Reinforced Galton-Watson processes II: Large-time behaviors

Jean Bertoin^{*} Bastien Mallein[†]

February 8, 2025

Abstract

Reinforced Galton-Watson processes have been introduced in [4] as population models with non-overlapping generations, such that reproduction events along genealogical lines can be repeated at random. We investigate here some of their sample path properties such as asymptotic growth rates and survival, for which the effects of reinforcement on the evolution appear quite strikingly.

Keywords: Galton-Watson process, stochastic reinforcement, growth rate, survival.

MSC subject classifications: 60J80; 60J85.

1 Introduction and main results

Reinforcement is a fundamental concept in many sciences, including notably Behavioral Psychology (as a key part of Skinner's behavioral theory of learning [10]) and Artificial Intelligence [7], where this terminology refers to methods that increase the likelihood of certain evolutions in both natural and artificial systems. It has been introduced in the setting of stochastic processes by Coppersmith and Diaconis in an influential unpublished article, where these authors modified step after step the dynamics of a random walk on a graph, in such a way that transitions which have already been often made in the past are more likely to occur again in the future.

It is a recent development in Demography that a reinforcement feature can be detected in the genealogy of human populations. In many human populations, the fertility levels of parents and children are positively correlated [3, 9]; this may be explained e.g. by socioeconomic, cultural or inherited factors. This observation, sometimes referred to as Intergenerational Transmission of Fertility, invalidates population models in which individuals reproduce independently one from the others, and provides an incentive for the study of reinforced versions which incorporate a dependency structure of fertility levels along lineages.

The Galton-Watson branching process is a population model in which every individual reproduces independently of the others according a probability distribution $(\nu(k))_{k>0}$; any individual has probability $\nu(k)$ of having k children at the

^{*}Institute of Mathematics, University of Zurich, Switzerland, jean.bertoin@math.uzh.ch. [†]Institut de Mathématiques de Toulouse, Université de Toulouse, France, bastien.mallein@math.univ-toulouse.fr.

next generation. This population model was originally introduced to study the evolution of human demography, namely the possible disappearance of family names. It is well-known that this process almost surely dies out if and only if the mean number of children of a given individual satisfies $m_{\rm GW} := \sum k\nu(k) \leq 1$ and $\nu(1) < 1$. We write

$$S_{\boldsymbol{\nu}} \coloneqq \{k \ge 1 : \nu(k) > 0\} \tag{1.1}$$

for the strictly positive part of the support of the reproduction law ν , and assume throughout this work that S_{ν} has a maximal element $\mathbf{k}_* \geq 2$. In other words, \mathbf{k}_* is the maximal possible number of children; observe that $\mathbf{k}_* \geq 2$ immediately implies $\nu(1) < 1$.

We recently introduced in [4] a reinforced version of Galton-Watson processes that involves random repetitions of reproduction events and depends on a parameter $q \in (0, 1)$. The evolution can be depicted as follows. Each individual at any generation $n \geq 1$ picks a forebear uniformly at random on its ancestral lineage, independently of the other individuals. Then either with probability q, this individual begets the same number of children as the selected forebear, or with complementary probability 1 - q, the number of its children is an independent sample from the reproduction law $\boldsymbol{\nu}$. Reproduction events that occurred at early stages of the process are thus more likely to be repeated in the future, as they are common to a larger number of genealogical lines, and informally speaking, the reinforced process thus keeps some memory of its past. Clearly, the usual Galton-Watson evolution corresponds to the boundary case q = 0 without reinforcement (i.e. without memory).

Let $Z = (Z(n))_{n\geq 0}$ be a reinforced Galton-Watson process with reproduction law $\boldsymbol{\nu}$ and reinforcement parameter $q \in (0, 1)$, started from a single ancestor. More precisely, Z(n) denotes the size of the *n*-th generation, and for $k \geq 1$, we write \mathbb{P}_k for the distribution of Z conditioned on Z(1) = k. Let us recall a simplified version of our main result about the asymptotic behavior of the averaged population size for large generations; see [4, Theorem 7.3]. As $n \to \infty$, we have

$$\mathbb{E}_{\mathbf{k}_{*}}(Z(n)) \sim \frac{m_{\boldsymbol{\nu},q}^{n}}{q + (1-q)\nu(\mathbf{k}_{*})},\tag{1.2}$$

whereas

$$\mathbb{E}_k(Z(n)) = o\left(m_{\nu,q}^n\right) \quad \text{for any } k \in S_{\nu} \text{ with } k \neq k_*, \tag{1.3}$$

where

$$m_{\boldsymbol{\nu},q} \coloneqq \frac{q}{\int_0^{1/\mathbf{k}_*} \Pi(t) \mathrm{d}t},\tag{1.4}$$

and

$$\Pi(t) \coloneqq \prod_{k \in S_{\nu}} (1 - tk)^{\nu(k)(1 - q)/q}, \qquad 0 \le t \le 1/k_*.$$
(1.5)

We now present the motivations for the present work and our two main results. Recall that for the usual Galton-Watson process $Z_{\rm GW}$ with reproduction law ν , not only the mean population size at the *n*-th generation is¹

$$\mathbb{E}(Z_{\rm GW}(n)) = m_{\rm GW}^n \qquad \text{for all } n \ge 1,$$

¹This has to be compared and contrasted with (1.2) and (1.3); recall also from [4, Proposition 7.1(i)] that $m_{\rm GW} = \lim_{q \to 0+} m_{\nu,q}$.

but in the supercritical case where $m_{\rm GW} > 1$, there is the much more precise pathwise convergence

$$\lim_{n \to \infty} m_{\rm GW}^{-n} Z_{\rm GW}(n) = W_{\rm GW} \qquad \text{a.s. and in } L^1.$$

Furthermore, in the original Galton-Watson process, the random variable $W_{\rm GW}$ vanishes exactly on the event of eventual extinction of $Z_{\rm GW}$. Our first main result shows that reinforcement may change dramatically this property.

In this direction, we need to introduce some more notation. Consider the sub-process $Z_* = (Z_*(n))_{n\geq 0}$ of the reinforced Galton-Watson process which results by suppressing every progeny (together with its descent) that has size strictly less than k_* . Since by construction, all the forebears of an individual in this sub-process have exactly k_* children, Z_* is a true Galton-Watson process under \mathbb{P}_{k_*} (recall that the latter denotes the distribution of the reinforced Galton-Watson process when the ancestor has the maximal number k_* of children). Specifically, all individuals in this sub-process, with the exception of the ancestor, have either k_* children with probability $q + (1 - q)\nu(k_*)$, or 0 child with complementary probability, independently one of each other's. We write

$$m_{*,q} \coloneqq \mathbf{k}_*(q + (1 - q)\nu(\mathbf{k}_*)) \tag{1.6}$$

for the mean reproduction number of Z_* and stress that $m_{*,q} \leq m_{\nu,q}$; see [4, Proposition 7.2].

Theorem 1.1. We have:

(i) If $m_{*,q} \leq 1$, then

$$\lim_{n \to \infty} m_{\nu,q}^{-n} Z(n) = 0 \qquad in \ \mathbb{P}_{\mathbf{k}_*} \text{-} probability.$$

(*ii*) If $m_{*,q} > 1$, then

$$\lim_{n \to \infty} m_{\boldsymbol{\nu}, q}^{-n} Z(n) = \lim_{n \to \infty} m_{*, q}^{-n} Z_*(n) \coloneqq W_* \qquad in \ L^1(\mathbb{P}_{\mathbf{k}_*}).$$

Moreover the events $\{W_* = 0\}$ and $\{\exists n \ge 1 : Z_*(n) = 0\}$ coincide \mathbb{P}_{k_*} -a.s.

Remark 1.2. Comparing Theorem 1.1(i) with (1.2) shows that the main contribution to the mean population size $\mathbb{E}_{k_*}(Z(n))$ when $m_{*,q} \leq 1$ is actually due to exceptional events for which the population is much larger than expected, i.e. $Z(n) \gg m_{\nu,q}^n$. When $m_{*,q} > 1$, this is no longer the case. However, we expect the survival event of Z to be strictly larger than $\{W_* \neq 0\}$, more precisely, that $\mathbb{P}_{k_*}(\inf_n Z_n > 0, W_* = 0) > 0.$

Remark 1.3. Consider in this remark only the case when the reproduction law ν has unbounded support. If we define a reinforced Galton-Watson process Z with these parameters, then for all $k \in \mathbb{Z}_+$, we can define $Z^{(k)}$ counting individuals in the reinforced process such that none of their ancestor had k ancestors or more. Setting

$$\forall 1 \le j \le k, \nu^{(k)}(j) = \nu(j) \text{ and } \nu^{(k)}(0) = 1 - \nu([1,k]),$$

we observe that $Z^{(k)}$ is a $(\boldsymbol{\nu}^{(k)}, q)$ reinforced Galton-Watson process, and that

$$m_{\boldsymbol{\nu}^{(k)},q} \ge k(q + (1-q)\boldsymbol{\nu}(k)) \ge kq.$$

As a result, we deduce from Theorem 1.1 that $\liminf_{n\to\infty} Z_n/A^n > 0$ with positive probability for all A > 0, and $\mathbb{E}(Z_n)$ grows super-exponentially fast as $n \to \infty$. This justifies our choice of only considering compactly supported measures ν .

Our second main result concerns survival probabilities. For usual Galton-Watson processes, $Z_{\rm GW}$ survives with strictly positive probability if and only it is super-critical, that is $m_{\rm GW} > 1$. In the reinforced setting, Theorem 1.1(i) entails that Z becomes eventually extinct a.s. whenever $m_{\nu,q} \leq 1$; however we conjectured in [4] that there should exist reproduction laws ν and reinforcement parameters q such that $m_{\nu,q} > 1$ and nonetheless, the reinforced Galton-Watson process becomes extinct eventually almost surely.

