YAGLOM’S LIMIT FOR CRITICAL GALTON-WATSON PROCESSES
IN VARYING ENVIRONMENT: A PROBABILISTIC APPROACH

NATALIA CARDONA-TOBÓN AND SANDRA PALAU

Abstract. A Galton-Watson process in varying environment is a discrete time branching process where the offspring distributions vary among generations. Based on a two-spine decomposition technique, we provide a probabilistic argument of a Yaglom-type limit for this family processes. The result states that, in the critical case, a suitable normalisation of the process conditioned on non-extinction converges in distribution to an exponential random variable. Recently, this result has been established by Kersting [J. Appl. Probab. 57(1), 196–220, 2020] using analytic techniques.

Key words and phrases: Galton-Watson processes; varying environment; Yaglom’s limit; spines decompositions.

MSC 2010: 60J80; 60F05; 60K37.

1. Introduction

A Galton-Watson process in varying environment (GWVE) is a discrete time branching process where the offspring distributions vary among generations, in other words individuals give birth independently and their offspring distributions coincide within each generation. More precisely, a varying environment is a sequence $Q = (q_1, q_2, \ldots)$ of probability measures on $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. A Galton-Watson process $Z^Q = \{Z^Q_n : n \geq 0\}$ in a varying environment $Q$ is a Markov chain defined recursively as follows

$$Z^Q_0 = 1 \quad \text{and} \quad Z^Q_n = \sum_{i=1}^{Z^Q_{n-1}} Y^{(n)}_i, \quad n \geq 1,$$

where $\{Y^{(n)}_i : i, n \geq 1\}$ is a sequence of independent random variables such that

$$\mathbb{P}(Y^{(n)}_i = k) = q_n(k), \quad k \in \mathbb{N}_0, \quad i, \ n \geq 1.$$

The variable $Y^{(n)}_i$ denotes the offspring of the $i$-th individual in the $(n-1)$-th generation. Its generating function is given by

$$f_n(s) := \mathbb{E}\left[s^{Y^{(n)}_i}\right] = \sum_{k=0}^{\infty} s^k q_n(k), \quad 0 \leq s \leq 1, \ n \geq 1.$$

Hence, by applying the branching property recursively, we deduce that the generating function of $Z^Q_n$ is given in terms of $(f_1, f_2, \ldots)$ as follows

$$(1) \quad \mathbb{E}\left[s^{Z^Q_n}\right] = f_1 \circ \cdots \circ f_n(s), \quad 0 \leq s \leq 1, \ n \geq 1,$$

where $f \circ g$ denotes the composition of $f$ with $g$. 

arXiv:2005.10186v1 [math.PR] 20 May 2020
Moreover, by differentiating in \( s \), we obtain
\[
(2) \quad \mathbb{E}[Z_n^Q] = \mu_n, \quad \text{and} \quad \mathbb{E}[Z_n^Q(Z_n^Q - 1)] = \mu_n^2 \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k}, \quad n \geq 1,
\]
where \( \mu_0 := 1 \) and for any \( n \geq 1, \)
\[
(3) \quad \mu_n := f'_1(1) \cdots f'_n(1), \quad \text{and} \quad \nu_n := \frac{f''_n(1)}{f'_n(1)^2} = \frac{\text{Var}[Y_i^{(n)}]}{\mathbb{E}[Y_i^{(n)}]^2} + \left( 1 - \frac{1}{\mathbb{E}[Y_i^{(n)}]} \right),
\]
where \( \text{Var}[Y_i^{(n)}] \) is the variance of the variable. For further details about GWVEs, we refer to the monograph of Kersting and Vatutin [7].

According with Kersting, [6], we say that a GWVE is \textit{regular} if there exists a constant \( c > 0 \) such that for all \( n \geq 1, \)
\[
\mathbb{E}\left[ (Y_i^{(n)})^2 1_{\{Y_i^{(n)} \geq 2\}} \right] \leq c \mathbb{E}\left[ Y_i^{(n)} 1_{\{Y_i^{(n)} \geq 2\}} \right] \mathbb{E}\left[ Y_i^{(n)} 1_{\{Y_i^{(n)} \geq 1\}} \right].
\]
He proved that a regular GWVE has extinction a.s (i.e. \( \mathbb{P}(Z_n^Q = 0 \text{ for some } n) = 1 \)) if and only if \( \sum_{k=0}^{\infty} \frac{\nu_{k+1}}{\mu_k} = \infty \) or \( \mu_n \to 0 \) as \( n \to \infty \), [6] Theorem 1]. In addition, he gave the following classification.

A regular GWVE is

i. \textit{supercritical} if and only if \( \sum_{k=0}^{\infty} \frac{\nu_{k+1}}{\mu_k} < \infty \) and \( \lim_{n \to \infty} \mu_n = \infty, \)

ii. \textit{asymptotically degenerate} if and only if \( \sum_{k=0}^{\infty} \frac{\nu_{k+1}}{\mu_k} < \infty \) and \( 0 < \lim_{n \to \infty} \mu_n < \infty, \)

iii. \textit{critical} if and only if \( \sum_{k=0}^{\infty} \frac{\nu_{k+1}}{\mu_k} = \infty \) and \( \lim_{n \to \infty} \mu_n \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} = \infty, \)

iv. \textit{subcritical} if and only if \( \lim_{n \to \infty} \mu_n = 0 \) and \( \liminf_{n \to \infty} \mu_n \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} < \infty. \)

Kersting’s definition is an extension of the classical categorisation of branching processes. Indeed, when the environment is constant, we have \( \mu_k = \mu^k \) and \( \nu_k = \sigma^2 \), for \( k \geq 1, \) where \( \mu \) and \( \sigma^2 \) are the mean and normalised second factorial moment of the offspring distribution, respectively; we recover the original classification. We observe that in this case, the asymptotically degenerate case is not possible.

Given a varying environment \( Q \), we define the sequence \( \{a_n^Q : n \geq 0\} \) as follows
\[
a_0^Q = 1, \quad \text{and} \quad a_n^Q = \frac{\mu_n}{2} \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k}, \quad n \geq 1.
\]
Kersting, [6] Theorem 4], showed that in the critical regime, \( a_n^Q \to \infty \) and that
\[
(4) \quad \lim_{n \to \infty} \frac{a_n^Q}{\mu_n} \mathbb{P}(Z_n^Q > 0) = 1.
\]

In the rest of the paper, we work with regular critical GWVE. Moreover, we assume the following condition
\[
(\text{A}) \quad \text{there exists } c > 0 \text{ such that } f''_n(1) \leq c f''_n(1)(1 + f'_n(1)), \quad \text{for any } n \geq 1.
\]
Kersting proved that this condition implies that the GWVE is regular, see [6, Proposition 2]. Moreover, he explained that condition (A) is a rather mild condition. Indeed, it is satisfied by most common probability distributions, for instance the Poisson, binomial, geometric, hypergeometric, and negative binomial distributions. Another important example satisfying condition (A) are random variables that are a.s. uniformly bounded by a constant.

We are ready to present our main result, which is in accordance with Yaglom's theorem for classical Galton-Watson processes.

**Theorem 1 (Yaglom’s limit).** Let \( \{Z_n^Q : n \geq 0\} \) be a critical GWVE that satisfies condition (A). Then

\[ \left( \frac{Z_n^Q}{a_n^Q}; \mathbb{P}(\cdot | Z_n^Q > 0) \right) \xrightarrow{(d)} (Y; \mathbb{P}), \quad \text{as} \quad n \to \infty, \]

where \( Y \) is a standard exponential random variable.

In the classical theory with constant environment, this result has several proofs, the first one was given by Yaglom [10]. In [8], a probabilistic proof via a characterisation of the exponential distribution was presented. Later on, Geiger characterised the exponential random variable by a distributional equation and he presented another proof of Yaglom’s limit based on that equation (see [2, 3]). Recently, Ren et al. [9], developed yet a new proof using a two-spine decomposition technique.

When the environment is varying, Jagers [5] proved the convergence under extra assumptions. Afterwards, Bhattacharya and Perlman [1] obtained the same result with weaker assumptions than Jagers (but stronger than ours). Kersting [6] provided yet another proof in a similar framework to ours, that we will explain below. An extension in the presence of immigration and the same setting as Kersting’s has been established in [4]. All these authors established the exponential convergence using an analytical approach. The condition in Kersting [6] is the following. For every \( \epsilon > 0 \) there is a constant \( c_\epsilon < \infty \) such that

\[ \mathbb{E} \left[ (Y_i^{(n)})^2 \mathbf{1}_{\{Y_i^{(n)} > c_\epsilon(1 + \mathbb{E}[Y_i^{(n)}])\}} \right] \leq \epsilon \mathbb{E} \left[ (Y_i^{(n)})^2 \mathbf{1}_{\{Y_i^{(n)} \geq 2\}} \right], \quad \text{for any} \quad n \geq 1. \]

He explained that a direct verification of his assumptions can be cumbersome. Therefore, he introduced condition (A) as an assumption easier to handle that implies the latter condition. For this reason, we prefer to work directly under the assumption (A), which is good enough for our purposes.

In this manuscript, we give a probabilistic argument of Yaglom’s limit for GWVE. It is based on a two spine decomposition method and a characterisation of the exponential distribution via a size-biased transform and is close in spirit to that of [9]. The authors in [9] created a two-spine decomposition technique for Galton-Watson processes in constant environment that cannot be translated directly into our settings. Here, associated to each \( Z_n^Q \), we construct a Galton-Watson tree in varying environment up to time \( n \) with two marked genealogical lines. This tree can be decomposed in subtrees along these lines. A key point is the distribution of the generation of the most recent common ancestor of these genealogical lines, \( K_n \). When the environment is constant, \( K_n \) has uniform distribution in \( \{0, \ldots, n-1\} \) and the subtrees are independent Galton-Watson trees. When the environment varies, this last property does not hold anymore. In order to match the
above decomposition with that at the exponential distribution, it is fundamental to know the law of $K_n$ explicitly. Thus, we determine the distribution of $K_n$ that makes the method work. Moreover, we identify the subtrees with Galton-Watson trees in a modified environment. In the next section, we explains this in further detail.

