THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

A Multitype Branching Processes Approach to the Evolutionary Dynamics of Escape

Maria Conceição Serra

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and Göteborg University
Göteborg, Sweden, 2007

A Multitype Branching Processes Approach to the Evolutionary Dynamics of Escape
Maria Conceição Serra
ISBN 978-91-7291-997-6

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie Nr 2678
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000

Printed at the Department of Mathematical Sciences
Göteborg, Sweden, 2007

A Multitype Branching Processes Approach to the Evolutionary Dynamics of Escape

Maria Conceição Serra

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology
Göteborg University

Abstract

Evolutionary dynamics of escape is a recent development in theoretical biology. It is an attempt to predict possible patterns of population dynamics for a certain strain of viruses placed in a hostile environment. The only way to escape extinction for the virus is to find a new form better adapted to the new environment. This is usually achieved by mutations in certain positions of the genome.

In this thesis we use multitype Galton-Watson branching processes to model the evolution of such virus populations and provide answers to some of the most relevant questions arising in them.

We determine the asymptotic probability of escape for a population stemming from a single progenitor. The calculations are obtained assuming mutations are rare events and generalize results previously known for particular reproduction laws.

We also give a description of the random path to escape, that is the chain of mutations leading to the escape form of the virus. Using this description, we also study the waiting time to escape, i.e., the time it takes to produce the escape form of the virus. We start by deriving results for simple populations allowing for two-types of individuals and simple mutation schemes. Later we perform asymptotic analysis, again assuming mutations are rare, for populations with quite general reproduction and mutation schemes.

Keywords: Galton-Watson branching processes, multitype, decomposable processes, population dynamics, extinction, mutation, path to escape, waiting time to escape.

The research presented in this work was financially supported by the *Fundação para a Ciência e Tecnologia* through the doctoral scholarship SFRH/BD/18995/2004, funded by the programme POCI 2010 and the European Social Fund.

Ciência.Inovação
2010
Programa Operacional Ciência e Inovação 2010
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

UNIÃO EUROPEIA
Fundo Social Europeu

Acknowledgments

First and most I would like to thank my advisers Serik Sagitov and Peter Jagers for taking the time to help me, for many interesting discussions, ideas and continuous encouragement. The opportunity to travel and exchange knowledge with researchers from other centers is also acknowledged.

Many thanks go also to Patsy Haccou for being so patient with my questions regarding biological matters. Thank you also for the hospitality while I was visiting Leiden.

I also would like to thank Prof. Maria Lucília Carvalho, from Lisbon University, for introducing me to the field of branching processes and for encouraging me to start the studies in Sweden. Probably, without her this work would not have not been possible.

During the years I spent in Sweden, I had the opportunity to find many new friends. Thank you for sharing the good and the not so good moments with me. Special thanks go to Gwenaëlle and Anastassia for such a good friendship and great support.

I also want to thank my friends back home for their continuous friendship, encouragement and for shortening, at least for me, the distance between Portugal and Sweden.

My family has always been supportive of my decisions no matter how strange they seemed and even when they required a long absence. Thank you for doing this, not only during the last years, but also my entire adult life.

Obrigada!

Contents

List of Papers

This thesis is based on the following papers, which will be referred throughout the text by the corresponding letters:

Paper A M.C. Serra (2006). On the waiting time to escape.
In: *Journal of Applied Probability*, **43**, 296-302.

Paper B M.C. Serra and P. Haccou (2007). Dynamics of escape mutants.
In: *Theoretical Population Biology*, **72**, 167-178.

Paper C S. Sagitov and M.C. Serra (2007). Multitype Galton-Watson processes escaping extinction.
Preprint 2007:25, Chalmers University of Technology and Göteborg University (submitted)

1 Introduction

Evolutionary dynamics of escape is a recent development in theoretical biology mainly due to Iwasa and coauthors. It is an attempt to predict possible patterns of population dynamics for a certain strain of viruses placed in a hostile environment. The only way to escape extinction for the virus is to find a new form better adapted to the new environment. This is usually achieved by mutations in certain positions of the genome.

The different forms of the virus can be labeled by binary sequences, $\bar{s} = (s_1, \ldots, s_L)$ of a given length, L. The initial form, of the virus placed in the new environment, is conveniently described by sequence $\bar{1} = (1, \ldots, 1)$ and the escape form (the one allowing for non-extinction) is then denoted by $\bar{0} = (0, \ldots, 0)$. Hence, in total there will be 2^L possible forms of the virus. With the exception of $\bar{0}$, all the forms are vulnerable in the sense that, if no mutations occur, a reproduction process stemming from any of them is doomed to extinction. So, in order to survive, the virus has to perform a chain of mutations starting from $\bar{1}$ and leading to $\bar{0}$. This can be illustrated with a graph, see Figure 1 below, where the vertices represent the different forms of the virus and the edges stand for the possible mutations.

The major issues arising in this setting are:

Question I What is an appropriate reproduction-mutation model?

Question II Within the chosen model compute (asymptotically) the probability of escape for a population stemming from a single progenitor.

Question III Describe the random path to escape, that is the chain of mutations leading to the escape form of the virus.

Question IV Assess the time to escape, that is the waiting time to produce the escape form of the virus.

When addressing the principal Question I, Iwasa et al. have chosen the framework of branching processes to model biological populations of interest, see [3] and [4]. This choice seems relevant for two reasons. Firstly, the key assumption of no interaction between individuals is not restrictive, since viruses reproduce asexually and, under the hostile environment, the population size is expected to be low enough to neglect possible interaction caused by overpopulation. The other obvious reason is that the theory of branching processes is well developed and provides with various deep results concerning a broad range of population models. In particular, the use of

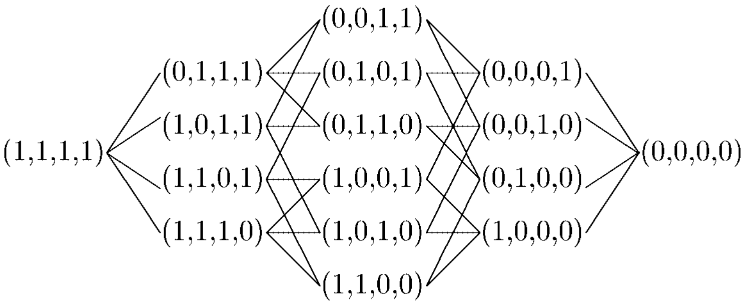

Figure 1: The network of sequences of length $L = 4$. Here the edges represent single point mutations of probability $\mu(1-\mu)^{L-1}$, where μ is the mutation probability per site. In the full network of possible mutations every pair of sequences is connected by an edge. The mutation probability between any two sequences $\bar{s}$ and $\bar{t}$ equals $\mu^{h_{st}}(1-\mu)^{L-h_{st}}$, where h_{st} is the Hamming distance between $\bar{s}$ and $\bar{t}$.

multitype branching processes is appropriate since the different forms of the virus can be associated with different types of individuals.

The branching model considered by Iwasa et al. was limited to two particular reproduction laws, Poisson and geometric. They addressed Question II and touched slightly Question III; see Section 2 for a description of their work.

One of the goals of this thesis is to extend the scope of the branching processes leading to the same kind of results. We believe this is important since in most of the applications the reproduction law is not really known. Another goal is to perform deeper analysis of Questions II-III and use the results to address Question IV for the first time.

Our approach is restricted to discrete time processes, known as multitype Galton-Watson branching processes (GWBP). We assume that each individual lives one unit of time and gives birth to a random number of offspring that may be of a different type. Their descendants will form the next generation and so on. Besides independence, individuals of the same type are assumed to reproduce according to a common distribution.

The use of discrete time models is questionable and this time structure seems artificial. Yet, mathematically discrete time models are much easier to handle than the ones in continuous time and for some problems, like extinction, it is possible to deduce results that hold for the continuous time processes.

Obviously, at a later stage, it would be interesting to extend the results of this thesis to continuous time branching processes. It is also desirable that other models allowing for density dependence, recombination and gene transfer are explored.

The example of the virus population introduced in the beginning of this section captures the essence of the problem of the evolution to escape, but it is not unique; see [3], [4] and Paper B of this thesis. Next we mention some more examples.

A typical example is that of cancer cells submitted to chemotherapy. The aim of the therapy is to reduce the basic reproductive ratio of the infected cells to less than one and eliminate cancer from the patient. Unfortunately, mutations in the cells often provide resistance to the therapy. This new type of cells has a higher reproductive ratio and can lead to failure of the treatment.

Another interesting example concerns the RNA first hypothesis. It claims that in the primordial "soup" there existed free-floating nucleotides which formed bounds with one another. The majority of these chains broke down and were not capable of replication. However certain sequences of base pairs had catalytic properties that allow them to stay together for longer periods of time. These chains are proposed to be the first primitive forms of life. Hence, in the RNA world, different forms of RNA competed with each other for free nucleotides and were subject to natural selection. The most efficient molecules of RNA, the ones capable of of efficiently catalyze their own reproduction, survived and evolved, forming todays RNA.

An important environmental problem is the spread of insecticide resistance due to hybridization between resistant and susceptible insects. Although matings between different biotypes are rare, hybrid formation can occur. The initial hybrids have low fitness, but after several backcrosses between resistant and sensitive biotypes, a successful resistant hybrid may be produced.

Finally, in agriculture there are situations where introgression may occur between genetically modified or cultivated organisms with the established wild populations. Usually the first backcrosses are not viable or fertile, but eventually a successful type may be produced.

The examples above have the following common features. Due to a small reproductive ratio of the individuals, the populations are in principle doomed to extinction. Yet, changes occurring during the reproduction process may lead to the appearance of individuals with higher reproductive ratio, making it possible for the population to escape extinction. In a general way, these are the kind of populations we had in mind while developing this work.

The outline of the thesis is the following. In Section 2 we give an overview of previous work done by Iwasa et al. Some background on GWBP is given in Section 3. In Section 4 we summarize the results of the three papers forming the main body of this thesis.

2 Principles on Evolutionary Dynamics of Escape

This section is dedicated to a detailed description of the results in Iwasa et al. papers, [3] and [4], that are relevant for the thesis. When possible and appropriate, we will also indicate the connection with the results in the appended papers.

Iwasa et al. use multitype GWBP where the offspring distribution is either geometric or Poisson, but the results apply also to continuous time branching processes where individuals either die or produce just one child.

In the more general setting, they consider a population where it is possible to distinguish between $k+1$ types of individuals, labeled $0, 1, \ldots, k$, with reproductive ratios $m_0, m_1, \ldots, m_k$, respectively. Suppose the number of offspring of a type i individual follows a geometric distribution with mean m_i. Types $1, \ldots, k$ have reproductive ratio less than 1 but $m_0 > 1$. The mutation scheme is the following: each one of the daughters of type i either mutates to a type $j \neq i$ with probability $0 \leq u_{ij} < 1$ or remains in type i with probability

$$u_{ii} = 1 - \sum_{\substack{j=0 \\ j \neq i}}^{k} u_{ij}.$$

Mutations are assumed to be independent and therefore, if type i produces n daughters, the number of daughters which are of type j has binomial distribution with parameters n and u_{ij}.

For the asymptotic analysis a key parameter u, characterizing the overall mutation rate, is introduced. It tends to 0 reflecting the fact that mutations are rare. In the virus population described earlier this parameter can be replaced by μ, the probability of mutation per site. Assuming that $u_{ij} = O(u)$, as $u \to 0$, it is proved that the column vector $\Xi = [\xi_1, \ldots, \xi_k]^{\mathrm{T}}$, where ξ_i denotes the probability of non-extinction of a population starting with a single individual of type i, satisfies approximately the following linear system of equations

$$\Xi = \mathbb{D}\mathbb{U}\Xi + \mathbb{D}\mathbf{u}_0\,\xi_0 \tag{1}$$

with $\mathbf{u}_0 = [u_{10}, \ldots, u_{k0}]^{\mathrm{T}}$, $\mathbb{D}$ the diagonal matrix with diagonal elements $\frac{m_1}{1-m_1}, \ldots, \frac{m_k}{1-m_k}$, $\mathbb{U} = [U_{ij}]_{i,j=1}^{k}$ the $k \times k$ matrix with elements

$$U_{ij} = u_{ij}1_{i \neq j}$$

and ξ_0 the probability of non-extinction if the processes starts with a single type 0 individual. Observe that ξ_0 is approximately equal to $1 - \frac{1}{m_0}$, which is precisely the non-extinction probability of a single-type GWBP with a geometric reproduction law with mean m_0. This means that, asymptotically, there will be no mutation from type 0 to any other types.

If the number of offspring follows a Poisson distribution, the linear system (1) still holds but ξ_0 is now, approximately, the unique solution of equation

$$\log(1-s) = -sm_0$$

in interval $(0,1)$.

The solution of system (1) is given by the formula

$$\Xi = [\mathbb{D} + \mathbb{D}\mathbb{U}\mathbb{D} + \mathbb{D}\mathbb{U}\mathbb{D}\mathbb{U}\mathbb{D} + \ldots]\,\mathbf{u}_0\,\xi_0 \tag{2}$$

which provides a clear picture of how a population starting from a single individual manages to escape extinction. Suppose the population starts with one type i individual. The i-th diagonal element of matrix $\mathbb{D}$ when multiplied by u_{ij} represents, asymptotically, the expected number of mutants of type j produced during the progenitor's lifetime. The first vector, $\mathbb{D}\mathbf{u}_0\,\xi_0$, represents the direct mutation from type i to the escape type 0. Hence, on average there will be $\frac{m_i}{1-m_i}u_{i0}$ straight mutations and at least one of them should lead to non-extinction, ξ_0. Second vector, $\mathbb{D}\mathbb{U}\mathbb{D}\,\mathbf{u}_0\,\xi_0$, stands for the possibility of mutation from i to an intermediate type $j \neq 0$ followed by direct mutation from j to 0. Again, on average, there will be $\frac{m_i}{1-m_i}u_{ij}$ individuals mutating from i to j and each one of them produces, on average, $\frac{m_j}{1-m_j}u_{j0}$ escape type individuals. Therefore, the i-th element of vector $\mathbb{D}\mathbb{U}\mathbb{D}\,\mathbf{u}_0\,\xi_0$

$$\sum_{\substack{j=1 \\ j \neq i,0}}^{k} \frac{m_i}{1-m_i}u_{ij}\frac{m_j}{1-m_j}u_{j0}\xi_0$$

is the asymptotic probability of escaping extinction with one intermediate mutation. The third vector represents the possibility of reaching the escape type trough two intermediate mutations and so on. Summation over all the vectors represents summation over the number of intermediate types visited before reaching the escape type and gives the total probability of escape.

Looking at (1) and (2), we see that the approximate values of the probabilities of escape depend only on the reproduction means and on the mutation probabilities. Hence, it seems that such approximations should hold for any offspring distribution; not only geometric and Poisson. In Paper B we prove that this is true if the variance of the offspring distribution is finite.

Particular attention is given to the so-called network mutation model where individuals are described as binary sequences of length L; just like the virus example given in the Introduction. If two sequences differ in at least one site, the corresponding individuals are considered to be of different types and therefore there exists a total of 2^L different types. Only type 0, described by sequence $\bar{0}$ is an escape mutant. Hence the reproductive ratios are $m_i < 1$, $i = 1, \dots, 2^L - 1$, and $m_0 > 1$. The mutation probabilities between the types are expressed through a parameter $\mu \in (0,1)$ which is the probability of mutation per site. Assuming that each site mutates independently of the others, then this is a particular model of the previous one, with the mutation probabilities being $u_{ij} = \mu^{h_{ij}}(1-\mu)^{L-h_{ij}}$, where h_{ij} is the Hamming distance between any two sequences i and j. The Hamming distance between two sequences is just the number of sites at which they differ.

