

SKELETONS OF NEAR-CRITICAL BIENAYMÉ–GALTON–WATSON BRANCHING PROCESSES

SERIK SAGITOV,* *Chalmers University of Technology and Gothenburg University*

MARIA CONCEIÇÃO SERRA,** *Minho University*

Abstract

Skeletons of branching processes are defined as trees of lineages characterized by an appropriate signature of future reproduction success. In the supercritical case a natural choice is to look for the lineages that survive forever (O’Connell (1993)). In the critical case it was suggested that the particles with the total number of descendants exceeding a certain threshold could be distinguished (see Sagitov (1997)). These two definitions lead to asymptotic representations of the skeletons as either pure birth process (in the slightly supercritical case) or critical birth–death processes (in the critical case conditioned on the total number of particles exceeding a high threshold value). The limit skeletons reveal typical survival scenarios for the underlying branching processes. In this paper we consider near-critical Bienaymé–Galton–Watson processes and define their skeletons using marking of particles. If marking is rare, such skeletons are approximated by birth and death processes, which can be subcritical, critical, or supercritical. We obtain the limit skeleton for a sequential mutation model (Sagitov and Serra (2009)) and compute the density distribution function for the time to escape from extinction.

Keywords: Bienaymé–Galton–Watson process; birth and death process; decomposable multitype branching process; escape from extinction

2010 Mathematics Subject Classification: Primary 60J80

Secondary 92D25

1. Introduction

Imagine a population of viruses trying to establish itself in a new environment. Suppose that the currently dominating type is nearly critical, in that its mean offspring number is close to 1. One can think of two main factors which may lead to survival of this population: reproductive success or an advantageous mutation (a mutation producing a new type of particle forming a strictly supercritical process); see, for instance, [4], [5], and [10]. While a reproductive success is possible in the slightly supercritical case, ‘survival due to an advantageous mutation’ is the only way to escape extinction for a slightly subcritical branching system.

The typical survival scenarios of such branching processes can be studied in terms of the so-called skeleton trees formed by lineages characterized by an appropriate signature of future reproduction success. In the supercritical case a natural choice is to look for the lineages that survive forever [6]. In the critical case it was suggested that to distinguish the particles with the total number of descendants exceeding a certain threshold could be distinguished (see [7]).

Received 11 April 2013; revision received 15 April 2014.

* Postal address: Mathematical Sciences, Chalmers and Gothenburg University, SE-41296 Gothenburg, Sweden.

Email address: serik@chalmers.se

** Postal address: Center of Mathematics, Minho University, Campus de Gualtar, 4710-057 Braga, Portugal.

Email address: mcserra@math.uminho.pt

These two definitions lead to asymptotic representations of the skeletons as either pure birth processes (in the slightly supercritical case) or critical birth–death processes (in the critical case conditioned on the total number of particles exceeding a high threshold value).

In this paper we suggest an alternative approach to defining a skeleton that relies on a random marking of the lineages in the family tree of a Bienaymé–Galton–Watson (BGW) process. The skeleton is then defined as the subtree formed by the infinite lineages together with the marked lineages. In Section 2 we describe the skeleton of infinite lineages and recall the result from [6] concerning a sequence of a single-type slightly supercritical BGW process. It says that, conditioned on the event that the skeleton is not empty, the skeleton is approximated by the standard Yule process (a linear pure birth process).

In Section 3 we consider an exactly critical BGW process with marking: each particle in the family tree is marked, independently of the others, with a small probability. Here the skeleton is formed by the lineages leading to a marked particle. We show that, conditioned on the event that it is not empty, the skeleton is approximated by a critical linear birth–death process.

In Section 4 we state the main result of the paper, which concerns a sequence of near-critical BGW processes with marked particles. The definition of a skeleton is adapted in order to include both infinite lineages and marked lineages. The marking is also performed in a more general way than in Section 3. Our main result states that, conditioned on the event that it is not empty, the skeleton is approximated by a linear birth–death process which can be either supercritical, critical, or subcritical, depending on the parameters of the model. The proof of this main result, Theorem 4.1, is given in Section 8.

In Section 5 we study a decomposable two-type BGW process with irreversible mutations, starting from a single wild-type individual. Each daughter of a wild-type individual becomes a mutant, independently of the others, with a small probability. We look at this binomial mutation model as a particular case of the processes treated in Section 4 by considering the following marking procedure: a wild-type individual is marked if it gives birth to at least one mutant daughter.

In Section 6 we deal with a sequential mutation model, considered in [8], for a viral population that escapes extinction due to a sequence of irreversible mutations that lead to a target-type. It is assumed that mutations appear according to a binomial mutation model and results from Section 5 are used to obtain the asymptotic shape of the limit skeleton. Finally, in Section 7 we address the important question of the time to escape from extinction in a sequential mutation model. By ‘the time to escape from extinction’ we mean the first generation where a particle of the target-type appears. Due to the shape of the skeleton, we are able to show that the time to escape from extinction is asymptotically equal to the time of the first death occurring in the limit skeleton. An explicit equation for the density distribution function of the time to the first death in the limit skeleton is derived.

2. Infinite lineages

Consider a sequence of branching processes $\{Z_m(n)\}_{n=0}^{\infty}$, $m = 1, 2, \dots$ with offspring distributions $(p_m(0), p_m(1), \dots)$ and starting from one particle $Z_m(0) = 1$ (see [1], [3],

and [9]). Assume that the processes are nearly critical with

$$\sum_{k=1}^{\infty} k p_m(k) = 1 + \varepsilon_m, \quad \varepsilon_m \rightarrow 0, m \rightarrow \infty, \tag{2.1}$$

$$\sum_{k=2}^{\infty} k(k-1) p_m(k) \rightarrow \sigma^2, \quad m \rightarrow \infty, \text{ for some } \sigma > 0, \tag{2.2}$$

$$\sup_m \sum_{k=n}^{\infty} k^2 p_m(k) \rightarrow 0, \quad n \rightarrow \infty. \tag{2.3}$$

Condition (2.3) requires uniform integrability for the sequence of squared offspring numbers and implies that the following equality

$$\sum_{k=0}^{\infty} s^k p_m(k) = 1 - (1 + \varepsilon_m)(1 - s) + \left(\frac{\sigma^2}{2} - R_m(s) \right) (1 - s)^2, \quad s \in [0, 1] \tag{2.4}$$

holds with $R_m(s) \rightarrow 0$ uniformly in m as $s \uparrow 1$.

