Defective Galton-Watson processes in a varying environment

Götz Kersting∗† Carmen Minuesa∗‡§

October 1, 2021

Abstract

We study an extension of the so-called defective Galton-Watson processes obtained by allowing the offspring distribution to change over the gene rations. Thus, in these processes, the individuals reproduce independently of the others and in accordance to some possibly defective offspring distribution depending on the generation. Moreover, the defect \(1 - f_n(1)\) of the offspring distribution at generation \(n\) represents the probability that the process hits an absorbing state \(\Delta\) at that generation. We focus on the asymptotic behaviour of these processes. We establish the almost sure convergence of the process to a random variable with values in \(\mathbb{N}_0 \cup \{\Delta\}\) and we provide two characterisations of the duality extinction-absorption at \(\Delta\). We also state some results on the absorption time and the properties of the process conditioned upon its non-absorption, some of which require us to introduce the notion of defective branching trees in varying environment.

Keywords: branching process; varying environment; defective distribution; absorption; family tree.

MSC 2020: 60J80.

1 Introduction

A special family of branching processes, known as defective Galton-Watson processes (DG-WPs), was studied in [20]. The definition of these processes is similar to classic Galton-Watson processes (GWPs), with the difference of considering defective offspring distributions. In this

∗This is a plain version of the paper accepted in Bernoulli for publication (see http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal). Both authors contributed equally to this work.

†Institut für Mathematik, Goethe Universität, Frankfurt am Main, Germany. E-mail address: kersting@math.uni-frankfurt.de

‡Department of Mathematics, Autonomous University of Madrid, Madrid, Spain. E-mail address: carmen.minuesa@uam.es

§Department of Mathematics, University of Extremadura, Badajoz, Spain. E-mail address: cminuesaa@unex.es ORCID: 0000-0002-8858-3145.
paper, we generalize the notion of defective Galton-Watson process by letting the defective offspring distribution change along the generations; the resulting process is called defective Galton-Watson processes in a varying environment (DGWPVE). The dynamics of the populations described by these processes is as follows. Under the common assumption of the independence in the reproduction, each individual at generation \( n \) produces \( k \) offspring with probability \( f_n[k] \), \( n \in \mathbb{N}, k \in \mathbb{N}_0 \). Moreover, the probability generating function (p.g.f.) of the offspring at generation \( n \), \( f_n \), satisfies \( f_n(1) = \sum_{k=0}^{\infty} f_n[k] \leq 1 \). Thus, these processes not only extend the family of DGWPs, but also the family of branching processes in a varying environment (BPVEs).

Similarly to BPVEs (see [15]), while the presence of the varying environment is a natural assumption for modelling practical situations, it entails an added difficulty in the study of these processes. In this framework, the increase in the complexity is more noticeable since the p.g.f. of the process at the \( n \)-th generation, \( f_{0,n} = f_1 \circ \cdots \circ f_n \), does not longer satisfy \( f_{0,n}(1) = 1 \). Due to all the aforementioned issues, the study of DGWPVE is challenging. Nevertheless, they constitute appropriate models for the description biological systems where each individual can show a certain trait (such as a physical feature, a mutation, or even suffering from some disease) and we are only interested in how the system evolves until the first individual shows this feature. Thus, we introduce the absorbing state \( \Delta \) to denote the presence of the trait in some individual of the population and the defect of the distribution \( f_n \), \( 1 - f_n(1) \), represents the probability that an individual at generation \( n \) develops this feature. Examples were these models can be applied are discussed in [11]. For instance, they consider the situation where a virus develops in a vaccinated host and a mutation leads to other type of individuals which are not affected by the vaccine and therefore, they can cause an epidemic outbreak. A second example is a population of cancer cells under some treatment and a mutation might turn them into resistant cells. The reader is referred to [11] for more further genetic problems were these models are of interest.

To describe these situations more precisely, let us consider a two-type branching process \( \{(Y_n(1), Y_n(2))\}_{n \in \mathbb{N}_0} \) defined as follows. The variable \( Y_n(1) \) represents the number of individuals of type 1 that are free of the mutation or disease at generation \( n \), whereas \( Y_n(2) \) denotes the number of individuals of type 2, mutants or individuals with the disease at generation \( n \). We assume that individuals of type 1 are the only ones that produce individuals of this type, but they can also give birth to individuals of type 2. Individuals of type 2, however, only produce individuals of the same type. If we also assume independence in the reproduction and that the population starts only with individuals of the first type, then the process is defined as

\[
(Y_0(1), Y_0(2)) = (N, 0), \quad (Y_{n+1}(1), Y_{n+1}(2)) = \left( \sum_{i=1}^{Y_n(1)} \xi_{ni}(1), \sum_{i=1}^{Y_n(1)+Y_n(2)} \xi_{ni}(2) \right), \quad n \in \mathbb{N}_0,
\]

where \( N \in \mathbb{N} \), and the variables of the family \( \{\xi_{ni}(1), \xi_{ni}(2) : i \in \mathbb{N}, n \in \mathbb{N}_0\} \) are independent distributed. If we additionally require that the distribution of the random vectors \( (\xi_{ni}(1), \xi_{ni}(2)) \) depend only on \( n \), we are dealing with a (degenerated) two-type branching process in a varying environment. Then, as is easy to see, up to the moment \( \tau_\Delta = \min\{n \in \mathbb{N} : Y_n(2) > 0\} \) the first
component \( \{Y_n(1)\}_{n \in \mathbb{N}_0} \) constitutes a DGWPVE, with offspring distributions given by

\[
\begin{align*}
  f_n[k] &= P[\xi_{n1}(1) = k, \xi_{n1}(2) = 0], \quad k \in \mathbb{N}_0, \\
  f_n(1) &= 1 - \sum_{k=0}^{\infty} P[\xi_{n1}(1) = k, \xi_{n1}(2) = 0] = 1 - P[\xi_{n1}(2) = 0] = P[\xi_{n1}(2) > 0].
\end{align*}
\]

For a constant environment this was already observed and applied in [13] motivated by a genetic problem in [19].

Despite the great interest of DGWPVEs and the number of papers dealing with branching processes in a varying environment (see, for instance, [12], [17], [1], [14], [4], [7], [5], [3], [18], or more recently, [21], [9], and [15]), none of them analyses this type of populations and consequently, our work provides the first results for this problem. Bearing in mind the previous interpretation, we focus on the study of the limiting behaviour of the process. In our first theorem we provide a necessary and sufficient condition for the process to escape the explosion and the absorption (at 0 or \( \Delta \)), as occurs for the BPVEs. However, we emphasize that the presence of defective distributions in the model together with their change over time (note that we identify time with generation in this framework) makes the difference between their asymptotic behaviour and the aforementioned models. First, we show that unlike the BPVEs, these processes do not become extinct almost surely. Second, we prove that these processes can avoid absorption with a positive probability, contrary to the duality extinction-absorption at \( \Delta \) that holds for DGWPs. Indeed, we establish necessary and sufficient conditions for the process to avoid its absorption and we also state the growth rate of the process in the non-absorption set. Our next results describe the asymptotic regimes, as \( n \to \infty \), of the probability that the process avoids absorption at generation \( n \) and of the expected number of individuals at generation \( n \). To determine the size of the process conditionally on its non-absorption at generation \( n \), as \( n \to \infty \), we introduce the defective branching trees. They are the counterpart of the branching trees in the non-defective case, which have been used in [6] to analyse the limiting behaviour of GWPs, or in [16] for BPVEs. For the readers convenience we explain in full detail how to adapt this construction of branching trees and state their relationship with the DGWPVEs. By this means, we provide an upper bound for the expectation of the process conditionally on its non-absorption at that generation.

Apart from this introduction, the paper is organised in three sections and one appendix. In Section 2 we provide the description of the probability model and introduce the notation that we use throughout this paper. In Section 3 we give several results on the behaviour of the process regarding its absorption either at the state 0 or \( \Delta \). Section 4 is devoted to DGWPVEs conditioned on non-absorption. In order to ease the reading, we collect the proofs of the results in a final appendix.
2 Description of the model

Mathematically, a defective Galton-Watson process in a varying environment $v = \{f_n\}_{n \in \mathbb{N}}$ is a discrete time stochastic process $\{Z_n\}_{n \in \mathbb{N}_0}$ defined recursively as:

$$Z_0 = 1, \quad Z_n = \sum_{j=1}^{Z_{n-1}} X_{nj}, \quad n \in \mathbb{N},$$

where $\{X_{nj} : n, j \in \mathbb{N}\}$ is a family of independent random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ and with range $\mathbb{N}_\Delta = \mathbb{N}_0 \cup \{\Delta\}$, and for each $n \in \mathbb{N}$, the $X_{nj}, j \in \mathbb{N}$, have a common distribution $f_n$ with weights $P[X_{nj} = k] = f_n[k], k \in \mathbb{N}_0$, and generating function

$$f_n(s) = E[s^{X_{n1}}] = \sum_{k=0}^{\infty} f_n[k] s^k, \quad s \in [0, 1],$$

satisfying $f_n(1) \leq 1$, thus $P[X_{nj} = \Delta] = 1 - f_n(1)$. Here we use the conventions

$$\Delta + k = \Delta, \quad k \in \mathbb{N}_\Delta, \quad s^\Delta = 0, \quad s \in [0, 1], \quad \sum_{j=1}^{\Delta} k_j = \Delta, \quad k_j \in \mathbb{N}_\Delta. \quad (2)$$

Without further mention in the sequel, we require that

$$0 < f_n'(1) < \infty$$

for all $n \geq 1$.

The possible defect

$$\delta_n = 1 - f_n(1)$$

of the distribution $f_n$ is interpreted as the probability with which any particle at generation $n - 1$ may send the whole process to an absorbing graveyard state $\Delta$ at generation $n$, where it stays forever. Intuitively, this process is appropriate to represent the evolution of populations where individuals reproduce independently of the others or show certain feature with a positive probability, and the distribution governing the reproduction in each generation is the same for all the individuals. More specifically, the process models populations that evolves as described and that are free of a feature developed by the individuals until the first time that one of them shows it. This event is represented by the state $\Delta$. DGWPVEs have also proved to be a useful tool in analysing other branching processes, as multitype branching processes (see [2], p.8, where $\Delta$ is identified with the state $\infty$).

The process $\{Z_n\}_{n \in \mathbb{N}_0}$ is an inhomogeneous Markov chain with state space $\mathbb{N}_\Delta$, and two of these states are absorbing: 0 and $\Delta$. Let us denote $f_{k,n} = f_{k+1} \circ \ldots \circ f_n$, for $k = 0, \ldots, n$, with the convention that $f_{n,n}(s) = s, s \in [0, 1]$. From (1), and by considering the absorption properties (2) at the graveyard state, one obtains that $E[s^{Z_n}] = f_{0,n}(s)$, for $s \in [0, 1]$.

