The coalescent tree of a Markov branching process with generalised logistic growth

David Cheek

Received: 1 January 2021 / Revised: 2 March 2022 / Accepted: 12 March 2022 / Published online: 5 April 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
We consider a class of density-dependent branching processes which generalises exponential, logistic and Gompertz growth. A population begins with a single individual, grows exponentially initially, and then growth may slow down as the population size moves towards a carrying capacity. At a time while the population is still growing superlinearly, a fixed number of individuals are sampled and their coalescent tree is drawn. Taking the sampling time and carrying capacity simultaneously to infinity, we prove convergence of the coalescent tree to a limiting tree which is in a sense universal over our class of models.

Mathematics Subject Classification 60J80

1 Introduction

A coalescent tree is a random tree which models the genealogical structure of individuals sampled from a population, serving as a stepping stone between genetic data and models of population evolution. The most famous example is Kingman’s coalescent (Kingman 1982, 2000). Kingman’s coalescent has been incredibly helpful in the statistical analysis of genetic data (Wakeley 2008), one reason being its robustness: it arises as the limiting coalescent tree from a wide class of models of population evolution, including the Wright-Fisher and Moran processes (Möhle 1998). In this paper we look for a robustness result on the coalescent tree of a supercritical branching process.

Growing populations beginning with a single individual are found everywhere in biology. Examples are a virus spreading through a population, an animal colonising a new habitat, or a growing population of cancer cells. Such populations are naturally and
commonly modelled by supercritical branching processes. However there is a question mark on the realism of supercritical branching processes. Although exponential growth is often realistic while a population is small, exponential growth is never indefinitely realistic. Instead, due to competition for limited resources, populations see their growth slow down as they become larger.

For greater biological realism, density-dependent branching processes have been studied (Klebaner 1984; Höpfner 1985; Jagers and Klebaner 2000) where individuals reproduce and die at rates which depend on the population size. Branching processes with logistic growth, where the per-individual death rate is linear with the population size, are especially popular (Lambert 2005; Campbell 2003). The downside to density-dependence is of course greatly increased complexity of analysis. Coalescent trees from density-dependent branching processes are apparently unknown.

In this paper we consider a class of density-dependent branching processes which includes exponential, logistic and Gompertz growth models as special cases. Our focus is the continuous-time Markov setting. We refer to the class of models as a ‘branching process with generalised logistic growth’ or a ‘generalised logistic branching process’. The assumptions defining this class are minimal, namely that the population begins with one individual, that the offspring distribution has a uniformly bounded second moment, and that growth is exponential while the population size is small relative to some carrying capacity. Growth may be progressively constrained as the population size moves towards its carrying capacity. At a time while growth is still superlinear, a fixed number of individuals are sampled from the population and their coalescent tree is drawn. We consider the limit as the sampling time and carrying capacity simultaneously tend to infinity. We prove that the coalescent tree of a generalised logistic branching process coincides, in the limit, with the coalescent tree of a density-independent supercritical branching process, whose structure was recently determined (Johnston 2019).

Our result can be viewed as a universality or robustness result for the coalescent tree of a supercritical branching process, where the robustness is with respect to changes in the reproduction law while the population is large. However the tree crucially depends upon the reproduction law while the population is small, whose parameter space is infinite-dimensional. Thus the ‘supercritical coalescent’ is necessarily robust in a weaker sense than Kingman’s coalescent.

The paper is organised as follows. In Sect. 2, we introduce the class of density-dependent branching processes and present examples. In Sect. 3, we state convergence of the coalescent tree and describe its limiting form with reference to recent branching process coalescence works. In Sect. 4 through to 7, we prove convergence of the coalescent tree. In Sect. 4, we couple the population genealogies of a generalised logistic branching process and a supercritical branching process. In Sect. 5, we show that genetic drift is negligible while the population is large and growing superlinearly, and hence the coupled population genealogies asymptotically coincide with respect to a certain representation. In Sect. 6, we translate the viewpoint from the population level to the sample level. In Sect. 7, we show convergence of the sampled individuals’ ancestral identities and coalescence times.
2 A generalised logistic branching process

2.1 Reproduction rates

The population comprises an integer number of individuals. The individuals reproduce and die at rates which depend on the population size and on some index $\kappa \in \mathbb{N}$, whose purpose is that taking $\kappa \to \infty$ will represent the large carrying capacity limit. For a given $\kappa \in \mathbb{N}$ and while the population size is $n \in \mathbb{N}$, each individual in the population is replaced by $i \in \mathbb{N} \cup \{0\}$ individuals at rate $\Lambda_i^{\kappa,n} \geq 0$. We shall soon describe the population evolution in more detailed language. First we state assumptions on the $\Lambda_i^{\kappa,n}$.

Assumption 2.1 The reproduction rate distribution has a uniformly bounded second moment: $\sup_{\kappa,n} \sum_i i^2 \Lambda_i^{\kappa,n} < \infty$.

Assumption 2.2 There are ‘intrinsic’ reproduction rates which give positive growth. That is, there are $\Lambda_i^\infty \geq 0$ for $i \in \mathbb{N} \cup \{0\}$ with

$$\sum_i (i - 1) \Lambda_i^\infty > 0,$$

and there exists $(n_\kappa)_{\kappa \in \mathbb{N}} \in \mathbb{N}^\mathbb{N}$ with

$$\lim_{\kappa \to \infty} n_\kappa = \infty \quad \text{and} \quad \lim_{\kappa \to \infty} \sup_{n \leq n_\kappa} \sum_{i=0}^{\infty} |\Lambda_i^{\kappa,n} - \Lambda_i^\infty| = 0.$$  

Assumption 2.2 defines the sense in which $\kappa \to \infty$ is the large carrying capacity limit, which will be illuminated by examples in Sect. 2.4.

2.2 Genealogical structure

Following standard notation, individuals in the population are denoted by elements of

$$\mathcal{T} = \cup_{m \in \mathbb{N} \cup \{0\}} \mathbb{N}^m.$$  

The set $\mathcal{T}$ has a partial ordering $<$ defined by $(u_i)_{i=1}^m < (v_i)_{i=1}^n$ if and only if $m < n$ and $(u_i)_{i=1}^m = (v_i)_{i=1}^m$. We write $\preceq$ for $<$ or $=$. The possible daughters of $u = (u_1, \ldots, u_m) \in \mathcal{T}$ are denoted $ui = (u_1, \ldots, u_m, i)$ for $i \in \mathbb{N}$.

For each $\kappa \in \mathbb{N}$ the population evolution is described by a continuous-time Markov process, $(\mathcal{N}_\kappa(t))_{t \geq 0}$, whose state space is the finite subsets of $\mathcal{T}$. The process has initial condition

$$\mathcal{N}_\kappa(0) = \{\emptyset\} = \mathbb{N}^0.$$
and sees transitions
\[
\mathcal{N}'_k(t) \mapsto \mathcal{N}_k(t) \cup \left( \bigcup_{i=1}^{j} \{ui\} \right) \setminus \{u\} \quad \text{at rate } \Lambda_j^{k, \mathcal{N}_k(t)} I\{u \in \mathcal{N}_k(t)\}
\]
for \( j \in \mathbb{N} \cup \{0\} \), where \( \bigcup_{i=1}^{0} \{ui\} = \emptyset \) and \( I \) is the indicator function. We shall write
\[
N_k(t) = \#\mathcal{N}_k(t)
\]
for the population size at time \( t \geq 0 \). Assumption 2.1 ensures that \( N_k(t) \) does not explode in finite time, so \( \mathcal{N}_k(t) \) is always well defined.

2.3 Sampling time

The sampling time \( T_k \) is a positive random variable which satisfies the following.