In the converse direction, observe first that if $q\mathbf{k}_* \geq 1$, then the true Galton-Watson process Z_* is supercritical, and the reinforced Galton-Watson process Z obviously survives with strictly positive probability. In particular, observe that if $\boldsymbol{\nu}$ had unbounded support, then the reinforced Galton-Watson process would always survive with positive probability. Assuming now that $q\mathbf{k}_* < 1$, we point at the following sufficient condition for survival of Z.

Theorem 1.4. Suppose that $qk_* < 1$, and that

$$\sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{1-qj} > 1.$$
(1.7)

Then the reinforced Galton-Watson process survives forever with strictly positive probability, that is for any $k \in S_{\nu}$, we have

$$\mathbb{P}_k(Z(n) \ge 1 \text{ for all } n \ge 1) > 0.$$

Remark 1.5. (i) Condition (1.7) obviously holds whenever there exists some $j \in S_{\nu}$ such that

$$\frac{(1-q)j\nu(j)}{1-qj} > 1.$$

This inequality can be written as $j(q + (1 - q)\nu(j)) > 1$, and we observe that the left-hand side is the mean reproduction number of the true Galton-Watson sub-process that results from the reinforced one by suppressing every progeny (together with its descent) with size different from j. That is, the true Galton-Watson sub-process is supercritical, hence it survives with strictly positive probability. A fortiori the same holds for the reinforced Galton-Watson process.

(ii) In the same vein, (1.7) also clearly holds when the usual Galton-Watson process $Z_{\rm GW}$ is supercritical, viz. $m_{\rm GW} > 1$. In the converse direction, it is easy to construct reproduction laws $\boldsymbol{\nu}$ with

$$m_{\rm GW} < 1 < \sum_{j=1}^{\mathbf{k}_*} \frac{(1-q)j\nu(j)}{1-qj}$$

This provides further examples of reinforced process Z that may survive in situations where usual Galton-Watson process Z_{GW} becomes eventually extinct a.s. (iii) The sufficient condition (1.7) corresponds exactly to the necessary and sufficient condition for the survival of the branching process \tilde{Z} in which parents beget children according to the law $\boldsymbol{\nu}$ with probability 1-q, or beget exactly as many children as their own parent with probability q. The process \tilde{Z} being a multitype Galton-Watson process, this claim is easily checked by studying the Perron-Frobenius eigenvalue of the mean reproduction matrix $(qj\delta_{i,j} + (1-q)j\nu(j))_{i,j}$, where $\delta_{\cdot,\cdot}$ is the Kronecker symbol, see [2, Chapter 4].

Theorem 1.1 implies that $m_{\nu,q} > 1$ is a necessary condition for survival. Theorem 1.4 gives a different sufficient condition for survival, namely (1.7). As a consequence, (1.7) entails $m_{\nu,q} > 1$; however, we have not been able to check this fact by purely analytic considerations. Despite these two conditions being quite close to one another, the two are not identical as is being illustrated in Figure 1.



Figure 1: Phase diagram of the survival of a reinforced Galton-Watson process with parameter $(\boldsymbol{\nu}_p, q)$, where we have set $\boldsymbol{\nu}_p = (1-4p)\delta_0 + p(\delta_1 + \delta_2 + \delta_3 + \delta_4)$, for $q \in [0, 0.25]$ and $p \in [0, 0.1]$. The blue line corresponds to the set of parameters such that $m_{\boldsymbol{\nu}_p,q} = 1$, the orange one to parameters such that $\sum (1-q)j\boldsymbol{\nu}_p(j)/(1-qj) = 1$. The grey domain corresponds to the set of parameters for which $m_{\boldsymbol{\nu}_p,q} > 1$ but (1.7) does not hold. Note that $m_{\rm GW} > 1$ for p > 0.1 and that $q\mathbf{k}_* = 4q \ge 1$ for $q \ge 0.25$, which explains the boundary points of the curves.

The rest of this work is organized as follows. Section 2 dwells on the observation that reinforced Galton-Watson processes can be viewed as multitype branching processes, where the type of an individual is defined by the sequence of the numbers of children of its forebears. This enables us to recover the branching property, of course at the cost of working with a rather large space of types. In Section 3, we establish some bounds for such multitype branching processes, which follow rather directly from (1.2) and (1.3). These bounds are then used in Section 4 to establish Theorem 1.1.

Section 5 is in turn devoted to the proof of Theorem 1.4; let us briefly sketch our approach which is fully independent of [4]. We start in Subsection 5.1 by constructing a natural nonnegative martingale M for the reinforced Galton-Watson process, which starts from $M_0 = 1$ and vanishes as soon as Z becomes extinct. This leads us to investigate whether M is uniformly integrable, since in that case, the reinforced Galton-Watson process obviously survives with strictly positive probability. In this direction, we use M to define a tilted probability law in Subsection 5.2, and describe the evolution of the reinforced Galton-Watson process under the tilted law in terms of the so-called spinal decomposition. Next, in Subsection 5.3, we interpret the evolution of the types along the spine in terms of a generalized Pólya urn process. This permits us to determine the asymptotic behavior of those types. The proof of Theorem 1.4 can then be completed in Subsection 5.4 by applying classical arguments involving the spinal decomposition and Durrett's criterion for the uniform integrability of a nonnegative martingale.

2 Reinforced Galton-Watson processes as multitype branching processes

The reproduction law of an individual in a reinforced Galton-Watson process depends on its entire ancestral lineage. Different individuals may partly share the same ancestral lineage, and then their respective evolutions are not independent. The Markov and the branching properties are therefore lost in this process, but can nonetheless be recovered by endowing each individual with a type that records the reproduction numbers of its forebears.

To start with, we briefly adapt to our setting the Ulam–Harris–Neveu framework which enables to encode any population model with non-overlapping generations started from a single ancestor by its the genealogical tree \mathcal{T} . Since in our model, an individual has never more than k_* children, we use the space of finite (possibly empty) sequences in $\{1, \ldots, k_*\}$,

$$\mathcal{U} \coloneqq \bigcup_{\ell \ge 0} \{1, \dots, k_*\}^\ell,$$

as the universal tree that contains \mathcal{T} .

The empty sequence \emptyset is assigned to the ancestor of the population and viewed as the root of the genealogical tree \mathcal{T} . An individual at the ℓ -th generation is represented by some $u = (u_1, \ldots, u_\ell) \in \mathcal{T}$, and if this individual has $k \ge 1$ children, the latter are represented in turn by the sequences $uj \coloneqq (u_1, \ldots, u_\ell, j)$ for $j = 1, \ldots, k$. As in our setting, an individual can have k children only for $k \ge 0$ such that $\nu(k) > 0$, the outer-degree d(u) of any individual $u \in \mathcal{T}$ always belongs to the support of the reproduction law ν ; moreover for any $u \in \mathcal{T}$, d(u) = 0 if and only if u is a leaf of \mathcal{T} . For $u = (u_1, \ldots, u_\ell) \in \mathcal{U}$ and $j \le |u| = \ell$, we write $u(j) = (u_1, \ldots, u_j)$ for the prefix consisting of the first j letters of u, which corresponds to the ancestor of u alive at generation j.

We further assign types to individuals that record outer-degree sequences along ancestral lines. Recall from (1.1) the notation S_{ν} for the strictly positive part of the support of ν . We call type a (possibly empty) finite sequence $\mathbf{t} = (d_1, \ldots, d_\ell)$ in S_{ν} , where d_j should be thought of as the number of children of the forebear *j*-generations backwards. We agree to write simply d_1 instead of (d_1) when the type has just one element. We denote the space of types by

$$\mathcal{W} \coloneqq \bigcup_{\ell \ge 0} S^{\ell}_{\boldsymbol{\nu}}$$

Types are associated to vertices of \mathcal{T} using the following (deterministic) algorithm. If an individual $u \in \mathcal{T}$ has type $\mathbf{t} = (d_1, \ldots, d_\ell)$ and has $k \ge 1$ children

at the next generation, then all its children receive the type $k\mathbf{t} = (k, d_1, \ldots, d_\ell)$. The assignation of types is thus completely determined once the initial type $\mathbf{t}(\emptyset)$ of the ancestor and the genealogical tree \mathcal{T} are known. The type $\mathbf{t}(\emptyset)$ of the ancestor may be either empty or a sequence of positive length, $|\mathbf{t}(\emptyset)| = \ell \ge 1$. In the latter case, say $\mathbf{t}(\emptyset) = (d_1, \ldots, d_\ell)$, we may imagine that the ancestor has also forebears before the origin of time, at generations $-1, \ldots, -\ell$. In all cases, the length $|\mathbf{t}(u)|$ of the type of an individual $u \in \mathcal{T}$ is at least as large as its generation |u|.

Counting occurrences of each $d \in S_{\nu}$ in types is important in our analysis. This incites us to introduce the space \mathcal{M}_{ν} of finite integer-valued measures on S_{ν} and define for all $\mathbf{t} = (d_1, \ldots, d_{\ell}) \in \mathcal{W}$

$$\boldsymbol{\varpi}_{\boldsymbol{t}} = \sum_{j=1}^{\ell} \delta_{d_j}$$

We simply call $\boldsymbol{\varpi}_t$ the ancestral reproduction measure of an individual of type t. Plainly, if $u \in \mathcal{T}$ has $k \geq 1$ children at the next generation, then all its children have the same ancestral reproduction measure $\boldsymbol{\varpi}_{t(u)} + \delta_k$. The assignation of ancestral reproduction measures to individuals is also completely determined by the ancestral reproduction measure $\boldsymbol{\varpi}_{t(\emptyset)}$ of the ancestor and the genealogical tree \mathcal{T} .