Let us also denote the probability of extinction and absorption at the graveyard state $\Delta$ by $q = P[Z_n \to 0]$ and $\widehat{q} = P[Z_n \to \Delta]$, respectively, and let $\tau_0 = \min\{n \in \mathbb{N} : Z_n = 0\}$ be the
extinction time, \( \tau_{\Delta} = \min\{n \in \mathbb{N} : Z_n = \Delta\} \) the time of absorption at the graveyard state \( \Delta \) and \( \tau_a = \tau_0 \land \tau_{\Delta} \) the ultimate absorption time. It is straightforward to see that

\[
P[\tau_0 \leq n] = f_{0,n}(0), \quad P[\tau_{\Delta} \leq n] = 1 - f_{0,n}(1), \quad P[\tau_a \leq n] = 1 + f_{0,n}(0) - f_{0,n}(1),
\]

and

\[
P[\tau_0 < \infty] = q, \quad P[\tau_{\Delta} < \infty] = \hat{q}, \quad P[\tau_a < \infty] = q + \hat{q}.
\]

Observe that if \( \hat{q} = 1 - q \), then the process becomes absorbed at time \( \tau_a \) with probability 1. This was the case for the DGWPs with offspring distribution independent of the generation (see [20], p.2); however, as we prove below, the introduction of a varying environment in the model allows the process to avoid absorption and consequently, the explosion of the population might occur. Moreover, note that in the case \( f_n(1) = 1 \), for each \( n \in \mathbb{N} \), one obtains the classical branching process in a varying environment. Otherwise, since \( \delta_n > 0 \), for some \( n \in \mathbb{N} \), \( 0 < \hat{q} \leq 1 \) and consequently \( 0 \leq q < 1 \); thus, contrary to the case of Galton-Watson processes in a varying environment (see [15, Theorem 1]), the almost sure extinction of the population is impossible.

Henceforth, we simply write \( E[h(X)] \) to refer \( E[h(X); X \neq \Delta] \) for any variable \( X \geq 0 \) and any function \( h(\cdot) \). A main obstacle in the treatment of DGWPVEs compared to the non-defective case is the difficulty to access the moments of \( Z_n \). Indeed, we have the formulas

\[
E[Z_n] = \prod_{j=1}^{n} f'_j(f_{j,n}(1)), \quad \tag{3}
\]

\[
\frac{E[Z_n^2]}{E[Z_n]^2} = \frac{1}{E[Z_n]} + \sum_{j=1}^{n} \frac{f''_j(f_{j,n}(1))}{f'_j(f_{j,n}(1))\mu_{j,n}}, \quad \tag{4}
\]

with

\[
\mu_{j,n} = \prod_{i=1}^{j} f'_i(f_{i,n}(1)), \quad j = 0, \ldots, n,
\]

(see [15, Lemma 4]). These expressions, derived in the Appendix, are in general difficult to control since in the defective case the quantities \( f_{i,n}(1) \) can no longer be just replaced with 1. We come back to this issue.

DGWPVEs can be written in a similar way to controlled branching processes with both the offspring and control distributions depending on the generation as follows:

\[
Z_0 = 1, \quad Z_n = \sum_{j=1}^{\phi_{n-1}(Z_{n-1})} \tilde{X}_{nj}, \quad n \in \mathbb{N}, \quad \tag{5}
\]

where the random variables \( \tilde{X}_{nj}, n, j \in \mathbb{N} \), are independent and for each \( n \in \mathbb{N} \), the \( \tilde{X}_{nj}, j \in \mathbb{N} \), are identically distributed according to

\[
g_n(s) = \frac{f_n(s)}{f_n(1)}, \quad s \in [0, 1].
\]
Moreover, the family \( \{ \phi_n(k) : n \in \mathbb{N}_0, k \in \mathbb{N}_\Delta \} \) is a family of independent variables, which is also independent of the family \( \{ X_{nj} : n, j \in \mathbb{N} \} \) and such that for each \( n \in \mathbb{N} \) and \( k \in \mathbb{N}_\Delta \), the random variable \( \phi_n(k) \) is defined as

\[
\phi_n(k) = \begin{cases} 
  k & \text{with probability } f_{n+1}(1)^k, \\
  \Delta & \text{with probability } 1 - f_{n+1}(1)^k.
\end{cases}
\]

Note that from (2), one has that \( \phi_n(\Delta) = \Delta \) a.s., for each \( n \in \mathbb{N}_0 \). This observation may be used to construct a DGWPVE out of a non-defective process: define the branching process \( \{ \tilde{Z}_n \}_{n \in \mathbb{N}_0} \) in a varying environment \( \{ g_n \}_{n \in \mathbb{N}} \),

\[
\tilde{Z}_0 = 1, \quad \tilde{Z}_n = \sum_{j=1}^{\tilde{Z}_{n-1}} X_{nj}, \quad n \in \mathbb{N},
\]

with values in \( \mathbb{N}_0 \), and the stopping time

\[
\tau_\Delta = \min\{ n \in \mathbb{N} : \phi_{n-1}(\tilde{Z}_{n-1}) = \Delta \}.
\]

Then

\[
Z_n = \begin{cases} 
  \tilde{Z}_n, & \text{if } \tau_\Delta > n, \\
  \Delta, & \text{if } \tau_\Delta \leq n,
\end{cases}
\]

is a DGWPVE \( v = \{ f_n \}_{n \in \mathbb{N}} \). This coupling will be useful in different proofs below.

### 3 Some results on absorption and explosion

In this section, we provide some results related to the absorption and explosion of a DGWPVE. The first result establishes the almost sure convergence of the process \( \{ Z_n \}_{n \in \mathbb{N}_0} \) to a random variable \( Z_\infty \) with values in the set \( \mathbb{N}_\Delta \cup \{ \infty \} \). As mentioned before, for a DGWPVE the duality explosion-absorption (either at the state 0 or \( \Delta \)) does not always hold true, and similarly to the Galton-Watson process in varying environment (see Lindvall [17]), the event \( \{ 0 < Z_\infty < \infty \} \) may have positive probability; the next theorem also establishes for that situation the same necessary and sufficient condition as was detected by Lindvall in the non-degenerate case.

**Theorem 1.** Let \( \{ Z_n \}_{n \in \mathbb{N}_0} \) be a DGWPVE \( v = \{ f_n \}_{n \in \mathbb{N}} \). There exists a random variable \( Z_\infty \) with values in \( \mathbb{N}_\Delta \cup \{ \infty \} \) such that as \( n \to \infty \),

\[
Z_n \to Z_\infty \quad \text{a.s.}
\]

Moreover,

\[
P[0 < Z_\infty < \infty] > 0 \quad \iff \quad \sum_{n=1}^{\infty} (1 - f_n[1]) < \infty.
\]

6
Observe that from Theorem 1, one has that a necessary condition for the event \( \{0 < Z_\infty < \infty\} \) to have positive probability is that \( f_n[1] \to 1 \), as \( n \to \infty \) (and consequently, \( \delta_n \to 0 \)); indeed, those convergences need to be fast in the terms described in (7). These facts are illustrated in the examples below. Let us remark that the necessary and sufficient condition provided in Theorem 1 is equivalent to the notion of complete convergence of the offspring random variables \( X_{n1} \) to 1, as introduced by Hsu and Robbins [8] and meaning that \( \sum_{n=1}^{\infty} P[|X_{n1} - 1| > \epsilon] < \infty \) for all \( \epsilon > 0 \). Indeed, if \( \epsilon \in (0, 1) \), then
\[
\sum_{n=1}^{\infty} (1 - f_n[1]) = \sum_{n=1}^{\infty} P[|X_{n1} - 1| > \epsilon].
\]
Taking into account that the variables of the family \( \{X_{n1} : n \in \mathbb{N}\} \) are independent, by using the first and second Borel Cantelli lemmas, we have that the previous condition is equivalent to the almost sure convergence \( X_{n1} \to 1 \) as \( n \to \infty \). This renders some plausibility to the equivalence (7). Let us mention that the somewhat stronger condition \( \sum_{n=1}^{\infty} E[|X_{n1} - 1|] < \infty \) has been used by Ispány [10] in the context of branching processes with immigration in a varying environment.

**Examples 1.**

(a) Let us consider a DGWPVE with environment satisfying that \( f_1(s) = 2^{-1} s \) and \( f_n(s) = (1 - n^{-1}) s \), for \( n \geq 2 \) and \( s \in [0, 1] \). Then,
\[
\sum_{n=1}^{\infty} (1 - f_n[1]) = \sum_{n=1}^{\infty} \delta_n = \frac{1}{2} + \sum_{n=2}^{\infty} \frac{1}{n} = \infty,
\]
and consequently, \( P[0 < Z_\infty < \infty] = 0 \).

(b) Let us consider a DGWPVE such that for \( s \in [0, 1] \), \( f_1(s) = 2^{-1} s \) and \( f_n(s) = (1 - n^{-2}) s \), for \( n \geq 2 \). Then,
\[
\sum_{n=1}^{\infty} (1 - f_n[1]) = \sum_{n=1}^{\infty} \delta_n = \frac{1}{2} + \sum_{n=2}^{\infty} \frac{1}{n^2} < \infty,
\]
and consequently, \( P[0 < Z_\infty < \infty] = P[Z_\infty = 1] > 0 \).

In view of Theorem 1, it is interesting to study the conditions for having the duality extinction-absorption at \( \Delta \) that holds for DGWPs; Theorem 2 and Proposition 1 deal with that issue. Let us first introduce the notation:
\[
\mu_0 = 1, \quad \mu_n = \prod_{i=1}^{n} f_i'(1), \quad n \in \mathbb{N}.
\]

**Theorem 2.** Let \( \{Z_n\}_{n \in \mathbb{N}_0} \) be a DGWPVE \( v = \{f_n\}_{n \in \mathbb{N}} \). Assume that there exists \( c > 0 \) such that for all \( n \in \mathbb{N} \),
\[
E[X_{n1}^2; X_{n1} \geq 2] \leq cE[X_{n1}; X_{n1} \geq 2]E[X_{n1} \mid X_{n1} \geq 1]. \tag{8}
\]
Then

\[ P[\tau_a = \infty] > 0 \quad \iff \quad \inf_{n \in \mathbb{N}} \mu_n > 0, \quad \sum_{n=1}^{\infty} \delta_n \mu_{n-1} < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{f_n''(1)}{f_n'(1)} \mu_n < \infty. \]

Obviously the regularity condition in (8) is equivalent to

\[ E[\tilde{X}_{n_1}; \tilde{X}_{n_1} \geq 2] \leq cE[\tilde{X}_{n_1}; \tilde{X}_{n_1} \geq 2]E[\tilde{X}_{n_1}] \tilde{X}_{n_1} \geq 1]. \] (9)

It is a uniformity assumption introduced in [15] for the study of branching processes in a varying environment and it is satisfied by a wide class of distributions.