A natural way of defining a skeleton for branching processes was proposed in [6], where such processes were considered with $\varepsilon_m > 0$ for all m (slightly supercritical case). Its survival probability, Q_m , according to Lemma 3.3 of [6] satisfies the following well-known approximation formula

$$Q_m \sim 2\varepsilon_m \sigma^{-2}, \quad m \rightarrow \infty. \tag{2.5}$$

It is a well-known fact that a supercritical branching process can be viewed as a two-type branching process by distinguishing between particles with an infinite line of descent and particles having a finite number of descendants. If we concentrate only on the number of particles with an infinite line of descent, we arrive at the so-called skeleton process. Conditioning the supercritical process on nonextinction and focusing on infinite lineages, we obtain a new sequence of supercritical branching process $\{X_m(n)\}_{n \geq 0}$ with $X_m(0) = 1$ (see [1]). Additionally, under conditions (2.1), (2.2), and (2.3), $\{X_m(n)\}_{n \geq 0}$ is weakly approximated as

$$\left\{ X_m \left(\frac{t}{\varepsilon_m} \right) \right\}_{t \geq 0} \rightarrow \{Y_1(t)\}_{t \geq 0}, \quad m \rightarrow \infty \tag{2.6}$$

by the Yule process, see Theorem 3.2 of [6]. Recall that the Yule process is a continuous-time Markov branching process with particles living exponential times with mean 1, and, at the moment of death, the particles are replaced by two new particles. The key part of the proof of (2.6) is to show, using (2.4), that

$$\mathbb{P}\{X_m(1) = 1 \mid X_m(0) = 1\} = 1 - \varepsilon_m + o(\varepsilon_m), \tag{2.7a}$$

$$\mathbb{P}\{X_m(1) = 2 \mid X_m(0) = 1\} = \varepsilon_m + o(\varepsilon_m), \tag{2.7b}$$

where \mathbb{P} is the probability measure. Then it remains to check the convergence of the generator of this Markov chain against the generator of the Yule process after the time is scaled accordingly.

The limitation of this definition of a skeleton is that it has no direct extension to the critical or subcritical branching processes. Theorem 2.1 from [6] shows that in the critical case if the branching process is conditioned ‘on very late extinction’ then the limiting skeleton (without any scaling) is a trivial discrete-time process $Y(n) \equiv 1, n = 0, 1, \dots$

In this paper we suggest an alternative approach of defining a skeleton relying on a random marking of the lineages in the family tree. We start by studying, in the next section, a simple case of exactly critical reproduction.

3. Critical branching processes with independently marked particles

Consider a single-type BGW process $Z(n)$ such that its offspring distribution (p_0, p_1, \dots) has mean $\sum_{k=1}^{\infty} kp_k = 1$ and finite variance $\sigma^2 = \sum_{k=2}^{\infty} k(k-1)p_k$. Consider the corresponding family tree and suppose that each vertex in the tree is independently marked with a small probability $\mu_m \rightarrow 0$, as $m \rightarrow \infty$. Any path connecting the root with a marked vertex will be considered as a part of the skeleton. Thus, the skeleton is the subtree of the family tree formed by the paths of the skeleton. Adapting Proposition 2.1 from [8] to this case one can show that for a given μ_m the sequence $\{X_m(n)\}_{n \geq 0}$ of numbers of branches in the skeleton forms a BGW process. Next, we find the conditional asymptotic structure of the skeleton.

Let again $Q_m = \mathbb{P}\{X_m(0) = 1\}$ stand for the probability that the skeleton is not empty (at least one particle is marked). Due to the branching property we have

$$1 - Q_m = (1 - \mu_m)\phi(1 - Q_m), \tag{3.1}$$

where $\phi(s) = \sum_{k=0}^{\infty} p_k s^k$. Indeed, (3.1) simply says that the skeleton is empty if and only if the root is not marked and all the daughter subtrees, if any, have empty skeletons. Using the Taylor expansion of ϕ around point 1, we obtain

$$\phi(1 - Q_m) = 1 - Q_m + \frac{Q_m^2 \sigma^2}{2} + o(Q_m^2),$$

and deduce from (3.1) that

$$Q_m \sim \sigma^{-1} \sqrt{2\mu_m}.$$

Denoting by ξ_m the indicator of the event that the ancestral particle is marked we obtain

$$\begin{aligned} \mathbb{E}\{r^{\xi_m} s^{X_m(1)}; X_m(0) = 1\} &= \mathbb{E}\{r^{\xi_m}\} \mathbb{E}\{s^{X_m(1)}\} - \mathbb{E}\{r^{\xi_m} s^{X_m(1)}; X_m(0) = 0\} \\ &= (r\mu_m + 1 - \mu_m)\phi(sQ_m + 1 - Q_m) - \mathbb{P}\{X_m(0) = 0\}, \end{aligned}$$