**Assumption 2.3** The sampling time is a stopping time with respect to the filtration \( (\mathcal{F}_k(t))_{t \geq 0} \) where \( \mathcal{F}_k(t) \) is the sigma-algebra generated by \( (\mathcal{N}_k(s))_{s \in [0, t]} \).

**Assumption 2.4** The sampling time is large: \( \lim_{k \to \infty} T_k = \infty \) in distribution.

**Assumption 2.5** Growth is superlinear before the sampling time: there exist \( \alpha \in (0, 1) \) and \( c > 0 \) such that
\[
\lim_{k \to \infty} \mathbb{P}\left[ \forall t \in [0, T_k], \tilde{\Lambda}^{N, \mathcal{N}_k(t)} \geq \frac{c}{N_k(t)^{\alpha}} \right] = 1,
\]
where \( \tilde{\Lambda}^{N, \mathcal{N}_k(t)} = \sum_{i=0}^{\infty} (i - 1) \Lambda_i^{k, \mathcal{N}_k(t)} \) is the per-individual growth rate. Assumption 2.5 can be made sense of by reference to a deterministic growth model. Consider a population of size \( n(t) \geq 0 \) at time \( t \geq 0 \), with \( n(0) = 1 \) and \( \frac{dn}{dt} \geq cn^{1-\alpha} \). Then \( n(t) \geq (1 + c\alpha t)^{1/\alpha} \), which is superlinear.

2.4 Examples

**Example 2.6** Exponential growth: \( \Lambda_i^{k,n} = \Lambda_i^{\infty} \), \( T_k = \kappa \).

**Example 2.7** Logistic growth with death due to competition:
\[
\Lambda_i^{k,n} = \begin{cases} 
\Lambda_0^{\infty} + \frac{n}{k} \sum_{j=0}^{\infty} (j - 1) \Lambda_j^{\infty}, & i = 0, \\
\Lambda_i^{\infty}, & i \geq 1,
\end{cases}
\]
and \( T_k = \min\{t \geq 0 : N_k(t) \geq x\kappa\} \) for some \( x \in (0, 1) \). Note the corresponding deterministic growth model: \( \frac{dn}{dt} = \bar{\Lambda}^{\infty} n \left( 1 - \frac{n}{k} \right) \), where \( \bar{\Lambda}^{\infty} = \sum_{j=0}^{\infty} (j - 1) \Lambda_j^{\infty} \) is the per-individual growth rate in the absence of a carrying capacity.
Example 2.8 Gompertz growth with death due to competition:

\[
\Lambda_i^{k,n} = \begin{cases} 
\Lambda_0^\infty + \frac{\log(n)}{\log(\kappa)} \sum_{j=0}^\infty (j-1) \Lambda_j^\infty , & i = 0, \\
\Lambda_i^\infty , & i \geq 1,
\end{cases}
\]

and \(T_k = \min\{t \geq 0 : \log(N_k(t)) \geq x \log(\kappa)\}\) for some \(x \in (0, 1)\). Note the corresponding deterministic growth model: \(\frac{dn}{dt} = \tilde{\Lambda}^\infty n \left(1 - \frac{\log(n)}{\log(\kappa)}\right)\).

Example 2.9 Logistic growth with reduced reproduction due to competition:

\[
\begin{align*}
\Lambda_0^{k,n} &= \Lambda_0^\infty, \\
\Lambda_2^{k,n} &= \Lambda_2^\infty + \frac{2}{\kappa} \left(-\Lambda_2^\infty + \Lambda_0^\infty\right), \\
\Lambda_i^{k,n} &= \Lambda_i^\infty = 0, & i \notin \{0, 2\},
\end{align*}
\]

and \(T_k = \min\{t \geq 0 : N_k(t) \geq x\kappa\}\) for some \(x \in (0, 1)\).

Example 2.10 Exponential-to-polynomial growth with reduced reproduction due to competition:

\[
\begin{align*}
\Lambda_0^{k,n} &= \Lambda_0^\infty, \\
\Lambda_2^{k,n} &= \Lambda_0^\infty + \frac{\kappa}{\kappa + n^a} (\Lambda_2^\infty - \Lambda_0^\infty), \\
\Lambda_i^{k,n} &= \Lambda_i^\infty = 0, & i \notin \{0, 2\},
\end{align*}
\]

for some \(a \in [0, 1)\). Here \(T_k\) may be anything which converges to infinity.

Remark 2.11 For Examples 2.7, 2.8 and 2.9, \(\kappa\) is the population’s carrying capacity in the sense that the population has zero growth rate while the population size is equal to \(\kappa\).

3 Coalescent tree

Let \(m \in \mathbb{N}\). On the event \(\{N_k(T_k) \geq m\}\) that at least \(m\) individuals are alive at time \(T_k\), let \(U_1^k, \ldots, U_m^k\) be a uniform sample without replacement of \(m\) individuals from the population \(N_k(T_k)\). For \(t \geq 0\) and \(\kappa \in \mathbb{N}\), let \(\sim_{t,\kappa}\) be the equivalence relation on \(\{1, \ldots, m\}\) defined by

\[i \sim_{t,\kappa} j \iff \exists u \in N_k(t), u \leq U_i^k \text{ and } u \leq U_j^k.\]

Let \(\pi_k(t)\) be the partition of \(\{1, \ldots, m\}\) induced by \(\sim_{t,\kappa}\). That is, \(i, j \in \{1, \ldots, m\}\) are in the same block of \(\pi_k(t)\) if and only if \(U_i^k\) and \(U_j^k\) share an ancestor at time \(t\). The coalescent tree is the partition-valued process \((\pi_k(t))_{t \geq 0}\). Representing coalescent trees as partition-valued processes is standard (Berestycki 2009). Less standard is that we view the tree forwards in time, with elements
of the partition breaking up. This contrasts with the traditional view of coalescent trees backwards in time, where elements of the partition eponymously coalesce.

### 3.1 Exponential growth

Before stating our main result which is convergence of the coalescent tree, we describe the limiting coalescent tree with reference to recent results for density-independent branching processes.

The coalescent trees of branching processes have been extensively studied (Popovic 2004; Lambert and Stadler 2013; Lambert and Popovic 2013; Bertoin and Le Gall 2000; Hong 2013; Athreya and Hong 2018; Athreya 2012; Harris et al. 2020; Lambert 2018; Johnston and Lambert 2019). In particular, we note that Johnston (2019) recently gave a broad account of the coalescent tree arising from a continuous-time Galton-Watson branching process, including a limit result for the supercritical regime - which is our interest here. The coalescent tree determined by Johnston requires a little more notation. Let 

\[
(N_\infty(t))_{t \geq 0} \text{ be a continuous-time Markov process on } \mathbb{N} \cup \{0\} \text{ with transitions } N_\infty(t) \mapsto N_\infty(t) + i \text{ at rate } N_\infty(t) \Lambda^\infty_i \text{ for } i \in \mathbb{N} \cup \{0\}, \text{ and with initial condition } N_\infty(0) = 1. \text{ In words, } N_\infty(t) \text{ is the population size if the carrying capacity were infinite. The per-individual growth rate is denoted } \bar{\Lambda}^\infty = \sum_i (i - 1) \Lambda^\infty_i. \]

