bar{S}_{N-1} = \bar{U}_{N-1})}.
\end{aligned}$$

Therefore, the probability that at least two of the random variables $U_{k,r,N}$, for $k \in \{1, 2, 3\}$, are less than δa_N is bounded above by

$$3\left(r - \delta - \frac{1}{N}\right)^{-4} \cdot \frac{(rN + 1)^2 c_N^2 (1 + 4\delta)^4}{N^2 P(\bar{S}_{N-1} = \bar{U}_{N-1})^2}.$$

This expression tends to zero faster than c_N because $\delta < 1/3$. $\qquad \square$

Lemma 2.6. *Suppose* $\lim_{N \to \infty} c_N = 0$. *For all* $\varepsilon > 0$ *we have*

$$\lim_{N \to \infty} \frac{N^2}{c_N} P\left(\left|\frac{\nu_{k,N}}{N} - \frac{X_{k,N}}{S_N}\right| > \varepsilon \text{ for some } k \in [N]\right) = 0.$$

Proof. Suppose Y is the number of red balls drawn, when n balls are chosen from an urn containing b balls, of which r are red. Chvátal [11] showed that if $\varepsilon > 0$, then

$$P\left(Y \geq \left(\frac{r}{b} + \varepsilon\right)n\right) \leq e^{-2\varepsilon^2 n}. \tag{2.18}$$

Conditional on $X_{k,N}$ and S_N, we see that $\nu_{k,N}$ can be interpreted as the number of red balls drawn, when N balls are chosen from an urn containing S_N balls, of which $X_{k,N}$

are red. By applying (2.18) on the event that $S_N > N$, and noting that on the event that $S_N = N$, we have $\nu_{k,N} = X_{k,N} = 1$ for all k, we get

$$P\left(\left|\frac{\nu_{k,N}}{N} - \frac{X_{k,N}}{S_N}\right| > \varepsilon \text{ for some } k \in [N]\right) \leq 2Ne^{-2\varepsilon^2 N}P(S_N > N).$$

Conditional on $S_N > N$, the probability that two randomly chosen individuals have the same ancestor is at least $2/(N(N+1))$, so

$$c_N \geq \frac{2}{N(N+1)}P(S_N > N). \tag{2.19}$$

The result follows. $\qquad\square$

Lemma 2.7. *Suppose* $\lim_{N\to\infty} c_N = 0$. *For all* $\varepsilon > 0$, *we have*

$$\lim_{N\to\infty} \frac{N^2}{c_N}P\big(\nu_{1,N} \geq N\varepsilon,\, \nu_{2,N} \geq N\varepsilon\big) = 0. \tag{2.20}$$

This in turn implies

$$\lim_{N\to\infty} \frac{E[(\nu_{1,N})_2(\nu_{2,N})_2]}{N^2 c_N} = 0. \tag{2.21}$$

Proof. Let $0 < \varepsilon < 1$. Let Φ_N be the set of all ordered pairs (j,k) with $1 \leq j < k \leq N$ such that $\nu_{j,N} \geq N\varepsilon$ and $\nu_{k,N} \geq N\varepsilon$. Let $|\Phi_N|$ be the cardinality of Φ_N. Note that at most $1/\varepsilon$ of the random variables $\nu_{k,N}$ can exceed $N\varepsilon$, and so

$$|\Phi_N| \leq 1/\varepsilon^2. \tag{2.22}$$

Note that

$$P\big(\nu_{1,N} \geq N\varepsilon,\, \nu_{2,N} \geq N\varepsilon\big) = \frac{2}{N(N-1)}E\big[|\Phi_N|\big]. \tag{2.23}$$

Define the events

$$A_N = \left\{\left|\frac{\nu_{k,N}}{N} - \frac{X_{k,N}}{S_N}\right| \leq \frac{\varepsilon}{2} \text{ for all } k \in [N]\right\} \tag{2.24}$$

and

$$B_N = \left\{S_N - \max_{1 \leq k \leq N} X_{k,N} \geq \delta a_N\right\}, \tag{2.25}$$

where δ is the constant from Lemma 2.5. Also, note that $X_{k,N}/S_N > \varepsilon/2$ is equivalent to $X_{k,N} > (S_N - X_{k,N})\varepsilon/(2-\varepsilon)$, which on B_N implies that $X_{k,N} > \varepsilon\delta a_N/2$. Therefore, using (2.22),

$$E\big[|\Phi_N|\big] = E\big[|\Phi_N|\mathbb{1}_{A_N^c}\big] + E\big[|\Phi_N|\mathbb{1}_{B_N^c}\big] + E\big[|\Phi_N|\mathbb{1}_{A_N \cap B_N}\big]$$
$$\leq \frac{P(A_N^c) + P(B_N^c)}{\varepsilon^2} + \binom{N}{2}P\left(X_{1,N} > \frac{\varepsilon\delta a_N}{2}\right)^2.$$

Now $P(A_N^c) \ll N^{-2}c_N$ by Lemma 2.6 and $P(B_N^c) \ll c_N$ by part 2 of Lemma 2.5. Also, by (2.8) and (2.11), we have $P(X_{1,N} > \varepsilon\delta a_N/2) \lesssim N^{-1}c_N$. Combining these observations, and using that $c_N \to 0$ by assumption, we get

$$\lim_{N\to\infty} c_N^{-1}E\big[|\Phi_N|\big] = 0.$$

The claim (2.20) now follows from (2.23).

To prove (2.21), note that

$$\frac{E[(\nu_{1,N})_2(\nu_{2,N})_2]}{N^2 c_N} \leq \frac{E\big[(\nu_{1,N})_2(\nu_{2,N})_2 \mathbb{1}_{\{\nu_{1,N} \geq N\varepsilon,\, \nu_{2,N} \geq N\varepsilon\}}\big]}{N^2 c_N}$$

$$+ \frac{E\big[(\nu_{1,N})_2(\nu_{2,N})_2 \mathbb{1}_{\{\nu_{1,N} < N\varepsilon\}}\big]}{N^2 c_N} + \frac{E\big[(\nu_{1,N})_2(\nu_{2,N})_2 \mathbb{1}_{\{\nu_{2,N} < N\varepsilon\}}\big]}{N^2 c_N}$$

$$\leq \frac{N^4 P(\nu_{1,N} \geq N\varepsilon,\, \nu_{2,N} \geq N\varepsilon)}{N^2 c_N} + \frac{2(N\varepsilon) E[(\nu_{1,N})_2 \nu_{2,N}]}{N^2 c_N}.$$

By exchangeability and the fact that $\nu_{1,N} + \cdots + \nu_{N,N} = N$,

$$E[(\nu_{1,N})_2 \nu_{2,N}] = \frac{1}{N-1} E[(\nu_{1,N})_2(\nu_{2,N} + \cdots + \nu_{N,N})] \leq \frac{N}{N-1} E[(\nu_{1,N})_2] = N c_N.$$

It now follows, using (2.20), that

$$\limsup_{N \to \infty} \frac{E[(\nu_{1,N})_2(\nu_{2,N})_2]}{N^2 c_N} \leq \limsup_{N \to \infty} \frac{N^2}{c_N} P(\nu_{1,N} \geq N\varepsilon,\, \nu_{2,N} \geq N\varepsilon) + 2\varepsilon \leq 2\varepsilon.$$

Because $\varepsilon > 0$ was arbitrary, the result (2.21) follows. $\qquad\square$

Lemma 2.8. *Suppose* $\lim_{N \to \infty} c_N = 0$. *Suppose* Λ *is a probability measure on* $[0,1]$. *Then, the ancestral processes* $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ *in the Cannings model described above converge in* $\mathbb{D}_{\mathcal{P}_n}[0,\infty)$ *to the* Λ-*coalescent for all* n *if and only if for all* $x \in (0,1)$ *such that* $\Lambda(\{x\}) = 0$, *we have, defining* a_N *as in (2.10),*

$$\lim_{N \to \infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{a_N} > \frac{x}{1-x}\right) = \int_x^1 y^{-2} \Lambda(\mathrm{d}y). \qquad (2.26)$$

Proof. We need to check the three conditions of Theorem 2.1. Condition 1 holds by assumption, and condition 2 holds by Lemma 2.7.

It remains to show that (2.26) is equivalent to condition 3 of Theorem 2.1. Define the events A_N as in (2.24) with ε in place of $\varepsilon/2$, and define B_N as in (2.25). Let $\bar{S}_{N-1} = S_N - X_{1,N}$. Let $x \in (0,1)$. We have

$$P(\nu_{1,N} > Nx) \geq P\left(A_N \cap \left\{\frac{X_{1,N}}{S_N} > x + \varepsilon\right\}\right)$$

$$\geq P\left(A_N \cap \left\{\bar{S}_{N-1} \leq (1+\delta)a_N\right\} \cap \left\{\frac{X_{1,N}}{a_N} \geq \frac{(x+\varepsilon)(1+\delta)}{1-x-\varepsilon}\right\}\right)$$

$$\geq P(\bar{S}_{N-1} \leq (1+\delta)a_N) P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x+\varepsilon)(1+\delta)}{1-x-\varepsilon}\right) - P(A_N^c).$$

Note that $P(A_N^c) \ll N^{-2} c_N$ by Lemma 2.6. Also, because (2.8) and (2.11) imply $X_{1,N}/a_N \to_p 0$, it follows from Lemma 2.5 that $\bar{S}_{N-1}/a_N \to_p 1$. Therefore,

$$\liminf_{N \to \infty} \frac{N}{c_N} P(\nu_{1,N} > Nx) \geq \liminf_{N \to \infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x+\varepsilon)(1+\delta)}{1-x-\varepsilon}\right) \qquad (2.27)$$

and

$$\limsup_{N \to \infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x+\varepsilon)(1+\delta)}{1-x-\varepsilon}\right) \leq \limsup_{N \to \infty} \frac{N}{c_N} P(\nu_{1,N} > Nx). \qquad (2.28)$$

The other direction is more involved. By exchangeability and the inclusion-exclusion formula,

$$NP(\nu_{k,N} > Nx) - \binom{N}{2} P(\nu_{1,N} > Nx,\, \nu_{2,N} > Nx) \leq P\left(\bigcup_{k=1}^N \{\nu_{k,N} > Nx\}\right). \qquad (2.29)$$

Now

$$P\left(\bigcup_{k=1}^{N}\{\nu_{1,N} > Nx\}\right) \leq P(A_N^c) + P(B_N^c) + P\left(B_N \cap \bigcup_{k=1}^{N}\left\{\frac{X_{k,N}}{S_N} > x - \varepsilon\right\}\right)$$
$$\leq P(A_N^c) + P(B_N^c) + NP\left(B_N \cap \left\{\frac{X_{1,N}}{S_N} > x - \varepsilon\right\}\right). \qquad (2.30)$$