Our contribution is that our proof provides further understanding on why the limit must be an exponential random variable. An important part of our approach relies in study random trees and adequately select inside them two marked genealogical lines. We believe that one can adapt this decomposition technique to establish a Yaglom’s limit for branching processes in random environment. The construction has to be the same but, for the two genealogical lines, one has to find the distribution of the generation of their most recent common ancestor that makes the method work. This possible application highlight the potential and relevance of this new approach.

The remainder of the paper is organised as follows. In Section 2, we introduce the one-spine and two-spines decompositions. With this in hand, we give an intuitive explanation of the result and we explain why the limit must be exponential. In Section 3, we give some properties of the measures associated with these decompositions and we characterise them via their Laplace transform. Finally, Section 4 contains the proof.

2. Outline of the proof

In this section, we provide an intuitive explanation of the result and explain why the limit must be an exponential random variable. First, we explain the one-spine and two-spines decompositions. Then, we relate them with a size-biased characterisation of the exponential random variable.

Recall that given a random variable $X$ and a Borel function $g$ such that $\mathbb{P}(g(X) \geq 0) = 1$, and $\mathbb{E}[g(X)] \in (0, \infty)$, we say that $W$ is a $g(X)$-transform of $X$ if

$$
\mathbb{E}[f(W)] = \frac{\mathbb{E}[f(X)g(X)]}{\mathbb{E}[g(X)]},
$$

for each positive Borel function $f$. If $g(x) = x$, we also call it the size-biased transform.

Observe that the law of a non-negative random variable $X$ conditioned on being positive can be described in terms of its size-biased transform. More precisely, for each $\lambda \geq 0$,

$$
\mathbb{E}[1 - e^{-\lambda X} | X > 0] = \int_0^\lambda \frac{\mathbb{E}[X e^{-s X}]}{\mathbb{P}(X > 0)} \, ds = \mathbb{E}[X | X > 0] \int_0^\lambda \mathbb{E}[e^{-s X}] \, ds,
$$

where $\hat{X}$ is the size-biased transform of $X$. Recall that a sequence of non-negative random variables converges in distribution if and only if their Laplace transforms converge. As a consequence, we obtain the following lemma

**Lemma 1.** Let $\{X_n : n \geq 0\}$ be a sequence of random variables. Then the variables conditioned on being positive $\{X_n : \mathbb{P}(\cdot | X_n > 0)\}_{n \geq 0}$ converges in distribution to a positive random variable $Y$ if and only if $\mathbb{E}[X_n | X_n > 0] \to \mathbb{E}[Y]$ and $\hat{X}_n$ converges in distribution to $\hat{Y}$, where $\hat{X}_n$ and $\hat{Y}$ are the size-biased transforms of $X_n$ and $Y$, respectively.

By Lemma 1 in order to prove Theorem 1 we need to study the size-biased process $\hat{Z}^Q := \{\hat{Z}^Q_n : n \geq 0\}$. Recall that there is a relationship between Galton-Watson processes
in environment \( Q \) and Galton-Watson trees in environment \( Q \). In the tree, any particle or individual in generation \( i \) gives birth to particles in generation \( i + 1 \) according to \( q_{i+1} \). The variable \( Z^Q_n \) is the number of particles at generation \( n \) in the tree. In a similar way, \( \dot{Z}^Q_n \) is the population size at generation \( n \) of some random tree. According to Kersting and Vatutin [7, Sections 1.4.1 and 1.4.2], the tree associated to \( \dot{Z}^Q_n \) is a size-biased tree in varying environment \( Q \). More precisely, for each \( i \geq 1 \), let \( \dot{q}_i \) be the size-biased transform of \( q_i \),

\[
\dot{q}_i(k) = \frac{k}{f'_i(1)} q_i(k), \quad k \in \mathbb{N}_0.
\]

The size-biased tree in varying environment \( Q \) is constructed as follows:

(i) We first establish an initial marked particle,
(ii) the marked particle in generation \( i \in \mathbb{N}_0 \) gives birth to particles in generation \( i + 1 \) according to \( \dot{q}_{i+1} \). Uniformly, we select one of these particles as the marked particle. All the others particles are unmarked,
(iii) any unmarked particle in generation \( i \in \mathbb{N}_0 \) gives birth to unmarked particles in generation \( i + 1 \) according to \( q_{i+1} \), independently of other particles.

The marked genealogical line is called spine. This construction is known as the one-spine decomposition; see Figure 1a below. The constant environment case was done by Lyons, Pemantle and Peres [8]. According to Kersting and Vatutin, \( Z^Q_n \) is the number of particles at generation \( n \) in the previous tree.

![Figure 1. Spine decompositions](image)

Now, we want to construct a random tree up to generation \( n \) with two marked genealogical lines or spines. Denote by \( K_n \) the generation of the most recent common ancestor of the lines. Note that before \( K_n \), there is only one spine and in generation \( K_n + 1 \) a second spine is created. Since the offspring distribution is varying among generations, \( K_n \) should depend on the environment. We assume that in this construction, \( K_n \) has the following distribution

\[
P(K_n = r) := \frac{\nu_{r+1}}{\mu_r} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1}, \quad 0 \leq r \leq n - 1,
\]

where \( \mu_n \) and \( \nu_n \) are defined in (3). Thus, by (3), generations with larger offspring mean or larger offspring variance are more probably to be chosen as \( K_n \). In generation
$K_n$, we need to have an offspring distribution with two or more individuals. We denote by $\tilde{q}_i$ the $q_i(q_i - 1)$-transform of $q_i$ given by

$$
\tilde{q}_i(k) = \frac{k(k-1)q_i(k)}{\nu_i f_i(1)^2}, \quad k \in \mathbb{N}_0, \quad i = 1, \ldots, n.
$$

We define a $X(X - 1)$-type size-biased tree in environment $Q$ up to time $n$ as the tree constructed as follows:

(i) we first establish an initial marked particle,
(ii) select $K_n$ according to [7],
(iii) the marked particle in generation $K_n$ gives birth to particles according to $\tilde{q}_{K_n+1}$.

Uniformly, we select two of these particles as the marked particles in generation $K_n + 1$. The other particles are unmarked,
(iv) any marked particle in generation $i \in \{0, \ldots, n-1\} \setminus K_n$ gives birth to particles in generation $i + 1$ according to $\tilde{q}_{i+1}$. Uniformly, select one of these as the marked particle. All the other particles are not marked,
(v) any unmarked particle in generation $i \in \{0, \ldots, n-1\}$ gives birth to unmarked particles in generation $i + 1$ according to $q_{i+1}$, independently of other particles.

We call this construction as the two-spine decomposition; see Figure 1b. Ren et. al [9] provided a two spine decomposition for Galton-Watson processes in a constant environment. In this case, the distribution of $K_n$ is uniform in $\{0, \ldots, n-1\}$. Using that the environment is constant we can recover their construction.

For any $0 \leq k \leq n$, let $\tilde{Z}_k^n$ be the population size at the $k$-th generation in the previous tree. From the constructions of the size-biased trees (see Figure 1), we see that we can decompose the particles associated to $\tilde{Z}_k^n$ into descendants attached to the longer spine and descendants attached to the shorter spine. The descendants attached to the longer spine can be seen as the population in the $n$-th generation of a size-biased tree with environment $Q$, while the descendants of the shorter spine can be seen as the population in generation $n - (K_n + 1)$ of a size-biased tree with environment $Q \circ \theta_{K_n+1}$ := $(q_{K_n+2}, q_{K_n+3}, \ldots)$. By construction, the two subpopulations are independent. Therefore,

$$
\tilde{Z}_n^Q \overset{(d)}{=} \tilde{Z}_n^Q + \tilde{Z}_n^Q \circ \theta_{n-(K_n+1)}, \quad n \geq 1,
$$

where the right-hand side of the equation is an independent sum. If we normalise with $a_n^Q$, we obtain

$$
\tilde{Z}_n^Q \overset{(d)}{=} \tilde{Z}_n^Q \circ \theta_{n-(K_n+1)} \overset{(d)}{\sim} U, \quad \text{as } n \to \infty,
$$

where $U$ is an uniform random variable on $[0, 1]$ (see Proposition 3), and that $\tilde{Z}_n^Q / a_n^Q$ converges in distribution to a random variable $\tilde{Y}$ (see Proposition 4). Since $\tilde{Z}_n^Q$ is the $(\tilde{Z}_n^Q - 1)$-transform of $\tilde{Z}_n^Q$, we have that $\tilde{Z}_n^Q / a_n^Q$ converges in distribution to $\tilde{Y}$, the $\tilde{Y}$-transform of $\tilde{Y}$. Hence, by Lemma 1 if we take limits in (9), we see...
that \( Z_n^Q/a_n^Q \) converges in distribution to a random variable \( Y \) that satisfies

\[
\tilde{Y} \overset{(d)}{=} \tilde{Y} + U \cdot \tilde{Y}'
\]

where \( \tilde{Y} \) and \( \tilde{Y}' \) are both \( Y \)-transforms of \( Y \), \( \tilde{Y} \) is a \( Y^2 \)-transform of \( Y \), and \( U \) is an uniform random variable on \([0, 1]\) independent of \( \tilde{Y} \) and \( \tilde{Y}' \). Ren et. al. [9, Lemma 1.3], showed that a variable \( Y \) is exponentially distributed if and only if (10) holds. Therefore, \( Z_n^Q/a_n^Q \) must converge in distribution to an exponential random variable.