**Examples 2.** As an application of the previous theorem, in the following examples we show two DGWPVEs in which the explosion set has null probability and positive probability, respectively.

(a) Let us consider a DGWPVE with environment satisfying that for \( s \in [0, 1], \) \( f_n(s) = (1 - n^{-2}2^{-n})s^2, \ n \in \mathbb{N}. \) It is easy to check that condition (8) holds for \( c \geq 2 \) and since \( f_n[1] = 0 \) for \( n \in \mathbb{N}, \) \( P[0 < Z_{\infty} < \infty] = 0. \) Moreover, due to the fact that \( f_n[0] = 0, \) \( n \in \mathbb{N}, \) one has \( P[Z_n \to 0] = 0, \) and since \( \prod_{i=1}^{n}(1 - i^{-2}2^{-i}) \) converges to a non-zero number,

\[ \sum_{n=1}^{\infty} \delta_n \mu_{n-1} = \frac{1}{2} \sum_{n=1}^{\infty} \left( \frac{1}{n} \prod_{i=1}^{n-1} \left( 1 - \frac{1}{i^2} \right) \right) = \infty, \]

implying \( P[\tau_a = \infty] = 0. \) In particular, this process gets absorbed at \( \Delta \) a.s.

(b) Let us consider a DGWPVE such that for \( s \in [0, 1], \) \( f_n(s) = (1 - n^{-2}2^{-n})s^2, \ n \in \mathbb{N}. \) With the same arguments as in the previous example, it is easy to verify that \( P[0 < Z_{\infty} < \infty] = 0 \) and \( P[Z_n \to 0] = 0, \) and condition (8) holds true for \( c \geq 1. \) Now, we use the fact that \( \prod_{i=1}^{n}(1 - i^{-2}2^{-i}) \) converges to a non-zero number to get that \( \mu_n \to \infty, \) as \( n \to \infty, \)

\[ \sum_{n=1}^{\infty} \delta_n \mu_{n-1} = \frac{1}{2} \sum_{n=1}^{\infty} \left( \frac{1}{n^2} \prod_{i=1}^{n-1} \left( 1 - \frac{1}{i^2} \right) \right) < \infty, \]

and

\[ \sum_{n=1}^{\infty} \frac{f_n''(1)}{f_n'(1)} \mu_n = \sum_{n=1}^{\infty} \frac{1}{2n} \prod_{i=1}^{n} \left( \frac{1}{1 - i^{-2}} \right) < \infty, \]

and consequently, \( P[\tau_a = \infty] = P[Z_{\infty} = \infty] > 0. \)

The coupling in (5) also enables to describe the asymptotic behaviour of the defective process \( \{Z_n\}_{n \in \mathbb{N}_0} \) on the non-absorption set \( \{\tau_a = \infty\}. \) Observe that the sequence of random variables \( \{W_n\}_{n \in \mathbb{N}_0}, \) with \( \tilde{W}_n = \tilde{Z}_n \prod_{i=1}^{n} \frac{f_i(1)}{f_i'(1)} \), is a non-negative martingale, and hence, there exists an integrable and non-negative random variable \( \tilde{W} \) such that \( \tilde{W}_n \) converges a.s. to \( \tilde{W}, \) as \( n \to \infty. \) Moreover, under the equivalent conditions formulated in Theorem 2, it follows that
\[\sum_{n=1}^{\infty} \delta_n < \infty, \text{ thus } \prod_{i=1}^{n} f_i(1) = \prod_{i=1}^{n} (1 - \delta_i) \text{ converges to a strictly positive limit and then, } Z_n/\mu_n \text{ converges to a non-negative random variable } W. \] Also, because of (9) we have from [15] Theorem 2] that \( \{\tilde{Z}_\infty > 0\} = \{\tilde{W} > 0\} = \{W > 0\} \) a.s. The following result is straightforward from the fact that on the set \( \{\tau_a = \infty\} \), both processes \( \{Z_n\}_{n \in \mathbb{N}_0} \) and \( \{\tilde{Z}_n\}_{n \in \mathbb{N}_0} \) visit the same states and consequently \( \{\tau_a = \infty\} \subset \{\tilde{Z}_\infty > 0\} = \{W > 0\} \) a.s.

**Corollary 1.** Under the assumption (8) from Theorem [2] if \( P[\tau_a = \infty] > 0 \), then there exists a non-negative random variable \( W \) such that, as \( n \to \infty \),

\[
\frac{Z_n}{\mu_n} \to W > 0 \quad \text{a.s. on } \{\tau_a = \infty\}.
\]

Theorem 2 shows that at least \( \delta_n \to 0 \) as \( n \to \infty \) is required in order to have a strictly positive probability \( P[\tau_a = \infty] \) of non-absorption. Conversely, in the situation, where e.g. \( \inf_{n \geq 1} \delta_n > 0 \) holds, the question arises how to estimate the absorption probabilities \( P[\tau_a > n] \). Applying the mean-value theorem to the formula \( P[\tau_a > n] = f_{0,n}(1) - f_{0,n}(0) \) yields

\[
\prod_{i=1}^{n} f'_i(f_{i,n}(0)) \leq P[\tau_a > n] \leq \prod_{i=1}^{n} f'_i(f_{i,n}(1)), \quad n \in \mathbb{N}.
\]

The lower bound can be 0 and the upper bound greater than or equal to 1, therefore one may ask whether one can do better. In a first step, we provide the following result.

**Proposition 1.** For each \( n \in \mathbb{N} \),

\[
P[\tau_a > n] \leq \inf_{i \leq n} \mu_i.
\]

Moreover, if there exists a constant \( c > 0 \) such that for all \( n \in \mathbb{N} \),

\[
E[X_{n1}^2; X_{n1} \geq 2] \leq cE[X_{n1}; X_{n1} \geq 2],
\]

then, there is a constant \( c' > 0 \) such that as \( n \to \infty \)

\[
\frac{E[Z_n]^2}{E[Z_n^2]} \leq P[\tau_a > n] \leq c'E[Z_n]^2.
\]

Condition (10) is stronger than (9), it implies that \( E[X_{n1}^2] \) is bounded uniformly in \( n \) (see formula (18) below).

Combining these bounds on \( P[\tau_a > n] \) with the equation (4) confronts us with the problem to obtain estimates for the terms \( f_{i,a}(1) \) from the given distributions \( f_n \). In general this appears to be difficult. The following theorem considers the case where all solutions \( \theta_n \) of the equations \( f_n(s) = s \) belong to some interval \([\rho, \sigma]\) with \( 0 < \rho < \sigma < 1 \). For \( 0 < s \leq 1 \) denote

\[
\mu_n(s) = \prod_{i=1}^{n} f'_i(s), \quad \nu_n(s) = \sum_{i=1}^{n} \frac{1}{\mu_i(s)}.
\]
Theorem 3. Under the assumption (10), then for any $\rho \in (0, 1)$ and $n_0 \in \mathbb{N}$ such that $f_n(\rho) \geq \rho$ for all $n \geq n_0$ we have

$$\liminf_{n \to \infty} \frac{E[Z_n]}{\mu_n(\rho)} > 0, \quad \text{and} \quad \liminf_{n \to \infty} \nu_n(\rho)P[\tau_a > n] > 0.$$

If, in addition, $\inf_{n \geq 1} \delta_n > 0$, $\inf_{n \geq 1} f_n''(\rho) > 0$, then for any $\sigma \in (\rho, 1)$ such that $f_n(\sigma) \leq \sigma$ for all $n \geq n_0$, and for any $\varepsilon > 0$ we have

$$\limsup_{n \to \infty} \frac{E[Z_n]}{\mu_n(\sigma + \varepsilon)} < \infty, \quad \text{and} \quad \limsup_{n \to \infty} \nu_n(\sigma + \varepsilon)P[\tau_a > n] < \infty.$$

Examples 3. (a) If $\inf_{n \geq 1} f_n[0] > 0$ then we may set $\rho = \inf_{n \geq 1} f_n[0]$. If $\sup_{n \geq 1} f_n[1] < 1$ then we may set $\sigma = \sup_{n \geq 1} f_n[1]$.

(b) In some important cases the numbers $\rho$ and $\sigma$ can be explicitly evaluated. First let us look at the case of defective binary offspring meaning that $f_n[k] = 0$ for $k > 2$ and all $n \geq 1$. Denote $p_n = f_n[2]$, $q_n = f_n[1]$ and $r_n = f_n[0]$ and suppose that there are $0 < \theta_n < 1$ fulfilling

$$f_n(\theta_n) = \theta_n.$$

These equations are quadratic with two roots, and the $\theta_n$ is the smaller root, given by

$$\theta_n = \frac{1 - q_n}{2p_n} - \sqrt{\frac{(1 - q_n)^2}{4p_n^2} - \frac{r_n}{p_n}}.$$

It follows

$$\rho = \inf_{n \geq 1} \theta_n, \quad \sigma = \sup_{n \geq 1} \theta_n,$$

whenever $\inf_{n \geq 1} \delta_n > 0$ and $\inf_{n \geq 1} r_n > 0$.

Second we consider the case of defective linear fractional distributions. In this case the generating functions are of the form

$$f_n(s) = q_n + \frac{r_n}{1 - p_n s},$$

and the equation $f_n(\theta_n) = \theta_n$ with $0 < \theta_n < 1$ reduces once more to a quadratic equation resulting in

$$\theta_n = \frac{1 + p_n q_n}{2p_n} - \sqrt{\frac{(1 + p_n q_n)^2}{4p_n^2} - \frac{r_n + q_n}{p_n}}.$$

Again, $\rho$ and $\sigma$ are the infimum and supremum of these quantities, whenever $\inf_{n \geq 1} \delta_n > 0$ and $\inf_{n \geq 1} (q_n + r_n) > 0$. 

10
(c) Let $f$ be a distribution with defect $1 - f(1) > 0$ and let $f_n$ be a sequence of defective distributions such that $f_n(s) \to f(s)$ as $n \to \infty$ for all $0 \leq s \leq 1$. Let $\theta$ be the unique solution of the equation $f(\theta) = \theta$ and assume that $\theta > 0$. Then, letting $\rho = \theta - \eta$ and $\sigma = \theta + \eta$, all assumptions of Theorem 3 are satisfied for any $0 < \eta < \min\{\theta, 1 - \theta\}$.