implying that the offspring distribution of the skeleton particles satisfies

$$\begin{aligned} \mathbb{E}\{r^{\xi_m} s^{X_m(1)} \mid X_m(0) = 1\} &= 1 - \frac{1 - \phi(1 - Q_m(1 - s))}{Q_m} - (1 - r)\frac{\mu_m}{Q_m} + o(\sqrt{\mu_m}) \\ &= s + \frac{\sigma^2}{2} Q_m(1 - s)^2 - \sigma\sqrt{2\mu_m} \frac{1 - r}{2} + o(\sqrt{\mu_m}) \\ &= (1 - \sigma\sqrt{2\mu_m})s + \sigma\sqrt{2\mu_m}(\frac{1}{2}r + \frac{1}{2}s^2) + o(\sqrt{\mu_m}), \end{aligned}$$

where \mathbb{E} is the expectation measure. It follows that using $\tau_m = \sigma\sqrt{2\mu_m}$ we can write

$$\mathbb{P}\{\xi_m = 0, X_m(1) = 1 \mid X_m(0) = 1\} = 1 - \tau_m + o(\tau_m), \tag{3.2a}$$

$$\mathbb{P}\{\xi_m = 1, X_m(1) = 0 \mid X_m(0) = 1\} \sim \frac{\tau_m}{2}, \tag{3.2b}$$

$$\mathbb{P}\{\xi_m = 0, X_m(1) = 2 \mid X_m(0) = 1\} \sim \frac{\tau_m}{2}. \tag{3.2c}$$

Comparing (3.2) to (2.7a) we conclude that, if the original branching process produces at least one marked particle, there holds a weak convergence in the Skorokhod sense (see [2])

$$\left\{ X_m \left(\frac{t}{\tau_m} \right) \right\}_{t \geq 0} \rightarrow \{ Y_{1/2}(t) \}_{t \geq 0}, \quad m \rightarrow \infty. \tag{3.3}$$

Here, the limit process is a continuous-time Markov branching process with the critical binary splitting:

- particles live exponential times with parameter 1,
- at the moment of its death each particle with probability $\frac{1}{2}$ leaves no children and with probability $\frac{1}{2}$ produces two children.

The limit process, $Y_{1/2}(\cdot)$, being a critical branching process, will eventually become extinct.

Relation (3.2) gives an enhanced interpretation of the limit skeleton (3.3). All marked particles appearing in the branching process can be associated with the tips of the limit skeleton. In particular, the total number of the marked particles, W_m , conditioned on $W_m > 0$, is asymptotically distributed as the total number of leaves, W , in the family tree of the limit skeleton. Due to the branching property we have

$$W \stackrel{D}{=} \mathbf{1}_{\{v=0\}} + (W' + W'') \mathbf{1}_{\{v=2\}},$$

where v is the number of offspring of the initial particle in the skeleton, $\mathbf{1}$ is the indicator function, and W' and W'' are independent and identically distributed (i.i.d.) with the same law as W . In terms of the generating functions we obtain an equation $\mathbb{E}\{s^W\} = (s + [\mathbb{E}\{s^W\}]^2)/2$ leading to $\mathbb{E}\{s^W\} = 1 - \sqrt{1-s}$.

4. Main result

In this section we combine and further develop the two approaches presented in Sections 2 and 3 for a more general model. Consider a nearly critical reproduction law $\{p_m(k)\}_{k=0}^\infty$ satisfying (2.1), (2.2), (2.3), and allowing for negative ε_m . Furthermore, assume that a particle with k offspring is marked with probability $A_m(k)$. The marking event may depend not only on the number of offspring (i.e. on the number of first generation daughters) but also on the whole daughter branching process. For example, the marking rule could be to mark all particles whose total number of descendants exceeds m (including the particles with infinite lineages, if any), see [7]. Observe that for the marked near-critical BGW process the total probability for a particle to be marked is given by

$$\mu_m = \sum_{k=0}^\infty p_m(k) A_m(k).$$

Clearly, the case of Section 2 corresponds to the zero marking probability, $\mu_m \equiv 0$, and the case of Section 3 corresponds to $\varepsilon_m \equiv 0$ and $A_m(k) \equiv \mu_m$.

Reconciling the two different definitions of a skeleton given in Sections 2 and 3 we next introduce a new definition.

Definition 4.1. For a given family tree of a BGW process with marking, the subtree formed by the marked lineages together with infinite lineages will be called the skeleton.

Clearly, if $\mu_m = 0$ the skeleton is formed only by the infinite lineages. If $\mu_m > 0$ any infinite lineage becomes marked and we can think that the skeleton is formed only by the marked lineages.

In this paper we study the asymptotic behavior of the skeleton assuming

$$\mu_m \rightarrow 0, \quad m \rightarrow \infty, \tag{4.1}$$

restricting ourselves to the cases when the mean offspring number for the marked particles

$$M_m = \mu_m^{-1} \sum_{k=1}^{\infty} k p_m(k) A_m(k)$$

satisfies

$$\limsup_{m \rightarrow \infty} M_m < \infty. \tag{4.2}$$

By this we exclude such extreme situations as, for example, when $A_m(k)$ is of order $\mu_m k^2$ for large k and $\sum k^3 p_m(k) \rightarrow \infty$. Observe also that, given (2.3) and (4.1),

$$\sum_{k=2}^{\infty} k(k-1) p_m(k) A_m(k) \rightarrow 0, \tag{4.3}$$

which is obtained by using the inequality

$$\sum_{k=2}^{\infty} k(k-1) p_m(k) A_m(k) \leq n^2 \mu_m + \sum_{k=n}^{\infty} k^2 p_m(k).$$

Let, as before, Q_m stand for the probability that the skeleton is not empty. Now we can state our main result, claiming that, conditioned on the event that the skeleton is not empty, a weak convergence of the following form holds

$$\left\{ X_m \left(\frac{t}{\tau_m} \right) \right\}_{t \geq 0} \rightarrow \{ Y_\lambda(t) \}_{t \geq 0}, \quad m \rightarrow \infty, \tag{4.4}$$

for a convenient sequence $(\tau_m)_{m \geq 0}$ and convenient $\lambda \in [0, 1]$, generalizing both (2.6) and (3.3). Here, for a given $\lambda \in [0, 1]$ the limit process is a continuous-time Markov branching process with binary splitting:

- particles live exponential times with parameter 1,
- at the moment of its death each particle with probability $1 - \lambda$ leaves no children and each particle with probability λ produces two children.