This setting enables us to view a reinforced Galton-Watson process as a multitype branching process; see [5] for background. Recall that $q \in (0, 1)$ is a reinforcement parameter and ν a probability vector on $\{0, 1, \ldots, k_*\}$. We just need to specify the reproduction law π_t of an individual as a function of its type t. First, when this type is empty (which can only be the case for the ancestor), we agree² that the individual has almost surely k_* children all with type k_* . Next, when the type t is non-empty, the probability that this individual begets no child equals

$$\pi_{\boldsymbol{t}}(0) = (1-q)\nu(0),$$

and the probability that it begets exactly $k \ge 1$ children all of type kt and no further children equals

$$\pi_{\boldsymbol{t}}(k) = (1-q)\nu(k) + q\frac{\varpi_{\boldsymbol{t}}(k)}{|\boldsymbol{t}|}, \qquad k \in S_{\boldsymbol{\nu}},$$

where the length |t| of the type coincides with the total mass of its ancestral reproduction measure $\boldsymbol{\varpi}_t$. In particular, $\boldsymbol{\varpi}_t/|t|$ is the empirical distribution of the numbers of children of the forebears of an individual of type t.

For every tree \mathcal{T} and vertex $u \in \mathcal{T}$, we write \mathcal{T}_u for the subtree that stems from u. Specifically a vertex v belongs to \mathcal{T}_u if and only if uv (the concatenation of the words u and v) is a vertex of \mathcal{T} , and then we agree that the type of v in \mathcal{T}_u is the one of uv in \mathcal{T} . We also write

$$\mathcal{T}_{|\ell} \coloneqq \{ u \in \mathcal{T} : |u| \le \ell \}$$

for the restriction of the genealogical tree to the first ℓ generations.

²Recall that in the first part of this work, we are essentially concerned with the law \mathbb{P}_{k_*} when the ancestor has k_* children, who all have type k_* . If we were working under law \mathbb{P}_k for $k \in S_{\nu}$, then we would have imposed that an individual with the empty type has exactly k children all with type k.

For every type $t \in \mathcal{W}$, we write \mathcal{P}_t for the law of the genealogical tree \mathcal{T} of the multitype branching process with the reproductions laws described above conditioned on the ancestor having type t. We also write \mathcal{E}_t for the corresponding expectation. With this formalism at hand, we can now recover the branching property. The basic Markov-branching property for the multitype branching process can be stated as follows.

Lemma 2.1. Let T be a fixed tree of height ℓ (i.e. such that $\sup_{v \in T} |v| = \ell$) and an arbitrary initial type $\mathbf{t} \in \mathcal{W}$. We denote by $\mathbf{t}(u)$ the type of the vertex $u \in T$ that results from fixing $\mathbf{t}(\emptyset) = \mathbf{t}$. Then, under the conditioned probability measure $\mathcal{P}_{\mathbf{t}}(\cdot|\mathcal{T}_{|\ell} = T)$, the subtrees at level ℓ ($\mathcal{T}_u : u \in \mathcal{T}$ and $|u| = \ell$) are independent and each \mathcal{T}_u has the law $\mathcal{P}_{\mathbf{t}(u)}$.

We will also use in the sequel the following consequence of Lemma 2.1. For any initial type $t \in \mathcal{W}$, every vertex $u \in \mathcal{U}$, and every $d \in S_{\nu}$, under the conditional probability measure \mathcal{P}_t given that $u \in \mathcal{T}$, that u has type t' and outer degree d(u) = d in \mathcal{T} , the d subtrees $\mathcal{T}_{u1}, \ldots, \mathcal{T}_{ud}$ are independent and each has the law $\mathcal{P}_{dt'}$.

We further point at a useful domination property which should be intuitively obvious.

Lemma 2.2. Consider two types $\mathbf{t} = (d_1, \ldots, d_\ell)$ and $\mathbf{t}' = (d'_1, \ldots, d'_\ell)$ with the same length $\ell \geq 1$. We say that \mathbf{t} dominates \mathbf{t}' if there is a permutation σ of $\{1, \ldots, \ell\}$ such that

$$d_{\sigma(j)} \ge d'_j \qquad \text{for all } j = 1, \dots, \ell. \tag{2.1}$$

In that case, we can construct two random genealogical trees \mathcal{T} and \mathcal{T}' , the first with the law \mathcal{P}_t and the second with the law $\mathcal{P}_{t'}$, such that \mathcal{T} is a subtree of \mathcal{T}' .

Observe that the existence of a permutation such that (2.1) holds can be rephrased in terms of the empirical distribution of the number of children of the forebears for individuals with types t and t'. More precisely, (2.1) holds if and only if $\boldsymbol{\varpi}_t/|t|$ stochastically dominates $\boldsymbol{\varpi}_{t'}/|t'|$.

Proof. We immediately see by inspection of the reproduction laws of the ancestor as a function of its type, that we can couple the number of children Z(1) of the ancestor under law \mathcal{P}_t with that Z'(1) under $\mathcal{P}_{t'}$, such that $Z(1) \leq Z'(1)$. Then the types Z(1)t of the individuals at the first generation under \mathcal{P}_t are also dominated by the types Z'(1)t' under $\mathcal{P}_{t'}$. We can then complete the proof by induction using the branching property of Lemma 2.1.

3 Some useful bounds

The purpose of this section is to establish some inequalities that will be needed in the proof of Theorem 1.1. These follow readily from the framework develop in the preceding section and the key estimates (1.2) and (1.3). For every $n \ge 0$, the integer-valued measure

$$\mathcal{Z}_n \coloneqq \sum_{|u|=n, u \in \mathcal{T}} \delta_{t(u)}, \qquad n \ge 0,$$

that counts the number of individuals at the *n*-th generation as a function of their types, can be viewed as an enriched version of the reinforced Galton-Watson process for which the types of individuals are recorded. In particular, the total mass of $\|\mathcal{Z}_n\| \coloneqq \mathcal{Z}_n(\mathcal{W})$ under $\mathcal{P}_{\varnothing}$ coincides with Z(n) under \mathbb{P}_{k_*} .

We start with the following generalization of the bounds (1.2) and (1.3) to arbitrary fixed initial types. Recall the notation (1.4) and (1.6).

Lemma 3.1. Consider a type $\mathbf{t} = (d_1, \ldots, d_\ell)$ with length $|\mathbf{t}| = \ell$. As $n \to \infty$, we have

$$\mathcal{E}_{t}(\|\mathcal{Z}_{n}\|) \sim \left(\frac{m_{\boldsymbol{\nu},q}}{m_{*,q}}\right)^{\ell} m_{\boldsymbol{\nu},q}^{n}$$

when $d_j = k_*$ for all $j = 1, \ldots, \ell$, whereas $\mathcal{E}_t(||\mathcal{Z}_n||) = o(m_{\nu,q}^n)$ otherwise.

We mention that the approach in [4] would yield a much sharper estimate in the second case. However, calculations would be technically rather demanding and we prefer to establish the weaker result using only a much simpler argument, as this suffices for our purpose.

Proof. We consider first the case when type of the ancestor of the reinforced Galton-Watson process has length 1 and is given by $j \in S_{\nu}$ with $j \neq k_*$, and work under \mathcal{P}_j . We may imagine that the ancestor had a parent at generation -1 and j-1 siblings. This yields the identity

$$j\mathcal{E}_j\left(\|\mathcal{Z}_n\|\right) = \mathbb{E}_j(Z(n+1)),$$

and (1.3) entails that for any fixed $\ell \geq 0$,

$$\mathcal{E}_j\left(\|\mathcal{Z}_{\ell+n}\|\right) = o\left(m_{\nu,q}^n\right) \text{ as } n \to \infty.$$

The probability under \mathcal{P}_j that the individual $(1, \ldots, 1)$ at generation ℓ is present in the population and has type $(\mathbf{k}_*, \ldots, \mathbf{k}_*, j)$ (i.e. all its forebears at generations $0, \ldots, \ell - 1$ had \mathbf{k}_* children) is no less than $((1-q)\nu(\mathbf{k}_*))^{\ell} > 0$. It follows from the branching property that

$$\mathcal{E}_{(\mathbf{k}_*,\dots,\mathbf{k}_*,j)}\left(\|\mathcal{Z}_n\|\right) \le \left((1-q)\nu(\mathbf{k}_*)\right)^{-\ell}\mathcal{E}_j\left(\|\mathcal{Z}_{\ell+n}\|\right) = o\left(m_{\nu,q}^n\right)$$

An application of Lemma 2.2 completes the second claim of the statement. We next work under $\mathcal{P}_{\varnothing}$; recall from (1.2) that

$$\mathcal{E}_{\varnothing}\left(\left\|\mathcal{Z}_{n+\ell}\right\|\right) = \mathbb{E}_{\mathbf{k}_{*}}(Z(n+\ell)) \sim \frac{m_{\nu,q}^{n+\ell}}{q+(1-q)\nu(\mathbf{k}_{*})}.$$

We apply the branching property under \mathcal{P}_{\emptyset} at the ℓ -th generation. Recall that the number of individuals at generation ℓ of type (k_*, \ldots, k_*) (the sequence with ℓ terms, all equal to k_*) is

$$\mathcal{Z}_{\ell}(\mathbf{k}_*,\ldots,\mathbf{k}_*)=Z_*(\ell),$$

and that Z_* is a usual Galton-Watson process with mean reproduction number $m_{*,q}$ given that the ancestor has k_* children. In particular, the mean number of these individuals is

$$\mathbb{E}_{\mathbf{k}_{*}}(Z_{*}(\ell)) = \mathbf{k}_{*}m_{*,q}^{\ell-1}.$$
(3.1)

There are also $Z(\ell) - Z_*(\ell) \leq k_*^{\ell}$ individuals with types different from (k_*, \ldots, k_*) , and we know from the first part of this proof that as $n \to \infty$, the average descent of each of them at generation $n + \ell$ is $o(m_{\nu,q}^n)$. Therefore, the branching property yields

$$\mathcal{E}_{\varnothing}\left(\left\|\mathcal{Z}_{n+\ell}\right\|\right) \sim \mathbb{E}_{\mathbf{k}_{*}}(Z_{*}(\ell))\mathcal{E}_{(\mathbf{k}_{*},\ldots,\mathbf{k}_{*})}\left(\left\|\mathcal{Z}_{n}\right\|\right),$$

and we conclude that

$$\mathcal{E}_{(\mathbf{k}_*,\ldots,\mathbf{k}_*)}\left(\|\mathcal{Z}_n\|\right) \sim \frac{1}{q+(1-q)\nu(\mathbf{k}_*)} \frac{m_{*,q}}{\mathbf{k}_*} \left(\frac{m_{\boldsymbol{\nu},q}}{m_{*,q}}\right)^{\ell} m_{\boldsymbol{\nu},q}^n = \left(\frac{m_{\boldsymbol{\nu},q}}{m_{*,q}}\right)^{\ell} m_{\boldsymbol{\nu},q}^n,$$

sing (1.6) for the last equality.

using (1.6) for the last equality.

