Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 123 (2013) 2940–2956

stochastic processes and their applications

www.elsevier.com/locate/spa

# Linear-fractional branching processes with countably many types

Serik Sagitov*

*Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Sweden*

Received 20 November 2011; received in revised form 13 March 2013; accepted 14 March 2013
Available online 22 March 2013

**Abstract**

We study multi-type Bienaymé–Galton–Watson processes with linear-fractional reproduction laws using various analytical tools like the contour process, spinal representation, Perron–Frobenius theorem for countable matrices, and renewal theory. For this special class of branching processes with countably many types we present a transparent criterion for $R$-positive recurrence with respect to the type space. This criterion appeals to the Malthusian parameter and the mean age at childbearing of the associated linear-fractional Crump-Mode-Jagers process.


*Keywords:* Multivariate linear-fractional distribution; Contour process; Spinal representation; Bienaymé–Galton–Watson process; Crump-Mode-Jagers process; Malthusian parameter; Perron–Frobenius theorem; $R$-positive recurrence; Renewal theory

## 1. Introduction

Branching processes is a steadily growing body of mathematical research having applications in various areas, primarily in theoretical population biology [9,16,19]. A basic version of branching processes, called the Bienaymé–Galton–Watson (BGW) process, describes populations of particles which live one unit of time and at the moment of death give birth to a random number of new particles independently of each other. In the single type setting the consecutive population sizes $\{Z^{(n)}\}_{n\geq 0}$ form a Markov chain with the state space $\{0, 1, 2, \ldots\}$. An important analytical tool for studying branching processes is the probability generating

* Tel.: +46 317725351.
*E-mail address:* serik@chalmers.se.

0304-4149/$ - see front matter 
http://dx.doi.org/10.1016/j.spa.2013.03.008

functions. Given $\phi(s) = \mathbb{E}(s^{Z^{(1)}})$, the $n$-th generation's size is characterized by the $n$-fold iteration of $\phi(\cdot)$

$$\phi^{(n)}(s) = \mathbb{E}(s^{Z^{(n)}}) = \phi(\ldots(\phi(s))\ldots).$$

Here and elsewhere in this paper *we always assume that a branching process starts from a single particle.*

The case of linear-fractional generating functions

$$\phi(s) = h_0 + \frac{h_1 s}{1 + m - ms} \tag{1}$$

with $h_0 \in [0, 1], h_1 = 1 - h_0$ and $m > 0$ is of special interest as their iterations are again linear-fractional functions allowing for explicit calculations of various entities of importance (see [3], p. 7). Such explicit results, although being specific, illuminate the known asymptotic results concerning more general branching processes, and, on the other hand, may bring insight into less investigated aspects of the theory of branching processes.

In the multi-type setting particles still reproduce independently but now the number of offsprings may depend on the mother's type. A flexible family of population models is obtained by means of BGW-processes with countably many types [4,10,16,23]. These are infinitely dimensional Markov chains

$$\mathbf{Z}^{(n)} = (Z_1^{(n)}, Z_2^{(n)}, \ldots), \quad n = 0, 1, 2, \ldots,$$

whose $i$-th component $Z_i^{(n)}$ gives the number of particles of type $i$ existing at time $n$. In this paper we study the class of such branching processes with the generating functions for vectors $\mathbf{Z}^{(n)}$ all being linear-fractional. As shown in Section 2 a linear-fractional BGW-process with countably many types is fully specified by a triplet of parameters $(\mathbf{H}, \mathbf{g}, m)$, where $\mathbf{H} = (h_{ij})_{i,j=1}^{\infty}$ is a sub-stochastic matrix, $\mathbf{g} = (g_1, g_2, \ldots)$ is a proper probability distribution, and $m$ is a positive constant.

For a given triplet $(\mathbf{H}, \mathbf{g}, m)$ the particles in the linear-fractional BGW-process have the following reproduction law. A particle of type $i$ has no offspring with probability $h_{i0} = 1 - \sum_{j \geq 1} h_{ij}$. Given that this particle has at least one offspring, the type of its first daughter is $j$ with probability $h_{ij}/(1 - h_{i0})$, and the number of subsequent daughters has a geometric distribution with mean $m$. With the exception of the first daughter the types of all other offspring particles follow the same distribution $\mathbf{g}$ independently of each other and *independently of mother's type.*

The countable matrix of the mean offspring numbers

$$\mathbf{M} = (m_{ij})_{i,j=1}^{\infty}, \quad m_{ij} = \mathbb{E}(Z_j^{(1)} | \mathbf{Z}^{(0)} = \mathbf{e}_i),$$

where $\mathbf{e}_i = (1_{\{i=1\}}, 1_{\{i=2\}}, \ldots)$, in the linear-fractional case is found as

$$\mathbf{M} = \mathbf{H} + m\mathbf{H}\mathbf{1}^{\mathrm{t}}\mathbf{g}, \tag{2}$$

where $\mathbf{1}^{\mathrm{t}}$ is the transpose of the row-vector $\mathbf{1}^{\mathrm{t}} = (1, 1, \ldots)$. Theorem 3 in Section 2 states that every vector $\mathbf{Z}^{(n)}$ has a multivariate linear-fractional distribution. The generating function of $\mathbf{Z}^{(n)}$ is explicitly expressed in terms of $(\mathbf{M}, \mathbf{g}, m)$. An important step in obtaining this formula, see Section 7.1, relies on a spinal representation argument making the derivation different from that used in [14] for the finite-dimensional case.

Assume that the type of the initial particle has distribution $\mathbf{g}$. Then as shown in Section 3, the total population sizes $Z^{(n)} = \mathbf{Z}^{(n)}\mathbf{1}^{\mathrm{t}}$ for the linear-fractional BGW-process form a single-type discrete time Crump-Mode-Jagers (CMJ) process [13], which we call a *linear-fractional CMJ-process*. This CMJ-model is not restrictive about the life length distribution, however, the point process of birth events must follow a very specific pattern: at each age a living individual produces independent and identically distributed (iid) geometric numbers of daughters. Making birth events in the linear-fractional CMJ-process very rare (by choosing the row sums of $\mathbf{H}$ to be close to zero) and rescaling time accordingly we arrive at a continuous time CMJ-process studied in [17]. In [17] special attention is paid to the properties of the so-called contour processes of the corresponding planar genealogical trees. The discrete time counterpart of these contour processes is the subject of Section 4.

Section 5 introduces a double classification of the BGW-processes with countably many types based on the Perron–Frobenius theorem for countable matrices. Besides the usual classification into subcritical, critical, and supercritical branching processes in the case of infinitely many types one has to distinguish among $R$-transient, $R$-null recurrent, and $R$-positively recurrent cases depending on the corresponding property of the mean matrix $\mathbf{M}$. The main result of this paper, Theorem 8, among other statements contains a transparent criterion for $R$-positive recurrence of a linear-fractional BGW-process. In terms of the associated CMJ-process this criterion requires that the corresponding Malthusian parameter is well defined and the mean age at childbearing is finite. Theorem 8 is proven in Section 7.3 using a renewal theory approach.

In Section 6 we present basic asymptotic results for subcritical, critical, and supercritical linear-fractional BGW-processes with countably many types in the positively recurrent case. All our results for the linear-fractional BGW-processes with parameters $(\mathbf{H}, \mathbf{g}, m)$ also apply to the case of finitely many, say $a$, types after putting $Z_i^{(0)} = 0$, $g_i = 0$ for $i \geq a+1$, and $h_{ij} = 0$ for $1 \leq i \leq a < j$. The transient and null recurrent cases will be addressed in a separate paper.

## 2. Linear-fractional distributions

We are using the following vector notation: $\mathbf{x} = (x_1, x_2, \ldots)$, $\mathbf{s}^{\mathbf{x}} = s_1^{x_1} s_2^{x_2} \ldots$, $\mathbf{0} = (0, 0, \ldots)$. Let $\mathbf{x}^{\mathrm{t}}$ stand for the transpose of the vector $\mathbf{x}$, and $\mathbf{I}$ denote the identity matrix $(1_{\{i=j\}})_{i\geq 1, j\geq 1}$. We denote by $\mathbb{Z}_+^\infty$ the set of vectors $\mathbf{k}$ with non-negative integer-valued components and finite $k = \mathbf{k}\mathbf{1}^{\mathrm{t}}$.