Also, since $f_n'(s) \to f'(s)$ for all $s < 1$ and since $f'(s)$ is strictly increasing, it follows that for any $\eta > 0$

$$\lim_{n \to \infty} \frac{\mu_n(s)}{f'(s - \eta)^n} = \infty \quad \text{and} \quad \lim_{n \to \infty} \frac{\mu_n(s)}{f'(s + \eta)^n} = 0,$$

or equivalently

$$\lim_{n \to \infty} \frac{1}{n} \log \mu_n(s) = \log f'(s).$$

If $f'(s) < 1$, we observe that $\mu_n(s)$ is exponentially decreasing in $n$ at the negative rate $\log f'(s)$, and then, as is easy to see, $\nu_n(s)$ is increasing at the rate $- \log f'(s)$, that is

$$\lim_{n \to \infty} \frac{1}{n} \log \nu_n(s) = - \log f'(s).$$

Now $f'(\theta) < 1$. Therefore, applying Theorem 3 we obtain

$$\lim_{n \to \infty} \frac{1}{n} \log E[Z_n] = \log f'(\theta) \quad \text{and} \quad \lim_{n \to \infty} \frac{1}{n} \log P[\tau_\alpha > n] = - \log f'(\theta).$$

**Remark 1.** Along similar lines we can get estimates on extinction probabilities. Let

$$q_l = \lim_{n \to \infty} f_{l,n}(0)$$

which is the probability of ultimate extinction given that there is one individual at generation $l$. Note that $q_l = f_{l,n}(q_n)$ for $0 < l \leq n$. Let us suppose again that there is a number $\sigma \in (0,1)$ such that

$$f_n(\sigma) \leq \sigma$$

for all $n \geq 1$. Then, along the same lines as above, it follows that $f_{l,n}(0) \leq f_{l,n}(\sigma) \leq \sigma$ and consequently

$$q_l \leq \sigma$$

for all $l \geq 1$. Now

$$P[n < \tau_0 < \infty] = q_0 - f_{0,n}(0) = f_{0,n}(q_n) - f_{0,n}(0) \leq f_{0,n}(\sigma) - f_{0,n}(0).$$

Applying the mean-value theorem we have $f_{0,n}(\sigma) - f_{0,n}(0) \leq \sigma \prod_{i=1}^n f_i'(f_{i,n}(\sigma))$ and finally

$$P[n < \tau_0 < \infty] \leq \sigma \prod_{i=1}^n f_i'(\sigma).$$

Similarly one derives the lower bound $P[n < \tau_\Delta < \infty] \geq (1 - \sigma) \prod_{i=1}^n f_i'(f_{i,n}(\sigma)).$
4 Conditioned DGWPVEs and defective branching trees

In this section we consider the random variables $Z_n$, conditioned on the event $\{\tau_a > n\}$. In particular, we obtain the following theorem on the conditional expectation of $Z_n$, which can be considered as a counterpart of Theorem 2 (a) in [20] in the case of varying environments. For other conditional moments similar estimates are also feasible but we omit them for sake of brevity.

Theorem 4. If $\inf_{n \in \mathbb{N}} f_n(0) = \alpha > 0$, and $\sup_{n \in \mathbb{N}} f_n(1) = \beta < 1$, then for some $c > 0$

$$E[Z_n \mid \tau_a > n] \leq 1 + cf_n(1) \left(1 + \sum_{j=0}^{n-1} \beta^j \left(1 + \frac{f''_{n-j}(1)}{f'_n(1)}\right) \right).$$

In particular, if condition (11) also holds, then the sequence $\{E[Z_n \mid \tau_a > n]\}_{n \in \mathbb{N}_0}$ is bounded.

It seems little promising to prove this result by those tools we used so far, at least Theorem 3 provides for this purpose insufficient control on the probability of the event $\{\tau_a > n\}$. Therefore we come up with a different means, namely a probabilistic construction of the conditional defective branching tree. Such branching trees have been applied to study properties of GWP (see [6]) or BPVEs (see [16, Chapter 2]). These structures are richer frameworks than the family of variables $\{X_{ni} : n, i \in \mathbb{N}\}$ associated with the branching process since they provide information on the genealogical relationship between any two individuals of the population.

Defective family trees

For the readers convenience we introduce this approach in full detail. First, we introduce the formal definitions of defective family tree. To that end, we make use of the Ulam-Harris labelling with some modifications in order to indicate the absorption at the state $\Delta$. Let us consider the two following sets

$$\mathcal{I} = \bigcup_{n=0}^{\infty} \mathbb{N}^n, \quad \text{and} \quad \mathcal{I}_\Delta = \bigcup_{n=0}^{\infty} (\mathbb{N} \cup \{\Delta\})^n,$$

with the convention $\mathbb{N}^0 = \emptyset$ and $(\mathbb{N} \cup \{\Delta\})^0 = \emptyset$. Individuals in the population correspond to elements of the set $\mathcal{I}$ written as finite strings, where the element $\emptyset$ is known as the founding ancestor or the root. Moreover, given an individual $i \in \mathcal{I}$, we denote $g(i)$ the generation of the individual $i$ and write $i\Delta$ to indicate that the individual $i$ sends the process to the graveyard state $\Delta$ at the next generation. Thus, the elements of type $i\Delta$ do not represent individuals of the population and we refer to them as defective elements. We also write $c(i)$ to refer to the number of children of the individual $i$ when $i$ does not send the process to $\Delta$ (including the case $c(i) = 0$), and write $c(i) = \Delta$ otherwise. Moreover, we make the following assumptions:

- If $g(i) = n$, then $i$ is a string $j_1 \ldots j_n$, with $j_1, \ldots, j_n \in \mathbb{N}$.
- If $i$ has $c = c(i) \in \mathbb{N}$ children, then the label of each child is $ij$, $1 \leq j \leq c$. 

12
• If \( n \in \mathbb{N} \), then the progenitor of \( i = j_1 \ldots j_n \) is the truncated string \( i' = j_1 \ldots j_{n-1} \).

Under the previous considerations, the population is included in a subset \( t \) of \( \mathcal{I}_\Delta \), whose elements satisfy the following properties:

(i) \( \emptyset \in t \).

(ii) Let \( i \in \mathcal{I} \), and \( j \in \mathbb{N}_\Delta \). If \( ij \in t \), then \( i \in t \).

(iii) If there exists \( i \in t \) equal to a string \( j_1 \ldots j_n \Delta \), with \( j_1, \ldots, j_n \in \mathbb{N} \), then all the elements of \( t \) are strings of length at most \( n + 1 \).

(iv) Let \( i \in \mathcal{I} \) and \( j \in \mathbb{N}_\Delta \). If \( ij \in t \), then either \( j \in \mathbb{N} \) and \( ij' \in t \) for all \( 1 \leq j' \leq j \), or else \( j = \Delta \) and \( ij' \notin t \) for all \( j' \in \mathbb{N} \).

(v) For \( i \in t \), there exists a \( j \in \mathbb{N} \) such that \( ij \notin t \).

A set \( t \) satisfying conditions (i)-(v) is called defective family tree.

![Figure 1: Example of a defective family tree with a defective element a t the third generation.](image)

The height is \( h(t) = 2 \) and the population sizes are \( Z_1(t) = 3 \), \( Z_2(t) = 4 \), and \( Z_3(t) = \Delta \).

We now define the height and generation sizes associated with a defective family tree. To that end, let us introduce the following sets

\[
\mathcal{I}^{(n)} = \bigcup_{l=0}^{n} \mathbb{N}^l, \quad \text{and} \quad \mathcal{I}_\Delta^{(n)} = \bigcup_{l=0}^{n} (\mathbb{N} \cup \{\Delta\})^l.
\]

Then, the height of the defective family tree \( t \) is

\[
h(t) = \max\{n \geq 0 : t \cap \mathcal{I}_\Delta^{(n)} \subseteq \mathcal{I}^{(n)} \text{ and } t \cap \mathcal{I}^{(n)} \neq t \cap \mathcal{I}^{(n-1)}\},
\]

and it can be finite or infinite. That is: if the population gets extinct or is absorbed at generation \( n \), then the height of the tree is \( n - 1 \), and else the height is infinite. The generation sizes of the tree \( t \) are defined as follows:

\[
z_n(t) = \begin{cases} 
\#\{i \in t : g(i) = n\}, & \text{if } n \leq h(t), \\
\Delta, & \text{if } n > h(t) \text{ and there exists } j_1 \ldots j_{h(t)} \Delta \in t, \\
0, & \text{if } n > h(t) \text{ and } j_1 \ldots j_{h(t)} \Delta \notin t, \text{ for any } j_1, \ldots, j_{h(t)} \in \mathbb{N},
\end{cases}
\]
Let \( t, t' \) be two defective family trees and \( h \in \mathbb{N}_0 \), we write \( t \overset{h}{=} t' \) if and only if
\[
\{ t \in \mathcal{T} : t \overset{h}{=} t' \}, \quad h \in \mathbb{N}_0, \ t \in \mathcal{T}
\]
and we say that \( t \) and \( t' \) coincide up to the height \( h \). Let us denote the set of all the defective family trees \( t \in \mathcal{T} \) by \( \mathcal{T} \) and consider the \( \sigma \)-field generated by all the sets of the form
\[
\{ t' \in \mathcal{T} : t \overset{h}{=} t' \}, \quad h \in \mathbb{N}_0, \ t \in \mathcal{T}
\]
Then, \( t \mapsto h(t) \) and \( t \mapsto z_n(t) \) are measurable mappings for \( n \in \mathbb{N}_0 \).