Remark 4.1. Importantly, as with (3.3) in Section 3, by claiming (4.4) we implicate that asymptotically there is a one-to-one correspondence between the marked particles appearing in the branching process and the tips of the limit skeleton $Y_\lambda(\cdot)$. In particular, in (2.6) the limit skeleton has no tips, implying that under the corresponding time scale we cannot expect to observe marked particles in the branching process.

Theorem 4.1. *Under conditions (2.1), (2.2), (2.3), (4.1), and (4.2), assuming that there exists a finite or infinite limit,*

$$c = \lim_{m \rightarrow \infty} \frac{\varepsilon_m}{\sqrt{\mu_m}}, \tag{4.5}$$

- (i) if $c = \infty$, then (2.5) and (2.6) hold,
- (ii) if $c \in (-\infty, \infty)$, then

$$Q_m \sim \sqrt{\mu_m} \frac{c + \sqrt{c^2 + 2\sigma^2}}{\sigma^2}, \tag{4.6}$$

also (4.4) holds with $\tau_m = \sqrt{\mu_m} \sqrt{c^2 + 2\sigma^2}$ and $\lambda = \frac{1}{2} + \frac{1}{2}c/\sqrt{c^2 + 2\sigma^2}$,

- (iii) if $c = -\infty$, then

$$Q_m \sim \frac{\mu_m}{|\varepsilon_m|},$$

also (4.4) holds with $\tau_m = |\varepsilon_m|$ and $\lambda = 0$.

According to Theorem 4.1 there are five different asymptotic regimes for the skeleton of a near-critical BGW process, depending on how the deviation from the critical reproduction, ε_m , relates to the square root of the marking probability, $\sqrt{\mu_m}$:

- in the supercritical case $c = \infty$ with a negligible marking probability, the limit skeleton is the Yule process which never dies out,
- in the supercritical case $c \in (0, \infty)$ with a balanced marking probability, the limit skeleton is a supercritical Markov branching process which dies out with probability $(\sqrt{c^2 + 2\sigma^2} - c)/(\sqrt{c^2 + 2\sigma^2} + c)$ and survives forever with probability $2c/(\sqrt{c^2 + 2\sigma^2} + c)$,
- if the reproduction law is very close to the purely critical, $c = 0$, then the limit skeleton is a critical Markov branching process which dies out with probability 1 although rather slowly,
- in the subcritical case $c \in (-\infty, 0)$ with a balanced marking probability, the limit skeleton is a subcritical Markov branching process which dies out with probability 1,
- in the subcritical case $c = -\infty$ with a very small marking probability, the limit skeleton is given by a single lineage that dies out after an exponential time.

5. Binomial mutation model

Here we present an important example of a marked branching process based on a decomposable two-type Bienaymé–Galton–Watson process modeling a population of individuals with irreversible mutations. The two-type branching process stems from a single wild-type individual which produces k offspring with probability $q_m(k)$. Suppose that each daughter of a wild-type individual becomes a mutant with probability π_m independently of the other daughters.

To introduce a marked BGW process we focus only on the wild-type individuals and mark those wild-type individuals who have at least one mutant offspring. The reproduction law for the marked branching process is given by the distribution for the number of wild-type offspring,

$$p_m(k) = \sum_{l=0}^{\infty} q_m(k+l) \binom{k+l}{l} (1-\pi_m)^k \pi_m^l, \tag{5.1}$$

and the conditional marking probabilities $A_m(k)$ are computed using the following relations

obtained by splitting (5.1) in two parts:

$$p_m(k)(1 - A_m(k)) = q_m(k)(1 - \pi_m)^k, \tag{5.2}$$

$$p_m(k)A_m(k) = \sum_{l=1}^{\infty} q_m(k+l) \binom{k+l}{l} (1 - \pi_m)^k \pi_m^l. \tag{5.3}$$

To ensure that we can use the results from the previous section, we need conditions (2.1)–(2.3) to hold. Therefore, we assume that the reproduction law with a mutant offspring satisfies

$$\sum_{k=1}^{\infty} k q_m(k) = 1 + \eta_m, \quad \eta_m \rightarrow 0, \tag{5.4}$$

$$\sum_{k=2}^{\infty} k(k-1) q_m(k) \rightarrow \sigma^2, \tag{5.5}$$

$$\sup_m \sum_{k=n}^{\infty} k^2 q_m(k) \rightarrow 0, \quad n \rightarrow \infty \tag{5.6}$$

for some $\sigma \in (0, \infty)$. Assume also

$$\pi_m \rightarrow 0, \quad m \rightarrow \infty. \tag{5.7}$$

Lemma 5.1. *Conditions (5.4), (5.5), (5.6), and (5.7) imply (2.1), (2.2), (2.3), (4.1), and (4.2) with*

$$\mu_m \sim \pi_m, \quad m \rightarrow \infty, \tag{5.8}$$

$$M_m \rightarrow \sigma^2, \quad m \rightarrow \infty, . \tag{5.9}$$

Remark 5.1. Relation (5.9) has an interesting implication for our two-type branching process: the mean number of wild-type siblings in a family with at least one mutant, asymptotically equals the variance of the total offspring number.