**Definition 1.** Let $(h_0, h_1, h_2, \ldots)$ be a probability distribution on $\{0, 1, 2, \ldots\}$, $(g_1, g_2, \ldots)$ be a probability distribution on $\{1, 2, \ldots\}$, and $m$ be a positive constant. Put $\mathbf{h} = (h_1, h_2, \ldots)$, $\mathbf{g} = (g_1, g_2, \ldots)$. We say that a random vector $\mathbf{Z}$ has a linear-fractional distribution LF $(\mathbf{h}, \mathbf{g}, m)$ if

$$\mathbb{P}(\mathbf{Z} = \mathbf{0}) = h_0, \qquad \mathbb{P}(\mathbf{Z} = \mathbf{k} + \mathbf{e}_i) = \frac{h_i m^k}{(1+m)^{k+1}} \binom{k}{k_1, k_2, \ldots} \mathbf{g}^{\mathbf{k}}$$

for all $\mathbf{k} \in \mathbb{Z}_+^\infty$, where $k = \mathbf{k}\mathbf{1}^{\mathrm{t}}$.

The name of the distribution is explained by the linear-fractional form of its multivariate generating function

$$\mathbb{E}(\mathbf{s}^{\mathbf{Z}}) = h_0 + \frac{\sum_{i=1}^{\infty} h_i s_i}{1 + m - m \sum_{i=1}^{\infty} g_i s_i}$$

which is an extension of its one-dimensional version (1). Slightly modifying Theorem 1 from [14] (devoted to the finite-dimensional case) one can demonstrate that Definition 1 covers all possible linear-fractional probability generating functions. A linear-fractional distribution is a geometric distribution modified at zero. Indeed, if $\mathbf{Z}$ has distribution LF$(\mathbf{h}, \mathbf{g}, m)$, then it can be represented as

$$\mathbf{Z} = \mathbf{X} + (\mathbf{Y}_1 + \cdots + \mathbf{Y}_N) \cdot 1_{\{\mathbf{X} \neq \mathbf{0}\}}$$

in terms of mutually independent random entities $(\mathbf{X}, N, \mathbf{Y}_1, \mathbf{Y}_2, \ldots)$. Here vectors $\mathbf{X}$ and $\mathbf{Y}_j$ have multivariate Bernoulli distributions

$$\mathbb{P}(\mathbf{X} = \mathbf{0}) = h_0, \qquad \mathbb{P}(\mathbf{X} = \mathbf{e}_i) = h_i, \qquad \mathbb{P}(\mathbf{Y}_j = \mathbf{e}_i) = g_i, \quad i \geq 1, \; j \geq 1,$$

and $N$ is a geometric random variable with distribution

$$\mathbb{P}(N = k) = m^k (1 + m)^{-k-1}, \quad k \geq 0.$$

Observe that $\mathbf{Z}$ conditionally on $\mathbf{Z} \neq \mathbf{0}$ has a multivariate shifted geometric distribution

$$\mathbb{E}\big(\mathbf{s}^{\mathbf{Z}} | \mathbf{Z} \neq \mathbf{0}\big) = \frac{(1 - h_0)^{-1} \sum_{j=1}^{\infty} h_j s_j}{1 + m - m \sum_{j=1}^{\infty} g_j s_j}.$$

**Definition 2.** Let $\mathbf{H} = (h_{ij})_{i,j=1}^{\infty}$ be a sub-stochastic matrix with rows $\mathbf{h}_i = (h_{i1}, h_{i2}, \ldots)$ having non-negative elements such that $h_{i0} := 1 - h_{i1} - h_{i2} - \cdots$ take values in $[0, 1]$. Let $\mathbf{g} = (g_1, g_2, \ldots)$ be a probability distribution on $\{1, 2, \ldots\}$, and $m$ be a positive constant. A multi-type BGW-process will be called linear-fractional with parameters $(\mathbf{H}, \mathbf{g}, m)$, if for all $i = 1, 2, \ldots$ particles of type $i$ reproduce according to the LF$(\mathbf{h}_i, \mathbf{g}, m)$ distribution.

Notice the strong limitation on the reproduction law requiring parameters $(\mathbf{g}, m)$ to be ignorant of mother's type. This is needed for the generating functions

$$\phi_i^{(n)}(\mathbf{s}) = \mathbb{E}(\mathbf{s}^{\mathbf{Z}^{(n)}} | \mathbf{Z}^{(0)} = \mathbf{e}_i), \quad i = 1, 2, \ldots$$

to be also linear-fractional. It is easy to see that if the denominators in

$$\phi_i(\mathbf{s}) \equiv \phi_i^{(1)}(\mathbf{s}) = h_{i0} + \frac{\sum_{j=1}^{\infty} h_{ij} s_j}{1 + m - m \sum_{j=1}^{\infty} g_j s_j}$$

were different for different $i$, then the iterations of these generating functions $\phi_i(\phi_1(\mathbf{s}), \phi_2(\mathbf{s}), \ldots)$ would lose the linear-fractional property. The following key result is proven in Sections 4.1 and 7.1.

**Theorem 3.** *Consider a linear-fractional BGW-process with parameters* $(\mathbf{H}, \mathbf{g}, m)$ *starting from a type i particle. Its n-th generation size vector* $\mathbf{Z}^{(n)}$ *has a linear-fractional distribution LF* $(\mathbf{h}_i^{(n)}, \mathbf{g}^{(n)}, m^{(n)})$ *whose parameters satisfy*

$$m^{(n)} = m \sum_{k=0}^{n-1} \mathbf{g}\mathbf{M}^k \mathbf{1}^{\mathrm{t}}, \tag{3}$$

$$m^{(n)}\mathbf{g}^{(n)} = m\mathbf{g}(\mathbf{I} + \mathbf{M} + \cdots + \mathbf{M}^{n-1}), \tag{4}$$

$$\mathbf{H}^{(n)} = \mathbf{M}^n - \frac{m^{(n)}}{1+m^{(n)}}\mathbf{M}^n\mathbf{1}^{\mathrm{t}}\mathbf{g}^{(n)}, \tag{5}$$

*where* $\mathbf{H}^{(n)}$ *is the matrix with the rows* $(\mathbf{h}_i^{(n)})_{i=1}^{\infty}$.

Multiplying (5) by $\mathbf{1}^{\mathrm{t}}$ we obtain

$$\mathbb{P}(\mathbf{Z}^{(n)} \neq \mathbf{0}) = (1 + m^{(n)})^{-1}\mathbf{M}^n\mathbf{1}^{\mathrm{t}}, \tag{6}$$

where $\mathbb{P}(\mathbf{Z}^{(n)} \neq \mathbf{0})$ is a column vector with elements $\mathbb{P}(\mathbf{Z}^{(n)} \neq \mathbf{0}|\mathbf{Z}^{(0)} = \mathbf{e}_i)$. Furthermore, Theorem 3 entails that conditionally on non-extinction $\mathbf{Z}^{(n)}$ has a multivariate shifted geometric distribution

$$\mathbb{E}[\mathbf{s}^{\mathbf{Z}^{(n)}}|\mathbf{Z}^{(n)} \neq \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i] = \frac{(1-h_{i0}^{(n)})^{-1}\sum_{j=1}^{\infty} h_{ij}^{(n)}s_j}{1 + m^{(n)} - m^{(n)}\sum_{j=1}^{\infty} g_j^{(n)}s_j}. \tag{7}$$

## 3. The linear fractional CMJ-process

A discrete time single type CMJ-process $\{Z^{(n)}\}_{n\geq 0}$ describes stochastic changes in the population size for a reproduction model with overlapping generations [13]. Compared to BGW-populations consisting of *particles* it is more appropriate to speak of *individuals* building a CMJ-population. Individuals are assumed to live and reproduce independently according to a common life law specifying a reproduction point process $(N_1, \ldots, N_L)$, where $L$ is the life length of an individual and $N_i$ is the number of its daughters produced at age $i$. In this section we introduce a special class of such processes called linear-fractional CMJ-processes and discuss its close connection to the class of linear-fractional multi-type BGW-processes.

**Definition 4.** A discrete time single type CMJ-process is called linear-fractional, if its individual life law satisfies the following property: $N_L = 0$ and given $L = k$ the random variables $N_1, \ldots, N_{k-1}$ are independent and have a common geometric distribution.