Let 

\[
H^j_t(s) = \frac{d}{ds} \mathbb{E} N_\infty(t)^s \text{ for } s, t \geq 0 \text{ and } j \in \mathbb{N} \cup \{0\}. \text{ Let } \psi(j) = \frac{d}{dv} \mathbb{E} e^{-v W} \text{ for } v \geq 0 \text{ and } j \in \mathbb{N} \cup \{0\}, \text{ where } W = \lim_{t \to \infty} e^{-\bar{\Lambda}^\infty t} N_\infty(t) \text{ (that the limit } W \in [0, \infty) \text{ exists almost surely is a classic branching process result (Athreya and Ney 2004). Finally let } (\pi_\infty(t))_{t \geq 0} \text{ be a process on the partitions of } \{1, \ldots, m\}, \text{ distributed as }
\]

\[
\mathbb{P}[\pi_\infty(t_1) = \gamma_1, \ldots, \pi_\infty(t_r) = \gamma_r] = \frac{(-1)^m e^{-m \bar{\Lambda}^\infty t_r}}{1 - \psi(\infty)} \int_0^\infty \frac{v^{m-1}}{(m-1)!} \prod_{i=0}^{r-1} \prod_{\Gamma \in \gamma_i} H^{b_i(\Gamma)}_{t_i-t_{i-1}} \left( \psi(e^{-\bar{\Lambda}^\infty t_{i+1} v}) \right) \prod_{\Gamma \in \gamma_r} \psi^\#(e^{-\bar{\Lambda}^\infty t_r v}) \, dv, \quad (1)
\]

where \(0 < t_1 < \ldots < t_r\); and \(\gamma_1, \ldots, \gamma_r\) are partitions of \(\{1, \ldots, m\}\) such that each \(\Gamma \in \gamma_i\) is the union of \(b_i(\Gamma) \geq 1\) elements of \(\gamma_{i+1}\), for \(i = 1, \ldots, r - 1\).

**Theorem 3.1** (Theorem 3.5 of Johnston (2019)) Consider Example 2.6, which is a density-independent supercritical branching process. As \(\kappa \to \infty\), \(\pi_\kappa\) on the event \(\{N_\kappa(T_\kappa) \geq m\}\) converges in finite dimensional distributions to \(\pi_\infty\).

**Remark 3.2** The coalescent tree \(\pi_\kappa\) is characterised by the coalescence times

\[
\tau_{\kappa}^{i,j} = \sup\{t \geq 0 : \exists u \in N_\kappa(t), u \leq U^i_k \text{ and } u \leq U^j_k\}, \quad 1 \leq i < j \leq m,
\]
and convergence of the coalescent tree is equivalent to convergence of the coalescence times.

While the distribution of the coalescent tree (1) may seem tricky at first glance, matters simplify considerably in the case of a binary branching process: $\Lambda_i^\infty = 0$ for $i \notin \{0, 2\}$. Here the coalescent tree is binary too and there are exactly $m - 1$ distinct coalescence times. We denote the unordered coalescence times, in the limit, as $\sigma_1^\infty, \ldots, \sigma_{m-1}^\infty$. The distribution of the $\sigma_i^\infty$ can be obtained from (1), or more quickly from Harris, Johnston and Roberts’ Theorem 1 (Harris et al. 2020). Taking the large time limit of their Theorem 1 in the supercritical case gives that the coalescence times are distributed as

$$
\Pr[\sigma_1^\infty \geq t_1, \ldots, \sigma_{m-1}^\infty \geq t_{m-1}] = m \prod_{i=1}^{m-1} \frac{e^{-\tilde{\Lambda}^\infty t_i}}{1 - e^{-\tilde{\Lambda}^\infty t_i}} + \sum_{j=1}^{m-1} \tilde{\Lambda}^\infty t_j e^{-\tilde{\Lambda}^\infty t_j} \left( \prod_{i=1}^{m-1} \frac{e^{-\tilde{\Lambda}^\infty t_i}}{1 - e^{-\tilde{\Lambda}^\infty t_i}} \right) \left( \prod_{i=1, i \neq j}^{m-1} \frac{e^{-\tilde{\Lambda}^\infty t_i}}{1 - e^{-\tilde{\Lambda}^\infty t_i}} \right)^2
$$

for $t_1, \ldots, t_{m-1} > 0$, and the coalescence times are independent of the tree topology which is that of Kingman’s coalescent (viewed backwards in time, each coalescence event involves a uniformly chosen pair of partition elements).

Lambert (2018) also determined the coalescent tree for a density-independent binary branching process. He elegantly described the coalescent tree in terms of a mixture of coalescent point processes, giving an effectively instantaneous method of tree simulation. Taking the large time limit of his result in the supercritical case recovers (2).

### 3.2 Generalised logistic growth

The paper’s main result generalises Theorem 3.1 to a branching process with generalised logistic growth.

**Theorem 3.3** As $\kappa \to \infty$, $\pi_\kappa$ on the event $\{N_\kappa(T_\kappa) \geq m\}$ converges in finite dimensional distributions to $\pi_\infty$, where $\pi_\infty$ is distributed as (1). That is, the limiting coalescent tree is parameterised by the intrinsic reproduction rates.

Theorem 3.3 serves two purposes. First, it demonstrates that the coalescent tree of a supercritical Markov branching process is robust with respect to deviations from exponential growth at large sizes, suggesting that the tree may be applied to genetic data from populations with a variety of growth laws. Second, it is to our knowledge a first complete description of the coalescent tree from a density-dependent branching process. Hopefully the result and proof will inspire more general explorations of coalescence in density-dependent branching processes in the future.
There are open questions. What if individuals are sampled after a carrying capacity is reached? What if the intrinsic offspring distribution is critical, sub-critical, or heavy-tailed? What if reproduction and death rates depend on the ages of individuals? Regarding the third question we note that there is an extensive history of research on age-dependent, density-independent branching processes (see Chapter 4 of Athreya and Ney (2004) for an introduction) and there are partial results on their coalescent trees (Lambert and Stadler 2013; Hong 2013; Athreya and Hong 2018). To expand coalescence theory to branching processes with both age-dependence and density-dependence (such as the model in Jagers and Klebaner (2000)) would do a great service for the aim of biological realism.

The remainder of the paper is concerned with Theorem 3.3’s proof.

4 Coupling with exponential growth

At the base of the proof is a coupling between generalised logistic growth and exponential growth. Exponential growth is described by \( \langle N_\infty(t) \rangle_{t \geq 0} \), a continuous-time Markov process whose state space is the finite subsets of \( T \). The process sees transitions

\[
N_\infty(t) \mapsto N_\infty(t) \cup \left( \bigcup_{i=1}^{j} \{ ui \} \right) \setminus \{ u \}
\]

at rate \( \Lambda \sum_{j \in \mathbb{N} \cup \{0\}} \mathbb{I}_{\{ u \in N_\infty(t) \}} \) for \( j \in \mathbb{N} \cup \{0\} \) and \( u \in T \). The initial condition is

\[
N_\infty(0) = \{ \emptyset \} = \mathbb{N}^0.
\]

Let

\[
N_\infty(t) = \# N_\infty(t)
\]

be the exponentially growing population size. Now put the \( N_\kappa(\cdot) \) for \( \kappa \in \mathbb{N} \cup \{ \infty \} \) on a single probability space. We define their joint distribution by the following.

