



HAL
open science

Reinforced Galton-Watson processes I: Malthusian exponents

Jean Bertoin, Bastien Mallein

► **To cite this version:**

Jean Bertoin, Bastien Mallein. Reinforced Galton-Watson processes I: Malthusian exponents. *Random Structures and Algorithms*, 2023, 65 (2), pp.387-410. 10.1002/rsa.21219 . hal-04328349

HAL Id: hal-04328349

<https://hal.science/hal-04328349v1>

Submitted on 7 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reinforced Galton-Watson processes I: Malthusian exponents

Jean Bertoin* Bastien Mallein†

December 7, 2023

Abstract

In a reinforced Galton-Watson process with reproduction law ν and memory parameter $q \in (0, 1)$, the number of children of a typical individual either, with probability q , repeats that of one of its forebears picked uniformly at random, or, with complementary probability $1 - q$, is given by an independent sample from ν . We estimate the average size of the population at a large generation, and in particular, we determine explicitly the Malthusian growth rate in terms of ν and q . Our approach via the analysis of transport equations owes much to works by Flajolet and co-authors.

Keywords: Galton-Watson process, Malthusian growth exponent, stochastic reinforcement, transport equation, singularity analysis of generating functions.

AMS subject classifications: 60J80; 60E10; 35Q49

1 Introduction and main result

We present first some motivation for this work, then our main result, and finally, we discuss techniques and sketch the plan of the rest of the article.

1.1 Motivations

Roughly speaking, stochastic reinforcement refers to step-by-step modifications of the dynamics of a random process such that transitions that already occurred often in the past are more likely to be repeated in the future. The evolution thus depends on the past of the process and not just on its current state, and is typically non-Markovian¹. We refer to [13] for a survey of various probabilistic models in this area, and merely recall that the concept of reinforcement notably lies at the heart of machine learning. The general question of describing how reinforcement impacts the long time behavior of processes has been intensively

*Institute of Mathematics, University of Zurich, Switzerland.

†Institut de Mathématiques de Toulouse, Université de Toulouse, France.

¹Nonetheless, it is a remarkable fact observed first by Coppersmith and Diaconis in an unpublished manuscript that linearly edge-reinforced random walks on a finite graph are actually mixtures of reversible Markov chains, with the mixing measure given by the so-called “magic formula”.

investigated for many years. In particular, there is a rich literature on so-called linear edge/vertex reinforced random walks, which culminates with the remarkable achievement of Sabot and Zeng [14] in which many important earlier references can also be found. In short, the purpose of the present work is to investigate the effects of reinforcement on branching processes.

More precisely, fix a reproduction law ν , that is $\nu = (\nu(k))$ is a probability measure on $\mathbb{Z}_+ = \{0, 1, \dots\}$. In a classical (Bienaymé-)Galton-Watson process, every individual reproduces independently of the others, and $\nu(k)$ is the probability that a typical individual has k children at the next generation. We consider a reinforced version of this model, in which the reinforcement depends on a memory parameter $q \in (0, 1)$. For every individual, say i , at a given generation $n \geq 1$, we pick a forebear, say $f(i)$, uniformly at random on its ancestral lineage. That is, $f(i)$ is the ancestor of i at generation $u(n)$, where $u(n)$ is a uniform random sample from $\{0, \dots, n-1\}$. Then, either with probability q , i begets the same number of children as $f(i)$, or with complementary probability $1 - q$, the number of children of i is given by an independent sample from the reproduction law ν . Implicitly, for different individuals i , we use independent uniform variables for selecting forebears $f(i)$ on their respective ancestral lineages, and the repetition events where given individuals i reproduce as their respective selected forebears $f(i)$ are also assumed to be independent. The reinforced version that we took interest in involves random repetitions of reproduction events, much in the same way as words are repeated in a well-known algorithm of Herbert Simon [15], or steps in a family of discrete time processes with memory known as step-reinforced random walks (see [2, 3, 4, 12] and references therein).

The reproduction of an individual in a reinforced Galton-Watson process depends on that along its entire ancestral lineage. It should be clear that, with the exception of the case of a few elementary reproduction laws, this invalidates fundamental features of Galton-Watson processes. Notably, different individuals partly share the same ancestral lineage, and thus their offspring is not independent. The Markov and the branching properties are lost, and even the simplest questions about Galton-Watson processes become challenging in the reinforced setting.

One naturally expects that reinforcement should enhance the growth of branching processes. Let us briefly analyze the elementary case of a binary branching when $\nu(2) = p$ and $\nu(0) = 1 - p$ for some $p \in (0, 1)$. Plainly, for every non-empty generation, all the forebears on any ancestral lineage had 2 children, and therefore the reinforced Galton-Watson process is in turn a binary Galton-Watson process, such that every individual aside from the ancestor has probability $q + (1 - q)p > p$ of having two children. In particular, the reinforced process survives forever with positive probability if and only if $q + (1 - q)p > 1/2$, which includes some subcritical reproduction laws with $p < 1/2$ (for example it can survive for all $p > 0$ whenever $q > 1/2$).

In the next section, we describe our main result on how reinforcement affects the Malthusian exponent, that is the growth rate of the averaged size of the population. In another forthcoming work, we shall investigate in turn the impact of reinforcement on survival probabilities.

1.2 Main result

We first introduce some notation; the memory parameter $q \in (0, 1)$ and the reproduction law ν have been fixed and thus most often are omitted from the notation. The degenerate case when $\nu(k) = 1$ for some $k \in \mathbb{Z}_+$ will be implicitly excluded. We write $Z = (Z(n))_{n \geq 0}$ for the reinforced Galton-Watson process, where $Z(n)$ is the size of the population at the n -th generation. We agree for simplicity that there is a single ancestor at generation 0 (i.e. $Z(0) = 1$) which has $Z(1)$ children, and denote the conditional distribution of the reinforced Galton-Watson process given $Z(1) = \ell$ by \mathbb{P}_ℓ . We are mainly interested in the situation where $Z(1)$ is random and distributed according to the reproduction law ν , and we shall then simply write $\mathbb{P} = \mathbb{P}_\nu$.

Perhaps the most basic question about Z is to estimate the average size of the population for large times; let us start with a few elementary observations. Reinforcement entails that the conditional probability for an individual to have k children, given that each of its ancestors also had k children, equals $q + (1 - q)\nu(k)$. By focussing on such individuals, we see that there is a sub-population of the reinforced Galton-Watson process that evolves as an ordinary Galton-Watson process with averaged reproduction size equal to $k(q + (1 - q)\nu(k)) > kq$. This points at the roles of the support of the reproduction law ν ,

$$\text{Supp}(\nu) := \{k \geq 0 : \nu(k) > 0\},$$

and of the maximal possible number of children which we denote by

$$k^* := \sup\{k \geq 0 : \nu(k) > 0\}.$$

Namely, if $qk^* \geq 1$, then not only does the reinforced Galton-Watson process survive with positive probability, but the averaged population size grows at least exponentially fast with $\mathbb{E}(Z(n)) \geq c(qk^*)^n$ for some $c > 0$. We stress that a high averaged growth rate is achieved as soon as the reproduction law has an atom at some large k , no matter how small the mass $\nu(k)$ of this atom is.

We assume henceforth that the support of the reproduction law is bounded, viz.

$$k^* < \infty,$$

and stress that otherwise, the averaged size of the population would grow super-exponentially fast. To state our main result, we introduce the function

$$\Pi(t) := \prod (1 - tk)^{\nu(k)(1-q)/q}, \quad t \leq 1/k^*, \quad (1.1)$$

where the product on the right-hand side is taken over all integers k (actually, only the positive k in $\text{Supp}(\nu)$ contribute to this product). Observe that $\Pi(0) = 1$, $1/k^*$ is the sole zero of Π , and Π is strictly decreasing on $[0, \frac{1}{k^*}]$. We can then define

$$m_{\nu,q} := \frac{q}{\int_0^{1/k^*} \Pi(t) dt}. \quad (1.2)$$

In particular, as $\Pi(t) < 1$ for all $0 < t \leq 1/k^*$, we remark that $m_{\nu,q} > qk^*$.

Theorem 1.1. *The Malthusian exponent of the reinforced Galton-Watson process Z equals $\log m_{\nu,q}$, in the sense that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}(Z(n)) = \log m_{\nu,q}.$$

Much more precisely, we have

$$m_{\nu,q}^{-n} \mathbb{E}(Z(n)) = \frac{\nu(k^*)}{q + \nu(k^*)(1-q)} + O(n^{-q/(q+(1-q)\nu(k^*))}).$$

Plainly, if $Z(n) = 0$ for some $n \geq 1$, then also $Z(n') = 0$ for all $n' \geq n$ and one says that the population becomes extinct eventually. Applying the Markov inequality to $Z(n)$, we remark that $m_{\nu,q} < 1$ implies that almost surely, the reinforced Galton-Watson process becomes extinct eventually. Therefore the condition $m_{\nu,q} \geq 1$ is necessary for the survival of the reinforced Galton-Watson process with strictly positive probability. We conjecture however that this condition should not be sufficient; we believe that there should exist reproduction laws ν and memory parameters q such that, almost surely, the reinforced Galton-Watson process becomes extinct eventually, even though $\mathbb{E}(Z(n)) \rightarrow \infty$ as $n \rightarrow \infty$. We give in a companion article [5] a sufficient condition for the survival of the reinforced Galton-Watson process, which does not match the necessary condition $m_{\nu,q} \geq 1$. However, under the condition $((1-q)\nu(k^*) + q)k^*$ (which implies $m_{\nu,q} > 1$), we show that the Malthusian exponent $m_{\nu,q}$ matches the a.s. exponential growth of Z .

