

EXPONENTIAL GROWTH OF BIFURCATING PROCESSES WITH ANCESTRAL DEPENDENCE

SANA LOUHICHI* AND

BERNARD YCART,** *Université Grenoble Alpes*

Abstract

Branching processes are classical growth models in cell kinetics. In their construction, it is usually assumed that cell lifetimes are independent random variables, which has been proved false in experiments. Models of dependent lifetimes are considered here, in particular bifurcating Markov chains. Under the hypotheses of stationarity and multiplicative ergodicity, the corresponding branching process is proved to have the same type of asymptotics as its classic counterpart in the independent and identically distributed supercritical case: the cell population grows exponentially, the growth rate being related to the exponent of multiplicative ergodicity, in a similar way as to the Laplace transform of lifetimes in the i.i.d. case. An identifiable model for which the multiplicative ergodicity coefficients and the growth rate can be explicitly computed is proposed.

Keywords: Branching process; bifurcating Markov chain; multiplicative ergodicity; cell kinetics

2010 Mathematics Subject Classification: Primary 60J85

Secondary 92D25

1. Introduction

Let \mathbb{T} denote the infinite complete binary tree where each vertex has exactly two descendants. Let $(T_v)_{v \in \mathbb{T}}$ be a bifurcating process, i.e. a family of positive random variables indexed by \mathbb{T} , defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The vertices of \mathbb{T} are interpreted as cells, and T_v as the lifetime of cell v . The root (ancestor) of the tree is born at time 0, and let N_t be the number of individuals alive at time t : $(N_t)_{t \geq 0}$ is a continuous time branching process (precise definitions will be given in Section 2). If the lifetimes are independent and identically distributed (i.i.d.), the population N_t grows exponentially in t : this is a particular case of one of the most basic results of the theory (see Bellmann and Harris [2], Harris [15, Chapter VI], and Athreya and Ney [1, Chapter IV]).

Theorem 1.1. *Assume that the lifetimes T_v are i.i.d. copies of an almost surely (a.s.) positive random variable T with nonlattice distribution. Then*

$$\lim_{t \rightarrow \infty} e^{-vt} N_t = W \quad \text{a.s.},$$

where

- W is a random variable with expectation C and finite variance,

Received 24 January 2014; revision received 21 May 2014.

* Postal address: Laboratoire Jean Kuntzmann, Université Grenoble Alpes, 51 rue des Mathématiques, 38041 Grenoble cedex 9, France.

** Email address: bernard.ycart@ujf-grenoble.fr

- the growth rate (also called Malthusian parameter) ν is such that

$$2\mathbb{E}\{e^{-\nu T}\} = 1, \quad (1.1)$$

- the proportionality constant C is

$$C = (4\nu\mathbb{E}\{Te^{-\nu T}\})^{-1}. \quad (1.2)$$

The aim of this paper is to extend Theorem 1.1 to models in which lifetimes may be dependent, and, in particular, to bifurcating Markov chains (BMCs). Our main result (Theorem 3.1) generalizes Theorem 1.1 to the case where $(T_\nu)_{\nu \in \mathbb{T}}$ is a multiplicatively ergodic, stationary BMC. The growth rate ν and the proportionality constant C in that case are related to the multiplicative ergodicity coefficients of birth dates.