1. The \( N_\kappa(\cdot) \), \( \kappa \in \mathbb{N} \), are conditionally independent given \( N_\infty(\cdot) \).
2. The \( \begin{pmatrix} N_\infty(t) \\ N_\kappa(t) \end{pmatrix} \) transitions to
The coalescent tree of a Markov branching process with ... Page 9 of 22

\begin{align*}
\begin{cases}
\left(N_\infty(t) \cup \bigcup_{i=1}^j \{ui\} \right) \setminus \{u\} \\
\mathcal{N}_k(t) \cup \left(\bigcup_{i=1}^j \{ui\} \right) \setminus \{u\}
\end{cases}
\text{at rate } \min \{\Lambda_\infty^\infty, \Lambda_j^{\kappa, \mathcal{N}_k(t)}\} I\{u \in N_\infty(t) \cap N_k(t)\}.
\end{align*}

\begin{align*}
\begin{cases}
\left(N_\infty(t) \cup \bigcup_{i=1}^j \{ui\} \right) \setminus \{u\} \\
\mathcal{N}_k(t)
\end{cases}
\text{at rate } (\Lambda_\infty^\infty - \Lambda_j^{\kappa, \mathcal{N}_k(t)}) I\{\Lambda_j^\infty > \Lambda_j^{\kappa, \mathcal{N}_k(t)}, u \in N_\infty(t) \cap N_k(t)\}
\begin{align*}
&+ \Lambda_j^\infty I\{u \in N_\infty(t) \setminus \mathcal{N}_k(t)\},
\end{align*}
\begin{align*}
\begin{cases}
N_\infty(t) \\
\mathcal{N}_k(t) \cup \left(\bigcup_{i=1}^j \{ui\} \right) \setminus \{u\}
\end{cases}
\text{at rate } (\Lambda_j^{\kappa, \mathcal{N}_k(t)} - \Lambda_j^\infty) I\{\Lambda_j^{\kappa, \mathcal{N}_k(t)} > \Lambda_j^\infty, u \in N_\infty(t) \cap N_k(t)\}
\begin{align*}
&+ \Lambda_j^{\kappa, \mathcal{N}_k(t)} I\{u \in \mathcal{N}_k(t) \setminus N_\infty(t)\},
\end{align*}
\end{align*}

for \( j \in \mathbb{N} \cup \{0\} \) and \( u \in \mathcal{T} \).

The conditional independence of the \( \mathcal{N}_k(\cdot) \) will not be important. We only specify the conditional independence so that some joint distribution is specified. On the other hand, the joint transition rates of \( \mathcal{N}_k(\cdot) \) and \( \mathcal{N}_\infty(\cdot) \) will be helpful for the proof. This coupling maximises the time for which \( \mathcal{N}_k(\cdot) \) and \( \mathcal{N}_\infty(\cdot) \) remain equal. Initially \( \mathcal{N}_k(\cdot) \) and \( \mathcal{N}_\infty(\cdot) \) are equal, but their state of equality is broken at rate

\[ N_\infty(t) \sum_{i=0}^\infty |\Lambda_i^\infty - \Lambda_i^{\kappa, \mathcal{N}_\infty(t)}|. \]

Note that

\[ P[N_\infty(s) = \mathcal{N}_k(s), \forall s \in [0, t]] = E \left[ \exp \left( - \int_0^t N_\infty(s) \sum_{i=0}^\infty |\Lambda_i^\infty - \Lambda_i^{\kappa, \mathcal{N}_\infty(s)}| ds \right) \right]. \quad (3) \]

The next result considers (3) in the view of the large carrying capacity limit. It says that the exponential and generalised logistic growth models coincide for a long initial time period.

**Lemma 4.1** There exists a sequence of times \( (\zeta_\kappa)_{\kappa \in \mathbb{N}} \in (0, \infty)^\mathbb{N} \) with

\[ \lim_{\kappa \to \infty} \zeta_\kappa = \infty \quad \text{and} \quad \lim_{\kappa \to \infty} P[N_\infty(s) = \mathcal{N}_k(s), \forall s \in [0, \zeta_\kappa]] = 1. \]
Proof of Lemma 4.1 Recall from Assumption 2.2 that \((n_κ)\) is some sequence of integers. Let

\[
δ_κ = \sup_{n ≤ n_κ} \sum_{i=0}^{∞} |Λ_i^{κ,n} - Λ_i^{∞}|.
\]

Then on the event \(\left\{ \sup_{s ∈ [0, t]} N_∞(s) ≤ \min\{δ_κ^{-1/2}, n_κ\} \right\}\), we see the bound

\[
\int_0^t N_∞(s) \sum_{i=0}^{∞} |Λ_j^{∞} - Λ_j^{κ,N_∞(s)}| ds ≤ tδ_κ^{1/2}.
\]

It follows that we can bound (3) below by

\[
\mathbb{P}\left[ N_∞(s) = N_κ(s), \forall s ∈ [0, t] \right] = \mathbb{E}\left[ \exp\left( -\int_0^t N_∞(s) \sum_{i=0}^{∞} |Λ_j^{∞} - Λ_j^{κ,N_∞(s)}| ds \right) \right]
\geq \mathbb{E}\left[ \exp\left( -tδ_κ^{1/2} \right) \mathbb{I}\left\{ \sup_{s ∈ [0, t]} N_∞(s) ≤ \min\{δ_κ^{-1/2}, n_κ\} \right\} \right]
= \exp\left( -tδ_κ^{1/2} \right) \mathbb{P}\left[ \sup_{s ∈ [0, t]} N_∞(s) ≤ \min\{δ_κ^{-1/2}, n_κ\} \right]
\geq \exp\left( -tδ_κ^{1/2} \right) \left[ 1 - \frac{e^{L_∞t}}{\min\{δ_κ^{-1/2}, n_κ\}} \right],
\]

where the last inequality is Doob’s martingale inequality \(e^{-L_∞t}N_∞(t)\) is a martingale). But Assumption 2.2 says that \(δ_κ\) converges to zero while \(n_κ\) converges to infinity. Thus by replacing \(t\) with

\[
ζ_κ = \log\left( \frac{\min\{δ_κ^{-1/2}, n_κ\}}{2L_∞} \right),
\]

the lower bound (4) converges to 1.

The explicit definition of \(ζ_κ\) seen in Lemma 4.1’s proof won’t be used. Only the properties of \(ζ_κ\) which are seen in the Lemma’s statement will be important.

Now for each member of the population we look at the abundance of their descendants. Let

\[
D_κ^u(t) = \#\{v ∈ N_κ(t) : u ≤ v\}
\]
be the number of individuals alive at time \( t \geq 0 \) which are descendants of \( u \in \mathcal{T} \) for \( \kappa \in \mathbb{N} \cup \{\infty\} \). Let

\[
F^u_\kappa(t) = I\{N_\kappa(t) > 0\} \frac{D^u_\kappa(t)}{N_\kappa(t)}
\]

be the fraction of the population at time \( t \) which descended from \( u \). We will refer to the \( F^u_\kappa(t) \) as the descendant fractions. The descendant fractions will play a crucial role in Theorem 3.3’s proof because they will give the probabilities that sampled individuals have particular ancestors. The next result shows convergence of the descendant fractions in the exponential model.

**Lemma 4.2** For \( u \in \mathcal{T} \), the limit

\[
\lim_{t \to \infty} F^u_\infty(t) =: F^u_\infty(\infty) \in [0, 1]
\]

exists almost surely.

**Proof** For any \( j, \kappa, n \in \mathbb{N} \),

\[
\sum_{i=1}^j i \log(i) \Lambda^\infty_i \leq j \log(j) \sum_{i=1}^j |\Lambda^\infty_i - \Lambda^\kappa,n_i| + \sum_{i=1}^\infty i^2 \Lambda^\kappa,n_i.
\]

So by Assumption 2.2,

\[
\sum_{i=1}^j i \log(i) \Lambda^\infty_i \leq \limsup_{\kappa \to \infty} j \log(j) \sum_{i=1}^j |\Lambda^\infty_i - \Lambda^\kappa,n_i| + \sup_{\kappa, n} \sum_{i=1}^\infty i^2 \Lambda^\kappa,n_i
\]

\[
= \sup_{\kappa, n} \sum_{i=1}^\infty i^2 \Lambda^\kappa,n_i.
\]

Then by Assumption 2.1, \( \sum_{i=1}^\infty i \log(i) \Lambda^\infty_i \) is finite, which gives the Kesten-Stigum condition for the supercritical branching process \( N^\infty(\cdot) \) (see for example Athreya and Ney (2004) Chapter 3.7, Theorem 2). Therefore, recalling that \( \bar{\Lambda}^\infty = \sum_i (i - 1) \Lambda^\infty_i \) denotes the exponential growth rate,

\[
\lim_{t \to \infty} e^{-\bar{\Lambda}^\infty t} N^\infty(t) = W \in [0, \infty)
\]

exists almost surely with

\[
\{W > 0\} = \{N^\infty(t) > 0, \forall t\}.
\]