In this particular case, the probability of escape starting from a single individual of type i is $O(\mu^{h_{i0}})$. In fact, the use of equation (2) yields, for any $i = 1, \dots, 2^L - 1$,

$$\xi_i = \xi_0 \sum_{p \in P_i} v(p) \tag{3}$$

where P_i is the set of all paths connecting sequence i and the escape sequence in a number of mutational steps equal to h_{i0} and $v(p)$ is the so-called value of the path p

$$v(p) = \mu^{h_{i0}} \prod_{j=1}^{g-1} \frac{m_{k_j}}{1 - m_{k_j}}, \tag{4}$$

when $p : i \equiv k_1 \to k_2 \to k_3 \to \dots \to k_g \equiv 0$. Observe that set P_i contains only paths formed by sequences with a strictly decreasing number of ones.

Let us have a closer look at the case $L = 2$. Individuals described by sequence $(0,0)$ are of the escape type and the other individuals, $(0,1), (1,0)$ and $(1,1)$, have reproductive ratios m_{01}, m_{10} and m_{11} which are less than one. To calculate the probability of escape, starting from an $(1,1)$ individual, one has to consider only paths connecting sequences $(1,1)$ and $(0,0)$ in 2 mutational steps. There are 3 such paths:

$$\begin{aligned} p_1 : (1,1) &\to (1,0) \to (0,0) \\ p_2 : (1,1) &\to (0,1) \to (0,0) \\ p_3 : (1,1) &\to (0,0) \end{aligned}$$

with the following path values,

$$\begin{aligned} v(p_1) &= \mu \frac{m_{11}}{1-m_{11}} \mu \frac{m_{10}}{1-m_{10}}, \\ v(p_2) &= \mu \frac{m_{11}}{1-m_{11}} \mu \frac{m_{01}}{1-m_{01}}, \\ v(p_3) &= \mu^2 \frac{m_{11}}{1-m_{11}}. \end{aligned}$$

Each one of these values is of order μ^2 so that the escape probability is approximately

$$\xi_{11} = \xi_0 [v(p_1) + v(p_2) + v(p_3)].$$

Observe that, because μ is small, the paths connecting sequence $(1,1)$ and the escape sequence in a number of steps larger than the Hamming distance lead to smaller terms in the probability of non-extinction and are therefore neglected. For instance, possible paths such us

$$\begin{aligned} p_4 : (1,1) &\to (1,0) \to (0,1) \to (0,0) \\ p_5 : (1,1) &\to (0,1) \to (1,0) \to (0,0) \end{aligned}$$

have 4 mutational steps and values of order μ^4

$$\begin{aligned} v(p_4) &= \frac{m_{11}}{1-m_{11}} \mu \frac{m_{10}}{1-m_{10}} \mu^2 \frac{m_{01}}{1-m_{01}} \mu, \\ v(p_5) &= \frac{m_{11}}{1-m_{11}} \mu \frac{m_{01}}{1-m_{01}} \mu^2 \frac{m_{10}}{1-m_{10}} \mu. \end{aligned}$$

An important conclusion to draw from here is that simultaneous mutation at all sites can be as important as the successive one step mutations since they have the same order of magnitude with respect to μ. The factors $\frac{m_i}{1-m_i}$, $i \neq 0$, favor the paths through the types with larger m_i.

In the particular network model considered in Paper C we show that, conditioned on the appearance of an escape mutant, the waiting time is a sum of a random number of independent geometric random variables. The factor above, $\frac{m_i}{1-m_i}$, $i \neq 0$, is actually the expectation of the time spent at type i on the path to escape. It is also worth to point out here that, when considering a sequential mutation model, formula (3) appears as (36) in Theorem 7.1 of Paper C.

3 Background on Galton-Watson Processes

From a theoretical point of view, branching processes have been intensively studied during the last decades; classical references are the books of Harris (1963), Sevastyanov (1971), Athreya and Ney (1972) and Jagers (1975). With a special emphasis on applications, recent books are Axelrood and Kimmel (2002) and Haccou, Jagers and Vatutin (2005). Throughout the literature it is possible to find examples of how these processes have been successfully used to solve important problems arising in different sciences such as medicine, biology, ecology, physics and even computer science.

In this section we describe the single-type and multitype Galton-Watson processes and give some background results for the appended papers. We will give some references but most of the material presented here can be found in the classical books referred above.

3.1 Single-Type Galton-Watson Processes

In the single type GWBP we assume that all individuals are of the same type. Although this is not the case of the populations we are interested in (they include different types of individuals), we can still use this model to a certain extent as well as many of its results. Besides, it is a model that deserves our attention in its own right and will help us understand the multitype models and more complicated structures.

A single-type GWBP is a sequence of random variables $\{Z_n, n \in \mathbb{N}_0\}$ defined by

$$\begin{cases} Z_0 = 1 \\ Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i}, \quad n \geq 0, \end{cases} \tag{5}$$

where $\xi_{n,i}, n, i \in \mathbb{N}_0$ are independent and identically distributed random variables with distribution $\{p_k, k \in \mathbb{N}_0\}$ (a sum from one to zero is assumed to be zero). To avoid trivialities, we assume $p_k \neq 1$, for all integers $k \geq 0$. Z_n denotes the size of the population in the n^{th} generation and condition $Z_0 = 1$ means that the process starts with just one individual. The size of the population in the first generation, Z_1, is the number of descendants of the initial progenitor (this is a realization of one r.v. with distribution $\{p_k, k \in \mathbb{N}_0\}$). These individuals reproduce, independently of each other and according to the same distribution, and their descendants will form the second generation, and so on. Due to its role, $\{p_k, k \in \mathbb{N}_0\}$ is the so-called offspring distribution or reproduction law.

$\{Z_n, n \in \mathbb{N}_0\}$ is a Markov chain with state space $\mathbb{N}_0$ and according to (5), for integers $j \geq 1$, $k \geq 0$,

$$P[Z_{n+1} = k | Z_n = j] = \sum_{i_1+i_2+\ldots+i_j=k} p_{i_1} p_{i_2} \cdots p_{i_j} = p_k^{*j}$$

where $\{p_k^{*j}, k \in \mathbb{N}_0\}$ denotes the j-fold convolution of $\{p_k, k \in \mathbb{N}_0\}$. Hence, for any $j \geq 0$, $k \geq 0$, the transition probabilities are

$$P[Z_{n+1} = k | Z_n = j] = \left\{ \begin{array}{ccl} p_k^{*j} & if & j \geq 1, k \geq 0 \\ 1 & if & j = 0, k = 0 \\ 0 & if & j = 0, k \geq 1 \end{array} \right. .$$

Since these probabilities do not depend on n, $\{Z_n, n \in \mathbb{N}_0\}$ is a homogeneous Markov chain. Plus, if $p_0 > 0$, 0 is an absorbing state and all the other states are transient. Regarding stationary distributions, these processes only have the trivial one: $(1, 0, 0, \ldots)$.

One of the most important properties of a GWBP is the additive property and it is used more or less explicitly throughout the thesis.

Additive Property: Let $\mathscr{B}(\mathbb{Z}_+^\infty)$ be the smallest σ-algebra of subsets of $\mathbb{Z}_+^\infty$ containing all finite dimensional sets and let also $\{Z_n^{(1)}\}, \{Z_n^{(2)}\}, \ldots,$ be independent GWBP with the same reproduction law as $\{Z_n\}$. Then, for any $r \in \mathbb{N}$ and for any $A \in \mathscr{B}(\mathbb{Z}_+^\infty)$

$$P[\{Z_n; n > r\} \in A | Z_r = k] = P\left[\left\{\sum_{j=1}^{k} Z_n^{(j)}; n \geq 1\right\} \in A\right].$$

This property tells us that, given that $Z_r = k$, the distribution of the process after generation r is the same as the distribution of the sum of k independent GWBP with the same reproduction law.

In the definition of a GWBP process we usually have $Z_0 = 1$. But this is not very realistic because we may have an initial population with more than one particle. If we assume that these particles reproduce independently of each other and according to the same reproduction law, the additive property allows us to make the appropriate adjustments to the case $Z_0 = k$, with $k > 1$.

An important tool in the analysis of these processes is the probability generating function (p.g.f.) of the reproduction law

$$f(s) = \sum_{k=0}^{\infty} p_k s^k, \qquad 0 \leq s \leq 1,$$

and its iterates

$$f_0(s) = s, \quad f_1(s) = f(s), \quad f_n(s) = \underbrace{(f \circ \ldots \circ f)}_{\text{n times}}(s).$$

From the nice properties of the p.g.f.'s it follows that random variable Z_n has p.g.f. given by

$$E[s^{Z_n}] = f_n(s), \quad 0 \leq s \leq 1.$$

Therefore,

$$E[Z_n] = m^n \qquad \text{and} \qquad Var[Z_n] = \begin{cases} \frac{\sigma^2 m^{n-1}(m^n - 1)}{m-1} & \text{if } m \neq 1 \\ n\sigma^2 & \text{if } m = 1 \end{cases}$$

where m and σ^2 are the mean and variance of the reproduction law

$$m = \sum_{k=0}^{\infty} k p_k \qquad \text{and} \qquad \sigma^2 = \sum_{k=0}^{\infty} (k - m)^2 p_k.$$

Next we consider the problem which originated the study of GWBP: the probability of extinction. This problem, posed by Francis Galton in 1873 and first studied by the Reverend Henry Watson, was in fact concerned with the extinction of family names in the British peerage. The event "extinction" is defined as

$$\bigcup_{n \in \mathbb{N}} [Z_n = 0].$$

The following theorem gives not only the probability of extinction but also the probability of the event $[Z_n \to \infty]$, to be called "escape". In the mathematical literature, $[Z_n \to \infty]$ is usually known as "explosion" but, in the biological context we consider, the name "escape" is natural because it means that the individuals (cancer cells, virus, etc...) were able develop resistance and escape extinction. The theorem can be found, for instance, in Jagers (1975).

Theorem 3.1 *The equation $f(s) = s$ has a unique solution in interval $[0,1)$ if $m > 1$, but if $m \leq 1$ and $p_1 \neq 1$ the unique solution in $[0,1]$ is $s = 1$. The extinction probability, denoted by q, is the smallest root of equation $f(s) = s$ and is such that*

$$\begin{aligned} m > 1 &\;\Rightarrow\; q < 1, \\ p_1 = 1 &\;\Rightarrow\; q = 0, \\ m \leq 1 \text{ and } p_1 < 1 &\;\Rightarrow\; q = 1. \end{aligned}$$

Further, if $p_1 \neq 1$, the probability of escape, denoted by r, is given by

$$r = 1 - q.$$

This theorem is frequently used in this thesis and, although it does not consider the possibility that $p_0 = 0$ it is easy to see that $q = 0$ if that is the case. It is also easy to understand that, if $p_1 = 1$ the process remains constant and equal to the initial population size, Z_0.

As we mentioned before, the discrete time structure is artificial but it is possible to determine the extinction probability of a continuous time branching process using the extinction probability of the corresponding embedded generation process. By embedded generation process we mean the discrete time process we obtain when we count the successive generation sizes in a continuous time branching process. It is easy to see that a population dies out if and only if its embedded generation process turns to zero.

As seen in the last theorem, the reproduction mean plays a crucial role in the study of a GWBP and therefore these processes are classified as *critical, subcritical* or *supercritical* if $m = 1$, $m < 1$ or $m > 1$, respectively.

In the populations we study we will often have individuals with reproduction mean smaller than one, and therefore they will be related with subcritical processes. Theorem 3.1 ensures that these processes die out almost surely and therefore the probability of survival until generation n decreases to 0, i.e., $P[Z_n > 0] \to 0$ as $n \to \infty$. Next theorem establishes the rate of convergence of this sequence to 0 and describes the limit behavior of a subcritical GWBP conditioned on survival until generation n. See Jagers (1975) for a slightly different version of the result.

Theorem 3.2 *In a subcritical process $\{Z_n, n \in \mathbb{N}_0\}$ with $Z_0 = 1$,*

$$\lim_{n\to\infty} \frac{P[Z_n > 0]}{m^n} = \begin{cases} 0 & if \quad \sum k\, p_k \log k = \infty \quad or \quad p_0 = 1 \\ C & if \qquad otherwise \end{cases}$$

where C is a positive constant. Furthermore, if $p_0 < 1$,

$$\lim_{n\to\infty} P[Z_n = k|Z_n > 0] = b_k$$

exists for $k \in \mathbb{N}$, with $\sum_{k=1}^{\infty} b_k = 1$, and, if $g(s) = \sum_{k=1}^{\infty} s^k\ b_k$, then g is the unique solution of equation

$$g \circ f = mg + 1 - m$$

among generating functions vanishing at zero. Plus,

$$\sum_{k=0}^{\infty} kb_k = \frac{1}{C} < \infty \quad \textit{iff} \quad \sum_{k=0}^{\infty} k\log(k)p_k < \infty.$$

The populations considered in this thesis usually start with an individual that has reproduction mean smaller than one and escape is possible only if individuals with reproduction mean larger than one are produced. The latter individuals will be related to supercritical processes. Next theorem establishes the limit behavior of such processes and can be found, for instance, in Haccou et al. (2005).

Theorem 3.3 *Let $\{Z_n, n \in \mathbb{N}_0\}$ be a supercritical process with $Z_0 = 1$ and such that the reproduction law satisfies $p_k \neq 1$, for any k, and*

$$\sum_{k=1}^{\infty} k\log(k)p_k < \infty \tag{6}$$

Then, as $n \to \infty$,

$$\frac{Z_n}{m^n} \longrightarrow W \ \textit{almost surely}$$

and W is such that $E[W] = 1$ and $P[W = 0] = q$. Furthermore, the Laplace transform of W, ϕ_W, satisfies

$$\phi_W(ms) = f(\phi_W(s)),\ \ s \geq 0.$$

Observe that, no matter the value of m, sequence Z_n/m^n is a martingale with respect to $\mathcal{B}_n = \sigma(Z_0, Z_1, \ldots, Z_n)$ and therefore converges almost surely to some random variable W. But this does not mean that W is not degenerated as it happens with the subcritical or critical case where

$P(W=0)=1$. Condition (6), which is slightly stronger than finite mean, is enough to ensure that $P(W=0)<1$ and that $E[W]=1$.

The last result of this section is the decomposition of a supercritical branching process in two parts: one that dies out almost surely and another one which is immortal; see Athreya and Ney (1972). In Paper C we obtain a different decomposition of a similar matter.

In a supercritical process, the extinction probability q is strictly less than 1, and $q>0$ if and only if $f(0)>0$. Assume $f(0)>0$ and let $\hat{f}$ be the following p.g.f.

$$\hat{f}(s)=\frac{f((1-q)s+q)-q}{1-q}, \quad 0\le s\le 1.$$

Let $\{Z_n, n\in\mathbb{N}_0\}$ and $\{\hat{Z}_n, n\in\mathbb{N}_0\}$ be supercritical GWBP on the same space $(\Omega,\mathcal{F})$ with generating functions f and $\hat{f}$, respectively, both starting with one individual. Denote by P and $\hat{P}$ the corresponding probability measures on $(\Omega,\mathcal{F})$. Observe that the first process has extinction probability $q>0$ but the second is immortal since $\hat{f}(0)=0$. Observe that $P(A\cup B)=1$ with

$$A=\{\omega\in\Omega:\ Z_n(\omega)\to\infty \text{ as } n\to\infty\}$$

and

$$B=\{\omega\in\Omega:\ Z_n(\omega)=0 \text{ for some } n>0\}.$$

By definition of B, we have $P(B)=q$ and from the assumption it follows that $0<P(A)<1$ (since $0<P(B)<1$). Now define

$$Z_n^{(1)}(\omega)=\begin{cases} 0 & \textit{if} \quad \omega\in B \\ \begin{array}{c}\text{the number of individuals of } Z_n(\omega)\\ \text{that have an infinite line of descent}\end{array} & \textit{if} \quad \omega\in A \end{cases}$$

As a consequence of this definition $Z_0^{(1)}(\omega)=1$, for $\omega\in A$.