The applications of branching processes to cell lineage studies has a long history, and the independence of lifetimes was questioned in the early 1950s [19]. Indeed, actual data show two types of correlation [34]: between the lifetimes of a mother and its two daughters, and between the two sisters conditioning on the mother; they will be referred to as mother correlation and sister correlation. In the mid 1950s Powell [30] remarked that sister correlations do not influence exponential growth (see also [6], [15, Section 28.2, p. 158], and [16]). The effect of mother correlation on growth rates was discussed by Harvey [16], yet finding any exact dependence in a given model has remained an open question since. Many models have been proposed to account for ancestry dependence, see [24], [26]. Here, lifetimes are seen as a stochastic process indexed by the binary tree; see Pemantle [28] and Benjamini and Peres [3] as general references on tree-indexed processes. Under a minimal hypothesis of stationarity, exponential growth for the mean population size $\mathbb{E}\{N_t\}$ is proved, and the growth rate ν as well as the proportionality constant C are expressed in terms of the Laplace transforms of cell birth dates (Theorem 2.1). Asymptotics of Laplace transforms for partial sums of a Markov chain are usually described by multiplicative ergodicity properties, which have been thoroughly studied by Meyn and his co-workers [22]; see [25, p. 519] for a short introduction. It is therefore natural to use a BMC as a model of lifetimes, see Benjamini and Peres [3] for tree-indexed Markov chains, Guyon [14] for applications to cell lineage data, and Bitseki Penda *et al.* [4] for a recent reference. Under a multiplicative ergodicity condition, Theorem 1.1 is generalized: $e^{-\nu t} N_t$ is shown to converge a.s.; moreover, the growth rate ν and the proportionality constant C are explicitly related to the multiplicative ergodicity coefficients (Theorem 3.1). The proof follows a classical scheme, already used by Bellman and Harris for the i.i.d. case of [2]. It consists of studying the first and second moments of N_t , then proving convergence in the quadratic mean, and finally deducing almost sure convergence. This is related to what Pemantle calls the ‘second-moment method’ [28, Section 2.3]. In applying it, we have tried to provide the weakest possible conditions at each step, starting with the stationarity hypothesis of Theorem 2.1. Proposition 6.1 provides sufficient conditions that ensure quadratic convergence of $e^{-\nu t} N_t$, Proposition 6.2 provides conditions for a.s. convergence. These conditions will be shown to hold under the hypotheses of Theorem 3.1.

An obvious drawback for applications is that the growth rate ν and the proportionality constant C cannot be computed in general. Therefore, an explicit model, potentially adjustable to observed data and for which ν and C can be computed in terms of the transition kernel, had to be proposed. It was constructed as a quadratic transformation of a bifurcating autoregressive process [5], [8], [9], [14]. It depends on five identifiable parameters, (location, scale, shape for lifetime distribution, plus mother and sister correlations) and can be fitted to actual data.

Having in mind the application to cell lineage studies, it was natural to write the results for the binary tree. Nevertheless, they extend quite straightforwardly to processes on the infinite complete k -ary tree for $k > 2$, at the only expense of heavier notation. Remarks in the text will make the generalization more precise. Further extensions are possible, in particular to the case where \mathbb{T} is a supercritical Galton–Watson tree and the lifetimes of the daughters are independent conditionally on their common mother; models with cell deaths, such as proposed in [9] are included. They will be the object of future work.

The paper is organized as follows. In Section 2 the branching process $(N_t)_{t \geq 0}$ associated to a bifurcating lifetime process $(T_v)_{v \in \mathbb{T}}$ is defined. Two notions of stationarity along lineages are introduced and the exponential growth of $\mathbb{E}\{N_t\}$ is proved. Section 3 is devoted to the definition of a BMC and the statement of Theorem 3.1. The explicit example of a BMC for which the multiplicative ergodicity coefficients can be computed is presented in Section 4. The relation between mother correlation and growth rate for a fixed marginal distribution of lifetimes is discussed in Section 5. Section 6 is devoted to conditions under which $e^{-\nu t} N_t$ converges in L^2 and almost surely. These conditions are verified for a multiplicatively ergodic BMC in Section 7.

2. Stationary bifurcating processes

In this section, notation on bifurcating processes is introduced. The birth date process $(S_v)_{v \in \mathbb{T}}$ and the branching process $(N_t)_{t \geq 0}$ associated to a bifurcating process $(T_v)_{v \in \mathbb{T}}$ are defined and related by Lemma 2.1. Two notions of stationarity are introduced: birth-stationarity (Definition 2.2) is the stationarity of birth dates in a given generation; fork-stationarity (Definition 2.3) is the stationarity of couples of birth dates when the generations of the two cells and their most recent common ancestor are fixed. Under birth-stationarity, the expectation of N_t is proved to grow exponentially, and the parameters of exponential growth ν and C are related to the Laplace transforms of birth dates (Theorem 2.1).