Let \( \tau = \inf\{t \geq 0 : u \in \mathcal{N}^\infty(t)\} \) be the first time at which \( u \) is alive and observe that, because \( \mathcal{N}^\infty(\cdot) \) is a branching process, the number of descendants of \( u \) follows the
same asymptotic behaviour as (5):

\[
\lim_{t \to \infty} I\{\tau < \infty\} e^{-\tilde{\lambda}^\infty_{\infty}(t-\tau)} D_{\infty}(t) = I\{\tau < \infty\} W' (6)
\]

almost surely, where \( W' \overset{d}{=} W \). Combining (5) and (6),

\[
\lim_{t \to \infty} F_{\infty}(t) = \lim_{t \to \infty} I\{\tau < \infty, N_{\infty}(t) > 0\} e^{-\tilde{\lambda}^\infty_{\infty}(t-\tau)} D_{\infty}(t)
\]

\[
= I\{\tau < \infty, W > 0\} e^{-\tilde{\lambda}^\infty_{\infty}} \frac{W'}{W}.
\]

Our final result of this section says that the descendant fractions viewed late during the exponential growth phase of the generalised logistic model converge, unsurprisingly, to the same limit as the descendant fractions in the exponential model.

**Lemma 4.3** For \( u \in \mathcal{T} \),

\[
\lim_{\kappa \to \infty} F_{\kappa}(\min(\xi_{\kappa}, T_{\kappa})) = F_{\infty}(\infty)
\]

in probability, where \( F_{\infty}(\infty) \) is defined in Lemma 4.2.

**Proof** Combine Lemmas 4.1 and 4.2.

This section has viewed the population’s initial exponential growth phase. The next section will look at later times when the population growth may be constrained.

### 5 Genetic drift

The quantity \( F_{\kappa}(T_{\kappa}) - F_{\kappa}(\min(\xi_{\kappa}, T_{\kappa})) \) can be thought of as ‘genetic drift’ on the time interval \( [\xi_{\kappa}, T_{\kappa}] \). In this section we will show that genetic drift is negligible. Towards this end, the next result bounds genetic drift by the integrated reciprocal of the population size. As an aside note that the Moran and Wright-Fisher processes see analogous behaviour, with genetic drift occurring at a rate proportional to the reciprocal of the population size.

**Lemma 5.1** Let \( \tilde{T}_{\kappa} \) be a stopping time with respect to \( (\mathcal{F}_{\kappa}(t))_{t \geq 0} \) (where \( \mathcal{F}_{\kappa}(t) \) is the sigma-algebra generated by \( (N_{\kappa}(s))_{s \in [0, t]} \)). There exists \( a > 0 \) such that for all \( \kappa \in \mathbb{N} \) and \( s \geq 0 \),

\[
\mathbb{E} \left[ (F_{\kappa}(\tilde{T}_{\kappa}) - F_{\kappa}(\min(s, \tilde{T}_{\kappa}))^2 \right] 
\leq a \int_{s}^{\infty} \mathbb{E} \left[ \frac{I\{N_{\kappa}(t) > 0, t < \tilde{T}_{\kappa}\}}{N_{\kappa}(t)} \right] dt.
\]
Proof We look at the process

\[ G(t) = F_u^\kappa(\min\{t, \tilde{T}_\kappa\}), \quad t \geq 0, \]

which sees the following transitions placed into three categories.

1. The mother \( u^- = (u_i)_{i=1}^{l-1} \) of \( u = (u_i)_{i=1}^l \) reproduces: for \( n, i > 0 \),

\[
\lim_{h \downarrow 0} h^{-1} \mathbb{P} \left[ G(t + h) = \frac{1}{n + i - 1} \left| N_\kappa(t) = n, t < \tilde{T}_\kappa, u^- \in N_\kappa(t) \right. \right] = \sum_{i=ul} \Lambda_i^{k,n}.
\]

2. A descendant of \( u \) (including \( u \)) reproduces or dies: for integers \( n \geq d > 0 \) and \( i \geq 0 \),

\[
\lim_{h \downarrow 0} h^{-1} \mathbb{P} \left[ G(t + h) = \frac{d + i - 1}{n + i - 1} \left| N_\kappa(t) = n, D_\kappa^u(t) = d, t < \tilde{T}_\kappa \right. \right] = \Lambda_i^{k,n} d
\]

(for \( d = n = 1, i = 0 \) in the above we use the convention \( 0/0 = 0 \)).

3. An individual who is not \( u \)’s mother nor a descendant of \( u \) reproduces or dies: for integers \( n > d > 0 \) and \( i \geq 0 \),

\[
\lim_{h \downarrow 0} h^{-1} \mathbb{P} \left[ G(t + h) = \frac{d}{n + i - 1} \left| N_\kappa(t) = n, D_\kappa^u(t) = d, t < \tilde{T}_\kappa \right. \right] = \Lambda_i^{k,n} (n - d).
\]

The transition rates give that

\[
\frac{d}{dt} \mathbb{E}[G(t)G(s)] = -\Lambda_0^{k,1} \mathbb{E}[I\{N_\kappa(t) = D_\kappa^u(t) = 1, t < \tilde{T}_\kappa\}G(s)]
\]

for \( t > s \), and

\[
\frac{d}{dt} \mathbb{E}[G(t)^2] = \sum_{i=ul} \mathbb{E} \left[ \Lambda_i^{k,N_\kappa(t)} I\{u^- \in N_\kappa(t), t < \tilde{T}_\kappa\} \right] \frac{(N_\kappa(t)+i-1)^2}{(N_\kappa(t)+i-1)^2} - \Lambda_0^{k,1} \mathbb{E}[I\{N_\kappa(t) = D_\kappa^u(t) = 1, t < \tilde{T}_\kappa\}]
\]
\[ + \sum_{i=0}^{\infty} \mathbb{E} \left[ \frac{\Lambda_i^{\kappa,N_k(t)}(i-1)(1-F_k^u(t))(1-F_k^u(t))N_k(t)}{(N_k(t) + i - 1)^2} \times I\{N_k(t) > D_k^u(t) > 0, t < \tilde{T}_k\} \right]. \]

It follows that for \( t \geq s \),

\[
\frac{d}{dt} \mathbb{E} \left[ (G(t) - G(s))^2 \right] \leq \sum_{i=1}^{\infty} \mathbb{E} \left[ \frac{\Lambda_i^{\kappa,N_k(t)} I\{u^- \in N_k(t), t < \tilde{T}_k\}}{(N_k(t) + i - 1)^2} \right] + \Lambda_0^{\kappa,1} \mathbb{E} \left[ I\{N_k(t) = 1, t < \tilde{T}_k\} \right] + \sum_{i=0}^{\infty} \mathbb{E} \left[ \frac{\Lambda_i^{\kappa,N_k(t)}(i-1)^2N_k(t)}{(N_k(t) + i - 1)^2} I\{N_k(t) > 1, t < \tilde{T}_k\} \right] \leq a' \mathbb{E} \left[ I\{N_k(t) > 0, t < \tilde{T}_k\} \right] \sum_{i=0}^{\infty} \Lambda_i^{\kappa,N_k(t)}(i-1)^2 \leq a \mathbb{E} \left[ I\{N_k(t) > 0, t < \tilde{T}_k\} \right],
\]

where \( a', a > 0 \) are constants independent of \( t \) and \( \kappa \) with \( a \) coming from Assumption 2.1. Finally

\[
\mathbb{E} \left[ (G(\infty) - G(s))^2 \right] = \int_s^\infty \frac{d}{dt} \mathbb{E} \left[ (G(t) - G(s))^2 \right] dt \leq a \int_s^\infty \mathbb{E} \left[ \frac{I\{N_k(t) > 0, t < \tilde{T}_k\}}{N_k(t)} \right] dt,
\]

which is the result. \( \square \)

For the next result, let

\[ M_\kappa = \{ n \in \mathbb{N} : \tilde{\Lambda}_n^{\kappa,n} \geq cn^{-\alpha} \} \]

be the set of population sizes for which growth is superlinear (where \( \alpha \in (0, 1) \) and \( c > 0 \) are seen in Assumption 2.5), and let

\[ S_\kappa = \min \{ t \geq 0 : N_k(t) \notin M_\kappa \} \quad (7) \]

be the first time at which the bound for superlinear growth is broken or the population becomes extinct.