### 3. Sized-biased trees

In this section, we study the sized-biased trees defined in the previous section. We associate them a probability measure in the set of rooted trees. For this purpose, we introduce the so-called Ulam-Harris labeling. Let \( U \) be the set of finite sequences of positive integers, including \( \emptyset \). For \( u \in U \), we define the length of \( u \) by \(|u| := n\), if \( u = u_1 \cdots u_n \), where \( n \geq 1 \) and by \(|\emptyset| := 0\) if \( u = \emptyset \). If \( u \) and \( v \) are two elements in \( U \), we denote by \( uv \) the concatenation of \( u \) and \( v \), with the convention that \( uv = u \) if \( v = \emptyset \). The genealogical line of \( u \) is denoted by \( [\emptyset, u] = \{\emptyset\} \cup \{u_1 \cdots u_j : j = 1, \ldots, n\} \).

Let \( s \subset U \), its most recent common ancestor is the unique element \( v \in \cap_{n \in s}[\emptyset, u] \) with maximal length and its generation is denoted by \( K_s \).

A rooted tree \( t \) is a subset of \( U \) that satisfies \( \emptyset \in t \), \([\emptyset, u] \subset t \) for any \( u \in t \), and if \( u \in t \) and \( i \in \mathbb{N} \) satisfy that \( u_i \in t \) then, \( u_j \in t \) for all \( 1 \leq j \leq i \). Denote by \( T = \{t : t \text{ is a tree}\} \), the subspace of rooted trees. The vertex \( \emptyset \) is called the root of the tree. For any \( u \in t \), we define the number of offspring of \( u \) by \( l_u(t) = \max\{i \in \mathbb{Z}^+ : u_i \in t\} \). The height of \( t \) is defined by \(|t| = \sup\{|u| : u \in t\} \).

For any \( n \in \mathbb{N} \) and \( t, \bar{t} \) trees, we write \( t \sim n \bar{t} \) if they coincide up to height \( n \). The population size in the \( n \)-th generation of the tree \( t \) is denoted by \( X_n(t) = \#\{u \in t : |u| = n\} \).

A Galton-Watson tree in the environment \( Q = (q_1, q_2, \ldots) \) is a \( T \)-valued random variable \( T \) such that

\[
G_n(t) := \mathbb{P}(T = t) = \prod_{u \in t : |u| < n} q_{|u|+1}(l_u(t)),
\]

for any \( n \geq 0 \) and any tree \( t \). As we said before, the process \( Z = \{Z_n^Q : n \geq 0\} \) defined as \( Z_n^Q = X_n(T) \) is a Galton-Watson process in environment \( Q \).

Now, we deal with the one-spine decomposition. This construction builds a tree along a distinguished path. More precisely, a spine or distinguished path \( v \) on a tree \( t \) is a sequence \( \{v(k) : k = 0, 1, \ldots, |t|\} \subset t \) such that \( v(0) = \emptyset \) and \( v(k) = v(k-1)j \) for some \( j \in \mathbb{N} \), for any \( 1 \leq k \leq |t| \). We denote by \( \bar{T} \), the subspace of trees with one spine

\[
\bar{T} = \{(t, v) : t \text{ is a tree and } v \text{ is a spine on } t\}
\]

and by \( \bar{T}_n = \{t \in \bar{T} : |t| = n\} \) and \( \bar{T}_n = \{(t, v) \in \bar{T} : |t| = n\} \) the restriction of \( \bar{T} \) and \( \bar{T} \) to trees with height \( n \).

We are going to construct the probability distribution of the size-biased tree in the environment \( Q \) on the state space \( \bar{T} \). First, we need to define a probability distribution on \( \bar{T} \). Recall the construction of the size-biased tree in the previous section; individuals along the spine, \( \{u \in t : u \in v\} \), have offspring distribution \( \hat{q}_{|u|+1} \) given by (3), from its offspring we select one uniformly as the spine individual in the next generation. Individuals
outside the spine, \( \{u \in t : u \notin v\} \), have offspring distribution \( q_{|u|+1} \). Then, the size-biased tree can be seen as a \( \mathcal{T} \)-valued random variable \((\mathcal{T}, \mathcal{V})\) with distribution

\[
\mathbb{P}(\langle \mathcal{T}, \mathcal{V} \rangle \overset{n}{=} (t, v)) := \prod_{u \in v : |u| < n} \hat{q}_{|u|+1}(l_u(t)) \frac{1}{l_u(t)} \prod_{u \in t \setminus v : |u| < n} q_{|u|+1}(l_u(t)),
\]

for any \( n \geq 0 \) and any \((t, v) \in \mathcal{T}_n\). One readily checks that this measure is a probability on \( \mathcal{T} \) by using the definition of \( \hat{q} \) and the fact that \( G_n \) is a probability measure. In a similar way, we can write

\[
\mathbb{P}(\langle \mathcal{T}, \mathcal{V} \rangle \overset{n}{=} (t, v)) = \frac{1}{\mu_n} \cdot G_n(t), \quad (t, v) \in \mathcal{T}.
\]

Hence, by summing over all the possible spines, we obtain the distribution of the size-biased Galton-Watson tree in environment \( Q \) on \( \mathcal{T} \)

\[
\mathcal{G}_n(t) := \mathbb{P}(\mathcal{T} \overset{n}{=} t) = \sum_{v : (t, v) \in \mathcal{T}_n} \mathbb{P}(\langle \mathcal{T}, \mathcal{V} \rangle \overset{n}{=} (t, v)) = \frac{1}{\mu_n} X_n(t) \cdot G_n(t),
\]

for any \( n \geq 0 \) and any \( t \in \mathcal{T}_n \). (see also [7, Lemma 1.2]). Define the process \( \hat{Z}_n = \{ \hat{Z}_n^Q : n \geq 0 \} \) as \( \hat{Z}_n^Q = X_n(\mathcal{T}) \), for each \( n \geq 1 \). Then, by using the measure \( \mathcal{G}_n \) we can see that the process \( \{ \hat{Z}_m^Q : 0 \leq m \leq n \} \) is a \( Z_n^Q \)-transform of \( \{ Z_m^Q : 0 \leq m \leq n \} \), in other words

\[
\mathbb{E} \left[ g(\hat{Z}_1^Q, \ldots, \hat{Z}_n^Q) \right] = \frac{\mathbb{E} \left[ Z_n^Q g(Z_1^Q, \ldots, Z_n^Q) \right]}{\mathbb{E} \left[ Z_n^Q \right]}, \quad \text{for all bounded function } g.
\]

Now we consider the probability distribution associated to the \( X(X-1) \)-type size-biased tree up to time \( n \) on the state space \( \mathcal{T}_n \). As we did before, we define a measure on

\[
\mathcal{T}_n := \left\{ (t, v, \tilde{v}) : (t, v), (t, \tilde{v}) \in \mathcal{T}_n, v \neq \tilde{v} \right\}, \quad n \in \mathbb{N},
\]

the subspace of trees with height \( n \) and two different spines. Given a \((t, v, \tilde{v}) \in \mathcal{T}_n\), we denote by \( K_{v, \tilde{v}} \) the generation of the most recent common ancestor of \( v \cup \tilde{v} \).

Recall the construction of a \( X(X-1) \)-type size-biased tree in the previous section; (i) consider an initial spine individual, (ii) select the generation of the most recent common ancestor, \( K_{v, \tilde{v}} \), according to \( [7] \), (iii) the spine individual \( u \) in that generation has offspring distribution \( \hat{q}_{|u|+1} \) given by \( [8] \). From its offspring we select uniformly two as spine individuals in the next generation, (iv) the spine individuals in the other generations, \( \{ u \in v \cup \tilde{v} : |u| \neq K_{v, \tilde{v}} \} \), have offspring distribution \( \hat{q}_{|u|+1} \) given by \( [6] \). From its offspring we select uniformly one as the spine individual in the next generation, (v) finally, individuals outside the spine, \( \{ u \in t : u \notin v \cup \tilde{v} \} \), have offspring distribution \( q_{|u|+1} \). Then, the \( X(X-1) \)-type size-biased tree up to time \( n \) can be seen as a \( \mathcal{T}_n \)-valued random
variable \((\mathbf{T}, \mathbf{V}, \mathbf{\tilde{V}})\) with distribution
\[
\mathbb{P}( (\mathbf{T}, \mathbf{V}, \mathbf{\tilde{V}}) \sim (\mathbf{t}, \mathbf{v}, \mathbf{\tilde{v}}) ) = \frac{\nu_{K_{\mathbf{v},\mathbf{v}}+1}}{\mu_{K_{\mathbf{v},\mathbf{v}}}} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \prod_{\mathbf{u} \in \mathbf{\tilde{V}} : K_{\mathbf{v},\mathbf{v}} \neq |\mathbf{u}| < n} \hat{q}_{|\mathbf{u}|+1}(l_u(\mathbf{t})) \frac{2}{l_u(\mathbf{t})(l_u(\mathbf{t})-1)} \prod_{\mathbf{u} \in \mathbf{\tilde{V}} : K_{\mathbf{v},\mathbf{v}} = |\mathbf{u}| < n} \tilde{q}_{|\mathbf{u}|+1}(l_u(\mathbf{t})) \prod_{\mathbf{u} \in \mathbf{\tilde{V}} : K_{\mathbf{v},\mathbf{v}} = |\mathbf{u}| = n} q_{|\mathbf{u}|+1}(l_u(\mathbf{t}))
\]

for any \((\mathbf{t}, \mathbf{v}, \mathbf{\tilde{v}}) \in \mathcal{T}_n\). Here, the first two terms in the right-hand side of the equation are associated with step (ii). The first product is associated with step (iii). Then, in the second line, the first product is obtained with (iv). Finally, we use (v) to obtain the last product. By using the definition of \(q\), \(\hat{q}\) and \(\tilde{q}\), one can readily verify that the previous expression defines a probability measure on \(\mathcal{T}_n\). Moreover, we have
\[
\mathbb{P}( (\mathbf{T}, \mathbf{V}, \mathbf{\tilde{V}}) \sim (\mathbf{t}, \mathbf{v}, \mathbf{\tilde{v}}) ) = \frac{2}{\mu_n^2} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} G_n(\mathbf{t}),
\]

for any \((\mathbf{t}, \mathbf{v}, \mathbf{\tilde{v}}) \in \mathcal{T}_n\). Then, by summing over all the possible two spines, we obtain that the \(X(X-1)\)-type size-biased tree up to time \(n\) is a \(\mathcal{T}_n\)-valued random variable \(\mathbf{T}\) with law
\[
\dot{G}_n(\mathbf{t}) := \mathbb{P}(\mathbf{T} = \mathbf{t}) = \frac{1}{\mu_n^2} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} X_n(\mathbf{t})(X_n(\mathbf{t})-1) \cdot G_n(\mathbf{t}),
\]

for any \(\mathbf{t} \in \mathcal{T}_n\). Define the process \(\dot{\mathbf{Z}}^Q = \{\dot{\mathbf{Z}}_m^Q : 0 \leq m \leq n\}\) by \(\dot{\mathbf{Z}}_m^Q = X_m(\mathbf{T})\).