Some classical notation for infinite trees will be recalled first; see [28]. The infinite rooted complete binary tree is defined by \mathbb{T} and its root by 0. If v is a vertex of \mathbb{T} , the number of edges connecting v to the root is denoted by $|v|$. If v and w are two vertices of \mathbb{T} , $v \preceq w$ is the order relation that holds if v is in the path from 0 to w ; $v \wedge w$ is the most recent common ancestor of v and w , i.e. the vertex at which the paths from 0 to v and w diverge. If $v \neq 0$, \tilde{v} is the vertex such that $\tilde{v} \preceq v$ and $|\tilde{v}| = |v| - 1$ (referred to as the mother of v). For $n \geq 0$, the n th generation Γ_n is the set of vertices v such that $|v| = n$ (vertices at distance n from the root). One simple way to explicitly construct \mathbb{T} is to identify Γ_n to the set of binary vectors of length $n + 1$ with first coordinate 0. With that identification, $v \preceq w$ if and only if v coincides with the $|v| + 1$ first coordinates of w . The mother of v , \tilde{v} is deduced from v by removing its last coordinate. The two daughters of v are obtained by appending to v a new coordinate 0 or 1: they will be denoted by $v0$ and $v1$. The concatenation of n zeros will be denoted by $0^n \in \Gamma_{n-1}$. Besides algorithmic considerations, one advantage of this construction is to naturally endow \mathbb{T} with the alphabetical order.

A bifurcating process is a set of random variables $(T_v)_{v \in \mathbb{T}}$ indexed by the binary tree \mathbb{T} . Here, the T_v 's are a.s. positive: T_v is the lifetime of cell v . The birth date process $(S_v)_{v \in \mathbb{T}}$ is also a bifurcating process: S_v is the sum of cell lifetimes from 0 to \tilde{v} . The branching process $(N_t)_{t \geq 0}$ is the counting process of living cells at time t .

Definition 2.1. Let $(T_v)_{v \in \mathbb{T}}$ be a bifurcating process.

(i) For $v \in \mathbb{T}$, the birth date of cell v is defined by $S_0 = 0$ and for $|v| > 0$,

$$S_v = S_{\bar{v}} + T_{\bar{v}}.$$

(ii) For $t \geq 0$, the number of living cells at time t is defined by

$$N_t = \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_v \leq t\}} - \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_{v0} \leq t\}}, \tag{2.1}$$

where $\mathbf{1}_A$ denotes the indicator of event A . As a convention at this point, $N_t = +\infty$ if the sums are infinite.

If S_v is the birth date of cell v , the common birth date of its two daughters $S_{v0} = S_{v1}$ is also the death date of v . So (2.1) expresses the fact that cells alive at time t are the set difference of cells born no later than t with cells dead no later than t . A simpler expression will be used.

Lemma 2.1. *With the notation above,*

$$N_t = \frac{1}{2} + \frac{1}{2} \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_v \leq t\}}. \tag{2.2}$$

Proof. From (2.1), and using the relation $S_{v0} = S_{v1}$,

$$N_t = \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_v \leq t\}} - \frac{1}{2} \sum_{w \in \mathbb{T}, w \neq 0} \mathbf{1}_{\{S_w \leq t\}} = 1 + \frac{1}{2} \sum_{v \in \mathbb{T}, v \neq 0} \mathbf{1}_{\{S_v \leq t\}},$$

hence, (2.2) holds.

Remark 2.1. On the k -ary tree, (2.2) becomes

$$N_t = \frac{1}{k} + \frac{k-1}{k} \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_v \leq t\}}.$$

Consider the particular case where lifetimes in a given generation are constant,

$$T_v = T_{0^{n+1}} \quad \text{for all } v \in \Gamma_n.$$

Define by S_n the common birth date of all cells in generation Γ_n , and assume that a law of large numbers is satisfied, i.e.

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \bar{t} > 0 \quad \text{a.s.}$$

The rank of the generation alive at time t , denoted by G_t , is the counting process associated to the sequence $(S_n)_{n \in \mathbb{N}}$, and $N_t = 2^{G_t}$. Since N_t doubles at S_n , $e^{-\nu t} N_t$ never converges, although

$$\lim_{t \rightarrow \infty} \frac{\log(N_t)}{t} = \frac{\log(2)}{\bar{t}} \quad \text{a.s.}$$

Consider now,

$$\frac{\log(\mathbb{E}\{N_t\})}{t} = \frac{\log(\mathbb{E}\{e^{G_t \log 2}\})}{t}.$$

