Multitype Galton-Watson processes escaping extinction

Serik Sagitov and Maria Conceição Serra[1]

School of Mathematical Sciences

Chalmers University of Technology

Abstract

A multitype Galton-Watson process describes populations of particles that live one season and are then replaced by a random number of children of possibly different types. Biological interpretation of the event that the daughter's type differs form the mother's type is that a mutation has occurred. We study a situation when mutations are rare and, among the types connected in a network, there is a supercritical type allowing the system to escape from extinction. We establish a neat asymptotic structure for the Galton-Watson process escaping extinction due to a sequence of mutations towards the supercritical type. The conditional limit process is a GW process with a multitype immigration stopped after a sequence of geometric times.

---

[1]Current address: Department of Mathematics, University of Minho, Campus de Gualtar, 4710-057 Braga Portugal

(0,0,1,1)

(0,1,1,1) (0,1,0,1) (0,0,0,1)

(1,0,1,1) (0,1,1,0) (0,0,1,0)

(1,1,1,1) (0,0,0,0)

(1,1,0,1) (1,0,0,1) (0,1,0,0)

(1,1,1,0) (1,0,1,0) (1,0,0,0)

(1,1,0,0)
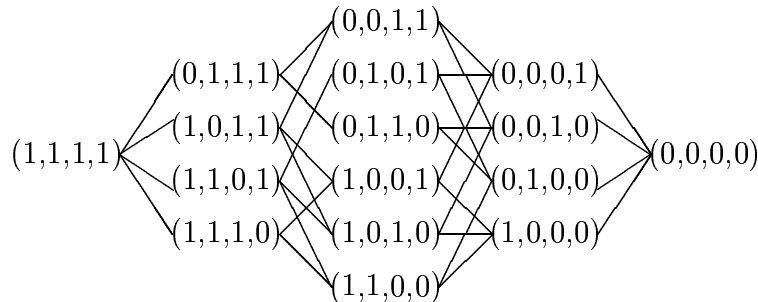
Figure 1: The network of 0-1 sequences of length $L = 4$. Here the edges represent single point mutations of probability $\mu(1-\mu)^{L-1}$. More generally, mutation between two sequences which differ in i sites has probability $\mu^i(1-\mu)^{L-i}$.

1 Introduction

This work is motivated by a series of papers by Iwasa et al [4], [5] studying the probability that a virus placed in a hostile environment escapes extinction via a chain of mutations. In their basic population model the virus is coded by a vector of zeros and ones of length L, so that the set of the sequences can be viewed as a graph, illustrated by Figure 1, with 2^L vertices representing different forms of the virus.

Let all the sequences with exactly i ones have the same fitness m_i defined as the mean offspring number. If we further assume that $0 < m_i < 1$ for $i = 1, \ldots, L$ and $m_0 > 1$, then a reproduction process stemming from the subcritical form $(1, \ldots, 1)$ is doomed to get extinct unless a sequence of mutations results in the supercritical form $(0, \ldots, 0)$. Suppose that all point mutations have the same probability μ per site per generation. Then according to [5] the escape probability is a small number of order μ^L as $\mu \to 0$. This means that asymptotically we can disregard the possibility of backward mutations on the path from the initial sequence $(1, \ldots, 1)$ towards the escape sequence $(0, \ldots, 0)$.

A relevant Markov chain describing such a virus population is a multi-type Galton-Watson (GW) process (see [3], [8]) with 2^L types of particles. Here time is measured in generations and particles are assumed to reproduce asexually. The offspring numbers are assumed to be independent random variables whose distributions are common for all particles of the same type. The aim of this paper is to give the asymptotic GW process describing the number and the types of viruses in a population escaping extinction.

We start our study with the case $L = 1$, when the GW process has just two types of particles 0 and 1. In Section 2 we consider a general two-type GW

process which starts with a type 1 particle and where all the type 0 particles are killed. Such a one type population can be viewed as a decomposable two-type GW process or, as shown in Section 3, as a decomposable three-type GW process. These decompositions are our basis for obtaining the limit structures of the GW processes escaping extinction like that in Section 4.

In the two-type case mutation is modelled in a more general way than discussed so far. Namely, in the limit theorem of Section 4 we will allow the mutation probability for a newborn type 1 particle to depend on its family size. We show that in the case when the initial type 1 is subcritical and the other type 0 is supercritical, the conditional limit process is a GW process with immigration. The immigration source is turned on during a geometric time T_1 with mean $E(T_1) = \frac{1}{1-m_1}$. The immigration source corresponds to a stem lineage leading to the successful mutation. The numbers of immigrants have the size-biased distribution.

Careful asymptotic analysis of the two type case demonstrates that the backward mutations on the path to escape are negligible. Therefore when analyzing the case with the number of types larger than two we simply disregard the possibility of backward mutations. In Section 6 we study a sequential mutation model, where mutations may occur along an interval of types starting from type L and ending at type 0. This is a natural intermediate step between the two-type case towards the network model. Indeed, if we treat each subset of sequences with exactly i ones as a single type i, then we arrive at a sequential model with the mutation probability between types i and j being asymptotically equivalent to $\binom{i}{j}\mu^{i-j}$, given $j < i$.

The asymptotic results in Section 7 address a wide class of sequential mutation models. Section 8 discusses the asymptotic distribution of the total time to escape and its expected value. Finally, in Section 9 we apply the results of Section 7 to the network mutation model.

2 A two-type GW process focussed on one type

Consider a GW process with two types of particles labelled by 0 and 1. If $Z_i(n)$ is the number of type i particles in generation n, then the vector $\{Z_0(n), Z_1(n)\}_{n\geq 0}$ forms a Markov chain describing the population size and type structure evolving generationwise. Proposition 2.1 below describes the fate of type 1 particles as a decomposable GW process recognizing two subtypes of type 1 particles. This construction reminds the well-known decomposition of the supercritical GW process into particles with infinite and

finite lines of descent (see [1], p. 47).

We are going to distinguish between two kinds of type 1 particles: subtype 10 - those who have type 0 particles among their descendants, and subtype 11 - those whose total progeny consists only of type 1 particles. Given the process starts from a single particle, called a *progenitor*, there are two possibilities to choose the type of the progenitor

$$
\begin{aligned}
B_0 &= \{Z_0(0) = 1, Z_1(0) = 0\}, \\
B_1 &= \{Z_0(0) = 0, Z_1(0) = 1\}.
\end{aligned}
$$

The event B_1, in turn, is the union of two disjoint events

$$
\begin{aligned}
B_{10} &= \{Z_{10}(0) = 1, Z_{11}(0) = 0\}, \\
B_{11} &= \{Z_{10}(0) = 0, Z_{11}(0) = 1\},
\end{aligned}
$$

revealing the subtype of the progenitor. We shall denote by P_0, P_1, P_{10}, P_{11} the conditional probabilities specifying the type or subtype of the progenitor and by E_0, E_1, E_{10}, E_{11} the respective expectation operators. Put

$$
Q_{10} = P_1(B_{10}), \ Q_{11} = P_1(B_{11}) = 1 - Q_{10},
$$

and define the two-type reproduction law in terms of the generating functions

$$
f_i(s_0, s_1) = E_i\left(s_0^{Z_0(1)} s_1^{Z_1(1)}\right), \ i = 0, 1.
$$

Consider $k = k_0 + k_1$ daughters of the progenitor assuming that $Z_0(1) = k_0$ among them have type 0 and $Z_1(1) = k_1$ have type 1. We label the type 1 daughters by the numbers $1, \ldots, k_1$ and the type 0 daughters by $k_1 + 1, \ldots, k$. Within the types the labeling is done uniformly at random. For an event A concerning the original GW process we will denote by $A^{(i)}$ its analog associated with the GW process stemming from the i-th daughter of the progenitor. Let R denote the smallest label among the type 1 children having the subtype 10,

$$
R(\omega) = \min\{1 \leq i \leq Z_1(1) : \omega \in B_{10}^{(i)}\},
$$

with the usual convention that the minimum of the empty set is infinity. The next lemma is analogous to Lemma 2.1 in [2].