Theorem 3.4 *Let $(A,\mathcal{F}_A,P_A)$ be the probability space with*

$$\mathcal{F}_A=\{A\cap F: F\in\mathcal{F}\}$$

σ-algebra over A and P_A the probability measure that assigns the value $P(E)/P(A)$ for any $E\in\mathcal{F}_A$. The stochastic processes $\{\hat{Z}_n, n\in\mathbb{N}_0\}$ on $(\Omega,\mathcal{F},\hat{P})$ and $\{Z_n^{(1)}, n\in\mathbb{N}_0\}$ on $(A,\mathcal{F}_A,P_A)$ are equivalent in the sense of finite dimensional distributions.

3.2 Multitype Galton-Watson Processes

In the multitype GWBP we allow for the existence of distinguishable individuals with different probabilistic behavior. This will be the case for the populations considered in the thesis since, during the reproduction process, individuals may give birth to individuals of their own type and also of other types.

To each individual we assign a type in a set $T = \{0, 1, \ldots, k\}$, the set of types, that is assumed to be finite. Each individual, say of type $r \in T$, is associated with a random vector $\xi_r = (\xi_r^0, \xi_r^1, \ldots, \xi_r^k)$, where ξ_r^j is a random variable that represents the number of children of type j born from a type r individual. Then, a multitype GWBP is a sequence of random vectors $\{(Z_n(0), Z_n(1), \ldots, Z_n(k)), n \in \mathbb{N}_0\}$, where $Z_n(j)$, $j \in T$, represents the number of individuals of type j in generation n.

Writing $m_{rj} = E[\xi_r^j]$, $r, j \in T$, we construct the first moment matrix in the following way:

$$\mathbb{A} = \begin{bmatrix} m_{00} & m_{01} & m_{02} & \ldots & m_{0K} \\ m_{10} & m_{11} & m_{12} & \ldots & m_{1K} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{k0} & m_{k1} & m_{k2} & \ldots & m_{kk} \end{bmatrix}.$$

If there exists $n \in \mathbb{N}$ such that $\mathbb{A}^n > 0$, the multitype GWBP is said to be positively regular. Most of the results available in the mathematical literature are concerned with these processes. However, multitype processes of interest for applications are often not positively regular and this will be the case in some of the populations we meet in the thesis, namely in Papers A and B. Nevertheless, in the more general setting of Paper C we allow for a positively regular process.

In these processes, the spectral radius of the matrix $\mathbb{A}$, denoted by ρ, plays the same crucial role as the reproduction mean, m, in the single-type processes. In fact, a positively regular multitype GWBP is classified as *critical, subcritical* or *supercritical* if $\rho = 1$, $\rho < 1$ or $\rho > 1$, respectively.

One of the reasons for such a classification is well illustrated in the next theorem, which is the analogue of Theorem 3.1 for multitype processes, and it can also be found in Jagers (1975). Before we state the theorem we need to introduce some notation. Let

$$\mathbf{s} = (s_1, \ldots, s_K) \in [0, 1]^K,$$

$$f^{(r)}(\mathbf{s}) = E\left[s_1^{\xi_r^1} s_2^{\xi_r^2} \ldots s_k^{\xi_r^K}\right], \quad r \in T,$$

$$\mathbf{f}(\mathbf{s}) = \left[f^{(1)}(s), \ldots, f^{(K)}(s) \right].$$

Theorem 3.5 *Assume* $\{(Z_n(0), Z_n(1), \ldots, Z_n(k)), n \in \mathbb{N}_0\}$ *is positively regular and nonsingular and let* q_r *denote the extinction probability when the process starts with a single individual of type* r, $r \in T$. *Then, the probability of extinction* $\mathbf{q} = [q_1, \ldots, q_K]$ *is the solution of equation* $\mathbf{f}(\mathbf{s}) = \mathbf{s}$ *that is closest to the origin in the unite cube* $[0,1]^K$. *If* $\rho \leq 1$ *then all* $q_r = 1$ *and* $\rho > 1$ *then all* $q_r < 1$.

The theorem excludes the singular processes which are processes such that $f(\mathbf{s}) = \mathbb{A}\mathbf{s}^{\mathrm{T}}$. In this processes there is no branching; each individual has exactly one daughter that can be of any type. Therefore, the process can be thought of an individual wondering between the types. This is simply a usual finite Markov chain with state space being the set of types T and hence $\mathbf{q} = [0, \ldots, 0]$.

4 Summary of Papers

4.1 Paper A

Paper A addresses essentially Question IV in a population where it is possible to distinguish between two kinds of individuals, 0 and 1, type 0 being of the escape type. Type 1 individuals have reproduction mean $0 < m < 1$ and each one of their daughters either mutates, with probability $u \in (0,1)$, to a type 0 individual or stays in the same type with probability $1-u$. Type 0 individuals have reproduction mean $1 < m_0 < \infty$ and produce only individuals of the same type. The population starts with a single individual of type 1. Therefore, the corresponding mathematical model to this population is a two-type decomposable GWBP, $\{(Z_n(0), Z_n(1)), n \in \mathbb{N}_0\}$, with $Z_0(1) = 1$ and $Z_0(0) = 0$. The results in this paper are exact; there is no asymptotic analysis as $u \to 0$.

The first results are concerned with the random variable T representing the waiting time to produce the first successful mutant. By successful mutant we mean an individual of type 0, whose mother is of type 1, that produces a lineage that does not get extinct. Using probability generating functions, we derive exact expressions for the tail probabilities of T and for the expectation, conditioned on T being finite.

The paper finishes with a result comparing the limit behavior of this process with the limit of a single-type supercritical GWBP. The result is an analogue of Theorem 3.3 for sequence $Z_n(0), n \geq 0$, representing the number of type 0 individuals in generation n. We prove that, under certain conditions, as $n \to \infty$,

$$\frac{Z_n(0)}{m_0^{n-\tau}} \to U,$$

almost surely and in L^1, with $E[U] = 1$ and

$$\tau = \left| \log_{m_0} \left(\frac{um}{m_0 - m(1-u)} \right) \right| .$$

One of the conclusions to draw is that sequence $Z_n(0)$ exhibits the same limit behavior as a single-type GWBP, consisting only of individuals of type 0, with a delay given by τ.

4.2 Paper B

Paper B is, in a certain way, an extension of the previous paper and is essentially concerned with Questions II and IV. Most of the results are asymptotic and address the basic two-type model considered in Paper A and, unless stated otherwise, in this subsection we will always be referring to this model.

We start with a generalization of the system of equations (1) originally obtained for two particular reproduction laws. This generalization holds for populations with any number of types, a general mutation scheme, only one escape type and any reproduction law with finite variance. This is an important result since in most of the applications the offspring distribution is not known.

From the results of Paper A it is not always easy, if possible, to obtain simple expressions for the distribution function of T. Hence, approximations are necessary for application purposes. A part of Paper B is dedicated to deriving such approximations, again under the assumption that the mutation rate u is small. It turns out that, conditioned on escape, the waiting time T has approximately a geometric distribution with success parameter $1 - m$, where m is as in Paper A.

In Paper B we also propose an alternative version of the hazard function of the waiting time to produce the first successful mutant. We want to quantify the immediate risk of escape, i.e., the probability of producing a successful mutant in the next generation given that it has not been produced yet. The usual definition of hazard function is not really suitable in this situation since a risk of producing a mutant exists only if there is at least one type 1 individual is alive. Hence, to simply condition on the non-appearance of a successful mutant, does not give an accurate idea of the risk. It seems that, for very old populations, this function becomes a constant reflecting the limit behavior of a subcritical GWBP conditioned on non-extinction; see Theorem 3.2.

In the last part of the paper we present some results concerning the time it takes for the number of individuals of the escape type to reach a high level, say x. For application purposes it is extremely important to obtain estimates of this time.

We start by showing that when the mutation rate is small, the waiting time to cross level x can be decomposed as a sum of two independent random variables, $T + L_x$. The first one has already been studied. The second, L_x, represents the time it takes for a single-type GWBP, stemming from one type 0 individual, to cross level x. This decomposition may not hold due to the contribution of several mutants, but we take care of this by that proving the probability of such event can be neglected when u is small.

It remains now to study the distribution of L_x. We reviewed some results on this problem due to Nagaev (1971) and Rösler et al. (2001). It turns out that, conditioned on $L_x < \infty$, the asymptotic distribution of L_x, as $x \to \infty$, depends heavily on the distribution of $(W|W > 0)$ where W is the limiting random variable of Theorem 3.3. Since not much is known about this random variable, we performed simulations for processes with the most used offspring distributions: Poisson, geometric and also binary splitting.

4.3 Paper C

The goal of Paper C is to give a detailed description of the path towards the escape for the network model already referred in Sections 1 and 2. We additionally assume that individuals described by sequences with the same number of ones have equal reproduction mean and will be considered of the same type. Hence, a sequence with i ones is assigned with reproduction mean $0 < m_i < 1$, $i = 1, \dots L$. Sequence with all zeroes is assigned with a reproduction mean $1 < m_0 < \infty$. The results are asymptotic, obtained by taking the limit as $\mu \to 0$, and address Questions II, III and IV in a more general setting than in Paper B.

The paper starts by considering the case $L = 1$ where a two-type GWBP is used. Here mutation is modeled in a more general way than before in that we allow the probability of mutation for a newborn type 1 to depend on its family size. We show that, conditioned on escape, the limit process is a GWBP with immigration. The immigration source corresponds to a stem (escape) lineage leading to the successful mutation and is turned on during a geometric number of generations with parameter $1-m_1$, as the results in Paper B indicated. The number of immigrants has the size-biased distribution of the reproduction law of type 1 individuals, except at the generation where the stem lineage stops. At this time the number of immigrants may have a different distribution due to the assumption that mutation probability can be family size dependent.

The next step is to extend the two-type model to a sequential mutation model where we distinguish between $L+1$ types of individuals, labeled $0,\ldots,L$. This is a convenient intermediate step towards the network model. In fact, since individuals described by sequences with the same number of ones are considered to be of the same type, the network model can be seen as a sequential mutation model where the mutation probabilities between types i and j are asymptotically equivalent to $\binom{i}{j}\mu^{i-j}$, $j<i$. We show that, conditioned on the appearance of a type 0 individual, the limit process is now a GWBP with a multitype immigration stopped after a sequence of geometric times.

From the description of the latter process it is possible to obtain a detailed answer to Question IV. In fact, the asymptotic waiting time to escape is decomposed as a sum of a random number of independent geometric random variables. It is formulated in terms of the waiting time until absorption at state 0 of a Markov chain $\{Y_n, n\geq 0\}$ describing the types visited by the stem lineage on the path to escape.

Finally, a description of the limit process for the network model is given using the result from the sequential mutation model. In terms of sequences, the process starts with sequence $\bar{1}$ initiating a stem lineage. In the next generation the stem sequence either remains in type $\bar{1}$, with probability m_L, or turns to an 0-1 sequence $\bar{u}_1=(u_{11},\ldots,u_{1L})$ with probability

$$P(\bar{1}\to\bar{u}_1)=\frac{D_{L,i_1}}{\binom{L}{i_1}},\quad i_1=u_{11}+\ldots+u_{1L},$$

where D_{L,i_1} are the transition probabilities of Markov chain $\{Y_n, n\geq 0\}$. During a geometric number of generations, $T_L\sim Geom(1-m_L)$, including the time it switches to $\bar{u}_1$, sequence $\bar{1}$ produces, according to the size-biased distribution of the reproduction law of type L individuals, a number of side lineages of mutation-free $\bar{1}$-viruses. 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.

Hence, the sequence dynamics from $\bar{1}$ towards $\bar{0}$ is described by a random path $\bar{1} \to \bar{u}_1 \to \ldots \to \bar{u}_k \equiv \bar{0}$ of a random length $1 \leq k \leq L+1$ through intermediate sequences with a 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

$$P(\bar{u}_j \to \bar{u}_l) = \frac{D_{i_j, i_l}}{\binom{i_j}{i_l}}.$$

The stem lineage spends at the type $\bar{u}_j$ a geometric number of generations with parameter $\lambda(i_j)$, where $\lambda(i) = 1 - m_i$. During this time it generates mutation-free $\bar{u}_j$-lineages. The number of such lineages per generation has the size-biased distribution of the reproduction law of type i_j individuals. Each mutation-free $\bar{u}_j$-lineage is a single-type GWBP process with the reproduction law of type i_j.

References

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

[2] Axelrod, D. E and Kimmel, M. (2002). Branching Processes in Biology. Springer, Berlin.

[3] Iwasa, Y., Michor, F., Nowak, M. A. (2003). Evolutionary dynamics of escape from biomedical intervention. Proceedings of the Royal Society London B 270, 2573-2578.

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

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

[6] Harris, T. E. (1963). The theory of Branching Processes. Springer, Berlin.

[7] Jagers, P. (1975). Branching Processes with Biological Applications. John Wiley, London.

[8] Nagev, A.V. (1971) A limit theorem for supercritical branching processes. Matematiceskie Zametki 9 (5), 585-592 (In Russian)

[9] Rösler, U., Topchii, V.A., Vatutin, V.A. (2001). High-level overshoot for a class of random sequences. Siberian Advances in Mathematics 11, (3), 1-13.

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

Paper A

J. Appl. Prob. **43,** 296–302 (2006)
Printed in Israel

ON THE WAITING TIME TO ESCAPE

MARIA CONCEIÇÃO SERRA,* *Chalmers University of Technology*

Abstract

The mathematical model we consider here is a decomposable Galton–Watson process with individuals of two types, 0 and 1. Individuals of type 0 are supercritical and can only produce individuals of type 0, whereas individuals of type 1 are subcritical and can produce individuals of both types. The aim of this paper is to study the properties of the waiting time to escape, i.e. the time it takes to produce a type-0 individual that escapes extinction when the process starts with a type-1 individual. With a view towards applications, we provide examples of populations in biological and medical contexts that can be suitably modeled by such processes.

Keywords: Decomposable Galton–Watson branching process; probability generating function

2000 Mathematics Subject Classification: Primary 60J85
Secondary 60J80

1. Introduction

In many biological and medical contexts we find populations that, due to the small reproductive ratio of the individuals, will become extinct after some time. However, sometimes changes occur during the reproduction process that lead to an increase of the reproductive ratio, making it possible for the population to escape extinction. In this work we use the theory of branching processes to model the evolution of this kind of population.

Cancer cells subjected to chemotherapy are an example of such a population. When the cells are subjected to chemotherapy, their capacity for division is reduced, hopefully leading to the extinction of tumour cells. Yet mutations may lead to another kind of cell that is resistant to the chemotherapy. Thus, the population of this new type of cell has a larger reproductive ratio and might escape extinction.

Another example can be found in epidemics like HIV or SARS. Imagine a virus of one host species that is transferred to another host species where it has a small reproductive mean and, therefore, the extinction of its lineage is certain. Mutations occurring during the reproduction process could still lead to a virus capable of initiating an epidemic in the new host species.

The goal of this article is to use a two-type Galton–Watson branching processes (GWBP) to study properties of populations of this sort. We assume that the process starts with a single subcritical individual that gives birth to individuals of the same type, but whose descendents, through mutation, can become supercritical and are therefore capable of establishing a population that has a positive probability of escaping extinction.

In Section 2 we introduce the model, the main reproduction parameters of the process, and give some references to theoretical and applied works. Section 3 contains the main results and proofs. Using probability generating functions, we derive properties of the distribution of the

---

Received 11 October 2005.

* Postal address: Department of Mathematical Statistics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden. Email address: mcserra@math.chalmers.se

waiting time to produce an individual that escapes extinction. We prove that it has a point mass at ∞ and compute the tail probabilities and its expectation (conditioned on being finite). We also show that, in the long run, the population size of this process grows like that of a single-type GWBP with a delay.