**Defective branching trees in varying environment**

Now, we introduce the definition of defective branching trees in a varying environment emulating those for the non-defective case in [16, Chapter 2]. A defective branching tree in varying environment (DBTVE) \( v = \{ f_n \}_{n \in \mathbb{N}} \) is a \( \mathcal{T} \)-valued random variable \( T \) with distribution given as follows:

(i) If \( h \in \mathbb{N}_0 \) and \( t \leq \mathcal{I} \), then
\[
P_v \left[ T \overset{h}{=} t \right] = \prod_{i \in \mathcal{I} : g(i) < h} f_{g(i)+1}(c(i)).
\]

(ii) If \( h \in \mathbb{N}_0 \), \( t \in \mathcal{T} \) is such that \( t \nsubseteq \mathcal{I} \) and \( h \leq h(t) \), then
\[
P_v \left[ T \overset{h}{=} t \right] = \prod_{i \in \mathcal{I} : g(i) < h} f_{g(i)+1}(c(i)).
\]

(iii) If \( h \in \mathbb{N}_0 \), \( t \in \mathcal{T} \) is such that \( t \nsubseteq \mathcal{I} \) and \( h > h(t) \), then
\[
P_v \left[ T \overset{h}{=} t \right] = \prod_{i \in \mathcal{I} : g(i) < h - 1} f_{g(i)+1}(c(i)) \prod_{i \in \mathcal{I} : g(i) = h(t), \ c(i) \neq \Delta} f_{g(i)+1}(c(i)) \cdot 
\]
\[
\prod_{i \in \mathcal{I} : g(i) = h(t), \ c(i) = \Delta} (1 - f_{g(i)+1}(1)).
\]

Given a DBTVE \( T \), we can define a DGWPVE \( v = \{ f_n \}_{n \in \mathbb{N}} \) associated with \( T \); this is the process \( \{ Z_n \}_{n \in \mathbb{N}_0} \) given by
\[
Z_n(\omega) = z_n(T(\omega)), \quad \text{for each} \ \omega \in \Omega, \ n \in \mathbb{N}_0.
\]
By using the relation between the DBTVE and the DGWPVE associated with it, we define the extinction time, absorption time at $\Delta$, and ultimate absorption time of a DBTVE as $T$ point out that the absorption times stem from the random tree that for $j, n \in \mathbb{N}_0$.

Then, $T$ is a DBTVE

$$P_{v_j}[\tau_0 \leq n] = f_{j,n+j}(0),$$

$$P_{v_j}[\tau_\Delta \leq n] = 1 - f_{j,n+j}(1),$$

$$P_{v_j}[\tau_a \leq n] = 1 - f_{j,n+j}(1) + f_{j,n+j}(0),$$

where $v_j = \{f_{j+1}, f_{j+2}, \ldots \}$ (note that $v = v_0$). We write $\tau_0(T), \tau_\Delta(T)$ and $\tau_a(T)$, if we like to point out that the absorption times stem from the random tree $T$.

Let us also denote the subtrees of the random tree $T$ founded by the children of the root as $T_i, 1 \leq i \leq z_1(T)$, and let $R_n$ be the rank of the left-most child of the root that has a descendant in generation $n$,

$$R_n = \min \{1 \leq i \leq z_1(T) : 0 < z_{n-1}(T_i) \neq \Delta \}, \quad n \in \mathbb{N}_0,$$

with $\min \emptyset = \infty$. We note that $R_n < \infty$ a.s. on the set $\{\tau_a > n\}$. Similarly to DBTVEs, a defective tree $t$ satisfying $h(t) \geq 1$ can be divided into $c$ subtrees $t_1, \ldots, t_c$, where $c = c(\emptyset)$ is the number of offspring of the founding ancestor. Moreover, these subtrees can be written as $t_j = \{i \in \mathcal{I}_\Delta : ji \in t\}$, for $j = 1, \ldots, c$.

Now we like to investigate defective branching trees conditioned to achieve a certain height. They can be built up by a stepwise probabilistic construction as introduced by Geiger [6] for standard Galton-Watson processes. The single steps proceed according to the following lemma, which gives insight into the structure of such conditional defective trees.

**Lemma 1.** Let $n \in \mathbb{N}$ and $v = \{f_n\}_{n \in \mathbb{N}}$ be a varying environment. Let the random defective tree $T$ be composed of the defective subtrees $T_1, \ldots, T_C$ in generation one, where $C$ is the random variable denoting the number of children of the founding ancestor and let $D$ be a $\mathbb{N}$-valued random variable with $1 \leq D \leq C$. Assume that:

(i) The distribution of $(D, C)$ is a proper probability distribution given by $\{g_n[d, c]\}_{1 \leq d \leq C; c \in \mathbb{N}}$, with

$$g_n[d, c] = \frac{f_{1,n}(1) - f_{1,n}(0)}{f_{0,n}(1) - f_{0,n}(0)} f_1[c] f_{1,n}(0)^{d-1} f_{1,n}(1)^{c-d}.$$

(ii) Given $(D, C)$, the random trees $T_1, \ldots, T_C$ are independent and satisfy that

(a) $T_i, 1 \leq i < D$, are DBTVEs $v_1 = \{f_2, f_3, \ldots \}$ conditioned on $\{\tau_0(T_i) \leq n - 1\}$.

(b) $T_D$ is a DBTVE $v_1 = \{f_2, f_3, \ldots \}$ conditioned on $\{\tau_a(T_D) > n - 1\}$.

(c) $T_i, D < i \leq C$, are DBTVEs $v_1 = \{f_2, f_3, \ldots \}$ conditioned on $\{\tau_\Delta(T_i) > n - 1\}$.

Then, $T$ is a DBTVE $v$ conditioned on $\{\tau_a > n\}$, that is, to have height at least $n$. 

15
Given a defective tree $t$ with height $h(t) \geq l$, then we call distinguished path or spine of length $l$ to a sequence $d_1, \ldots, d_l \in \mathbb{N}$ such that any string of length $k \leq l$ satisfies $d_1 \ldots d_k \in t$. Analogously, a distinguished path of length $n$ in a defective branching tree $T$ results from the numbers $D_1, \ldots, D_n$ and the distinguished individuals $\Lambda_1, \ldots, \Lambda_n$, with $\Lambda_l = D_1 \ldots D_l$, $l = 1, \ldots, n$. It starts with the founding ancestor and it finishes in the distinguished individual $\Lambda_n$ at generation $n$.

The previous lemma allows us to construct a DBTVE conditioned on non-absorption up to height $n$ as indicated below.

**Construction 1:** Let us consider the probability distributions $\{g_{l,n}[d, c]\}_{1 \leq d \leq c \in \mathbb{N}}$, with $l = 1, \ldots, n$ and

$$g_{l,n}[d, c] = \frac{f_{l,n}(1) - f_{l,n}(0)}{f_{l-1,n}(1) - f_{l-1,n}(0)} f_l[c] f_{l,n}(0)^{d-1} f_{l,n}(1)^{c-d}, \quad d = 1, \ldots, c; c \in \mathbb{N}.$$ 

Then, the following procedure provides a DBTVE $v = \{f_n\}_{n \in \mathbb{N}}$ conditioned upon the event $\{\tau_a > n\}$:

1. Establish the distinguished path containing the founding ancestor $\emptyset$ and $n$ distinguished individuals $\Lambda_1, \ldots, \Lambda_n$ from generation 1 to $n$.

2. Generate independent random variables $(D_1, C_1), \ldots, (D_n, C_n)$ with distributions $g_{l,n}, \ldots, g_{n,n}$. Supply the distinguished individual in generation $l$, $\Lambda_l$, with $C_l - 1$ siblings, $D_l - 1$ of them to the left and $C_l - D_l$ to the right.

3. Given these random variables, generate independent defective branching trees $T_{1,l}, \ldots, T_{D_l-1,l}, T_{D_l+1,l}, \ldots, T_{C_l,l}$ in varying environment $v_l = \{f_{l+1,} \ldots\}$ satisfying:
   
   (a) $T_{i,l}$, $1 \leq i \leq D_l$, are DBTVEs conditioned on $\{\tau_0(T_{i,l}) < n - l\}$,
   
   (b) $T_{i,l}$, $D_l < i \leq C_l$, are DBTVEs conditioned on $\{\tau_{\Delta}(T_{i,l}) \geq n - l\}$.

   Attach the trees $T_{1,l}, \ldots, T_{D_l-1,l}$ to the siblings to the left of the distinguished path and $T_{D_l+1,l}, \ldots, T_{C_l,l}$ to the siblings to the right in generation $l$.

4. Complete the tree by adding an independent, unconditioned defective branching tree $T_{D_{n},n}$ in varying environment $v_n = \{f_{n+1,} \ldots\}$ on the top of the distinguished individual in generation $n$.

Then, by this previous construction, an induction argument and Lemma lead to the following result.

**Proposition 2.** Let $\tilde{T}_n$ be the branching tree resulting from the previous construction. Then:

$$\mathcal{L}(\tilde{T}_n) = \mathcal{L}(T|\tau_a(T) > n), \quad n \in \mathbb{N}_0,$$

where $T$ is a DBTVE $v = \{f_n\}_{n \in \mathbb{N}}$. 

16
Appendix

Before providing the proofs of the results of the paper, we derive the expressions (3) and (4). The first one follows immediately from the fact that
\[ E_r Z_n = \cdots = f_{0,n}(1) = \prod_{j=1}^{n} f_j'(f_{j,n}(1)) = \mu_{n,n}. \]

Next, to show the expression in (4) we write
\[ \frac{E[Z_n^2]}{E[Z_n]^2} = \frac{1}{E[Z_n]} + \frac{E[Z_n(Z_n - 1)]}{E[Z_n]^2} = \frac{1}{E[Z_n]} + \frac{f_{0,n}'(1)}{f_{0,n}(1)^2}, \]
and use the fact that
\[ \frac{f_{0,n}'(1)}{f_{0,n}(1)} = \left[ \log(f_{0,n}(1)) \right]' = \left[ \sum_{j=1}^{n} \log(f_j'(f_{j,n}(1))) \right]' = \sum_{j=1}^{n} \frac{f_j''(f_{j,n}(1))}{f_j'(f_{j,n}(1))} \prod_{i=j+1}^{n} f_i'(f_{i,n}(1)) = \sum_{j=1}^{n} \frac{f_j''(f_{j,n}(1))}{f_j'(f_{j,n}(1))} \times \frac{f_{0,n}(1)}{\mu_{j,n}}. \]

Proof of Theorem 1

We use the coupling from formula (6). By Lindvall’s theorem (see [17]) we have that \( \hat{Z}_n \) converges a.s. to a random variable \( \hat{Z}_\infty \) with values in \( \mathbb{N}_0 \cup \{\infty\} \). Thus from (6) it follows that \( Z_n \) has the a.s. limit
\[ Z_\infty = \begin{cases} \hat{Z}_\infty, & \text{if } \tau_\Delta = \infty, \\ \Delta, & \text{if } \tau_\Delta < \infty. \end{cases} \]