Lemma 2.1 *The conditional distribution of R is given by*

$$
P_1(R = j | Z_0(1) = k_0, Z_1(1) = k_1) = \begin{cases} Q_{10} Q_{11}^{j-1}, & 1 \leq j \leq k_1, \\ Q_{11}^{k_1}, & j = \infty, \end{cases} \tag{1}
$$

4

implying

$$E[s^R] = \frac{sQ_{10}}{1 - sQ_{11}}(1 - f_1(1, sQ_{11})).$$

and

$$Q_{11} = f_1(0, Q_{11}). \tag{2}$$

PROOF Let $\{A_1, \ldots, A_k\}$ be some random events concerning the GW process generated by a single progenitor particle. For $k = k_0 + k_1$ due to independence of particle lives we have

$$P_1(Z_0(1) = k_0, Z_1(1) = k_1, A_i^{(i)}, 1 \le i \le k) \tag{3}$$

$$= P_1(Z_0(1) = k_0, Z_1(1) = k_1) \prod_{i=1}^{k_1} P_1(A_i) \prod_{i=k_1+1}^{k} P_0(A_i).$$

With a specific choice of $A_i = C_i$, where

$$C_i = \begin{cases} B_{11}, & 1 \le i \le j - 1, \\ B_{10}, & i = j, \\ B_1, & j + 1 \le i \le k_1, \\ B_0, & k_1 + 1 \le i \le k, \end{cases} \tag{4}$$

and $j \in [1, k_1]$, we get

$$\{R = j, Z_0(1) = k_0, Z_1(1) = k_1\} = \{Z_0(1) = k_0, Z_1(1) = k_1, C_i^{(i)}, 1 \le i \le k\},$$

Thus according to (3)

$$P_1(R = j | Z_0(1) = k_0, Z_1(1) = k_1) = Q_{11}^{j-1} Q_{10}. \tag{5}$$

Now it remains only to note that

$$P_1(R = \infty | Z_0(1) = k_0, Z_1(1) = k_1) = Q_{11}^{k_1},$$

and that equation (2) follows from

$$B_{11} = \{Z_0(1) = 0, R = \infty\}. \tag{6}$$

$\square$

Lemma 2.2 *Let $\{A_i, i = 1, 2, 3, \ldots\}$ be random events concerning the GW process generated by a single progenitor particle. The daughter versions of*

5

theses events $\{A_i^{(i)}, i = 1, \dots, k\}$, $k = k_0 + k_1$ *are conditionally independent given* $R = j, Z_0(1) = k_0, Z_1(1) = k_1$ *with*

$$P_1(A_i^{(i)}|R = j, Z_0(1) = k_0, Z_1(1) = k_1) = \begin{cases} P_{11}(A_i), & 1 \le i \le j - 1 \wedge k_1, \\ P_{10}(A_i), & i = j, \\ P_1(A_i), & j + 1 \le i \le k_1, \\ P_0(A_i), & k_1 + 1 \le i \le k. \end{cases}$$

PROOF Let $j \in [1, k_1]$. In terms of the system of events (4)

$$\{R = j, Z_0(1) = k_0, Z_1(1) = k_1, A_i^{(i)}, 1 \le i \le k\}$$
$$= \{Z_0(1) = k_0, Z_1(1) = k_1, A_i^{(i)} \cap C_i^{(i)}, 1 \le i \le k\}.$$

The stated conditional independence follows now from (3) and (5)

$$P_1(A_i^{(i)}, 1 \le i \le k, R = j|Z_0(1) = k_0, Z_1(1) = k_1)$$
$$= \prod_{i=1}^{j-1} P_1(A_i \cap B_{11})P_1(A_j \cap B_{10}) \prod_{i=j+1}^{k_1} P_1(A_i) \prod_{i=k_1+1}^{k} P_0(A_i)$$
$$= P_1(R = j|Z_0(1) = k_0, Z_1(1) = k_1)$$
$$\times \prod_{i=1}^{j-1} P_{11}(A_i)P_{10}(A_j) \prod_{i=j+1}^{k_1} P_1(A_i) \prod_{i=k_1+1}^{k} P_0(A_i).$$

$\square$

Proposition 2.1 *Let the two-type GW process start from a 1-particle and kill every 0-particle appearing in the population. The resulting process can be treated as a decomposable GW process* $\{Z_{10}(n), Z_{11}(n)\}_{n \ge 0}$ *with two types of particles: 10 and 11. The progenitor's type has distribution* (Q_{10}, Q_{11}), *and the new two-type reproduction law is defined by*

$$E_{11}\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right) = \frac{f_1(0, Q_{11}s_1)}{Q_{11}}, \tag{7}$$

$$E_{10}\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right) = \frac{f_1(1, Q_{10}s_0 + Q_{11}s_1) - f_1(0, Q_{11}s_1)}{Q_{10}}. \tag{8}$$

PROOF Since

$$E_{11}\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right) Q_{11} + E_{10}\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right) Q_{10} = E_1\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right)$$

and due to the basic branching property

$$E_1\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right) = f_1(1, Q_{10}s_0 + Q_{11}s_1),$$

6

to prove (7) and (8) it suffices to verify that

$$E_1\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}; B_{11}\right) = f_1(0, Q_{11}s_1).$$

But this follows from (6) and (1):

$$
\begin{aligned}
E_1\left(s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}; Z_0(1) = 0, R = \infty\right) &= E_1\left(s_1^{Z_1(1)}; Z_0(1) = 0, R = \infty\right) \\
&= f_1(0, Q_{11}s_1).
\end{aligned}
$$

Finally, the claimed independence of particle lives in the framework of the new two-type system follows from Lemma 2.2.

$\square$

3 A refined structure of the two-type GW process focussed on one type

With the same two-type GW process let us now distinguish between two kinds of type 0 particles which will be labelled as 00 and 01 depending on the explosion (00) or extinction (01) of the GW process stemming from a 0-particle in question. This allows us to refine our earlier classification of the type 1 particles after the subtype 10 is further divided into subtypes 100 and 101. A subtype 100 particle is a 10-particle with at least one 00-descendant. It means that the other subtype 101 must have 01-descendants but never 00-descendants. Let $Z_{100}(n)$ and $Z_{101}(n)$ stand for the number of particles of subtypes 100 and 101 in generation n. Put

$$
\begin{aligned}
Q_{100} &= P_1\left(Z_{100}(0) = 1, Z_{101}(0) = 0, Z_{11}(0) = 0\right), \\
Q_{101} &= P_1\left(Z_{100}(0) = 0, Z_{101}(0) = 1, Z_{11}(0) = 0\right),
\end{aligned}
$$

so that $Q_{100} + Q_{101} = Q_{10}$.