Opposite to what happens with \((\mathbf{G}_n : n \geq 1)\), by construction, the measures \((\dot{\mathbf{G}}_n : n \geq 1)\) are not consistent in the sense that \(\dot{\mathbf{G}}_n\) is not a restriction of \(\dot{\mathbf{G}}_{n+1}\) to the tree with size \(n\). Then, the change of measure in the next proposition is not a martingale change of measure. However, it allows us to conclude that \(\{\dot{\mathbf{Z}}_m^Q : 0 \leq m \leq n\}\) is a \(\mathcal{Z}_n^Q(\mathcal{Z}_n^Q-1)\)-transform of \(\{\mathbf{Z}_m^Q : 0 \leq m \leq n\}\).

**Proposition 1.** Let \(\{\mathbf{Z}_n^O : n \geq 0\}\) be a GWVE and for any \(n \in \mathbb{N}_0\), let \(\dot{\mathbf{Z}}^Q = (\dot{\mathbf{Z}}_m^Q : 0 \leq m \leq n)\) be the process associated with the \(X(X-1)\)-type size-biased tree up to time \(n\). Then, for any bounded function \(g : \mathcal{Z}_n^O \to \mathbb{R}\),
\[
\mathbb{E}[g(\dot{\mathbf{Z}}^Q_1, \ldots, \dot{\mathbf{Z}}^Q_n)] = \mathbb{E}[\mathbf{Z}_n^Q(\mathcal{Z}_n^O-1)g(\mathbf{Z}_1^O, \ldots, \mathbf{Z}_n^O)] / \mathbb{E}[\mathbf{Z}_n^Q(\mathcal{Z}_n^O-1)].
\]

**Proof.** Fix \(n \geq 0\) and recall that for each \(m \leq n\), \(\mathbf{Z}_m^Q = X_m(\mathbf{T})\) under the measure \(\mathbf{G}_n\) and \(\dot{\mathbf{Z}}_m^Q = X_m(\mathbf{T})\) under the measure \(\dot{\mathbf{G}}_n\). Hence, by (11)
\[
\mathbb{E}[g(\dot{\mathbf{Z}}^Q_1, \ldots, \dot{\mathbf{Z}}^Q_n)] = \dot{\mathbf{G}}_n[g(X_1(\mathbf{T}), \ldots, X_n(\mathbf{T}))]
\]
\[
= \frac{1}{\mu_n^2} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \mathbf{G}_n[X_n(\mathbf{T})(X_n(\mathbf{T})-1)g(X_1(\mathbf{T}), \ldots, X_n(\mathbf{T}))]
\]
\[
= \frac{1}{\mu_n^2} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \mathbb{E}[\mathbf{Z}_n^Q(\mathcal{Z}_n^O-1)g(\mathbf{Z}_1^O, \ldots, \mathbf{Z}_n^O)].
\]
By taking \( g \equiv 1 \), we deduce that
\[
\mathbb{E}[Z_n^Q(Z_n^Q - 1)] = \mu_n^2 \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k},
\]
which implies the result. \( \square \)

In the remainder of this section, we study some properties of the previous decompositions. We first introduce the notation to refer to shifted environments. Let \( q \) be a probability measure on \( \mathbb{N}_0 \) such that \( q(\{0, 1, \ldots, r-1\}) = 0 \) for some \( r \in \mathbb{N} \). We define the probability measure \([q - r] \) in \( \mathbb{N}_0 \) by \([q - r](i) = q(i + r)\) for all \( i \in \mathbb{N}_0 \). Given a probability measure \( q \) and an environment \( Q = (q_1, q_2, \ldots) \), we denote
\[
q \oplus Q := (q, q_1, q_2, \ldots).
\]
For any \( m \in \mathbb{N}_0 \), we set
\[
Q \circ \theta_m := (q_{m+1}, q_{m+2}, \ldots).
\]

We can compute the Laplace transform of \( \hat{Z}_n^Q \) in terms of the Laplace transform of \( \hat{Z}_n^{Q_0\theta_{m+1}} \) as indicated below. The proof follows similar arguments as those used in \([9, \text{Proposition 2.1}] \), however the presence of varying environment turns out into significant changes.

**Proposition 2.** Fix \( n \geq 1 \). Let \( \{\hat{Z}_m : m \leq n\} \) and \( \{\hat{Z}_m : m \leq n\} \) be the population size of the size-biased tree and the \( X(X - 1)\)-type size-biased tree up to time \( n \). Then, we have the following decomposition, for each \( \lambda \geq 0 \)
\[
\mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^Q\right\}\right] = \mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^{Q_0\theta_{m+1}}\right\}\right] \mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^{Q_0\theta_{m+1}}\right\}\right] g(n, m, \lambda),
\]
where the function \( g \) is defined as follows
\[
g(n, m, \lambda) := \frac{\mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^{Q_0\theta_{m+1}}\right\}\right]}{\mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^{Q_0\theta_{m+1}}\right\}\right]}, \quad 0 \leq m \leq n - 1, \quad 0 \leq \lambda.
\]

**Proof.** Let \( \hat{T} \) be a size-biased Galton-Watson tree in environment \( Q \) up to time \( n \). We can decompose \( \hat{T} \) in subtrees with roots along the spine \( V \); see Figure 2a. More precisely, for every \( 0 \leq k \leq n \), there is a \( v^{(k)} \in V \) with \( |v^{(k)}| = k \) and a random tree \( t_k \in \mathcal{T} \) such that
\[
v^{(k)} \cdot t_k = \{u \in \hat{T} : |[0, u] \cap V| = k\} \quad \text{and} \quad \hat{T} = \bigsqcup_{k=0}^{n} v^{(k)} \cdot t_k,
\]
where \( \bigsqcup \) denotes the disjoint union. Note that \( X_n(\hat{T}) = \sum_{k=0}^{n} X_{n-k}(t_k) \). In the size-biased tree, individual along the spine gives birth according to \( \hat{q} \) and one of its offspring is the spine individual in the next generation. Then, it follows that the subtrees \( t_k, \ 0 \leq k \leq n \), are independent Galton-Watson trees with environment \([q_{k+1} - 1] \oplus Q \circ \theta_{k+1} \). Therefore,
\[
\mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^Q\right\}\right] = \prod_{k=0}^{n} \mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_{n-k}^{q_{k+1} - 1} \oplus Q \circ \theta_{k+1}\right\}\right], \quad \lambda \leq 0, \ n \in \mathbb{N}_0.
\]
Let $\mathbf{T}$ be a $X(X - 1)$-type size-biased Galton-Watson tree up to time $n$. In a similar way, we can decompose $\mathbf{T}$ in subtrees with roots along the spines; see Figure 2b. Denote by $V$ and $\bar{V}$, the associated spines and recall that $K_n = \max\{r < n : V = \bar{V}\}$. We can form a partition of $\mathbf{T}$ in the sense that

$$
\mathbf{T} = \bigcup_{k=0}^{n} (v^{(k)} \mathbf{t}_k) \sqcup \bigcup_{k=1+K_n}^{n} (\bar{v}^{(k)} \bar{\mathbf{t}}_k)
$$

and $X_n(\mathbf{T}) = \sum_{k=0}^{n} X_{n-k}(\mathbf{t}_k) + \sum_{k=1+K_n}^{n} X_{n-k}(\bar{\mathbf{t}}_k)$,

where, for every $0 \leq k \leq K_n$, $v^{(k)} \in V \cap \bar{V}$ and $\mathbf{t}_k \in \mathcal{T}$ are such that $|v^{(k)}| = k$ and $v^{(k)} \mathbf{t}_k = \{u \in \mathbf{T} : |(\emptyset, u) \cap (V \cup \bar{V})| = k\}$;

and, for every $K_n < k \leq n$, $v^{(k)} \in V$, $\bar{v}^{(k)} \in \bar{V}$ and $\mathbf{t}_k, \bar{\mathbf{t}}_k \in \mathcal{T}$ satisfy $|v^{(k)}| = k = |\bar{v}^{(k)}|$, $v^{(k)} \mathbf{t}_k = \{u \in \mathbf{T} : |(\emptyset, u) \cap V| = k\}$ and $\bar{v}^{(k)} \bar{\mathbf{t}}_k = \{u \in \mathbf{T} : |(\emptyset, u) \cap \bar{V}| = k\}$.