1.3 Techniques and plan

Although reinforcement invalidates fundamental properties of Galton-Watson processes, our strategy for establishing Theorem 1.1 nonetheless uses some classical tools from the theory of branching processes, and notably analytic methods. These techniques can be pushed further and yield an asymptotic expansion of $\mathbb{E}(Z(n))$ with an arbitrary order. It is for the sake of readability only that we just consider the first order in this work.

In Section 2, we start by relating the expected population size for the reinforced Galton-Watson process to the probability generating functions of a multitype Yule process; see Lemma 2.1. In short, this combines the basic many-to-one formula with the classical embedding of urn schemes into continuous time multitype branching processes (see, for instance, [1, Section V.9]).

In Section 3, we observe that those probability generating functions satisfy a multidimensional ODE; see Lemma 3.3 there. This is similar to the use of analytic methods that were introduced in the pioneering works of Flajolet *et al.* [6, 7] for determining distributions in certain Pólya type urn processes.

In Section 4, we shall perform several transformations to reduce the multidimensional ODE obtained above to a one-dimensional inhomogeneous transport equation with growth. We then analyze characteristic curves and in particular, we determine the parameters for which the probability generating functions explode; see Proposition 4.3. This points at a critical case, which corresponds to the situation when the quantity $m_{\nu,q}$ in (1.2) equals 1.

Section 5 is then devoted to the study of the asymptotic behavior of the probability generating functions in the critical case. Our main result there, Proposition 5.1, shows that this problem fits the framework of singular analysis of generating functions; see [8, 9].

Theorem 1.1 is established in Section 6; it suffices to combine the elements developed in the preceding sections. Last, some comments and further results are presented in Section 7. We notably discuss a sharper version of Theorem 1.1

under the law \mathbb{P}_ℓ , that is given that the ancestor has ℓ children. We also give some explicit bounds on $m_{\nu,q}$ and show it to be continuous as $q \rightarrow 0$ and $q \rightarrow 1$.

2 A multitype Yule process

We introduce a standard Yule process $Y = (Y(t))_{t \geq 0}$, that is Y is a pure birth process started from $Y(0) = 1$ and such that the rate of jump from k to $k + 1$ equals k for all $k \geq 1$. We think of Y as the process recording the size of a population evolving in continuous time, such that individuals are immortal and each individual gives birth to a child at unit rate and independently of the other individuals. The interpretation in terms of population models is that at any birth event, the parent that begets is chosen uniformly at random in the population immediately before the birth event. We enumerate individuals according to the increasing order of their birth times and then record the genealogy as a combinatorial tree \mathbf{T} on $\mathbb{N} = \{1, \dots\}$. The distribution of \mathbf{T} is that of an infinite random recursive tree, that is such that for every fixed $n \in \mathbb{N}$, its restriction to $\{1, \dots, n\}$ has the uniform distribution on the set of trees rooted at 1 and such that the sequence of vertices along any branch from the root is increasing. Clearly enough, the tree \mathbf{T} and the Yule process Y are independent (the latter only records the size of the population as time passes and ignores the genealogy).

Next, we randomly and recursively assign types to the individuals in the population, that is to the vertices of \mathbf{T} , where the space of types is $\text{Supp}(\nu)$, the support of the reproduction law ν . For every $n \geq 2$, conditionally on the genealogical tree \mathbf{T} and the types already assigned to the $n - 1$ first individuals, if the type of the parent of n is j , then the probability that the type of n is k equals

$$q\mathbf{1}_{\{k=j\}} + (1 - q)\nu(k).$$

We should think of the assignment of types as a random process on \mathbf{T} , such that for each individual $i \geq 2$ (the ancestor 1 is excluded), with probability q the type of i merely repeats that of its parent, and with complementary probability $1 - q$, it is chosen at random according to the reproduction law ν . We stress that the assignment of types is performed independently of the Yule process Y . If originally, the type ℓ is assigned to the ancestor 1, then we write \mathbf{P}_ℓ for the distribution of the population process with types, and \mathbf{P}_ν , or simply \mathbf{P} , when the type of 1 is random with law ν .

For every $t \geq 0$, let us write $Y_j(t)$ for the number of individuals of type j in the population at time t , and $\mathbf{Y}(t) = (Y_j(t))$ for the sequence indexed by $j \in \text{Supp}(\nu)$. Then the process $\mathbf{Y} = (\mathbf{Y}(t))_{t \geq 0}$ is a multitype Yule process such that every individual begets a child bearing the same type with rate q , and a child with type chosen independently according to ν with rate $1 - q$.

We point out the similarity, but also the difference, with the reinforcement dynamic that motivates this work. In both cases, the newborn individual repeats with probability q the type, or the offspring number, of a forebear, and is otherwise assigned an independent type, or offspring number, distributed according to the fixed law ν . However, the forebear is merely the parent of the new individual in the multitype Yule setting, whereas it is picked uniformly at random on the ancestral lineage in the reinforced Galton-Watson setting.

We can now state the connection between the reinforced Galton-Watson process and the multitype Yule process which is the first step of our analysis. In this direction, recall that \mathbb{P}_ℓ stands for the conditional distribution of the reinforced Galton-Watson process given $Z(1) = \ell$.

Lemma 2.1. *For every $t \geq 0$, $c > 0$, and $\ell \in \text{Supp}(\nu)$, there is the identity*

$$\mathbf{E}_\ell \left(\prod (cj)^{Y_j(t)} \right) = e^{-t} \sum_{n=1}^{\infty} (1 - e^{-t})^{n-1} c^n \mathbb{E}_\ell(Z(n)),$$

where in the left-hand side, the product implicitly runs over $j \in \text{Supp}(\nu)$ and we use the convention $0^0 = 1$.

Proof. Consider any deterministic genealogical tree \mathcal{T} and construct iteratively a sequence of random variables (ζ_n) recording the sibling sizes along a randomly chosen descent lineage. Specifically, we first let ζ_1 be the number of children of the root of \mathcal{T} . If $\zeta_1 = 0$, then we also set $\zeta_n = 0$ for all $n \geq 2$. Otherwise, we select an individual chosen uniformly at random in the first generation, and write ζ_2 for the number of its children at the second generation, and iterate in an obvious way. That is, if $\zeta_2 = 0$, then $\zeta_n = 0$ for all $n \geq 3$. Otherwise, we select a child uniformly at random among these siblings. We continue this procedure, defining ζ_3, ζ_4 , etc. for the number of children of the individual selected at the second, third, etc., generation. We then observe that for any $n \in \mathbb{N}$, writing $\mathcal{Z}(n)$ for the number of individuals at the n th generation of \mathcal{T} , we have

$$\mathcal{Z}(n) = \mathbb{E}(\zeta_1 \times \cdots \times \zeta_n).$$

Indeed, observe that the probability for a given individual in the n th generation to be chosen is equal to the inverse of the product of the sibling sizes along its ancestral path.

In particular, in the case of the reinforced Galton-Watson process, we get

$$\mathbb{E}_\ell(Z(n)) = \mathbb{E}_\ell(\zeta_1 \times \cdots \times \zeta_n). \quad (2.1)$$

The reinforcement dynamics entail that, provided that $\zeta_n \geq 1$, the conditional probability that $\zeta_{n+1} = k$ given the first n values $(\zeta_i)_{1 \leq i \leq n}$ is given by

$$qn^{-1} \sum_{i=1}^n \mathbf{1}_{\{\zeta_i=k\}} + (1-q)\nu(k).$$

Comparing with the evolution of the multitype Yule process \mathbf{Y} , we realize that for any $n \geq 1$, the conditional distribution of $(\zeta_i)_{1 \leq i \leq n+1}$ given $\zeta_1 = \ell$ and $\zeta_n \neq 0$ is the same as the conditional distribution under \mathbf{P}_ℓ of the sequence of the types of the first $n+1$ individuals of the multitype Yule process given that no individuals amongst the first n have type 0. Moreover, we also have

$$\mathbb{P}_\ell(\zeta_n = 0) = \mathbf{P}_\ell(Y_0(\beta_n) \geq 1),$$

where $\beta_n = \inf\{t \geq 0 : Y(t) = n\}$ stands for the birth time of the n -th individual in the multitype Yule process. As a consequence, we have

$$\mathbb{E}_\ell(\zeta_1 \times \cdots \times \zeta_n) = \mathbf{E}_\ell \left(\prod j^{Y_j(\beta_n)} \right),$$

where, as usual, the product on the right-hand side runs over $j \in \text{Supp}(\nu)$. On the other hand, recalling that the monotype Yule process $Y = \sum Y_j$ is independent of the assignment of types, we get that for any c and $t > 0$, there is the identity

$$c^n \mathbf{E}_\ell \left(\prod j^{Y_j(\beta_n)} \right) = \mathbf{E}_\ell \left(\prod (cj)^{Y_j(t)} \mid Y(t) = n \right),$$

Putting the pieces together, we arrive at

$$\mathbf{E}_\ell \left(\prod (cj)^{Y_j(t)} \mid Y(t) = n \right) = c^n \mathbb{E}_\ell(Z(n)).$$

We can now conclude the proof by recalling the basic fact that $Y(t)$ has the geometric distribution with parameter e^{-t} . \square

We conclude this section by pointing out a monotonicity property in the memory parameter q , which may be rather intuitive even though its proof is not completely obvious. We also stress that this property is valid under $\mathbb{P} = \mathbb{P}_\nu$, but might fail under some \mathbb{P}_ℓ .