It is well-known that the extinction probabilities

$$q_i = P_i\left(\lim_{n\to\infty}\left(Z_0(n) + Z_1(n)\right) = 0\right)$$

satisfy the following pair of equations (see [1], p.186)

$$
\begin{aligned}
q_0 &= f_0(q_0, q_1), \\
q_1 &= f_1(q_0, q_1).
\end{aligned}
$$

Clearly, $Q_{100} = 1 - q_1$, since for the progenitor of type 1 to be classified as a subtype 100 particle the corresponding branching process must survive forever. This yields the following equation

$$1 - Q_{100} = f_1(q_0, 1 - Q_{100}). \tag{9}$$

Proposition 3.1 *Let the two-type GW process start from a 1-particle and kill each 0-particle appearing in the population. The resulting process can be treated as a decomposable GW process $\{Z_{100}(n), Z_{101}(n), Z_{11}(n)\}_{n \geq 0}$ with three types of particles 100, 101, and 11.*

The progenitor's type has distribution $(Q_{100}, Q_{101}, Q_{11})$ and the new three-type reproduction law is defined by

$$E_{11}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) = \frac{f_1(0, Q_{11}s_1)}{Q_{11}}, \tag{10}$$

$$E_{101}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) = \frac{f_1(q_0, Q_{101}s_{01} + Q_{11}s_1) - f_1(0, Q_{11}s_1)}{Q_{101}}, \tag{11}$$

$$E_{100}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) \tag{12}$$

$$= \frac{f_1(1, Q_{100}s_{00} + Q_{101}s_{01} + Q_{11}s_1) - f_1(q_0, Q_{101}s_{01} + Q_{11}s_1)}{Q_{100}}.$$

PROOF It is easy to adjust Lemma 2.1 and Lemma 2.2 to verify the branching property of the three-type process, and it directly follows from Proposition 2.1 that (10) holds. Now, in view of the branching property of the three-type process

$$E_1\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) = f_1(1, s_{00}Q_{100} + s_{01}Q_{101} + s_1Q_{11})$$

and because of

$$\begin{aligned} E_1\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) &= Q_{100}E_{100}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) \\ &\quad + Q_{101}E_{101}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) \\ &\quad + Q_{11}E_{11}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right) \end{aligned}$$

to prove (11) and (12) it suffices to show that the sum of the last two terms equals $f_1(q_0, Q_{11}s_1 + Q_{101}s_{01})$, or equivalently

$$\begin{aligned} E_1\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}; \text{the process dies out}\right) \\ = E_1\left(s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}; \text{the process dies out}\right) \\ = f_1(q_0, Q_{101}s_{01} + Q_{11}s_1). \end{aligned}$$

But the last equality follows from the branching property saying that for the process to die out all the daughter processes should die out independently, and when it comes to type one daughters there are two possible ways toward extinction: either with or without type 0 descendants.

$\square$

8

4 Limit theorem in the two-type case

In the previous two sections we considered a two-type GW process with a general reproduction law described by a pair of generating functions $f_0(s_0, s_1)$ and $f_1(s_0, s_1)$. In this section we deal with a family of the two-type GW processes labelled by a parameter $0 < \mu < 1$ regulating communication rates between types 0 and 1. We will assume a particular kind of the reproduction law for the type 1 particles:

$$
\begin{aligned}
f_1^{(\mu)}(s_0, s_1) &= \sum_{k=0}^{\infty} p_1(k) \left(s_1(1 - \mu a_{10}^{(\mu)}(k)) + s_0 \mu a_{10}^{(\mu)}(k) \right)^k \\
&= \sum_{k=0}^{\infty} p_1(k) \left(s_1 + (s_0 - s_1)\mu a_{10}^{(\mu)}(k) \right)^k .
\end{aligned}
\tag{13}
$$

Here $\{p_1(k)\}_{k=0}^{\infty}$ is the distribution of the total offspring number for a type 1 particle. Notice that the offspring number is independent of the parameter μ which controls mutation or change of type.

According to (13) each out of k offspring independently chooses its type: type 1 with probability $(1 - \mu a_{10}^{(\mu)}(k))$ or type 0 (mutation event) with probability $\mu a_{10}^{(\mu)}(k)$. In our asymptotic analysis μ goes to zero making mutations rare events. We will assume the uniform convergence

$$
\sup_{k \geq 0} | a_{10}^{(\mu)}(k) - a_{10}(k)| \to 0, \ \mu \to 0,
\tag{14}
$$

where the limit sequence is uniformly bounded

$$
\sup_{k \geq 0} a_{10}(k) < \infty.
\tag{15}
$$

Obviously, $f_1^{(\mu)}(s_0, s_1) \to \phi_1(s_1)$, where

$$
\phi_1(s) = \sum_{k=0}^{\infty} p_1(k)s^k.
$$

We assume a similar convergence for the offspring numbers of the type 0 particles

$$
f_0^{(\mu)}(s_0, s_1) \to \phi_0(s_0),
\tag{16}
$$

where the limit generating function $\phi_0(s)$ describes the limit reproduction regime of type 0 with no mutation to type 1.

9

The limiting mean offspring numbers $m_i = \phi_i'(1)$, $i = 0, 1$ are always supposed to be positive and finite. Condition (16) implies $q_0^{(\mu)} \to q$, where $q = 1$ if $m_0 \leq 1$, and $q \in [0, 1)$, $q = \phi_0(q)$ if $m_0 > 1$. In terms of the generating function

$$\psi_{10}(s) = \sum_{k=1}^{\infty} k p_1(k) a_{10}(k) s^{k-1}$$

condition (15) ensures $\psi_{10}(1) < \infty$.

Theorem 4.1 *Consider the μ-labelled two-type GW process stemming from a type 1 particle which satisfies conditions (13)-(16). If $m_1 < 1$ and $\psi_{10}(1) > 0$, then the probability of the mutation event has asymptotics*

$$\frac{Q_{10}^{(\mu)}}{\mu} \to \frac{\psi_{10}(1)}{1 - m_1}, \ \ \mu \to 0, \tag{17}$$

and conditioned on the mutation event, the process $\{Z_{10}(n), Z_{11}(n)\}_{n \geq 0}$ converges in distribution to a limit process $\{X_{10}(n), X_{11}(n)\}_{n \geq 0}$, which is a decomposable two-type GW process described below.

If furthermore, $m_0 > 1$, then the probability of the escape event has asymptotics

$$\frac{Q_{100}^{(\mu)}}{\mu} \to (1 - q)\frac{\psi_{10}(1)}{1 - m_1}, \ \ \mu \to 0, \tag{18}$$

and conditioned on the escape event, the process $\{Z_{100}(n), Z_{101}(n), Z_{11}(n)\}_{n \geq 0}$ converges in distribution to $\{X_{10}(n), 0, X_{11}(n)\}_{n \geq 0}$.

In view of Propositions 2.1 and 3.1 this theorem is a consequence of three convergences as $\mu \to 0$ proven in the next section

$$E_{11}^{(\mu)}\left[s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right] \ \ \to \ \ \phi_1(s_1), \tag{19}$$

$$E_{10}^{(\mu)}\left[s_0^{Z_{10}(1)} s_1^{Z_{11}(1)}\right] \ \ \to \ \ m_1 s_0 \frac{\phi_1'(s_1)}{\phi_1'(1)} + (1 - m_1)\frac{\psi_{10}(s_1)}{\psi_{10}(1)}, \tag{20}$$

$$E_{100}^{(\mu)}\left[s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)}\right] \ \ \to \ \ m_1 s_{00} \frac{\phi_1'(s_1)}{\phi_1'(1)} + (1 - m_1)\frac{\psi_{10}(s_1)}{\psi_{10}(1)}. \tag{21}$$

The limit generating functions in (19)-(21) imply the following reproduction rules in the limit process.

The limit $\{X_{10}(n), X_{11}(n)\}_{n \geq 0}$ is a GW process with two types 10^* and 11^* corresponding to the escape (stem) lineage and extinct (side) lineages respectively. It starts with a single 10^*-particle, whose reproduction law is

described by the limit generation function in (20)-(21). At the time of death this particle either produces one 10^*-particle with probability m_1, or zero 10^*-particles with probability $(1 - m_1)$. In both cases it also produces a random number of 11^*-particles: in the former case the generating function for the number of 11^*-offspring is $\phi_1'(s)/m_1$, and in the latter case it is $\psi_{10}(s)/\psi_{10}(1)$. It follows that asymptotically the stem lineage stays alive for a geometric time T_1 with mean $E(T_1) = \frac{1}{1-m_1}$ (cf. [7]).