Observe that by the branching property, the subtrees are independent. The spine individual at generation $K_n = m$, has offspring distribution $\tilde{q}_{m+1}$, from its offspring we select two as the spine individuals in the next generation. Then, the subtree $\mathbf{t}_m$ is a Galton-Watson tree with environment $[\tilde{q}_{m+1} - 2] \oplus Q \circ \theta_{m+1}$. The other subtrees $\{\mathbf{t}_k : 0 \leq k \leq n, k \neq m\}$ and $\{\bar{\mathbf{t}}_k : m < k \leq n\}$ are Galton-Watson trees with environment $[\tilde{q}_{k+1} - 1] \oplus Q \circ \theta_{k+1}$. Therefore, by using the decomposition (15), we have

$$
\mathbb{E}\left[\exp\left\{-\lambda Z_n^Q\right\}\right] = \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \mathbb{E}\left[\exp\left\{-\lambda Z^{|\tilde{q}_{m+1} - 2| \oplus Q \circ \theta_{m+1}}\right\}\right]
$$

$$
\times \prod_{k=0, k \neq m}^{n} \mathbb{E}\left[\exp\left\{-\lambda Z^{|\tilde{q}_{k+1} - 1| \oplus Q \circ \theta_{k+1}}\right\}\right] \prod_{k=m+1}^{n} \mathbb{E}\left[\exp\left\{-\lambda Z^{|\tilde{q}_{k+1} - 1| \oplus Q \circ \theta_{k+1}}\right\}\right].
$$
Finally, if we apply equation (14) for environments $Q$ and $Q \circ \theta_{m+1}$, we obtain the result. In other words,

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_n^Q \right\} \right] = \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-m}^{[\hat{q}_{m+1}-1] \oplus Q \circ \theta_{m+1}} \right\} \right]
\]

\[
\times \frac{\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_n^Q \right\} \right]}{\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-(m+1)}^Q \right\} \right]} \mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-(m+1)}^{Q \circ \theta_{m+1}} \right\} \right].
\]

\[\square\]

The distribution of the previous processes can be expressed via the generating functions $(f_1, f_2, \ldots)$ associated to $Q = (q_1, q_2, \ldots)$. For each $0 \leq m \leq n$ we define

\[f_{m,n}(s) := [f_{m+1} \circ \ldots \circ f_n](s),\]

and $f_{n,n} := s$. The generating function of $Z_n^Q$ is equal to $f_{0,n}$. For the others, we note that for every $s \in [0, 1]$ and $0 \leq m < n$,

\[
f'_{m,n}(s) = \prod_{l=m+1}^{n} f'_l(f_{l,n}(s)), \quad f''_{m,n}(s) = f'_{m,n}(s)^2 \sum_{l=m+1}^{n} \frac{f''_l(f_{l,n}(s))}{f'_l(f_{l,n}(s))^2} \prod_{j=m+1}^{l-1} f'_j(f_{j,n}(s)),
\]

where $f'_{n,n}(s) = 1$ and $f''_{n,n}(s) = 0$.

**Lemma 2.** Let $n \geq 1$ and $Q$ be a varying environment. Let $(Z_n^Q : 0 \leq m \leq n)$ and $(\hat{Z}_m^Q : 0 \leq m \leq n)$ be a GWVE, a sized-biased GWVE and a $X(X-1)$-type sized-biased GWVE up to time $n$. Then, for any $0 \leq m < n$ and $\lambda \geq 0$,

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-m}^{[\hat{q}_{m+1}-1] \oplus Q \circ \theta_{m+1}} \right\} \right] = \frac{1}{f'_{m+1}(1)} f'_{m+1}(f_{m+1,n}(e^{-\lambda})),
\]

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-m}^{[\hat{q}_{m+1}-2] \oplus Q \circ \theta_{m+1}} \right\} \right] = \frac{1}{\nu_{m+1} f'_{m+1}(1)^2} f''_{m+1}(f_{m+1,n}(e^{-\lambda})),
\]

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_n^Q \right\} \right] = \frac{1}{\mu_n} f'_{0,n}(e^{-\lambda}) e^{-\lambda},
\]

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-(m+1)}^{Q \circ \theta_{m+1}} \right\} \right] = \frac{\mu_{m+1}}{\mu_n} f'_{m+1,n}(e^{-\lambda}) e^{-\lambda},
\]

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_n^Q \right\} \right] = \frac{1}{\mu_n} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} f''_{0,n}(e^{-\lambda}) e^{-2\lambda}.
\]

**Proof.** Denote by $(g_{m+1}, f_{m+2}, f_{m+3}, \ldots)$ the generating functions of the environment $[\hat{q}_{m+1} - 1] \oplus Q \circ \theta_{m+1} = (\hat{q}_{m+1} - 1, q_{m+2}, q_{m+3}, \ldots)$, where $\hat{q}$ is given in (6). Note that,

\[
g_{m+1}(s) = \frac{1}{f'_{m+1}(1)} \sum_{k=1}^{\infty} k s^{k-1} q_{m+1}(k) = \frac{1}{f'_{m+1}(1)} f'_{m+1}(s).
\]

Then, we can deduce (16), i.e.

\[
\mathbb{E}\left[ \exp \left\{ -\lambda \hat{Z}_{n-m}^{[\hat{q}_{m+1}-1] \oplus Q \circ \theta_{m+1}} \right\} \right] = g_{m+1} \circ f_{m+2} \circ \ldots \circ f_n(e^{-\lambda}) = \frac{1}{f'_{m+1}(1)} f'_{m+1}(f_{m+1,n}(e^{-\lambda})),
\]
where we use that the Laplace transform of a GWVE given in (1). The proof of (17) follows similar arguments. Recall the definition of $\hat{q}_{m+1}$ in (8). It is enough to see that the generating function of $[\hat{q}_{m+1} - 2]$, denoted by $h_{m+1}$, is

$$h_{m+1}(s) = \frac{1}{\nu_{m+1}f_{m+1}(1)^2} \sum_{k=2}^{\infty} k(k-1)s^{k-2}q_{m+1}(k) = \frac{1}{\nu_{m+1}f_{m+1}(1)^2} f''_{m+1}(s).$$

In order to prove (18), note that $Z_n^Q$ is a size-biased transform of $Z_n^Q$. Then, by (5)

$$\int_{0}^{\lambda} \mathbb{E}\left[\exp\left\{-s\hat{Z}_n^Q\right\}\right] ds = \mathbb{E}\left[1 - \exp\left\{-\lambda Z_n^Q \mid Z_n^Q > 0\right\}\right] = \frac{\mathbb{E}\left[1 - \exp\left\{-\lambda Z_n^Q\right\}\right]}{\mathbb{E}\left[Z_n^Q \mid Z_n^Q > 0\right]}.$$ 

for all $\lambda \geq 0$. Differentiating the previous equation with respect to $\lambda$ and using the generating function of $Z_n^Q$, we obtain

$$\mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_n^Q\right\}\right] = \frac{1}{\mu_n} \frac{d}{d\lambda} \left(1 - f_{0,n}(e^{-\lambda})\right) = \frac{1}{\mu_n} f'_{0,n}(e^{-\lambda})e^{-\lambda}.$$ 

The identity (19) is obtained similarly as (18) but instead of working with the original environment $Q$ we use the shift environment $Q \circ \theta_{m+1}$.

Finally, in order to obtain (20) we use the decomposition presented in Proposition 2

$$\mathbb{E}\left[e^{-\lambda \hat{Z}_n^Q}\right] = \mathbb{E}\left[e^{-\lambda \hat{Z}_{n-m}^Q}\right] \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_{n-(m+1)}^Q\right\}\right] \mathbb{E}\left[\exp\left\{-\lambda \hat{Z}_{n-m}^Q \oplus \theta_{m+1}\right\}\right].$$

Remember that $K_n$ has distribution (7). Hence, substituting the previous Laplace transforms (i.e. equations (16), (17) and (18)) and simplifying, we get

$$\mathbb{E}\left[e^{-\lambda \hat{Z}_n^Q}\right] = \frac{f'_{0,n}(e^{-\lambda})}{\mu_n} \sum_{m=0}^{n-1} \frac{\nu_{m+1}}{\mu_m} \left(\sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k}\right)^{-1} \frac{f'_{m+1,n}(e^{-\lambda})e^{-\lambda}f''_{m+1}(f_{m+1,n}(e^{-\lambda}))}{\mu_n} \frac{\mu_{m+1}f'_{m+1,n}(e^{-\lambda})f''_{m+1}(f_{m+1,n}(e^{-\lambda}))}{\mu_n f''_{m+1}(f_{m+1,n}(e^{-\lambda}))}.$$ 

Note that for all $s \in [0,1]$ and $0 \leq m < n$,

$$f'_{m+1,n}(s) = \prod_{l=m+2}^{n} f'_{l}(f_{l,n}(s)) = \prod_{l=1}^{n} f'_{l}(f_{l,n}(s)) = \frac{f'_{0,n}(s)}{f''_{m+1}(f_{m+1,n}(s))} \prod_{l=1}^{m} f'_{l}(f_{l,n}(s)).$$

Then,

$$\mathbb{E}\left[e^{-\lambda \hat{Z}_n^Q}\right] = \frac{f'_{0,n}(e^{-\lambda})^2}{\mu_n^2} \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \left(\sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k}\right)^{-1} f''_{0,n}(e^{-\lambda})^2 \frac{f''_{m+1}(f_{m+1,n}(e^{-\lambda}))}{\prod_{l=1}^{m} f'_{l}(f_{l,n}(e^{-\lambda}))}.$$ 

This completes the proof. 