Proof. First observe that (5.3) entails a useful expression for the marking probability

$$\mu_m = \sum_{k=0}^{\infty} p_m(k) A_m(k) = \sum_{k=1}^{\infty} q_m(k) (1 - (1 - \pi_m)^k). \tag{5.10}$$

Clearly, (5.10) and (5.4) yield

$$0 \leq \mu_m - \pi_m(1 + \eta_m) \leq \pi_m^2 \sum_{k=1}^{\infty} k^2 q_m(k), \tag{5.11}$$

and (5.8) follows from (5.4), (5.6) and (5.11). Next, due to (5.2) we have

$$\begin{aligned} 0 &\leq \sum_{k=2}^{\infty} k(k-1) q_m(k) - \sum_{k=2}^{\infty} k(k-1) p_m(k) (1 - A_m(k)) \\ &= \sum_{k=2}^{\infty} k(k-1) q_m(k) (1 - (1 - \pi_m)^k) \\ &\leq n^3 \pi_m + \sup_j \sum_{k=n}^{\infty} k^2 q_j(k) \end{aligned}$$

for any $n \geq 2$. Here, first letting $m \rightarrow \infty$ and then $n \rightarrow \infty$, due to (5.5) and (5.6), we arrive at

$$\sum_{k=2}^{\infty} k(k-1)p_m(k)(1-A_m(k)) \rightarrow \sigma^2.$$

This together with (4.3) implies (2.2). Now, according to (5.3) we have

$$\begin{aligned} \mu_m M_m \pi_m^{-1} &= \sum_{k=1}^{\infty} k \sum_{l=1}^{\infty} q_m(k+l) \binom{k+l}{l} (1-\pi_m)^k \pi_m^{l-1} \\ &= \sum_{j=2}^{\infty} q_m(j) \sum_{l=1}^j (j-l) \binom{j}{l} (1-\pi_m)^{j-l} \pi_m^{l-1} \\ &= (1-\pi_m)^{-1} \sum_{j=2}^{\infty} j(j-1)p_m(j)(1-A_m(j)) \\ &\quad + \pi_m \sum_{j=2}^{\infty} q_m(j) \sum_{l=2}^j (j-l) \binom{j}{l} (1-\pi_m)^{j-l} \pi_m^{l-2}. \end{aligned}$$

At this point we easily obtain (5.9), and, therefore, (4.2), using (4.3) and

$$\sum_{j=2}^{\infty} q_m(j) \sum_{l=1}^j (j-l) \binom{j}{l} (1-\pi_m)^{j-l} \pi_m^{l-2} \leq \sum_{j=2}^{\infty} j(j-1)q_m(j).$$

In view of

$$\sum_{k=1}^{\infty} k p_m(k)(1-A_m(k)) = 1 + \varepsilon_m - \mu_m M_m$$

we derive the following from (5.2) and (5.4):

$$0 \leq \eta_m - \varepsilon_m + \mu_m M_m \leq \pi_m \sum_{k=1}^{\infty} k^2 q_m(k). \tag{5.12}$$

Combining (4.2) and (5.12) we obtain (2.1).

To prove (2.3) we turn to (5.1) and note that

$$\begin{aligned} \sum_{k=n}^{\infty} k^2 p_m(k) &= \sum_{k=n}^{\infty} k^2 \sum_{l=0}^{\infty} q_m(k+l) \binom{k+l}{l} (1-\pi_m)^k \pi_m^l \\ &= \sum_{j=n}^{\infty} q_m(j) \sum_{l=0}^{j-n} (j-l)^2 \binom{j}{l} (1-\pi_m)^{j-l} \pi_m^l \\ &\leq \sum_{j=n}^{\infty} j^2 q_m(j). \end{aligned}$$

Thus, (2.3) is an immediate consequence of (5.6).

Corollary 5.1. *For the binomial mutation model satisfying equations (5.4), (5.5), (5.6), (5.7), and $\eta_m/\sqrt{\pi_m} \rightarrow c$, the statements (i), (ii), (iii) of Theorem 4.1 are valid after (ε_m, μ_m) are replaced by (η_m, π_m) .*

6. The sequential mutation model

Our next illustration of Theorem 4.1 deals with the sequential mutation model [8] for a viral population with irreversible mutations which escapes extinction as soon as a target-type of viruses is produced. To simplify the discussion we focus mainly on the two-step mutation model, extending the one-step model from Section 5.

Suppose we have a population of viruses stemming from a single virus which is able to reproduce and mutate giving rise to what we call the intermediate-type of viruses. The viruses of intermediate-type reproduce according to a common law and by mutation generate a new type of viruses, which we call the target-type. The marking rule for the intermediate-type is straightforward: we mark mothers with at least one daughter of the target-type. The wild-type marking rule is a bit more complicated: we mark a mother which has at least one *successful mutant daughter* (that is a mutant of the intermediate-type which has at least one marked descendant in the whole line of descent).

We will assume that the reproduction laws and marking probabilities for both wild-type and intermediate-type branching processes satisfy conditions of type (5.4), (5.5), (5.6), (5.7), and are described by triplets $(\eta_m, \sigma^2, \pi_m)$ and $(\hat{\eta}_m, \sigma_2^2, \hat{\pi}_m)$, respectively. Furthermore, we assume for simplicity that

$$\frac{\hat{\pi}_m}{\pi_m} \rightarrow \alpha, \quad m \rightarrow \infty \text{ for some } \alpha \in [0, \infty).$$