Moreover, we have
\[ P[Z_\infty \neq \Delta | \hat{Z}_n, n \in \mathbb{N}_0] = P[\hat{Z}_n = Z_n \text{ for all } n \in \mathbb{N}_0 | \hat{Z}_n, n \in \mathbb{N}_0] = \prod_{n=1}^{\infty} f_n(1) \hat{Z}_{n-1}, \]
and \( \prod_{n=1}^{\infty} f_n(1) \hat{Z}_{n-1} = 0 \) a.s. on the event \( \{0 < \hat{Z}_\infty < \infty\} \) if and only if \( \prod_{n=1}^{\infty} f_n(1) = 0 \). Hence
\[ P[0 < Z_\infty < \infty] = P[Z_\infty \neq \Delta, 0 < \hat{Z}_\infty < \infty] = E[P[Z_\infty \neq \Delta | \hat{Z}_n, n \in \mathbb{N}_0]; 0 < \hat{Z}_\infty < \infty] > 0 \]
if and only if \( \prod_{n=1}^{\infty} f_n(1) > 0 \) and \( P[0 < \hat{Z}_\infty < \infty] > 0 \). On the one hand, it is well known that \( \prod_{n=1}^{\infty} f_n(1) > 0 \) if and only if
\[ \sum_{n=1}^{\infty} (1 - f_n(1)) = \sum_{n=1}^{\infty} \delta_n < \infty. \]
On the other hand, in view of Lindvall’s theorem $P[0 < \tilde{Z}_\infty < \infty] > 0$ if and only
\[
\sum_{n=1}^{\infty} (1 - g_n[1]) = \sum_{n=1}^{\infty} \frac{f_n(1) - f_n[1]}{f_n(1)} < \infty, \tag{12}
\]
where recall that $g_n(s) = f_n(s)/f_n(1)$. Now, (11) implies $f_n(1) \to 1$, and therefore we obtain that (11) and (12) is equivalent to
\[
\sum_{n=1}^{\infty} (1 - f_n[1]) = \sum_{n=1}^{\infty} \delta_n + \sum_{n=1}^{\infty} (f_n(1) - f_n[1]) < \infty.
\]
This is our claim.

\[\square\]

**Proof of Theorem 2**

Throughout this proof all the limits are taken as $n \to \infty$ unless specified otherwise.

First of all, by coupling in (6), one sees that $\tau n \in N_0$ and $\tilde{Z}_n \in N_0$ visit the same states until the process $\{Z_n\}_{n \in N_0}$ goes to the graveyard state.

Moreover, $\{\tilde{W}_n\}_{n \in N_0}$, with $\tilde{W}_n = \tilde{Z}_n\tilde{\mu}_n^{-1}$, and $\tilde{\mu}_n = \prod_{i=1}^{n} f_i(1)f_i(1)^{-1}$ is a non-negative martingale and hence $\tilde{W}_n$ converges a.s. to an integrable random variable $\tilde{W} \geq 0$. By applying Theorem 2 in [15] one has that under condition (9), $P[\tilde{Z}_n \to 0] = P[\tilde{W} = 0]$. From the proof of Theorem 1 we have
\[
P[Z_n = \tilde{Z}_n \text{ for all } n \in N_0 | \tilde{Z}_n, n \in N_0] = \prod_{n=1}^{\infty} f_n(1)\tilde{Z}_n^{-1} \text{ a.s.},
\]
and then,
\[
P[Z_\infty \neq \Delta] = E \left[ \prod_{n=1}^{\infty} f_n(1)\tilde{Z}_n^{-1} \right]
= E \left[ \prod_{n=1}^{\infty} f_n(1)\tilde{Z}_n^{-1} ; \tilde{Z}_n \to 0 \right] + E \left[ \prod_{n=1}^{\infty} f_n(1)\tilde{Z}_n^{-1} ; \tilde{W} > 0 \right]
= P[Z_n \to 0] + E \left[ \prod_{n=1}^{\infty} f_n(1)\tilde{Z}_n^{-1} \mid \tilde{W} > 0 \right] P[\tilde{W} > 0],
\]
As a result,
\[
P[\tau_a = \infty] = P[Z_\infty \neq \Delta] - P[Z_\infty = 0] = E \left[ \exp \left( \sum_{n=1}^{\infty} \tilde{Z}_n^{-1} \log f_n(1) \right) \mid \tilde{W} > 0 \right] P[\tilde{W} > 0].
\]
Thus to have extinction or absorption in $\Delta$ with positive probability, both right-hand terms have to be positive.
First let us consider the right-hand conditional expectation. Rewriting it as
\[ E \left[ \exp \left( \sum_{n=1}^{\infty} \frac{\tilde{Z}_{n-1}}{\tilde{\mu}_{n-1}} \log f_n(1) \right) \mid \tilde{W} > 0 \right] \]
and remembering that \( \tilde{Z}_n/\tilde{\mu}_n \) converges to \( \tilde{W} \) a.s. we see that this expectation is positive if and only if
\[ \sum_{n=1}^{\infty} \tilde{\mu}_{n-1} \log f_n(1) > -\infty. \]

Second, again due to \( P[\tilde{W} > 0] = 1 - P[\tilde{Z}_n \to 0] \), from Theorem 1 (iv) in [15] it follows that \( P[\tilde{W} > 0] > 0 \) if and only if
\[ \sum_{n=1}^{\infty} \frac{g_n''(1)}{g_n'(1)^2 \tilde{\mu}_{n-1}} < \infty, \quad \text{and} \quad \limsup_{n \to \infty} \tilde{\mu}_n > 0, \]
which from Theorem 1 (viii) in [15] is also equivalent to
\[ \sum_{n=1}^{\infty} \frac{g_n''(1)}{g_n'(1)^2 \tilde{\mu}_{n-1}} < \infty \quad \text{and} \quad \tilde{\mu}_n \to r; \]
for some \( 0 < r \leq \infty \). This implies that \( P[\tilde{W} > 0] > 0 \) is as well equivalent to the condition
\[ \sum_{n=1}^{\infty} \frac{g_n''(1)}{g_n'(1) \tilde{\mu}_n} < \infty \quad \text{and} \quad \inf_{n \geq 1} \tilde{\mu}_n > 0 \]
convenient for our purpose.

Altogether we see that \( P[\tau_n = \infty] > 0 \) holds if and only if
\[ \inf_{n \geq 1} \tilde{\mu}_n > 0, \quad \sum_{n=1}^{\infty} \tilde{\mu}_{n-1} \log f_n(1) > -\infty \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{g_n''(1)}{g_n'(1) \tilde{\mu}_n} < \infty. \]

Concluding the proof we rephrase this condition. First the two left-hand conditions imply that \( \log f_n(1) \to 0 \). Therefore we have \( -\log f_n(1) \sim 1 - f_n(1) = \delta_n \) (where \( a_n \sim b_n \) means that there exist \( c, C > 0 \) such that \( ca_n \leq b_n \leq Ca_n \)), leading to the equivalent set of conditions
\[ \inf_{n \geq 1} \tilde{\mu}_n > 0, \quad \sum_{n=1}^{\infty} \delta_n \tilde{\mu}_{n-1} < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{g_n''(1)}{g_n'(1) \tilde{\mu}_n} < \infty. \]

Second, these conditions entail \( \sum_{n=1}^{\infty} \delta_n < \infty \) which means that \( \prod_{i=1}^{n} (1 - \delta_n) = \prod_{i=1}^{n} f_i(1) \) converges to a positive limit. Because of \( \tilde{\mu}_n = \mu_n/\prod_{i=1}^{n} f_i(1) \) we end up with the conditions
\[ \inf_{n \geq 1} \mu_n > 0, \quad \sum_{n=1}^{\infty} \delta_n \mu_{n-1} < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{f_n''(1)}{f_n'(1) \mu_n} < \infty, \]
which in turn by the same reasoning entail the former sets of conditions. This finishes the proof.
Proof of Proposition 1

The proof of the first inequality is as follows. For each \( k \leq n \) we have

\[
P[\tau_a > n] \leq P[Z_k \neq 0, Z_k \neq \Delta] \leq E[Z_k] = \prod_{j=1}^{k} f_j(f_j,k(1)) \leq \mu_k.
\]

For the second part, for each \( t \in (0, 1) \) and \( n \in \mathbb{N} \), let us denote as \( \xi_{nt} \) the random variable with the distribution \( g_{nt}(u) = f_n(ut)/f_n(t), \ u \in [0, 1] \). Then, under condition (10), it satisfies the following condition: there exists \( c > 0 \) such that

\[
E[\xi_{nt}^2; \xi_{nt} \geq 2] \leq cE[\xi_{nt}; \xi_{nt} \geq 2] E[\xi_{nt}; \xi_{nt} \geq 1].\tag{13}
\]

Indeed, if we consider the probability measures \( h_t \) with weights

\[
h_t[0] = h_t[1] = 0, \quad h_t[k] = \frac{\sum_{j=2}^{\infty} j f_n[j]}{\sum_{j=2}^{\infty} j f_n[j] t^j}, \quad k \geq 2,
\]

and apply Lemma 3 in [15] for \( h_{t_1} \) and \( h_{t_2} \), with \( t_1 < t_2 \), we obtain that

\[
\sum_{k=2}^{\infty} k h_t[k] = \frac{E[\xi_{nt}^2; \xi_{nt} \geq 2]}{E[\xi_{nt}; \xi_{nt} \geq 2]}
\]

is an increasing function in \( t \). Thus,

\[
E[\xi_{nt}^2; \xi_{nt} \geq 2] \leq cE[\xi_{nt}; \xi_{nt} \geq 2] \leq cE[\xi_{nt}; \xi_{nt} \geq 2] E[\xi_{nt}; \xi_{nt} \geq 1],
\]

with the last factor being greater than one.