Note that $\bar{S}_{N-1} \geq \delta a_N$ on B_N. Then

$$P\left(B_N \cap \left\{\frac{X_{1,N}}{S_N} > x - \varepsilon\right\}\right)$$
$$\leq P\left(\{\delta a_N \leq \bar{S}_{N-1} \leq (1-\delta)a_N\} \cap \left\{\frac{X_{1,N}}{X_{1,N} + \delta a_N} > x - \varepsilon\right\}\right)$$
$$+ P\left(\{\bar{S}_{N-1} > (1-\delta)a_N\} \cap \left\{\frac{X_{1,N}}{X_{1,N} + (1-\delta)a_N} > x - \varepsilon\right\}\right)$$
$$\leq P(\bar{S}_{N-1} \leq (1-\delta)a_N)P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x-\varepsilon)\delta}{1-x+\varepsilon}\right) + P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x-\varepsilon)(1-\delta)}{1-x+\varepsilon}\right),$$
$$(2.31)$$

and plugging this back into (2.29) and (2.30), we get

$$NP(\nu_{1,N} > Nx) \leq P(A_N^c) + P(B_N^c) + \binom{N}{2}P(\nu_{1,N} > Nx, \nu_{2,N} > Nx)$$
$$+ NP(\bar{S}_{N-1} \leq (1-\delta)a_N)P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x-\varepsilon)\delta}{1-x+\varepsilon}\right)$$
$$+ NP\left(\frac{X_{1,N}}{a_N} \geq \frac{(x-\varepsilon)(1-\delta)}{1-x+\varepsilon}\right). \qquad (2.32)$$

We now show that the first four terms on the right-hand side of (2.32) are small. We have $P(A_N^c) \ll N^{-2}c_N$ by Lemma 2.6, and $P(B_N^c) \ll c_N$ by part 2 of Lemma 2.5. We have $P(\nu_{1,N} > Nx, \nu_{2,N} > Nx) \ll N^{-2}c_N$ by Lemma 2.7. Recall also that $\bar{S}_{N-1}/a_N \to_p 1$, and therefore $P(\bar{S}_{N-1} \leq (1-\delta)a_N) \to 0$, while we have $P(X_{1,N} \geq \theta a_N) \lesssim N^{-1}c_N$ for all $\theta > 0$ by (2.8) and (2.11). Thus, we get

$$\limsup_{N \to \infty} \frac{N}{c_N}P(\nu_{1,N} > Nx) \leq \limsup_{N \to \infty} \frac{N}{c_N}P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x-\varepsilon)(1-\delta)}{1-x+\varepsilon}\right) \qquad (2.33)$$

and

$$\liminf_{N \to \infty} \frac{N}{c_N}P\left(\frac{X_{1,N}}{a_N} \geq \frac{(x-\varepsilon)(1-\delta)}{1-x+\varepsilon}\right) \geq \liminf_{N \to \infty} \frac{N}{c_N}P(\nu_{1,N} > Nx). \qquad (2.34)$$

Recall that we need to show that condition 3 of Theorem 2.1 is equivalent to (2.26). First, suppose condition 3 of Theorem 2.1 holds. Choose $x \in (0,1)$ such that $\Lambda(\{x\}) = 0$. Because δ and ε can be arbitrarily small, equations (2.28) and (2.34) imply that for all $\theta > 0$ for which $\Lambda(\{x-\theta\}) = \Lambda(\{x+\theta\}) = 0$, we have

$$\int_{x+\theta}^{1} y^{-2}\,\Lambda(\mathrm{d}y) = \liminf_{N \to \infty} \frac{N}{c_N}P(\nu_{1,N} > N(x+\theta)) \leq \liminf_{N \to \infty} \frac{N}{c_N}P\left(\frac{X_{1,N}}{a_N} \geq \frac{x}{1-x}\right)$$
$$\leq \limsup_{N \to \infty} \frac{N}{c_N}P\left(\frac{X_{1,N}}{a_N} \geq \frac{x}{1-x}\right) \leq \limsup_{N \to \infty} P(\nu_{1,N} > N(x-\theta))$$
$$= \int_{x-\theta}^{1} y^{-2}\,\Lambda(\mathrm{d}y).$$

Because $\Lambda(\{x\}) = 0$, equation (2.26) follows. Conversely, suppose (2.26) holds for all x such that $\Lambda(\{x\}) = 0$. Reasoning as above, equations (2.27) and (2.33) imply that for all $\theta > 0$ for which $\Lambda(\{x - \theta\}) = \Lambda(\{x + \theta\}) = 0$, we have

$$
\begin{aligned}
\int_{x+\theta}^1 y^{-2}\,\Lambda(\mathrm{d}y) = \liminf_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{a_N} > \frac{x+\theta}{1-(x+\theta)}\right) &\leq \liminf_{N\to\infty} \frac{N}{c_N} P(\nu_{1,N} > Nx) \\
&\leq \limsup_{N\to\infty} \frac{N}{c_N} P(\nu_{1,N} > Nx) \leq \limsup_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{a_N} \geq \frac{x-\theta}{1-(x-\theta)}\right) \\
&= \int_{x-\theta}^1 y^{-1}\,\Lambda(\mathrm{d}y).
\end{aligned}
$$

Therefore, condition 3 of Theorem 2.1 holds, and the proof is complete. □

3 Results for two-point distributions

We would like to prove the results for the simple model with a two-point distribution for the activation times introduced in Section 1.2. To this end, we begin with an observation for the asymptotic behaviour of the moments of the geometric distributions that will govern the numbers of offspring of the early bird in each of the three regimes. The more general results from Section 2 will be very useful.

Let $G_{1,N}$ be a geometric random variable with parameter $e^{-\lambda_N T_N}$. The following lemma provides, under condition (1.3), the asymptotic behavior of the nth moments

$$
M_N(n) := E\left[\left(\frac{G_{1,N}}{G_{1,N} + N - 1}\right)^n\right].
$$

Lemma 3.1. *Let $n \in \mathbb{N}$. Assume that (1.3) holds with $\kappa, \beta > 0$.*

1. *If $\beta > 1$, $\lim_{N\to\infty} M_N(n) = 1$.*

2. *If $\beta = 1$, $\lim_{N\to\infty} M_N(n) = E[Y_\kappa^n]$, where Y_κ is a random variable on $[0,1]$ with distribution*
$$
P(Y_\kappa > x) = e^{-\frac{x}{\kappa(1-x)}}, \qquad x \in (0,1).
$$

3. *If $\beta < 1$, $M_N(n) \sim n! \kappa^{n\beta} N^{-n(1-\beta)}$ as $N \to \infty$.*

Proof. Note first that

$$
M_N(n) = \int_0^1 P\left(\frac{G_{1,N}}{G_{1,N} + N - 1} > x^{1/n}\right)\mathrm{d}x = \int_0^1 P\left(G_{1,N} > \frac{(N-1)x^{1/n}}{1 - x^{1/n}}\right)\mathrm{d}x. \quad (3.1)
$$

Assume now that $\beta > 1$ and let us show part 1 of the lemma. Note that, for any $y \geq 0$

$$
P\left(G_{1,N} > (N-1)y\right) = \left(1 - \frac{1}{(\kappa N)^\beta}\right)^{\lfloor (N-1)y \rfloor} \xrightarrow[N\to\infty]{} 1.
$$

Thus, part 1 of the lemma follows from the dominated convergence theorem.

Assume now that $\beta = 1$ and let us prove part 2. Note that, for any $y \geq 0$

$$
P\left(G_{1,N} > (N-1)y\right) = \left(1 - \frac{1}{\kappa N}\right)^{\lfloor (N-1)y \rfloor} \xrightarrow[N\to\infty]{} e^{-\frac{y}{\kappa}}.
$$

Thus, part 2 follows using the dominated convergence theorem and making the substitution $y = x^{1/n}$.

In the remainder of the proof we assume that $\beta < 1$. Making the change of variable $y = x^{1/n}/(1 - x^{1/n})$ in (3.1) and using standard properties of the floor function, we obtain

$$M_N(n) \sim \int_0^\infty \frac{ny^{n-1}}{(1+y)^{n+1}} \left(1 - \frac{1}{(\kappa N)^\beta}\right)^{(N-1)y} \mathrm{d}y = \int_0^\infty \frac{ny^{n-1}}{(1+y)^{n+1}} e^{-\alpha_N y} \mathrm{d}y,$$

as $N \to \infty$, where $\alpha_N = -(N-1)\log(1 - (\kappa N)^{-\beta})$. Making the substitution $z = \alpha_N y$, we obtain

$$M_N(n) \sim \alpha_N^{-n} \int_0^\infty \frac{nz^{n-1}}{\left(1 + \frac{z}{\alpha_N}\right)^{n+1}} e^{-z} \mathrm{d}z \sim \alpha_N^{-n} \int_0^\infty nz^{n-1} e^{-z} \mathrm{d}z = \alpha_N^{-n} n!,$$

and part 3 follows since $\alpha_N \sim \kappa^{-\beta} N^{1-\beta}$ as $N \to \infty$. $\qquad\square$

Lemma 3.2. *Under assumption* (1.2), *we have*

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

Moreover, if (1.3) *holds for*

1. $\beta > 1$, *then* $c_N \sim N\omega_N$ *as* $N \to \infty$.
2. $\beta = 1$, *then* $c_N \sim N\omega_N E[Y_\kappa^2]$ *as* $N \to \infty$.
3. $\beta < 1$, *then* $c_N \sim 2\kappa^{2\beta}\omega_N N^{2\beta-1}$ *as* $N \to \infty$.

Proof. We begin with a few general observations. Let $\bar{S}_{N-1} = S_N - X_{1,N} = \sum_{i=2}^N X_{i,N}$, and note that, thanks to Lemma 2.4, we have

$$c_N = NE\left[\frac{(X_{1,N})_2}{(S_N)_2}\right] = N\omega_N E\left[\frac{(G_{1,N})_2}{(G_{1,N} + \bar{S}_{N-1})_2}\right], \tag{3.2}$$

Since the expectation in the previous expression is smaller than one, the first statement follows from the assumption that $N\omega_N \to 0$ as $N \to \infty$.

Let us now have a closer look at the expectation in (3.2). Splitting on the event that there exists $b_{i,N} = 1$ with $i > 1$ and its complement, and defining

$$I_N^+ := E\left[\frac{(G_{1,N})_2}{(G_{1,N} + \bar{S}_{N-1})_2} \;\middle|\; \sum_{i=2}^N b_{i,N} > 0\right] \leq 1 \quad \text{and} \quad I_N^0 := E\left[\frac{(G_{1,N})_2}{(G_{1,N} + N - 1)_2}\right],$$

we obtain

$$E\left[\frac{(G_{1,N})_2}{(G_{1,N} + \bar{S}_{N-1})_2}\right] = (1 - (1-\omega_N)^{N-1})I_N^+ + (1-\omega_N)^{N-1}I_N^0, \tag{3.3}$$

Clearly, $0 \leq I_N^+ \leq I_N^0$, and because $N\omega_N \to 0$ it follows that I_N^0 is the leading term in (3.3) and

$$c_N = N\omega_N E\left[\frac{(G_{1,N})_2}{(G_{1,N} + \bar{S}_{N-1})_2}\right] \sim N\omega_N I_N^0, \quad \text{as } N \to \infty. \tag{3.4}$$

In addition, we have

$$I_N^0 = M_N(2) - E\left[\frac{(N-1)G_{1,N}}{(G_{1,N} + N - 1)(G_{1,N} + N - 1)_2}\right], \tag{3.5}$$

and the second term is smaller than $1/N$. Note that we have not used any assumptions on the distribution of $G_{1,N}$ up to this point.