The next lemma provides the uniform convergence of the function $g$ defined in (13). The reader will find its importance in the next Section.
Lemma 3. Suppose that condition [A] is fulfilled. Then, for any $\lambda \geq 0$,
\[
\lim_{n \to \infty} \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \left( 1 - g \left( n, m, \frac{s}{a_n^Q} \right) \right) = 0.
\]

Proof. By applying Lemma 2, we have that for any $s \in [0, \lambda]$ and $0 \leq m \leq n - 1$,
\[
g \left( n, m, \frac{s}{a_n^Q} \right) = \frac{f'_{m+1}(1)}{f'^{m+1}_{m+1}(e^{-s/a_n^Q})} \frac{f'^{m+1}_{m+1}(e^{-s/a_n^Q})}{f'_{m+1}(1)}.
\]

The proof is thus complete as soon as we can show the following uniform convergences
\[
\lim_{n \to \infty} \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \left( 1 - \frac{f'_{m+1}(f_{m+1,n}(e^{-s/a_n^Q}))}{f'_{m+1}(1)} \right) = 0,
\]
\[
\lim_{n \to \infty} \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \left( 1 - \frac{f''_{m+1}(f_{m+1,n}(e^{-s/a_n^Q}))}{f''_{m+1}(1)} \right) = 0.
\]

We shall start with (21). With the help of the Mean Value Theorem for $f'_{m+1}$ and using
\[
0 \leq \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \left( 1 - \frac{f'_{m+1}(f_{m+1,n}(e^{-s/a_n^Q}))}{f'_{m+1}(1)} \right) \leq \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \frac{f''_{m+1}(1)}{f'_{m+1}(1)} \left( 1 - f_{m+1,n}(e^{-s/a_n^Q}) \right). 
\]

Kersting [6] Equation 23] showed that under condition [A], there exists $c > 0$ such that
\[
(f_{k+1}(1) - c f_{k+1}(1)) (1 + f_{k+1}(1)), \quad \text{for all } k \geq 1.
\]

Then,
\[
\sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \left( 1 - \frac{f'_{m+1}(f_{m+1,n}(e^{-s/a_n^Q}))}{f'_{m+1}(1)} \right) \leq \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} c(1 + f'_{m+1}(1)) \left( 1 - f_{m+1,n}(e^{-s/a_n^Q}) \right).
\]

For similar argument to those given above, using condition [A], and upon an adjustment
of the value of the constant, we can get the same upper bound for the left-hand side
supremums in (22). Therefore, it is enough to prove
\[
\lim_{n \to \infty} \sup_{0 \leq m < n} \sup_{s \in [0, \lambda]} \left( 1 + f'_{m+1}(1) \right) \left( 1 - f_{m+1,n}(e^{-s/a_n^Q}) \right) = 0.
\]

Let $\lambda \geq 0$. By the Mean Value Theorem for $f_{m+1,n}$ and using that $f'_{m+1,n}$ is an increasing
function, we get for any $0 \leq s \leq \lambda$ and $0 \leq m < n$
\[
0 \leq (1 + f'_{m+1}(1)) \left( 1 - f_{m+1,n}(e^{-s/a_n^Q}) \right) \leq (1 + f'_{m+1}(1)) f'_{m+1,n}(1)(1 - e^{-s/a_n^Q}).
\]

Observe that by Taylor’s approximation, $e^{-s/a_n^Q} = 1 - \frac{s}{a_n^Q} + y_n$ where $y_n \geq 0$ is the
remainder error term. Then, for $s \in [0, \lambda]$ and $0 \leq m < n$
\[
0 \leq (1 + f'_{m+1}(1)) \left( 1 - f_{m+1,n}(e^{-s/a_n^Q}) \right) \leq (1 + f'_{m+1}(1)) \frac{\mu_n}{\mu_{m+1} a_n^Q} \frac{\lambda}{\mu_{m+1} a_n^Q} = \left( \frac{1}{\mu_{m+1}} + \frac{1}{\mu_m} \right) \frac{\mu_n}{a_n^Q} \lambda.
\]

Now, we decompose the left-hand side of (24) into two limits where the supremum is
taken over two separate sets. Recall that in the critical case, given an $\epsilon > 0$ there exists
$N > 0$ such that \((a^Q_k)^{-1} \leq \epsilon\) for any $k \geq N$. Then, we take the two sets as \(\{m < N\}\) and \(\{N \leq m < n\}\). For the first limit, we observe
\[
\sup_{0 \leq m < N} \sup_{s \in [0, \lambda]} (1 + f_{m+1}'(1)) \left(1 - f_{m+1,n}(e^{-\lambda/a^Q_n})\right) \leq \frac{\mu_n}{a^Q_n} \max_{0 \leq m < N} \left(\frac{1}{\mu_{m+1}} + \frac{1}{\mu_m}\right).
\]
By criticality, $\mu_n/a^Q_n \to 0$ as $n \to 0$. Then,
\[
(26) \quad \lim_{n \to \infty} \sup_{0 \leq m < N} \sup_{s \in [0, \lambda]} (1 + f_{m+1}'(1)) \left(1 - f_{m+1,n}(e^{-\lambda/a^Q_n})\right) = 0.
\]
For the second limit, we note that for any $0 \leq m \leq n$,
\[
a^Q_m = \frac{1}{2} \sum_{k=0}^{m-1} \nu_k + 1 \leq \frac{1}{2} \sum_{k=0}^{n-1} \nu_k + 1 \mu_k = a^Q_n / \mu_n.
\]
Then, by (25) and using that $N \leq m < n$ we get
\[
\sup_{N \leq m < n} \sup_{s \in [0, \lambda]} (1 + f_{m+1}'(1)) \left(1 - f_{m+1,n}(e^{-\lambda/a^Q_n})\right) \leq \lambda \sup_{N \leq m < n} \left(\frac{1}{a^Q_{m+1}} + \frac{1}{a^Q_m}\right) \leq 2 \epsilon \lambda.
\]
Therefore,
\[
\lim_{n \to \infty} \sup_{N \leq m < n} \sup_{s \in [0, \lambda]} (1 + f_{m+1}'(1)) \left(1 - f_{m+1,n}(e^{-\lambda/a^Q_n})\right) = 0,
\]
which together with the limit (26) gives us (24). This concludes the proof. \(\square\)

4. PROOF OF THE MAIN RESULT

As we explained in the outline of the proof, in this manuscript we provide a probabilistic argument of a Yaglom-type limit for critical GWVEs. In the previous section we deduced that $\hat{Z}^Q_n$ is the $Z^Q_n(Z^Q_n - 1)$-transform of $Z^Q_n$. Here, we prove the other remaining steps, contained in Proposition 3 and Proposition 4. First, we present these propositions. Then, using all the tools that we created, we provide a proof for our main result. Finally, we prove the two propositions.

Recall the definition of $K_n$ in (7). Given the environment $Q$, we define
\[
A_{n,m} := \frac{a^Q_{n-1} a^Q_{m+1}}{a^Q_m}, \quad \text{for } 0 \leq m < n.
\]

**Proposition 3.** Let $Z^Q$ be a critical GWVE satisfying condition (A). Then
\[
A_{n,K_n} \xrightarrow{(d)} U, \quad \text{as } n \to \infty,
\]
where $U$ is a uniform random variable on $[0, 1]$.

Using the previous proposition, we can show the following.

**Proposition 4.** Let $\hat{Z}^Q = \{\hat{Z}^Q_n: n \geq 0\}$ be a size-biased GWVE. Then,
\[
(a^Q_n)^{-1} \hat{Z}^Q_n \xrightarrow{(d)} \hat{Y} \quad \text{as } n \to \infty,
\]
where $\hat{Y}$ is the size-biased transform of an exponential random variable.

We have all the ingredients to prove Yaglom’s Theorem under assumption (A).
the partition is defined by $\Pi_n$. Then, by definition

$$
\mu_n = \frac{\mathbb{E}[Z_n^Q]}{a_n^Q \mathbb{P}(Z_n^Q > 0)} = \frac{\mathbb{E}[Z_n^Q]}{a_n^Q \mathbb{P}(Z_n^Q > 0)},
$$

which goes to 1 according to (4). Therefore, Theorem 1 holds.

This manuscript is complete as soon as we prove Propositions 3 and 4. We start with Proposition 3.

**Proof of Proposition 3.** In order to obtain this result, it is enough to deduce

$$
\lim_{n \to \infty} \mathbb{P}(A_{n,K_n} \leq y) = y, \quad y \in [0, 1].
$$

Denote by $(\tilde{f}_1, \tilde{f}_2, \ldots)$ the generating functions associated with the environment $Q \circ \theta_{m+1}$. They can be written in term of the original environment as follows $\tilde{f}_k = f_{m+1+k}$, for $k \geq 1$. Then, by definition

$$
\mu_k = f_{m+1+k}(1) \cdot \cdots \cdot f_{m+k}(1) = \frac{\mu_{m+1+k}}{\mu_{m+1}} \quad \text{and} \quad \nu_k = \frac{f_{m+1+k}(1)}{f_{m+k}(1)^2} = \nu_{m+1+k}.
$$

Hence,

$$
a^{Q\circ \theta_{m+1}}_{n-(m+1)} = \frac{\mu_{n-(m+1)}}{2} \sum_{k=0}^{n-(m+1)-1} \frac{\nu_{k+1}}{\mu_k} = \frac{\mu_n}{2} \sum_{j=m+1}^{n-1} \frac{\nu_{j+1}}{\mu_j},
$$

and

$$
A_{n,m} = a^{Q\circ \theta_{m+1}}_{n-(m+1)} = \frac{\sum_{j=m+1}^{n-1} \nu_{j+1}}{\mu_j} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} = 1 - \sum_{j=0}^{m} \frac{\nu_{j+1}}{\mu_j} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1},
$$

where in the last equality, we completed the sum. Then,

$$
\mathbb{P}
\left(A_{n,K_n} = 1 - \sum_{j=0}^{m} \frac{\nu_{j+1}}{\mu_j} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1}
\right) = \mathbb{P}(A_{n,K_n} = A_{n,m}) = \mathbb{P}(K_n = m)
$$

(28)

$$
= \frac{\nu_{m+1}}{\mu_m} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1}.
$$

Note that $\{A_{n,m} : m = 0, \ldots, n-1\} \subset [0, 1]$ is a decreasing sequence with $A_{n,n-1} = 0$. Then, we can associate it to the partition $P^{(n)} = \{0 = \Pi_0^{(n)} < \Pi_1^{(n)} < \cdots < \Pi_{n-1}^{(n)} < \Pi_n^{(n)} = 1\}$ defined by $\Pi_k^{(n)} = A_{n,n-k-1}$, for any $0 \leq k < n$, with $\Pi_n^{(n)} = 1$. The norm of the partition is defined by

$$
||P^{(n)}|| = \max_{1 \leq k \leq n} \left\{ \Pi_k^{(n)} - \Pi_{k-1}^{(n)} \right\} = \max_{0 \leq m \leq n-1} \left\{ \frac{\nu_{m+1}}{\mu_m} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \right\}.
$$
Since $P^{(n)}$ is a partition, for each $y \in [0, 1]$ there exists $l_n := l(y, n) \in \{0, 1, \ldots, n-1\}$ such that $\Pi^{(n)}_{l_n} \leq y < \Pi^{(n)}_{l_{n+1}}$. Then, by (28)

$$\mathbb{P}(A_n, K_n \leq y) = \sum_{k=0}^{l_n} \mathbb{P}(A_n, K_n = \Pi^{(n)}_k) = \sum_{m=n-l_n}^{n-1} \frac{\nu_{m+1}}{\mu_m} \left( \frac{\sum_{k=0}^{n-1} \nu_{k+1}}{\mu_k} \right)^{-1} = \Pi^{(n)}_{l_{n+1}}.$$

It is easy to deduce that in order to prove (27), we have to prove that $\Pi^{(n)}_{l_{n+1}} \to y$ as $n \to \infty$. We always choose $l_n$ such that $y \in [\Pi^{(n)}_{l_n}, \Pi^{(n)}_{l_{n+1}})$. Therefore, it is enough to show that $||P^{(n)}|| \to 0$ as $n \to \infty$.