Relation (19) says that the 11^*-particles reproduce themselves according to the generating function $\phi_1(s)$. Therefore, the process $X_{11}(n)$ can be viewed as the number of particles in a GW process with a stopped immigration. Think of the stem lineage described above as the immigration source, with every immigrant initiating an independent GW process with the offspring generating function $\phi_1(s)$. At times $1, \ldots, T_1 - 1$, the independent numbers of immigrants have a common distribution with the generating function $\phi_1'(s)/m_1$. At the time T_1, when the stem lineage stops, the number of immigrants has a possibly different distribution with the generating function $\psi_{10}(s)/\psi_{10}(1)$.

Notice that if mutation probability is independent of the family size $a_{10}(k) \equiv c$, then $\psi_{10}(s) = c\phi_1'(s)$ and $\psi_{10}(1) = cm_1$, so that even the last number of immigrants has the generating function $\phi_1'(s)/m_1$. Observe that this generating function corresponds to the so-called size-biased version of the offspring distribution $\phi_1(s)$, see for example [6]. In this case $\{X_{10}(n), X_{11}(n)\}_{n \geq 0}$ becomes a size-biased version of the single type GW process with the offspring generating function $\phi_1(s)$, whose distinguished line is stopped at the geometric time T_1.

5 Proof of Theorem 4.1

Throughout this section we assume (13)-(16) and $m_1 < 1$. We prove (17)-(21) (where relations (18) and (21) additionally require that $m_0 > 1$) using the following lemma.

Lemma 5.1 *As $\mu \to 0$ uniformly over $(s_0, s_1) \in [0, 1]^2$*

$$
\begin{aligned}
f_1^{(\mu)}(s_0, s_1) &= \phi_1(s_1) + \mu (s_0 - s_1)\psi_{10}(s_1) \\
&+ o(\mu(s_0 - s_1)) + O\left(\mu^2(s_0 - s_1)^2 \frac{\eta(s_1)}{1 - s_1}\right),
\end{aligned}
$$

where $\eta(s) = m_1 - \phi_1'(s)$ is such that $\eta(s) \searrow 0$ as $s \to 1$.

11

PROOF If $0 \leq a \leq a + b \leq 1$, then

$$0 \leq (a+b)^k - a^k - ka^{k-1}b \leq b^2 \sum_{i=1}^{k-1}(k-i)a^{i-1} \leq k\, b^2 \frac{1-a^{k-1}}{1-a} \qquad (22)$$

and therefore

$$\left| \sum_{k=1}^{\infty} p_1(k)[s_1 + \mu(s_0 - s_1)a_{10}(k)]^k - \phi_1(s_1) - \mu(s_0 - s_1)\psi_{10}(s_1) \right|$$

$$\leq \sum_{k=1}^{\infty} p_1(k) \left| [s_1 + \mu(s_0 - s_1)a_{10}(k)]^k - s_1^k - k\mu a_{10}(k)(s_0 - s_1)s_1^{k-1} \right|$$

$$\leq \sum_{k=1}^{\infty} p_1(k) \left\{ \mu^2(s_0 - s_1)^2 a_{10}^2(k) \sum_{i=1}^{k-1}(k-i)s_1^{i-1} \right\}$$

$$\leq C^2 \mu^2 (s_0 - s_1)^2 \sum_{k=1}^{\infty} p_1(k)k \frac{1 - s_1^{k-1}}{1 - s_1}$$

$$= C^2 \mu^2 (s_0 - s_1)^2 \frac{m_1 - \phi_1'(s_1)}{1 - s_1}.$$

On the other hand, (13) implies

$$\left| f_1^{(\mu)}(s_0, s_1) - \sum_{k=1}^{\infty} p_1(k)[s_1 + \mu(s_0 - s_1)a_{10}(k)]^k \right|$$

$$\leq \mu |s_0 - s_1| \sum_{k=1}^{\infty} p_1(k)k |a_{10}^{(\mu)}(k) - a_{10}(k)|$$

which gives the $o(\mu(s_0 - s_1))$ term due to the uniform convergence condition.
$\square$

PROOF OF (17)-(18)
The probability $Q_{10}^{(\mu)}$ that a 1-particle will have at least one 0 type descendant is estimated from below by

$$p_1(k)k\mu a_{10}^{(\mu)}(k)[1 - \mu a_{10}^{(\mu)}(k)]^{k-1}$$

whatever is $k = 1, 2, 3, \ldots$. Since $\psi_{10}(1) > 0$, there exists such a k that $p_1(k)a_{10}(k) > 0$. Thus in view of the condition (15) we can conclude that

$$\limsup_{\mu \to 0} \mu / Q_{10}^{(\mu)} < \infty. \qquad (23)$$

By Lemma 5.1

$$f_1^{(\mu)}(0, Q_{11}^{(\mu)}) - \phi_1(Q_{11}^{(\mu)}) + \mu Q_{11}^{(\mu)} \psi_{10}(Q_{11}^{(\mu)}) = o(\mu) + O\left(\mu^2 \frac{\eta(Q_{11}^{(\mu)})}{Q_{10}^{(\mu)}}\right)$$

which combined with (2) and (23) yields

$$Q_{11}^{(\mu)} - \phi_1(Q_{11}^{(\mu)}) + \mu Q_{11}^{(\mu)} \psi_{10}(Q_{11}^{(\mu)}) = o(\mu) + O\left(\mu \eta(Q_{11}^{(\mu)})\right).$$

It follows immediately that $Q_{11}^{(\mu)} \to 1$ and therefore

$$\frac{\phi_1(Q_{11}^{(\mu)}) - Q_{11}^{(\mu)}}{\mu} \to \psi_{10}(1).$$

This implies (17), since $\phi_1(s) - s \sim (1 - m_1)(1 - s)$ as $s \to 1$.

Applying Lemma 5.1 once again we obtain

$$f_1^{(\mu)}(q_0^{(\mu)}, 1 - Q_{100}^{(\mu)}) - \phi_1(1 - Q_{100}^{(\mu)})$$
$$= \mu(q_0^{(\mu)} - 1 + Q_{100}^{(\mu)})\psi_{10}(1 - Q_{100}^{(\mu)}) + o(\mu) + O\left(\frac{\mu^2 \eta(1 - Q_{100}^{(\mu)})}{Q_{100}^{(\mu)}}\right),$$

where $q_0^{(\mu)} \to q$ with $q \in [0, 1)$ given $m_0 > 1$. Using (9) we can derive

$$1 - Q_{100}^{(\mu)} - \phi_1(1 - Q_{100}^{(\mu)})$$
$$= \mu(q - 1 + Q_{100}^{(\mu)})\psi_{10}(1 - Q_{100}^{(\mu)}) + o(\mu) + O\left(\mu\eta(1 - Q_{100}^{(\mu)})\right),$$

since $Q_{100}^{(\mu)}/\mu$ is bounded away from zero. Now it is obvious how to finish the proof of (18).