2. Description of the model

Consider a two-type GWBP $\{(Z_n^{(0)}, Z_n^{(1)}),\ n \in \mathbb{N}_0\}$, where $Z_n^{(0)}$ and $Z_n^{(1)}$ respectively denote the number of individuals of type 0 and of type 1 in the nth generation, and $\mathbb{N}_0$ is the set of nonnegative integers. Suppose that individuals of type 1 are subcritical, i.e. have reproduction mean $m,\ 0 < m < 1$, and that each one of their descendents can mutate, independently of the others, to type 0 with probability $u,\ 0 < u < 1$. Individuals of type 0 are supercritical, i.e. have reproduction mean $m_0,\ 1 < m_0 < \infty$, and there is no backward mutation. For this particular two-type GWBP, the first moment matrix is of the form

$$A = \begin{bmatrix} m_0 & 0 \\ mu & m(1-u) \end{bmatrix}.$$

Unless stated otherwise, we assume that the process starts with just one individual, of type 1, i.e. $Z_0^{(0)} = 0$ and $Z_0^{(1)} = 1$. The probability generating function of the reproduction law of type-i individuals will be denoted by $f_i,\ i \in \{0, 1\}$, and the joint probability generating function of $(Z_1^{(0)}, Z_1^{(1)})$ is given by

$$\begin{aligned} F(s_0, s_1) &= \mathrm{E}\Big[s_0^{Z_1^{(0)}} s_1^{Z_1^{(1)}}\Big] = \sum_{k=0}^{\infty} p_k^{(1)} \sum_{j=0}^{k} \binom{k}{j} s_0^j u^j s_1^{k-j} (1-u)^{k-j} \\ &= f_1(s_0 u + (1-u)s_1), \qquad (s_0, s_1) \in [0, 1]^2, \end{aligned} \tag{2.1}$$

where $\{p_k^{(1)},\ k \in \mathbb{N}_0\}$ represents the reproduction law of type-1 individuals.

Branching processes have been intensively studied during the last decades; classical references are the books of Harris (1963), Athreya and Ney (1972), Jagers (1975), and Mode (1971). For recent books, with emphasis on applications, see Axelrod and Kimmel (2002) and also Haccou *et al.* (2005). For a nice example of how branching processes can be used to solve important problems in biology and medicine, the reader is referred to the papers of Iwasa *et al.* (2003), (2004).

3. Main results

3.1. Number of mutants and the probability of extinction

Consider the sequence of random variables $\{I_n,\ n \in \mathbb{N}_0\}$, with I_n being the total number of mutants produced until generation n (inclusive), and let I be the random variable that represents the number of mutants in the whole process. By *mutant* we mean an individual of type 0 whose mother is of type 1.

It is obvious that the sequence I_n converges pointwise to the random variable I. In our first theorem, we use this convergence to establish a functional equation for the probability generating function of I, denoted by f_I.

Theorem 3.1. *The probability generating function of I satisfies the functional equation*

$$f_I(s) = f_1(us + (1-u) f_I(s)), \tag{3.1}$$

for all $s \in [0, 1]$.

Proof. First we establish a recursive relation for the probability generating functions of the random variables I_n, denoted by f_{I_n}. We find that, for all $n \geq 1$,

$$\begin{aligned} f_{I_n}(s) &= \mathrm{E}[s^{I_n}] \\ &= \mathrm{E}[\mathrm{E}[s^{I_n} \mid Z_1^{(0)}, Z_1^{(1)}]] \\ &= \mathrm{E}\Big[\mathrm{E}\Big[s^{Z_1^{(0)}+\sum_{i=1}^{Z_1^{(1)}} I_{n-1}^i} \;\Big|\; Z_1^{(0)}, Z_1^{(1)}\Big]\Big] \\ &= \mathrm{E}\Big[s^{Z_1^{(0)}} \mathrm{E}[s^{I_{n-1}}]^{Z_1^{(1)}}\Big] \\ &= F(s, f_{I_{n-1}}(s)) \\ &= f_1(su + (1-u) f_{I_{n-1}}(s)), \end{aligned} \tag{3.2}$$

where the I_{n-1}^i are independent, identically distributed copies of the random variable I_{n-1}, the function F is as defined in (2.1), and $f_{I_0}(s) = 1$.

By taking the limit in relation (3.2) we obtain the functional equation (3.1).

We now proceed to determine the probability of extinction. Using the notation

$$\begin{aligned} q_0 &= \mathrm{P}[Z_n^{(0)} = Z_n^{(1)} = 0 \text{ for some } n \geq 1 \mid Z_0^{(0)} = 1,\ Z_0^{(1)} = 0], \\ q_1 &= \mathrm{P}[Z_n^{(0)} = Z_n^{(1)} = 0 \text{ for some } n \geq 1 \mid Z_0^{(0)} = 0,\ Z_0^{(1)} = 1], \end{aligned}$$

it follows, from the classical result on the extinction of branching processes, that q_0 is the smallest root of

$$q_0 = f_0(q_0)$$

in the interval $[0, 1]$. To determine q_1, notice that extinction of the process occurs if and only if all the supercritical single-type GWBPs starting from the mutants die out. Since there are I such processes, we have

$$q_1 = \mathrm{E}[q_0^I] = f_I(q_0).$$

Obtaining an explicit expression for q_1 is not always possible; therefore, approximations are necessary for application purposes. Assuming there to be small mutation rate u, Iwasa *et al.* (2003), (2004) provided these approximations for particular reproduction laws, namely for Poisson and geometric distributions. Their results extend to an even more complex scheme of mutations leading to branching processes with more than two types of individual.

3.2. Waiting time to produce a successful mutant

Consider the random variable T, which represents the time to escape, i.e. the first generation in which a successful mutant is produced. By *successful mutant* we mean a mutant that is able to start a single-type GWBP that escapes extinction. This variable takes values in the set $\{1, 2, \dots, \infty\}$, with $T = \infty$ if no successful mutant is produced.

Theorem 3.2. *The distribution of T has the following properties:*

(i) $\mathrm{P}[T > k] = f_{I_k}(q_0)$ *for all* $k \geq 0$,

(ii) $\mathrm{P}[T = \infty] = q_1$,

(iii) $\mathrm{E}[T \mid T < \infty] = \sum_{k=0}^{\infty}(f_{I_k}(q_0) - q_1)/(1 - q_1)$.

Proof. To prove (i), observe that $T > k$ means that all I_k mutants were unsuccessful. Therefore,

$$\mathrm{P}[T > k] = \mathrm{E}[q_0^{I_k}] = f_{I_k}(q_0).$$

To prove (ii), observe that $(T > k)_{k \geq 0}$ is a nonincreasing sequence of events and that

$$\mathrm{P}[T = \infty] = \mathrm{P}\left[\bigcap_{k=0}^{\infty}(T > k)\right] = \lim_{k\to\infty} \mathrm{P}[T > k] = \lim_{k\to\infty} f_{I_k}(q_0) = f_I(q_0) = q_1.$$

To prove (iii), observe that $T > 0$ and, therefore,

$$\begin{aligned}\mathrm{E}[T \mid T < \infty] &= \sum_{k=0}^{\infty} \frac{\mathrm{P}[T > k, T < \infty]}{\mathrm{P}[T < \infty]} \\ &= \sum_{k=0}^{\infty} \frac{\mathrm{P}[T < \infty] - \mathrm{P}[T \leq k]}{1 - q_1} \\ &= \sum_{k=0}^{\infty} \frac{f_{I_k}(q_0) - f_I(q_0)}{1 - q_1},\end{aligned}$$

with the f_{I_k} as recursively defined in (3.2).

A similar problem was considered in Bruss and Slavtchova-Bojkova (1999), where a single-type GWBP with immigration to the state 0 was used to model the repopulation of an environment. The idea is the following. Consider a population starting with a supercritical individual and let it grow according to a GWBP. If extinction occurs at time t then immigration takes place immediately after, i.e. one individual of the same kind is introduced and a new process, independent of and identically distributed to the first one, restarts. Among other results, Bruss and Slavtchova-Bojkova derived properties of the last instant of immigration, i.e. of the generation into which was introduced an immigrant that started a process that escaped extinction.

In the applications we consider, the mutants appear at random times as descendents of the subcritical individuals, and the model described above therefore does not apply.

3.3. Comparison with a single-type supercritical GWBP

In this section we prove a result that will allow us to compare the limit behavior of the sequence $Z_n^{(0)}$ with the limit behavior of a single-type supercritical GWBP. First, we recall a result on single-type GWBPs. The proof can be found in any of the classical books referred to in Section 2.

Theorem 3.3. *Let* $\{Y_n,\ n \in \mathbb{N}_0\}$ *be a single-type supercritical GWBP with reproduction law* $\{p_k^{(0)},\ k \in \mathbb{N}_0\}$, *and suppose that* $Y_0 = 1$. *If*

$$\sum_{k=0}^{\infty} k \log k p_k^{(0)} < \infty \tag{3.3}$$

then $Y_n/\mu^n \to W$ *almost surely and in* L^1, *where* $\mu = \sum_{k=0}^{\infty} k p_k^{(0)}$ *and* $\mathrm{E}[W] = 1$. *Furthermore, the Laplace transform of* W, ϕ_W, *satisfies*

$$\phi_W(\mu s) = f_0(\phi_W(s)), \qquad s \geq 0.$$

Our result is as follows.

Theorem 3.4. *If the reproduction law of type-0 individuals satisfies condition (3.3), then*

$$\frac{Z_n^{(0)}}{m_0^n} \to U \quad \textit{almost surely and in } L^1,$$

with $\mathrm{E}[U] = um/(m_0 - m(1-u)) < 1$. *Furthermore, the Laplace transform of* U, ϕ_U, *satisfies the functional equation*

$$\phi_U(m_0 s) = f_1(u\phi_W(s) + (1-u)\phi_U(s)),$$

where ϕ_W *is as in Theorem 3.3.*

Proof. Consider the sequence of random variables $\{J_n,\ n \geq 1\}$, where J_n represents the number of mutants in generation n, i.e. $J_n = I_n - I_{n-1}$. Using these variables, $Z_n^{(0)}$, $n \geq 1$, can be decomposed in the following way:

$$Z_1^{(0)} = J_1, \qquad Z_n^{(0)} = \sum_{k=1}^{n-1} \sum_{i=1}^{J_k} Y_{n-k}^i, \quad n \geq 2. \tag{3.4}$$

Here, the random variable Y_{n-k}^i represents the number of individuals in generation $n-k$ of the single-type supercritical GWBP initiated by the ith mutant of generation k. These processes are independent of each other and have the same reproduction law, namely $\{p_k^{(0)},\ k \in \mathbb{N}_0\}$.

By dividing (3.4) by m_0^n and taking expectations, we obtain

$$\begin{aligned}
\mathrm{E}\left[\frac{Z_n^{(0)}}{m_0^n}\right] &= \sum_{k=1}^{n-1} \frac{1}{m_0^k} \mathrm{E}\left[\sum_{i=1}^{J_k} \frac{Y_{n-k}^i}{m_0^{n-k}}\right] \\
&= \sum_{k=1}^{n-1} \frac{1}{m_0^k} \mathrm{E}[J_k] \\
&= \sum_{k=1}^{n-1} \frac{1}{m_0^k} um[m(1-u)]^{k-1} \\
&\to \frac{um}{m_0 - m(1-u)} < 1 \quad \text{as } n \to \infty.
\end{aligned} \tag{3.5}$$

The expectation of J_k is obtained by differentiation of the recursive relation (3.2). Since $\{m_0^{-n} Z_n^{(0)},\ n \geq 0\}$ is a submartingale with respect to the σ-algebra $F_n = \sigma\{Z_m^{(0)}, Z_m^{(1)}, 0 \leq m \leq n\}$ and, from (3.5), we have

$$\sup \mathrm{E}\left[\frac{Z_n^{(0)}}{m_0^n}\right] < \infty,$$

the martingale convergence theorem ensures that the sequence converges almost surely to a random variable U with $\mathrm{E}[U] < \infty$.

To prove L^1-convergence, it remains to show that

$$\mathrm{E}[U] = \frac{um}{m_0 - m(1-u)}. \tag{3.6}$$

Observe that, given $(Z_1^{(0)}, Z_1^{(1)})$, the following decomposition holds:

$$\frac{Z_n^{(0)}}{m_0^n} = \frac{1}{m_0} \sum_{i=1}^{Z_1^{(0)}} \frac{Y_{n-1}^i}{m_0^{n-1}} + \frac{1}{m_0} \sum_{j=1}^{Z_1^{(1)}} \frac{X_{n-1,j}^{(0)}}{m_0^{n-1}}. \tag{3.7}$$

In this expression the Y_{n-1}^i are as described in decomposition (3.4) and the $X_{n-1,j}^{(0)}$ are the random variables that represent the number of type-0 individuals in generation $n-1$ of the jth two-type GWBP initiated in generation 1. There are $Z_1^{(1)}$ such processes and they are independent of each other. Taking the limit in (3.7) (the existence of the limits of the sequences involved was already proved) gives

$$U = \frac{1}{m_0} \sum_{i=1}^{Z_1^{(0)}} W_i + \frac{1}{m_0} \sum_{j=1}^{Z_1^{(1)}} U_j, \tag{3.8}$$

where W_i are independent, identically distributed copies of W, as defined in Theorem 3.3, and U_j are independent, identically distributed copies of U. It is now a matter of taking expectations in (3.8) to obtain the desired result, (3.6).

Finally, proving the functional equation for the Laplace transform of U is just a matter of using (3.8). Indeed,

$$\begin{aligned}
\phi_U(s) &= \mathrm{E}[\mathrm{e}^{-sU}] \\
&= \mathrm{E}[\mathrm{E}[\mathrm{e}^{-sU} \mid Z_1^{(0)}, Z_1^{(1)}]] \\
&= \mathrm{E}\Bigg[\mathrm{E}\Bigg[\exp\Bigg(-\frac{s}{m_0}\sum_{i=1}^{Z_1^{(0)}} W_i\Bigg) \Bigg| Z_1^{(0)}, Z_1^{(1)}\Bigg] \mathrm{E}\Bigg[\exp\Bigg(-\frac{s}{m_0}\sum_{j=1}^{Z_1^{(1)}} U_j\Bigg) \Bigg| Z_1^{(0)}, Z_1^{(1)}\Bigg]\Bigg] \\
&= \mathrm{E}\Bigg[\Bigg(\phi_W\Bigg(\frac{s}{m_0}\Bigg)\Bigg)^{Z_1^{(0)}} \Bigg(\phi_U\Bigg(\frac{s}{m_0}\Bigg)\Bigg)^{Z_1^{(1)}}\Bigg] \\
&= f_1\Bigg(u\phi_W\Bigg(\frac{s}{m_0}\Bigg) + (1-u)\phi_U\Bigg(\frac{s}{m_0}\Bigg)\Bigg).
\end{aligned}$$

With

$$\tau = \left|\log_{m_0}\left(\frac{um}{m_0 - m(1-u)}\right)\right|,$$

we conclude that there exists a random variable U^* such that

$$\frac{Z_n^{(0)}}{m_0^{n-\tau}} \to U^* \quad \text{almost surely and in } L^1,$$

with $\mathrm{E}[U^*] = 1$. This indicates that the sequence $Z_n^{(0)}$ exhibits the same limit behavior as a single-type supercritical GWBP, except with a delay τ. It remains to investigate the relation between the constant τ and the random variable that represents the delay between the two processes.

In applications, it is not only important to study the time taken to produce a successful mutant, but also the time taken for the number of type-0 individuals to reach high levels. Theorem 3.4 provides a first step in determining this.

Acknowledgements

The author would like to thank Patsy Haccou, Peter Jagers, and Serik Sagitov for many helpful discussions and careful reading of this manuscript. She is also grateful to the Center of Mathematics of the University of Minho, Portugal, and to the Portuguese Foundation for Science and Technology for supporting her stay at Chalmers University.

References

ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes.* Springer, Berlin.

AXELROD, D. E. AND KIMMEL, M. (2002). *Branching Processes in Biology.* Springer, New York.

BRUSS, F. T. AND SLAVTCHOVA-BOJKOVA, M. (1999). On waiting times to populate an environment and a question of statistical inference. *J. Appl. Prob.* **36,** 261–267.