From inequality (23), we see that for each $1 \leq n$,

$$\frac{\nu_n}{\mu_{n-1}} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} = \frac{f_n''(1)}{f_n'(1)} \left( \frac{\mu_n}{\nu_n} \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \leq c(1 + f_n'(1)) \left( \frac{\mu_n}{\nu_n} \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} = c \left( \frac{\mu_n}{\nu_n} \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} + c \left( \frac{\mu_{n-1}}{\nu_n} \sum_{k=0}^{n-2} \frac{\nu_{k+1}}{\mu_k} \right)^{-1}.$$

Since we are in the critical regime and $\nu_n \geq 0$ for all $n \geq 1$, both summands in the right-hand side of the last equality go to zero as $n \to \infty$. In other words, given $\epsilon > 0$ there exists $N \geq 1$ such that

$$\sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \leq \epsilon,$$  \hspace{1cm} (29)

On the other hand, by criticality, for any fixed $m \leq N$, there is a $M_m \in \mathbb{N}$ such that

$$\frac{\nu_{m+1}}{\mu_m} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \leq \epsilon,$$  \hspace{1cm} (30)

We define $M = N \vee \max\{M_m : m \leq N\}$. Then, by (29) and (30), for any $n \geq M$

$$||P^{(n)}|| = \max_{0 \leq m < n} \left\{ \frac{\nu_{m+1}}{\mu_m} \left( \sum_{k=0}^{n-1} \frac{\nu_{k+1}}{\mu_k} \right)^{-1} \right\} \leq \epsilon,$$

and the claim is true. \hfill \Box

Now, we present a result whose relevance will be notice in the proof of Proposition 4

**Lemma 4.** Let $Q$ be a varying environment satisfying condition (A) and $\{\hat{Z}_n : n \geq 0\}$ be a size-biased GWVE. Define

$$B_1^{(n)} = \int_0^\lambda \left( \mathbb{E} \left[ \exp \left\{ -s \frac{\hat{Z}_n}{a_n} \right\} \right] - \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \mathbb{E} \left[ \exp \left\{ -s A_{n,m} \frac{\hat{Z}_n}{a_n} \right\} \right] \right) \, ds,$$

$$B_2^{(n)} = \int_0^\lambda \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \left( \mathbb{E} \left[ \exp \left\{ -s A_{n,m} \frac{\hat{Z}_n}{a_n} \right\} \right] - \mathbb{E} \left[ \exp \left\{ -s \frac{\hat{Z}_{n-(m+1)}}{a_n} \right\} \right] \right) \, ds,$$

$$B_3^{(n)} = \int_0^\lambda \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \mathbb{E} \left[ \exp \left\{ -s \frac{\hat{Z}_{n-(m+1)}}{a_n} \right\} \right] \left( 1 - g \left( n, m, \frac{s}{a_n} \right) \right) \, ds,$$
where $U$ is an uniform random variable on $[0, 1]$ independent of $\dot{Z}^Q$. Then,

$$\limsup_{n \to \infty} B_1^{(n)} = \limsup_{n \to \infty} B_2^{(n)} = \limsup_{n \to \infty} B_3^{(n)} = 0.$$ 

Proof. We start with $B_1^{(n)}$. Recall the partition $P^{(n)} = \{\Pi_0^{(n)} < \Pi_1^{(n)} < \ldots < \Pi_{n-1}^{(n)} < \Pi_n^{(n)}\}$ given in the proof of Proposition $3$ and that $P(K_n = m) = \Pi_{n-m}^{(n)} - \Pi_{n-m-1}^{(n)}$. Then

$$b_1^{(n)}(s) := \mathbb{E}\left[ \exp\{-sU \frac{\dot{Z}^Q_n}{a_n}\}\right] - \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \mathbb{E}\left[ \exp\{-sA_{m,n} \frac{\dot{Z}^Q_n}{a_n}\}\right]$$

$$= \int_0^1 \mathbb{E}\left[ \exp\{-su \frac{\dot{Z}^Q_n}{a_n}\}\right] du - \sum_{m=0}^{n-1} (\Pi_{n-m}^{(n)} - \Pi_{n-m-1}^{(n)}) \mathbb{E}\left[ \exp\{-sA_{m,n} \frac{\dot{Z}^Q_n}{a_n}\}\right].$$

By decomposing $[0, 1]$ in the subintervals $[\Pi_{n-m-1}^{(n)}, \Pi_n^{(n)}]$, $m = 0, \ldots, n - 1$, we get

$$b_1^{(n)}(s) = \sum_{m=0}^{n-1} \int_{\Pi_{n-m-1}^{(n)}}^{\Pi_n^{(n)}} \mathbb{E}\left[ \exp\{-su \frac{\dot{Z}^Q_n}{a_n}\}\right] du - \sum_{m=0}^{n-1} \int_{\Pi_{n-m-1}^{(n)}}^{\Pi_n^{(n)}} \mathbb{E}\left[ \exp\{-sA_{m,n} \frac{\dot{Z}^Q_n}{a_n}\}\right] du.$$

Now, by Lemma $2$, the Laplace transform of $\dot{Z}^Q_n$ can be expressed in terms of $f_{0,n}'$. Since $x \mapsto f_{0,n}'(e^{-\lambda x})e^{-\lambda x}$ is a decreasing function ($f_{0,n}'$ is convex), and $u, A_{m,n} \in [\Pi_{n-m-1}^{(n)}, \Pi_n^{(n)}]$ for $m = 0, \ldots, n - 1$, we deduce

$$|b_1^{(n)}(s)| \leq \frac{1}{\mu_n} \sum_{m=0}^{n-1} \int_{\Pi_{n-m-1}^{(n)}}^{\Pi_n^{(n)}} \left| f_{0,n}' \left( e^{-su/a_n^Q} \right) e^{-su/a_n^Q} - f_{0,n}' \left( e^{-sA_{m,n}/a_n^Q} \right) e^{-sA_{m,n}/a_n^Q} \right| du$$

$$\leq \frac{1}{\mu_n} \sum_{m=0}^{n-1} (\Pi_{n-m}^{(n)} - \Pi_{n-m-1}^{(n)})$$

$$\times \left( f_{0,n}' \left( e^{-sPi_{n-m-1}/a_n^Q} \right) e^{-sPi_{n-m-1}/a_n^Q} - f_{0,n}' \left( e^{-sPi_{n-m}/a_n^Q} \right) e^{-sPi_{n-m}/a_n^Q} \right).$$

The last sum can be bounded by the norm of the partition multiplied by a telescopic sum with $\Pi_0^{(n)} = 0$ and $\Pi_n^{(n)} = 1$. Therefore

$$|b_1^{(n)}(s)| \leq \frac{1}{\mu_n} ||P^{(n)}|| \left( f_{0,n}'(1) - f_{0,n}' \left( e^{-s/a_n^Q} \right) e^{-s/a_n^Q} \right) \leq \frac{1}{\mu_n} f_{0,n}'(1) ||P^{(n)}|| = ||P^{(n)}||.$$

Since the norm of the partition goes to zero as $n \to \infty$ (see the proof of Proposition $3$), we get the result for $B_1^{(n)}$,

$$\limsup_{n \to \infty} |B_1^{(n)}| \leq \limsup_{n \to \infty} \int_0^\lambda |b_1^{(n)}(s)| ds \leq \limsup_{n \to \infty} \lambda ||P^{(n)}|| = 0.$$

Now we deal with $B_2^{(n)}$. By Lemma $2$, the Laplace transform of $\dot{Z}^Q_n$ and $\dot{Z}^{Q\phi_{m+1}}_{n-m-1}$ can be expressed in terms of $f_{0,n}'$ and $f_{m+1,n}'$, respectively. Then,

$$b_2^{(n,m)}(s) := \mathbb{E}\left[ \exp\{-sA_{m,n} \frac{\dot{Z}^Q_n}{a_n}\}\right] - \mathbb{E}\left[ \exp\{-sA_{m,n} \frac{\dot{Z}^{Q\phi_{m+1}}_{n-m-1}}{a_n}\}\right]$$

$$= \frac{1}{\mu_n} f_{0,n}' \left( e^{-sA_{m,n}/a_n^Q} \right) e^{-sA_{m,n}/a_n^Q} - \frac{\mu_{m+1}}{\mu_n} f_{m+1,n}' \left( e^{-s/a_n^Q} \right) e^{-s/a_n^Q}.$$
Using first the Fundamental Theorem of Calculus and then the Mean Value Theorem in
the functions $f_{0,n}$ and $f_{m+1,n}$, we deduce that

$$
\int_0^\lambda b_2^{(n,m)}(s) \, ds \\
= \frac{1}{\mu_n} \left( \frac{\alpha_n^Q}{A_{n,m}} \left( f_{0,n}(1) - f_{0,n} \left( e^{-\lambda A_{n,m}/a_n^Q} \right) \right) - \mu_{m+1} a_n^Q \left( f_{m+1,n}(1) - f_{m+1,n} \left( e^{-\lambda/a_n} \right) \right) \right) \\
= \frac{1}{\mu_n} \left( \frac{\alpha_n^Q}{A_{n,m}} \left( 1 - e^{-\lambda A_{n,m}/a_n^Q} \right) f'_{0,n}(\xi) - \mu_{m+1} a_n^Q \left( 1 - e^{-\lambda/a_n} \right) f'_{m+1,n}(\eta) \right),
$$