$\square$

PROOF OF (19)-(21)
In view of Proposition 2.1 relation (19) is obvious. The other two relations have similar proofs - here we give a proof of (21) based on the next observation. If $0 \le s_i \le s_i + \delta_i \le 1$ for $i = 0, 1$, then according to (22)

$$0 \le f_1^{(\mu)}(s_0 + \delta_0, s_1 + \delta_1) - f_1^{(\mu)}(s_0, s_1) - R^{(\mu)}(s_0, s_1, \delta_0, \delta_1) \tag{24}$$

$$\le \sum_{k=1}^{\infty} k p_1(k) \left(\delta_1 + \mu(\delta_0 - \delta_1)a_{10}^{(\mu)}(k)\right)^2 \frac{1 - \left(s_1 + \mu(s_0 - s_1)a_{10}^{(\mu)}(k)\right)^{k-1}}{1 - s_1 - \mu(s_0 - s_1)a_{10}^{(\mu)}(k)}$$

13

where

$$R^{(\mu)}(s_0, s_1, \delta_0, \delta_1)$$
$$= \sum_{k=1}^{\infty} k p_1(k) \left(s_1 + \mu(s_0 - s_1) a_{10}^{(\mu)}(k) \right)^{k-1} \left(\delta_1 + \mu(\delta_0 - \delta_1) a_{10}^{(\mu)}(k) \right).$$

Relations (24) and (12) yield

$$0 \leq Q_{100}^{(\mu)} E_{100}^{(\mu)} \left[s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_1^{Z_{11}(1)} \right] - R^{(\mu)}(q_0^{(\mu)}, s^{(\mu)}, 1 - q_0^{(\mu)}, Q_{100}^{(\mu)} s_{00})$$
$$\leq \sum_{k=1}^{\infty} k p_1(k) \left(Q_{100}^{(\mu)} s_{00} + \mu(1 - q_0^{\mu} - Q_{100}^{(\mu)} s_{00}) a_{10}^{(\mu)}(k) \right)^2$$
$$\times \ \frac{1 - \left(s^{(\mu)} + \mu(q_0^{(\mu)} - s^{(\mu)}) a_{10}^{(\mu)}(k) \right)^{k-1}}{1 - s^{(\mu)} - \mu(q_0^{(\mu)} - s^{(\mu)}) a_{10}^{(\mu)}(k)}$$

with $s^{(\mu)} = Q_{11}^{(\mu)} s_1 + Q_{101}^{(\mu)} s_{01} \to s_1$. It remains to observe that the right hand side is $O(\mu^2)$ and

$$\frac{R^{(\mu)}(q_0^{(\mu)}, s^{(\mu)}, 1 - q_0^{(\mu)}, Q_{100}^{(\mu)} s_{00})}{Q_{100}^{(\mu)}}$$
$$= \sum_{k=1}^{\infty} k p_1(k)$$
$$\times \ \left(s^{(\mu)} + \mu(q_0^{(\mu)} - s^{(\mu)}) a_{10}^{(\mu)}(k) \right)^{k-1} \left(s_{00} + \mu \left(\frac{1 - q_0^{(\mu)}}{Q_{100}^{(\mu)}} - s_{00} \right) a_{10}^{(\mu)}(k) \right)$$
$$\to \ m_1 s_{00} \frac{\phi_1'(s_1)}{\phi_1'(1)} + (1 - m_1) \frac{\psi_{10}(s_1)}{\psi_{10}(1)}.$$

$\square$

6   The sequential mutation model

Suppose we can distinguish between $L+1$ types of particles, labelled $0, \ldots, L$. Type i particles can only produce particles of the types $0, \ldots, i$, whatever is $i \in [0, L]$. Notice that this sequential mutation model only partially extends the previous two-type model. We prohibit the reverse mutations for the sake of simplicity. As the asymptotic analysis of the two type case shows, the more general sequential model with reversed mutations should lead to the same asymptotic behavior.

Let $Z_j(n)$ be the number of type j particles existing at time n given that the branching process stems from a single particle whose type is specified by the index of the probability measure P_i. Adjusting the notation of the two-type case put

$$f_i(s_0, s_1, \ldots, s_i) = E_i\left(s_0^{Z_0(1)}\, s_1^{Z_1(1)}\, \ldots\, s_i^{Z_i(1)}\right)$$

and

$$Q_{i1} = P_i[Z_0(n) = 0, \text{ for all } n \geq 0], \ \ Q_{i0} = 1 - Q_{i1}.$$

Then

$$Q_{i1} = f_i(0, Q_{11}, Q_{21}, \ldots, Q_{i1}) \tag{25}$$

since to avoid descendants of type 0 nor the progenitor itself can have daughters of type 0, neither the progenitor's daughters can have descendants of type 0. We split each type in two subtypes in a way similar to our decomposition of Section 2. Consider the future of a particle of type i: with probability Q_{i0} it will eventually manage to produce a particle of type 0, in which case the particle is labelled $i0$, and otherwise with probability Q_{i1} it is labelled $i1$. Arguing as in Lemma 2.1 we see that the process $\{Z_{L0}(n), Z_{L1}(n), \ldots, Z_{10}(n), Z_{11}(n)\}$ is a decomposable $2L$-type GW process.

To describe the reproduction law in the $2L$-type GW process observe that

$$E_i\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right)$$

$$= Q_{i0} E_{i0}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right)$$

$$+ Q_{i1} E_{i1}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right)$$

where the left hand side is

$$f_i(1, Q_{10}s_{10} + Q_{11}s_{11}, \ldots, Q_{i0}s_{i0} + Q_{i1}s_{i1})$$

and

$$Q_{i1} E_{i1}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) = f_i(0,\, s_{11}Q_{11},\, s_{21}Q_{21},\, \ldots,\, s_{i1}Q_{i1}).$$

It follows

$$E_{i1}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) = \frac{f_i(0,\, s_{11}Q_{11},\, s_{21}Q_{21},\, \ldots,\, s_{i1}Q_{i1})}{Q_{i1}},$$

$$\tag{26}$$

$$E_{i0}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) \tag{27}$$

$$= \frac{f_i(1, Q_{10}s_{10} + Q_{11}s_{11}, \ldots, Q_{i0}s_{i0} + Q_{i1}s_{i1}) - f_i(0,\, s_{11}Q_{11},\, \ldots,\, s_{i1}Q_{i1})}{Q_{i0}}.$$

Our forthcoming asymptotic analysis turns to an extension of the two-type case (13)

$$
\begin{aligned}
f_i^{(\mu)}(s_0, s_1, \ldots, s_i) &= E_i^{(\mu)}\left(s_0^{Z_0(1)} s_1^{Z_1(1)} \ldots s_i^{Z_i(1)}\right) \\
&= \sum_{k=0}^{\infty} p_i(k)\left[s_i\left(1 - \sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)\right) + \sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)s_j\right]^k \\
&= \sum_{k=0}^{\infty} p_i(k)\left[s_i + \sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)(s_j - s_i)\right]^k,
\end{aligned}
$$

(28)

which says that each out of the k offspring of an i-particle independently chooses its type: it mutates to a type $j \in [0, i-1]$ with probability $\mu^{i-j}a_{ij}^{(\mu)}(k)$ or retains the maternal type type i with probability $1 - \sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)$. Here again parameter μ controls mutation rates so that as $\mu \to 0$ mutations become rare

$$
f_i^{(\mu)}(s_0, s_1, \ldots, s_i) \to \phi_i(s_i), \ \ \phi_i(s) = \sum_{k=0}^{\infty} p_i(k)s^k.
$$

(29)

Put $m_i = \sum_{k=1}^{\infty} kp_i(k)$. We will assume that all types, possibly except 0, are asymptotically subcritical, i.e.,

$$
0 < m_i < 1, \ i = 1, \ldots, L, \ 0 < m_0 < \infty.
$$

(30)

As in Section 2 we will assume the uniform convergence

$$
\sup_{0 \le j < i \le L} \sup_{k \ge 0} |a_{ij}^{(\mu)}(k) - a_{ij}(k)| \to 0, \ \mu \to 0,
$$

(31)

where the limit sequences are uniformly bounded

$$
\sup_{0 \le j < i \le L} \sup_{k \ge 0} a_{ij}(k) < \infty.
$$

(32)

Put

$$
\psi_{ij}(s) = \sum_{k=1}^{\infty} kp_i(k)a_{ij}(k)s^{k-1}, \ 0 \le j < i \le L
$$

(33)

and define a matrix $\mathbb{A} = [A_{ij}]_{i,j=0}^{L}$ by

$$
A_{ij} = \begin{cases} 1, & i = j = 0, \\ \frac{\psi_{ij}(1)}{1-m_i}, & 0 \le j \le i - 1, \\ 0, & \text{otherwise.} \end{cases}
$$

Define a vector $(\chi_0, \ldots, \chi_L)$ recursively

$$\chi_i = \sum_{j=0}^{i-1} A_{ij}\chi_j, \quad \chi_0 = 1. \tag{34}$$

If all $\psi_{ij}(1) > 0$, then all components of this vector are strictly positive.