HACCOU, P., JAGERS, P. AND VATUTIN, V. A. (2005). *Branching Processes: Variation, Growth and Extinction of Populations.* Cambridge University Press.

HARRIS, T. E. (1963). *The Theory of Branching Processes.* Springer, Berlin.

IWASA, Y., MICHOR, F. AND NOWAK, M. (2003). Evolutionary dynamics of escape from biomedical intervention. *Proc. R. Soc. London B* **270,** 2573–2578.

IWASA, Y., MICHOR, F. AND NOWAK, M. (2004). Evolutionary dynamics of invasion and escape. *J. Theoret. Biol.* **226,** 205–214.

JAGERS, P. (1975). *Branching Processes with Biological Applications.* John Wiley, London.

MODE, C. (1971). *Multitype Branching Processes: Theory and Applications.* Elsevier, New York.

Paper B

Dynamics of escape mutants

Maria Conceição Serra [a,b,*] , Patsy Haccou [c]

[a] *Department of Mathematics, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal*

[b] *School of Mathematical Sciences, Chalmers University of Technology and Göteborg University, SE-41296 Sweden*

[c] *Section Theoretical Biology, Institute of Biology, Leiden University, Kaiserstraat 63, 2311 GP Leiden, The Netherlands*

Abstract

We use multi-type Galton-Watson branching processes to model the evolution of populations that, due to a small reproductive ratio of the individuals, are doomed to extinction. Yet, mutations occurring during the reproduction process, may lead to the appearance of new types of individuals that are able to escape extinction. We provide examples of such populations in medical, biological and environmental contexts and give results on *i)* the probability of escape/extinction, *ii)* the distribution of the waiting time to produce the first individual whose lineage does not get extinct and *iii)* the distribution of the time it takes for the number of mutants to reach a high level. Special attention is dedicated to the case where the probability of mutation is very small and approximations for *i)-iii)* are derived.

Key words: Multi-type Galton-Watson branching process, probability generating function, mutation, escape and extinction probability

---

1 Introduction

In many medical, biological and environmental contexts we find populations that, due to a small reproductive ratio of the individuals, will go extinct after some time. Yet, sometimes changes can occur during the reproduction process

---

* Corresponding author at: Department of Mathematics, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. Fax:+351253678982

Email address: `mcserra@math.uminho.pt` (Maria Conceição Serra).

that lead to an increase of the reproductive ratio, making it possible for the population to escape extinction.

Cancer cells submitted to chemotherapy are an example of such populations(c.f. Michor et al. (2004) and Nowak et al. (2004)). During chemotherapy the capacity of division of the cancer cells is reduced, which should lead to the destruction of tumors. Yet, sometimes mutations in the cells provide resistance to the therapy. This new type of cells has a higher reproduction and can escape extinction.

Another example can be found in viruses. A virus adapted to one host species that switches to another host usually has a small reproductive mean and, therefore, the extinction of its lineage is certain. Mutations can lead to a virus capable of initiating an epidemic in the new host species. This happened for instance in HIV and SARS viruses.

More generally, in many evolutionary processes mutants may appear which are initially less viable than the resident type and thus are doomed for extinction. Additional mutations, however, may lead to a more successful type that can outcompete the resident. In this way, an evolutionary process can cross a fitness valley.

In agriculture we find situations where introgression may occur between genetically modified or cultivated organisms with wild populations, Ellstrand et al. (1999) and Maan (1987). Usually the first backcrosses are not very viable and fertile, but eventually a successful type may be produced.

Finally, an important environmental problem is the spread of insecticide resistance due to hybridization between resistant and susceptible insects. Some of the biotypes of the sweet potato whitefly, for instance, have developed resistance to certain insecticides, Byrne et al. (1994), De Barro et al. (2000) and Guirao et al. (1997). Although matings between different biotypes are rare, hybrid formation can occur. The initial hybrids have low fitness. Yet, after several backcrosses between resistant and sensitive biotypes, a successful resistant hybrid may be produced.

In all the examples given above it is of vital importance to have good estimates of the probability of successful escape and related aspects, such as the distribution of numbers of escape mutations, the distribution of waiting times until escape, and the time it takes for the number of escape mutants to reach a high level. In this paper we will derive such estimates.

We will use a discrete time branching process, known as the Galton-Watson branching process (GWBP) as a model, and therefore we assume that each individual of the population belongs to a unique generation, lives one unit of time and gives birth to a random number of individuals. The descendants will form the next generation and so on. Since we consider different kinds of individuals, with different reproductive ratio, we will use a multi-type GWBP to model their dynamics.

The assumption of discrete non-overlapping generations is obviously a simplification. Yet, mathematically, discrete models are much easier to handle than their continuous time analogues, and in some problems, like extinction, we can deduce results that also hold for the analogous continuous time process. Furthermore, our results will serve as a first step, and may be generalized at a later stage.

The structure of the paper is as follows. In Section 2 we describe the basic model for two types and derive the probability generating function of the numbers of mutations to the escape type. We demonstrate how this function can be used to determine moments of the distribution of the numbers of mutations. In Section 3 we derive expressions for the probability of extinction (and escape) of this process. Furthermore, we present approximations of these probabilities for multi-type processes with a general reproduction and mutation scheme, assuming that the mutation rates are small. We show that the results derived by Iwasa et al. (2003, 2004), for the Poisson and geometric offspring distributions can be applied more generally, and we provide a bound for the error term in the approximation. In Section 4 we study the waiting time to produce a successful mutant in a population with two types of individuals and give approximations for its distribution. Finally, in Section 5 we consider the time until the mutant population has grown to a high level.

2 The basic model

Consider a population with two types of individuals, type 0 and type 1, such that each descendant of an individual of type 1 can mutate, with probability $u \in (0,1)$, to type 0. Mutations from type 0 to type 1 are assumed not to occur. Furthermore, we assume that individuals reproduce independently of each other and that the occurrence of mutations is also independent. We assume that individuals of type 1 have reproduction mean $0 < m < 1$ and we call them subcritical individuals. Individuals of type 0 will be called escape type and we assume they have reproduction mean $1 < m_0 < \infty$.

Notice that a single-type GWBP, starting with one individual of type 0, has positive probability of escaping extinction and this justifies the choice of calling them "escape" individuals. The choice to call type 1 “subcritical” individuals follows established mathematical terminology.

We will use a two-type GWBP, $\{(Z_n^{(0)}, Z_n^{(1)}), n \in \mathbb{N}_0\}$, to model the number of individuals in this kind of population. As usual, $Z_n^{(0)}$ and $Z_n^{(1)}$ denote the number of individuals of type 0 and of type 1, respectively, in the n^{th} generation. Unless stated otherwise, we assume that the population starts with a single individual of type 1, i.e., $Z_0^{(0)} = 0$ and $Z_0^{(1)} = 1$. The joint probability generating function (p.g.f.) of $(Z_1^{(0)}, Z_1^{(1)})$ is given by

$$F(s_0, s_1) = f(s_0 u + (1-u)s_1), \quad (s_0, s_1) \in [0,1]^2, \tag{1}$$

where f denotes the p.g.f. of the reproduction law of type 1 individuals.

Unless mutations occur, such a process will be a single-type subcritical GWBP and it is the appearance of mutants that makes the study of such populations an interesting task. Therefore it is important to study, for instance, the total number of mutations that occur in the whole process. This random quantity will play a crucial role in determining the extinction probability of the process. Let I be the random variable (r.v.) that counts the total number of mutations to the escape type in the whole process. Using (1), we can deduce that the p.g.f. of I, which will be denoted by h, satisfies the following functional equation:

$$h(s) = F(s, h(s)) = f(su + (1-u)h(s)), \ \ \forall s \in [0,1]. \tag{2}$$

From this functional equation we can derive the moments of I. We start by computing the mean value. Differentiating (2), we get

$$h'(s) = f'(su + (1-u)h(s))(u + (1-u)h'(s)). \tag{3}$$

and replacing s by 1 in (3), and solving for $h'(1) = E[I]$, yields

$$E[I] = \frac{mu}{1 - m(1-u)} \tag{4}$$

Differentiating (3), we can obtain the variance in a similar way, which leads to

$$Var[I] = h''(1) + E[I](1 - E[I]) = \frac{um(1-u)(1-m)^2 + u^2\sigma^2}{[1 - m(1-u)]^3} \tag{5}$$

where σ^2 denotes the variance of ξ.

From (4) it can be seen that, on average, less than $m$mutants are produced before the original type is extinct. It may come as a surprise that this upper limit holds for all u and $m < 1$, considering the fact that, if no mutations were allowed, the expectation of the total progeny of the initial individual would be $m/(1-m)$ which is always larger than m, and larger than 1 when $m > 1/2$. On the other hand, we expect to have the largest number of mutants when $u = 1$ and this corresponds to the case where the total number of mutants equals the number of individuals in the first generation and so $E[I] = m$. Further, as can be expected, $E[I]$ increases with the expected number of offspring of the subcritical individuals, m, and with the mutation chance, u.

3 The probability of escape

In the model introduced in the previous section, let q_0 and q_1 denote the probability of extinction of the process, when it starts with one individual of type 0 or of type 1, i.e.,

$$q_0 = P[Z_n^{(0)} = Z_n^{(1)} = 0, \text{ for some } n \in \mathbb{N} \mid Z_0^{(0)} = 1, Z_0^{(1)} = 0]$$

and

$$q_1 = P[Z_n^{(0)} = Z_n^{(1)} = 0, \text{ for some } n \in \mathbb{N} \mid Z_0^{(0)} = 0, Z_0^{(1)} = 1].$$

respectively. Since individuals of the escape type cannot mutate to the subcritical type, q_0 is just the probability of extinction of a single-type supercritical GWBP. Therefore q_0 is the unique solution in the interval $[0, 1)$ of equation

$$s = f_0(s)$$

where f_0 denotes the p.g.f. of the offspring distribution of the escape individuals. If we assume $f_0(0) > 0$, we have $0 < q_0 < 1$.

To determine q_1 we need only to remember that the process is extinct if and only if the lineage of individuals that mutated from type 1 to type 0 is also extinct. Since the number of such individuals is given by the r.v. I, we have

$$q_1 = E[q_0^I] = h(q_0), \tag{6}$$

where h is given by (2).

To determine the probability of escape we just have to use the extinction/explosion dichotomy. If we define

$$r_0 = P[Z_n^{(0)} \to \infty \mid Z_0^{(0)} = 1, Z_0^{(1)} = 0]$$

and

$$r_1 = P[Z_n^{(0)} \to \infty \mid Z_0^{(0)} = 0, Z_0^{(1)} = 1]$$

i.e., r_0 and r_1 denote the probability of escape of the two-type GWBP, when it starts with one individual of type 0 or of type 1, respectively, we have

$$r_0 = 1 - q_0 \quad \text{and} \quad r_1 = 1 - q_1 = 1 - h(q_0).$$

Observe that in the definition of r_0 and r_1 we mention only the infinite number of individuals of type 0 because, due to their "subcriticality", the number of individuals of type 1 will tend to zero.

Obtaining an explicit expression for the escape probability is not always possible, even in such a simple model as studied above; therefore, approximations are necessary for application purposes. Since mutations are rare events, Iwasa et al. (2003, 2004) assumed there to be small mutation rates and developed approximations for populations with particular reproduction laws, namely Poisson and geometric, and with any number of types. We show that the same approximations hold for populations with any reproduction law, as long as the variance is finite.

From now on we consider a population with many different types of individuals with labels in the set $\{0, 1, \ldots, K\}$, $K \geq 2$. Individuals of type j have reproduction mean $0 < m_j < 1$, if $j \in \{1, \ldots K\}$, but $1 < m_0 < \infty$. Therefore we will say that individuals of type 0 are of the escape type and the others are of the subcritical type. Suppose that mutations from type j to type i occur with probability $0 \leq u_{ji} < 1$, $j \neq i$, and let

$$u_{jj} = 1 - \sum_{\substack{i=0 \\ i \neq j}}^{K} u_{ji}. \tag{7}$$

Mutations from type 0 to the other types cannot occur and so $u_{0j} = 0$, for $j \in \{1, \ldots, K\}$. Because figures are better than words, below we draw an example of such a population with 4 different types of individuals.

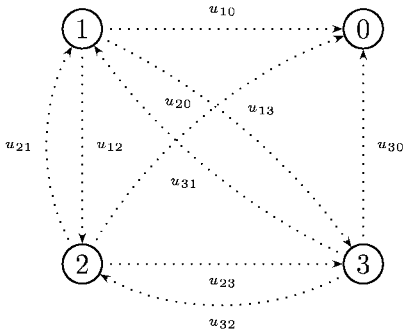

Fig. 1. Mutation scheme with 4 types of individuals ($K = 3$)

This kind of population is suitably modeled with a multi-type GWBP $\{(Z_n^{(0)}, Z_n^{(1)}, \ldots, Z_n^{(K)}),\ n \in \mathbb{N}_0\}$, where $Z_n^{(j)}$ represents the number of individuals of type j in generation n. The probability of escape when the process starts with a single individual of type j is defined as

$$r_j = P[Z_n^{(0)} \to \infty |\ Z_0^{(j)} = 1, Z_0^{(i)} = 0, \forall i \neq j].$$

In the Appendix A.1 it is shown that, if we assume that $u_{ij} = O(u)$, with u small, and neglect terms of order bigger or equal than u^2 , we can use the following approximation

$$r_j \approx \frac{m_j}{1 - m_j} \sum_{\substack{i=0 \\ i \neq j}}^{K} u_{ji} r_i\ . \tag{8}$$

We also show that, when the reproduction law of the subcritical individuals has finite variance, the error is $O\left(u^2\right)$.

At this point we are able to interpret quite well the approximations proposed in (8). The escape probability, r_j, is expressed through

$$\frac{m_j}{1-m_j}\, u_{ji}, \quad i \neq j.$$

From (4), we see that this quantity is approximately the expectation of the total number of mutants of type i produced by the initial individual of type j. This quantity is then multiplied by r_i, the escape probability of a process that starts with a type i individual, meaning that at least one of the mutants has to escape extinction. Summing up over the different types gives the different mutation paths to escape.

Next we show that the vector of the approximations of the escape probabilities, $r = [r_1, \ldots, r_K]^\top$, is the solution of a linear system of equations. Let D be the diagonal matrix with elements $\frac{m_1}{1-u_{11}m_1}, \ldots, \frac{m_K}{1-u_{KK}m_K}$, U the matrix of the mutations rates between subcritical individuals, i.e.,

$$U = \begin{bmatrix} 0 & u_{12} & \ldots & u_{1K} \\ u_{21} & 0 & \ldots & u_{2K} \\ \vdots & \vdots & \vdots & \vdots \\ u_{K1} & u_{K2} & \ldots & 0 \end{bmatrix}$$

and u_0 the vector of the mutation rates from the subcritical types to the escape type, i.e., $u_0 = [u_{10}, \ldots, u_{K0}]^\top$. From (8), it follows that the vector r satisfies approximately the system of equations

$$r = DUr + Du_0 r_0$$

which has the following solution

$$r = [I - DU]^{-1} D u_0 r_0. \tag{9}$$

The approximations (8) and the solution (9) of the corresponding system of equations were already obtained by Iwasa et al. (2003, 2004), for populations with Poisson (A.14) and geometric (A.16) offspring distribution. We showed that they can be used in a much more general situation. Since in most applications the offspring distribution is unknown, such a generalization is quite important.

In Table 1 we compare the exact values of the escape probabilities with the approximations, when the offspring distribution is binary splitting (A.13) which, together with the Poisson and the geometric distribution already considered by Iwasa *et al.*, is often used in applications. To obtain the values in the table we considered a population with 4 different types of subcritical individuals, labeled 1, 2, 3, 4. The mutation scheme is the following: type i individuals, $i = 1, 2, 3, 4$, can only mutate to the previous type, $i - 1$, with type 0 being the escape type, and mutations occur always with the same probability u. The splitting probabilities are 0.375, 0.35, 0.325 and 0.3 for type 1, 2, 3 and 4 respectively which yields $m_1 = 0.75$, $m_2 = 0.7$, $m_3 = 0.65$ and $m_4 = 0.6$ (cf Appendix A.3). For supercritical individuals, the splitting probability is 2/3 which yields $m_0 = 4/3$ and $q_0 = 0.5$. The mutation rates are 10^{-1}, 10^{-2} and 10^{-3} and these will be the typical values we will consider throughout the paper.