Now, from (13) and Lemma 1 in [15], one has for each \( t \in (0, 1) \) and \( n \in \mathbb{N} \)

\[
\frac{1}{1 - g_{nt}(u)} = \frac{1}{g_{nt}(1)(1-u)} + \varphi_{nt}(u), \quad u \in [0, 1],
\]

for some function \( \varphi_{nt} (\cdot) \) satisfying

\[
\frac{f_n''(t)f_n(t)}{2c f_n'(t)^2} \leq \varphi_{nt}(0) = \frac{f_n''(t)f_n(t)}{2f_n'(t)^2} \leq \varphi_{nt}(1) = \frac{f_n''(t)f_n(t)}{2f_n'(t)^2}, \quad u \in [0, 1],
\]

and consequently, for \( 0 < s < t < 1 \),

\[
\frac{1}{f_n(t) - f_n(s)} = \frac{1}{f_n'(t)(t-s)} + \frac{\varphi_{nt}(s/t)}{f_n(t)}, \tag{15}
\]

with

\[
\frac{f_n''(t)}{2c f_n'(t)^2} \leq \frac{\varphi_{nt}(s/t)}{f_n(t)} \leq \frac{f_n''(t)}{f_n'(t)^2}, \tag{16}
\]

20
By iterating the formula (15), one obtains

\[
\frac{1}{P[\tau_a > n]} = \frac{1}{f_{0,n}(1) - f_{0,n}(0)} = \prod_{j=1}^{n} \frac{f_j'(f_{j,n}(1))}{f_j'(f_{j,n}(1)) + \sum_{j=1}^{n} \varphi_{j,f_j,n}(1) (f_{j,n}(0)/f_{j,n}(1))}
\]

and from (16),

\[
\prod_{j=1}^{n} \frac{f_j'(f_{j,n}(1))}{f_j'(f_{j,n}(1)) + \frac{1}{2c} \sum_{j=1}^{n} \frac{f_j''(f_{j,n}(1))}{\prod_{i=1}^{j} f_i'(f_{i,n}(1))}} \leq \frac{1}{P[\tau_a > n]} \leq \prod_{j=1}^{n} \frac{f_j'(f_{j,n}(1))}{f_j'(f_{j,n}(1)) + \sum_{j=1}^{n} \frac{f_j''(f_{j,n}(1))}{\prod_{i=1}^{j} f_i'(f_{i,n}(1))}}.
\]

(17)

Now, the proof finishes by using the formulas in (3) and (4).

\[\square\]

**Proof of Theorem 3**

First note that by the Cauchy-Schwartz inequality and by (10) one has

\[E[X_{n1}; X_{n1} \geq 2]^2 \leq E[X_{n1}^2; X_{n1} \geq 2] \leq c E[X_{n1}; X_{n1} \geq 2],\]

and therefore \(E[X_{n1}; X_{n1} \geq 2] \leq c\). Consequently

\[f_n'(1) \leq c + 1 \quad \text{and} \quad f_n''(1) \leq E[X_{n1}^2; X_{n1} \geq 2] \leq c E[X_{n1}; X_{n1} \geq 2] \leq c^2.\]

(18)

Also, by means of the probability measure \(h_t\) from (14) (with \(n\) replaced with \(i\) and \(t = f_{i,n}(1)\))

\[
\frac{f_i''(t)}{f_i'(t)} \leq \frac{\sum_{k=2}^{\infty} k^2 t^{-k} f_i[k]}{\sum_{k=2}^{\infty} k t^{-k} f_i[k]} \leq \frac{1}{t} \sum_{k=2}^{\infty} k h_t[k] = \frac{1}{t} \sum_{k=2}^{\infty} k h_1[k] = \frac{1}{t} E[X_{i1}^2; X_{i1} \geq 2] \leq c.
\]

(19)

where we again used that \(\sum_{k=2}^{\infty} k h_t[k]\) is increasing in \(t\).

Coming to our first claim note that under the condition \(f_n(\rho) \geq \rho\) for \(n \geq n_0\) we have for \(n_0 \leq i \leq n\)

\[f_{i,n}(1) \geq f_{i,n}(\rho) \geq \rho.\]

(20)

The proof proceeds via induction on \(i\). For \(i = n\) this follows from \(f_{n,n}(s) = s\), and if the claim holds true for \(i > n_0\), then

\[f_{i-1,n}(1) \geq f_{i-1,n}(\rho) = f_i(f_{i,n}(\rho)) \geq f_i(\rho) \geq \rho.\]

Additionally, for \(0 \leq i \leq n_0\)

\[f_{i,n}(1) = f_{i,n_0}(f_{n_0,n}(1)) \geq f_{i,n_0}(\rho) > 0.\]

(21)

For \(n_0 \leq j \leq n\) we obtain

\[
\mu_{j,n} = \prod_{i=1}^{j} f_i'(f_{i,n}(1)) \geq \prod_{i=1}^{n_0} f_i'(f_{i,n_0}(\rho)) \prod_{i=n_0+1}^{j} f_i'(\rho).
\]

21
This implies that there is a constant $c_1 > 0$ such that
\[ \mu_{j,n} \geq c_1 \mu_j(\rho). \]
(For $j = n$ this is an estimate for $E[Z_n]$). With $c_1$ sufficiently small this holds for $n \leq n_0$, too. Using this estimate together with Proposition 1 and the formulas (17), (19), (20), (21) yields
\[
\frac{1}{P[\tau_a > n]} \leq \frac{1}{E[Z_n]} + \sum_{j=1}^{n} \frac{f''_j(f_{j,n}(1))}{f'_j(f_{j,n}(1))} \mu_{j,n}
\leq \frac{1}{E[Z_n]} + \sum_{j=1}^{n} \frac{c}{f_{j,n}(1)} \mu_{j,n}
\leq \frac{1}{c_1 \mu_n(\rho)} + \sum_{j=1}^{n_0} \frac{c}{f_{j,n_0}(\rho) c_1 \mu_j(\rho)} + \sum_{j=n_0+1}^{n} \frac{c}{\rho c_1 \mu_j(\rho)}
\leq \nu(\rho) \left( \frac{1}{c_1} + \sum_{j=1}^{n_0} \frac{c}{f_{j,n_0}(\rho) c_1} + \frac{c}{\rho c_1} \right).
\]
These estimates give the first part of the theorem.

As to the second part we notice that by our assumptions and by convexity the generating functions $f_n(s)$ take their values for $n \geq n_0$ and $s \geq \sigma$ below the straight line connecting the points $(\sigma, \sigma)$ and $(1, \zeta)$ in the plane, with $\zeta = \sup_{n \geq n_0} f_n(1)$. Since by assumption $\zeta < 1$, it follows that for any $\varepsilon > 0$ there is a natural number $m = m_\varepsilon$ such that $f_n(s) \leq s - m_\varepsilon^{-1}$ for $s \geq \sigma + \varepsilon$. Also, $f_n(s) \leq \sigma + \varepsilon$ for $s \leq \sigma + \varepsilon$. Therefore, iterating these estimates, if $n \geq n_0$ and $n - i \geq m$, then
\[ f_{i,n}(1) \leq \sigma + \varepsilon. \]
This implies for $n \geq j > m + n_0$
\[ \mu_{j,n} = \prod_{i=1}^{j} f'_i(f_{i,n}(1)) \leq \prod_{i=1}^{n_0} f'_i(1) \prod_{i=n_0+1}^{j-m} f'_i(\sigma + \varepsilon) \prod_{i=j-m+1}^{j} f'_i(1). \]
Furthermore, from (19) we have $f''_i(\rho) \leq c f'_i(\rho) / \rho$, hence by assumption $\inf_{i \geq 1} f''_i(\rho) > 0$ also $\inf_{i \geq 1} f'_i(\rho) > 0$ and consequently $\inf_{i \geq 1} f'_i(\sigma + \varepsilon) > 0$. Taking also (18) into account it follows that there is a constant $c_2 > 0$ such that
\[ \mu_{j,n} \leq c_2 \mu_j(\sigma + \varepsilon). \] (22)
Again, by further enlarging $c_2$, this estimate is valid for all $n \geq 1$. Next, from (17) and (20), for $n \geq n_0$
\[ \frac{1}{P[\tau_a > n]} \geq \frac{1}{E[Z_n]} + \frac{1}{2c} \sum_{j=n_0}^{n} \frac{f''_j(\rho)}{f'_j(1)} \mu_{j,n}, \]
Using the assumption $\inf_{n \geq 1} f''(\rho) > 0$, (18) and (22) we obtain
\[ \frac{1}{P[\tau_a > n]} \geq c_3 \sum_{j=n_0}^{n} \frac{1}{\mu_j(\sigma + \varepsilon)} \]
for some constant $c_3 > 0$. This estimate implies the assertion of the theorem, and the proof is finished.

\[ \square \]

**Proof of Lemma 1**

We follow the steps in the proof of Lemma 1.1 in [16], thereby correcting a notation error of that proof.

First, we check that $\{g_n[d, c]\}_{1 \leq d \leq c; c \in \mathbb{N}}$ is indeed a non-defective (proper) probability distribution.

\[
\sum_{1 \leq d \leq c < \infty} g_n[d, c] = \frac{1}{f_{0,n}(1) - f_{0,n}(0)} \left[ \sum_{c=1}^{\infty} f_1[c] f_{1,n}(1)^c - \sum_{c=1}^{\infty} f_1[c] f_{1,n}(0)^c \right] = \frac{f_1(f_{1,n}(1)) - f_1(f_{1,n}(0))}{f_{0,n}(1) - f_{0,n}(0)} = 1.
\]

Let $t$ be a tree of height at least $n$ and $c = c(\emptyset)$; we divide the tree $t$ into $c$ subtrees $t_1, \ldots, t_c$. Analogously, we divide the defective branching tree $T$ into subtrees $T_1, \ldots, T_c$. Thus, if $h' = h - 1$, then

\[
\{T \equiv t\} = \{C = c, T_1 \equiv t_1, \ldots, T_c \equiv t_c\},
\]

and bearing in mind the independence in the reproduction among individuals

\[
P_v \left[ T \equiv t \right] = f_1[c] \prod_{j=1}^{c} P_{v_1} \left[ T \equiv t_j \right],
\]

since $P_v \left[ T_j \equiv t_j \right] = P_{v_1} \left[ T_j \equiv t_j \right]$, for all $1 \leq j \leq c$. Now, note that if the tree $t$ has height at least $n$, then there is a distinguished individual $d$ in generation one, $d \in \{1, \ldots, c\}$, such that $t_d$ has height at least $n - 1$, but the trees $t_j$ have heights less than $n - 1$ for $j \in \{1, \ldots, d - 1\}$, and the trees $t_j$, for $j \in \{d + 1, \ldots, c\}$ may have height less than, equal to or greater than $n - 1$. As a result

\[
P_v \left[ T \equiv t \right] = f_1[c] \prod_{j=1}^{d-1} P_{v_1} \left[ T \equiv t_j \right] P_{v_1} \left[ T \equiv t_d \right] \prod_{j=d+1}^{c} P_{v_1} \left[ T \equiv t_j \right].
\]

For $h \geq n$ this equation may be rewritten as

\[
P_v \left[ T \equiv t, \tau_a > n \right] = f_1[c] \prod_{j=1}^{d-1} P_{v_1} \left[ T \equiv t_j, \tau_0 \leq n - 1 \right] \times P_{v_1} \left[ T \equiv t_d, \tau_a > n - 1 \right] \prod_{j=d+1}^{c} P_{v_1} \left[ T \equiv t_j, \tau_\Delta > n - 1 \right].
\]
or equivalently
\[
P_v \left[ T = t \mid \tau_a > n \right] = g_n[d, c] \prod_{j=1}^{d-1} P_{v_1} \left[ T = t_j \mid \tau_0 \leq n - 1 \right]
\]
\[
\times P_{v_1} \left[ T = t_d \mid \tau_a > n - 1 \right] \prod_{j=d+1}^{c} P_{v_1} \left[ T = t_j \mid \tau_D > n - 1 \right].
\]