In terms of the matrix powers $\mathbb{A}^n = [A_{ij}^{(n)}]_{i,j=0}^{L}$ we can write

$$
\begin{aligned}
\chi_i &= A_{i0} + \sum_{j=1}^{i-1} A_{ij}\chi_j \\
&= A_{i0} + \sum_{j=1}^{i-1} A_{ij}\left(A_{j0} + \sum_{k=1}^{j-1} A_{jk}\chi_k\right) \\
&= A_{i0}^{(2)} + \sum_{j=1}^{i-2} A_{ij}^{(2)}\chi_j = \ldots \\
&= A_{i0}^{(i)} \\
&= \sum_{k=0}^{i-1} \sum_{0=j_0<j_1<j_2<\ldots<j_k<i} A_{ij_k}\ldots A_{j_10}.
\end{aligned}
$$

It follows from (34) that the i-th row of the matrix $\mathbb{B} = [B_{ij}]_{i,j=0}^{L}$ with

$$B_{ij} = \frac{\chi_j}{\chi_i}A_{ij} \tag{35}$$

defines a probability distribution on the set $\{0, \ldots, i-1\}$. Notice that the matrix powers $\mathbb{A}^n$ and $\mathbb{B}^n$ are connected by $B_{ij}^{(n)} = \frac{\chi_j}{\chi_i}A_{ij}^{(n)}$.

7   Limit theorem for the sequential model

The following result partially extends the two-type Theorem 4.1. It is clear how a full extension would look like.

Theorem 7.1 *Consider the μ-labelled $2L$-type process GW process described in Section 6 that starts from a type L particle and satisfies conditions (29)-(32). Let all $\psi_{ij}(1) > 0$. The probability that the process produces at least one particle of type 0 has asymptotics*

$$\frac{Q_{L0}}{\mu^L} \to \chi_L, \quad \mu \to 0. \tag{36}$$

Conditioned on the event that a 0-particle is produced, the process $\{Z_{L0}(n), Z_{L1}(n), \ldots, Z_{10}(n), Z_{11}(n)\}_{n \geq 0}$ converges in distribution to a limit process $\{X_{L0}(n), X_{L1}(n), \ldots, X_{10}(n), X_{11}(n)\}_{n \geq 0}$ which is a decomposable $2L$-type GW process described below.

The limit process $\{X_{L0}(n), X_{L1}(n), \ldots, X_{10}(n), X_{11}(n)\}_{n \geq 0}$ starts with a single particle of type $L0^*$. This particle lives a geometric number T_L of generations with $P(T_L = n) = m_L^{n-1}(1 - m_L)$. At times $1, 2, \ldots, T_L - 1$ the stem $L0^*$-particle gives birth to particles of type $L1^*$ according to a size-biased distribution with generating function $\phi_L'(s)/m_L$. Each particle of type $L1^*$ initiates an independent single-type subcritical GW process with the offspring generating function $\phi_L(s)$. At time T_L the stem particle is replaced by a stem particle of type $i0^*$, where index i chosen from the set $\{0, 1, \ldots, L-1\}$ according to the distribution $\{B_{L0}, \ldots, B_{L,L-1}\}$, see (35). The number of $L1^*$ particles born at time T_L has a different distribution: the p.g.f. is given by $\psi_{Li}(s)/\psi_{Li}(1)$.

After time T_L the scheme above is repeated with L being replaced by i. The particle of type $i0^*$ lives a geometric time T_i with mean $\frac{1}{1-m_i}$, in that

$$X_{i0}(n) = \begin{cases} 1, & n \in \{T_L, \ldots, T_L + T_i - 1\} \\ 0, & \text{otherwise} \end{cases} .$$

At times $T_L + 1, \ldots, T_L + T_i - 1$ particles of type $i1^*$ appear from the stem particle according to the size-biased distribution $\phi_i'(s)/m_i$ and each one of them initiates an independent single-type subcritical GW processes with offspring generating function $\phi_i(s)$. At time $T_L + T_i$ the stem particle changes its type to $j0^*$, where j is chosen from $\{0, 1, \ldots, i-1\}$ according to the probability measure $\{B_{i0}, \ldots, B_{i,i-1}\}$. The distribution of the number of type $i1^*$ particles produced at time $T_L + T_i$ has generating function $\psi_{ij}(s)/\psi_{ij}(1)$. And the whole process restarts from this type j particle until the type 0 particle is produced.

PROOF The proof of Theorem 7.1 is similar to the proof of Theorem 4.1 therefore here we only outline the major changes. Lemma 5.1 can be extended to

$$\left| f_i^{(\mu)}(s_0, s_1, \ldots, s_i) - \phi_i(s_i) - \sum_{j=0}^{i-1} \mu^{i-j}(s_j - s_i)\psi_{ij}(s_i) \right| = \qquad (37)$$

$$= o\left(\sum_{j=0}^{i-1} \mu^{i-j}|s_j - s_i| \right) + O\left(\mu^{2i} \frac{\eta_i(s_i)}{1 - s_i} \left(\sum_{j=0}^{i-1} \frac{|s_j - s_i|}{\mu^j} \right)^2 \right)$$

18

where the ψ_{ij} are the functions defined in (33) and $\eta_i(s) = m_i - \phi_i'(s)$.

Convergence (36) is proven by induction over L. The case $L = 1$ is covered by Theorem 4.1. Now assume that, for any $j \in \{1, \ldots, i-1\}$ it is known that $\frac{Q_{j0}}{\mu^j} \to \chi_j$. We prove that $\frac{Q_{i0}}{\mu^i} \to \chi_i$ using (25). First observe that $\frac{Q_{i0}}{\mu^i}$ is bounded away from 0, since $\psi_{i,i-1}(1) > 0$, there exists a $k \geq 1$ such that $p_i(k)a_{i,i-1}(k) > 0$ and the inequality

$$Q_{i0} \geq p_i(k)k\mu a_{i,i-1}^{(\mu)}(k)Q_{i-1,0}(1 - \mu a_{i,i-1}^{(\mu)}(k))^{k-1}$$

implies $\limsup \mu^i/Q_{i0} < \infty$ due to the induction assumption. Therefore, (37) gives

$$\left| f_i^{(\mu)}(0, Q_{11}, \ldots, Q_{i1}) - \phi_i(Q_{i1}) - \sum_{j=0}^{i-1} \mu^{i-j}(Q_{j1} - Q_{i1})\psi_{ij}(Q_{i1}) \right| =$$

$$= o\left(\sum_{j=0}^{i-1} \mu^{i-j}|Q_{j1} - Q_{i1}| \right) + O\left(\mu^i \eta^{(i)}(Q_{i1}) \left[\sum_{j=0}^{i-1} \frac{|Q_{j1} - Q_{i1}|}{\mu^j} \right]^2 \right)$$

which combined with (25) yields that $Q_{i1} \to 1$ and

$$1 - Q_{i0} - \phi_i(1 - Q_{i0}) - \sum_{j=0}^{i-1} \mu^{i-j}(Q_{i0} - Q_{jo})\psi_{ij}(1) =$$

$$= o\left(\sum_{j=0}^{i-1} \mu^{i-j}|Q_{i0} - Q_{j0}| \right) + o\left(\mu^i \left[\sum_{j=0}^{i-1} \frac{|Q_{i0} - Q_{j0}|}{\mu^j} \right]^2 \right).$$

It follows

$$\phi_i(1 - Q_{i0}) - 1 + Q_{i0} = \mu^i \sum_{j=0}^{i-1} \chi_j \psi_{ij}(1) + O(\mu Q_{i0})$$

and (36) for $L = i$ is derived from (34).