Lemma 2.2. *For every fixed $n \geq 1$, the average population size $\mathbb{E}(Z(n))$ is a monotone increasing function of the memory parameter q .*

Proof. We will use a coupling argument for a simple variation of a Pólya urn model with random replacement and balls labelled $\textcircled{1}, \textcircled{2}, \dots$. Imagine an urn which contains initially just one ball $\textcircled{1}$, and let $(\varepsilon_n)_{n \geq 1}$ be a sequence of i.i.d. Bernoulli variables with parameter q . At each step $n = 1, \dots$, we draw a ball uniformly at random from the urn and return it together with either a copy of the sampled ball if $\varepsilon_n = 1$, or a new ball bearing the first label which was not already present in the urn if $\varepsilon_n = 0$. For any $k \geq 1$, denote by $N_k(n)$ the number of balls \textcircled{k} in the urn after $n - 1$ steps, i.e. when the urn contains in total n balls.

Then consider also a sequence $(\xi_k)_{k \geq 1}$ of i.i.d. samples from ν , independent of the urn process. We get from the many-to-one formula (2.1) (recall the convention $0^0 = 1$) that

$$\mathbb{E}(Z(n)) = \mathbb{E} \left(\prod_{k \geq 1} \xi_k^{N_k(n)} \right) = \mathbb{E} \left(\prod_{k \geq 1} m_\nu(N_k(n)) \right),$$

where on the right-hand side, we wrote

$$m_\nu(\ell) := \mathbb{E}(\xi^\ell) = \sum_j j^\ell \nu(j), \quad \ell \geq 0$$

for the ℓ -th moment of the reproduction law.

Now let $q' > q$. By an elementary coupling, we get a sequence $(\varepsilon'_n)_{n \geq 1}$ of i.i.d. Bernoulli variables with parameter q' such that $\varepsilon'_n \geq \varepsilon_n$ for all $n \geq 1$. We can use these to couple the urn processes with respective memory parameters q and q' , so that the second one is obtained from the first by some (random) non-injective relabelling of the balls. Since we know from Jensen's inequality that

$$m_\nu(\ell) \times m_\nu(\ell') \leq m_\nu(\ell + \ell'),$$

we deduce that

$$\prod_{k \geq 1} m_{\nu}(N_k(n)) \leq \prod_{k \geq 1} m_{\nu}(N'_k(n)),$$

where on the right-hand side, $N'_k(n)$ stands for the number of balls \textcircled{k} in the urn with memory parameter q' after $n - 1$ steps. Taking expectations completes the proof. \square

3 Probability generating functions of the multi-type Yule process

Lemma 2.1 incites us to investigate the probability generating functions (i.e. the joint factorial moment generating functions) $\mathbf{M} = (M_{\ell})$ of the multitype Yule process \mathbf{Y} , where

$$M_{\ell}(\mathbf{a}, t) := \mathbf{E}_{\ell} \left(\prod a_j^{Y_j(t)} \right), \quad t \geq 0 \text{ and } \mathbf{a} = (a_j) \in \mathbb{R}_+^{\text{Supp}(\nu)}.$$

Clearly, $M_{\ell}(\mathbf{a}, t) \equiv 0$ if $a_{\ell} = 0$, and otherwise it is a strictly positive quantity which may be infinite when $\|\mathbf{a}\|_{\infty} > 1$ and t is sufficiently large, with the notation $\|\cdot\|_{\infty}$ for the maximum norm on $\mathbb{R}^{\text{Supp}(\nu)}$. We first make the following simple observation in this direction. From now on, the trivial case when $\mathbf{a} = 0$ will be systematically ruled out.

Lemma 3.1. *For every $\mathbf{a} \in \mathbb{R}_+^{\text{Supp}(\nu)}$, there exists $\varrho(\mathbf{a}) \in (0, \infty]$ such that for every $\ell \in \text{Supp}(\nu)$ with $a_{\ell} > 0$,*

$$M_{\ell}(\mathbf{a}, t) \begin{cases} < \infty & \text{for all } 0 \leq t < \varrho(\mathbf{a}), \\ = \infty & \text{for all } t > \varrho(\mathbf{a}). \end{cases}$$

Moreover, $t \mapsto M_{\ell}(\mathbf{a}, -\log(1 - t))$ is an entire function with radius of convergence no less than $1 - e^{-\varrho(\mathbf{a})}$.

Remark 3.2. For instance, if $a_j = c > 1$ for all j , then the probability generating functions of the multitype Yule process \mathbf{Y} reduce to that of the monotype Yule process Y , and we immediately get

$$M_{\ell}(\mathbf{a}, t) = \frac{ce^{-t}}{1 - c(1 - e^{-t})} \quad \text{for } t < \varrho_{\ell}(\mathbf{a}) = -\log(1 - 1/c).$$

Proof. Using the independence of the Yule process and the assignment of types as at the end of the proof of Lemma 2.1, we see that

$$M_{\ell}(\mathbf{a}, t) = e^{-t} f(1 - e^{-t}),$$

where $f(x) = \sum_{n=0}^{\infty} a_n x^n$ is some power series with positive coefficients (for notational simplicity, we do not indicate the dependency in ℓ). Writing $r > 0$ for the radius of convergence of f and then setting

$$\varrho_{\ell}(\mathbf{a}) := \begin{cases} -\log(1 - r) & \text{if } r < 1, \\ \infty & \text{if } r \geq 1, \end{cases}$$

yields our claims provided that $\varrho_{\ell}(\mathbf{a})$ does not depend on the index ℓ .

It thus only remains to check that $\varrho_\ell(\mathbf{a}) \geq \varrho_j(\mathbf{a})$ for any $j \in \text{Supp}(\boldsymbol{\nu})$ with $a_j > 0$. We may assume that $\varrho_\ell(\mathbf{a}) < \infty$ since otherwise there is nothing to prove. Take any $t' > t > \varrho_\ell(\mathbf{a})$, and work under the law \mathbf{P}_j when the type of ancestor is j . The probability $p(t' - t)$ that at time $t' - t$ the population has exactly two individuals, one with type j and the other with type ℓ , is strictly positive. The branching property then shows that

$$M_j(\mathbf{a}, t') \geq p(t' - t)M_\ell(\mathbf{a}, t)M_j(\mathbf{a}, t) = \infty.$$

Hence $\varrho_j(\mathbf{a}) \leq t'$, and the proof is now complete. \square

We shall henceforth refer to $\varrho(\mathbf{a})$ as the explosion time. The problem of determining its value is at the heart of the proof of Theorem 1.1. For instance, recall from the Cauchy–Hadamard theorem and Lemma 2.1 that, taking $a_j = cj$ for all $j \in \text{Supp}(\boldsymbol{\nu})$ and any c sufficiently large so that $\varrho(\mathbf{a}) < \infty$, we obtain the identity

$$\limsup_{n \rightarrow \infty} n^{-1} \log \mathbb{E}(Z(n)) = -\log(1 - e^{-\varrho(\mathbf{a})}) - \log c.$$

The first step for its resolution is closely related to the analytic approach² of two-color Pólya urn models by Flajolet *et al.* [6, 7], see also [10] for a version with random replacement schemes. It relies on the observation that the probability generating functions solves a multidimensional ODE.

We use the notation $\langle \cdot; \cdot \rangle$ for the scalar product on $\mathbb{R}^{\text{Supp}(\boldsymbol{\nu})}$ and set

$$\varphi(t) := (1 - q) \langle \boldsymbol{\nu}; \mathbf{M}(\mathbf{a}, t) \rangle - 1. \quad (3.1)$$

For every $\ell \in \text{Supp}(\boldsymbol{\nu})$ and $t < \varrho(\mathbf{a})$, we also write

$$M'_\ell(\mathbf{a}, t) := \frac{\partial M_\ell}{\partial t}(\mathbf{a}, t).$$

Lemma 3.3. *For every \mathbf{a} , the function $\mathbf{M}(\mathbf{a}, \cdot)$ is the unique solution on $[0, \varrho(\mathbf{a}))$ to*

$$M'_\ell(\mathbf{a}, t) = M_\ell(\mathbf{a}, t) (qM_\ell(\mathbf{a}, t) + \varphi(t)), \quad \ell \in \text{Supp}(\boldsymbol{\nu}),$$

with initial condition $\mathbf{M}(\mathbf{a}, 0) = \mathbf{a}$.

Observe that the differential equation satisfied by $\log M$ is analogous to the one satisfied by the Laplace exponent associated to continuous state branching processes; see e.g. [11, Theorem 12.1].