Table 1
Exact values and approximations for r_i, $i = 1, 2, 3, 4$

i	Result	$u = 0.1$	$u = 0.01$	$u = 0.001$
1	Approximation	1.5000e-01	1.5000e-02	1.5000e-03
	Exact	9.4375e-02	1.4043e-02	1.4896e-03
2	Approximation	3.5000e-02	3.5000e-04	3.5000e-06
	Exact	1.7265e-02	3.1996e-04	3.4676e-06
3	Approximation	6.5000e-03	6.5000e-06	6.5000e-09
	Exact	2.6906e-03	5.8336e-06	6.4279e-09
4	Approximation	9.7500e-04	9.7500e-08	9.7500e-12
	Exact	3.5073e-04	8.6211e-08	9.6273e-12

4 The waiting time to produce a successful mutant

In this section we consider the time T until a type 0 individual arises whose lineage escapes extinction. Such an individual is called a *successful mutant.* We restrict our attention to the two-type process $\{(Z_n^{(0)}, Z_n^{(1)}), n \in \mathbb{N}_0\}$ introduced in Section 2. The r.v. T takes values in the set $\{1, 2, \ldots, \infty\}$, with $T = \infty$ if the population goes extinct before a successful mutant is produced. Serra (2006), derived exact expressions for the tail of the distribution of T, and for its expectation conditioned on escape (i.e., conditioned on T being finite). We will use these results to obtain approximations for the distribution function of T, under the assumption that the mutation rate is small. The main theorem of Serra (2006) is stated below with the notation introduced in the last two sections.

Theorem 4.1 *Let f be the p.g.f. of the offspring distribution of the subcritical individuals. The distribution of T satisfies the following:*

(i) $P[T > k] = Q_k$, *for all* $k \geq 0$,

(ii) $P[T = \infty] = q_1$,

(iii)

$$E[T|T < \infty] = \sum_{k=0}^{\infty} \frac{Q_k - q_1}{1 - q_1}$$

where the Q_k *are defined recursively by*

$$Q_k = f(uq_0 + (1-u)Q_{k-1}), k \geq 1, \tag{10}$$

with $Q_0 = 1$.

In Appendix A.2 we use this theorem to develop approximations for the distribution function of T. There we show that, if the mutation rate u is small and we neglect terms of order bigger than or equal than u^2, the following first order approximation holds

$$P[T \leq k] \approx \frac{mu(1-q_0)}{1-m}(1-m^k), \quad \text{for all} \quad k \geq 1. \tag{11}$$

We also show that, when the reproduction law of type 1 individuals has finite variance, the error is $O(u^2)$.

In Figure 2 we can see how good these approximations are when the mutation rate becomes small. In each graph we plotted 3 curves: the solid line corresponds to the approximations (11); the dashed and the dotted lines correspond to the distribution function of T, determined according to Theorem 4.1, for Poisson (A.14) and binary splitting (A.13) offspring distribution, respectively. To produce the graphics, we considered the same mutation rates as in Table 1. The type 0 individuals are assumed to have extinction probability $q_0 = 0.5$ and type 1 individuals have reproduction mean $m = 0.75$. Therefore, $p = 0.375$ in the binary splitting case and $\lambda = 0.75$ in the Poisson case (cf Appendix A.3). The distribution function of T is a step function but, just for convenience, we plotted it as a continuous function.

From Figure 2 we can also see that the approximation for $P[T \leq k]$ is especially good when k is small, but, due to the accumulation of errors, it deviates from the exact value as k increases. Note that, when the mutation rate is small, mutants either appear quite early or they do not appear at all. Therefore, we need good approximations particularly for small values of k. Nevertheless it is

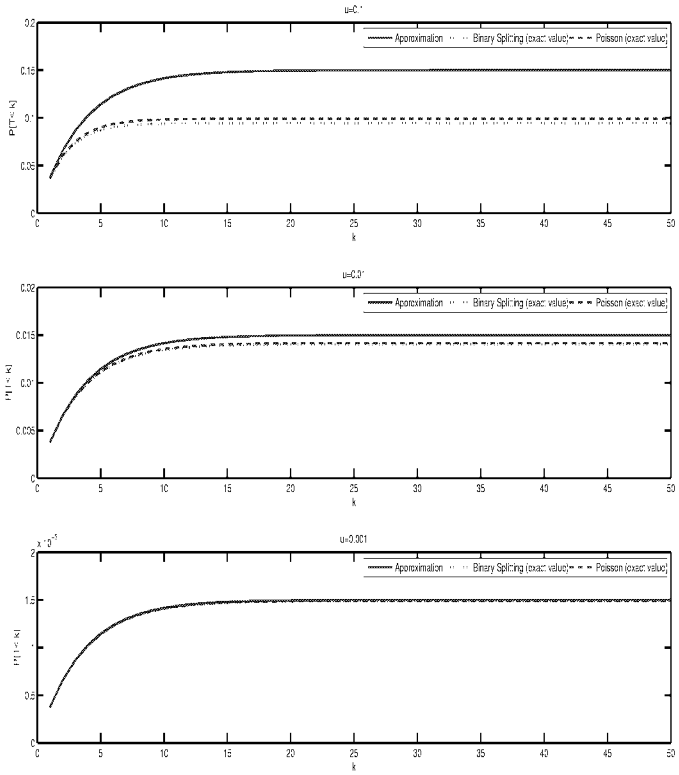

Fig. 2. Distribution Function of T

important to guarantee that this does not affect the order of the error when k is large. In fact, the use of Theorem 4.1, *(ii)*, yields

$$\lim_{k\to\infty} P[T \leq k] = r_1 \tag{12}$$

and, taking the limit as $k \to \infty$ in (11), gives

$$\frac{mu(1-q_0)}{1-m}(1-m^k) \quad \nearrow \quad \frac{mu(1-q_0)}{1-m}, \tag{13}$$

which is the first order approximation for r_1 proposed in (8) with error $O\left(u^2\right)$, as desired.

The distribution of T, conditioned on escape, has a quite nice and intriguing property. The use of $(A.11)$ and $(A.2)$ yields the following:

$$P[T > k | T < \infty] \approx [m(1-u)]^k, \quad k \geq 0, \tag{14}$$

and

$$E[T|T<\infty] \approx \frac{1}{1-m(1-u)}. \tag{15}$$

This indicates that, conditioned on escape, the number of generations up to the production of the first successful mutant has approximately a geometric distribution (A.16) with success parameter $1-m(1-u)$, for small values of u. Apparently, the fact that the mutation rate is small gives some kind of independence structure between the generations of the process, i.e., given that a successful mutation occurs, at each generation and independently of the others, a successful mutant is produced with probability $1-m(1-u)$. The use of (11) and (8) yields the following first order approximations

$$P[T>k|T<\infty] \approx m^k \quad \text{and} \quad E[T|T<\infty] \approx \frac{1}{1-m}. \tag{16}$$

In the applications it is useful to have tools to characterize the immediate risk of escape, i.e., the probability of producing a successful mutant in the next generation given that it has not been produced yet. In general one could use the hazard function of variable T, defined as $P[T=k|T>k-1]$ for $k \geq 1$, to quantify such risk, but here we propose a different version of this function. The fact that variable T has a defective distribution ($T=\infty$ when no successful mutant is produced) and that if there are no subcritical individuals alive at generation $k-1$ the probability of producing a successful mutant in generation k is zero, requires a different definition of the hazard function. So we propose the following:

$$g(k) = P[T=k|T>k-1, Z_{k-1}^{(1)}>0], \quad k \geq 1. \tag{17}$$

With this new definition we obtain the probability that a successful mutant will be produced in the next generation, given that a successful mutant was not produced before and the subcritical population is not extinct at the current generation.

In this section we provide the theoretical expression of (17) and present its form for particular offspring distributions. We have:

$$\begin{aligned} g(k) &= \frac{P[T=k]}{P[T>k-1, Z_{k-1}^{(1)}>0]} \\ &= \frac{P[T=k]}{P[T>k-1] - P[T>k-1, Z_{k-1}^{(1)}=0]}. \end{aligned} \tag{18}$$

The probabilities $P[T=k]$ and $P[T>k-1]$ are computed using Theorem 4.1. The second term in the denominator satisfies the following recursive formula:

$$
\begin{aligned}
P[T > k\ ,\ Z_k^{(1)} = 0] &= \sum_{i=0}^{\infty} p_i (uq_0 + (1-u)P[T > k-1, Z_{k-1}^{(1)} = 0])^i \\
&= f(uq_0 + (1-u)P[T > k-1, Z_{k-1}^{(1)} = 0]),\ k \geq 1,
\end{aligned}
$$

where p_i, $i \geq 0$, denotes the probability that a type 1 individual gives birth to i children and $P[T > 0, Z_0^{(1)} = 0] = 0$. So, the hazard function is given by

$$
g(k) = \begin{cases} \dfrac{P[T = k]}{P[T > k-1] - f(uq_0 + (1-u)P[T > k-2, Z_{k-2}^{(1)} = 0])} & \text{if } k \geq 2, \\ \\ 1 - f(uq_0 + 1 - u) & \text{if } k = 1. \end{cases}
$$

In Figure 3 we plot the hazard function for the offspring distributions $(i)-(iii)$ listed in Appendix A.3, with the same parameters as in Figure 2. In the linear fractional case we chose $(b, c) = (0.1875, 0.5)$. We observe that the hazard function has the same general shape for the different values of the mutation rate: there is an initial increase after which it remains more-or-less constant for a long period. The mutation rate seems to have a strong effect on the duration of the initial phase of increase: when u is big this phase is short, but for small u it is quite long. We would also like to point out that the subcritical population dies out quite early and so the probability that there are still subcritical individuals alive after a large number of generations is very small.

5 Attaining High Levels

Besides the time until the production of a successful mutant, the time that it takes for the number of individuals of the escape type to reach high levels contains important information for applications. In this section we present some results on this statistic for the two-type process introduced in Section 2.

In the following, let $\{Y_n, n \in \mathbb{N}_0\}$ be a single-type supercritical GWBP process, with the same reproduction law as the escape type individuals, starting with just one individual, i.e., $Y_0 = 1$. For $x \geq 2$, let L_x be the r.v. that represents the first generation where this process crosses level x, i.e.,

$$
L_x = \left\{ \inf_n : Y_n \geq x \right\}.
$$

Observe that L_x may be infinite because there is a positive probability that the process goes extinct before level x is attained. However, if we condition

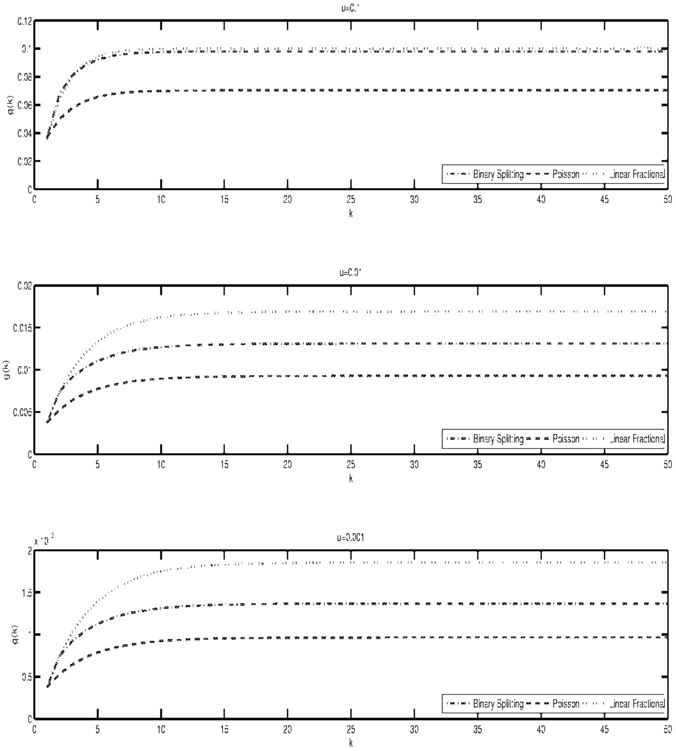

Fig. 3. Hazard Function

the process on non-extinction, L_x will be finite for all values of x, since then $Y_n \to \infty$, as $n \to \infty$.

Consider now the two-type process, $\{(Z_n^{(0)}, Z_n^{(1)}),\ n \in \mathbb{N}_0\}$, introduced in Section 2. Let T_x be the r.v. that represents the time for the number of escape type individuals to cross level x, i.e.,

$$T_x = \left\{ \inf_n : Z_n^{(0)} \geq x \right\}.$$

Again, T_x may be infinite but, if we condition on $T < \infty$, it will be finite for all values of x because, when a successful mutant is produced, it starts a process identically distributed with $\{Y_n, n \in \mathbb{N}_0\}$ conditioned on non-extinction. Our goal here is to show that, for small values of u, the r.v. $[T_x \,|\, T < \infty]$ can be approximated by the sum of two independent random variables:

$$[T \,|\, T < \infty] \quad \text{and} \quad [L_x \,|\, Y_n \to \infty].$$

Observe that it is not possible to write it as an exact sum because the number of supercritical individuals may cross level x due to the contribution of several mutants. Below we prove that, when the mutation rate is small, we can neglect the probability of such an event. This is formalized in the following lemma.

Lemma 5.1 *The distribution of the r.v. I, defined in Section 2, satisfies*

$$\lim_{u\to 0} P[I=1|T<\infty]=1 \quad \textit{and} \quad P[I\geq 2]=O(u^2),\ u\to 0.$$

The distribution of T, conditioned on $T<\infty$, was already studied in the previous section, in particular for small values of the mutation rate u. In this section we review some results on the distribution of L_x and provide simulation results for $E[L_x \mid L_x<\infty]$.

The problem of attaining high levels in a supercritical single-type GWBP was already studied in Nagaev (1971) and Rösler et al. (2001). There it is shown that, if the reproduction law has finite variance and mean $m_0>1$, as $x\to\infty$,

$$\sup_{k=0,\pm1,\pm2,\dots} \left| P\left(L_x=\lfloor a(x)\rfloor+1+k\right) - r_0 P\left(m_0^{\{a(x)\}-k-1} < W^* < m_0^{\{a(x)\}-k}\right)\right| \to 0$$

where

$$r_0 = P\left[Y_n\to\infty\right],\quad a(x)=\frac{\log(x)}{\log(m_0)} \quad \text{and} \quad W^*=\lim_{n\to\infty}\left(\frac{Y_n}{m_0^n}\mid Y_n>0\right).$$

The symbols $\lfloor z\rfloor$ and $\{z\}$ represent the integral and fractional parts of z, respectively, i.e.,

$$\lfloor z\rfloor = \left\{\sup_{n\in\mathbb{Z}} : n\leq z\right\} \quad \text{and} \quad \{z\}=z-\lfloor z\rfloor.$$

From this we can deduce that

$$\sup_{k=0,\pm1,\pm2,\dots}\left|\frac{1}{r_0}P\left(L_x-1-\lfloor a(x)\rfloor=k\right)-P\left(k\leq\{a(x)\}-\frac{\log(W^*)}{\log(m_0)}<k+1\right)\right|\to 0$$

which is equivalent to

$$\sup_{k=0,\pm1,\pm2,\dots}\left|P\left(L_x-1-\lfloor a(x)\rfloor=k|Y_n\to\infty\right)-P\left(\left\lfloor\{a(x)\}-\frac{\log(W^*)}{\log(m_0)}\right\rfloor=k\right)\right|\to 0.$$

This means that the difference that the distributions of the r.v.'s

$$\left(L_x-1-\lfloor a(x)\rfloor \mid Y_n\to\infty\right) \quad \text{and} \quad \left\lfloor\{a(x)\}-\frac{\log(W^*)}{\log(m_0)}\right\rfloor$$

assign to any event converges to zero, as $x\to\infty$. Therefore, from the distribution of $\log(W^*)$ we can, in principle, obtain limiting results on the distribution of L_x.