Parts 1 and 2 of the lemma now follow directly using Lemma 3.1. In the remainder of the proof we assume that (1.3) holds for $\beta < 1$. In order to prove part 3, we use Lemma 3.1 to see that

$$M_N(2) \sim \frac{2\kappa^{2\beta}}{N^{2(1-\beta)}},$$

as $N \to \infty$. To complete the proof, we need to show that the second term in (3.5) converges faster to 0. Note that, using Lemma 3.1 with $n = 1$ in the last step, we get

$$E\left[\frac{G_{1,N}}{(G_{1,N} + N - 1)^2}\right] \leq \frac{1}{N} M_N(1) \sim \frac{\kappa^{\beta}}{N^{2-\beta}},$$

as $N \to \infty$. This completes the proof. $\qquad\square$

Lemma 3.3. *Assume that condition (1.2) holds. If in addition (1.3) holds for $\beta < 1$, then*

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

In particular, in this case, the processes $(\Psi_{n,N}(t/c_N))_{t\geq 0}$ converge to Kingman's coalescent.

Proof. The second statement follows from the first one using [25, Thm. 4(b)]. Let us prove the first statement. Thanks to Lemma 2.4, and using that $(X_{1,N})_3 = 0$ if $b_{1,N} = 0$, we obtain

$$\frac{E[(\nu_{1,N})_3]}{N^2 c_N} = \frac{(N-1)(N-2)}{N c_N} E\left[\frac{(X_{1,N})_3}{(S_N)_3}\right] \leq \frac{N\omega_N}{c_N} E\left[\frac{(G_{1,N})_3}{(G_{1,N} + \bar{S}_{N-1})_3}\right],$$

with $\bar{S}_{N-1} = \sum_{k=2}^{N} X_{i,N}$. Moreover, from part 3 in Lemma 3.2, we see that it suffices to show that

$$\lim_{N\to\infty} N^{2(1-\beta)} E\left[\frac{(G_{1,N})_3}{(G_{1,N} + \bar{S}_{N-1})_3}\right] = 0.$$

Now, using that $\bar{S}_{N-1} \geq N - 1$ and that, for $a \leq b$ and $b \geq 3$, we have $(a)_3/(b)_3 \leq (a/b)^3$, we obtain

$$E\left[\frac{(G_{1,N})_3}{(G_{1,N} + \bar{S}_{N-1})_3}\right] \leq E\left[\left(\frac{G_{1,N}}{G_{1,N} + N - 1}\right)^3\right] = M_N(3),$$

for $N \geq 3$, and the result follows using Lemma 3.1 with $n = 3$. $\qquad\square$

Proof of Theorem 1.1. The case $\beta < 1$ is already covered by Lemma 3.3.

We want to apply Theorem 2.1 for the case of $\beta \geq 1$. Conditions 1 and 2 hold by Lemma 3.2 and Lemma 2.7 respectively. Hence we are left to check condition 3.

Using Lemma 2.6, we obtain that for any $x \in (0,1)$ and any $\varepsilon \in (0, x \wedge (1-x))$, for sufficiently large N

$$\frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > x + \varepsilon\right) \leq \frac{N}{c_N} P\left(\nu_{1,N} > Nx\right) \leq \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > x - \varepsilon\right). \qquad (3.6)$$

Let A be the event that the individual with label 1 woke up early, i.e. $A := \{\tau_{1,N} = 0\} = \{b_{1,N} = 1\}$, and let B be the event that at least two individuals woke up early, i.e. $B := \{\sum_{i=1}^{N} b_{i,N} = 2\}$. For any $y \in (0,1)$ we can split the probability we are interested in into

$$P\left(\frac{X_{1,N}}{S_N} > y\right) = P\left(\frac{X_{1,N}}{S_N} > y \mid A^c\right) P(A^c) + P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B^c\right) P(A \cap B^c)$$

$$+ P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B\right) P(A \cap B).$$

Note that $P\left(X_{1,N}/S_N > y \mid A^c\right) \leq P(1/N > y) = 0$ for any given y and N sufficiently large. In addition we can bound $P\left(A \cap B\right) \leq N\omega_N^2$. By Lemma 3.2, we have $c_N \sim N\omega_N C(\beta)$, where $C(\beta) = 1$ if $\beta > 1$ and $C(1) = E[Y_\kappa^2]$ if $\beta = 1$. Hence,

$$\limsup_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B\right) P\left(A \cap B\right) = 0.$$

Observe that we can rewrite

$$P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B^c\right) = P\left(\frac{G_{1,N}}{G_{1,N} + N - 1} > y\right) = P\left(G_{1,N} > \frac{y}{1-y}(N-1)\right)$$

$$= \left(1 - \frac{1}{(\kappa N)^\beta}\right)^{\left\lfloor \frac{y}{1-y}(N-1)\right\rfloor},$$

since $G_{1,N}$ has a geometric distribution with parameter $(\kappa\beta)^{-1}$.

If $\beta > 1$, combining the observations above, we obtain

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y\right) = \lim_{N\to\infty} \frac{N\omega_N(1-\omega_N)^{N-1}}{N\omega_N}\left(1 - \frac{1}{(\kappa N)^\beta}\right)^{\left\lfloor \frac{y}{1-y}(N-1)\right\rfloor} = 1$$

for any $y \in (0,1)$. Setting $\Lambda_{ss} := \delta_1$, the Dirac-measure on 1, we conclude from (3.6) that

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\nu_{1,N} > Nx\right) = 1 = \int_x^1 z^{-2} \Lambda_{ss}(\mathrm{d}z),$$

for any $x \in (0,1)$. Theorem 2.1 then implies that the genealogy converges to the star-shaped coalescent.

Analogously, if $\beta = 1$, then defining Y_κ as in part 2 of Theorem 1.1,

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y\right) = \lim_{N\to\infty} \frac{N\omega_N(1-\omega_N)^{N-1}}{c_N}\left(1 - \frac{1}{\kappa N}\right)^{\left\lfloor \frac{y}{1-y}(N-1)\right\rfloor} = \frac{1}{E[Y_\kappa^2]}e^{-\frac{1}{\kappa}\frac{y}{1-y}}.$$

Since (3.6) holds for all $\varepsilon > 0$ and $\sup_{\varepsilon > 0} \frac{x-\varepsilon}{1-(x-\varepsilon)} = \inf_{\varepsilon>0} \frac{x+\varepsilon}{1-(x+\varepsilon)} = \frac{x}{1-x}$ this implies that

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\nu_{1,N} > Nx\right) = \frac{1}{E[Y_\kappa^2]}e^{-\frac{1}{\kappa}\frac{x}{1-x}} = \frac{1}{E[Y_\kappa^2]}P(Y_\kappa > x)$$

Hence, setting $\Lambda_\kappa(\mathrm{d}y) := \frac{y^2}{E[Y_\kappa^2]}P(Y_\kappa \in \mathrm{d}y)$, we obtain

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

for all $x \in (0,1)$, which ends the proof. $\qquad\square$

4 Exponentially increasing rates of exit from dormancy

This section is devoted to the proof of Theorem 1.3, which characterizes the asymptotic genealogies in the model described in Section 1.3.

4.1 A comparison between the genealogies of two models

The main ingredient in the proof of Theorem 1.3 is a result that allows us to compare the genealogies of two populations constructed from the same sequence $(\zeta_i)_{i=1}^N$ of i.i.d. positive random variables. The first model is the one described in Section 1.1, with no summer (i.e. $t_N = T_N$), with $\lambda_N = \lambda > 0$, and where $T_N - \tau_{i,N} = \zeta_i \wedge T_N$ (the model

in Section 1.3 is the special case where ζ_1 is exponentially distributed with parameter γ). In the second model, the family sizes $(\tilde{X}_1, \ldots, \tilde{X}_N)$ at the end of the year are i.i.d. and such that, conditionally on ζ_i, $\tilde{X}_i$ is geometrically distributed with parameter $e^{-\lambda\zeta_i}$ (i.e. days start at time $-\infty$). The vector of family sizes $(\tilde{\nu}_{1,N}, \ldots, \tilde{\nu}_{N,N})$ is obtained by sampling N individuals without replacement among the $\tilde{S}_N := \tilde{X}_1 + \cdots + \tilde{X}_N$ present at the end of the year.

The next result will be useful for comparing the genealogies of the previously described models.

Lemma 4.1. *Let* $f : \mathbb{R}_+^N \to \mathbb{R}_+$ *be a positive bounded function. Then,*

$$|E[f(X_{1,N}, \ldots, X_{N,N})] - E[f(\tilde{X}_1, \ldots, \tilde{X}_N)]| \leq 2\|f\|_\infty (1 - P(\zeta_1 \leq T_N)^N).$$

Proof. It follows directly from the fact that $(X_{1,N}, \ldots, X_{N,N})$ and $(\tilde{X}_1, \ldots, \tilde{X}_N)$ are equal if $\zeta_i \leq T_N$ for all $i \in [N]$. □

Let us now set

$$c_N := \frac{E[(\nu_{1,N})_2]}{N-1} \quad \text{and} \quad \tilde{c}_N := \frac{E[(\tilde{\nu}_{1,N})_2]}{N-1}.$$

The next result provides sufficient conditions for the limiting genealogies of the two models to coincide.

Proposition 4.2. *Assume that*

$$\lim_{N \to \infty} N^3 P(\zeta_1 > T_N) = 0. \tag{4.1}$$

Then $c_N \sim \tilde{c}_N$ *as* $N \to \infty$. *Moreover, for* $r \geq 1$, *the condition*

$$\lim_{N \to \infty} N^{r+2} P(\zeta_1 > T_N) = 0 \tag{4.2}$$

implies that, for all $k_1, \ldots, k_r \geq 2$,

$$\lim_{N \to \infty} \frac{E[(\nu_{1,N})_{k_1} \cdots (\nu_{r,N})_{k_r}]}{N^{k_1 + \cdots + k_r - r} c_N} = \lim_{N \to \infty} \frac{E[(\tilde{\nu}_{1,N})_{k_1} \cdots (\tilde{\nu}_{r,N})_{k_r}]}{N^{k_1 + \cdots + k_r - r} \tilde{c}_N},$$

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

Proof. Using Lemma 2.4 for the two models we obtain

$$c_N = NE\left[\frac{(X_{1,N})_2}{(S_N)_2}\right] \quad \text{and} \quad \tilde{c}_N = NE\left[\frac{(X_1)_2}{(\tilde{S}_N)_2}\right].$$

Thus, using Lemma 4.1 with f defined via $f(x_1, \ldots, x_N) = N(x_1)_2/(x_1 + \cdots + x_N)_2$, we obtain

$$|c_N - \tilde{c}_N| \leq 2N(1 - P(\zeta_1 < T_N)^N),$$

and the first result follows from (4.1) and (2.4).