Proof. We denote by T the first branching time of the multitype Yule process and by L the random label of the newly added vertex. Applying the branching property at time T , we have

$$\begin{aligned} M_\ell(\mathbf{a}, t) &= a_\ell \mathbb{P}(T > t) + \mathbb{E}(M_\ell(\mathbf{a}, t - T)M_L(\mathbf{a}, t - T)) \\ &= a_\ell e^{-t} + \int_0^t e^{-s} M_\ell(\mathbf{a}, t - s) (qM_\ell(\mathbf{a}, t - s) + \varphi(t - s) + 1) ds. \end{aligned}$$

²Note however that these works deal with exponential generating functions, whereas we rather consider here their factorial moment version as in Lemma 2.1.

Hence, by change of variable $u = t - s$, we obtain

$$e^t M_\ell(\mathbf{a}, t) = a_\ell + \int_0^t e^u M_\ell(\mathbf{a}, u) (q M_\ell(\mathbf{a}, u) + \varphi(u) + 1) du.$$

Taking the derivative of this expression in t completes the proof. \square

We finish this section by pointing out simple monotonicity properties of the probability generating functions and their explosion times.

Corollary 3.4. (i) For every $j, \ell \in \text{Supp}(\nu)$ with $a_j \leq a_\ell$, the ratio function

$$t \mapsto \frac{M_\ell(\mathbf{a}, t)}{M_j(\mathbf{a}, t)}$$

is monotone non-decreasing on $[0, \varrho(\mathbf{a})]$. In particular, if $a_j = a_\ell$, then for all $t \geq 0$ we have $M_j(\mathbf{a}, t) = M_\ell(\mathbf{a}, t)$.

(ii) Suppose that

$$\sup_{t \geq 0} M_\ell(\mathbf{a}, t) < \infty.$$

Then for all $j \in \text{Supp}(\nu)$ with $a_j < a_\ell$, we have

$$\lim_{t \rightarrow \infty} M_j(\mathbf{a}, t) = 0.$$

Proof. The claims readily derive from the observation from Lemma 3.3 that the logarithmic derivative of the ratio function equals $q(M_\ell(\mathbf{a}, t) - M_j(\mathbf{a}, t))$. \square

4 Determining the probability generating function

The purpose of this section is to show that the probability generating functions $\mathbf{M}(\mathbf{a}, t)$ can be determined rather explicitly, and this will be sufficient for our purposes. To do so, it is convenient to introduce a bivariate moment-generating function G defined by

$$G(t, s) := \sum_j \frac{M_j(\mathbf{a}, t)}{1 - s M_j(\mathbf{a}, t)} \nu(j) \quad (4.1)$$

for any $0 \leq t < \varrho(\mathbf{a})$ and $0 \leq s < 1/\|\mathbf{M}(\mathbf{a}, t)\|_\infty$. Note that

$$G(t, 0) = \langle \nu; \mathbf{M}(\mathbf{a}, t) \rangle = \mathbf{E}_\nu \left(\prod_j a_j^{Y_j(t)} \right)$$

is the function with explosion time $\varrho(\mathbf{a})$ that we would like to determine, and that

$$G(0, s) = \sum_j \frac{a_j}{1 - s a_j} \nu(j) \quad (4.2)$$

is a known function of the variable s .

The motivation for introducing G is that it satisfies an inhomogeneous transport equation with growth in dimension 1.

Lemma 4.1. For all $0 \leq t < \varrho(\mathbf{a})$ and $0 \leq s < 1/\|\mathbf{M}(\mathbf{a}, t)\|_\infty$, the function G satisfies

$$\partial_t G(t, s) = (q + s\varphi(t))\partial_s G(t, s) + \varphi(t)G(t, s),$$

where $\varphi(t)$ has been defined in (3.1).

Proof. For all $k \geq 1$, we introduce the function $R_k : [0, \varrho(\mathbf{a})] \rightarrow \mathbb{R}_+$ by

$$R_k(t) := \sum_j \nu(j) M_j(\mathbf{a}, t)^k.$$

The function $G(t, \cdot)$ arises in this setting as the moment-generating function of the sequence $(R_k(t))_{k \geq 1}$. Namely, we have for every $0 \leq s < 1/\|\mathbf{M}(\mathbf{a}, t)\|_\infty$ that

$$\sum_{k=1}^{\infty} s^{k-1} R_k(t) = \sum_j \nu(j) \sum_{k=1}^{\infty} s^{k-1} M_j(\mathbf{a}, t)^k = G(t, s).$$

Note also that for $k = 1$, we have

$$R_1(t) = \langle \boldsymbol{\nu}; \mathbf{M}(\mathbf{a}, t) \rangle = G(t, 0).$$

We compute the derivative of R_k using Lemma 3.3 in the second identity below, and get

$$\begin{aligned} R'_k(t) &= k \sum_j \nu(j) M_j(\mathbf{a}, t)^k \frac{M'_j(\mathbf{a}, t)}{M_j(\mathbf{a}, t)} \\ &= k (qR_{k+1}(t) + \varphi(t)R_k(t)). \end{aligned}$$

This yields

$$\begin{aligned} \partial_t G(t, s) &= \sum_{k=1}^{\infty} k s^{k-1} (qR_{k+1}(t) + \varphi(t)R_k(t)) \\ &= \partial_s \left(\sum_{k=1}^{\infty} s^k (qR_{k+1}(t) + \varphi(t)R_k(t)) \right) \\ &= \partial_s (qG(t, s) + \varphi(t)sG(t, s)). \end{aligned}$$

We complete the proof by applying the chain rule. \square

Lemma 4.1 may not look quite satisfactory, because the coefficients of the PDE there involve the unknown function φ defined in (3.1), which is precisely the one we would like to determine. However, its main input is that this PDE can be analyzed in terms of a trajectory on the real line, as opposed to the multidimensional ODE of Lemma 3.3. Lemma 4.1 shows indeed that the function G defined in (4.1) satisfies an inhomogeneous transport equation, and as a consequence,

$$\varphi(t) = (1 - q)G(t, 0) - 1$$

can be obtained as the solution of an integral equation as we shall now see.

Lemma 4.2. *Introduce the functions*

$$A : [0, \varrho(\mathbf{a})) \rightarrow \mathbb{R}_+, \quad A(t) = \int_0^t \exp\left(\int_0^s \varphi(u) du\right) ds,$$

and

$$F : [0, 1/(q\|\mathbf{a}\|_\infty)) \rightarrow \mathbb{R}_+, \quad F(s) = (1-q) \sum_j \frac{a_j}{1-qs a_j} \nu(j). \quad (4.3)$$

Then for $t_0 > 0$ sufficiently small, A is the unique solution on $[0, t_0]$ of the Cauchy problem

$$A''(t) = A'(t)(A'(t)F(A(t)) - 1) \quad \text{with } A(0) = 0 \text{ and } A'(0) = 1. \quad (4.4)$$

Proof. The method of characteristics applied to the PDE of Lemma 4.1 incites us to construct a pair of functions

$$\gamma : [0, t_0] \rightarrow \mathbb{R}_+ \quad \text{and} \quad \sigma : [0, t_0] \rightarrow \mathbb{R}_+,$$

such that $\sigma(t) < 1/\|\mathbf{M}(\mathbf{a}, t)\|_\infty$ and

$$\gamma(t)G(t, \sigma(t)) = \gamma(0)G(0, \sigma(0)) \quad \text{for all } 0 \leq t \leq t_0, \quad (4.5)$$

where $0 < t_0 < \varrho(\mathbf{a})$ will be chosen later on. That is, we require the function $t \mapsto \gamma(t)G(t, \sigma(t))$ to have derivative identical to 0. Thanks to Lemma 4.1, this translates into

$$\begin{aligned} & \gamma'(t)G(t, \sigma(t)) + \gamma(t)\partial_t G(t, \sigma(t)) + \gamma(t)\sigma'(t)\partial_s G(t, \sigma(t)) \\ &= G(t, \sigma(t))(\gamma'(t) + \varphi(t)\gamma(t)) + \gamma(t)\partial_s G(t, \sigma(t))(\sigma'(t) + q + \varphi(t)\sigma(t)) \\ &= 0. \end{aligned}$$

We deduce that (4.5) is satisfied provided that γ and σ solve the differential equations

$$\gamma' + \varphi\gamma = 0 \quad \text{and} \quad \sigma' + \varphi\sigma + q = 0.$$

These linear differential equations can be solved explicitly in terms of φ and q . Setting $\gamma(0) = 1$ and $\sigma(t_0) = 0$, in the notation of the statement, we can write

$$\gamma(t) = 1/A'(t) \quad \text{and} \quad \sigma(t) = q \int_t^{t_0} \exp\left(\int_t^s \varphi(u) du\right) ds.$$

We now choose $t_0 > 0$ sufficiently small so that $\sigma(t) < 1/\|\mathbf{M}(\mathbf{a}, t)\|_\infty$ for all $t \leq t_0$. Observe that $\sigma(0) = qA(t_0)$, and we then get from (4.5)

$$G(t_0, 0) = A'(t_0)G(0, qA(t_0)).$$

This allows us to rewrite the identity

$$\varphi(t_0) = (1-q)G(t_0, 0) - 1 = (1-q)A'(t_0)G(0, qA(t_0)) - 1.$$

Since $\varphi = A''/A'$, we conclude from (4.2) that A satisfies (4.4). \square

We can now solve the Cauchy problem (4.4). Introduce first

$$\Pi_{\mathbf{a}}(x) := \prod (1 - xa_j)^{\nu(j)(1-q)/q}, \quad 0 \leq x \leq 1/\|\mathbf{a}\|_{\infty}, \quad (4.6)$$

and its integral

$$I_{\mathbf{a}}(x) := \int_0^x \Pi_{\mathbf{a}}(y) dy, \quad 0 \leq x \leq 1/\|\mathbf{a}\|_{\infty}. \quad (4.7)$$

If we set

$$i_{\mathbf{a}} := I_{\mathbf{a}}(1/\|\mathbf{a}\|_{\infty}),$$

then $I_{\mathbf{a}}$ defines a bijection from $[0, 1/\|\mathbf{a}\|_{\infty}]$ to $[0, i_{\mathbf{a}}]$, and we write

$$I_{\mathbf{a}}^{-1} : [0, i_{\mathbf{a}}] \rightarrow [0, 1/\|\mathbf{a}\|_{\infty}]$$

for the inverse bijection.