Given $\mathbb{P}(L > n) = d_n$ and $m = \mathbb{E}(N_1|L > 1)$ a linear-fractional CMJ-process is fully characterized by a pair $(\mathbf{d}, m)$: the vector $\mathbf{d} = (d_1, d_2, \ldots)$ with the non-negative components satisfying $1 \geq d_1 \geq d_2 \geq \ldots$ and the positive constant $m$. In particular, the total offspring number $N_1 + \cdots + N_{L-1}$ has mean $\mu = m(\lambda - 1)$, where $\lambda = \mathbb{E}(L)$.

**Definition 5.** Consider a linear-fractional CMJ-process with parameters $(\mathbf{d}, m)$. Let $f(s) = \sum_{n\geq 1} d_n s^n$ and put

$$R_f = \inf\{s > 0 : f(s) = \infty\}.$$

If $f(R_f) \geq 1/m$, we define the Malthusian parameter $\alpha$ of the CMJ-process as the unique real solution of the equation $mf(e^{-\alpha}) = 1$. If $f(R_f) < 1/m$, we put $\alpha = -\infty$.

From $\lambda = 1 + f(1)$ we find $\mu = mf(1)$ and it easy to see that conditions $\mu < 1, \mu = 1, \mu > 1$ are equivalent to $\alpha < 0, \alpha = 0, \alpha > 0$. In the framework of CMJ-processes [11] a branching process is called subcritical if the Malthusian parameter is negative $\alpha < 0$, critical, if $\alpha = 0$, or supercritical, if $\alpha > 0$. We point out that given $\alpha > -\infty$

$$mf(se^{-\alpha}) = \sum_{n=1}^{\infty} \hat{d}_n s^n, \quad \hat{d}_n = md_n e^{-\alpha n},$$

is the generating function for the so-called regeneration age of the immortal individual [13]. (Notice that in the critical case we have $\hat{d}_n = \frac{\mathbb{P}(L>n)}{\mathbb{E}(L)-1}$.) The corresponding mean value

$$\beta = m \sum_{n=1}^{\infty} nd_n e^{-\alpha n}$$

is usually called the mean age at childbearing [11]. For $\alpha = -\infty$ we put $\beta = \infty$.

**Example 1.** Let $d_n = c_n n^{-k} e^{-\gamma n}$ for some constants $\gamma \geq 0$, $k \geq 0$, and assume $0 < \liminf_n c_n < \limsup_n c_n < \infty$. Clearly, in this case $R_f = e^{\gamma}$ and putting $A = \sum_{n=1}^{\infty} c_n n^{-k}$ we get

- if $A > 1/m$, then $-\gamma < \alpha < \infty$ and $\beta < \infty$,
- if $A = 1/m$, then $\alpha = -\gamma$ and $\beta < \infty$ iff $k > 2$,
- if $A < 1/m$, then $\alpha = -\infty$ and $\beta = \infty$.

It turns out that for a given triplet $(\mathbf{H}, \mathbf{g}, m)$ the corresponding linear-fractional BGW-process $\mathbf{Z}^{(n)}$ can be associated with a linear-fractional CMJ-process $Z^{(n)} = \mathbf{Z}^{(n)}\mathbf{1}^{\mathrm{t}}$ characterized by a pair $(\mathbf{d}, m)$, where the vector $\mathbf{d}$ has components

$$d_n = \mathbf{g}\mathbf{H}^n\mathbf{1}^{\mathrm{t}}, \quad n \geq 1. \tag{8}$$

To see this let the initial particle of the BGW-process have distribution $\mathbf{g}$. If the initial particle dies without producing any offspring we say that the initial individual in the associated CMJ-process had the life length $L = 1$ and, if the initial particle produced at least one offspring we say $L > 1$. Obviously, $\mathbb{P}(L > 1) = \mathbf{g}\mathbf{H}\mathbf{1}^{\mathrm{t}}$. The key idea of defining the associated CMJ-process is to view an individual as a sequence of first-born descendants of a particle which itself is either the progenitor particle or a particle which is not first-born. Following this idea given $L > 1$, we say $L > 2$ if the first-born daughter of the progenitor produces at least one offspring particle. This explains (8) for $n = 2$. Continuing in the same manner we see that (8) indeed gives us the distribution of the life length of the progenitor individual. The fact that all daughter individuals behave in the way prescribed by the linear-fractional CMJ-model is a straightforward consequence of the particle properties of the linear-fractional BGW-process.

Turn for visual help to Fig. 1(C) which gives the individual based picture of the same genealogical tree as in Fig. 1(A). Each vertical arrowed branch in Fig. 1(C) represents an individual dying before the observation time $n = 5$. In particular, the initial individual lives two units of time producing two daughters: one of them lives two units of time and the other only one. We see also that the first granddaughter of the initial individual produces two daughters at different ages.

The associated CMJ-process $Z^{(n)}$ tracks only the total number of BGW-particles at time $n$ ignoring the information on the types of the particles. To recover this information we may introduce additional labeling of individuals using the types of underlying BGW-particles. The evolution of a labeled individual over the type space can be modeled by a Markov chain whose state space $\{0, 1, 2, \ldots\}$ is the type space $\{1, 2, \ldots\}$ of the BGW-process augmented with a graveyard state $\{0\}$. The transition probabilities of such a chain are given by a stochastic matrix

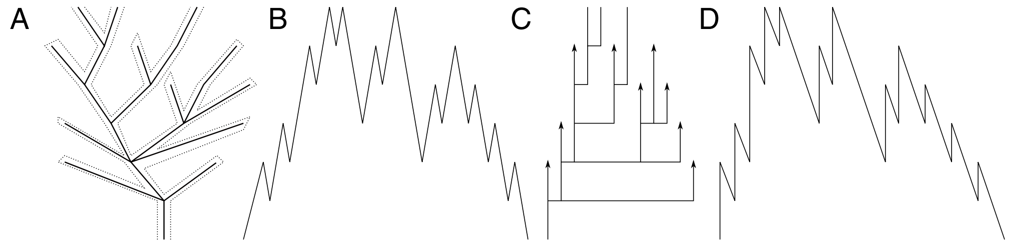


Fig. 1. (A) A BGW-tree (thick line) stopped at level $n = 5$ and supplied with a contour (dashed) line. (B) The corresponding unfolded contour process. (C) The CMJ-view of the same tree when depicted in terms of individuals forming a branching process with overlapping generations. The vertices marked by arrows represent individuals which are dead at that time. The stopped tree gives no information about the fate of the three tip vertices. (D) The modified contour process of a constant speed descent with iid upward jumps.

$\mathring{\mathbf{H}} = (\mathring{h}_{ij})_{i,j\geq 0}$ with $\mathring{h}_{ij} = h_{ij}$ for $i \geq 1$, $j \geq 0$, $\mathring{h}_{0j} = 0$ for $j \geq 1$, and $\mathring{h}_{00} = 1$. In terms of this Markov chain the life length $L$ is the time until absorption at $\{0\}$ starting from a state $j \in \{1, 2, \ldots\}$ with probability $g_j$. (Distributions describing absorption times in Markov chains are called *phase-type distributions*; see for example [1].) A labeled individual is able to visit all elements of the type space except phantom types defined next.

**Definition 6.** Consider a linear-fractional multi-type BGW-process with parameters $(\mathbf{H}, \mathbf{g}, m)$. If the $j$-th element of the vector $\mathbf{g}\mathbf{H}^n$ is zero for all $n \geq 0$, we call $j$ a phantom type of this BGW-process.

Among many pairs $(\mathbf{H}, \mathbf{g})$ resulting in the same life length distribution vector $\mathbf{d}$ it is worth emphasizing the next one

$$\mathbf{H} = \begin{pmatrix} 0 & d_1 & 0 & 0 & 0 & \ldots \\ 0 & 0 & d_2/d_1 & 0 & 0 & \ldots \\ 0 & 0 & 0 & d_3/d_2 & 0 & \ldots \\ 0 & 0 & 0 & 0 & d_4/d_3 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix}, \qquad \mathbf{g} = \mathbf{e}_1.$$

For this particular choice of $(\mathbf{H}, \mathbf{g})$ it easy to verify that (8) holds. In this case the particle type can be viewed as the age of the corresponding individual. Obviously, there are no phantom types if $d_n > 0$ for all $n \geq 1$. If $d_a > 0$ and $d_{a+1} = 0$ for some natural $a$, then to avoid the phantom types we must restrict the state space to $\{1, \ldots, a\}$. Observe that with $\mathbf{g} = \mathbf{e}_2$ type $j = 1$ becomes a phantom type.