In order to obtain the generating functions of the reproduction law of the limit process we need an extension of (24): if $0 \leq s_j \leq s_j + \delta_j \leq 1$, then

$$0 \leq f_i^{(\mu)}(s_0 + \delta_0, \ldots, s_i + \delta_i) - f_i^{(\mu)}(s_0, \ldots, s_i) - R_i^{(\mu)}(s_0, \ldots, s_i, \delta_0, \ldots, \delta_i)$$

$$\leq \sum_{k=0}^{\infty} kp_i(k)d_{k,\mu}^2(\delta_0, \ldots, \delta_i)\frac{1 - d_{k,\mu}^{k-1}(s_0, \ldots, s_i)}{1 - d_{k,\mu}(s_0, \ldots, s_i)} \tag{38}$$

where $d_{k,\mu}(s_0, \ldots, s_i) = s_i + \sum_{j=0}^{i-1} \mu^{i-j}(s_j - s_i) a_{ij}^{(\mu)}(k)$ and

$$R_i^{(\mu)}(s_0, \ldots, s_i, \delta_0, \ldots, \delta_i) = \sum_{k=1}^{\infty} k p_i(k) d_{k,\mu}^{k-1}(s_0, \ldots, s_i) d_{k,\mu}(\delta_0, \ldots, \delta_i).$$

From (27) and (38) it follows that

$$Q_{i0} E_{i0} \left[s_{i0}^{Z_{i0}(1)} s_{i1}^{Z_{i1}(1)} \cdots s_{10}^{Z_{10}(1)} s_{11}^{Z_{11}(1)} \right]$$
$$= R_i^{(\mu)}(0, Q_{11} s_{11}, \ldots, Q_{i1} s_{i1}, 1, Q_{10} s_{10}, \ldots, Q_{i0} s_{i0}) + O(\mu^{2i})$$

and it remains to check that

$$Q_{i0}^{-1} R_i^{(\mu)}(0, Q_{11} s_{11}, \ldots, Q_{i1} s_{i1}, 1, Q_{10} s_{10}, \ldots, Q_{i0} s_{i0})$$
$$\rightarrow s_{i0} \phi_i'(s_{i1}) + \sum_{j=0}^{i-1} s_{j0} \frac{\chi_j}{\chi_i} \psi_{ij}(s_{i1})$$
$$= m_i s_{i0} \frac{\phi_i'(s_{i1})}{m_i} + (1 - m_i) \sum_{j=0}^{i-1} B_{ij} s_{j0} \frac{\psi_{ij}(s_{i1})}{\psi_{ij}(1)}.$$

$\square$

8 The total time to escape

For application purposes, it is important to study the waiting time W_L to produce the escape type along an asymptotically viable path of mutations. For the sequential mutation model studied in Sections 6 and 7, W_L is a sum of a random number of independent geometric random variables. In terms of a Markov chain $\{Y(n)\}_{n \geq 0}$ with the transition matrix

$$\mathbb{D} = [D_{ij}]_{i,j=0}^{L}, \quad D_{ij} = (1 - m_i) B_{ij} + m_i 1_{\{i=j\}}$$

this is the waiting time until absorption at state 0

$$P(W_L \leq n) = P(Y(n) = 0 | Y(0) = L).$$

The last probability is the element $D_{L0}^{(n)}$ of the nth step transition matrix $\mathbb{D}^n$ which can be computed from the Chapman-Kolmogorov equation

$$\begin{aligned} D_{L0}^{(n)} &= D_{L0} D_{00}^{(n-1)} + \ldots + D_{LL} D_{L0}^{(n-1)} \\ &= (1 - m_L) \left(B_{L0} + B_{L1} D_{10}^{(n-1)} + \ldots + B_{L,L-1} D_{L-1,0}^{(n-1)} \right) + m_L D_{L0}^{(n-1)}. \end{aligned}$$

Subtracting a similar formula for $D_{L0}^{(n-1)}$ we get a recursion for the probability $P_L(n) = P(W_L = n)$

$$P_L(n) = m_L P_L(n-1) + (1-m_L)\sum_{j=1}^{L-1} B_{Lj} P_j(n-1).$$

Turning to the expected waiting time

$$M_L = E(W_L) = \sum_{n=1}^{\infty} n P_L(n)$$

we derive

$$
\begin{aligned}
M_L &= \frac{1}{1-m_L} + \sum_{j=1}^{L-1} B_{Lj} M_j \\
&= \frac{1}{1-m_L} + \sum_{j=1}^{L-1} \frac{B_{Lj}}{1-m_j} + \sum_{j=1}^{L-2} B_{Lj}^{(2)} M_j \\
&= \frac{1}{1-m_L} + \sum_{j=1}^{L-1} \frac{B_{Lj} + B_{Lj}^{(2)} + \ldots + B_{Lj}^{(L-j)}}{1-m_j} \\
&= \frac{1}{1-m_L} + \sum_{j=1}^{L-1} \frac{\chi_j(A_{Lj} + \ldots + A_{Lj}^{(L-j)})}{\chi_L(1-m_j)}.
\end{aligned}
$$

Observe that the last formula is a weighted sum of the individual waiting times $E(T_j) = \frac{1}{1-m_j}$. The corresponding weight

$$
\begin{aligned}
\frac{\chi_j}{\chi_L}(A_{Lj} + \ldots + A_{Lj}^{(L-j)}) &= \frac{A_{Lj} A_{j0}^{(j)} + \ldots + A_{Lj}^{(L-j)} A_{j0}^{(j)}}{A_{L0}^{(L)}} \\
&= P(Y(n) = j \text{ for some } n) \qquad (39)
\end{aligned}
$$

gives the probability that the chain $Y(n)$ visits the state j before it is absorbed at 0. Notice that in the case of "neutral mutation" with $m_j = m$, $j = 1, \ldots, L$ we get

$$
\begin{aligned}
M_L &= \frac{1}{1-m} + \frac{1}{1-m}\sum_{j=1}^{L-1} \frac{\chi_j(A_{Lj} + \ldots + A_{Lj}^{(L-j)})}{\chi_L} \\
&= \frac{1}{1-m}\left(1 + \frac{\chi_L - A_{L0} + \ldots + \chi_L - A_{L0}^{(L-1)}}{\chi_L}\right) \\
&= \frac{1}{1-m}\left(L - \frac{A_{L0} + \ldots + A_{L0}^{(L-1)}}{\chi_L}\right).
\end{aligned}
$$

21

Finally, we describe a case where there is a simple formula for the coefficients χ_i. Suppose that $a_{ij}(k) \equiv a_i(k)$ is the same for all daughter types j given the mother type i. Then with simplified notation $\psi_{ij}(1) = c_i$ we obtain

$$\chi_i = \frac{c_i}{1-m_i}\left(1 + \sum_{k=1}^{i-1}\sum_{0<j_1<j_2<\dots<j_k<i}\frac{c_{j_k}}{1-m_{j_k}}\cdots\frac{c_{j_1}}{1-m_{j_1}}\right)$$

$$= \frac{c_i}{1-m_i}\left(1 + \frac{c_{i-1}}{1-m_{i-1}}\right)\cdots\left(1+\frac{c_1}{1-m_1}\right).$$