\( \square \) Springer
Lemma 5.2 There exists $b > 0$ such that for all $\kappa \in \mathbb{N}$ and $t > 0$,

$$
E \left[ \frac{I \{ t < S_\kappa \}}{N_\kappa(t)} \right] \leq bt^{-1/\alpha}.
$$

Proof Observe that

$$
I \{ n \in M_\kappa \} \sum_i \frac{(i-1)\Lambda_i^{\kappa,n}}{n+i-1}
$$

$$
\geq n^{-1} I \{ n \in M_\kappa \} \sum_i (i-1)\Lambda_i^{\kappa,n} - \left| \sum_i \Lambda_i^{\kappa,n} \frac{(i-1)}{n+i-1} - \sum_i \Lambda_i^{\kappa,n} \frac{(i-1)}{n} \right|
$$

$$
= n^{-1} I \{ n \in M_\kappa \} \sum_i (i-1)\Lambda_i^{\kappa,n} - \sum_i \Lambda_i^{\kappa,n} \frac{(i-1)^2}{n(n+i-1)}
$$

$$
\geq cn^{-1-\alpha} - dn^{-2}
$$

for all $\kappa$ and $n$, where $d > 0$ is some constant which exists by Assumption 2.1. So

$$
\inf_\kappa \left( I \{ n \in M_\kappa \} \sum_i \frac{(i-1)\Lambda_i^{\kappa,n}}{n+i-1} \right) \geq cn^{-1-\alpha} - dn^{-2} \quad (8)
$$

The left hand side of (8) is positive for all $n$ while the right hand side approaches $cn^{-1-\alpha}$ for large $n$. Hence there exists $\beta > 0$ such that

$$
\inf_\kappa \left( I \{ n \in M_\kappa \} \sum_i \frac{(i-1)\Lambda_i^{\kappa,n}}{n+i-1} \right) \geq \beta n^{-1-\alpha}. \quad (9)
$$

Now

$$
\frac{d}{dt} \mathbb{E} \left[ \frac{I \{ t < S_\kappa \}}{N_\kappa(t)} \right]
$$

$$
= -\sum_{i=0}^\infty \mathbb{E} \left[ \Lambda_i^{N_\kappa(t),\kappa} I \{ t < S_\kappa \} \right. 

\times \left( \frac{I \{ N_\kappa(t) + i - 1 \notin M_\kappa \} (i-1) N_\kappa(t) + i - 1 + I \{ N_\kappa(t) + i - 1 \in M_\kappa \} \right) 

\leq -\mathbb{E} \left[ I \{ t < S_\kappa \} \sum_{i=0}^\infty \frac{(i-1)\Lambda_i^{\kappa,N_\kappa(t)}}{N_\kappa(t) + i - 1} \right]
$$

$$
\leq -\beta \mathbb{E} \left[ \frac{I \{ t < S_\kappa \}}{N_\kappa(t)^{1+\alpha}} \right]. \quad (10)
$$
where (10) is due to (9). Next using Jensen’s inequality, (10) becomes

$$\frac{d}{dt} \mathbb{E} \left[ I\{ t < S_\kappa \} \right] \leq -\beta \mathbb{E} \left[ I\{ t < S_\kappa \} \right]^{1+\alpha}$$  \hspace{1cm} (11)$$

which, combined with the initial condition

$$\mathbb{E} \left[ \frac{I\{ t < S_\kappa \} \} {N_\kappa(t)} \right] \leq 1,$$

implies that

$$\mathbb{E} \left[ \frac{I\{ t < S_\kappa \} \} {N_\kappa(t)} \right] \leq r(t)$$

where \( \frac{d}{dt} r(t) = -\beta r(t)^{1+\alpha} \) and \( r(0) = 1 \). But \( r(t) = (1 + \alpha \beta t)^{-1/\alpha} \) so we are done. \( \Box \)

Lemma 5.1 says that genetic drift is bounded by the expected reciprocal of the population size, integrated over time. Lemma 5.2 says that the expected reciprocal of the population size decays faster than the reciprocal of time provided that the growth rate is superlinear. Assumption 2.5 says that the growth rate is indeed superlinear before the sampling time. The next result ties these strands together, saying that genetic drift during the time interval \([\zeta_\kappa, T_\kappa]\) is negligible.

**Lemma 5.3** For \( u \in \mathcal{T} \),

$$\lim_{\kappa \to \infty} \left( F^u_\kappa(T_\kappa) - F^u_\kappa(\min\{\zeta_\kappa, T_\kappa\}) \right) = 0$$

in probability.

**Proof** Let \( \tilde{T}_\kappa = \min\{S_\kappa, T_\kappa\} \). By Lemmas 5.1 and 5.2,

\[
\mathbb{E} \left[ (F_\kappa(\tilde{T}_\kappa) - F^u_\kappa(\min\{\zeta_\kappa, \tilde{T}_\kappa\}))^2 \right] \\
\leq a \int_\zeta \mathbb{E} \left[ I\{N_\kappa(t) > 0, t < \tilde{T}_\kappa \} \right] dt \\
\leq a \int_\zeta \mathbb{E} \left[ \frac{I\{ t < S_\kappa \} \} {N_\kappa(t)} \right] dt \\
\leq ab \int_\zeta t^{-1/\alpha} dt \\
\to 0
\]

(12)
as $\kappa \to \infty$. Then

$$
|F_k^u(T_\kappa) - F_k^u(\min\{\zeta_\kappa, T_\kappa\})| \\
= |F_k^u(T_\kappa) - F_k^u(\min\{\zeta_\kappa, T_\kappa\})|I\{T_\kappa > S_\kappa\} \\
+ |F_k^u(\bar{T}_\kappa) - F_k^u(\min\{\zeta_\kappa, \bar{T}_\kappa\})|I\{T_\kappa \leq S_\kappa\} \\
\leq I\{T_\kappa > S_\kappa\} + |F_k^u(\bar{T}_\kappa) - F_k^u(\min\{\zeta_\kappa, \bar{T}_\kappa\})|
$$

converges to zero in probability by Assumption 2.5 and (12).

**Proposition 5.4** For $u \in T$,

$$
\lim_{\kappa \to \infty} F_k^u(T_\kappa) = F_\infty^u(\infty)
$$

in probability.

**Proof** Combine Lemmas 4.3 and 5.3.

Proposition 5.4 is at the heart of Theorem 3.3’s proof. It says that the generalised logistic and exponential models see the same limiting descendant fractions. The descendant fractions view a particular aspect of the population’s genealogical structure. The next section translates our viewpoint from the population level to the sample level.

**6 Sampling**

We will sample from the population in a way which makes use of the fact that the descendant fractions converge in probability. Our sampling method will ultimately ensure that the coalescent tree converges in probability too. The sampling structure is defined via subintervals of $[0, 1]$.

Let

$$
\Theta_k^\emptyset(t) = \left[0, F_k^\emptyset(t)\right) \subset [0, 1)
$$

and then recursively for $u \in T$,

$$
\Theta_k^{ui}(t) = \left[\inf \Theta_k^u(t) + \sum_{i=1}^{j-1} F_k^{ui}(t), \inf \Theta_k^u(t) + \sum_{i=1}^{j} F_k^{ui}(t)\right) \subset [0, 1),
$$

for $j \in \mathbb{N}$, $\kappa \in \mathbb{N} \cup \{\infty\}$, and $t \in [0, \infty]$.