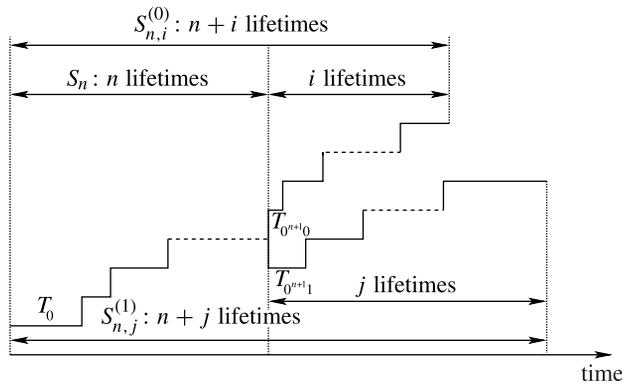


FIGURE 1: Birth dates $S_{n,i}^{(0)}$ and $S_{n,j}^{(1)}$ of the first two cells of generations $n + i$ and $n + j$ whose most recent common ancestor is in generation n .

The convergence of $\log(\mathbb{E}\{e^{\theta G_t}\})/t$ is a Gärtner–Ellis condition on G_t . Glynn and Whitt [13] proved that it is equivalent to the analogous condition on S_n . But the convergence of $\log(\mathbb{E}\{N_t\})/t$ does not imply that of $e^{-\nu t}\mathbb{E}\{N_t\}$. A law of large numbers, even strengthened by large deviation inequalities, does not suffice to prove our results: additional hypotheses are needed. We begin with the stationarity requirements.

The notion of stationarity that seems the most natural is invariance through automorphisms of the tree; see [27]. It will be satisfied by the BMC models of the next two sections. Weaker hypotheses will suffice for our preliminary convergence results. The first hypothesis says that birth dates of cells in a given generation have the same distribution. For $n \geq 0$, we will define by S_n the birth date of the first cell in generation Γ_n , by alphabetical order,

$$S_n = S_{0^{n+1}} = T_0 + T_{0^2} + \dots + T_{0^n}.$$

Definition 2.2. The bifurcating process $(T_v)_{v \in \mathbb{T}}$ is birth-stationary if for all $n \in \mathbb{N}$ and for all $v \in \Gamma_n$,

$$S_v \stackrel{D}{=} S_n,$$

where ‘ $\stackrel{D}{=}$ ’ means equality in distribution.

Observe that birth-stationarity does not imply that lifetimes T_v are identically distributed, even in a given generation. It will be used to prove the Cesàro convergence of $\mathbb{E}\{e^{-\nu t} N_t\}$ in Theorem 2.1 below. For the convergence in quadratic mean and a.s., a stronger notion will be used: the joint distribution of the birth dates of two cells in generations $n + i$ and $n + j$ with most recent common ancestor in generation n , should depend only on n , i , and j . The first such couple in alphabetical order is $(0^{n+1+i}, 0^{n+1}10^{j-1})$. The corresponding birth dates will be denoted by $S_{n,i}^{(0)}$ and $S_{n,j}^{(1)}$ (see Figure 1),

$$S_{n,i}^{(0)} = S_{n+i} = S_{0^{n+1+i}} \quad \text{and} \quad S_{n,j}^{(1)} = S_{0^{n+1}10^{j-1}}.$$

Definition 2.3. The bifurcating process $(T_v)_{v \in \mathbb{T}}$ is fork-stationary if for all $(n, i, j) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^*$ and for all $(v, w) \in \Gamma_{n+i} \times \Gamma_{n+j}$ such that $v \wedge w \in \Gamma_n$,

$$(S_v, S_w) \stackrel{D}{=} (S_{n,i}^{(0)}, S_{n,j}^{(1)}).$$

For $i = 0$, the definition includes the $v \prec w$ case: for all $n, j \in \mathbb{N} \times \mathbb{N}^*$, for all $(v, w) \in \Gamma_n \times \Gamma_{n+j}$ such that $v \prec w$,

$$(S_v, S_w) \stackrel{D}{=} (S_n, S_{n+j}).$$

In particular, all couples (S_v, S_{v0}) are identically distributed for $v \in \Gamma_n$, hence lifetimes in a given generation have the same distribution.

Remark 2.2. On the k -ary tree, one might think that forks should have k prongs. Yet fork-stationarity is used in Proposition 6.1 to express $\mathbb{E}\{N_t N_{t+\tau}\}$ in terms of the joint distribution of $(S_{n,i}^{(0)}, S_{n,j}^{(1)})$. This remains the same on the k -ary tree.