For the second statement, let us assume that (4.2) holds. Note first that Lemma 2.4 yields

$$\frac{E[(\tilde{\nu}_{1,N})_{k_1} \cdots (\tilde{\nu}_{r,N})_{k_r}]}{N^{k_1 + \cdots + k_r - r} \tilde{c}_N} = \frac{(N)_{k_1 + \ldots + k_r}}{N^{k_1 + \ldots + k_r}} \frac{N^r}{\tilde{c}_N} E\left[\frac{(\tilde{X}_1)_{k_1} \cdots (\tilde{X}_r)_{k_r}}{(\tilde{S}_N)_{k_1 + \cdots + k_r}}\right], \tag{4.3}$$

and

$$\frac{E[(\nu_{1,N})_{k_1} \cdots (\nu_{r,N})_{k_r}]}{N^{k_1 + \cdots + k_r - r} c_N} = \frac{(N)_{k_1 + \ldots + k_r}}{N^{k_1 + \ldots + k_r}} \frac{N^r}{c_N} E\left[\frac{(X_{1,N})_{k_1} \cdots (X_{r,N})_{k_r}}{(S_N)_{k_1 + \cdots + k_r}}\right]. \tag{4.4}$$

Consider the function $f : \mathbb{N}^N \to \mathbb{R}_+$ defined via

$$f(x_1, \ldots, x_N) = \frac{(x_1)_{k_1} \cdots (x_r)_{k_r}}{(x_1 + \cdots + x_N)_{k_1 + \cdots + k_r}}.$$

Note that $f(x_1, \ldots, x_N)$ gives the probability that, if we sample $k_1 + \cdots + k_r$ balls from an urn containing, for each $i \in [N]$, x_i balls with label i, the first k_1 of them have label 1, the next k_2 balls have label 2, and so on. In particular, $\|f\|_\infty \leq 1$. Thus, applying Lemma 4.1 with f, we obtain, for N sufficiently large,

$$\frac{N^r}{\tilde{c}_N} \left| E\left[\frac{(\tilde{X}_1)_{k_1} \cdots (\tilde{X}_r)_{k_r}}{(\bar{S}_N)_{k_1+\cdots+k_r}} \right] - E\left[\frac{(X_{1,N})_{k_1} \cdots (X_{r,N})_{k_r}}{(S_N)_{k_1+\cdots+k_r}} \right] \right| \leq \frac{2N^r}{\tilde{c}_N}(1 - P(\zeta_1 < T_N)^N). \quad (4.5)$$

Therefore, using (2.4) and (4.2), we conclude that the left-hand side in (4.5) converges to zero as $N \to \infty$. Moreover, since (4.2) implies (4.1), we have $c_N \sim \tilde{c}_N$. Hence, the result follows from (4.3) and (4.4). $\qquad\square$

4.2 Exponential model

In this section, we come back to the model described in Section 4, i.e. we assume that ζ_1 satisfies (1.6), that its law has no mass at zero, and that T_N satisfies (1.7).

The next result provides the asymptotic behaviour of the tails of $\tilde{X}_1$.

Lemma 4.3. *Assume that* (1.6) *holds, and set* $a = \gamma/\lambda$. *Then,*

$$P(\tilde{X}_1 > k) \sim c\,\Gamma(1 + a)k^{-a}, \quad \text{as } k \to \infty. \quad (4.6)$$

Proof. Let μ denote the law of ζ_1. Consider the bijective function $f_\lambda : (0, \infty) \to (0, \infty)$ defined via

$$f_\lambda(u) := -\ln(1 - e^{-\lambda u}), \quad u \in (0, \infty),$$

and the measure $\mu_\lambda := \mu \circ f_\lambda^{-1}$, i.e. μ_λ is the push forward measure of μ by f_λ. With these definitions, we have

$$P(\tilde{X}_1 > k) = \int_0^\infty (1 - e^{-\lambda u})^k P(\zeta_1 \in \mathrm{d}u) = \int_0^\infty e^{-ku} \mu_\lambda(\mathrm{d}u),$$

i.e. the tail of $\tilde{X}_1$ is the Laplace transform of the measure μ_λ. The asymptotic behaviour of this Laplace transform around ∞ relates to the behaviour around zero of the distribution function of μ_λ via a Tauberian theorem. Let us make this precise. Set $\bar{\mu}(x) := \mu((y, \infty))$, $y > 0$, and note that

$$\mu_\lambda((0, x]) = \bar{\mu}(f_\lambda^{-1}(x)).$$

Hence, using the substitution $y = f_\lambda^{-1}(x)$, we obtain

$$\lim_{x \to 0+} \frac{\mu_\lambda((0, x])}{x^a} = \lim_{y \to \infty} \frac{\bar{\mu}(y)}{(f_\lambda(y))^a} = \lim_{y \to \infty} \frac{\bar{\mu}(y)}{e^{-\lambda a y}} = c.$$

The result then follows directly from Karamata's Tauberian theorem (Theorem 1.7.1 of [2]). $\qquad\square$

Proof of Theorem 1.3. Let us denote by $(\tilde{\Psi}_{n,N}(d))_{d=0}^\infty$ the ancestral process associated with the model whose family sizes (before sampling) are $(\tilde{X}_1, \ldots, \tilde{X}_N)$, and set $a := \gamma/\lambda$.

Since the law of ζ_1 satisfies (1.6), we conclude from Lemma 4.3 that the tails of $\tilde{X}_1$ satisfy (4.6). Hence, Theorem 4 in [31] implies that: i) if $a \geq 1$, the processes $(\tilde{\Psi}_{n,N}(\lfloor t/\tilde{c}_N \rfloor))_{t \geq 0}$ converge as $N \to \infty$ to the Λ-coalescent in which Λ is δ_0 if $a \geq 2$ or the Beta$(2 - a, a)$ distribution if $a \in [1, 2)$, and ii) if $a \in (0, 1)$, the processes $(\tilde{\Psi}_{n,N}(d))_{d=0}^\infty$ converge as $N \to \infty$ to the Ξ_a-coalescent described in Theorem 4(d) of [31].

It remains to prove that that the processes $(\Psi_{n,N}(d))_{d=0}^\infty$ and $(\tilde{\Psi}_{n,N}(d))_{d=0}^\infty$, in the appropriate time scale, have the same limiting genealogy. For this, note that, under (1.6) and (1.7), conditions (4.1) and (4.2) are satisfied. Hence, the result follows combining Proposition 4.2 with Theorem 2.1 in [26]. $\qquad\square$

5 Spring and summer

In this section we prove Proposition 1.5, which, informally speaking, states that under mild hypotheses, the length of the summer $T_N - t_N$ does not change the genealogy of the population. The key to the argument will be the following lemma about Pólya urns.

Lemma 5.1. *Consider a Pólya urn in which there are initially M balls, all of different colors. We repeatedly draw a ball at random from the urn and return it to the urn along with another ball of the same color, until there are N balls in the urn. Then if we choose two balls at random from the urn, the probability that they are the same color is at most $2/(M+1)$.*

Proof. Let D be the event that both balls have the same color. For $k \in \{0, 1, 2\}$, let R_k be the event that k out of the two balls chosen are among the M balls that were in the urn at the beginning. Then $P(R_2|D) = 0$ because the initial M balls all have different colors. By symmetry, we have $P(R_1|D) = 1/M$. Finally, the well-known exchangeability of Pólya urns (see, for example, [17]) implies that $P(R_0|D)$ is the same as the probability that the first two balls added to the urn are the same color, which is $2/(M+1)$. It follows that $P(D) = \sum_{k=0}^{2} P(R_k)P(D|R_k) \leq 2/(M+1)$. $\square$

Proof of Proposition 1.5. For the population with a summer, we want to bound the probability q_N that two individuals chosen at random at time T_N have the same ancestor at time t_N. Let $M = Y_{1,N} + \cdots + Y_{N,N}$ be the number of individuals alive at time t_N. Between times t_N and T_N, the sizes of the families started by the M individuals at time t_N evolve as independent Yule processes. Using a well-known connection between Yule processes and Pólya urns, one can see that the sequence keeping track of the families into which successive individuals are born follows the same dynamics as the sequence keeping track of the colors of successive balls added to a Pólya urn starting from M balls of different colors. Thus, by Lemma 5.1, the probability, conditional on M, that two individuals chosen at random at time T_N have the same ancestor at time t_N is at most $2/(M+1)$, which means the unconditional probability satisfies

$$q_N \leq 2E\left[\frac{1}{Y_{1,N} + \cdots + Y_{N,N}}\right].$$

Consider the canonical coupling of $\Psi_{n,N}$ and $\hat{\Psi}_{n,N}$ using the same activation times $\tau_{i,N}^{(d)}$ and the same birth times during the spring for both processes. Let D_N be the first day (or year) that at least one pair of individuals in the sample in $\Psi_{n,N}$ finds a common ancestor during the summer, that is, in $(t_N, T_N]$. By the coupling, the processes $\Psi_{n,N}$ and $\hat{\Psi}_{n,N}$ coincide until time D_N, and by the observations above and condition (1.8), for all $K > 0$ we have

$$P\big(\Psi_{n,N}(d) = \hat{\Psi}_{n,N}(d) \ \forall d = 0, \ldots, \lfloor K\rho_N \rfloor\big)$$
$$\geq P(D_N > K\rho_N) \geq \left(1 - 2\binom{n}{2}E\left[\frac{1}{Y_{1,N} + \cdots + Y_{N,N}}\right]\right)^{K\rho_N} \xrightarrow{N \to \infty} 1,$$

which implies the result. $\square$

6 Classifying the possible limits

In this section, we prove Theorem 1.7 and Proposition 1.8. The first one classifies all possible Λ-coalescents that can arise as limits in the model introduced in Section 1.1. The second one provides an alternate characterization of the measures Λ' appearing in Theorem 1.7.