## 4. Jumping contour processes in discrete time

In this section we remind the concept of a contour process generated by a planar BGW-tree. Then we show that in the multivariate linear-fractional framework a jumping version of the contour process (in the spirit of [17]) has a nice Markovian structure of a constant speed descent with independent and identically distributed upward jumps. In the end of this section we apply the contour process method to give a short proof of the first statement in Theorem 3.

Both the contour process and the spinal representation methods (the latter discussed in Section 7.1) rely on a planar genealogical tree connecting the particles of the branching processes appeared up to the time of observation. For the current setting of linear-fractional BGW-processes it is important to use a particular planar version of the genealogical tree: given a group of siblings stemming from the same particle, the *leftmost branch* should connect the mother to its *first daughter* (that one whose type may depend on mother's type).

For a given planar tree, its contour profile is defined by the depth-first search procedure. Fig. 1 illustrates the basic definition of the contour process for a finite tree supplied with a path around the tree. The contour process is simply the seesaw line graph (panel B) representing the height of the location of a virtual car driving with a constant speed along the path outlined on panel A. Notice that the $x$-axis in panel A is introduced just to distinguish among different branches on the same level; hence the speed of the car is meant along the $y$-axis. The resulting contour B of the tree A is an excursion of a random walk starting and ending at level $-1$. Even if the realization of the genealogical tree is infinite, one can still work with the contour processes after cutting off the branches above level $n$ corresponding to the observation time, as shown in Fig. 1.

It is easy to reconstruct the tree on panel A from the contour process on the panel B. As we said, the whole tree A is represented by the excursion B starting and ending at the bottom level $-1$. Raising the bottom level from $-1$ to 0 will split the tree A into 3 subtrees stemming from 3 daughters of the progenitor particle. At the same time the excursion B becomes split into 3 sub-excursions starting and ending at level 0. Proceeding in this way by moving up the bottom level and observing how the excursions are decomposed in sub-excursions allows us to fully reconstruct the branching history of the original genealogical tree.

The contour process approach has proven to be very useful in the theory of branching processes (see for example [8] and references therein). In the single type linear-fractional case (1) the contour process has a simple structure of an alternating random walk. Alternating upward and downward stretches have independent lengths following shifted geometric laws having mean $h_0^{-1}$ for the upward stretches and mean $\frac{1+m}{m}$ for the downward stretches.

In the multi-type linear-fractional setting, one can ensure a Markov property of the contour process by introducing additional labeling of the vertices in the contour path. Each vertex will be labeled by a pair of integers $(l, i)$ with $l \geq -1$ and $i \geq 0$. The current state $(l, i)$ with $l \geq 0, i \geq 1$ tells three things about the contour process: the current level is $l$, the last move was upward, and the underlying BGW-particle is of type $i$. If the contour process is at the vertex labeled $(l, 0)$, then again its current level is $l$ but now we know that this level was attained after a downward step. Such a labeled contour process (if we ignore the first compulsory move from level $-1$ to 0) can be viewed as a Markov chain with the following transition probabilities

$$\mathbb{P}\{(l,i) \to (l+1,j)\} = h_{ij}, \qquad \mathbb{P}\{(l,i) \to (l-1,0)\} = h_{i0},$$
$$\mathbb{P}\{(l,0) \to (l+1,j)\} = \frac{m}{1+m} g_j, \qquad \mathbb{P}\{(l,0) \to (l-1,0)\} = \frac{1}{1+m},$$
$$\mathbb{P}\{(-1,0) \to (-1,0)\} = 1,$$

for all $i \geq 1,\ j \geq 1,\ l \geq 0$.

The following alternative way of introducing Markovian structure in the contour process of a linear-fractional multi-type branching process does not require additional labeling. What we call here the *jumping contour process* (cf. [17]) has a trajectory of a constant speed descent with independent upward jumps each distributed as the individual life length $L$. The process starts from level $-1$ with an instantaneous jump and proceeds as follows. From any given current level $l$ the jumping contour process moves one level down to $l-1$ and either settles there with probability $\frac{1}{1+m}$ or, with probability $\frac{m}{1+m}$, it instantaneously jumps say $k$ levels up coming to the level $k+l-1$. Fig. 1(D) clearly illustrates the last construction.

Notice that supercriticality of the underlying branching process can be identified via a positive drift for the contour process. The drift $\lambda - 1 - m^{-1}$ of the jumping contour process is computed as the difference between the mean jump size $\lambda = \mathbb{E}(L)$ and the average length of a downward stretch $1 + m^{-1}$. Clearly, the inequality $\lambda - 1 - m^{-1} > 0$ is equivalent to $\mu > 1$.

### *4.1. Proof of Theorem 3, part 1*

Consider a linear-fractional BGW-process with parameters $(\mathbf{H}, \mathbf{g}, m)$. Next we show that if the BGW-process stems from a particle of type $i$, then its vector of $n$-th generation sizes $\mathbf{Z}^{(n)}$ has a linear-fractional joint distribution with some parameters $(\mathbf{h}_i^{(n)}, \mathbf{g}^{(n)}, m^{(n)})$ so that $(\mathbf{g}^{(n)}, m^{(n)})$ are independent of $i$. Consider the genealogical tree of the linear-fractional BGW-process stopped at level $n$ and denote by $j$ the type of the leftmost tip in the tree if any. The random vector $\mathbf{Z}^{(n)}$ counts the tips of various types and $Z^{(n)} = \mathbf{Z}^{(n)}\mathbf{1}^{\mathrm{t}}$ gives the total number of the tree tips. We have to verify that conditioned on $Z^{(n)} \geq 1$ the following two properties hold:

1. the number $Z^{(n)} - 1$ of the tree tips to the right of the leftmost tip has a geometric distribution which is independent from the types $(i, j)$,
2. the types of these $Z^{(n)} - 1$ tips are iid and independent from $(i, j)$.

These properties are simple consequences of the following Markov features of the contour process described above:

- the number of particles alive at time $n$, if any, is 1 plus the number of excursions of the contour process starting at level $n$ downwards and coming back to the level $n$ escaping absorption at level $-1$,
- the future of the contour process that just made a downward move depends only on the current level and has no memory of the earlier path.

It follows that $Z^{(n)} - 1$ has a geometric distribution whose parameter is the probability that the jumping contour process starting downwards from level $n$ will be absorbed at level $-1$ without visiting level $n$ once again.

## 5. Classification of branching processes with countably many types

Multi-type BGW-processes are classified according to the asymptotic properties of the mean matrices $\mathbf{M}^{(n)} = (m_{ij}^{(n)})_{i,j=1}^{\infty}$ with elements

$$m_{ij}^{(n)} = \mathbb{E}(Z_j^{(n)} | \mathbf{Z}^{(0)} = \mathbf{e}_i)$$

as $n \to \infty$. The assumed independence of particles implies a recursion $\mathbf{M}^{(n)} = \mathbf{M}\mathbf{M}^{(n-1)}$, where $\mathbf{M} = \mathbf{M}^{(1)}$. It follows that $\mathbf{M}^{(n)} = \mathbf{M}^n$. Given that all powers $\mathbf{M}^n$ are element-wise finite (which is always true in the linear-fractional case) the asymptotic behavior of these powers is described by the Perron–Frobenius theory for countable matrices (see Chapter 6 in [22]).

Next we remind some crucial conclusions from this theory holding for an *irreducible and aperiodic* countable matrix $\mathbf{M}$. Recall that a non-negative matrix $\mathbf{M}$ is called irreducible, if for any pair of indices $(i, j)$ there is a natural number $n$ such that $m_{ij}^{(n)} > 0$. The period of an index $i$ in an irreducible matrix $\mathbf{M}$ is defined as the greatest common divisor of all natural numbers $n$ such that $m_{ii}^{(n)} > 0$. In the irreducible case all such indices have the same period which is called the period of $\mathbf{M}$. When this period equals one the matrix $\mathbf{M}$ is called aperiodic.