where $\xi \in \left( e^{-\lambda A_{n,m}/a_n^Q}, 1 \right)$ and $\eta \in \left( e^{-\lambda/a_n}, 1 \right)$. The fact that $f'_{0,m}$ and $f'_{m,n}$ are
increasing functions and that

$$
\mu_{m+1} f'_{m+1,n} \left( e^{-\lambda/a_n} \right) \geq \prod_{l=1}^{m+1} f'_{l,m+1} \left( e^{-\lambda/a_n} \right)
$$

implies

$$
B_2^{(n)} = \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \int_0^\lambda b_2^{(n,m)}(s) \, ds \in \left[ \hat{B}_2^{(n)}, \tilde{B}_2^{(n)} \right],
$$

where

$$
\hat{B}_2^{(n)} = \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \left( \frac{\alpha_n^Q}{A_{n,m}} \left( 1 - e^{-\lambda A_{n,m}/a_n^Q} \right) f'_{0,n}(e^{-\lambda A_{n,m}/a_n^Q}) \right. \\
- \left. \mu_{m+1} a_n^Q \left( 1 - e^{-\lambda/a_n} \right) f'_{m+1,n}(1) \right)
$$

and

$$
\tilde{B}_2^{(n)} = \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \left( \frac{\alpha_n^Q}{A_{n,m}} \left( 1 - e^{-\lambda A_{n,K_n}/a_n^Q} \right) f'_{0,n}(1) - a_n^Q \left( 1 - e^{-\lambda/a_n} \right) f'_{0,n}(e^{-\lambda/a_n}) \right)
$$

Recall that $0 \leq A_{n,K_n} \leq 1$ and $a_n^Q \to \infty$ as $n \to \infty$, then as $n \to \infty$

$$
a_n^Q \left( 1 - e^{-\frac{\lambda}{a_n^Q}} \right) \to \lambda,
$$

and

$$
\frac{a_n^Q}{A_{n,K_n}} \left( 1 - e^{-\lambda A_{n,K_n}/a_n^Q} \right) \to \lambda,
$$

and

$$
\frac{f'_{0,n}(e^{-\lambda/a_n})}{f'_{0,n}(1)} \to 1
$$

a.s.
By Dominated Convergence Theorem, we have that $\hat{B}_2^{(n)} \to 0$ and $\tilde{B}_2^{(n)} \to 0$ as $n \to \infty$. Therefore, $B_2^{(n)}$ has the same behaviour.

Finally, we deal with $B_3^{(n)}$. Given an $\epsilon > 0$, by Lemma 3 there exists $M > 0$ such that for $n \geq M$, $0 \leq s \leq \lambda$ and $0 \leq m < n$

$$\left| g \left( n, m, \frac{s}{a_n^Q} \right) - 1 \right| \leq \epsilon.$$ 

Hence, for $n \geq M$,

$$|B_3^{(n)}| \leq \int_0^\lambda \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) \left| g \left( n, m, \frac{s}{a_n^Q} \right) - 1 \right| ds \leq \epsilon \lambda.$$ 

Since $\epsilon$ is arbitrary, we get that $\limsup_{n \to \infty} B_3^{(n)} = 0$. □

For the proof of Proposition 4, we need the following two lemmas, the reader can find them in [9, Lemma 3.1 and Lemma 3.2]. The first lemma compares the generating functions of two variables with the generating functions of their size-biased transforms. The second lemma is similar to Grönwall’s Lemma.

**Lemma 5.** Let $X$ and $W$ be two non-negative random variables with mean $\mu$. Let $F$ and $G$ be functions such that $E \left[ e^{-\lambda X} \right] = E \left[ e^{-\lambda X} \right] F(\lambda)$ and $E \left[ e^{-\lambda W} \right] = E[e^{-\lambda W}]G(\lambda)$, where $\hat{X}$ and $\hat{W}$ are the size-biased transforms of $X$ and $W$. Then,

$$\left| E \left[ e^{-\lambda X} \right] - E \left[ e^{-\lambda W} \right] \right| \leq \mu \left| \int_0^\lambda (F(s) - G(s)) ds \right|, \quad \lambda \geq 0.$$ 

**Lemma 6.** Suppose that a non-negative bounded function $F$ on $[0, \infty)$ and a constant $c > 0$ satisfying

$$F(\lambda) \leq c \int_0^1 du \int_0^\lambda F(us) ds,$$  

for $\lambda \geq 0$, then, $F \equiv 0$.

Finally, we present the last proof in this manuscript.

**Proof of Proposition 4.** We define the bounded function

$$M(\lambda) = \limsup_{n \to \infty} \mathbb{E}[e^{-\lambda \hat{Y}}] - \mathbb{E} \left[ \exp \left\{ -\lambda \frac{\hat{Z}_n^Q - 1}{a_n^Q} \right\} \right], \quad \text{for } \lambda \geq 0.$$ 

We will use Lemma 5 with $X = \hat{Y}$ and $W = (\hat{Z}_n^Q - 1)/a_n^Q$. Since $Y$ is an exponential random variable and (2) holds, we get

$$\mathbb{E}[\hat{Y}] = 2 = \mathbb{E} \left[ \frac{\hat{Z}_n^Q - 1}{a_n^Q} \right].$$ 

Thanks to characterisation (10), we see that $F(\lambda) = \mathbb{E}[e^{-\lambda U \hat{Y}}]$, where $U$ is an uniform variable on $[0, 1]$ independent of $Y$. From Proposition 1 it turns out that $(\hat{Z}_n^Q - 1)/a_n^Q$ is a $(\hat{Z}_n^Q - 1)/a_n^Q$-transform of $(\hat{Z}_n^Q - 1)/a_n^Q$. Then, by Proposition 2 we have

$$G(\lambda) = \sum_{m=0}^{n-1} \mathbb{P}(K_n = m) g \left( n, m, \frac{\lambda}{a_n^Q} \right) \mathbb{E} \left[ \exp \left\{ -\lambda \frac{\hat{Z}_n^Q - (m+1)}{a_n^Q} \right\} \right].$$
where \( g \) is given in (13). Hence, by Lemma 5 and the triangle inequality,

\[
\left| \mathbb{E} \left[ e^{-\lambda \hat{Y}} \right] - \mathbb{E} \left[ \exp \left\{ -\lambda \frac{\hat{Z}_n^Q - 1}{a_n^Q} \right\} \right] \right| \\
\leq 2 \int_0^\lambda \left( \mathbb{E} \left[ e^{-sU \hat{Y}} \right] - \sum_{m=0}^{n-1} \mathbb{P}(K_m = m) g \left( n, m, \frac{s}{a_n^Q} \right) \mathbb{E} \left[ \exp \left\{ -s \frac{\hat{Z}_n^Q - 1}{a_n^Q} \right\} \right] \right) ds \\
\leq 2 \left( |B_1^{(n)}| + |B_2^{(n)}| + |B_3^{(n)}| + |B_4^{(n)}| \right),
\]

where \( B_1^{(n)}, B_2^{(n)} \) and \( B_3^{(n)} \) are defined in Lemma 4 and

\[
B_4^{(n)} = \int_0^\lambda \left( \mathbb{E} \left[ e^{-sU \hat{Y}} \right] - \mathbb{E} \left[ \exp \left\{ -s \frac{\hat{Z}_n^Q}{a_n^Q} \right\} \right] \right) ds,
\]

with \( U \) a uniform random variable on \([0,1]\) independent of \( \hat{Y} \) and \( \hat{Z}_n^Q \). Then, by Lemma (4) and the Dominated Convergence Theorem, we obtain

\[
M(\lambda) \leq 2 \limsup_{n \to \infty} \int_0^\lambda \left| \mathbb{E} \left[ e^{-sU \hat{Y}} \right] - \mathbb{E} \left[ \exp \left\{ -s \frac{\hat{Z}_n^Q}{a_n^Q} \right\} \right] \right| ds \\
\leq 2 \int_0^\lambda \int_0^1 M(us) du ds.
\]

By Lemma 6, \( M \equiv 0 \) which implies that \( \hat{Z}_n^Q / a_n^Q \) converges weakly to \( \hat{Y} \). \( \square \)

**Acknowledgements**

N.C.-T. acknowledges support from CONACyT-MEXICO grant no. 636133. S.P. is a Newton International Fellow Alumnus (AL191032). S.P. would like to thank Yan-Xia Ren and Zhenyao Sun for a discussion of their paper. We are grateful to Juan Carlos Pardo for his careful reading of an earlier version of this manuscript. This is a research supported by UNAM-DGAPA-PAPIIT grant no. IA103220.

**References**

[1] N. Bhattacharya and M. Perlman. Time-inhomogeneous branching processes conditioned on non-extinction. *arXiv preprint arXiv:1703.00337*, 2017.

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

[3] J. Geiger. A new proof of Yaglom’s exponential limit law. In *Mathematics and Computer Science*, pages 245–249. Springer, 2000.

[4] M. González, G. Kersting, C. Minuesa, and I. del Puerto. Branching processes in varying environment with generation-dependent immigration. *Stochastic Models*, 35(2):148–166, 2019.

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

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

[7] G. Kersting and V. A. Vatutin. *Discrete time branching processes in random environment*. Wiley Online Library, 2017.

[8] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of L log L criteria for mean behavior of branching processes. *The Annals of Probability*, pages 1125–1138, 1995.

[9] Y.-X. Ren, R. Song, and Z. Sun. A 2-spine decomposition of the critical Galton-Watson tree and a probabilistic proof of Yaglom’s Theorem. *Electron. Commun. Probab.*, 23(42):12, 2018.

[10] A. M. Yaglom. Certain limit theorems of the theory of branching random processes. In *Doklady Akad. Nauk SSSR (NS)*, volume 56, page 3, 1947.