Proposition 4.3. *The explosion time can be identified as*

$$\varrho(\mathbf{a}) = \begin{cases} -\log(1 - i_{\mathbf{a}}/q) & \text{if } i_{\mathbf{a}} < q, \\ \infty & \text{if } i_{\mathbf{a}} \geq q, \end{cases}$$

and for $x \in [0, \varrho(\mathbf{a}))$, we have

$$qA(x) = I_{\mathbf{a}}^{-1}(q(1 - e^{-x})). \quad (4.8)$$

Proof. Write A^{-1} for the inverse of the bijection $A : [0, \varrho(\mathbf{a})) \rightarrow [0, A(\varrho(\mathbf{a}))]$, and set $H := A' \circ A^{-1}$, so H is the reciprocal of the derivative of A^{-1} . Using (4.4), we compute the derivate of H on some small neighborhood of the origin as

$$H' = \frac{A'' \circ A^{-1}}{H} = HF - 1. \quad (4.9)$$

Plainly, we also have the initial value $H(0) = 1$.

It is now straightforward to solve this first order linear ODE for H . The function F defined in (4.3) bears a close connection to the logarithmic derivative of the function $\Pi_{\mathbf{a}}$ in (4.6), namely

$$F(x) = -\frac{q\Pi'_{\mathbf{a}}(qx)}{\Pi_{\mathbf{a}}(qx)}.$$

Using the notation (4.7), this yields

$$H(x) = \frac{q - I_{\mathbf{a}}(qx)}{q\Pi_{\mathbf{a}}(qx)}; \quad (4.10)$$

more precisely, this identity is valid for all sufficiently small $x > 0$.

Now observe from (4.10) that $1/H(x) = (A^{-1})'(x)$ is the derivative of

$$B(x) := -\log(1 - I_{\mathbf{a}}(qx)/q), \quad 0 \leq x < 1/(q\|\mathbf{a}\|_{\infty}).$$

The functions A^{-1} and B both vanish at $x = 0$ and have the same derivative on some right neighborhood of 0, on which they therefore coincide. Computing the inverse function B^{-1} , we arrive at (4.8), provided that $x > 0$ is sufficiently small. Tracing the definition, we infer from Lemma 3.1 that A is analytic on $[0, \varrho(\mathbf{a}))$, and $I_{\mathbf{a}}^{-1}$ is also an analytic function on $[0, i_{\mathbf{a}}]$. The radii of convergence of A and B^{-1} are the same, which yields the formula for the explosion time. We complete the proof by uniqueness of analytic continuation. \square

It is now an easy matter to express the probability generating function $M(\mathbf{a}, \cdot)$ in terms of the function A that has just been determined.

Corollary 4.4. *For every \mathbf{a} and $\ell \in \text{Supp}(\boldsymbol{\nu})$ with $a_\ell > 0$, we have*

$$M_\ell(\mathbf{a}, t) = \frac{a_\ell A'(t)}{1 - qa_\ell A(t)}, \quad t \in [0, \varrho(\mathbf{a})].$$

Proof. We re-visit the equation in Lemma 3.3 and now view it as a Bernoulli differential equation involving the known function φ . The latter can be solved explicitly using the transformation $y(t) = 1/M_\ell(\mathbf{a}, t)$; this yields the linear ODE

$$-y' = q + \varphi y.$$

The solution for the initial condition $y(0) = 1/a_\ell$ is given by

$$y(t) = \frac{a_\ell \exp\left(\int_0^t \varphi(s) ds\right)}{1 - qa_\ell \int_0^t \left(\exp\left(\int_0^r \varphi(s) ds\right)\right) dr}, \quad t \in [0, \varrho(\mathbf{a})].$$

Recalling the notation in Lemma 4.2, we arrive at the stated formula. \square

5 Singularity analysis of the critical case

This section focuses on the critical case where

$$i_{\mathbf{a}} = q, \tag{5.1}$$

so by Proposition 4.3, the explosion time $\varrho(\mathbf{a})$ is infinite. For the sake of simplicity, we also require that there exists a unique $j_1 \in \text{Supp}(\boldsymbol{\nu})$ such that $a_{j_1} = \|\mathbf{a}\|_\infty$, i.e.

$$a_{j_1} > a_j \quad \text{for all } j \neq j_1. \tag{5.2}$$

Although the situation where the maximum is attained for two or more indices can be treated similarly, only (5.2) will be relevant for the proof of Theorem 1.1.

Using the notation $\log : \mathbb{C} \setminus \mathbb{R}_- \rightarrow \mathbb{C}$ for the principal determination of the complex logarithm and

$$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$$

for the open unit disk in \mathbb{C} , we know from (3.1) and the proof of Lemma 3.1 that $\varphi(-\log(1-z))$ defines an analytic function on \mathbb{D} . In short, our main purpose is show that Φ can be extended analytically to a domain

$$\Delta(R, \theta) := \{z \in \mathbb{C} : |z| < R, z \neq 1 \text{ and } |\arg(z-1)| > \theta\}$$

for some $R > 1$ and acute angle $\theta \in (0, \pi/2)$, and to analyze the asymptotic behavior as z approaches 1 in Δ .

Proposition 5.1. *The following holds under the assumptions (5.1) and (5.2). For any acute angle $\theta \in (0, \pi/2)$, there exists some $R = R(\theta) > 1$ such that the*

function $z \mapsto \varphi(-\log(1-z))$ can be extended analytically to $\Delta(R, \theta)$. Moreover, as z approaches 1 in $\Delta(R, \theta)$, one has

$$\varphi(-\log(1-z)) + 1/\beta = O(|1-z|^{1/\beta}),$$

where

$$\beta := 1 + (1-q)\nu(j_1)/q > 1. \quad (5.3)$$

The most of rest of this section is devoted to the proof of Proposition 5.1; we will also present some direct consequences at the end. The assumptions (5.1) and (5.2) being implicitly enforced. We further write

$$a_{j_2} := \max\{a_j : j \in \text{Supp}(\nu), j \neq j_1\} < a_{j_1}.$$

Introduce the complex planes slit along the real half-line $[1/a_{j_i}, \infty)$,

$$\mathbb{C}_i := \{z \in \mathbb{C} : \Im z \neq 0 \text{ or } \Re z < 1/a_{j_i}\}, \quad i = 1, 2.$$

We now view $\Pi_{\mathbf{a}}$ defined in (4.6) as a holomorphic function on \mathbb{C}_1 that has no zeros on this domain. Recalling (4.7), we also write $I_{\mathbf{a}}$ for its primitive on \mathbb{C}_1 with $I_{\mathbf{a}}(0) = 0$. The following elementary facts are the keys of our analysis.

Lemma 5.2. (i) *The function*

$$\frac{q - I_{\mathbf{a}}(z)}{q\Pi_{\mathbf{a}}(z)}, \quad z \in \mathbb{C}_1$$

can be extended analytically to \mathbb{C}_2 . Moreover one has

$$\frac{q - I_{\mathbf{a}}(z)}{q\Pi_{\mathbf{a}}(z)} \sim \frac{1 - za_{j_1}}{a_{j_1}(q + (1-q)\nu(j_1))} \quad \text{as } z \rightarrow 1/a_{j_1}.$$

(ii) *For some sufficiently small $\varepsilon > 0$ and any acute angle $\theta \in (0, \pi/2)$, there is an open set $\Theta \subset \mathbb{C}_1$ that contains a small open real segment with right end $1/a_{j_1}$, such that the restriction*

$$\frac{1}{q}I_{\mathbf{a}} : \Theta \rightarrow \{z \in \mathbb{C} : 0 < |z-1| < \varepsilon \text{ and } |\text{Arg}(z-1)| > \theta\}$$

is bijective.

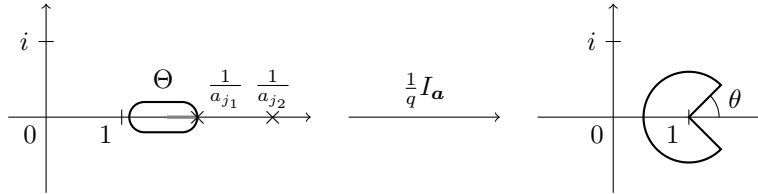


Figure 1: Representation of the domain Θ and its image via the function $\frac{1}{q}I_{\mathbf{a}}$.