Remark 6.1. Suppose the ancestral processes $(\Psi_{n,N}(\lfloor \rho_N t \rfloor))_{t \geq 0}$ converge to the Λ-coalescent for all n, where Λ is a finite nonzero measure. Because $\Lambda([0,1])$ is the rate at which two randomly chosen lineages merge in the Λ-coalescent and c_N is the probability that two individuals have the same parent in the previous generation in the Cannings model, we must have $\rho_N c_N \sim \Lambda([0,1])$ as $N \to \infty$. It follows that if we replace ρ_N by $1/c_N$, then all of the transition rates in the limit will be multiplied by $1/\Lambda([0,1])$. Therefore, the ancestral processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ will converge for all n to the $\tilde{\Lambda}$-coalescent, where $\tilde{\Lambda} = \Lambda/\Lambda([0,1])$ is a probability measure.

Conversely, suppose the ancestral processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge for all n to the Λ-coalescent, where Λ is a probability measure. Let $a > 0$. Then, if we choose $\rho_N = a/c_N$, the ancestral processes $(\Psi_{n,N}(\lfloor \rho_N t \rfloor))_{t \geq 0}$ will converge for all n to the $a\Lambda$-coalescent. We can also, of course, obtain convergence to the zero measure by choosing ρ_N such that $\rho_N c_N \to 0$ as $N \to \infty$.

In view of Remark 6.1, we may restrict our attention to the case in which $\rho_N = 1/c_N$, and to show that the possible Λ-coalescents that can arise as limits with this choice of scaling are precisely the probability measures satisfying (1.9) and (1.10). Proposition 6.2 shows how to obtain all of these Λ-coalescents as limits.

Proposition 6.2. *Let Λ be a probability measure on $[0,1]$ of the form*

$$\Lambda = b_0 \delta_0 + b_1 \delta_1 + \Lambda' \tag{6.1}$$

for nonnegative real numbers b_0 and b_1 and Λ' a measure on $(0,1)$ with density h given by (1.10). Then there exist choices for birth-rates $(\lambda_N)_{N=1}^{\infty}$, the times $(t_N)_{N=1}^{\infty}$, $(T_N)_{N=1}^{\infty}$, and distributions of the wake-up times $(\tau_{1,N})_{N=1}^{\infty}$ such that the ancestral processes $(\Psi_{n,N}(\lfloor t/c_N \rfloor))_{t \geq 0}$ converge for all n to the Λ-coalescent.

Proof. We will construct the approximating Cannings model as a mixture of the simple two-point models discussed in Section 1.2 and deduce the convergence of their ancestral processes to the desired Λ-coalescent from Theorem 2.1 using the analogous observations for the simple models made in Section 3.

For any $N \in \mathbb{N}$, let the length of spring be $t_N = T_N = \log(N^2)$. This is an arbitrary choice that ensures that the length of spring is finite in the Cannings model, but grows to infinity as $N \to \infty$. Let

$$\omega_N = N^{-2},$$

and let $b_{1,N}, \ldots, b_{N,N}$ be i.i.d. Bernoulli random variables with parameter ω_N. Choose the wake-up times to be of the form

$$\tau_{i,N} = (1 - b_{i,N})T_N + b_{i,N}\tilde{\tau}_{i,N}, \quad i \in [N],$$

where $\tilde{\tau}_{1,N}, \ldots, \tilde{\tau}_{N,N}$ are i.i.d. with a distribution that will depend on the specific form of Λ in (6.1), and $(b_{i,N})_{i \in [N]}$ and $(\tilde{\tau}_{i,N})_{i \in [N]}$ are independent, for any fixed $N \in \mathbb{N}$. Here $\tilde{\tau}_{i,N}$ describes the wake-up time of the individual i given that it woke up early. Assuming $\lambda_N \equiv 1$, the number of offspring of individual i at the end of spring before resampling is thus given by

$$X_{i,N} = (1 - b_{i,N}) + b_{i,N}G_{i,N}$$

where $G_{i,N}$ has a geometric distribution on $\mathbb{N}$ with parameter $e^{-(T_N - \tilde{\tau}_{i,N})}$. Abbreviate $\sigma_{i,N} := T_N - \tilde{\tau}_{i,N}$.

Like in (3.4) in the simple cases, we immediately see

$$c_N \sim N\omega_N E\left[\frac{(G_{1,N})_2}{(G_{1,N} + N - 1)_2}\right]. \tag{6.2}$$

Note that this observation holds for any distribution of $\sigma_{i,N}$. In particular, if we denote by $c_N^{(1)}$, $c_N^{(2),\kappa}$ and $c_N^{(3)}$ the analogs of c_N in the simple model with $\beta = 1+r$, $\beta = 1$ and $\beta = 1-r$, respectively, by Lemma 3.2 we know

$$N\omega_N E\left[\frac{(G_{1,N})_2}{(G_{1,N}+N-1)_2}\,\bigg|\,\sigma_{1,N}=\log(N^{1+r})\right] \sim c_N^{(1)} \sim N\omega_N, \tag{6.3}$$

$$N\omega_N E\left[\frac{(G_{1,N})_2}{(G_{1,N}+N-1)_2}\,\bigg|\,\sigma_{1,N}=\log(\kappa N)\right] \sim c_N^{(2),\kappa} \sim N\omega_N E[Y_\kappa^2], \quad \kappa > 0,$$

$$N\omega_N E\left[\frac{(G_{1,N})_2}{(G_{1,N}+N-1)_2}\,\bigg|\,\sigma_{1,N}=\log(N^{1-r})\right] \sim c_N^{(3)} \sim 2\omega_N N^{1-2r}.$$

Assume more specifically that $r \in (0,1/2)$, and define

$$\alpha_N := \int_{N^{-r/2}}^N \eta(\mathrm{d}\kappa).$$

Since we assume $\int_0^\infty (\kappa^2 \wedge 1)\eta(\mathrm{d}\kappa) < \infty$, we can estimate

$$\alpha_N = \int_{N^{-r/2}}^N \eta(\mathrm{d}\kappa) = \int_{N^{-r/2}}^N \frac{1}{(\kappa^2 \wedge 1)}(\kappa^2 \wedge 1)\eta(\mathrm{d}\kappa) \le N^r \int_0^\infty (\kappa^2 \wedge 1)\eta(\mathrm{d}\kappa). \tag{6.4}$$

In particular, $N^{-2r}\alpha_N \to 0$ as $N \to \infty$.

We now distinguish different cases for the possible choices of Λ in (6.1). Let us first treat the case where $b_0 > 0$. Using $X :=_d \mu$ to denote that we are defining a random variable X to have the distribution μ, for sufficiently large N we can define

$$\begin{aligned}
\sigma_{1,N} :=_d\; & \frac{2N^{-2r}\alpha_N}{b_0} \int_{N^{-r/2}}^N \delta_{(\log(\kappa N)\wedge T_N)}\frac{\eta(\mathrm{d}\kappa)}{\alpha_N} \\
& + \frac{2N^{-2r}b_1}{b_0}\delta_{(\log(N^{1+r})\wedge T_N)} + \left(1 - \frac{2N^{-2r}(\alpha_N+b_1)}{b_0}\right)\delta_{(\log(N^{1-r})\wedge T_N)} \\
=\; & \frac{2N^{-2r}\alpha_N}{b_0} \int_{N^{-r/2}}^N \delta_{\log(\kappa N)}\frac{\eta(\mathrm{d}\kappa)}{\alpha_N} \\
& + \frac{2N^{-2r}b_1}{b_0}\delta_{\log(N^{1+r})} + \left(1 - \frac{2N^{-2r}(\alpha_N+b_1)}{b_0}\right)\delta_{\log(N^{1-r})}.
\end{aligned}$$

If η is not the zero-measure, then α_N is positive and the integral in the first term is an integral with respect to the probability measure $\mathbb{1}_{[N^{-r/2},N]}(\kappa)\eta(\mathrm{d}\kappa)/\alpha_N$. Otherwise the first term is simply zero.

Note that this $\sigma_{1,N}$ is precisely a mixture of the simple cases discussed above. We now check the three conditions in Theorem 2.1 to obtain the desired convergence. Let $G_{1,N}^\kappa$ be a random variable having a geometric distribution on $\mathbb{N}$ with parameter $(\kappa N)^{-1}$. Conditioning on the possible values of $\sigma_{1,N}$ in (6.2), and using (6.3) in the last step,

which determines the asymptotic behavior of the last two summands, we then get

$$
c_N \sim N\omega_N \left\{ \frac{2N^{-2r}\alpha_N}{b_0} \int_{N^{-r/2}}^{N} E\left[\frac{(G_{1,N})_2}{(G_{1,N}+N-1)_2} \,\middle|\, \sigma_{1,N} = \log(\kappa N) \right] \frac{\eta(\mathrm{d}\kappa)}{\alpha_N} \right.
$$
$$
+ \frac{2N^{-2r}b_1}{b_0} E\left[\frac{(G_{1,N})_2}{(G_{1,N}+N-1)_2} \,\middle|\, \sigma_{1,N} = \log(N^{1+r}) \right]
$$
$$
\left. + \left(1 - 2N^{-2r}\frac{\alpha_N + b_1}{b_0}\right) E\left[\frac{(G_{1,N})_2}{(G_{1,N}+N-1)_2} \,\middle|\, \sigma_{1,N} = \log(N^{1-r}) \right] \right\}
$$
$$
\sim N\omega_N \frac{2N^{-2r}}{b_0} \int_{N^{-r/2}}^{N} E\left[\frac{(G_{1,N}^\kappa)_2}{(G_{1,N}^\kappa+N-1)_2} \right] \eta(\mathrm{d}\kappa)
$$
$$
+ \frac{2N^{-2r}b_1}{b_0} c_N^{(1)} + \left(1 - 2N^{-2r}\frac{\alpha_N + b_1}{b_0}\right) c_N^{(3)} \tag{6.5}
$$

Due to the integral, the first summand requires a bit more care. Equation (6.3) yields

$$
\lim_{N\to\infty} E\left[\frac{(G_{1,N}^\kappa)_2}{(G_{1,N}^\kappa+N-1)_2} \right] = E\left[Y_\kappa^2 \right].
$$

By (1.5), we have $Y_\kappa = \kappa W/(\kappa W + 1)$, where W has an exponential distribution with parameter 1. Therefore, using (1.4) and the fact that $E[Y_\kappa^2] = \int_0^\infty P(Y_\kappa > \sqrt{x})\,\mathrm{d}x$, we estimate

$$
\int_0^1 e^{-\frac{1}{\kappa}\frac{\sqrt{x}}{1-\sqrt{x}}}\,\mathrm{d}x = E\left[Y_\kappa^2 \right] = E\left[\left(\frac{\kappa W}{\kappa W + 1}\right)^2 \right] \le \kappa^2 E[W^2] = 2\kappa^2 \tag{6.6}
$$

for any $\kappa > 0$. Like in the simple cases (3.5), we obtain

$$
E\left[\frac{(G_{1,N}^\kappa)_2}{(G_{1,N}^\kappa+N-1)_2} \right] = E\left[\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa+N-1}\right)^2 \right] - E\left[\frac{(N-1)G_{1,N}^\kappa}{(G_{1,N}^\kappa+N-1)(G_{1,N}^\kappa+N-1)_2} \right],
$$

and the second term is smaller than $1/N$. Using (6.4)

$$
\int_{N^{-r/2}}^{N} E\left[\frac{(N-1)G_{1,N}^\kappa}{(G_{1,N}^\kappa+N-1)(G_{1,N}^\kappa+N-1)_2} \right] \eta(\mathrm{d}\kappa) \le N^{-1}N^r \int_0^\infty (\kappa^2 \wedge 1)\eta(\mathrm{d}\kappa) \xrightarrow{N\to\infty} 0. \tag{6.7}
$$