A similar argument also yields the following uniform bounds.

Lemma 3.2. There is some finite constant $c_{\nu,q}$ depending only on the reinforcement parameter and the reproduction law, such that for any type $t \in \mathcal{W}$ and any $n \ge 0$, one has

$$\mathcal{E}_{\boldsymbol{t}}\left(\|\mathcal{Z}_{n}\|\right) \leq c_{\boldsymbol{\nu},q} \left(\frac{m_{\boldsymbol{\nu},q}}{m_{*,q}}\right)^{|\boldsymbol{t}|} m_{\boldsymbol{\nu},q}^{n}.$$

Proof. Consider first the case of the empty type, so |t| = 0. Recall that

 $\mathcal{E}_{\varnothing}\left(\|\mathcal{Z}_n\|\right) = \mathbb{E}_{\mathbf{k}_n}(Z(n)),$

so by (1.2), we can find some finite constant c such that

$$\mathcal{E}_{\varnothing}\left(\|\mathcal{Z}_n\|\right) \le cm_{\boldsymbol{\nu},q}^n \quad \text{for all } n \ge 0$$

Next, as it was already discussed previously, by focussing on individuals at a given generation $\ell \geq 1$ whose forebears all had k_* children and applying the branching property of Lemma 2.1, we get the inequality

$$\mathcal{E}_{\varnothing}\left(\left\|\mathcal{Z}_{n+\ell}\right\|\right) \geq \mathbb{E}_{\mathbf{k}_{*}}(Z_{*}(\ell))\mathcal{E}_{(\mathbf{k}_{*},\ldots,\mathbf{k}_{*})}\left(\left\|\mathcal{Z}_{n}\right\|\right).$$

The identity (3.1) entails our claim whenever $t = (k_*, \ldots, k_*)$ with $c_{\nu,q} =$ $cm_{*,q}/k_*$. Finally, the general case for a type t follows from above by an application of Lemma 2.2, since t is dominated by (k_*, \ldots, k_*) .

When the usual Galton-Watson process Z_* is supercritical, we can also bound the second moment of $\|Z_n\|$ by combining these inequalities with a combinatorial argument.

Lemma 3.3. If $m_{*,q} > 1$, then there exists a constant $C_{\nu,q}$ such that for all $n \geq 1$, and $t \in \mathcal{W}$, we have

$$\mathcal{E}_{\boldsymbol{t}}\left(\|\mathcal{Z}_n\|^2\right) \leq C_{\boldsymbol{\nu},q}\left(\frac{m_{\boldsymbol{\nu},q}}{m_{*,q}}\right)^{2|\boldsymbol{t}|} m_{\boldsymbol{\nu},q}^{2n}.$$

Proof. Observe that we can write

$$\|\mathcal{Z}_n\|^2 = \|\mathcal{Z}_n\| (\|\mathcal{Z}_n\| - 1) + \|\mathcal{Z}_n\|.$$

It is therefore enough to bound $\mathcal{E}_t(||\mathcal{Z}_n|| (||\mathcal{Z}_n|| - 1))$ (the mean number of couples of distinct individuals alive at generation n) to complete the proof.

For $v \in \mathcal{T}$ with |v| < n, we denote by

$$\mathcal{Z}_n^v = \sum_{|u|=n} \mathbf{1}_{\{u(|v|)=v\}} \delta_{\boldsymbol{t}(u)}$$

the empirical counting measure of the types in the subpopulation at generation n descending from the individual v. We can now write

$$\begin{aligned} \mathcal{E}_{t} \left(\| \mathcal{Z}_{n} \| \left(\| \mathcal{Z}_{n} \| - 1 \right) \right) &= \mathcal{E}_{t} \left(\sum_{|u|=n} \mathbf{1}_{\{u \in \mathcal{T}\}} \left(\| \mathcal{Z}_{n} \| - 1 \right) \right) \\ &= \sum_{|u|=n} \mathcal{E}_{t} \left(\mathbf{1}_{\{u \in \mathcal{T}\}} \sum_{j=0}^{n-1} \sum_{k=1}^{d(u(j))} \mathbf{1}_{\{u(j)k \neq u(j+1)\}} \| \mathcal{Z}_{n}^{u(j)k} \| \right), \end{aligned}$$

In words, we decompose the $\|Z_n\| - 1$ individuals alive at generation n barring u according to their most recent common ancestor with u. We obtain from the branching property of Z that

$$\mathcal{E}_{t} \left(\| \mathcal{Z}_{n} \| \left(\| \mathcal{Z}_{n} \| - 1 \right) \right)$$

= $\sum_{|u|=n} \sum_{j=0}^{n-1} \mathcal{E}_{t} \left(\mathbf{1}_{\{u \in \mathcal{T}\}} (d(u(j)) - 1) \mathcal{E}_{t(u(j+1))} (\| \mathcal{Z}_{n-j-1} \|) \right)$
 $\leq \sum_{|u|=n} \sum_{j=0}^{n-1} \mathcal{P}_{t} (u \in \mathcal{T}) \mathbf{k}_{*} c_{\boldsymbol{\nu}, q} m_{\boldsymbol{\nu}, q}^{n-j-1} \left(\frac{m_{\boldsymbol{\nu}, q}}{m_{*, q}} \right)^{|t|+j+1},$

where we used that $d(u(j)) \leq k_*$ and $|t(u(j)k)| = |t| + j + 1 \mathcal{P}_t$ -a.s., and we applied Lemma 3.2. As $m_{*,q} > 1$, we immediately obtain that

$$\mathcal{E}_{\boldsymbol{t}}\left(\left\|\mathcal{Z}_{n}\right\|\left(\left\|\mathcal{Z}_{n}\right\|-1\right)\right) \leq \frac{\mathbf{k}_{*}c_{\boldsymbol{\nu},q}}{m_{*,q}-1}\left(\frac{m_{\boldsymbol{\nu},q}}{m_{*,q}}\right)^{|\boldsymbol{t}|}\mathcal{E}_{\boldsymbol{t}}\left(\left\|\mathcal{Z}_{n}\right\|\right)m_{\boldsymbol{\nu},q}^{n}.$$

Applying again Lemma 3.2, the proof is now complete.

4 Proof of Theorem 1.1

This section is devoted to the proof of Theorem 1.1. We first show that under the assumption $m_{*,q} \leq 1$, $Z(n)/\mathbb{E}_{k_*}(Z(n))$ converges to zero in probability. We next prove that if $m_{*,q} > 1$, then $Z(n)/\mathbb{E}_{k_*}(Z(n))$ converges in L^1 to a non-degenerate random variable.

(i) We assume here that $m_{*,q} \leq 1$, that is the usual Galton-Watson process Z_* is critical or sub-critical, and therefore becomes eventually extinct \mathbb{P}_{k_*} -almost surely. For any given $\varepsilon > 0$, we can choose a generation $\ell \geq 1$ sufficiently large so that $\mathbb{P}_{k_*}(Z_*(\ell) \geq 1) \leq \varepsilon^2$. In other words, with $\mathcal{P}_{\varnothing}$ -probability at least $1-\varepsilon^2$,

all individuals at generation ℓ have at least one forebear that had strictly less than k_* children. Plainly, there are at most k_*^{ℓ} individuals at generation ℓ , and the set of possible types for these individuals is also bounded by k_*^{ℓ} . The branching property yields

$$\begin{split} & \mathbb{E}_{\mathbf{k}_*}(m_{\boldsymbol{\nu},q}^{-n-\ell}Z(n+\ell), Z_*(\ell)=0) \\ & \leq m_{\boldsymbol{\nu},q}^{-\ell}\mathbf{k}_*^{-\ell}\max\{m_{\boldsymbol{\nu},q}^{-n}\mathcal{E}_{\boldsymbol{t}}\left(\|\mathcal{Z}_n\|\right): |\boldsymbol{t}|=\ell, \boldsymbol{t}\neq(\mathbf{k}_*,\ldots,\mathbf{k}_*)\}, \end{split}$$

and we now see from Lemma 3.1 that the right-hand side can be bounded from above by ε^2 for all sufficiently large n. An application of the Markov inequality now gives

$$\lim_{n \to \infty} \mathbb{P}_{\mathbf{k}_*}(m_{\boldsymbol{\nu},q}^{-n-\ell}Z(n+\ell) \ge \varepsilon) \le 2\varepsilon,$$

and Theorem 1.1(i) is proven.