This formula implies the lemma’s assertion.

\[\square\]

**Proof of Theorem 4**

First, for each \( n \in \mathbb{N} \), by the construction we have

\[
E[Z_n \mid \tau_a > n] = 1 + \sum_{l=1}^{n} E[C_l - D_l] \cdot E_{v_1} [Z_{n-l}(T_{D_l+1}) \mid \tau_D > n - l]
\]
\[
= 1 + \sum_{l=1}^{n} E[C_l - D_l] \cdot \frac{E_{v_1} [Z_{n-l}I_{(\tau_D > n-l)}]}{P_{v_1} [\tau_D > n - l]}.
\]

On the one hand,

\[
E_{v_1} [Z_{n-l}I_{(\tau_D > n-l)}] = E_{v_1} \left[ \tilde{Z}_{n-l} \prod_{i=1}^{n-l} f_{i+l}(1) \tilde{z}_{i-1} \right]
\]
\[
= E_{v_1} \left[ E_{v_1} \left[ \tilde{Z}_{n-l} \prod_{i=1}^{n-l} f_{i+l}(1) \tilde{z}_{i-1} \mid \tilde{Z}_{n-l-1} \right] \right]
\]
\[
= E_{v_1} \left[ E_{v_1} \left[ \tilde{Z}_{n-l} \mid \tilde{Z}_{n-l-1} \right] \prod_{i=1}^{n-l} f_{i+l}(1) \tilde{z}_{i-1} \right]
\]
\[
= E_{v_1} \left[ \tilde{Z}_{n-l-1} \frac{f''(1)}{f(1)} \prod_{i=1}^{n-l} f_{i+l}(1) \tilde{z}_{i-1} \right]
\]
\[
\leq \prod_{i=1}^{n-l-1} f_{i+l}(1) f''(1) E_{v_1} \left[ \tilde{Z}_{n-l-1}f_n(1) \tilde{z}_{n-l-1-1} \right]
\]
\[
\leq \beta^{n-l-1} f''(1) E_{v_1} \left[ \tilde{Z}_{n-l-1}f_n(1) \tilde{z}_{n-l-1-1} \right],
\]

and since the function \( x \mapsto xf_n(1)x^{-1} \) has a maximum at \( -1/\log(f_n(1)) \), then

\[
E_{v_1} \left[ \tilde{Z}_{n-l-1}f_n(1) \tilde{z}_{n-l-1-1} \right] \leq \frac{1}{f_n(1)^{1+ \log(f_n(1))} \log \left( \frac{1}{f_n(1)} \right)}.
\]
On the other hand,

\[ P_{\ell}[\tau_{\Delta} > n - l] \geq P_{\ell}[\tau_{\Delta} = \infty] \geq P_{\ell}[\tau_{0} = 1] = f_{1+1}(0) \geq \alpha > 0. \]

Now, we use the inequalities \( E[C_{l} - D_{l}] \leq E[C_{l}], \ l = 1, \ldots, n \). To compute the last expectation, we first determine the distribution of \( C_{l} \), which is given by

\[ P[C_{l} = c] = \sum_{d=1}^{c} g_{l,n}[d, c] = \frac{f_{l,n}(1) - f_{l,n}(0)}{f_{l-1,n}(1) - f_{l-1,n}(0)} f_{l}[c] \sum_{d=1}^{c} f_{l,n}(0)^{d-1} f_{l,n}(1)^{c-d} \]

\[ = f_{l}[c] \frac{f_{l,n}(1)c - f_{l,n}(0)c}{f_{l-1,n}(1) - f_{l-1,n}(0)}, \ c \in \mathbb{N}. \]

Therefore, for each \( l = 1, \ldots, n \), and \( c \in \mathbb{N} \), using the mean value theorem, we have that for some \( \xi \) between \( f_{l,n}(0)c \) and \( f_{l,n}(1)c \)

\[ P[C_{l} = c + k] = \frac{f_{l}[c + k]}{f_{l}[c]} \cdot \frac{f_{l,n}(1)c^{k} - f_{l,n}(0)c^{k}}{f_{l,n}(1)c - f_{l,n}(0)c} \]

\[ = f_{l}[c + k] \cdot \frac{\left( f_{l,n}(1)c \right)^{1+k/c} - \left( f_{l,n}(0)c \right)^{1+k/c}}{f_{l,n}(1)c - f_{l,n}(0)c} \]

\[ = f_{l}[c + k] \cdot \left( 1 + \frac{k}{c} \right)^{c^{k/c}} \]

\[ \leq \frac{(c + k)f_{l}[c + k]/f_{l}'(1)}{cf_{l}[c]/f_{l}'(1)}. \]

Since \( \{cf_{l}[c]/f_{l}'(1)\}_{c \in \mathbb{N}} \) are the weights of a probability distribution, we apply [15, Lemma 3] and obtain

\[ E[C_{l}] = \sum_{c=1}^{\infty} cP[C_{l} = c] \leq \sum_{c=1}^{\infty} c^{2} \frac{f_{l}[c]}{f_{l}'(1)} = 1 + \sum_{c=1}^{\infty} c(c - 1) \frac{f_{l}[c]}{f_{l}'(1)} = 1 + \frac{f_{l}''(1)}{f_{l}'(1)}. \]

Finally, combining all the above we get

\[ E[Z_{n} | \tau_{a} > n] \leq 1 + \frac{f_{n}(1)}{\alpha f_{n}(1)^{1+1/logf_{n}(1)}} \log \left( \frac{1}{f_{n}(1)} \right) \sum_{l=1}^{n} \left( 1 + \frac{f_{l}''(1)}{f_{l}'(1)} \right) \beta^{n-l-1}. \]

Using \( 0 < \alpha < f_{n}(1) < \beta < 1 \) and \( f_{n}(1)^{1/logf_{n}(1)} = e \) (the Euler constant) we see that the right-hand denominator is bounded away from 0, and our first claim follows with

\[ c = \frac{1}{e \alpha^{2} \beta \log \beta^{-1}}. \]

Finally, by [13] and with the same arguments as those in the proof of (19) we obtain the second part.
Acknowledgements

This manuscript was partially prepared while Carmen Minuesa was visiting the Institute of Mathematics, Goethe University Frankfurt, in Frankfurt am Main, and she is grateful for the hospitality and collaboration.

The authors would like to thank Serik Sagitov (Chalmers University of Technology and University of Gothenburg) for suggesting the topic of this research.

The authors also thank the anonymous referees for their valuable comments.

Funding

Carmen Minuesa’s research has been supported by the Ministerio de Economía y Competitividad (grant MTM2015-70522-P), Ministerio de Ciencia e Innovación (grant PID2019-108211GB-I00), the Junta de Extremadura and the European Regional Development Fund (grants IB16099 and GR18103).

References

[1] A. Agresti. On the extinction times of varying and random environment branching processes. *Journal of Applied Probability*, 12(1):39–46, 1975.

[2] P. Braunsteins and S. Hautphenne. Extinction in lower Hessenberg branching processes with countably many types. *Annals of Applied Probability*, 29(5):2782–2818, 2019.

[3] J. C. D'Souza and J. D. Biggins. The supercritical Galton–Watson processes in varying environment. *Stochastic Processes and their Applications*, 42:39–47, 1992.

[4] J. H. Foster and R. T. Goettge. The rates of growth of the Galton-Watson process in varying environment. *Journal of Applied Probability*, 13(1):144–147, 1976.

[5] T. Fujimagari. On the extinction time distribution of a branching process in varying environments. *Advances in Applied Probability*, 12(2):350–366, 1980.

[6] J. Geiger. Elementary new proofs of classical limit theorems for Galton-Watson processes. *Journal of Applied Probability*, 36(2):301–309, 1999.

[7] R. T. Goettge. Limit theorems for the supercritical Galton-Watson process in varying environments. *Mathematical Biosciences*, 28(1):171 – 190, 1976.

[8] P. L. Hsu and H. Robbins. Complete convergence and the law of large numbers. In *Proceedings of the National Academy of Sciences of the United States of America*, volume 33, pages 25–31, 1947.
[9] Y. Hu, W. Hu, and Y. Yin. The extinction of a branching process in a varying or random environment. In S. Li, X. Wang, Y. Okazaki, J. Kawabe, T. Murofushi, and L. Guan, editors, *Nonlinear Mathematics for Uncertainty and its Applications. Advances in Intelligent and Soft Computing*, volume 100, pages 309–315. Springer, Berlin, Heidelberg, 2011.

[10] M. Ispány. Some asymptotic results for strongly critical branching processes with immigration in varying environment. In I. del Puerto, M. González, C. Gutiérrez, R. Martínez, C. Minuesa, M. Molina, M. Mota, and A. Ramos, editors, *Branching Processes and Their Applications*, volume 219 of *Lecture Notes in Statistics*, pages 77–95. Springer, 2016.

[11] Y. Iwasa, F. Michor, and M. A. Nowak. Evolutionary dynamics of invasion and escape. *Journal of Theoretical Biology*, 226(2):205–214, 2004.

[12] P. Jagers. Galton-Watson processes in varying environments. *Journal of Applied Probability*, 11(1):174–178, 1974.

[13] S. Karlin and S. Tavaré. Linear birth and death processes with killing. *Journal of Applied Probability*, 19(3):477–487, 1982.

[14] N. Keiding and J. E. Nielsen. Branching processes with varying and random geometric offspring distributions. *Journal of Applied Probability*, 12(1):135–141, 1975.

[15] G. Kersting. A unifying approach to branching processes in a varying environment. *Journal of Applied Probability*, 57(1):196–220, 2020.

[16] G. Kersting and V. Vatutin. *Discrete Time Branching Processes in Random Environment*. ISTE Ltd and John Wiley and Sons, Inc., 2017.

[17] T. Lindvall. Almost sure convergence of branching processes in varying and random environments. *The Annals of Probability*, 2(2):344–346, 1974.

[18] I. M. MacPhee and H. J. Schuh. A Galton-Watson branching process in varying environments with essentially constant offspring means and two rates of growth. *Australian Journal of Statistics*, 25(2):329–338, 1983.

[19] A. Robertson. The time of detection of recessive visible genes in small populations. *Genetics Research*, 31:255–264, 1978.

[20] S. Sagitov and C. Minuesa. Defective Galton–Watson processes. *Stochastic Models*, 33(3):451–472, 2017.

[21] J. Yu and J. Pei. Extinction of branching processes in varying environments. *Statistics and Probability Letters*, 79(17):1872 – 1877, 2009.