Proof. We first make some observations which will be used for establishing each statement. Denote the open disk in \mathbb{C} centered at $1/a_{j_1}$ with radius $r_1 := 1/a_{j_2} - 1/a_{j_1}$ by $D(1/a_{j_1}, r_1)$, so that $\mathbb{C}_2 = \mathbb{C}_1 \cup D(1/a_{j_1}, r_1)$.

Using the notation (5.3), we express the function $\Pi_{\mathbf{a}}$ as the product

$$\Pi_{\mathbf{a}}(z) = (1 - za_{j_1})^{\beta-1} \times \Pi_{\mathbf{a}}^*(z),$$

where

$$\Pi_{\mathbf{a}}^*(z) := \prod_{j \neq j_1} (1 - za_j)^{\nu(j)(1-q)/q}$$

is a holomorphic function on \mathbb{C}_2 which has no zeros there. Thus $\Pi_{\mathbf{a}}^*$ is given in the neighborhood of $1/a_{j_1}$ by

$$\Pi_{\mathbf{a}}^*(z) = \sum_{n=0}^{\infty} b_n (1 - za_{j_1})^n, \quad z \in D(1/a_{j_1}, r_1),$$

for some sequence $(b_n)_{n \geq 0}$ of real numbers with

$$b_0 = \Pi_{\mathbf{a}}^*(1/a_{j_1}) > 0 \quad \text{and} \quad \limsup_{n \rightarrow \infty} |b_n|^{1/n} \leq 1/r_1.$$

On the other hand, we have from the criticality assumption (5.1) that

$$q - I_{\mathbf{a}}(z) = \int_{[z, 1/a_{j_1}]} \Pi_{\mathbf{a}}(z') dz', \quad z \in \mathbb{C}_1,$$

where we wrote $[z, 1/a_{j_1}]$ for the segment from z to $1/a_{j_1}$, which stays in \mathbb{C}_1 except for its right end.

(i) It follows from the observations above that for any $z \in D(1/a_{j_1}, r_1)$,

$$\begin{aligned} q - I_{\mathbf{a}}(z) &= \sum_{n=0}^{\infty} b_n \int_{[z, 1/a_{j_1}]} (1 - z'a_{j_1})^{n+(1-q)\nu(j_1)/q} dz' \\ &= h(z)(1 - za_{j_1})^{\beta}, \end{aligned} \tag{5.4}$$

where

$$h(z) := \frac{1}{a_{j_1}} \sum_{n=0}^{\infty} \frac{b_n}{1 + n + (1 - q)\nu(j_1)/q} (1 - za_{j_1})^n$$

is a holomorphic function on $D(1/a_{j_1}, r_1)$.

We can now write

$$\frac{q - I_{\mathbf{a}}(z)}{q\Pi_{\mathbf{a}}(z)} = (1 - za_{j_1}) \frac{h(z)}{q\Pi_{\mathbf{a}}^*(z)}, \quad z \in \mathbb{C}_1 \cap D(1/a_{j_1}, r_1),$$

which immediately entails the two claims in (i).

(ii) Recall (5.4). Since $h(1/a_{j_1})$ is a strictly positive real number, we can pick $r_2 > 0$ sufficiently small such that $g : z \mapsto (1 - za_{j_1})h(z)^{1/\beta}$ is a well-defined holomorphic function on the disk $D(a_{j_1}, r_2)$. Furthermore, since $g(1/a_{j_1}) = 0$ and $g'(1/a_{j_1}) \neq 0$, g is injective provided that r_2 has been chosen sufficiently small, and its image contains a small disk around the origin and radius, say $r_3 > 0$. Take an arbitrary angle $\alpha \in (0, \pi)$, and let

$$\Theta := \{g^{-1}(z') : 0 < |z'| < r_3 \text{ and } |\text{Arg}(z')| < \alpha/\beta\}.$$

Since the real function $q - I_{\mathbf{a}} : [0, 1/a_{j_1}] \rightarrow [0, q]$ is a decreasing bijection, Θ contains a real open segment with right end $1/a_{j_1}$. The claim (ii) is seen from the identity $q - I_{\mathbf{a}}(z) = g(z)^\beta$ for all $z \in \Theta$, taking $\theta = \pi - \alpha$. \square

We will also need the following.

Lemma 5.3. *There is a domain Θ' with*

$$(-\infty, 1/a_{j_1}) \subset \Theta' \subset \{z : \Re(z) < 1/a_{j_1}\},$$

such that the function

$$I_{\mathbf{a}} : \Theta' \rightarrow \{w \in \mathbb{C} : \Re(w) < q\}$$

is bijective.

Proof. We consider the differential equation

$$z'(t) = i/\Pi_{\mathbf{a}}(z(t)), \tag{5.5}$$

our main goal is to check that for any initial condition $z(0) = x$ in $(-\infty, 1/a_{j_1})$, (5.5) has a well-defined solution for all times $t \in \mathbb{R}$. In this direction, note first that the function $i/\Pi_{\mathbf{a}}$ is globally Lipschitz-continuous on any domain in \mathbb{C}_1 whose distance to the half-line $[1/a_{j_1}, \infty)$ is strictly positive. Thanks to the Cauchy-Lipschitz theorem, the existence of a global solution to (5.5) defined for all $t \in \mathbb{R}$ will be granted if we can check that for any local solution, $\Re(z'(t))$ has the opposite sign of t . Indeed, the function $t \mapsto \Re(z(t))$ is then non-decreasing on $(-\infty, 0]$ and non-increasing on $[0, \infty)$, and hence the distance of $z(t)$ to the half-line $[1/a_{j_1}, \infty)$ is never less than $1/a_{j_1} - x > 0$. For the sake of simplicity we focus on the case when $t > 0$ in the sequel, as the case when $t < 0$ follows from a similar argument.

In this direction, we consider the continuous determination of the argument of $\Pi_{\mathbf{a}}$ on \mathbb{C}_1 , which we denote by \arg_{Π} . Explicitly, writing $\arg(z) = \Im(\log(z)) \in (-\pi, \pi)$ for the argument of $z \in \mathbb{C} \setminus \mathbb{R}_-$, we have

$$\arg_{\Pi}(z) = \frac{1-q}{q} \sum_j \nu(a_j) \arg(1/a_j - z).$$

Note that \arg_{Π} vanishes on $(-\infty, 1/a_{j_1})$ and that $\arg_{\Pi}(z) < 0$ for all $z \in \mathbb{C}_1$ with $\Im(z) > 0$. By (5.5), the function $t \mapsto \pi/2 - \arg_{\Pi}(z(t))$ is the continuous determination of the argument of $z'(t)$, so in order to check that $\Re(z'(t)) < 0$, it suffices to prove that $\arg_{\Pi}(z(t))$ remains bounded from above by 0 and from below by $-\pi/2$ for any $t > 0$.

That $\arg_{\Pi}(z(t))$ is never 0 for any $t > 0$ is clear, since the zero set of \arg_{Π} is the real half-line $(-\infty, 1/a_{j_1})$. Assume that $\arg_{\Pi}(z(t)) = -\pi/2$ for some $t > 0$, and let $t_0 > 0$ be the smallest of such times. Then $\arg(z'(t_0)) = -\pi$, that is the derivative $z'(t_0)$ is a strictly negative real number. Since $\Re(z(t_0)) < 1/a_j$ for all $j \in \text{Supp}(\nu)$, it follows that $\arg(1/a_j - z(t)) > \arg(1/a_j - z(t_0))$ for all $t < t_0$ sufficiently close to t_0 , and *a fortiori* $\arg_{\Pi}(z(t)) > \arg_{\Pi}(z(t_0)) = -\pi/2$ for such t 's. This contradicts our assumption, since then there would exist $t < t_0$ with $\arg_{\Pi}(z(t)) = -\pi/2$. Thus $\arg(z'(t)) \in [\pi/2, \pi]$ for all $t \geq 0$, and we conclude

that (5.5) with arbitrary initial condition $z(0) = x \in (-\infty, 1/a_{j_1})$ has indeed a solution for all times $t \in \mathbb{R}$.

Our motivation for introducing (5.5) is that its integrated version can be expressed as

$$I_{\mathbf{a}}(z(t)) = I_{\mathbf{a}}(x) + it. \quad (5.6)$$

Since plainly, the function $I_{\mathbf{a}}$ is bijective from $(-\infty, 1/a_{j_1})$ to $(-\infty, q)$, we now see that the function from $(-\infty, 1/a_{j_1}) \times \mathbb{R}$ to \mathbb{C}_1 which maps a pair (x, t) to the value $z(t)$ of solution to (5.5) at time t for the initial condition x , is injective. Moreover, again by an argument of Lipschitz-continuity, this function is also continuous, and we denote its range by

$$\Theta' := \{z(t) : t \in \mathbb{R} \text{ and } z(0) = x \in (-\infty, 1/a_{j_1})\}.$$

The differential flow (5.5) thus induces a bijection between $(-\infty, 1/a_{j_1}) \times \mathbb{R}$ and Θ' . For any $w \in \mathbb{C}$ with $\Re(w) < q$, we can choose $x \in (-\infty, 1/a_{j_1})$ such that $I_{\mathbf{a}}(x) = \Re(w)$ and $t = \Im(w)$, and then $I_{\mathbf{a}}(z(t)) = w$. We conclude from (5.6) that the map $I_{\mathbf{a}} : \Theta' \rightarrow \{w \in \mathbb{C} : \Re(w) < q\}$ is bijective. \square

It is now an easy matter to establish Proposition 5.1 by combining Lemmas 5.2 and 5.3 with the expressions obtained in the preceding section.