(ii) We now suppose that $m_{*,q} > 1$, i.e. that the usual Galton-Watson process Z_* is supercritical. Since its reproduction law has bounded support, the process

$$W_*(n) \coloneqq m_{*,q}^{-n} Z_*(n), \qquad n \ge 0$$

is a martingale bounded in L^2 , and we write W_* for its terminal value. Recalling that the ancestor has k_* children, we have

$$\mathcal{E}_{\varnothing}(W_*) = \mathcal{E}_{\varnothing}(W_*(n)) = \mathbf{k}_* m_{*,q}^{-1} = \frac{1}{q + (1-q)\nu(\mathbf{k}_*)}$$

Note that by (1.2) and (1.3), we have

$$\lim_{n \to \infty} \mathbb{E}_{\mathbf{k}_*}(m_{\boldsymbol{\nu},q}^{-n} Z(n)) = \frac{1}{q + (1-q)\nu(\mathbf{k}_*)} = \mathcal{E}_{\varnothing}(W_*).$$

To prove that $m_{\nu,q}^{-n}Z(n)$ converges to W_* in L^1 , we use the following classical variant of Scheffé's lemma: a sequence of non-negative random variables (ξ_n) converges in $L^1(\mathbb{P})$ to some random variable ξ whenever $\lim_{n\to\infty} \mathbb{E}(\xi_n) = \mathbb{E}(\xi)$ and $\lim_{n\to\infty} \xi_n = \xi$ in probability. Note that the usual Scheffé's lemma makes the stronger requirement $\xi_n \to \xi$ a.s., but using that from any extraction of ξ_n one can find a subsequence converging almost surely to ξ , the result still holds.

We denote by \mathcal{T}_* the Galton-Watson subtree of \mathcal{T} obtained by only keeping elements of \mathcal{T} with outdegree k_* . Recall that $\#\{|u| = \ell : u \in \mathcal{T}_*\} = Z_*(\ell)$. We prove the convergence in probability of $m_{\nu,q}^{-n}Z(n)$ to W_* by decomposing \mathcal{Z}_n at an intermediate generation ℓ as $\|\mathcal{Z}_n\| = \|\mathcal{Z}_n^{\ell*}\| + R_n$, where $\mathcal{Z}_n^{\ell*}$ is the point measure on types associated to individuals at generation n in \mathcal{T} with ancestors at generation ℓ that belong to \mathcal{T}_* . Using the branching property at generation ℓ , we remark that $\|\mathcal{Z}_n^{\ell*}\|$ is the sum of $Z_*(\ell)$ independent copies of $\|\mathcal{Z}_{n-\ell}\|$ under law $\mathcal{P}_{(k_*,...,k_*)}$, while R_n is the sum of at most k_*^{ℓ} independent copies of $\|\mathcal{Z}_{n-\ell}\|$ starting from initial conditions such that $|t| = \ell$ and $t \neq (k_*,...,k_*)$.

As a result, for each fixed $\ell > 0$, we have by Lemma 3.1 that

$$\lim_{n \to \infty} \mathcal{E}_{\varnothing} \left(m_{\nu,q}^{-n} R_n \right) = 0.$$

In particular, $m_{\nu,q}^{-n}R_n$ converges to 0 in probability.

We now compute the mean and variance of $\|\mathcal{Z}_n^{\ell*}\|$ conditionally on the first ℓ generations of the process. Using the consequence of the branching property described above, we have

$$\mathcal{E}_{\varnothing}\left(\left|\mathcal{Z}_{n}^{\ell*}\right|\right|\mathcal{T}_{\left|\ell\right)}=Z_{*}(\ell)\mathcal{E}_{\left(\mathbf{k}_{*},\ldots,\mathbf{k}_{*}\right)}\left(\left\|\mathcal{Z}_{n-\ell}\right\|\right)\sim_{n\to\infty}m_{\nu,q}^{n}W_{*}(\ell)\quad\text{a.s.}$$

by Lemma 3.1. Similarly, we compute the conditional variance

$$\begin{aligned} \mathcal{E}_{\varnothing} \left(\left(\left\| \mathcal{Z}_{n}^{\ell*} \right\| - \mathcal{E}_{\varnothing} \left(\left\| \mathcal{Z}_{n}^{\ell*} \right\| \left| \mathcal{T}_{\left| \ell \right.} \right) \right)^{2} \left| \mathcal{T}_{\left| \ell \right.} \right) \right. \\ &= Z_{*}(\ell) \mathcal{E}_{\left(k_{*}, \dots, k_{*}\right)} \left(\left(\left\| \mathcal{Z}_{n-\ell} \right\| - \mathcal{E}_{\left(k_{*}, \dots, k_{*}\right)} \left(\left\| \mathcal{Z}_{n-\ell} \right\| \right) \right)^{2} \right) \\ &\leq C_{\boldsymbol{\nu}, q} Z_{*}(\ell) m_{\boldsymbol{\nu}, q}^{2(n-\ell)} \left(\frac{m_{\boldsymbol{\nu}, q}}{m_{*, q}} \right)^{2\ell}, \end{aligned}$$

by Lemma 3.3. Therefore, for all $\varepsilon > 0$, applying a conditional Bienaymé-Chebyshev inequality, we obtain that for all $0 \le \ell \le n$

$$\mathcal{P}_{\varnothing}\left(m_{\nu,q}^{-n}\left|\left\|\mathcal{Z}_{n}^{\ell*}\right\|-\mathcal{E}_{\varnothing}\left(\left\|\mathcal{Z}_{n}^{\ell*}\right\|\left|\mathcal{T}_{|\ell}\right)\right|>\varepsilon\right)\leq C_{\nu,q}\mathcal{E}_{\varnothing}(Z_{*}(\ell))m_{*,q}^{-2\ell}\varepsilon^{-2}$$

and we observe that this bounds converges to 0 as $\ell \to \infty$ uniformly in *n*. Let $\varepsilon > 0$, we fix $\ell > 0$ large enough such that

$$\mathcal{P}_{\varnothing}(|W_*(\ell) - W_*| > \varepsilon) < \varepsilon \text{ and } C_{\nu,q} \mathcal{E}_{\varnothing}(Z_*(\ell)) m_{*,q}^{-2\ell} \varepsilon^{-2} \le \varepsilon,$$

then $n \geq \ell$ large enough such that

$$\mathcal{P}_{\varnothing}\left(\left|\mathcal{E}_{\varnothing}\left(m_{\boldsymbol{\nu},q}^{-n}\|\mathcal{Z}_{n}^{\ell*}\|\left|\mathcal{T}_{\ell}\right)-W_{*}(\ell)\right|>\varepsilon\right)\leq\varepsilon\quad\text{and}\quad\mathcal{P}_{\varnothing}(m_{\boldsymbol{\nu},q}^{-n}R_{n}>\varepsilon)\leq\varepsilon,$$

we have

$$\mathcal{P}_{\varnothing}\left(|m_{\nu,q}^{-n}Z_n - W_*| > 4\varepsilon\right) \le 4\varepsilon,$$

which completes the proof.

5 Proof of Theorem 1.4

In this section, we will prove Theorem 1.4; let us briefly recall our approach using the notation we introduced. We work from the viewpoint of multitype branching processes under \mathcal{P}_d for some arbitrary fixed $d \in S_{\nu}$. This is equivalent to consider, under \mathbb{P}_d , the subpopulation generated by one of the *d* individuals at the first generation. We shall introduce a nonnegative martingale $M = (M_n)_{n\geq 1}$ starting from $M_0 = 1$, which is naturally related to the dynamics of the reinforced Galton-Watson process, and in particular vanishes as soon as *Z* becomes extinct. Therefore *Z* survives on the event that the terminal value M_{∞} is nonzero, and the latter happens with positive probability as soon as *M* is uniformly integrable.

5.1 A natural martingale

We start by introducing some notation. For every type t, we set

$$m_{\boldsymbol{t}} \coloneqq \mathcal{E}_{\boldsymbol{t}}(Z(1)) = \sum_{j} j \pi_{\boldsymbol{t}}(j),$$

where, using the notation in Section 2, π_t is the reproduction law of an individual with type t. More explicitly, for a non-empty type $t = (d_1, \ldots, d_\ell) \in \mathcal{W}$, we have

$$m_{t} = (1-q)m_{\rm GW} + \frac{q}{\ell} \sum_{j=1}^{\ell} d_j, \qquad (5.1)$$

where $m_{\text{GW}} = \sum_{j} j\nu(j)$ is the mean reproduction number for the usual Galton-Watson process with reproduction law ν . Recall that the type $\mathbf{t}((u(j)))$ of the forebear u(j) of u at generation $j \leq |u|$ is given by the suffix of $\mathbf{t}(u)$ with length $j + |\mathbf{t}(\emptyset)|$. We define, for $u \in \mathcal{T}$,

$$\Phi(u) \coloneqq \prod_{j=0}^{|u|-1} \frac{1}{m_{t(u(j))}},$$
(5.2)

with the convention that $\Phi(\emptyset) = 1$. In words, for any individual $u \in \mathcal{T}$, say at generation $|u| = k \ge 1$, $1/\Phi(u)$ is the product of the mean reproduction numbers of the forebears of this individual.

If follows immediately from the definition of the mean m_t and the function Φ that the process $M = (M_n)_{n \ge 1}$ given by

$$M_n = \sum_{u \in \mathcal{T}, |u|=n} \Phi(u) \tag{5.3}$$

is a martingale under \mathcal{P}_t for any initial type t. Indeed, using the branching property of Lemma 2.1 we have

$$\mathcal{E}_{\boldsymbol{t}}\left(M_{n+1}\big|\mathcal{T}_{|n}\right) = \sum_{u\in\mathcal{T}:|u|=n}\frac{\Phi(u)}{m_{\boldsymbol{t}(u)}}\mathcal{E}_{\boldsymbol{t}(u)}(|\mathcal{Z}_{1}|) = M_{n}.$$

In other words, Φ is mean-harmonic for the multitype branching process in the sense of [5]. The main purpose of this section is to study the asymptotic behavior of M_n as $n \to \infty$.