According to Corollary 5.1, if

$$\hat{\eta}_m \sim c_2 \sqrt{\pi_m}, \quad m \rightarrow \infty \text{ for some } c_2 \in (-\infty, \infty),$$

then, both if $\alpha > 0$ or $\alpha = 0$,

$$\hat{Q}_m \sim \sqrt{\pi_m} \frac{c_2 + \sqrt{c_2^2 + 2\alpha\sigma_2^2}}{\sigma_2^2},$$

where \hat{Q}_m is the counterpart of Q_m in Corollary 5.1 for the skeleton of the intermediate-type. If $c_2 > 0$ then the skeleton of the intermediate-type is supercritical. In particular, with $\alpha = 0$ it becomes the Yule process. The time scale for the intermediate-type skeleton is given by

$$\hat{t}_m = \sqrt{\pi_m} \sqrt{c_2^2 + 2\alpha\sigma_2^2}. \tag{6.1}$$

Our main interest in the two-step mutation model is, of course, the limit skeleton leading to the target-type. Therefore, we want to apply Corollary 5.1 once again to the branching system with the probability of a successful mutation for the wild-type viruses given by

$$\tilde{\pi}_m = \pi_m \hat{Q}_m \sim \pi_m^{3/2} \frac{c_2 + \sqrt{c_2^2 + 2\alpha\sigma_2^2}}{\sigma_2^2}.$$

Assuming

$$\eta_m \sim c_1 \pi_m^{3/4}, \quad m \rightarrow \infty \text{ for some } c_1 \in (-\infty, \infty),$$

we obtain

$$Q_m \sim \sqrt{\tilde{\pi}_m} \frac{c + \sqrt{c^2 + 2\sigma^2}}{\sigma^2}, \quad c = c_1 \sigma_2 \left[c_2 + \sqrt{c_2^2 + 2\alpha\sigma_2^2} \right]^{-1/2},$$

implying

$$Q_m \sim \pi_m^{3/4} \frac{c_1 \sigma_2 + \left[c_1^2 \sigma_2^2 + 2\sigma^2 \left(c_2 + \sqrt{c_2^2 + 2\alpha \sigma_2^2} \right) \right]^{1/2}}{\sigma^2 \sigma_2}. \tag{6.2}$$

The limit skeleton for the wild-type is supercritical if and only if $c_1 > 0$.

The time scale for the wild-type skeletons is given by

$$\tau_m = \pi_m^{3/4} \sqrt{c_1^2 + 2\sigma^2 \sigma_2^{-2} \left(c_2 + \sqrt{c_2^2 + 2\alpha \sigma_2^2} \right)},$$

which in the considered case is much slower than the time scale of the intermediate-type (6.1). Thus, the overall skeleton is given by the wild-type skeleton, and the first death in the limit skeleton corresponds to the time of escape from extinction when the first virus of the target-type appears.

The above considered case is one of the many possible combination of reproduction–mutation regimes for the two-step mutation model. Without analyzing each of the remaining cases we just point out that there is a situation when both parts of the skeleton live on the same time scale. This is the case when

$$\begin{aligned} \hat{\pi}_m &\sim \alpha \pi_m^\gamma, & m \rightarrow \infty \text{ for some } \gamma \in (1, 2), \\ \hat{\eta}_m &\sim -\beta \pi_m^{\gamma-1}, & m \rightarrow \infty \text{ for some } \beta \in (0, \infty), \\ \eta_m &\sim c_1 \pi_m, & m \rightarrow \infty \text{ for some } c_1 \in (-\infty, \infty). \end{aligned}$$

Here, both time scales are of order $1/\pi_m$.

Turning to the sequential model with $b - 1$ intermediate steps before the target-type, we extrapolate (6.2) to $Q_m \sim \text{const } \pi_m^{1-2^{-b}}$. This prediction should be compared with the strictly subcritical case where one expects $Q_m \sim \text{const } \pi_m^b$; see Theorem 7.1 of [8].

7. Time to escape

In the framework of the sequential mutation model of Section 6 it is crucial to be able to describe the time to escape from extinction.

Let T_m be the time until the first marked particle is observed and write

$$Q(t) = \lim_{m \rightarrow \infty} \mathbb{P} \left\{ T_m > \frac{t}{\sqrt{\mu_m (c^2 + 2\sigma^2)}} \right\}.$$

According to Theorem 4.1, given that the limit (4.5) is finite, $Q(t) = \mathbb{P}\{T > t\}$ is the tail probability of the time T to the first death in the limit skeleton $Y_\lambda(\cdot)$ with $\lambda = \frac{1}{2} + \frac{1}{2}c/\sqrt{c^2 + 2\sigma^2}$. The branching property of $Y_\lambda(\cdot)$ says that

$$T = L + \min(T', T'') \mathbf{1}_{\{v=2\}}, \tag{7.1}$$

where L is the exponential life length with mean 1, v is the number of offspring of the initial particle, T' and T'' are i.i.d. with T . Due to the branching property (7.1),

$$Q(t) = \mathbb{P}\{L > t\} + \mathbb{P}\{T > t, L \leq t\} = e^{-t} + \lambda \int_0^t Q^2(t - u) e^{-u} du.$$

It follows that

$$e^t Q(t) = 1 + \lambda \int_0^t Q^2(u)e^u du.$$

Differentiation over t yields a simple differential equation

$$Q'(t) + Q(t) = \lambda Q^2(t), \quad Q(0) = 1,$$

giving $Q(t) = 1/(\lambda + (1 - \lambda)e^t)$. Thus,

$$\mathbb{P}\left\{T_m > \frac{t}{\sqrt{\mu_m}}\right\} \rightarrow 2\left(1 + \frac{c}{\sqrt{c^2 + 2\sigma^2}} + \left(1 - \frac{c}{\sqrt{c^2 + 2\sigma^2}}\right)e^{t\sqrt{c^2 + 2\sigma^2}}\right)^{-1}$$

and we conclude that the scaled time $T_m\sqrt{\mu_m}$ has the limit density distribution function

$$\psi(t) = 2(\sqrt{c^2 + 2\sigma^2} - c)e^{t\sqrt{c^2 + 2\sigma^2}}\left(1 + \frac{c}{\sqrt{c^2 + 2\sigma^2}} + \left(1 - \frac{c}{\sqrt{c^2 + 2\sigma^2}}\right)e^{t\sqrt{c^2 + 2\sigma^2}}\right)^{-2}, \quad t \geq 0. \tag{7.2}$$