Like in (3.1) and below, we see

$$
E\left[\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa+N-1}\right)^2 \right] = \int_0^1 P\left(G_{1,N}^\kappa > \frac{(N-1)\sqrt{x}}{1-\sqrt{x}} \right) \mathrm{d}x
$$
$$
= \int_0^1 \left(1 - \frac{1}{\kappa N}\right)^{\left\lfloor (N-1)\frac{\sqrt{x}}{1-\sqrt{x}} \right\rfloor} \mathrm{d}x
$$
$$
\le \left(1 - \frac{1}{\kappa N}\right)^{-1} \int_0^1 e^{\frac{\sqrt{x}}{1-\sqrt{x}}(N-1)\log\left(1-\frac{1}{\kappa N}\right)} \mathrm{d}x
$$
$$
= \left(1 - \frac{1}{\kappa N}\right)^{-1} E\left[Y_{\frac{1}{c(\kappa,N)}}^2 \right],
$$

if we define

$$
c(\kappa, N) := -(N-1)\log\left(1 - \frac{1}{\kappa N}\right) > 0.
$$

Using (6.6), for any N sufficiently large we can bound

$$E\left[\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1}\right)^2\right] \le 3c(\kappa, N)^{-2}.$$

Standard calculus shows that $\kappa c(\kappa, N)$ is decreasing in κ, and we can therefore estimate

$$\inf_{N^{-r/2} \le \kappa \le 1} \kappa c(\kappa, N) = c(1, N) = -(N-1)\log\left(1 - \frac{1}{N}\right) \xrightarrow{N \to \infty} 1.$$

This yields a uniform upper bound for $c(\kappa, N)^{-1}$ because for all N sufficiently large

$$\forall \kappa \in [N^{-r/2}, 1], \quad \text{we have} \quad \kappa c(\kappa, N) \ge \frac{1}{2} \quad \text{and therefore} \quad c(\kappa, N)^{-1} \le 2\kappa.$$

This allows us to estimate

$$\mathbb{1}_{[N^{-r/2}, N]}(\kappa) E\left[\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1}\right)^2\right] \le \mathbb{1}_{[N^{-r/2}, 1]}(\kappa) 3c(\kappa, N)^{-2} + \mathbb{1}_{(1, N]}(\kappa)$$

$$\le \mathbb{1}_{[N^{-r/2}, 1]}(\kappa) 12\kappa^2 + \mathbb{1}_{(1, N]}(\kappa)$$

$$\le \mathbb{1}_{[N^{-r/2}, N]}(\kappa) 12(\kappa^2 \wedge 1). \tag{6.8}$$

Since we assume $\int_0^\infty (\kappa^2 \wedge 1)\eta(\mathrm{d}\kappa) < \infty$, we have found an integrable upper bound. Lebesgue's dominated convergence theorem and the fact that $c(\kappa, N)^{-1} \to \kappa$ as $N \to \infty$ yield

$$\lim_{N \to \infty} \int_0^\infty \mathbb{1}_{[N^{-r/2}, N]}(\kappa) E\left[\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1}\right)^2\right] \eta(\mathrm{d}\kappa) = \int_0^\infty E\left[Y_\kappa^2\right] \eta(\mathrm{d}\kappa).$$

If we combine this with (6.7), we obtain

$$\lim_{N \to \infty} \int_{N^{-r/2}}^N E\left[\frac{(G_{1,N}^\kappa)_2}{(G_{1,N}^\kappa + N - 1)_2}\right] \eta(\mathrm{d}\kappa) = \int_0^\infty E\left[Y_\kappa^2\right] \eta(\mathrm{d}\kappa). \tag{6.9}$$

It follows from (1.11) that

$$\int_0^\infty E[Y_\kappa^2]\,\eta(\mathrm{d}\kappa) = \Lambda'([0,1]) = 1 - b_0 - b_1, \tag{6.10}$$

and it was noted after (6.4) that $N^{-2r}\alpha_N \to 0$ as $N \to \infty$. Therefore, plugging (6.9) into (6.5) yields

$$c_N \sim N\omega_N \frac{2N^{-2r}}{b_0} \int_0^\infty E\left[Y_\kappa^2\right] \eta(\mathrm{d}\kappa) + \frac{2N^{-2r}b_1}{b_0} N\omega_N + \left(1 - 2N^{-2r}\frac{\alpha_N + b_1}{b_0}\right) 2\omega_N N^{1-2r}$$

$$\sim N^{1-2r}\omega_N\left(\frac{2}{b_0}(1 - b_0 - b_1) + \frac{2b_1}{b_0} + 2\right) = N^{1-2r}\omega_N \frac{2}{b_0}. \tag{6.11}$$

In particular, $c_N \to 0$ as $N \to \infty$ and the first condition of Theorem 2.1 holds. The second condition of Theorem 2.1 now follows directly from Lemma 2.7.

We are only left to verify the third condition. As in the proof of Theorem 1.1, let $A := \{\tau_{1,N} = 0\} = \{b_{1,N} = 1\}$ be the event that the individual with label 1 woke up early,

and let $B := \{\sum_{i=1}^{N} b_{i,N} = 2\}$ be the event that at least two individuals woke up early. Like in the simple case, regardless of the precise distribution of $\sigma_{1,N}$ we have

$$P\left(\frac{X_{1,N}}{S_N} > y\right) = P\left(\frac{X_{1,N}}{S_N} > y \mid A^c\right) P(A^c) + P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B\right) P(A \cap B)$$
$$+ P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B^c\right) P(A \cap B^c).$$

We have $P(X_{1,N}/S_N > y \mid A^c) \leq P(1/N > y) = 0$ for N sufficiently large, whence we may ignore this term in further considerations. Also, since $P(A \cap B) \leq N\omega_N^2$ and we chose $\omega_N = N^{-2}$,

$$\limsup_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y \mid A \cap B\right) P(A \cap B) \leq \lim_{N\to\infty} \frac{N}{c_N} N\omega_N^2 = \lim_{N\to\infty} \frac{b_0}{2} N^{-1+2r} = 0.$$

As before, we condition on the different values of $\sigma_{1,N}$. Letting $G_{1,N}^\kappa$, $G_{1,N}^+$, and $G_{1,N}^-$ be geometric random variables with parameters $(\kappa N)^{-1}$, $N^{-(1+r)}$ and $N^{-(1-r)}$ respectively, we obtain

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y\right) = \lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{G_{1,N}}{G_{1,N} + N - 1} > y\right) P(A_1 \cap B_2^c)$$
$$= \lim_{N\to\infty} \frac{N\omega_N}{c_N} \left\{ \frac{2N^{-2r}\alpha_N}{b_0} \int_{N^{-r/2}}^{N} P\left(\frac{G_{1,N}}{G_{1,N} + N - 1} > y \;\middle|\; \sigma_{1,N} = \log(\kappa N)\right) \frac{\eta(\mathrm{d}\kappa)}{\alpha_N} \right.$$
$$+ \frac{2N^{-2r}b_1}{b_0} P\left(\frac{G_{1,N}}{G_{1,N} + N - 1} > y \;\middle|\; \sigma_{1,N} = \log(N^{1+r})\right)$$
$$+ \left(1 - 2N^{-2r}\frac{\alpha_N + b_1}{b_0}\right) P\left(\frac{G_{1,N}}{G_{1,N} + N - 1} > y \;\middle|\; \sigma_{1,N} = \log(N^{1-r})\right) \right\}$$
$$= \lim_{N\to\infty} \int_{N^{-r/2}}^{N} P\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1} > y\right) \eta(\mathrm{d}\kappa)$$
$$+ \lim_{N\to\infty} b_1 P\left(\frac{G_{1,N}^+}{G_{1,N}^+ + N - 1} > y\right)$$
$$+ \lim_{N\to\infty} \frac{b_0}{2} N^{2r} P\left(\frac{G_{1,N}^-}{G_{1,N}^- + N - 1} > y\right).$$

Let us consider the three limits separately. Reasoning as in the proof of Lemma 3.1,

$$\lim_{N\to\infty} P\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1} > y\right) = \lim_{N\to\infty} \left(1 - \frac{1}{\kappa N}\right)^{\lfloor (N-1)\frac{y}{1-y} \rfloor} = e^{-\frac{1}{\kappa}\frac{y}{1-y}}.$$

From Chebychev's inequality and (6.8), we get $P(G_{1,N}^\kappa/(G_{1,N}^\kappa + N - 1) > y) \leq 12y^{-2}(\kappa^2 \wedge 1)$ when $\kappa \in [N^{-r/2}, 1]$. Therefore, we again obtain an integrable upper bound and can use Lebesgue's dominated convergence theorem to obtain

$$\lim_{N\to\infty} \int_{N^{-r/2}}^{N} P\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1} > y\right) \eta(\mathrm{d}\kappa) = \int_0^\infty e^{-\frac{1}{\kappa}\frac{y}{1-y}} \eta(\mathrm{d}\kappa). \qquad (6.12)$$

Likewise,

$$\lim_{N\to\infty} P\left(\frac{G_{1,N}^+}{G_{1,N}^+ + N - 1} > y\right) = \lim_{N\to\infty} \left(1 - \frac{1}{N^{1+r}}\right)^{\lfloor (N-1)\frac{y}{1-y} \rfloor} = 1. \qquad (6.13)$$

Lastly,

$$\lim_{N\to\infty} \frac{b_0}{2} N^{2r} P\left(\frac{G_{1,N}^-}{G_{1,N}^- + N - 1} > y\right) = \lim_{N\to\infty} \frac{b_0}{2} N^{2r} \left(1 - \frac{1}{N^{1-r}}\right)^{\lfloor (N-1)\frac{y}{1-y}\rfloor} = 0.$$

Combining this, we obtain

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y\right) = \int_0^\infty e^{-\frac{1}{\kappa}\frac{y}{1-y}} \eta(\mathrm{d}\kappa) + b_1 + 0 \qquad (6.14)$$

for every $y \in (0,1)$. Applying Lemma 2.6 as in (3.6) and using the identity

$$e^{-\ell x/(1-x)} = \int_x^1 \frac{\ell}{(1-y)^2} e^{-\ell y/(1-y)}\, \mathrm{d}y, \qquad (6.15)$$

equation (6.14) implies that for every $x \in (0,1)$, we have

$$\begin{aligned}
\lim_{N\to\infty} \frac{N}{c_N} P\left(\nu_{1,N} > Nx\right) &= \int_0^\infty e^{-\frac{1}{\kappa}\frac{x}{1-x}} \eta(\mathrm{d}\kappa) + b_1 + 0 \\
&= \int_0^\infty \int_x^1 \frac{1}{\kappa}\frac{1}{(1-y)^2} e^{-\frac{1}{\kappa}\frac{y}{1-y}} \mathrm{d}y\, \eta(\mathrm{d}\kappa) + b_1 + 0 \\
&= \int_x^1 y^{-2} h(y)\mathrm{d}y + b_1 + 0 \\
&= \int_x^1 y^{-2}(\Lambda' + b_1\delta_1 + b_0\delta_0)(\mathrm{d}y). \qquad (6.16)
\end{aligned}$$

With this we have verified the third condition of Theorem 2.1 and may thus conclude that the ancestral processes of the Cannings model we constructed do indeed converge to the Λ-coalescent.