Due to Theorem 6.1 from [22] all elements of the matrix power series $\mathbf{M}(s) = \sum_{n\geq 0} s^n \mathbf{M}^n$ have a common convergence radius $0 \leq R < \infty$, called the convergence parameter of the matrix $\mathbf{M}$. Furthermore, one of the two alternatives holds:

- $R$-transient case: $\sum_{n=0}^{\infty} m_{ii}^{(n)} R^n < \infty,\ i \geq 1$,
- $R$-recurrent case: $\sum_{n=0}^{\infty} m_{ii}^{(n)} R^n = \infty,\ i \geq 1$.

According to [22] (Theorem 6.2 and a remark afterwards) in the $R$-recurrent case there exist unique up to constant multipliers *positive* vectors $\mathbf{u}$ and $\mathbf{v}$ such that

$$R\mathbf{M}\mathbf{u}^{\mathrm{t}} = \mathbf{u}^{\mathrm{t}}, \qquad R\mathbf{v}\mathbf{M} = \mathbf{v}.$$

Using $Rv_j m_{ji}/v_i$ one can transform the matrix $\mathbf{M}$ into a stochastic matrix.

The $R$-recurrent case is further divided into two sub-cases: $R$-null, when $\mathbf{v}\mathbf{u}^{\mathrm{t}} = \infty$, and $R$-positive with $\mathbf{v}\mathbf{u}^{\mathrm{t}} < \infty$. In the $R$-null case (and clearly also in the $R$-transient case)

$$R^n m_{ij}^{(n)} \to 0 \quad \text{for all } i, j \geq 1. \tag{9}$$

In the $R$-positive case (Theorem 6.5 from [22]) one can scale the eigenvectors so that $\mathbf{v}\mathbf{u}^{\mathrm{t}} = 1$ and obtain

$$R^n m_{ij}^{(n)} \to u_i v_j \quad \text{for all } i, j \geq 1.$$

In the matrix notation the last element-wise convergence can be written either as

$$R^n \mathbf{M}^n \to \mathbf{u}^{\mathrm{t}}\mathbf{v}, \quad n \to \infty, \tag{10}$$

or as $\mathbf{M}^n \sim \rho^n \mathbf{u}^{\mathrm{t}}\mathbf{v}$, where $\rho = 1/R$ is the Perron–Frobenius eigenvalue of $\mathbf{M}$.

These results suggest a double classification of the BGW-processes with countably many types having a mean matrix $\mathbf{M}$. The usual classification of the multi-type BGW-processes distinguishes among subcritical, critical, or supercritical branching processes depending on whether $\rho < 1$, $\rho = 1$, or $\rho > 1$. In view of possibilities other than (10) an additional classification is needed to account for particles escaping to infinity across the type space.

**Definition 7.** A BGW-process with countably many types will be called subcritical (critical, supercritical) and transient {recurrent, null-recurrent, positively recurrent} in the type space, if its matrix of the mean offspring numbers $\mathbf{M}$ has a convergence radius $R > 1$ ($R = 1$, $R < 1$) and is $R$-transient {$R$-recurrent, $R$-null recurrent, $R$-positively recurrent}.

There are several published results for the BGW-processes with countably many types (see for example [2,15]). One of them is Theorem 1 in [18] dealing with the $R$-positively recurrent supercritical ($R < 1$) case. It states that if

$$\sum_{i=1}^{\infty} v_i \mathbb{E}\left( (\mathbf{Z}^{(1)}\mathbf{u}^{\mathrm{t}})^2 | \mathbf{Z}^{(0)} = \mathbf{e}_i \right) < \infty,$$

then for any $\mathbf{w}$ such that $\mathbf{w} \leq c\mathbf{u}$ for some positive constant $c$, the convergence $R^n \mathbf{Z}^{(n)} \mathbf{w}^{\mathrm{t}} \to Y\mathbf{v}\mathbf{w}^{\mathrm{t}}$ holds in mean square, where $Y \geq 0$ has a finite second moment. This statement is cited here just to illustrate the need for finding illuminating examples of branching processes, where conditions like $R$-positive recurrence could be verified and the values of $(R, \mathbf{u}, \mathbf{v})$ be computed in terms of the basic model parameters.

Returning to a linear-fractional branching process with parameters $(\mathbf{H}, \mathbf{g}, m)$ consider the matrix $\mathbf{M}$ of the mean offspring numbers given by (2). Clearly, irreducibility of $\mathbf{M}$ prohibits the phantom types; see Definition 6. The opposite is not true, if there exist so-called final types that never produce offspring, in other words, if $\mathbf{H}$ contains zero rows.

**Theorem 8.** *The matrix* $\mathbf{M}$ *given by* (2) *is irreducible if and only if there are no phantom types and* $\mathbf{H}$ *does not contain zero rows. If* $\mathbf{M}$ *is irreducible and aperiodic, the following three statements are valid*

(i) *the convergence parameter of* $\mathbf{M}$ *is computed as*

$$R = \begin{cases} e^{-\alpha}, & \text{if } \alpha > -\infty, \\ R_f, & \text{if } \alpha = -\infty, \end{cases} \tag{11}$$

*using Definition 5 and formula (8) for the components of* $\mathbf{d}$,

(ii) $\mathbf{M}$ *is* $R$*-recurrent if and only if* $\alpha > -\infty$,

(iii) $\mathbf{M}$ *is* $R$*-positively recurrent if and only if* $\beta < \infty$.

*In the positively recurrent case we have the element-wise convergence* (10), *where element-wise positive and finite vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are given by*

$$\mathbf{u}^{\mathrm{t}} = (1+m)\beta^{-1}\sum_{k=1}^{\infty} R^k \mathbf{H}^k \mathbf{1}^{\mathrm{t}}, \tag{12}$$

$$\mathbf{v} = \frac{m}{1+m}\sum_{k=0}^{\infty} R^k \mathbf{g}\mathbf{H}^k, \tag{13}$$

*and satisfy* $\mathbf{v}\mathbf{u}^{\mathrm{t}} = \mathbf{v}\mathbf{1}^{\mathrm{t}} = 1$ *as well as* $\mathbf{g}\mathbf{u}^{\mathrm{t}} = \frac{1+m}{m\beta}$.

**Example 2.** Assume that for some positive constant $r$ the pair $(\mathbf{H}, \mathbf{g})$ satisfies one or both of the following conditions

1. $\mathbf{gH} = r\mathbf{g}$ so that $\mathbf{gM} = (1+m)r\mathbf{g}$ and $\mathbf{v} = \mathbf{g}$,
2. $\mathbf{H1}^{\mathrm{t}} = r\mathbf{1}^{\mathrm{t}}$ so that $\mathbf{M1}^{\mathrm{t}} = (1+m)r\mathbf{1}^{\mathrm{t}}$ and $\mathbf{u} = \mathbf{1}$.

We have necessarily $r \le 1$ since $r = r\mathbf{g}\mathbf{1}^{\mathrm{t}} = \mathbf{gH1}^{\mathrm{t}} \le \mathbf{g}1^{\mathrm{t}} = 1$. In both cases we obtain $\rho = (1+m)r$, $\beta = \frac{1+m}{m}$, and $\mathbb{P}(L > n) = r^n$. Notice that $\mathbb{P}(L = \infty) = 1$ for $r = 1$.

## 6. *R*-positively recurrent case

Consider a linear-fractional BGW-process with an irreducible and aperiodic $\mathbf{M}$ assuming $\beta < \infty$. In this case according to Theorem 8 we have $\mathbf{M}^n \sim \rho^n \mathbf{u}^{\mathrm{t}}\mathbf{v}$, where $\rho = R^{-1} = e^{\alpha}$. It follows that the left eigenvector $\mathbf{v}$ describes the stable type distribution: $\mathbf{e}_i\mathbf{M}^n \sim u_i\rho^n\mathbf{v}$, and the right eigenvector $\mathbf{u}$ compares productivity of different types: $\mathbf{M}^n\mathbf{1}^{\mathrm{t}} \sim \rho^n\mathbf{u}^{\mathrm{t}}$ (so that $u_i$ can be interpreted as the "reproductive value" of type $i$). The next three propositions present basic asymptotic results for the linear-fractional BGW-processes extending similar statements for the finite-dimensional case obtained in [14] and [20].