In particular, with $c = 0$ we obtain

$$\psi(t) = 2\sqrt{2}\sigma^2 e^{t\sigma\sqrt{2}}(1 + e^{t\sigma\sqrt{2}})^{-2}.$$

We illustrate the asymptotic density function, ψ , by Figure 1. In the supercritical case the density curve reaches its a maximum value at

$$t_{\max} = \frac{1}{\sqrt{c^2 + 2\sigma^2}} \ln\left(1 + \frac{2c}{\sqrt{c^2 + 2\sigma^2} - c}\right),$$

making the most likely value for the time to escape T_m to be around

$$\hat{T}_m = \sqrt{\frac{1}{\mu_m(c^2 + 2\sigma^2)}} \ln\left(1 + \frac{2c}{\sqrt{c^2 + 2\sigma^2} - c}\right).$$

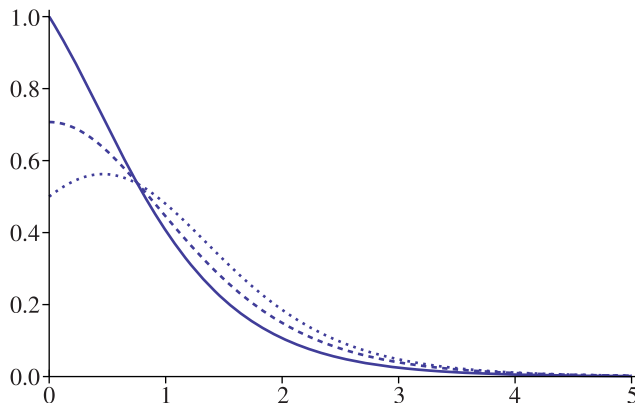


FIGURE 1: Three density curves given by (7.2) with $\sigma = 1$: subcritical case $c = -0.5$ (solid line), critical case $c = 0$ (dashed line), supercritical case $c = 0.5$ (dotted line).

8. Proof of Theorem 4.1

Our proof relies on the properties of the probability generating function

$$f_m(r, s) = \sum_{k=0}^{\infty} p_m(k)s^k(1 - A_m(k) + A_m(k)r)$$

jointly characterizing the marking status of a particle (through r) and its offspring number (through s).

Lemma 8.1. *Given conditions (2.1), (2.2), (2.3), and any sequence $u_m \in (0, 1)$ such that $u_m \rightarrow 0$ as $m \rightarrow \infty$, the following decomposition holds:*

$$1 - f_m(r, 1 - u_m) = \mu_m(1 - r) + u_m(1 + \varepsilon_m - \mu_m M_m(1 - r)) - \frac{u_m^2 \sigma_{r,m}^2}{2},$$

with $\sigma_{r,m}^2 \rightarrow \sigma^2$ as $m \rightarrow \infty$ uniformly over $r \in [0, 1]$.

Proof. This follows from a Taylor expansion around the point $(r, 1)$

$$f_m(r, 1 - u_m) = f_m(r, 1) - u_m \frac{\partial f_m}{\partial s}(r, 1) + \frac{u_m^2}{2} \left(\frac{\partial^2 f_m}{\partial s^2}(r, 1) + R_m(r) \right),$$

where

$$\begin{aligned} f_m(r, 1) &= \sum_{k=0}^{\infty} p_m(k)(1 - A_m(k) + A_m(k)r) = 1 - \mu_m(1 - r), \\ \frac{\partial f_m}{\partial s}(r, 1) &= \sum_{k=1}^{\infty} k p_m(k)(1 - A_m(k) + A_m(k)r) = 1 + \varepsilon_m - \mu_m M_m(1 - r), \\ \frac{\partial^2 f_m}{\partial s^2}(r, 1) &= \sum_{k=2}^{\infty} k(k - 1) p_m(k)(1 - A_m(k) + A_m(k)r), \end{aligned}$$

and

$$R_m(r) = \sum_{k=2}^{\infty} k(k - 1) p_m(k)(1 - A_m(k) + A_m(k)r)(1 - \theta_m^{k-2})$$

for some $\theta \in (1 - u_m, 1)$. Indeed, since, for any $n \geq 3$,

$$R_m(r) \leq \sum_{k=2}^{\infty} k^2 p_m(k)(1 - (1 - u_m)^{k-2}) \leq u_m n^3 + \sum_{k=n}^{\infty} k^2 p_m(k),$$

condition (2.3) implies $R_m(r) \rightarrow 0$. It remains to apply (2.2) and (4.3).

The skeleton is empty if the initial particle is not marked and all her children produce empty skeletons

$$1 - Q_m = f_m(0, 1 - Q_m).$$

Using Lemma 8.1, with $r = 0$ and $u_m = Q_m$, we obtain a quadratic equation

$$\sigma_{0,m}^2 Q_m^2 - 2Q_m(\varepsilon_m - \mu_m M_m) - 2\mu_m = 0,$$

entailing

$$Q_m = \frac{\varepsilon_m - \mu_m M_m + \sqrt{(\varepsilon_m - \mu_m M_m)^2 + 2\sigma_{0,m}^2 \mu_m}}{\sigma_{0,m}^2}, \tag{8.1}$$

where $\sigma_{0,m} \rightarrow \sigma$. Once again applying Lemma 8.1 now to the right-hand side of

$$\begin{aligned} \mathbb{E}\{r^{\xi_m} s^{X_m(1)}; X_m(0) = 1\} &= \mathbb{E}\{r^{\xi_m} s^{X_m(1)}\} - \mathbb{P}\{X_m(0) = 0\} \\ &= f_m(r, sQ_m + 1 - Q_m) - (1 - Q_m) \end{aligned}$$

with $u_m = Q_m(1 - s)$, we obtain

$$\begin{aligned} \mathbb{E}\{r^{\xi_m} s^{X_m(1)} \mid X_m(0) = 1\} &= (1 + \varepsilon_m - Q_m \sigma^2)s + \frac{Q_m s^2 \sigma^2}{2} + \mu_m Q_m^{-1} r \\ &\quad - \mu_m Q_m^{-1} - \varepsilon_m + \frac{Q_m \sigma^2}{2} + o(Q_m) + O(\mu_m). \end{aligned} \tag{8.2}$$

Now we are ready to verify the statements of Theorem 4.1 case by case.