Proposition 5.1. We assume that $qk_* < 1$, and consider an arbitrary $d \in S_{\nu}$.

- (i) If (1.7) holds, then the martingale M is uniformly integrable under \mathcal{P}_d .
- (ii) If

$$\sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{1-qj} < 1, \tag{5.4}$$

then the terminal value of the martingale M is $M_{\infty} = 0$, \mathcal{P}_d -a.s.

As it was already pointed out, Proposition 5.1(i) entails Theorem 1.4. The rest of this section is devoted to the proof of Proposition 5.1. The analysis relies on classical arguments involving a change of probability induced by the martingale M and a decomposition of the branching process along the spine. This leads us to investigate the asymptotic behavior of types along the spine, for which we will rely on classical results by Athreya and Karlin [1] and Janson [8] on Pólya urns with random replacements.

5.2 Spinal decomposition

Let $d \in S_{\nu}$. In this section, we introduce two distributions, $\overline{\mathcal{P}}_d$ and $\widehat{\mathcal{P}}_d$, the first on the space of marked genealogical trees, and the second on the richer space of marked genealogical trees with a distinguished infinite branch called the spine. The spinal decomposition then identifies $\overline{\mathcal{P}}_d$ as the projection of $\widehat{\mathcal{P}}_d$. As above, we will write $\overline{\mathcal{E}}_d$, respectively $\widehat{\mathcal{E}}_d$ for the expectations associated to the law $\overline{\mathcal{P}}_d$ and $\widehat{\mathcal{P}}_d$ respectively.

In this direction, recall from Section 2 that π_t stands for the reproduction law of an individual with type t. For any probability distribution π on \mathbb{Z}_+ with a finite and non-zero first moment, we also denote by $\hat{\pi}$ the size-biased distribution of π , defined by

$$\widehat{\pi}(k) = \frac{k\pi(k)}{\sum_{j=0}^{\infty} j\pi(j)}, \qquad k \ge 1.$$

To start with, $\overline{\mathcal{P}}_d$ is defined for every $n \ge 1$ by

$$\overline{\mathcal{P}}_d(A) = \mathcal{E}_d(M_n \mathbf{1}_A), \qquad \forall A \in \mathcal{F}_n$$

where $(\mathcal{F}_n)_{n\geq 1}$ stands for the natural filtration on the space of genealogical trees (with types) induced by generations. We next construct another distribution involving a spine $\varsigma = (\varsigma(n))_{n\geq 0}$, where the latter is a distinguished line of descent, that is, a sequence of individuals such that for every $n \ge 0$, $\varsigma(n+1)$ is a child of $\varsigma(n)$. Specifically, $\varsigma(0) = \emptyset$ and we let $\varsigma(0)$ reproduce according to the size-biased reproduction law $\hat{\pi}_d$. We then select an individual $\varsigma(1)$ uniformly at random amongst the, say, k children of $\varsigma(0)$ which have all the type $t(\varsigma(1)) = (k, d)$. Again we let $\varsigma(1)$ reproduce according to the size-biased reproduction law $\widehat{\pi}_{t(\varsigma(1))}$, whereas each other individual u in the sibling reproduce independently according to π_u . And so on, and so forth, that is by induction at every generation, individuals reproduce independently of one from the other's, in such a way that a non-spine individual with type t reproduce according to the law π_t , and the spine particle $\varsigma(n)$, say of type s, reproduces according to the law $\hat{\pi}_{s}$. The next individual of the spine $\varsigma(n+1)$ is chosen uniformly among the children of $\varsigma(n)$. The law of the resulting genealogical tree endowed with a spine is denoted by \mathcal{P}_d .

The following result, which is referred to as the spine decomposition of a multitype branching process, is a consequence of [5, Proposition 12.1 and Lemmas 12.2 and 12.3], our function Φ playing the role of the harmonic function h in that article, and with $\hat{\mathcal{P}}$ corresponding to \mathbb{Q} there.

Proposition 5.2 (Spine decomposition of the reinforced Galton-Watson process). For all $n \in \mathbb{N}$, we have

$$\overline{\mathcal{P}}_d\Big|_{\mathcal{F}_n} = \left.\widehat{\mathcal{P}}_d\right|_{\mathcal{F}_n}.$$

Moreover, for all $n \in \mathbb{N}$ and $u \in \mathcal{U}$ with |u| = n, we have

$$\widehat{\mathcal{P}}_d(\varsigma_n = u | \mathcal{F}_n) = \frac{\Phi(u) \mathbf{1}_{\{u \in \mathcal{T}\}}}{M_n}$$

Remark 5.3. As usual, this spine decomposition result gives rise to a many-toone type lemma, yielding in particular to an alternative proof of [4, Equation (2.1)]. We observe that

$$\mathcal{E}_{\boldsymbol{t}}(\|\mathcal{Z}_n\|) = \widehat{\mathcal{E}}_{\boldsymbol{t}}\left(\frac{\|\mathcal{Z}_n\|}{M_n}\right) = \widehat{\mathcal{E}}_{\boldsymbol{t}}\left(\sum_{|u|=n} \frac{\mathbf{1}_{\{u=\varsigma(n)\}}}{\Phi(u)}\right) = \widehat{\mathcal{E}}_{\boldsymbol{t}}\left(\prod_{j=0}^{n-1} m_{\boldsymbol{t}(\varsigma(j))}\right).$$

Using the definition of m_t , we conclude that $\mathcal{E}_t(||\mathcal{Z}_n||)$ can be computed as the mean of the product of the number of children sampled along a randomly selected line in the reinforced branching process.

In the next section, we study in more details the reproduction law of the spine particle in terms of an urn model. It allows to describe the reinforced Galton-Watson process under law $\hat{\mathcal{P}}_d$ as a multitype branching process with immigration, whose asymptotic behaviour can be studied. Then, using a classical argument due to Durrett [6, Theorem 4.3.5] (see also [5, Theorem 3]), we are able to provide necessary and sufficient conditions for the uniform integrability of the martingale M, which we translate into Proposition 5.1 in Section 5.4.

5.3 Dynamics of the spine as a generalized Pólya urn

For $n \geq 0$, let ξ_{n+1} denote the number of children of $\varsigma(n)$, the individual on the spine at generation n. We also agree that $\xi_0 \equiv d$ under $\widehat{\mathcal{P}}_d$; in particular, the type of $\varsigma(n)$ is given by $\boldsymbol{\tau}_n \coloneqq (\xi_n, \xi_{n-1}, \ldots, \xi_0)$. From the construction of $\widehat{\mathcal{P}}_d$ in the preceding subsection, $\xi_0 = d$, ξ_1 has the law $\widehat{\boldsymbol{\pi}}_d$, and we have for any $k \geq 1$ that

$$\widehat{\mathcal{P}}_{d}(\xi_{k+1} = j \mid \xi_{0}, \xi_{1}, \dots, \xi_{k}) = c_{k} j \left((1-q)\nu(j) + q \varpi_{\tau_{k}}(j) \right), \ j \ge 0,$$

where $c_k > 0$ is the constant of normalization.

In this section, we shall first identify these dynamics as those of a generalized Pólya urn. Next, we will determine the asymptotic behavior of the urn process by identifying the principal spectral elements of its mean replacement matrix, using classical results of Janson [8] in this field. This enables us to estimate the value of the mean-harmonic function Φ defined in (5.2) along the spine.

Recall from (1.1) that S_{ν} designates the strictly positive part of the support of the reproduction law ν . We think of any $k \in S_{\nu}$ as a color, and add a special color denoted by \star . We define an urn process with balls having colors in $S_{\nu} \cup \{\star\}$ as follows. Imagine that a ball with color $k \in S_{\nu}$ has activity qk, meaning that the probability that it is picked at some random drawing from the urn is proportional to qk, whereas a ball with color \star has activity $(1-q)m_{\text{GW}}$.

At the initial time n = 0, the urn contains one ball with color d and one ball with color \star . At each step $n \geq 1$, a ball is drawn at random in the urn with probability proportional to its activity. The ball is then replaced in the urn and two new balls are added to the urn. If the color of the sample ball is $j \in S_{\nu}$, then the first new ball has the color j and the second the color \star . If the sampled ball has color \star , then the first new ball has the color \star and the color of the second is sampled according to the law size-biased reproduction law $\hat{\nu}$. Write X_n the label of the non- \star ball added to the urn at time n. We also set $X_0 = d$ as we initiate the urn with a ball with color d and a ball with color \star . **Lemma 5.4.** The sequence $(X_n, n \ge 0)$ of colors added to the urn as above has the same law as $(\xi_n, n \ge 0)$ under $\widehat{\mathcal{P}}_d$.

Proof. For all $n \ge 0$ and $j \in S_{\nu} \cup \{\star\}$, we denote by $N_n(j)$ the number of balls with color j in the urn after n steps. It is plain

$$N_n(\star) = \sum_{j \in S_{\nu}} N_n(j) = n + 1.$$

Moreover, for any $j \in S_{\nu}$, we have

$$\mathbb{P}(X_{n+1}=j|X_0,\ldots,X_n) = \frac{(1-q)m_{\rm GW}N_n(\star)\widehat{\nu}(j) + jqN_n(j)}{(1-q)m_{\rm GW}N_n(\star) + q\sum_{i\in S_{\nu}}iN_n(i)}$$

since $\{X_{n+1} = j\}$ is the event that at time n + 1, either a ball labelled j was sampled, or a ball labelled \star and the extra ball added was labelled j. Since $j\nu(j) = m_{\rm GW}\hat{\nu}(j)$, this quantity can be rewritten as

$$\mathbb{P}(X_{n+1} = j | X_0, \dots, X_n) = \frac{(1-q)(n+1)j\nu(j) + qjN_n(j)}{(1-q)m_{\rm GW}(n+1) + q\sum_{i \in S_{\nu}} iN_n(i)}$$
$$= \widehat{\pi}_{\tau_n}(j),$$

where $\tau_n = (X_n, \dots, X_0)$. This proves that the dynamics of X are identical to those of ξ under law $\hat{\mathcal{P}}_d$.