**Proposition 9.** *In the subcritical positively recurrent case when* $\rho < 1$, *or equivalently* $\mu < 1$,

$$\mathbb{P}\big(\mathbf{Z}^{(n)} \ne \mathbf{0}\big) \sim \rho^n(1+m)^{-1}(1-\mu)\mathbf{u}^{\mathrm{t}}. \tag{14}$$

*Furthermore, for any initial type* $i$ *we get*

$$\mathbb{P}\big(\mathbf{Z}^{(n)} = \mathbf{k} | \mathbf{Z}^{(n)} \ne \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i\big) \to \mathbb{P}(\mathbf{Y} = \mathbf{k}) \quad \text{for all } \mathbf{k} \in \mathbb{Z}_+^{\infty},$$

*where* $\mathbf{Y}$ *has a distribution LF* $(\tilde{\mathbf{h}}, \tilde{\mathbf{g}}, \tilde{m})$ *with* $\tilde{m} = m\lambda(1-\mu)^{-1}$,

$$\tilde{\mathbf{h}} = (1+m)(1-\mu)^{-1}\mathbf{v} - m\mathbf{g}(\mathbf{I}-\mathbf{M})^{-1}, \quad \tilde{\mathbf{h}}\mathbf{1}^{\mathrm{t}} = 1,$$
$$\tilde{\mathbf{g}} = \lambda^{-1}(1-\mu)\mathbf{g}(\mathbf{I}-\mathbf{M})^{-1}.$$

**Proposition 10.** *In the critical positively recurrent case when* $\rho = 1$ *we have*

$$\mathbb{P}(\mathbf{Z}^{(n)} \neq \mathbf{0}) \sim n^{-1}(1+m)^{-1}\beta\mathbf{u}^{\mathrm{t}}.$$

*If a vector* $\mathbf{w}$ *has bounded components* $(\sup_{j\geq 1}|w_j| < \infty)$ *and* $\mathbf{v}\mathbf{w}^{\mathrm{t}} > 0$, *then for all* $x > 0$ *and* $i \geq 1$

$$\mathbb{P}\big(\mathbf{Z}^{(n)}\mathbf{w}^{\mathrm{t}} > nx | \mathbf{Z}^{(n)} \neq \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i\big) \to e^{-x/c_w}, \quad c_w = (1+m)\beta^{-1}\mathbf{v}\mathbf{w}^{\mathrm{t}}.$$

*In other words, conditionally on non-extinction* $n^{-1}\mathbf{Z}^{(n)}$ *weakly converges to* $X\mathbf{v}$, *where* $X$ *is exponentially distributed with mean* $(1+m)\beta^{-1}$.

**Proposition 11.** *In the supercritical positively recurrent case when* $\rho > 1$

$$\mathbb{P}(\mathbf{Z}^{(n)} \neq \mathbf{0}) \to (\rho-1)(1+m)^{-1}\beta\mathbf{u}^{\mathrm{t}}.$$

*Furthermore, for any* $\mathbf{w}$ *with bounded components and* $\mathbf{v}\mathbf{w}^{\mathrm{t}} > 0$

$$\mathbb{P}\big(\mathbf{Z}^{(n)}\mathbf{w}^{\mathrm{t}} > \rho^n x | \mathbf{Z}^{(n)} \neq \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i\big) \to e^{-x(\rho-1)/c_w}, \quad x > 0.$$

As straightforward corollaries of Propositions 9–11 we get the following asymptotic results for the linear-fractional CMJ-processes with $\beta < \infty$. The survival probability $\mathbb{P}(Z^{(n)} > 0) = \sum_{i\geq 1} g_i \mathbb{P}\big(\mathbf{Z}^{(n)} \neq \mathbf{0} | \mathbf{Z}^{(0)} = \mathbf{e}_i\big)$ satisfies a particularly transparent asymptotical formula

$$\mathbb{P}(Z^{(n)} > 0) \sim \begin{cases} e^{\alpha n}(1-\mu)(m\beta)^{-1}, & \text{if } \alpha < 0, \\ (nm)^{-1}, & \text{if } \alpha = 0, \\ (e^{\alpha}-1)m^{-1}, & \text{if } \alpha > 0. \end{cases}$$

Moreover, in the subcritical case we get a geometric conditional limit distribution

$$\mathbb{P}(Z^{(n)} = k | Z^{(n)} > 0) \to m^{k-1}(1+m)^{-k}, \quad k \geq 1,$$

in the critical case we have

$$\mathbb{P}(Z^{(n)} > nx | Z^{(n)} > 0) \to e^{-\beta x/(1+m)}, \quad x > 0,$$

and in the supercritical case

$$\mathbb{P}(Z^{(n)} > e^{\alpha n}x | Z^{(n)} > 0) \to \exp\{-x\beta(e^{\alpha}-1)/(1+m)\}, \quad x > 0.$$

These explicit results illuminate much more general limit theorems for the CMJ-processes available in [11,12,21].

## 7. Proofs of Theorems 3, 8 and Propositions 9–11

### *7.1. Proof of Theorem 3, part 2*

In Section 4 we have shown that $\mathbf{Z}^{(n)}$ has a linear-fractional distribution with unspecified parameters $(\mathbf{H}^{(n)}, \mathbf{g}^{(n)}, m^{(n)})$. Turning to the proof of relations (3), (4) and (5) observe first

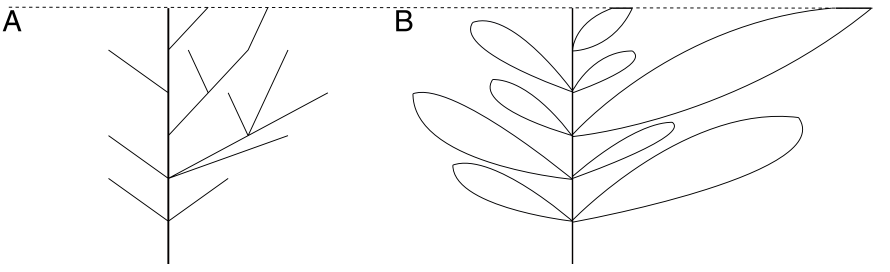


Fig. 2. A spinal representation of a BGW-tree reaching the observation level. (A) The spinal representation of the tree from Fig. 1. (B) A schematic view of the spinal representation, cf. [7].

that after multiplying (2) by $\mathbf{1}^{\mathrm{t}}$ we obtain $\mathbf{M1}^{\mathrm{t}} = (1+m)\mathbf{H1}^{\mathrm{t}}$, which leads to a useful reverse expression of $\mathbf{H}$ in terms of $\mathbf{M}$

$$\mathbf{H} = \mathbf{M} - \frac{m}{1+m}\mathbf{M1}^{\mathrm{t}}\mathbf{g}.$$

Clearly, relation (5) is a straightforward counterpart of the last relation applied to the linear-fractional distribution of $\mathbf{Z}^{(n)}$.

We prove (3) using the spinal representation of the BGW-tree illustrated in Fig. 2. Suppose that $Z^{(n)} \geq 1$. The corresponding spine of the planar BGW-tree is the leftmost lineage of particles reaching the level $n$. Recall that the unspecified parameter $m^{(n)}$ is the mean number of all branches present at level $n$ except the spinal one. Since this mean is the sum of contributions from all the lineages stemming to the right of the spine (see Fig. 2), to establish equality (3) it suffices to see that the average number of particles stemming from the spinal particle at time $k \in [0, n-1]$ equals $m\mathbf{gM}^{n-k-1}\mathbf{1}^{\mathrm{t}}$. The last assertion is a straightforward consequence of the memoryless property of geometric distribution:

- in the linear-fractional case at each level $k \in [0, n-1]$ there is a geometric with mean $m$ number of branches growing off the spine to the right of it,
- every one of such daughter branching processes produces on average $\mathbf{gM}^{n-k-1}\mathbf{1}^{\mathrm{t}}$ particles at time $n$.

Equality (4) is obtained using the same argument. It is just a detailed version of (3) taking into account the number of particles of various types existing at time $n$.

### 7.2. Renewal theory argument

This section contains two lemmata used in Section 7.3. The first lemma deals with two power series $\mathbf{M}(s) = \sum_{n\geq 0} s^n\mathbf{M}^n$ and $\mathbf{H}(s) = \sum_{n\geq 0} s^n\mathbf{H}^n$.