The $\Theta_k^u(t)$ have some important properties. First, $|\Theta_k^u(t)| = F_k^u(t)$. Second, $u \leq v$ implies $\Theta_k^u(t) \supset \Theta_k^v(t)$. Third, on the event $\{N_\kappa(t) > 0\}$ and for $s \leq t$, $\{\Theta_k^u(t) : u \in \mathcal{N}_\kappa(s)\}$ is a partition of $[0, 1)$. In other words, the $\Theta_k^u(t)$ provide a ‘stick-breaking’ view of the ancestral structure of all individuals alive at time $t$: the size of the interval corresponding to $u$ matches the fraction of living individuals at time $t$ which
descended from $u$, and this interval is partitioned by the intervals of $u$’s children (unless $u$ herself is alive at time $t$).

Let $X^1, \ldots, X^m$ be i.i.d. uniform random variables on $[0, 1)$ independently of everything else. On the event \{\(N_\kappa(T_\kappa) > 0\)\}, let $V^i_\kappa$ be the unique element of $\mathcal{N}_\kappa(T_\kappa)$ such that $X^i \in \Theta^{V^i_\kappa}_\kappa(T_\kappa)$. On the event of population extinction \{\(N_\kappa(T_\kappa) = 0\)\}, let $V^i_\kappa = \Delta$, where $\Delta \notin \mathcal{T}$ is a ghost element introduced for technical convenience.

According to the definition above $(V^i_\kappa)^m_{i=1}$ is, on the event

$$E_\kappa = \{N_\kappa(T_\kappa) > 0\},$$

a uniform sample from $\mathcal{N}_\kappa(T_\kappa)$ with replacement. Contrast this with the statement of Theorem 3.3 where $(V^i_\kappa)^m_{i=1}$ is, on the event \{\(N_\kappa(T_\kappa) \geq m\)\}, a uniform sample from $\mathcal{N}_\kappa(T_\kappa)$ without replacement. Note the connection: $(V^i_\kappa)^m_{i=1}$ conditioned on the event

$$\tilde{E}_\kappa = \{N_\kappa(T_\kappa) \geq m \text{ and } V^i_\kappa \neq V^j_\kappa, \forall i \neq j\}$$

is distributed as $(U^i_\kappa)^m_{i=1}$. The following result says that sampling with or without replacement are asymptotically equivalent to conditioning on

$$E_\infty = \{N_\infty(t) > 0, \forall t \geq 0\}.$$

**Lemma 6.1** \(\lim_{k \to \infty} I(E_\kappa) = \lim_{k \to \infty} I(\tilde{E}_\kappa) = I(E_\infty)\) in probability.

**Proof** The event \{\(F^\emptyset_\kappa(T_\kappa) = 1\)\} coincides with the event $E_\kappa$ for $\kappa \in \mathbb{N} \cup \{\infty\}$. Thus by Proposition 5.4,

$$\lim_{k \to \infty} I(E_\kappa) = I(E_\infty)$$

in probability. To see convergence of $I(\tilde{E}_\kappa)$ requires more work. A crucial missing ingredient is that the population size at sampling time is large or extinct; this may be intuitively clear but is so far unproven. Let $n \in \mathbb{N}$. Using Markov’s inequality

$$\mathbb{P}[1 \leq N_\kappa(T_\kappa) \leq n] = \mathbb{P} \left[ \frac{nI\{N_\kappa(T_\kappa) > 0\}}{N_\kappa(T_\kappa)} \geq 1 \right] \leq n\mathbb{E} \left[ \frac{I\{N_\kappa(T_\kappa) > 0\}}{N_\kappa(T_\kappa)} \right] \leq n\mathbb{P}[S_\kappa < T_\kappa] + n\mathbb{E} \left[ \frac{I\{N_\kappa(\min\{S_\kappa, T_\kappa\}) > 0\}}{N_\kappa(\min\{S_\kappa, T_\kappa\})} \right],$$

(13)
where $S_\kappa$ is the first time at which the growth rate breaks the superlinear lower bound, which was defined in (7). Adapt (11) to see that

$$
\left( \frac{I\{N_\kappa(\min\{t, S_\kappa, T_\kappa\}) > 0\}}{N_\kappa(\min\{t, S_\kappa, T_\kappa\})} \right)_{t \geq 0}
$$

is a supermartingale. Then (13) is bounded above by

$$
P[1 \leq N_\kappa(T_\kappa) \leq n] \leq nP[S_\kappa < T_\kappa] + nE \left[ \frac{I\{N_\kappa(\min\{\xi_\kappa, S_\kappa, T_\kappa\}) > 0\}}{N_\kappa(\min\{\xi_\kappa, S_\kappa, T_\kappa\})} \right]. \quad (14)
$$

Applying Assumption 2.5 and Lemma 4.1 to (14),

$$
\lim_{\kappa \to \infty} P[1 \leq N_\kappa(T_\kappa) \leq n] = 0.
$$

It follows that

$$
P[E_\kappa \setminus \tilde{E}_\kappa] = P[1 \leq N_\kappa(T_\kappa) < m]
$$

$$
+ \sum_{l=m}^{\infty} P[N_\kappa(T_\kappa) = l] \left( 1 - \frac{l(l-1)\ldots(l-m+1)}{l^m} \right)
$$

$$
\leq P[1 \leq N_\kappa(T_\kappa) < n] + \left( 1 - \frac{n(n-1)\ldots(n-m+1)}{n^m} \right)
$$

$$
\to 1 - \frac{n(n-1)\ldots(n-m+1)}{n^m}
$$

as $\kappa \to \infty$, for any integer $n \geq m$. Therefore $\lim_{\kappa \to \infty} P[E_\kappa \setminus \tilde{E}_\kappa] = 0$ and the proof is done. $\square$

### 7 Coalescence

The most recent common ancestor of $V^i_\kappa$ and $V^j_\kappa$ is defined as

$$
V^{i,j}_\kappa = \max\{u \in T : u \preceq V^i_\kappa \text{ and } u \preceq V^j_\kappa\}.
$$

for $\kappa \in \mathbb{N}$ and $i \neq j$, where the maximum is with respect to $\prec$ and $\max \emptyset = \Delta$. Equivalently,

$$
V^{i,j}_\kappa = \max\{u \in T : X^i \in \Theta^u_\kappa(T_\kappa), X^j \in \Theta^u_\kappa(T_\kappa)\}.
$$

We can extend the definition to $\kappa = \infty$ by setting

$$
V^{i,j}_\infty = \max\{u \in T : X^i \in \Theta^u_\infty(\infty), X^j \in \Theta^u_\infty(\infty)\}.
$$
Lemma 7.1 For \( i \neq j \),

\[
\lim_{\kappa \to \infty} \mathbb{P} \left[ V^i,j_\kappa = V^i,j_\infty \right] = 1.
\]

**Proof** By Proposition 5.4, the boundaries of the interval \( \Theta^u_\kappa(T_\kappa) \) converge in probability to the boundaries of \( \Theta^u_\infty(\infty) \). So

\[
I \{ u \leq V^i,j_\kappa \} = I \{ X^i, X^j \in \Theta^u_\kappa(T_\kappa) \}
\]

converges in probability to

\[
I \{ u \leq V^i,j_\infty \} = I \{ X^i, X^j \in \Theta^u_\infty(\infty) \}.
\]

The result follows. \( \square \)

The coalescence time of \( V^i_\kappa \) and \( V^j_\kappa \) is

\[
\tau^{i,j}_\kappa = \sup \{ t \in [0, T_\kappa] : V^{i,j}_\kappa \in \mathcal{N}_\kappa(t) \}
\]

for \( \kappa \in \mathbb{N} \cup \{ \infty \} \) and \( i \neq j \), where \( \sup \emptyset = \infty \).