In this case we can also compute the asymptotic probability (39) that the random path from type L towards type 0 visits type j

$$P(Y(n) = j \text{ for some } n)$$

$$= \frac{\chi_j}{\chi_L}\frac{c_L}{1-m_L}\left(1 + \sum_{k=1}^{L-j}\sum_{j<j_1<j_2<\dots<j_k<L}\frac{c_{j_k}}{1-m_{j_k}}\cdots\frac{c_{j_1}}{1-m_{j_1}}\right)$$

$$= \frac{c_j}{1-m_j}\left(1 + \frac{c_{L-1}}{1-m_{L-1}}\right)^{-1}\cdots\left(1+\frac{c_j}{1-m_j}\right)^{-1}$$

$$\times \left(1 + \frac{c_{L-1}}{1-m_{L-1}}\right)\cdots\left(1+\frac{c_{j+1}}{1-m_{j+1}}\right)$$

$$= \frac{c_j}{1+c_j-m_j}.$$

Thus the expected total time to escape becomes

$$M_L = \frac{1}{1-m_L} + \sum_{j=1}^{L-1}\frac{c_j}{(1+c_j-m_j)(1-m_j)}.$$

In particular, if $a_{ij}(k) \equiv 1$, then $c_j = m_j$ and

$$M_L = \frac{1}{1-m_L} + \sum_{j=1}^{L-1}\frac{m_j}{1-m_j}.$$

If furthermore $m_j \equiv m$, then $\chi_j = m(1-m)^{-j}$ and $P(Y(n) = j \text{ for some } n) = m$. In this special case the number of intermediate types has a binomial distribution $\text{Bin}(L-1, m)$ and

$$M_L = \frac{1+(L-1)m}{1-m}.$$

9   The network mutation model

We now return to the network model described in the introduction. This model was introduced in the papers by Iwasa et al [4] and [5]. Here particles are coded with binary sequences $\bar{u} = (u_1, \ldots, u_L)$ of length L. A mutation occurs if one of the L sites changes from 1 to 0 or 0 to 1. Therefore there are 2^L possible sequences which we will group into $L + 1$ types each containing sequences $\bar{u}$ with the same number of ones $|\bar{u}| = u_1 + \ldots + u_L$. Assuming that all sequences within a type i have the same offspring number distribution described by generating function $\phi_i(s)$, we arrive at an important example of the sequential mutation model allowing for backward mutations.

Let $m_i = \phi_i'(1) \in (0, 1)$ be the mean offspring number for the virus of type i whose sequence contains $i \in [1, L]$ ones and $(L - i)$ zeros. The sequence with all zeros $\bar{0} = (0, \ldots, 0)$ will be assigned a supercritical reproduction number $m_0 \in (1, \infty)$. Given mutation rate μ per site per generation the mutation probability between two sequences which differ in j sites becomes $\mu^j (1 - \mu)^{L-j}$. Clearly, for $j < i$ the mutation probability between types i and j is asymptotically equivalent to $\binom{i}{j} \mu^{i-j}$ and does not depend on the family size. Thus

$$\psi_{ij}(s) = \binom{i}{j} \phi_i'(s), \quad \psi_{ij}(1) = \binom{i}{j} m_i,$$

implying

$$
\begin{aligned}
A_{ij} &= \frac{m_i}{1 - m_i} \binom{i}{j} 1_{\{0 \le j \le i-1\}} + 1_{\{i=j=0\}}, \\
B_{ij} &= \frac{\chi_j}{\chi_i} A_{ij},
\end{aligned}
$$

where $(\chi_0, \ldots, \chi_L)$ is defined recursively by (34).

A proper extension of Theorem 7.1 allowing for backward mutations provides an asymptotic picture of the network mutation model conditional on escape. On the sequence level the limit process starts with the sequence $\bar{1} = (1, \ldots, 1)$ initiating a "stem lineage". Each next generation the stem sequence either remains to be $\bar{1}$ with probability m_L or turns to a 0-1 sequence $\bar{u}_1 = (u_{11}, \ldots, u_{1L})$ with probability

$$P(\bar{1} \to \bar{u}_1) = \frac{(1 - m_L) B_{L, i_1}}{\binom{L}{i_1}}, \quad i_1 = |\bar{u}_1|.$$

For a geometric number of generations $T_L \sim \text{Geom}(1 - m_L)$ including the time it switches to $\bar{u}_1$, the stem sequence $\bar{1}$ produces random numbers of side lineages of mutation-free $\bar{1}$-viruses. The number of such lineages per

generation has generating function $\phi'_L(s)/m_L$. Each mutation-free $\bar{1}$-lineage is a single type GW process with the offspring generating function $\phi_L(s)$.

The mutant stem sequence $\bar{u}_1$ follows the same pattern only with L replaced by i_1 which is strictly less than L, unless $i_1 = 0$ in which case the system stops after hitting the escape form $\bar{0}$ of the virus. Thus the sequence dynamics from $\bar{1}$ towards $\bar{0}$ is described by a random path $\bar{1} \to \bar{u}_1 \to \ldots \to \bar{u}_k = \bar{0}$ of a random length $k \in [1, L+1]$ through intermediate sequences with strictly decreasing numbers of ones $L > i_1 > \ldots > i_k = 0$, $i_j = |\bar{u}_{i_j}|$. The random path forms a Markov chain with transition probabilities

$$P(\bar{u}_j \to \bar{u}_l) = \frac{(1 - m_{i_j}) B_{i_j, i_l}}{\binom{i_j}{i_l}}.$$

The stem lineage spends at the type $\bar{u}_j$ a geometric number of generations with mean $\lambda(i_j)$, where $\lambda(i) = \frac{1}{1-m_i}$. During this time it generates mutation-free $\bar{u}_j$-lineages. The number of such lineages per generation has generating function $\phi'_{i_j}(s)/m_{i_j}$. Each mutation-free $\bar{u}_j$-lineage is a single type GW process with the offspring generating function $\phi_{i_j}(s)$.

References

[1] Athreya, K. and Ney, P. Branching processes. John Wiley & Sons, London-New York-Sydney, 1972.

[2] Geiger, J. Elementary new proofs of classical limit theorems for Galton-Watson processes. J. Appl. Probab. 36 (1999), 301-309.

[3] Haccou P., Jagers P., Vatutin V.A. Branching Processes: Variation, Growth and Extinction of Populations. Cambridge University Press, Cambridge, 2005.

[4] Iwasa, Y., Michor, F., Nowak, M.A., 2003. Evolutionary dynamics of escape from biomedical intervention. Proc. Roy. Soc. London B 270, 2573-2578.

[5] Iwasa, Y., Michor, F., Nowak, M.A., 2004. Evolutionary dynamics of invasion and escape. Journal of Theoretical Biology 226, 205-214.

[6] Lyons, R., Pemantle, R., and Peres, Y. Conceptual proofs of $l\log l$ criteria for mean behavior of branching processes. Ann. Probab. 23 (1995) 1125-1138.

[7] Serra, M.C. and Haccou, P. Dynamics of escape mutants. Theoretical Population Biology, 72, (1), (2007), 167-178.

[8] Sevastyanov, B. A. Vetvyashchiesya protsessy (in Russian). Nauka, Moscow, 1971. Also available in German: Sewastjanow B. A. Verzweigungsprozesse. Akademie-Verlag, Berlin, 1974.