We next study the asymptotic behaviour of X_n as $n \to \infty$ by applying general results of Athreya and Karlin [1] and Janson [8] on generalized Pólya urns. Recall that $N_n(j)$ denotes the number of balls with color $j \in S_{\nu} \cup \{\star\}$ in the urn after *n* steps. By [1, Section 4.2] or [8, Theorem 3.21], there exists a constant c > 0 such that for any $j \in S_{\nu} \cup \{\star\}$,

$$\lim_{n \to \infty} \frac{N_n(j)}{n} = c\lambda_1 v_1(j) \quad \text{a.s.},$$

where λ_1 is the leading eigenvalue and $\mathbf{v}_1 = (v_1(j))$ an associated left-eigenvector of the matrix $A = (A_{i,j})_{i,j \in S_{\nu} \cup \{\star\}}$ is defined for $i, j \in S_{\nu}$ by

$$A_{i,j} = qi\delta_{i,j}, \quad A_{i,\star} = qi, \quad A_{\star,j} = (1-q)j\nu(j), \quad A_{\star,\star} = (1-q)m_{\rm GW}.$$
 (5.5)

Roughly speaking, the matrix A is mean replacement matrix re-weighted by activities. Beware that Janson [8] rather uses the notation A for the transposed of our matrix; in particular left-eigenvectors \mathbf{v} in our setting correspond to right-eigenvectors in [8], and vice-versa. The following spectral properties of A are the key to the analysis.

Lemma 5.5. The eigenvalues of the matrix A defined in (5.5) are all simple, real and nonnegative. They are given by 0 and the $\#S_{\nu}$ positive solutions of the equation

$$\sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{x-qj} = 1, \qquad \frac{x}{q} \in \mathbb{R} \backslash S_{\boldsymbol{\nu}}.$$
(5.6)

Moreover, a left-eigenvector \mathbf{v}_{λ} associated to an eigenvalue $\lambda \geq 0$ is given by

$$v_{\lambda}(i) = \frac{(1-q)i\nu(i)}{\lambda - qi}$$
 for $i \in S_{\nu}$, and $v_{\lambda}(\star) = 1$,

and similarly a right-eigenvector \mathbf{u}_{λ} associated to an eigenvalue $\lambda \geq 0$ is given by

$$u_{\lambda}(i) = rac{qi}{\lambda - qi} \quad for \ i \in S_{\nu}, \ and \ u_{\lambda}(\star) = 1.$$

Proof. We first remark that for each consecutive elements i, j of S_{ν} , the function

$$x \mapsto \sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{x-qj}, \qquad \frac{x}{q} \in \mathbb{R} \backslash S_{\nu},$$

is decreasing from ∞ to $-\infty$ on the interval (qi, qj). Additionally, this function is also decreasing on $(-\infty, q \min S_{\nu})$ while staying non-positive, and is decreasing on (qk_*, ∞) from ∞ to 0. Consequently, there are exactly $\#S_{\nu}$ roots to the equation (5.6), all being positive.

We then consider any solution, say $\lambda > 0$, to (5.6) and check by immediate computations that $(\lambda, \mathbf{v}_{\lambda})$ as defined in the statement are eigenvalues and associated left-eigenvectors for the matrix A. Indeed, we have first for any $j \in S_{\nu}$,

$$\sum_{i \in S_{\nu} \cup \{\star\}} v_{\lambda}(i) A_{i,j} = qj \frac{(1-q)j\nu(j)}{\lambda - qj} + (1-q)j\nu(j) = \lambda v_{\lambda}(j),$$

and then for $j = \star$,

$$\sum_{i \in S_{\nu} \cup \{\star\}} v_{\lambda}(i) A_{i,\star} = (1-q) m_{\text{GW}} + \sum_{i=1}^{k_*} q_i \frac{(1-q)i\nu(i)}{\lambda - q_i}$$
$$= (1-q) \sum_{i=1}^{k_*} \left(i\nu(i) + q_i \frac{i\nu(i)}{\lambda - q_i} \right)$$
$$= (1-q) \sum_{i=1}^{k_*} i\nu(i) \frac{\lambda}{\lambda - q_i}$$
$$= \lambda,$$

where the ultimate equality uses that λ solves (5.6).

We check by similar calculations for $(0, \mathbf{v}_0)$ that

$$\sum_{i \in S_{\nu} \cup \{\star\}} v_0(i) A_{i,j} = 0 \quad \text{for all } j \in S_{\nu} \cup \{\star\}.$$

Finally, we verify in the same way that the \mathbf{u}_{λ} are also right-eigenvectors.

As a conclusion, we found $\#S_{\nu} + 1$ different real eigenvalues of the square matrix A of dimension $\#S_{\nu} + 1$. They are hence all simple and there exist no further eigenvalues.

We order the positive eigenvalues of the mean replacement matrix A in the decreasing order, $\lambda_1 > \lambda_2 > \ldots > \lambda_{\#S_{\nu}}$ and then write simply $\mathbf{v}_i = \mathbf{v}_{\lambda_i}$ and $\mathbf{u}_i = \mathbf{u}_{\lambda_i}$ for the corresponding left and right eigenvectors given in Lemma 5.5. Note that \mathbf{u}_1 and \mathbf{v}_1 have positive coordinates (as it should be expected from Perron-Frobenius' theorem), and that we did not impose the usual normalization that their scalar product should be 1, for the sake of simplicity. It is a well-known fact that the first order asymptotic behaviour of N_n is determined by

 λ_1 and \mathbf{v}_1 , whereas the fluctuations depend on the sign of $\lambda_2 - \lambda_1/2$. More precisely, we first apply [8, Theorem 3.21], to obtain the asymptotic behavior of the number $N_n(j)$ of balls with color j after n steps as $n \to \infty$.

Lemma 5.6. We have for all $j \in S_{\nu}$ that

$$\lim_{n \to \infty} \frac{N_n(j)}{n} = v_1(j) = \frac{(1-q)j\nu(j)}{\lambda_1 - qj}, \qquad \widehat{\mathcal{P}}_d\text{-}a.s.$$

Proof. The matrix A is irreducible and aperiodic, and Lemma 5.5 entails that the conditions (A1–6) of [8] are satisfied. By [8, Theorem 3.21], there exists c > 0 such that for all $j \in S_{\nu} \cup \{\star\}$

$$\lim_{n \to \infty} \frac{N_n(j)}{n} = c\lambda_1 v_1(j) \quad \text{a.s}$$

Moreover, as $N_n(\star) = n + 1$ a.s., we have $N_n(\star)/n \to 1$ a.s. As $v_1(\star) = 1$, we conclude that $c = 1/\lambda_1$, which completes the proof.

Corollary 5.7. Assume that $qk_* < 1$.

- (i) The principal eigenvalue $\lambda_1 > 1$ if and only if (1.7) holds. In that case, the series $\sum_{k=0}^{\infty} \Phi(\varsigma(n))$ converges a.s.
- (ii) The principal eigenvalue $\lambda_1 < 1$ if and only if (5.4) holds. In that case, we have $\lim_{n\to\infty} \Phi(\varsigma(n)) = \infty$ a.s.

Proof. The equivalences

$$\lambda_1 > 1 \Longleftrightarrow \sum_{j=1}^{\mathbf{k}_*} \frac{(1-q)j\nu(j)}{1-qj} > 1,$$

and

$$\lambda_1 < 1 \Longleftrightarrow \sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{1-qj} < 1.$$

should be plain from Lemma 5.5. Indeed, the only eigenvalue greater than $qk_* < 1$ is the Perron-Frobenius principal eigenvalue λ_1 , and we have seen that the function

$$x \mapsto \sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{1-qj}$$

decreases on $[1, \infty)$.

Next, recall that the type τ_n of the individual $\varsigma(n)$ is $\tau_n = (\xi_n, \ldots, \xi_0)$, where ξ_k stands the number of children of $\varsigma(k-1)$. We know moreover from Lemma 5.4 that the sequence $(\xi_k)_{k\geq 0}$ has the same distribution under $\widehat{\mathcal{P}}_d$ as the sequence in S_{ν} of the colors of the balls added to the urn at each step. We deduce from Lemma 5.6 that $\widehat{\mathcal{P}}_d$ -a.s., there is the convergence

$$\lim_{n \to \infty} \frac{\#\{k \le n : \xi_k = j\}}{|\boldsymbol{\tau}_n|} = v_1(j), \quad \text{for all } j \in S_{\boldsymbol{\nu}}.$$

We then get from (5.1) that

$$\lim_{n \to \infty} m_{\tau_n} = (1-q) \sum_{j=1}^{k_*} j\nu(j) + q \sum_{j=1}^{k_*} \frac{(1-q)j^2\nu(j)}{\lambda_1 - qj}$$
$$= (1-q) \sum_{j=1}^{k_*} \frac{\lambda_1 j\nu(j)}{\lambda_1 - qj}$$
$$= \lambda_1.$$

where the last equality is due to the fact that λ_1 solves (5.6). Last, we deduce from the very definition (5.2) of Φ that as $n \to \infty$,

$$\log \Phi(\varsigma(n)) \sim -n \log \lambda_1.$$

As a consequence, the series $\sum_{k=0}^{\infty} \Phi(\varsigma(n))$ converges a.s. whenever $\lambda_1 > 1$, whereas $\lim_{n\to\infty} \Phi(\varsigma(n)) = \infty$ a.s. whenever $\lambda_1 < 1$.

We have now all the ingredients needed to establish Proposition 5.1.