The case of $b_0 = 0$ just requires an adaptation of the distribution of $\sigma_{1,N}$. We define

$$\sigma_{1,N} :=_d N^{-2r}\alpha_N \int_{N^{-r/2}}^N \delta_{\log(\kappa N)} \frac{\eta(\mathrm{d}\kappa)}{\alpha_N} + N^{-2r}b_1\delta_{\log(N^{1+r})} + \left(1 - N^{-2r}(\alpha_N + b_1)\right)\delta_0.$$

As before, we calculate c_N by conditioning on the possible values of $\sigma_{1,N}$ and obtain

$$c_N \sim N\omega_N N^{-2r} \int_{N^{-r/2}}^N E\left[\frac{(G_{1,N}^\kappa)_2}{(G_{1,N}^\kappa + N - 1)_2}\right] \eta(\mathrm{d}\kappa) + N^{-2r}b_1 c_N^{(1)} + 0.$$

Using (6.3), (6.9), and (6.10), we get

$$c_N \sim N^{1-2r}\omega_N \left\{\int_0^\infty E\left[Y_\kappa^2\right] \eta(\mathrm{d}\kappa) + b_1\right\} = N^{1-2r}\omega_N \{1 - b_1 + b_1\} = N^{1-2r}\omega_N,$$

which converges to 0 as $N \to \infty$ and therefore the first condition in Theorem 2.1 holds. Again, the second condition follows directly from Lemma 2.7. To obtain the third condition, as before, we condition on the different values of $\sigma_{1,N}$ and obtain

$$\begin{aligned}
\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y\right) = {}&\lim_{N\to\infty} \int_{N^{-r/2}}^N P\left(\frac{G_{1,N}^\kappa}{G_{1,N}^\kappa + N - 1} > y\right) \eta(\mathrm{d}\kappa) \\
&+ \lim_{N\to\infty} b_1 P\left(\frac{G_{1,N}^+}{G_{1,N}^+ + N - 1} > y\right).
\end{aligned}$$

Using (6.12) and (6.13) we obtain

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\frac{X_{1,N}}{S_N} > y\right) = \int_0^\infty e^{-\frac{1}{\kappa}\frac{y}{1-y}} \eta(\mathrm{d}\kappa) + b_1$$

for every $y \in (0,1)$ as in (6.14) and therefore, for every $x \in (0,1)$,

$$\lim_{N\to\infty} \frac{N}{c_N} P\left(\nu_{1,N} > Nx\right) = \int_x^1 y^{-2}(\Lambda' + b_1\delta_1)(\mathrm{d}y)$$

as in (6.16). With this we have verified the third condition of Theorem 2.1, which completes the proof. □

Proof of Theorem 1.7. Proposition 6.2 and Remark 6.1 establish that all measures Λ that can be written as in (1.9), with the density of Λ' being given by (1.10), can arise as limits of the ancestral processes in the model introduced in Section 1.1. It remains to show that these are the only measures that can be obtained.

In view of Remark 6.1, it suffices to consider the scaling in which $\rho_N = 1/c_N$. Note that if we denote by μ_N the distribution of $\exp(-\lambda_N(T_N - \tau_{1,N}))$, then

$$P(X_{1,N} > m) = \int_0^1 (1-p)^m \mu_N(\mathrm{d}p), \qquad \text{for all } m \in \mathbb{N}. \tag{6.17}$$

That is, the distribution of $X_{1,N}$ is a mixture of geometric distributions. We need to show that if (6.17) holds, then the measure Λ that appears on the right-hand side of (2.26) must satisfy (1.9) and (1.10).

Let $(a_N)_{N=1}^\infty$ be the sequence introduced in Lemma 2.5. For $0 < x < 1$, we have

$$\frac{N}{c_N} P\left(\frac{X_{1,N}}{a_N} > \frac{x}{1-x}\right) = \frac{N}{c_N} P\left(X_{1,N} > \frac{a_N x}{1-x}\right) = \frac{N}{c_N} \int_0^1 (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p).$$

We first show that we get minimal contribution to the integral when $p \geq N^{-3/4}$. Using (2.19),

$$\frac{N}{c_N} \int_{N^{-3/4}}^1 (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p) \leq \frac{N^2(N+1)}{2P(S_N > N)} \int_{N^{-3/4}}^1 (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p)$$

$$\leq \frac{N^2(N+1)}{2P(X_{1,N} > 1)} \int_{N^{-3/4}}^1 (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p)$$

$$= \frac{N^2(N+1)}{2} \frac{\int_{N^{-3/4}}^1 (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p)}{\int_{N^{-3/4}}^1 (1-p)\, \mu_N(\mathrm{d}p)}.$$

Note that $a_N \geq N$, which can be seen from (2.10) because the $X_{k,n}$, and therefore the $W_{k,N}$, are always at least 1. Hence, for all N and all $p \geq N^{-3/4}$, we have

$$\frac{(1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor}}{1-p} \leq \left(1 - \frac{1}{N^{3/4}}\right)^{\frac{Nx}{1-x}-2},$$

and therefore

$$\limsup_{N\to\infty} \frac{N}{c_N} \int_{N^{-3/4}}^1 (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p) \leq \limsup_{N\to\infty} \frac{N^2(N+1)}{2}\left(1 - \frac{1}{N^{3/4}}\right)^{\frac{Nx}{1-x}-2} = 0.$$

It follows that for all $x \in (0,1)$ such that $\Lambda(\{x\}) = 0$, we have

$$\lim_{N\to\infty} \frac{N}{c_N} \int_0^{N^{-3/4}} (1-p)^{\lfloor \frac{a_N x}{1-x}\rfloor} \mu_N(\mathrm{d}p) = \int_x^1 y^{-2}\, \Lambda(\mathrm{d}y). \tag{6.18}$$

We claim that for $p < N^{-3/4}$, we can make the approximation $1 - p \approx e^{-p}$, and we therefore assume for now that

$$\lim_{N \to \infty} \frac{N}{c_N} \int_0^{N^{-3/4}} \left((1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} - e^{-\frac{a_N p x}{1-x}} \right) \mu_N(dp) = 0. \tag{6.19}$$

Then, (6.18) implies that for all $x \in (0,1)$ such that $\Lambda(\{x\}) = 0$, we have

$$\lim_{N \to \infty} \frac{N}{c_N} \int_0^{N^{-3/4}} e^{-\frac{a_N p x}{1-x}} \mu_N(dp) = \int_x^1 y^{-2} \, \Lambda(dy).$$

Define a new measure χ_N on $(0, \infty)$ to be the push-forward of the restriction of μ_N to $(0, N^{-3/4})$ by the map $p \mapsto a_N p$, multiplied by N/c_N. Writing $z = x/(1-x)$, we then have

$$\lim_{N \to \infty} \int_0^\infty e^{-\ell z} \, \chi_N(d\ell) = \int_x^1 y^{-2} \, \Lambda(dy) \tag{6.20}$$

for all $x \in (0,1)$ such that $\Lambda(\{x\}) = 0$. We claim that this convergence must hold for all $x \in (0,1)$. To see this, we assume, seeking a contradiction, that $\Lambda(\{x\}) = b > 0$ for some $x > 0$. Choose u and v such that $0 < u < v < z$. Choose C_1 and C_2 such that $e^{-\ell v} \ell \le C_1 e^{-\ell u}$ for all $\ell \ge 0$ and $\int_0^\infty e^{-\ell u} \chi_N(d\ell) \le C_2$ for sufficiently large N. Choose $0 < \delta < \min\{z - v, b/(4x^2 C_1 C_2)\}$. Then, (6.20) implies that for sufficiently large N, we have

$$\frac{b}{2x^2} < \int_0^\infty \left(e^{-\ell(z-\delta)} - e^{-\ell(z+\delta)} \right) \chi_N(d\ell)$$

$$\le 2\delta \int_0^\infty e^{-\ell(z-\delta)} \ell \, \chi_N(d\ell)$$

$$\le 2\delta C_1 \int_0^\infty e^{-\ell u} \, \chi_N(d\ell)$$

$$\le 2\delta C_1 C_2$$

$$\le \frac{b}{2x^2},$$

which is a contradiction. Therefore, $\Lambda(\{x\}) = 0$ for all $x > 0$, and thus by (6.20) the Laplace transforms of the measures χ_N converge pointwise to a limit on $(0, \infty)$. By Theorem 8.5 of [1], it follows that the measures χ_N converge vaguely to a limit measure χ, and the pointwise limit of the Laplace transforms of χ_N is the Laplace transform of χ. That is, for all $x \in (0,1)$, we have

$$\int_0^\infty e^{-\ell x/(1-x)} \, \chi(d\ell) = \int_x^1 y^{-2} \, \Lambda(dy).$$

We now use (6.15) and change the order of integration to get

$$\int_x^1 \left(\int_0^\infty \ell \left(\frac{y}{1-y} \right)^2 e^{-\ell y/(1-y)} \, \chi(d\ell) \right) y^{-2} \, dy = \int_x^1 y^{-2} \, \Lambda(dy).$$

Now letting η be the push-forward of χ by the map $x \mapsto 1/x$, we see that the restriction of Λ to $(0,1)$ must have density h, as given in (1.10). To obtain the integrability condition, we note that because we are assuming $\rho_N = 1/c_N$, the measure Λ must be a probability measure, and therefore

$$1 \ge \int_0^1 h(y) \, dy = \int_0^\infty \int_0^1 y^2 \cdot \frac{1}{\kappa(1-y)^2} e^{-\frac{y}{\kappa(1-y)}} \, dy \, \eta(d\kappa).$$

Letting W have an exponential distribution with mean 1, the inner integral is

$$E\left[\left(\frac{\kappa W}{\kappa W + 1}\right)^2\right],$$

which is easily seen to be bounded between $C_3(1 \wedge \kappa^2)$ and $C_4(1 \wedge \kappa^2)$ for some positive constants C_3 and C_4 for all $\kappa > 0$. This implies that $\int_0^\infty (1 \wedge \kappa^2)\, \eta(\mathrm{d}\kappa) < \infty$.