**Lemma 12.** *Let $f(s) = \sum_{n\geq 1} d_n s^n$ with $d_n$ given by (8). The vector $\mathbf{M}(s)\mathbf{1}^{\mathrm{t}}$ is element-wise finite if and only if $mf(s) < 1$. If $mf(s) < 1$, then*

$$\mathbf{M}(s) = \mathbf{H}(s) + \frac{m}{1 - mf(s)}(\mathbf{H}(s) - \mathbf{I})\mathbf{1}^{\mathrm{t}}\mathbf{gH}(s).$$

**Proof.** Due to (2) we have $\mathbf{M}^{n+1} = (\mathbf{H} + m\mathbf{HG})\mathbf{M}^n$, where $\mathbf{G} = \mathbf{1}^{\mathrm{t}}\mathbf{g}$. Using induction we obtain

$$\mathbf{M}^n = \mathbf{H}^n + m\sum_{i=1}^{n}\mathbf{H}^i\mathbf{GM}^{n-i}.$$

Putting $\bar{\mathbf{H}}(s) = \mathbf{H}(s) - \mathbf{I}$ and $M(s) = \mathbf{g}\mathbf{M}(s)\mathbf{1}^{\mathrm{t}}$ we derive first

$$\mathbf{M}(s) = \mathbf{H}(s) + m\bar{\mathbf{H}}(s)\mathbf{G}\mathbf{M}(s) \tag{15}$$

and then

$$M(s) = 1 + f(s) + mf(s)M(s).$$

Thus given $mf(s) < 1$, we have

$$M(s) = \frac{1 + f(s)}{1 - mf(s)}, \tag{16}$$

and we see that the vector $\mathbf{M}(s)\mathbf{1}^{\mathrm{t}} = \mathbf{H}(s)\mathbf{1}^{\mathrm{t}} + m\bar{\mathbf{H}}(s)\mathbf{1}^{\mathrm{t}}M(s)$ is element-wise finite if and only if $mf(s) < 1$.

According to (15) we have for all $n \geq 1$

$$\mathbf{M}(s) = \mathbf{H}(s) + \sum_{k=1}^{n-1}(m\bar{\mathbf{H}}(s)\mathbf{G})^k\mathbf{H}(s) + (m\bar{\mathbf{H}}(s)\mathbf{G})^n\mathbf{M}(s),$$

and moreover

$$\begin{aligned}
(\bar{\mathbf{H}}(s)\mathbf{G})^n &= \left(\sum_{i=1}^{\infty} s^i\mathbf{H}^i\mathbf{1}^{\mathrm{t}}\mathbf{g}\right)^n \\
&= \left(\sum_{i=1}^{\infty} s^i\mathbf{H}^i\mathbf{1}^{\mathrm{t}}\right)\left(\sum_{i=1}^{\infty} s^i\mathbf{g}\mathbf{H}^i\mathbf{1}^{\mathrm{t}}\right)\ldots\left(\sum_{i=1}^{\infty} s^i\mathbf{g}\mathbf{H}^i\mathbf{1}^{\mathrm{t}}\right)\mathbf{g} \\
&= f^{n-1}(s)\sum_{i=1}^{\infty} s^i\mathbf{H}^i\mathbf{1}^{\mathrm{t}}\mathbf{g} = f^{n-1}(s)\bar{\mathbf{H}}(s)\mathbf{G}.
\end{aligned}$$

It follows that for $s$ such that $mf(s) < 1$, the term

$$(m\bar{\mathbf{H}}(s)\mathbf{G})^n\mathbf{M}(s) = m^n f^{n-1}(s)\bar{\mathbf{H}}(s)\mathbf{G}\mathbf{M}(s)$$

vanishes as $n \to \infty$ and the previous two relations yield

$$\begin{aligned}
\mathbf{M}(s) &= \mathbf{H}(s) + \sum_{n=1}^{\infty}(m\bar{\mathbf{H}}(s)\mathbf{G})^n\mathbf{H}(s) \\
&= \mathbf{H}(s) + \sum_{n=1}^{\infty} m^n f^{n-1}(s)\bar{\mathbf{H}}(s)\mathbf{G}\mathbf{H}(s) \\
&= \mathbf{H}(s) + \frac{m}{1 - mf(s)}(\mathbf{H}(s) - \mathbf{I})\mathbf{1}^{\mathrm{t}}\mathbf{g}\mathbf{H}(s). \quad \square
\end{aligned}$$

The following well-known renewal theorem taken from Chapter XIII.4 in [6] will be used by us several times.

**Lemma 13.** *Let* $A(s) = \sum_{n=0}^{\infty} a_n s^n$ *be a probability generating function and* $B(s) = \sum_{n=0}^{\infty} b_n s^n$ *be a generating function for a non-negative sequence so that* $A(1) = 1$ *while* $B(1) \in (0, \infty)$. *Then the non-negative sequence defined by* $\sum_{n=0}^{\infty} t_n s^n = \frac{B(s)}{1-A(s)}$ *is such that* $t_n \to \frac{B(1)}{A'(1)}$ *as* $n \to \infty$.

*7.3. Proof of Theorem 8*

We show first that if there are no phantom types and $\mathbf{H}$ has no zero rows, then for any $j$ there exists such $n = n_j$ that $m_{ij}^{(n)} > 0$ for all $i$. This easily follows from the inequality $\mathbf{M}^n \geq m\mathbf{H}\mathbf{1}^{\mathrm{t}}\mathbf{g}\mathbf{H}^{n-1}$ which comes from (15). Indeed, on one hand, all components of the vector $\mathbf{H}\mathbf{1}^{\mathrm{t}}$ are positive. On the other hand, the absence of phantom types implies that for the given $j$ we can find such $n = n_j \geq 2$ that the $j$-th component of the vector $\mathbf{g}\mathbf{H}^{n-1}$ is positive.

Assume from now on that $\mathbf{M}$ is irreducible and aperiodic. Statements (i) and (ii) of Theorem 8 follow directly from Lemma 12.

Assertion (iii) is easily obtained by combining (16) and Lemma 13. Using

$$R^n\mathbf{g}\mathbf{M}^n\mathbf{1}^{\mathrm{t}} \to \frac{1+f(R)}{mf'(R)} = \frac{1+m}{\beta m}$$

we conclude that (9) holds if and only if $\beta = \infty$.

We show next that $\mathbf{H}(R)$ is element-wise finite provided $\alpha > -\infty$. First notice that $\mathbf{g}\mathbf{H}(R)\mathbf{1}^{\mathrm{t}} = f(R) = 1/m$. On the other hand, in the absence of phantom types for any $i \geq 1$ we can find a $k = k_i$ and a positive $c_i$ such that $\mathbf{g}\mathbf{H}^k \geq c_i\mathbf{e}_i$ implying

$$c_i\mathbf{e}_i\mathbf{H}(R) \leq \sum_{n=0}^{\infty} R^n\mathbf{g}\mathbf{H}^{k+n} \leq R^{-k}\sum_{n=0}^{\infty} R^n\mathbf{g}\mathbf{H}^n = R^{-k}\mathbf{g}\mathbf{H}(R)$$

so that $\mathbf{e}_i\mathbf{H}(R)\mathbf{1}^{\mathrm{t}} \leq c_i^{-1}R^{-k}m^{-1} < \infty$.

Now let $\beta < \infty$. Consider vectors $\mathbf{u}$ and $\mathbf{v}$ which, thanks to the just proved finiteness of $\mathbf{H}(R)$, are well-defined by (12) and (13). The claimed equality $\mathbf{v}\mathbf{u}^{\mathrm{t}} = 1$ follows from

$$m\mathbf{g}\mathbf{H}(R)(\mathbf{H}(R) - \mathbf{I})\mathbf{1}^{\mathrm{t}} = m\sum_{n=1}^{\infty} nR^n\mathbf{g}\mathbf{H}^n\mathbf{1}^{\mathrm{t}} = \beta,$$

which is a consequence of

$$\begin{aligned}\mathbf{H}(s)\mathbf{H}(s) - \mathbf{H}(s) &= \sum_{i=1}^{\infty}\mathbf{H}^i s^i\sum_{k=0}^{\infty}\mathbf{H}^k s^k \\ &= \sum_{i=1}^{\infty}\sum_{n=i}^{\infty}\mathbf{H}^n s^n = \sum_{n=1}^{\infty}\sum_{i=1}^{n}\mathbf{H}^n s^n = \sum_{n=1}^{\infty} n\mathbf{H}^n s^n.\end{aligned}$$

It remains to prove (10). To this end define a sequence of matrices $\mathbf{B}_n$ by

$$\sum_{n=0}^{\infty}\mathbf{B}_n s^n = \tfrac{m}{1-mf(Rs)}(\mathbf{H}(Rs) - \mathbf{I})\mathbf{1}^{\mathrm{t}}\mathbf{g}\mathbf{H}(Rs), \quad s \in [0, 1)$$

so that $R^n\mathbf{M}^n = R^n\mathbf{H}^n + \mathbf{B}_n$ due to Lemma 12. According to Lemma 13 we have an element-wise convergence $\mathbf{B}_n \to \mathbf{u}^{\mathrm{t}}\mathbf{v}$ as $n \to \infty$ and it remains to see that each element of $R^n\mathbf{H}^n$ converges to zero, since $\mathbf{H}(R) = \sum_{n\geq 0} R^n\mathbf{H}^n$ is element-wise finite.