Rösler et al. (2001) give special attention to the case of high levels of the form m_0^n wich is the expected value of Y_n. They show that, for any $b \in [1, m_0)$

$$L_{bm_0^n} - n \xrightarrow{a.e} \eta_b^* \tag{19}$$

where

$$\eta_b^* = \begin{cases} \infty & \text{on the set } \{V = 0\}, \\ \left\lceil -\frac{\log(V/b)}{\log(m_0)} \right\rceil & \text{on the set } \{V > 0\}. \end{cases}$$

with

$$V = \lim_{n \to \infty} \frac{Y_n}{m_0^n} \quad \text{and} \quad \lfloor z \rfloor = \left\{ \inf_{n \in \mathbb{Z}} : n \geq z \right\}.$$

Since in general not much is known about the distributions of W^* and V we cannot draw many more conclusions from here.

We finish this section with some simulation results concerning the r.v. L_x. In Figure 4 we plot the value of $E[L_x \mid L_x < \infty]$ for two high values of x, 10^4 and 10^5, and processes with Poisson, binary splitting and geometric reproduction laws. We considered different values of the reproduction mean in interval $[1.1, 2)$. In each case we simulated 10.000 processes. During the simulation of a supercritical G.W.B.P. it is not always possible to observe if extinction occurs. But we can always observe if the process crosses value x. Therefore the results are conditioned on $L_x < \infty$ and not conditioned on $Y_n \to \infty$, as above. These events do not have the same probability because extinction is still possible after crossing value x. This happens with probability less or equal than $(1 - r_0)^x$, but since we will consider only high values of x, we can neglect this quantity and assume that the two events have approximately the same probability.

From the figure we can conclude that the time to cross level x increases with x and decreases as the reproduction mean m_0 increase. This was expected obviously, but it is surprising to see that such different reproduction laws, like Poisson and binary splitting, have quite similar behavior regarding this problem. Yet, we observe that the geometric distribution exhibits lower average times and this may be related to the fact that this is the reproduction law with higher variance when m_0 takes values in the interval $[1.1, 2)$.

6 Discussion

In this paper we have presented statistics for quantification of several aspects of the dynamics of escape mutants: the probability of ultimate success; the distribution of the time until appearance of a successful mutant; the hazard

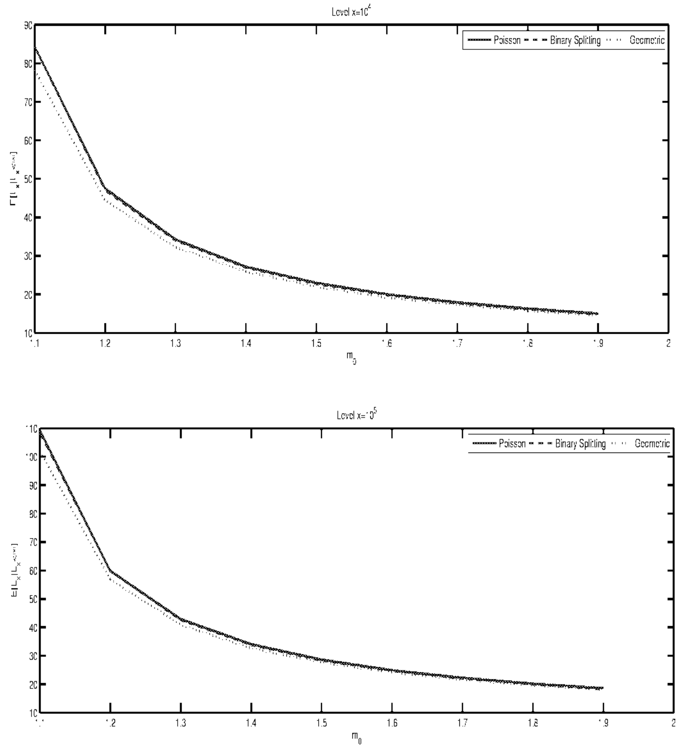

Fig. 4. Simulation results for $E[L_x|L_x < \infty]$

rate for appearance of a successful mutant; the time until escape individuals reach a certain level. Different aspects are important in different contexts. For instance, in evolutionary processes the rate of evolution is an important parameter. This can be calculated from the expected time until a successful mutant appears. The hazard rate is an insightful way to quantify the risk of introgression of introduced genes into wild populations. Another important aspect of such risks is the expected time until the number of individuals of the supercritical type reaches a specific level. This quantity gives an indication of the time until the situation becomes critical, once a successful invasion has occurred. Furthermore, this measure can be used to estimate the sensitivity of monitoring schemes to detect escape mutants. The previous remarks also apply to the detection of mutant tumor cells in chemotherapy patients. Before we proceed with a detailed discussion of the results obtained in each section, we would like to remark that results need to be derived for extensions of the model, with more than two types, and several transition schemes.

6.1 On the probability of escape

We provide a mathematically robust derivation and generalization of the results of Iwasa et al. (2003, 2004), to approximate the establishment success probability. We showed that the approximation is valid for any offspring distribution with finite variance.

Numerical results indicate that the approximation works quite well for small values of the mutation probability u. It is, however, not so good for larger u , especially in models with more than two types (see Table 1).

6.2 On the waiting time to produce a sucessful mutant

We derived an approximation for the distribution function of the time T until a successful escape mutant appears, $P[T \leq k]$. That approximation works very well for small k and u (Figure 2). It is especially important to have a good estimate of this distribution function for small values of k, since successful mutants appear early. We also derived an approximation for the distribution of T, conditioned on escape.

Once the subcritical population has gone extinct, the chance that successful mutants are produced is zero. Therefore, the hazard rate $g(k)$, derived in Section 4, is the probability per time unit that a successful mutant will be produced in generation k, given that there are still subcritical individuals alive at k. The probability that the subcritical population still exists at k will, however, go to zero in the long run.

Figure 3 shows the behavior of the hazard function for small values of k. It is natural to ask what happens to this function after a long period of time. Does it remain constant in time or does the risk of producing an escape mutant drop to zero in the long run? It is easy to see that, as k increases, both numerator and denominator in (18) decrease to 0. This leads to fluctuations of the hazard function that drop to 0 when we try to compute its value for large values of k. Nevertheless, we believe that the limit of $g(k)$, when $k \to \infty$, is not zero but equals the value that the function shows before the fluctuations start and that we can see in Figure 3. If this is the case, we can also conclude that the limit should depend on more parameters than just u, m and q_0, since in each graphic these values are common. The variance of the reproduction law of the subcritical individuals may also play a role here. For instance, the linear fractional case exhibits the larger values of the hazard function and it also has the largest reproduction variance. This question remains to be investigated.

There are also situations with a constant source of subcritical individuals. For instance, when there is a constant formation of hybrids, through gene flow from a crop. In that case we need a different model, with immigration in the subcritical population, and the resultant hazard rate will look different. This is an important generalization, which is the subject of future research.

6.3 On the attaining of high levels

We showed that the waiting time until the population of escape individuals has attained a certain level x can be approximated by the time T until a successful mutant is produced, plus the time L_x it takes the supercritical process, starting with one escape mutant, to attain x (where all distributions are conditioned on escape success). We reviewed some results on the distribution of L_x, obtained in Nagaev (1971). These results did not allow us to draw many conclusions and so we simulated supercritical processes GWBP with particular offspring distributions. The simulations suggested that the variance of the reproduction law may play a role here and this is also subject of further research.

For practical applications, it would be worth while to have limiting results for this distribution, especially for the case where the expected numbers of offspring of supercritical individuals are close to 1. Such approximations remain to be developed.

Acknowledgments: We thank Peter Jagers and Serik Sagitov for many helpful discussions and comments on previous versions of this manuscript. M. C. Serra also thanks the Center of Mathematics of the University of Minho, Portugal, and the Portuguese Foundation for Science and Technology for supporting the stay at Chalmers University. The research of P. Haccou was partly carried out during her appointment as Chalmers University jubilee professor, and was also funded by Dutch-Russian Research cooperation grant No. 047.016.013 of the Netherlands Organization for Scientific Research (NWO).

A Appendix

A.1 Approximations for the escape probabilities

We start by deriving a formula for the quantities $q_j = 1 - r_j, j \neq 0$.

$$\begin{aligned} q_j &= \sum_{m=0}^{\infty} p_m^j (u_{j0}q_0 + \ldots + u_{jj}q_j + \ldots + u_{jK}q_K)^m \\ &= f_j(u_{j0}q_0 + \ldots + u_{jj}q_j + \ldots + u_{jK}q_K) \end{aligned} \tag{A.1}$$

where $\{p_m^j, m \in \mathbb{N}_0\}$ is the offspring distribution of type j individuals and f_j is the corresponding p.g.f.. Observe that q_0 is known since it is the unique solution of equation $q_0 = f_0(q_0)$ in interval $[0, 1)$. Replacing $q_j = 1 - r_j$ in $(A.1)$ and using (7), we obtain

$$r_j = 1 - f_j\left(1 - \sum_{i=0}^{K} u_{ji}r_i\right).$$

A first order Taylor expansion of f_j around point 1 gives that, for some $0 < \theta < 1$,

$$r_j = \frac{m_j}{1 - u_{jj}m_j} \sum_{\substack{i=0 \\ i \neq j}}^{K} u_{ji}r_i - \frac{\eta(\theta, r_1, \ldots, r_K)}{1 - u_{jj}m_j} \tag{A.2}$$

where

$$\eta(\theta, r_1, \ldots, r_K) = \frac{1}{2} f''\left(1 - \theta \sum_{i=0}^{K} u_{ji}r_i\right)\left(\sum_{i=0}^{K} u_{ji}r_i\right)^2.$$

Observe that

$$\frac{\eta(\theta, r_1, \ldots, r_K)}{1 - u_{jj}m_j} = \frac{f_j''(1 - \theta y)y^2}{2(1 - u_{jj}m_j)} \leq y^2 \frac{f_j''(1)}{1 - u_{jj}m_j} \leq y^2 \frac{\sigma_j^2}{1 - m_j}$$

where $y = \sum_{i=0}^{K} u_{ji}r_i$ and σ_j^2 denotes the variance of $\{p_m^j, m \in \mathbb{N}_0\}$. In the first inequality we used the fact that f_j'' is a nondecreasing function and second inequality is a consequence of $m_j < 1$ for $j \neq 0$. Observe that y^2 is of order u^2 and, if we assume that σ_j^2 is finite, the quantity above is also of order u^2.

Using (7), another first order Taylor expansion around point 1 gives that, for some $0 < \beta < 1$,

$$\frac{m_j}{1 - u_{jj}m_j} = \frac{m_j}{1 - m_j} - \frac{m_j^2}{(1 - m_j)^2} \sum_{i \neq j} u_{ji} + \gamma(\beta) \tag{A.3}$$

where

$$\gamma(\beta) = \frac{m_j^3}{(1 - [1 - \beta \sum_{i \neq j} u_{ji}]\, m_j)^3} \left(\sum_{i \neq j} u_{ji}\right)^2 \leq \frac{m_j^3}{(1 - m_j)^3} \left(\sum_{i \neq j} u_{ji}\right)^2 \tag{A.4}$$

Assuming that $u_{ij} = O(u)$ and neglecting terms of order bigger or equal than

u^2, $(A.2)$ and $(A.3)$ yields the following approximation:

$$r_j \approx \frac{m_j}{1-m_j} \sum_{\substack{i=0 \\ i \neq j}}^{K} u_{ji} r_i \tag{A.5}$$

where $\alpha(u) \approx \delta(u)$ means that $\alpha(u) = \delta(u) + O(u^2)$. From $(A.4)$, we observe that $\gamma(\beta) = O(u^2)$ and therefore the error of the approximation will also be of order u^2.

A.2 Approximations for the distribution function of T

We will derive approximations for the distribution function of T, under the assumption that the mutation rate u is small. Let x_k, $k \geq 1$, be the probability that a successful mutant is produced before or at generation k. According to (i) of Theorem 4.1

$$x_k = P[T \leq k] = 1 - Q_k, \quad k \geq 1. \tag{A.6}$$

Using (10) we can write x_k in the following way:

$$x_k = 1 - f\left(1 - x_{k-1} - u[Q_{k-1} - q_0]\right) \tag{A.7}$$

Observe that, if u is small, the quantity

$$y_k = x_{k-1} + u[Q_{k-1} - q_0]$$

is also small because $x_{k-1} = O(u)$. In fact, the definition of x_{k-1} and of r_1 together with $(A.5)$ gives that, for all $k \geq 1$, there exists a constant $C > 0$ s.t.

$$x_{k-1} \leq r_1 \quad \Rightarrow \quad \frac{x_{k-1}}{u} \leq \frac{r_1}{u} = \frac{m\, r_0}{1-m} + \frac{O(u^2)}{u} < C.$$

Continuing with $(A.7)$, a first order Taylor expansion of f around point 1 gives that, for some $0 < \theta_1 < 1$,

$$x_k = f'(1) y_k - \eta_1(\theta_1,\ y_k) \tag{A.8}$$

with $\eta_1(\theta_1,\ y_k) = O\left(u^2\right)$. In fact

$$\eta_1(\theta_1,\ y_k) = \frac{1}{2} f''(1 - \theta_1\, y_k)\, y_k^2 \leq \frac{y_k^2}{2} f''(1) \leq \frac{y_k^2\, \sigma^2}{2} \tag{A.9}$$

where σ^2 denotes the variance of the offspring distribution of the subcritical individuals, which is assumed to be finite. Continuing with $(A.8)$ we obtain the following approximation

$$\begin{aligned} x_k &\approx f'(1)\,(x_{k-1} + u[1 - x_{k-1} - q_0]) \\ &= m\,x_{k-1} + m\,u(1 - x_{k-1} - q_0) \\ &= mu(1 - q_0) + m(1-u)x_{k-1} \end{aligned} \tag{A.10}$$

and repeating the recursion $(A.10)$ above we obtain:

$$x_k \approx mu(1-q_0) + mu(1-q_0)\,m(1-u) + \ldots + mu(1-q_0)[m(1-u)]^{k-1}.$$

Recalling the definition of x_k, we get

$$P[T \leq k] \approx mu(1-q_0)\sum_{n=0}^{k-1}[m(1-u)]^n = mu(1-q_0)\frac{1-[m(1-u)]^k}{1-m(1-u)} \tag{A.11}$$

and, the use of $(A.3)$, yields the following first order approximation

$$P[T \leq k] \approx \frac{mu(1-q_0)}{1-m}(1-m^k) \tag{A.12}$$

A.3 Most used offspring distributions

(1) Binary splitting with parameter p:

$$f(s) = 1 - p + ps^2, \qquad m = 2p, \qquad \sigma^2 = 4p(1-p) \tag{A.13}$$

(2) Poisson with parameter λ:

$$f(s) = e^{-\lambda(1-s)}, \qquad m = \lambda, \qquad \sigma^2 = \lambda \tag{A.14}$$

(3) Linear Fractional with parameters (b, c):

$$f(s) = 1 - \frac{b}{1-c} + \frac{bs}{1-cs}, \qquad m = \frac{b}{(1-c)^2}, \qquad \sigma^2 = \frac{b(1-c^2)-b^2}{(1-c)^4} \tag{A.15}$$

The case $(b, c) = (p(1-p), 1-p)$ corresponds to the Geometric distribution with success parameter p:

$$f(s) = \frac{p}{1-(1-p)s}, \qquad m = \frac{1-p}{p}, \qquad \sigma^2 = \frac{1-p}{p^2} \tag{A.16}$$

A.4 Proof of Lemma 5.1

From the mean value theorem, it follows that

$$\begin{aligned} P[I=1|T<\infty] &= \frac{P[T<\infty|I=1]P[I=1]}{P[T<\infty]} \\ &= \frac{r_0 h'(0)}{1-E[q_0^I]} = \frac{r_0\,h'(0)}{1-h(q_0)} \\ &= \frac{r_0\,h'(0)}{(1-q_0)\,h'(a)} \\ &= \frac{h'(0)}{h'(a)} \end{aligned}$$

for some $a \in (q_0, 1)$. From (3), we get

$$h'(s) = \frac{u\,f'(u\,s+(1-u)\,h(s))}{1-(1-u)\,f'(u\,s+(1-u)\,h(s))}, \qquad s \in [0,1]$$

and the first assertion is proved in the following way:

$$\begin{aligned} P[I=1|T<\infty] &= \frac{f'((1-u)h(0))[1-(1-u)f'(ua+(1-u)h(a))]}{f'(ua+(1-u)h(a))[1-(1-u)f'((1-u)h(0))]} \\ &\xrightarrow[u\to 0]{} \frac{f'(1)[1-f'(1)]}{f'(1)[1-f'(1)]} = 1. \end{aligned}$$

In the calculations above we used the fact that $h(s) \to 1$ as $u \to 0$ and this can be proved in the following way. Let I_u be the r.v. that represents the total number of mutations in a process with mutation rate u and denote by h_u its p.g.f., which, obviously, satisfies the functional equation (2), i.e., $h_u(s) = f(us + (1-u)h_u(s))$. First we show that

$$I_u \xrightarrow{d} 0 \quad \text{when} \quad u \to 0.$$

Observe that $P[I_u = 0] = h_u(0)$, is bounded and increases as u decreases to 0. Therefore it is convergent and the limit is the solution in $[0,1]$ of equation

$$\lim_{u\to 0} h_u(0) = f(\lim_{u\to 0} h_u(0)).$$

Since $f'(1) < 1$, we conclude that $\lim_{u\to 0} h_u(0) = 1$ and the convergence in distribution is proved. Finally, from the continuity theorem we conclude that

$$h_u(s) = E[s^{I_u}] \xrightarrow[u\to 0]{} E[s^0] = 1$$

which finishes the proof.