Proof of Proposition 5.1. To start with, we use (4.9) to write

$$\varphi(t) = \frac{A''(t)}{A'(t)} = H'(A(t)) = H(A(t))F(A(t)) - 1, \quad 0 \leq t \leq A(\infty) = 1/(qa_{j_1}).$$

Applying Proposition 4.3 in the second line below yields for $x \in [0, 1)$

$$\begin{aligned} \varphi(-\log(1-x)) &= H(A(-\log(1-x)))F(A(-\log(1-x))) - 1 \\ &= H(I_{\mathbf{a}}^{-1}(qx)/q)F(I_{\mathbf{a}}^{-1}(qx)/q) - 1. \end{aligned}$$

On the one hand, recall that the left-hand side above defines a holomorphic function on the open unit disk \mathbb{D} . On the other hand, we know from Lemma 5.3 that $I_{\mathbf{a}}^{-1}$ can be extended analytically to the half-plane $\{w \in \mathbb{C} : \Re(w) < q\}$ and then takes values in $\{z \in \mathbb{C} : \Re(z) < 1/a_{j_1}\} \subset \mathbb{C}_1$. Moreover, we see from (4.3) and (4.10) that HF is holomorphic on $\{z \in \mathbb{C} : qa_{j_1}\Re(z) < 1\}$, and we conclude that the function $z \mapsto \varphi(-\log(1-z))$ can be extended analytically to $\{z \in \mathbb{C} : \Re(z) < 1\}$.

Similarly, we know from Lemma 5.2(ii) that, given an acute angle $\theta \in (0, \pi/2)$, any z with $|\text{Arg}(z-1)| > \theta$ and $|z-1|$ sufficiently small, can be given in the form $z = 1 - I_{\mathbf{a}}(z')/q$ for a unique $z' \in \Theta$. On the other hand, we infer from (4.3), (4.10) and Lemma 5.2(ii) that the function $z' \mapsto H(z'/q)F(z'/q)$ is holomorphic on Θ , and since Θ contains an open real segment to the right of $1/a_{j_1}$, we deduce by uniqueness of holomorphic extensions that there is the identity

$$\varphi(-\log(1 - I_{\mathbf{a}}(z')/q)) = H(z'/q)F(z'/q) - 1, \quad z' \in \Theta.$$

This establishes the first claim, since we can choose $R > 1$ sufficiently close to 1 such that the domain $\Delta(R, \theta)$ is contained into the union

$$\{z \in \mathbb{C} : \Re(z) < 1\} \cup \{z \in \mathbb{C} : 0 < |z-1| < \varepsilon \text{ and } |\text{Arg}(z-1)| > \theta\}.$$

Finally, we express the right-hand side above using (4.3) and (4.10) as

$$(1-q) \frac{q - I_{\mathbf{a}}(z')}{q \Pi_{\mathbf{a}}(z')} \sum_j \frac{a_j \nu(j)}{1 - z' a_j} - 1 = -\frac{q}{q + (1-q)\nu(j_1)} + O(|1 - z' a_{j_1}|),$$

where the identity stems from Lemma 5.2(i). Since, thanks to (5.4),

$$|1 - I_{\mathbf{a}}(z')/q| \asymp |z' a_{j_1} - 1|^\beta \quad \text{as } z' \rightarrow 1/a_{j_1} \text{ and } z' \in \Theta,$$

where the notation $f \asymp g$ means that both $f = O(g)$ and $g = O(f)$, the proof is complete. \square

6 Proof of Theorem 1.1

We now take

$$a_j := cj, \quad j \in \text{Supp}(\nu),$$

where $c > 0$ will be chosen later on. Plainly, (5.2) holds with $j_1 = k^*$ and $a_{j_1} = ck^*$. Furthermore, comparing (1.1) and (4.6), we see that

$$\Pi_{\mathbf{a}}(t) = \Pi(ct), \quad 0 \leq t \leq 1/a_{j_1},$$

and then (4.7) reads

$$I_{\mathbf{a}}(1/a_{j_1}) = \frac{1}{c} \int_0^{1/k^*} \Pi(t) dt.$$

We now choose, in the notation (1.2),

$$c = 1/m_{\nu, q} = \frac{1}{q} \int_0^{1/k^*} \Pi(t) dt,$$

so that (5.1) holds.

It is convenient to introduce here the notation $\Phi = (\Phi_\ell)$, where for $\ell \in \text{Supp}(\nu)$,

$$\Phi_\ell(z) := M_\ell(\mathbf{a}, -\log(1-z)). \quad (6.1)$$

We set further

$$\Phi_{\nu} := \langle \nu; \Phi \rangle,$$

so that, by (3.1),

$$\Phi_{\nu}(z) = \frac{\varphi(-\log(1-z)) + 1}{1-q}. \quad (6.2)$$

Next, recall from Lemma 2.1 that for any $x \in [0, 1)$,

$$\begin{aligned} \Phi_{\nu}(x) &= \mathbf{E} \left(\prod (cj)^{Y_j(-\log(1-x))} \right) \\ &= (1-x) \sum_{n=1}^{\infty} x^{n-1} c^n \mathbf{E}(Z(n)). \end{aligned}$$

We have seen that this power series has radius of convergence equal to 1 and know from Proposition 5.1 that it defines a holomorphic function on some domain $\Delta(R, \theta)$.

We next re-express the above in the form

$$\begin{aligned} & \frac{\Phi_{\nu}(z) - \nu(j_1)/(q + (1-q)\nu(j_1))}{1-z} \\ &= \sum_{n=1}^{\infty} z^{n-1} \left(m_{\nu,q}^{-n} \mathbb{E}(Z(n)) - \frac{\nu(j_1)}{q + (1-q)\nu(j_1)} \right). \end{aligned} \quad (6.3)$$

The assumptions of Proposition 5.1 have been checked, and combining its conclusion with (6.2) and the observation from (5.3) that

$$\frac{\nu(j_1)}{q + (1-q)\nu(j_1)} = \frac{1 - 1/\beta}{1 - q},$$

enables us to apply to this generating function a basic transfer theorem from singularity analysis; see [8, Theorem 1], or [9, Theorem VI.3]. We get

$$m_{\nu,q}^{-n} \mathbb{E}(Z(n)) - \frac{\nu(j_1)}{q + (1-q)\nu(j_1)} = O(n^{-1/\beta}).$$

Recalling that $j_1 = k^*$, this is the claim of Theorem 1.1.

7 Three further results

7.1 On the dependence on the parameters

In random population models, the ratio $\mathbb{E}(Z(n+1))/\mathbb{E}(Z(n))$ is sometimes called the effective reproduction number at generation n , and Theorem 1.1 thus identifies its limit when $n \rightarrow \infty$ for the reinforced Galton-Watson process as $m_{\nu,q}$ in (1.2). It is natural to study how this quantity depends on the memory parameter q and on the reproduction law ν .

Proposition 7.1. (i) *The function $q \mapsto m_{\nu,q}$ is monotone increasing on $(0, q)$, with*

$$\lim_{q \rightarrow 0+} m_{\nu,q} = \sum j\nu(j) \quad \text{and} \quad \lim_{q \rightarrow 1-} m_{\nu,q} = k^*.$$

(ii) *Let $q \in (0, 1)$ be a fixed memory parameter and $k^* \geq 1$ a fixed integer. The function $\nu \mapsto m_{\nu,q}$ is log-concave on the subset of reproduction laws ν with $\sup\{k \geq 0 : \nu(k) > 0\} = k^*$.*

Proof. (i) The monotonicity assertion derives from Lemma 2.2 and Theorem 1.1. Next, we write

$$1/m_{\nu,q} = \int_0^{1/(qk^*)} \Pi(qt) dt, \quad (7.1)$$

and note that for every $t \geq 0$,

$$\lim_{q \rightarrow 0+} \Pi(qt) = \exp\left(-t \sum j\nu(j)\right).$$

Integrating over $t \geq 0$ yields

$$\lim_{q \rightarrow 0+} 1/m_{\nu,q} = 1/\sum j\nu(j).$$

On the other hand, the function

$$q \mapsto (1 - tk)^{\nu(k)(1-q)/q}$$

increases on $(0, 1)$. Its limit as $q \rightarrow 1-$ equals 1 for all $t < 1/k^*$, which yields $m_{\nu, 1-} = k^*$.