It remains only to establish (6.19). For $0 < p < 1$, the Taylor expansion $\log(1-p) = -\sum_{n=1}^\infty p^n/n$ yields $-p/(1-p) \leq \log(1-p) \leq -p$, and therefore $e^{-p/(1-p)} \leq 1 - p \leq e^{-p}$. It follows that

$$e^{-\frac{p^2}{1-p}\frac{a_N x}{1-x}} e^{-\frac{a_N p x}{1-x}} = e^{-\left(\frac{p}{1-p}\frac{a_N x}{1-x}\right)} \leq (1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} \leq e^{-p\left(\frac{a_N x}{1-x}-1\right)} = e^{-\frac{a_N p x}{1-x}} e^p. \qquad (6.21)$$

For $p \leq N^{-3/4}$, the upper bound in (6.21) gives

$$e^{-\frac{a_N p x}{1-x}} \geq (1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} e^{-N^{-3/4}},$$

which, in combination with the finiteness of the right-hand side of (6.18), implies that

$$\limsup_{N \to \infty} \frac{N}{c_N} \int_0^{N^{-3/4}} \left((1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} - e^{-\frac{a_N p x}{1-x}}\right) \mu_N(\mathrm{d}p) \leq 0. \qquad (6.22)$$

For the bound in the other direction, note that if $p \leq N^{-3/4}$ and $a_N \leq N^{5/4}$, then the lower bound in (6.21) gives

$$e^{-\frac{a_N p x}{1-x}} \leq (1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} e^{\frac{2x}{(1-x)}N^{-1/4}}.$$

Hence, if we assume that $a_N \leq N^{5/4}$ for all $N \in \mathbb{N}$, the finiteness of the right-hand side of (6.18), gives

$$\liminf_{N \to \infty} \frac{N}{c_N} \int_0^{N^{-3/4}} \left((1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} - e^{-\frac{a_N p x}{1-x}}\right) \mu_N(\mathrm{d}p) \geq 0. \qquad (6.23)$$

Now suppose instead that $a_N > N^{5/4}$ for all $N \in \mathbb{N}$. If $p \leq N^{1/4}/a_N$ and $p \leq N^{-3/4}$, then $p^2 a_N \leq N^{-1/2}$. Therefore, the lower bound in (6.21) gives

$$e^{-\frac{a_N p x}{1-x}} \leq (1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} e^{\frac{2x}{(1-x)}N^{-1/2}},$$

and therefore

$$\liminf_{N \to \infty} \frac{N}{c_N} \int_0^{N^{1/4}/a_N} \left((1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} - e^{-\frac{a_N p x}{1-x}}\right) \mu_N(\mathrm{d}p) \geq 0. \qquad (6.24)$$

If, on the other hand, $N^{1/4}/a_N < p \leq N^{-3/4}$, note that the assumption $a_N > N^{5/4}$ implies, by part 1 of Lemma 2.5, that $P(S_N \geq \frac{1}{2}N^{5/4}) \geq 1/2$ for sufficiently large N, and therefore there is a positive constant C_5 such that $c_N \geq C_5/N$. Therefore,

$$\liminf_{N \to \infty} \frac{N}{c_N} \int_{N^{1/4}/a_N}^{N^{-3/4}} \left((1-p)^{\lfloor \frac{a_N x}{1-x} \rfloor} - e^{-\frac{a_N p x}{1-x}}\right) \mu_N(\mathrm{d}p) \geq \liminf_{N \to \infty} -\frac{N^2}{C_5} e^{-\frac{x}{1-x}N^{1/4}} = 0. \quad (6.25)$$

Equations (6.24) and (6.25) imply that (6.23) holds also when $a_N > N^{5/4}$ for all $N \in \mathbb{N}$. Hence, along with (6.22), we have proven that (6.19) holds for any sequence $(a_N)_{N=1}^\infty$ such that either all elements are greater or all elements are less or equal than $N^{5/4}$. Finally, to handle the case of general $(a_N)_{N=1}^\infty$, note that any subsequence will in turn have a subsequence which falls into one of the two categories we have proven to converge. $\qquad \square$

Proof of Proposition 1.8. First note that, the relation $Tg(y) = (y/(1-y))^2 g(y/(1-y))$, $y \in (0,1)$, defines a bijective function T mapping positive functions on $(0,\infty)$ into positive functions on $(0,1)$. Moreover, making the substitution $v = y/(1-y)$, we see that

$$\int_0^1 \frac{y^2}{(1-y)^2} \, g\left(\frac{y}{1-y}\right) \, \mathrm{d}y = \int_0^\infty \frac{v^2}{(1+v)^2} g(v) \, \mathrm{d}v.$$

Hence, the integrability of Tg is equivalent to the integrability of $v \mapsto g(v)(1 \wedge v^2)$. In addition, if h is given by (1.10), then $h = Tg$ with

$$g(v) = \int_0^\infty \frac{1}{\kappa} e^{-\frac{v}{\kappa}} \, \eta(\mathrm{d}\kappa).$$

Clearly g is the Laplace transform of the measure $m(\mathrm{d}z) := z\eta \circ \phi^{-1}(\mathrm{d}z)$, where $\phi(\kappa) = 1/\kappa$, and thus, g is completely monotone. Conversely, assume that $h = Tg$ for some completely monotone function g. The latter is, by Bernstein's theorem, the Laplace transform of a measure m on $[0,\infty)$. It follows that h is given by (1.10) with $\eta(\mathrm{d}\kappa) := \kappa \, m \circ \phi^{-1}(\mathrm{d}\kappa)$, which ends the proof. □

References

[1] R. Bhattacharya and E. Waymire (2007). *A Basic Course in Probability Theory*. Springer, New York. MR2331066

[2] N. H. Bingham, C. M. Goldie, and J. L. Teugels (1987). *Regular Variation*. Encyclopedia of Mathematics and its Applications. Cambridge University Press. MR1015093

[3] J. Blath, A. González Casanova, N. Kurt, and D. Spanò (2013). The ancestral process of long-range seed bank models. *Journal of Applied Probability,* 50(3):741–759. MR3102512

[4] J. Blath, A. González Casanova, N. Kurt, and M. Wilke-Berenguer (2016). A new coalescent for seed-bank models. *Annals of Applied Probability,* **26**:857–891. MR3476627

[5] J. Blath, A. González Casanova, N. Kurt, and M. Wilke-Berenguer (2020). The seed bank coalescent with simultaneous switching. *Electronic Journal of Probability,* **25**(27):21pp. MR4073688

[6] J. Blath, F. Hermann, and M. Slowik (2021). A branching process model for dormancy and seed banks in randomly fluctuating environments. *Journal of Mathematical Biology,* **83**(17):40pp. MR4289473

[7] J. Blath and A. Tóbiás (2020). Invasion and fixation of microbial dormancy traits under competitive pressure. *Stochastic Processes and their Applications,* **130**:7363–7395. MR4167209

[8] F. Boenkost, A. González Casanova, C. Pokalyuk and A. Wakolbinger (2021). Haldane's formula in Cannings models: the case of moderately weak selection. *Electronic Journal of Probability,* **26**(4):36pp. MR4216517

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

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

[11] V. Chvátal (1979). The tail of the hypergeometric distribution. *Discrete Mathematics,* **25**:285–287. MR0534946

[12] P. Donnelly and T. G. Kurtz (1999). Particle representations for measure-valued population models. *Annals of Probability,* **27**:166–205. MR1681126

[13] S. N. Evans (1997). Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques,* **33**:339-358. MR1457055

[14] K. J. Faull, and C. R. Williams (2015). Intraspecific variation in desiccation survival time of Aedes aegypti (L.) mosquito eggs of Australian origin. *Journal of Vector Ecology,* **40**:292–300.

[15] A. González Casanova, E. Aguirre von Wobeser, G. Espín, L. Servin González, N. Kurt, D. Spanò, J. Blath, and G. Soberón Chavez (2014). Strong seed-bank effects in bacterial evolution. *Journal of Theoretical Biology,* **356**:62–70.

[16] L. Heinrich, J. Müller, A. Tellier, and D. Živcović (2018). Effects of population- and seed bank size fluctuations on neutral evolution and efficacy of natural selection. *Theoretical Population Biology,* **123**:45–69.

[17] B. M. Hill, D. Lane, and W. P. Sudderth (1987). Exchangeable urn processes. *Annals of Probability* **15**: 1586–1592. MR0905350

[18] S. E. Jones, and J. T. Lennon (2010). Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences USA,* **107**(13):5881–5886.

[19] I. Kaj, S. Krone, and M. Lascoux. Coalescent theory for seed bank models (2001). *Journal of Applied Probability,* **38**(2):285–300. MR1834743

[20] O. Kallenberg. *Foundations of Modern Probability.* 2nd edition. Springer-Verlag, New York, 2002. MR1876169

[21] J. F. C. Kingman (1982). The coalescent. *Stochastic Processes and their Applications,* **13**:235–248. MR0671034

[22] J. T. Lennon and S. E. Jones (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology,* **9**:119–130.

[23] R. E. Lenski, J. A. Mongold, P. D. Sniegowski, et al. (1998). Evolution of competitive fitness in experimental populations of *E. coli*: What makes one genotype a better competitor than another? *Antonie Van Leeuwenhoek,* **73**:35–47.

[24] J. Lin, M. Manhart, and A. Amir (2020). Evolution of microbial growth traits under serial dilution. *Genetics,* **215**(3):767–777.

[25] M. Möhle (2000). Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. *Advances in Applied Probability,* **32**:983–993. MR1808909

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

[27] Y. Morono, M. Ito, T. Hoshino, T. Terada, T. Hori, M. Ikehara, S. D'Hondt, and F. Inagaki (2020). Aerobic microbial life persists in oxic marine sediment as old as 101.5 million years. *Nature Communications,* **11**(1):3626.

[28] N. Perkowski (2009). Stochastic Population Models: Measure-Valued and Partition-Valued Formulations. Diplom thesis, Humboldt-Universität zu Berlin.

[29] J. Pitman (1999). Coalescents with multiple collisions. *Annals of Probability,* **27**:1870–1902. MR1742892

[30] S. Sagitov (1999). The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability,* **36**:1116–1125. MR1742154

[31] J. Schweinsberg (2003). Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Processes and their Applications,* **106**:107–139. MR1983046

[32] J. A. Souza-Neto, J. R. Powell, and M. Bonizzoni (2019). *Aedes aegypti* vector competence studies: A review. *Infection, Genetics and Evolution,* **67**:191–209.

[33] E. S. Wright and K. H. Vetsigian (2019). Stochastic exits from dormancy give rise to heavy-tailed distributions of descendants in bacterial populations. *Molecular Ecology,* **28**:3915–3928.