Case (i). If $c = \infty$, then $\mu_m = o(\varepsilon_m)$ and (8.1) yields (2.5). Furthermore, $\mu_m Q_m^{-1} = o(\varepsilon_m)$ and (8.2) gives

$$\mathbb{E}\{r^{\xi_m} s^{X_m(1)} \mid X_m(0) = 1\} = (1 - \varepsilon_m)s + \varepsilon_m s^2 + o(\varepsilon_m)$$

as desired.

Case (ii). If $\varepsilon_m \sim c\sqrt{\mu_m}$ for $c \in (-\infty, +\infty)$ then (8.1) yields (4.6) as we can neglect the terms involving $\mu_m M_m$. Furthermore, $\mu_m Q_m^{-1} \sim \sqrt{\mu_m} \sigma^2 / (c + \sqrt{c^2 + 2\sigma^2})$ and the last terms in (8.2) are negligible

$$\begin{aligned} & - \mu_m Q_m^{-1} - \varepsilon_m + \frac{Q_m \sigma^2}{2} + o(Q_m) + O(\mu_m) \\ &= - \frac{\sqrt{\mu_m} \sigma^2}{c + \sqrt{c^2 + 2\sigma^2}} - c\sqrt{\mu_m} + \sqrt{\mu_m} \frac{c + \sqrt{c^2 + 2\sigma^2}}{2} + o(\sqrt{\mu_m}) \\ &= o(\sqrt{\mu_m}). \end{aligned}$$

We conclude that

$$\begin{aligned} & \mathbb{E}\{r^{\xi_m} s^{X_m(1)} \mid X_m(0) = 1\} \\ &= (1 - \sqrt{\mu_m} \sqrt{c^2 + 2\sigma^2})s \\ &\quad + \sqrt{\mu_m} \sqrt{c^2 + 2\sigma^2} \left[\left(\frac{1}{2} - \frac{1}{2} \frac{c}{\sqrt{c^2 + 2\sigma^2}} \right) r + \left(\frac{1}{2} + \frac{1}{2} \frac{c}{\sqrt{c^2 + 2\sigma^2}} \right) s^2 \right] + o(\sqrt{\mu_m}) \end{aligned}$$

as desired.

Case (iii). If $c = -\infty$ then $a_m := \varepsilon_m - \mu_m M_m$ takes negative values for sufficiently large m , and from (8.1) we derive

$$Q_m = \frac{2\sigma_{0,m}^2 \mu_m}{\sigma_{0,m}^2 (\sqrt{a_m^2 + 2\sigma_{0,m}^2 \mu_m} - a_m)} \sim \frac{\mu_m}{|\varepsilon_m|}$$

as stated. It remains to see that (8.2) gives

$$\mathbb{E}\{r^{\xi_m} s^{X_m(1)} \mid X_m(0) = 1\} = (1 - |\varepsilon_m|)s + |\varepsilon_m|r + o(\varepsilon_m),$$

since $\mu_m Q_m^{-1} \sim |\varepsilon_m|$, $\mu_m = o(|\varepsilon_m|)$ and $Q_m = o(|\varepsilon_m|)$.

Acknowledgements

The authors would like to thank the careful revision and valuable comments of an anonymous referee. The research of Serik Sagitov was supported by the Bank of Sweden Tercentenary Foundation and the Swedish Research Council (grant no. 621-2010-5623). The research of Maria Conceição Serra was supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the Fundação para a Ciência e a Tecnologia, through the Project PEstOE/MAT/UI0013/2014.

References

- [1] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
- [2] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterisation and Convergence*. John Wiley, New York.
- [3] HACCOU, P., JAGERS, P. AND VATUTIN, V. A. (2007). *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press.
- [4] IWASA, Y., MICHOR, F. AND NOWAK, M. A. (2003). Evolutionary dynamics of escape from biomedical intervention. *Proc. R. Soc. London B* **270**, 2573–2578.
- [5] IWASA, Y., MICHOR, F. AND NOWAK, M. A. (2004). Evolutionary dynamics of invasion and escape. *J. Theoret. Biol.* **226**, 205–214.
- [6] O'CONNELL, N. (1993). Yule process approximation for the skeleton of a branching process. *J. Appl. Prob.* **30**, 725–729.
- [7] SAGITOV, S. (1997). Limit skeleton for critical Crump–Mode–Jagers branching processes. In *Classical and Modern Branching Processes* (IMA Vol. Math. Appl. **84**), Springer, New York, pp. 295–303.
- [8] SAGITOV, S. AND SERRA, M. C. (2009). Multitype Bienaymé–Galton–Watson processes escaping extinction. *Adv. Appl. Prob.* **41**, 225–246.
- [9] SEVAST'YANOV, B. A. (1971). *Vetvyashchiesya Protssesy*. Nauka, Moscow (in Russian). German translation: SEWASTJANOW, B. A. (1974). *Verzweigungsprozesse*. Akademie-Verlag, Berlin.
- [10] TAÏB, Z. (1992). *Branching Processes and Neutral Evolution* (Lecture Notes Biomath. **93**), Springer, Berlin.