Let's proceed with the proof of the second assertion.

$$P[I \geq 2] = 1 - h(0) - h'(0) = h'(c) - h'(q_0) = c\,h''(a) \leq c\,h''(1)$$

for some constants a and c satisfying $0 < a < c < 1$. The use of (5) and a few calculations gives the following

$$\frac{h''(1)}{u^2} = \frac{\sigma^2 - m(1-m)}{[1-m(1-u)]^3} \leq \frac{\sigma^2}{(1-m)^3}$$

and this finishes the proof of the lemma.

References

Byrne, D., Cahill, M., Denholm, I., Devonshire, A.L, 1994. Biochemical identification of interbreeding between B-type and non B-type strains of the tobacco whitefly Bemisia tabaci. Biochemical Genetics, 33, 13-23.

De Barro, P.J., Hart, P.J., 2000. Mating interactions between two biotypes of the whitefly, Bemisia tabaci, (Hemiptera: Aleyrodidae) in Australia. Bulletin of Entomological research 90, 103-112.

Ellstrand, N.C., Prentice, H.C., Hancock, J.F., 1999. Gene flow and introgression from domesticated plants into their wild relatives. Annual Review of Ecology and Systematics 30, 539-63.

Guirao, P., Beitia, F., Cenis, J.L., 1997. Biotype determination of Spanish populations of Bemisia tabaci (Hemiptera: Aleyrodidae). Bulletin of Entomological Research 87, 587-593.

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

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

Jagers, P., 1975. Branching Processes with Biological Applications. Wiley, London

Maan, S.S., 1987. Interspecific and intergeneric hybridisation in wheat. Heyne, E. G., (Ed.) Wheat and wheat improvement Vol. 13, 2nd ed. ASA, CSSA and SSSA, Madison, pp. 453-461.

Michor, F., Iwasa, Y., Nowak, M.A., 2004. Dynamics of cancer progression. Nature Review Cancer 4:197-205.

Nagaev, A.V., 1971. A limit theorem for supercritical branching processes. Matematiceskie Zametki 9, No. 5, 585-592. (in russian)

Nowak, M.A., Michor,F., Komarova, N.L. , Iwasa, Y., 2004. Evolutionary dynamics of tumor suppressor gene inactivation. Proc. Natl. Acad. Sci. USA 101, 10635-10638.

Rösler, U., Topchii, V.A., Vatutin, V.A., 2001. High-level overshoot for a class of random sequences. Siberian Advances in Mathematics, v. 11, No. 3, 1-13.

Serra, M.C., 2006. On the waiting time to escape. J. Appl. Prob. 43, No.1, 296-302.

Paper C

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.

Keywords: Galton-Watson process, multitype, decomposable, escape from extinction.

AMS 2000 Mathematics Subject Classification: Primary 60J80, Secondary 92D25

---

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

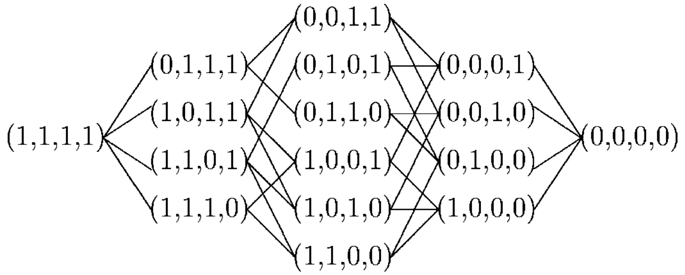

Figure 1: The network of 0-1 sequences of length $L = 4$. Here the edges represent single point mutations of probability $\mu(1-\mu)^{L-1}$. More generally, mutation between two sequences which differ in i sites has probability $\mu^i(1-\mu)^{L-i}$.

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 $i = 1, \dots, L$ and $m_0 > 1$, then a reproduction process stemming from the subcritical form $(1, \dots, 1)$ is doomed to get extinct unless a sequence of mutations results in the supercritical form $(0, \dots, 0)$. 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 $(1, \dots, 1)$ towards the escape sequence $(0, \dots, 0)$.

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,

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

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*

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

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

$$\begin{aligned} & P_1(Z_0(1) = k_0, Z_1(1) = k_1, A_i^{(i)}, 1 \le i \le k) \\ = \quad & P_1(Z_0(1) = k_0, Z_1(1) = k_1) \prod_{i=1}^{k_1} P_1(A_i) \prod_{i=k_1+1}^{k} P_0(A_i). \end{aligned} \tag{3}$$

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

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

□

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* *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)

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

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

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

□

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

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

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.

□

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

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

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*

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

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

$$\begin{aligned} &E_{100}\left(s_{00}^{Z_{100}(1)} s_{01}^{Z_{101}(1)} s_{1}^{Z_{11}(1)}\right) \\ &\quad = \frac{f_1(1, Q_{100}s_{00} + Q_{101}s_{01} + Q_{11}s_1) - f_1(q_0, Q_{101}s_{01} + Q_{11}s_1)}{Q_{100}}. \end{aligned} \tag{12}$$

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

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

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.

□

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

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

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.

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*

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

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

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

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 $s \to 1$.

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

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

and therefore

$$\begin{aligned}
&\left| \sum_{k=1}^{\infty} p_1(k)[s_1 + \mu(s_0 - s_1)a_{10}(k)]^k - \phi_1(s_1) - \mu\,(s_0 - s_1)\psi_{10}(s_1) \right| \\
\leq\ & \sum_{k=1}^{\infty} p_1(k) \left| [s_1 + \mu(s_0 - s_1)a_{10}(k)]^k - s_1^k - k\mu a_{10}(k)(s_0 - s_1)s_1^{k-1} \right| \\
\leq\ & \sum_{k=1}^{\infty} p_1(k) \left\{ \mu^2 (s_0 - s_1)^2 a_{10}^2(k) \sum_{i=1}^{k-1} (k-i) s_1^{i-1} \right\} \\
\leq\ & C^2 \mu^2 (s_0 - s_1)^2 \sum_{k=1}^{\infty} p_1(k) k \frac{1 - s_1^{k-1}}{1 - s_1} \\
=\ & C^2 \mu^2 (s_0 - s_1)^2 \, \frac{m_1 - \phi_1'(s_1)}{1 - s_1}.
\end{aligned}$$

On the other hand, (13) implies

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

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

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

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

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

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

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

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

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)

$$\begin{aligned} 0 \;&\le\; f_1^{(\mu)}(s_0+\delta_0, s_1+\delta_1) - f_1^{(\mu)}(s_0,s_1) - R^{(\mu)}(s_0,s_1,\delta_0,\delta_1) \qquad (24) \\ &\le\; \sum_{k=1}^{\infty} kp_1(k)\left(\delta_1 + \mu(\delta_0-\delta_1)a_{10}^{(\mu)}(k)\right)^2 \frac{1-\left(s_1+\mu(s_0-s_1)a_{10}^{(\mu)}(k)\right)^{k-1}}{1-s_1-\mu(s_0-s_1)a_{10}^{(\mu)}(k)} \end{aligned}$$

where

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

Relations (24) and (12) yield

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

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

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

□

6 The sequential mutation model

Suppose we can distinguish between $L+1$ types of particles, labelled $0, \dots, L$. Type i particles can only produce particles of the types $0, \dots, i$, 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

$$\begin{aligned} &E_i\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) \\ &= Q_{i0}E_{i0}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) \\ &+ Q_{i1}E_{i1}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) \end{aligned}$$

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

$$E_{i1}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) = \frac{f_i(0,\ s_{11}Q_{11},\ s_{21}Q_{21},\ \ldots,\ s_{i1}Q_{i1})}{Q_{i1}}, \tag{26}$$

$$\begin{aligned} &E_{i0}\left(s_{i0}^{Z_{i0}(1)}\, s_{i1}^{Z_{i1}(1)}\, \ldots\, s_{10}^{Z_{10}(1)}\, s_{11}^{Z_{11}(1)}\right) \\ &= \frac{f_i(1,\ Q_{10}s_{10} + Q_{11}s_{11},\ \ldots,\ Q_{i0}s_{i0} + Q_{i1}s_{i1}) - f_i(0,\ s_{11}Q_{11},\ \ldots\ ,s_{i1}Q_{i1})}{Q_{i0}}. \end{aligned} \tag{27}$$

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

$$\begin{aligned} f_i^{(\mu)}(s_0,s_1,\ldots,s_i) &= E_i^{(\mu)}\left(s_0^{Z_0(1)}\,s_1^{Z_1(1)}\,\ldots\,s_i^{Z_i(1)}\right) \qquad (28)\\ &= \sum_{k=0}^{\infty} p_i(k)\left[s_i\left(1-\sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)\right)+\sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)s_j\right]^k\\ &= \sum_{k=0}^{\infty} p_i(k)\left[s_i+\sum_{j=0}^{i-1}\mu^{i-j}a_{ij}^{(\mu)}(k)(s_j-s_i)\right]^k, \end{aligned}$$

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

$$f_i^{(\mu)}(s_0,s_1,\ldots,s_i)\to\phi_i(s_i),\ \ \phi_i(s)=\sum_{k=0}^{\infty}p_i(k)s^k. \qquad (29)$$

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

$$0<m_i<1,\ \ i=1,\ldots,L,\ \ 0<m_0<\infty. \qquad (30)$$

As in Section 2 we will assume the uniform convergence

$$\sup_{0\le j<i\le L}\ \sup_{k\ge 0}|\,a_{ij}^{(\mu)}(k)-a_{ij}(k)\,|\to 0,\ \ \mu\to 0, \qquad (31)$$

where the limit sequences are uniformly bounded

$$\sup_{0\le j<i\le L}\ \sup_{k\ge 0}a_{ij}(k)<\infty. \qquad (32)$$

Put

$$\psi_{ij}(s)=\sum_{k=1}^{\infty}kp_i(k)a_{ij}(k)s^{k-1},\ \ 0\le j<i\le L \qquad (33)$$

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

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

Define a vector $(\chi_0, \dots, \chi_L)$ 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

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

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 $\{0, \dots, i-1\}$. 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*

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

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

The limit process $\{X_{L0}(n), X_{L1}(n), \dots, X_{10}(n), X_{11}(n)\}_{n\geq 0}$ 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 $1, 2, \dots, T_L - 1$ 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 $\{0, 1, \dots, L-1\}$ according to the distribution $\{B_{L0}, \dots, B_{L,L-1}\}$, 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

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

At times $T_L+1, \dots, T_L+T_i-1$ 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 $\{0, 1, \dots, i-1\}$ according to the probability measure $\{B_{i0}, \dots, B_{i,i-1}\}$. 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

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

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

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 $j \in \{1, \dots, i-1\}$ 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

$$\left| f_i^{(\mu)}(0, Q_{11}, \dots, Q_{i1}) - \phi_i(Q_{i1}) - \sum_{j=0}^{i-1} \mu^{i-j} (Q_{j1} - Q_{i1}) \psi_{ij}(Q_{i1}) \right| =$$
$$= o\left(\sum_{j=0}^{i-1} \mu^{i-j} |Q_{j1} - Q_{i1}| \right) + O\left(\mu^i \eta^{(i)}(Q_{i1}) \left[\sum_{j=0}^{i-1} \frac{|Q_{j1} - Q_{i1}|}{\mu^j} \right]^2 \right)$$

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

$$\begin{aligned} 0 &\leq f_i^{(\mu)}(s_0 + \delta_0, \dots, s_i + \delta_i) - f_i^{(\mu)}(s_0, \dots, s_i) - R_i^{(\mu)}(s_0, \dots, s_i, \delta_0, \dots, \delta_i) \\ &\leq \sum_{k=0}^{\infty} k p_i(k) d_{k,\mu}^2(\delta_0, \dots, \delta_i) \frac{1 - d_{k,\mu}^{k-1}(s_0, \dots, s_i)}{1 - d_{k,\mu}(s_0, \dots, s_i)} \end{aligned} \tag{38}$$

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

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

From (27) and (38) it follows that

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

and it remains to check that

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

□

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 $P_L(n) = P(W_L = n)$

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

Turning to the expected waiting time

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

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

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

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

$$\begin{aligned}\chi_i &= \frac{c_i}{1-m_i}\left(1+\sum_{k=1}^{i-1}\sum_{0<j_1<j_2<\ldots<j_k<i}\frac{c_{j_k}}{1-m_{j_k}}\cdots\frac{c_{j_1}}{1-m_{j_1}}\right)\\ &= \frac{c_i}{1-m_i}\left(1+\frac{c_{i-1}}{1-m_{i-1}}\right)\ldots\left(1+\frac{c_1}{1-m_1}\right).\end{aligned}$$

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

$$\begin{aligned}&P(Y(n)=j \text{ for some } n)\\ =\ & \frac{\chi_j}{\chi_L}\frac{c_L}{1-m_L}\left(1+\sum_{k=1}^{L-j}\sum_{j<j_1<j_2<\ldots<j_k<L}\frac{c_{j_k}}{1-m_{j_k}}\cdots\frac{c_{j_1}}{1-m_{j_1}}\right)\\ =\ & \frac{c_j}{1-m_j}\left(1+\frac{c_{L-1}}{1-m_{L-1}}\right)^{-1}\ldots\left(1+\frac{c_j}{1-m_j}\right)^{-1}\\ \times\ & \left(1+\frac{c_{L-1}}{1-m_{L-1}}\right)\ldots\left(1+\frac{c_{j+1}}{1-m_{j+1}}\right)\\ =\ & \frac{c_j}{1+c_j-m_j}.\end{aligned}$$

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 $\mathrm{Bin}(L-1, m)$ 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 $\phi_{i_j}(s)$.

Acknowledgement. M.C. Serra thanks the Center of Mathematics of the University of Minho and the Portuguese Foundation for Science and Technology for supporting her stay at the Chalmers University of Technology.

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.