(ii) Let ν_1 and ν_2 be two reproduction laws with $\sup\{k \geq 0 : \nu_i(k) > 0\} = k^*$ for $i = 1, 2$. Write

$$\Pi_i(t) := \prod (1 - tk)^{\nu_i(k)(1-q)/q}, \quad t \leq 1/k^* \text{ and } i = 1, 2.$$

Let $c \in (0, 1)$, consider $\nu = c\nu_1 + (1 - c)\nu_2$. As $\Pi = \Pi_1^c \times \Pi_2^{1-c}$ in the obvious notation, we deduce from (7.1) and the Hölder inequality that

$$m_{\nu, q}^{-1} \leq m_{\nu_1, q}^{-c} m_{\nu_2, q}^{c-1}.$$

This shows the log-concavity statement. \square

7.2 Explicit upper and lower bounds

The purpose of this section is to discuss the following simple bounds for the asymptotic effective reproduction number $m_{\nu, q}$.

Proposition 7.2. *We have*

$$k^*(q + (1 - q)\nu(k^*)) \leq m_{\nu, q} \leq k^*q + (1 - q)\bar{\nu},$$

where on the right-hand side,

$$\bar{\nu} := \sum j\nu(j)$$

denotes the mean of ν .

We shall now give two different proofs of this claim, the first is purely probabilist whereas the second is purely analytic and based on the formula (1.2).

A probabilistic proof. Consider first the sub-process of the reinforced Galton-Watson process which results by suppressing every sibling (together with its possible descent) which has size strictly less than k^* . This sub-process is then a true Galton-Watson process in which individuals have either k^* children with probability $q + (1 - q)\nu(k^*)$ or 0 child with complementary probability. Therefore $Z(n)$ dominates the size of the n -th generation of a Galton-Watson process with averaged reproduction $k^*(q + (1 - q)\nu(k^*))$, and the lower-bound in the statement follows.

For the upper-bound consider another Galton-Watson process with the following reproduction law. With probability q , the number of children of a typical individual is k^* , and with complementary probability $1 - q$, it is given by a random variable with law ν . In words, at each repetition event, we systematically increase the number of children to k^* . Obviously, this second true Galton-Watson process dominates the reinforced Galton-Watson process, and this entails the upper-bound. \square

An analytic proof. For the lower-bound, we simply write for every $0 \leq t \leq 1$ the inequality

$$\prod (1 - tj/k^*)^{(1-q)\nu(j)/q} \leq (1 - t)^{(1-q)\nu(k^*)/q}.$$

By integration, we conclude that

$$\int_0^1 \prod (1 - tj/k^*)^{(1-q)\nu(j)/q} dt \leq \frac{q}{q + (1 - q)\nu(k^*)},$$

and now the first inequality of the statement can be seen from (1.2).

For the upper-bound, we first write

$$(1 - tj/k^*)^{(1-q)\nu(j)/q} = \left((1 - tj/k^*)^{k^*/j} \right)^{\frac{(1-q)j\nu(j)}{qk^*}},$$

and then use the convexity inequality

$$(1 - at)^{1/a} \geq 1 - t \quad \text{for } a \leq 1.$$

We get

$$\prod (1 - tj/k^*)^{(1-q)\nu(j)/q} \geq (1 - t)^{\frac{(1-q)\bar{\nu}}{qk^*}}.$$

By integration, we get

$$\int_0^1 \prod (1 - tj/k^*)^{(1-q)\nu(j)/q} dt \geq \frac{q}{q + (1 - q)\bar{\nu}/k^*},$$

and the second inequality of the statement can again be seen from (1.2). \square

7.3 On the dependence on the first offspring number

We now conclude this work presenting a stronger version of Theorem 1.1. More precisely, the latter is given for the probability measure \mathbb{P} under which the number of children $Z(1)$ of the ancestor is a random sample of ν . Here is the version under the conditional probabilities $\mathbb{P}_\ell = \mathbb{P}(\cdot \mid Z(1) = \ell)$, which again underlines the key role of the maximal possible number of children k^* .

Theorem 7.3. *There exists some constant $\gamma \in (0, \infty)$ such that for every $\ell \in \text{Supp}(\nu)$ with $\ell \neq k^*$, we have*

$$\lim_{n \rightarrow \infty} n^{1/\beta} m_{\nu, q}^{-n} \mathbb{E}_\ell(Z(k)) = \frac{\gamma}{\Gamma(1/\beta)(1/\ell - 1/k^*)}.$$

We also have

$$\lim_{n \rightarrow \infty} m_{\nu, q}^{-n} \mathbb{E}_{k^*}(Z(k)) = 1/(q + \nu(k^*)(1 - q)).$$

Remark 7.4. One can compute the value of γ explicitly in terms of the parameters. However, since its expression is not quite simple and chasing the constants in the calculation below would have been a bit boring, we favored simplicity of the argument over a more precise statement.

Proof. The guiding line of the argument is similar to that for the proof of Theorem 1.1 and will not be repeated. We shall just indicate the main steps of the calculation. Recall that we focus on the critical case when assumptions (5.1) and (5.2) hold.

Our starting point is now the formula for the probability generating function that been obtained in Corollary 4.4. Using the notation (6.1), this reads

$$\Phi_\ell(z) = \frac{a_\ell A'(-\log(1-z))}{1 - qa_\ell A(-\log(1-z))}.$$

Essentially, we have to check that this defines an analytic function on $\Delta(R, \theta)$ and then to determine its asymptotic behavior as $z \rightarrow 1$ in $\Delta(R, \theta)$.

Let us first discuss analyticity; recall that $A'(t) = \exp\left(\int_0^t \varphi(s) ds\right)$. Thanks to Proposition 5.1, both the numerator and the denominator of $\Phi_\ell(z)$ are analytic functions on $\Delta(R, \theta)$, so all that is needed is to check that the denominator does not vanish there. Assume that there is some $z_\ell \in \Delta(R, \theta)$ such that $qa_\ell A(-\log(1-z_\ell)) = 1$. Obviously, then $qa_j A(-\log(1-z_\ell)) \neq 1$ for any j with $a_j \neq a_\ell$, and we deduce that z_ℓ would also be a singularity of the linear combination Φ_ν . This contradicts Proposition 5.1(ii) and we conclude that Φ_ℓ is indeed analytic on $\Delta(R, \theta)$ for all ℓ .

Next, we know from Proposition 5.1 that

$$\varphi(w) = -1/\beta + O(e^{-w/\beta}) \quad \text{as } \Re(w) \rightarrow \infty \text{ with } 1 - e^{-w} \in \Delta(R, \theta).$$

We deduce by integration that in the same regime

$$A'(w) = \exp\left(\int_{[0,w]} \varphi(s) ds\right) \sim \gamma e^{-w/\beta} (1 + O(e^{-w/\beta})),$$

where

$$\gamma := \exp\left(\int_0^\infty (\varphi(t) + 1/\beta) dt\right) \in (0, \infty).$$

We now see that for any $\ell \neq j_1$, we have

$$\Phi_\ell(z) \sim \frac{\gamma(1-z)^{1/\beta}}{1/a_\ell - 1/a_{j_1}} \quad \text{as } z \rightarrow 1, z \in \Delta(R, \theta).$$

The rest of the proof of the statement for $\ell \neq k^*$ uses the same path as in Section 6, except that for the transfer, we use [9, Corollary VI.1] in place of [9, Theorem VI.3]. The second claim then follows by linearity from the first and Theorem 1.1. \square

Acknowledgment: We express our sincere gratitude to the reviewers for their insightful comments, which greatly contributed to enhancing the quality of this manuscript. J.B. would also like to thank Klaus Widmayer for discussions about transport equations, and notably Lemma 3.3.

References

- [1] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].
- [2] Marco Bertenghi and Alejandro Rosales-Ortiz. Joint invariance principles for random walks with positively and negatively reinforced steps. *J. Stat. Phys.*, 189(3):Paper No. 35, 31, 2022.
- [3] Jean Bertoin. Scaling exponents of step-reinforced random walks. *Probab. Theory Related Fields*, 179(1-2):295–315, 2021.
- [4] Jean Bertoin. Universality of noise reinforced Brownian motions. In *In and out of equilibrium 3. Celebrating Vladas Sidoravicius*, volume 77 of *Progr. Probab.*, pages 147–161. Birkhäuser/Springer, Cham, [2021] ©2021.
- [5] Jean Bertoin and Bastien Mallein. Reinforced Galton-Watson processes II: Large time behaviors. arXiv:2310.19030, 2023.
- [6] Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Discrete Math. Theor. Comput. Sci. Proc., AG, pages 59–118. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2006.
- [7] Philippe Flajolet, Joaquim Gabarró, and Helmut Pekari. Analytic urns. *Ann. Probab.*, 33(3):1200–1233, 2005.
- [8] Philippe Flajolet and Andrew Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.
- [9] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
- [10] Kiyoshi Inoue. Exactly solvable urn models under random replacement schemes and their applications. *Journal of Applied Probability*, pages 1–20, 2023.
- [11] Andreas E. Kyprianou. *Fluctuations of Lévy processes with applications. Introductory lectures*. Universitext. Berlin: Springer, 2nd ed. edition, 2014.
- [12] Lucile Laulin. *About the elephant random walk*. Ph.D. Thesis. Université de Bordeaux, 2022.
- [13] Robin Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79, 2007.
- [14] Christophe Sabot and Xiaolin Zeng. A random Schrödinger operator associated with the vertex reinforced jump process on infinite graphs. *J. Amer. Math. Soc.*, 32(2):311–349, 2019.
- [15] Herbert A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3/4):425–440, 1955.