Lemma 7.2 For \( i \neq j \),

\[
\mathbb{P} \left[ \tau^{i,j}_\infty < \infty | \forall t, N_\infty(t) > 0 \right] = 1.
\]

**Proof** On the event \( \{ N_\infty(\kappa) > 0 \} \), let \( \tilde{V}^i_\kappa \) and \( \tilde{V}^j_\kappa \) be the unique elements of \( \mathcal{N}_\infty(\kappa) \) such that \( X^i \in \Theta^{\tilde{V}^i}_\infty(\kappa) \) and \( X^j \in \Theta^{\tilde{V}^j}_\infty(\kappa) \) respectively; meanwhile on the event \( \{ N_\infty(\kappa) = 0 \} \), let \( \tilde{V}^i_\kappa = \tilde{V}^j_\kappa = \Delta \). That is, \( \tilde{V}^i_\kappa \) and \( \tilde{V}^j_\kappa \) are randomly sampled with replacement from the exponential population at time \( \kappa \). Their most recent common ancestor is

\[
\tilde{\tau}^{i,j}_\kappa = \max \{ u \in \mathcal{T} : u \leq \tilde{V}^i_\kappa \text{ and } u \leq \tilde{V}^j_\kappa \}
\]

\[
= \max \{ u \in \mathcal{T} : X^i, X^j \in \Theta^u_\infty(\kappa) \}.
\]

By Lemma 4.2 the boundaries of \( \Theta^u_\infty(\kappa) \) converge almost surely as \( \kappa \to \infty \) to the boundaries of \( \Theta^u_\infty(\infty) \), and therefore \( \tilde{V}^{i,j}_\kappa \) converges almost surely to \( V^{i,j}_\infty \). It follows that the coalescence time

\[
\tilde{\tau}^{i,j}_\kappa = \sup \{ t \in [0, \kappa] : \tilde{V}^{i,j}_\kappa \in \mathcal{N}_\kappa(t) \}
\]

as \( \kappa \to \infty \) converges almost surely to \( \tau^{i,j}_\infty \). Moreover

\[
\lim_{\kappa \to \infty} \tilde{\tau}^{i,j}_\kappa I \{ N_\infty(\kappa) > 0, \tilde{V}^i_\kappa \neq \tilde{V}^j_\kappa \} = \tau^{i,j}_\infty I \{ \forall t, N_\infty(t) > 0 \}
\]
The coalescent tree of a Markov branching process with... Page 21 of 22

almost surely. But by Theorem 3.1, \( \tilde{\tau}_{i,j}^{\kappa} \) on the event \( \{ N_{\infty}(\kappa) > 0, \hat{V}_{k}^{i} \neq \hat{V}_{k}^{j} \} \) converges in distribution to a finite random variable. Thus \( \tau_{\infty}^{i,j} \) on the event \( \{ \forall t, N_{\infty}(t) > 0 \} \) is finite with probability one.

\[ \text{Lemma 7.3} \quad \text{For } i \neq j, \]
\[ \lim_{\kappa \to \infty} \mathbb{P}[\tau_{\kappa}^{i,j} = \tau_{\infty}^{i,j}] = 1. \]

\[ \text{Proof} \quad \text{The event } \{ \tau_{\kappa}^{i,j} = \tau_{\infty}^{i,j} \} \quad \text{is a superset of} \]
\[ \{ N_{\kappa}(t) = N_{\infty}(t), \forall t \leq \zeta_{\kappa} \} \cap \{ V_{\kappa}^{i,j} = V_{\infty}^{i,j} \} \cap \{ \tau_{\infty}^{i,j} \leq \zeta_{\kappa} \text{ or } N_{\infty}(\zeta_{\kappa}) = 0 \}. \quad (15) \]

By Lemmas 4.1, 7.1 and 7.2 respectively, each of the events in (15) have probabilities which converge to 1.

\[ \text{Proof of Theorem 3.3} \quad \text{For } \kappa \in \mathbb{N} \cup \{ \infty \}, \text{the coalescent tree is } (\pi_{\kappa}(t))_{t \geq 0}, \text{where } \pi_{\kappa}(t) \text{ is the partition of } \{1, \ldots, m\} \text{ such that } i \text{ and } j \text{ are in the same block if and only if } \tau_{\kappa}^{i,j} \geq t. \]

Lemma 7.3 implies joint convergence of the coalescence times, which is equivalent to convergence in finite dimensional distributions of the coalescent tree. By Lemma 6.1 this convergence holds when sampling with or without replacement. The limiting coalescent tree is measurable with respect to the \( \mathcal{X} \) and \( N_{\infty}(\cdot) \). That is, the limiting coalescent tree is parameterised by the intrinsic reproduction rates.

\[ \text{Acknowledgements} \quad \text{I thank Alex Mcavoy for supportive and inspiring discussions, Michael Nicholson for shrewd comments on a earlier draft of the paper, and an anonymous reviewer for many thoughtful suggestions and corrections.} \]

\[ \text{References} \]

Athreya Krishna (2012) Coalescence in the recent past in rapidly growing populations. Stochast Process Appl 122(11):3757–3766

Athreya Krishna, Hong Jyy-i (2018) Coalescence on supercritical Bellman–Harris branching processes. Taiwanese J Math 22(1):245–261

Athreya Krishna, Ney Peter (2004) Branching Processes. Dover Publications, Mineola

Berestycki Nathanaël (2009) Recent progress in coalescent theory. Ensaios Matemáticos 16:1–193

Bertoin Jean, Le Gall Jean-François (2000) The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. Prob Theory Related Fields 117:249–266

Campbell RB (2003) A logistic branching process for population genetics. J Theor Biol 225(2):195–203

Harris Simon, Johnston Samuel, Roberts Matthew (2020) The coalescent structure of continuous-time Galton–Watson trees. Ann Appl Prob 30(3):1368–1414

Hong Jyy-i (2013) Coalescence in subcritical Bellman-Harris age-dependent branching processes. J Appl Prob 50(2):576–591

Höpfner Reinhard (1985) On some classes of population-size-dependent Galton–Watson processes. J Appl Prob 22(1):25–36

Jagers Peter, Klenbaner Fima (2000) Population-size-dependent and age-dependent branching processes. Stochast Process Appl 87(2):235–254

Johnston Samuel, Lambert Amaury (2019) The coalescent structure of uniform and Poisson samples from mutitype branching processes. arXiv:1912.00198

Johnston, Samuel (2019) The genealogy of Galton-Watson trees. Electronic J Prob 24(94)
Kingman John (1982) The coalescent. Stochast Process Appl 13(3):235–248
Kingman John (2000) Origins of the coalescent: 1974–1982. Genetics 156(4):1461–1463
Klebaner Fima (1984) On population-size-dependent branching processes. Adv Appl Prob 15(1):30–55
Lambert Amaury (2005) The branching process with logistic growth. Ann Appl Prob 15(2):1506–1535
Lambert Amaury (2018) The coalescent of a sample from a binary branching process. Theor Population Biol 122:30–35
Lambert Amaury, Popovic Lea (2013) The coalescent point process of branching trees. Ann Appl Prob 23(1):99–144
Lambert Amaury, Stadler Tanja (2013) Birth-death models and coalescent point processes: the shape and probability of reconstructed phylogenies. Theor Population Biol 90:113–128
Möhle Martin (1998) Robustness results for the coalescent. J Appl Prob 35(2):438–447
Popovic Lea (2004) Asymptotic genealogy of a critical branching process. Ann Appl Prob 14(4):2120–2148
Wakeley John (2008) Coalescent theory: an introduction. W. H. Freeman, New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.