*7.4. Proof of Propositions 9–11*

**Proof of Proposition 9.** From (3) and (16) we obtain

$$m^{(n)} \to m\frac{1+f(1)}{1-mf(1)} = \tilde{m}$$

which together with (6) implies (14). The statement on the convergence of the conditional distribution of $\mathbf{Z}^{(n)}$ follows from (7):

$$\mathbb{E}\big(\mathbf{s}^{\mathbf{Z}^{(n)}}|\mathbf{Z}^{(n)} \neq \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i\big) \to \frac{\sum_{j=1}^{\infty} \tilde{h}_j s_j}{1 + \tilde{m} - \tilde{m} \sum_{j=1}^{\infty} \tilde{g}_j s_j},$$

since $m^{(n)}\mathbf{g}^{(n)} \to m\mathbf{g}(\mathbf{I} - \mathbf{M})^{-1} = \tilde{m}\tilde{\mathbf{g}}$ and

$$\mathbf{H}^{(n)} \sim \rho^n(\mathbf{u}^{\mathrm{t}}\mathbf{v} - (1-\mu)(1+m)^{-1}\mathbf{u}^{\mathrm{t}}\tilde{m}\tilde{\mathbf{g}}) = \rho^n(1-\mu)(1+m)^{-1}\mathbf{u}^{\mathrm{t}}\tilde{\mathbf{h}}. \quad \square$$

**Proof of Proposition 10.** Lemma 13 and relations (3), (16) imply that in the critical case

$$m^{(n)} \sim n(1+m)\beta^{-1}.$$

Thus the stated asymptotics for the survival probability follows from (6). Using (7) we express the conditional moment generating function as

$$\mathbb{E}\left(e^{zn^{-1}\mathbf{Z}^{(n)}\mathbf{w}^{\mathrm{t}}}|\mathbf{Z}^{(n)} \neq \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i\right) = \frac{(1-h_{i0}^{(n)})^{-1} \sum_{j=1}^{\infty} h_{ij}^{(n)} e^{zw_j/n}}{1 + m^{(n)} - m^{(n)} \sum_{j=1}^{\infty} g_j^{(n)} e^{zw_j/n}}.$$

Since $m^{(n)}\mathbf{g}^{(n)} \sim n(1+m)\beta^{-1}\mathbf{v}$, we obtain

$$(1-h_{i0}^{(n)})^{-1} \sum_{j=1}^{\infty} h_{ij}^{(n)} e^{zw_j/n} \to 1,$$

and applying the monotone convergence theorem for series, we get

$$m^{(n)} \sum_{j=1}^{\infty} g_j^{(n)}(e^{zw_j/n} - 1) \to z(1+m)\beta^{-1}\mathbf{v}\mathbf{w}^{\mathrm{t}}.$$

Now, the asserted weak convergence follows from the convergence of moment generating functions

$$\mathbb{E}\left(e^{zn^{-1}\mathbf{Z}^{(n)}\mathbf{w}^{\mathrm{t}}}|\mathbf{Z}^{(n)} \neq \mathbf{0}, \mathbf{Z}^{(0)} = \mathbf{e}_i\right) \to \frac{1}{1 - z(1+m)\beta^{-1}\mathbf{v}\mathbf{w}^{\mathrm{t}}}$$

for all $z \in [0, z_0]$, where $z_0$ is some positive number (see [5]). $\square$

**Proof of Proposition 11.** Rewrite (3) as

$$R^{n-1}m^{(n)} = m \sum_{k=0}^{n-1} R^k R^{n-1-k}\mathbf{g}\mathbf{M}^k\mathbf{1}^{\mathrm{t}}$$

to obtain the following consequence of (16)

$$\sum_{n=1}^{\infty}(Rs)^{n-1}m^{(n)} = \frac{m(1+f(sR))}{(1-mf(sR))(1-Rs)}.$$

Thus Lemma 13 entails

$$m^{(n)} \sim \rho^n (1+m)\beta^{-1}(\rho-1)^{-1}.$$

This together with (6) and (10) gives the stated formula for the survival probability. The assertion on weak convergence is proved in a similar way as in the critical case above. □

## Acknowledgments

Several insightful comments and constructive suggestions of two anonymous referees have helped the author to significantly improve the presentation of the paper making the assertions more precise and the proofs more rigorous. This work was supported by the Swedish Research Council grant 621-2010-5623.

## References

[1] S. Asmussen, Applied Probability and Queues, Springer, New York, 2003.
[2] K. Athreya, H. Kang, Some limit theorems for positive recurrent Markov chains I and II, Adv. Appl. Probab. 30 (1998) 693–722.
[3] K. Athreya, P. Ney, Branching Processes, John Wiley & Sons, London–New York–Sydney, 1972.
[4] A. Barbour, M. Luczak, Laws of large numbers of epidemic models with countably many types, Ann. Appl. Probab. 18 (2008) 2208–2238.
[5] J.H. Curtiss, A note on the theory of moment generating functions, Ann. Math. Statist. 13 (1942) 430–433.
[6] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. I, second ed., John Wiley & Sons, London–New York–Sydney, 1959.
[7] J. Geiger, Elementary new proofs of classical limit theorems for Galton–Watson processes, J. Appl. Probab. 36 (1999) 301–309.
[8] J. Geiger, G. Kersting, Depth-first search of random trees, and Poisson point processes, in: Classical and Modern Branching Processes, (Minneapolis, MN, 1994), in: IMA Math. Appl. Vol. 84, Springer, New York, 1997, pp. 111–126.
[9] P. Haccou, P. Jagers, V.A. Vatutin, Branching Processes: Variation, Growth and Extinction of Populations, Cambridge University Press, Cambridge, 2005.
[10] FM. Hoppe, Coupling and the non-degeneracy of the limit in some plasmid reproduction models, Theor. Popul. Biol. 52 (1997) 27–31.
[11] P. Jagers, Branching Processes with Biological Applications, Wiley, New-York, 1975.
[12] P. Jagers, O. Nerman, The growth and composition of branching populations, Adv. Appl. Probab. 16 (1984) 221–259.
[13] P. Jagers, S. Sagitov, General branching processes in discrete time as random trees, Bernoulli 14 (2008) 949–962.
[14] A. Joffe, G. Letac, Multitype linear fractional branching processes, J. Appl. Probab. 43 (2006) 1091–1106.
[15] H. Kesten, Supercritical branching processes with countably many types and the sizes of random cantor sets, in: Probability, Statistics and Mathematics. Papers in Honor of Samuel Karlin, Academic Press, New York, 1989, pp. 108–121.
[16] M. Kimmel, D. Axelrod, Branching Processes in Biology, Springer, New York, 2002.
[17] A. Lambert, The contour of splitting trees is a Levy process, Ann. Probab. 38 (2010) 348–395.
[18] S.-T.C. Moy, Extensions of a limit theorem of Everett Ulam and Harrison multi-type branching processes to a branching process with countably many types, Ann. Math. Statist. 38 (1967) 992–999.
[19] A.G. Pakes, Biological Applications of Branching Processes, in: Handbook of Statistics vol. 21, Elsevier Science, Amsterdam, Netherlands, 2003, pp. 693–773.
[20] E. Pollak, Survival probabilities and extinction times for some multitype branching processes, Adv. Appl. Probab. 6 (1974) 446–462.
[21] S. Sagitov, A key limit theorem for critical branching processes, Stochastic Process. Appl. 56 (1995) 87–100.
[22] E. Seneta, Non-negative matrices and Markov chains, in: Springer Series in Statistics, No. 21, Springer, New-York, 2006.
[23] E. Seneta, S. Tavare, Some stochastic models for plasmid copy number, Theor. Popul. Biol. 23 (1983) 241–256.