5.4 Proof of Proposition 5.1

The starting point of the proof is the following well-known observation due to Durrett [6, Theorem 4.3.5] (see also [5, Theorem 12.1]). As $M_0 = 1$, and writing $M_{\infty} = \lim_{n \to \infty} M_n$, we have

$$\mathcal{E}_d(M_\infty) = \overline{\mathcal{P}}_d(M_\infty < \infty). \tag{5.7}$$

Recall that the tilted law $\overline{\mathcal{P}}_d$ has been introduce in Section 5.2, and that the spine decomposition of the reinforced Galton-Watson process has been described in Proposition 5.2

It is worth noting that M is a non-negative martingale under \mathcal{P}_d , and 1/M a non-negative super-martingale under $\overline{\mathcal{P}}_d$, hence the convergence of M is immediate under the original and tilted laws. As a result, to prove Proposition 5.1(i), it is enough to study the convergence of M under the richer law $\widehat{\mathcal{P}}_d$ by Proposition 5.2. Specifically, almost sure finiteness of M_∞ under $\widehat{\mathcal{P}}_d$ implies uniform integrability under \mathcal{P}_d , while if $M_\infty = \infty \ \widehat{\mathcal{P}}_d$ -a.s., then $M_\infty = 0 \ \mathcal{P}_d$ -a.s.

(i) We denote by $\mathcal{Y} = \sigma(\xi_k, k \ge 0)$ the sigma-field of the information on the number of children of all spine particles. The spinal decomposition yields

$$\widehat{\mathcal{E}}_d(M_n|\mathcal{Y}) = \Phi(\varsigma(n)) + \sum_{k=0}^{n-1} (\xi_k - 1) \Phi(\varsigma(k+1)), \qquad \widehat{\mathcal{P}}_d\text{-a.s.},$$

using that the individual $\varsigma(k)$ at generation k < n on the spine has ξ_k children, one of those chosen as the individual on the spine at generation k and the $\xi_k - 1$ other –that have the same type $\mathbf{t}(\varsigma(k))$ as the spine individual– evolve independently according to the law $\mathcal{P}_{\mathbf{t}(\varsigma(k+1))}$. Since $\xi_k - 1 \leq k_*$, we have a fortiori that

$$\widehat{\mathcal{E}}_d(M_n|\mathcal{Y}) \leq \mathbf{k}_* \sum_{k=0}^{\infty} \Phi(\varsigma(k)), \qquad \widehat{\mathcal{P}}_d\text{-a.s.}$$

We can now conclude from Corollary 5.7 and the conditional version of the Fatou lemma that if

$$\sum_{j=1}^{\kappa_*} \frac{(1-q)j\nu(j)}{1-qj} > 1,$$

then $\liminf_{n\to\infty} M_n < \infty$, $\widehat{\mathcal{P}}_d$ -a.s. As a result, by Proposition 5.2 and (5.7) we have $\mathcal{E}_d(M_\infty) = 1$, and by the Scheffé lemma, M is uniformly integrable under \mathcal{P}_d .

(ii) The simple observation that $M_n \ge \Phi(\varsigma(n))$ combined with Corollary 5.7(ii) enables us to conclude that $\widehat{\mathcal{P}}_d(M_\infty = \infty) = 1$ whenever (5.4) holds. Using again Proposition 5.2 and (5.7), the proof is now complete.

5.5 Short discussion of the critical case

The proof of Proposition 5.1 in the preceding section relies on the study of the quantity $\sum_{k=0}^{\infty} \Phi(\varsigma(n))$, which can be thought of as a particular case of *reinforced* perpetuity. We have seen above that this series converges when $\lambda_1 > 1$ whereas $\lim_{n\to\infty} \Phi(\varsigma(n)) = \infty$ when $\lambda_1 < 1$. We now conclude this work by discussing briefly and a bit informally the critical case when the principal eigenvalue is $\lambda_1 = 1$, which presents interesting complexities. Recall from Corollary 5.7 that this is equivalent to

$$\sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{1-qj} = 1$$

We shall need the following result (that however does not requires $\lambda_1 = 1$) about the fluctuations of the convergence in Lemma 5.6.

Lemma 5.8. (i) (Heavy urn) If $\lambda_2 > \lambda_1/2$, then there exists a square integrable non-degenerate random variable W_2 such that for all $j \in S_{\nu}$,

$$\lim_{n \to \infty} n^{-\lambda_2/\lambda_1} (N_n(j) - n\lambda_1 v_1(j)) = W_2 v_2(j), \quad \widehat{\mathcal{P}}_d\text{-}a.s.$$

(ii) (Light urn) If $\lambda_2 < \lambda_1/2$, then there is the joint convergence in distribution for all $j \in S_{\nu}$

$$\lim_{n \to \infty} n^{-1/2} \left(N_{[nt]}(j) - nt\lambda_1 v_1(j)) \right)_{t \ge 0} = (G_t(j))_{t \ge 0}, \quad in \ law_t$$

where $\mathbf{G} = (G(j))_{j \in S_{\nu}}$ is some continuous centered $\#S_{\nu}$ -dimensional Gaussian process.

Proof. This is a direct application of [8, Theorem 3.24 and Theorem 3.31], using that λ_2 is a simple eigenvalue with left-eigenvector \mathbf{v}_2 , and that A has no non-real eigenvalues.

In our setting, we now see that if $\lambda_1 = 1$ and $\lambda_2 > 1/2$, then as $n \to \infty$,

$$m_{\tau_n} = 1 + q n^{\lambda_2 - 1} W_2 \sum_{j=1}^{k_*} \frac{(1 - q)j^2 \nu(j)}{\lambda_2 - qj} + o(n^{\lambda_2 - 1}), \qquad \widehat{\mathcal{P}}_d\text{-a.s.}$$

Moreover, we remark that

$$\sum_{j=1}^{k_*} \frac{(1-q)j^2\nu(j)}{\lambda_2 - qj} = \frac{1}{q} \left(\sum_{j=1}^{k_*} \frac{(1-q)j(qj-\lambda_2)\nu(j)}{\lambda_2 - qj} + \lambda_2 \sum_{j=1}^{k_*} \frac{(1-q)j\nu(j)}{\lambda_2 - qj} \right)$$
$$= \frac{\lambda_2 - (1-q)m_{\rm GW}}{q},$$

using that λ_2 solves (5.6). As a consequence, on the event

$$\{W_2(\lambda_2 - (1-q)m_{\rm GW}) > 0\},\$$

the series $\sum_{k=0}^{\infty} \Phi(\varsigma(n))$ converges and the terminal value $M_{\infty} < \infty$. At the opposite, on the event

$$\{W_2(\lambda_2 - (1-q)m_{\rm GW}) < 0\},\$$

the series $\sum_{k=0}^{\infty} \Phi(\varsigma(n))$ diverges.

We remark that W_2 is a random variable that is positive, respectively negative, with positive probability. Indeed, we observe from [8] that W_2 is constructed as the almost sure limit of an uniformly integrable martingale $(W_2(n))$ with $W_2(0) = 0$. Since $W_2(n)$ can be positive or negative with positive probability, depending on the values of X_1, \ldots, X_n , we conclude that $\mathbb{P}(W_2 > 0)\mathbb{P}(W_2 < 0) > 0$. Therefore, assuming that $\lambda_2 - (1-q)m_{\rm GW} \neq 0$, the reinforced perpetuity converges, respectively diverges, with positive probability. In particular the martingale M converges to a non-degenerate random variable, with positive probability under law \mathcal{P}_d , while being non-uniformly integrable.

In the light case when $\lambda_1 = 1$ and $\lambda_2 < 1/2$, $n^{1/2}(m_{\tau_n} - 1)$ rather converges in distribution to some centered Gaussian variable. This suggests that in most (if not all) cases, one should have $\limsup_{n\to\infty} \Phi(\varsigma(n)) = \infty$. This would imply that the terminal value $M_{\infty} = \infty$, $\widehat{\mathcal{P}}_d$ -a.s., and therefore also $M_{\infty} = 0$, \mathcal{P}_d -a.s.

The possible behaviors of critical reinforced perpetuity hence appear richer than for regular perpetuities. This also suggests that there exist reinforced branching processes such that (5.4) holds but the process survives with positive probability.

Acknowledgments. The authors wish to thank the anonymous referees for their comments that helped improved previous versions of this article.

Funding. The first author was partially supported by the project SNSF 200020 212115: Random processes with genealogy dependent dynamics.

The second author was partially supported by the MITI interdisciplinary program 80PRIME GEx-MBB and the ANR MBAP-P (ANR-24-CE40-1833) project.

References

 Krishna B. Athreya and Samuel Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.

- [2] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].
- [3] Eva Beaujouan and Anne Solaz. Is the family size of parents and children still related? revisiting the cross-generational relationship over the last century. *Demography*, 56(2):595–619, 2019.
- [4] Jean Bertoin and Bastien Mallein. Reinforced Galton-Watson processes I: Malthusian exponents. *Random Structures Algorithms*, 65(2):387–410, 2024.
- [5] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. Adv. Appl. Probab., 36(2):544–581, 2004.
- [6] Rick Durrett. Probability. Theory and examples, volume 49 of Camb. Ser. Stat. Probab. Math. Cambridge: Cambridge University Press, 5th edition edition, 2019.
- [7] Sven Gronauer and Klaus Diepold. Multi-agent deep reinforcement learning: a survey. Artificial Intelligence Review, 55(2):895–943, 2022.
- [8] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes Appl.*, 110(2):171–245, 2004.
- [9] Michael Murphy. Cross-national patterns of intergenerational continuities in childbearing in developed countries. *Biodemography and social biology*, 59(2):101–126, 2013.
- [10] Burrhus Frederic Skinner. Science and human behavior. Number 92904. Simon and Schuster, 1965.