The main result of this section concerns the exponential growth of $\mathbb{E}\{N_t\}$; it relates the growth rate ν and the proportionality constant C to the Laplace transform of S_n . In order to enhance the link with Theorem 1.1, we chose to express our results in terms of Laplace transforms instead of characteristic functions or logarithmic moment generating functions, as is customary in large deviations theory [10]. Throughout this paper, the Laplace transforms evaluated at $\gamma \geq 0$ of S_n , and of S_n conditioned on $T_0 = u$, will be denoted by $\mathcal{L}_n(\gamma)$ and $\mathcal{L}_n(\gamma, u)$,

$$\mathcal{L}_n(\gamma) = \mathbb{E}\{e^{-\gamma S_n}\} \quad \text{and} \quad \mathcal{L}_n(\gamma, u) = \mathbb{E}\{e^{-\gamma S_n} \mid T_0 = u\}.$$

Theorem 2.1. *Let $(T_v)_{v \in \mathbb{T}}$ be a birth-stationary bifurcating process. Assume that ν and C given below are well defined, positive, and finite,*

$$\nu := \inf \left\{ \gamma > 0, \sum_{n=1}^{\infty} 2^n \mathcal{L}_n(\gamma) < \infty \right\}, \tag{2.3}$$

$$C := \lim_{\gamma \searrow 0} \frac{\gamma}{\gamma + \nu} \sum_{n=1}^{\infty} 2^{n-1} \mathcal{L}_n(\gamma + \nu). \tag{2.4}$$

Then, for all $t \geq 0$, $\mathbb{E}\{N_t\} < \infty$ and

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t e^{-\nu s} \mathbb{E}\{N_s\} ds = C. \tag{2.5}$$

In the particular case where the lifetimes T_v are i.i.d. nonlattice random variables, $\mathcal{L}_n(\gamma) = (\mathbb{E}\{e^{-\gamma T_0}\})^n$; it is straightforward to check that (2.3) and (2.4) reduce to (1.1) and (1.2).

Proof. From (2.2),

$$N_t = \frac{1}{2} + \frac{1}{2} \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_v \leq t\}} = 1 + \frac{1}{2} \sum_{n=1}^{\infty} \sum_{v \in \Gamma_n} \mathbf{1}_{\{S_v \leq t\}}.$$

By birth-stationarity,

$$\mathbb{E}\{N_t\} = 1 + \sum_{n=1}^{\infty} 2^{n-1} \mathbb{P}\{S_n \leq t\}. \tag{2.6}$$

By Markov’s inequality, for all $\gamma > 0$ and $t \geq 0$,

$$\mathbb{P}\{S_n \leq t\} \leq e^{\gamma t} \mathbb{E}\{e^{-\gamma S_n}\}.$$

The hypotheses of Theorem 2.1 imply that there exists $\gamma > \nu$ such that

$$\sum_{n=1}^{\infty} 2^{n-1} \mathbb{E}\{e^{-\gamma S_n}\} < \infty.$$

Hence, $\mathbb{E}\{N_t\} < \infty$ for all $t \geq 0$. Consider now:

$$A_\nu(t) = e^{-\nu t} \sum_{n=1}^{\infty} 2^{n-1} \mathbb{P}\{S_n \leq t\}.$$

Let $\tilde{A}_\nu(\gamma)$ be the Laplace transform of A_ν . The Laplace transform of $\mathbb{P}\{S_n \leq t\}$ evaluated at $\gamma > 0$ is $(1/\gamma)\mathcal{L}_n(\gamma)$. The Laplace transform of $e^{-\nu t}\mathbb{P}\{S_n \leq t\}$ is $(1/(\gamma + \nu))\mathcal{L}_n(\gamma + \nu)$. Therefore,

$$\tilde{A}_\nu(\gamma) = \frac{1}{\gamma + \nu} \sum_{n=1}^{\infty} 2^{n-1} \mathcal{L}_n(\gamma + \nu).$$

By (2.4),

$$\lim_{\gamma \searrow 0} \gamma \tilde{A}_\nu(\gamma) = C.$$

If both $\lim_{t \rightarrow +\infty} A_\nu(t)$ and $\lim_{\gamma \searrow 0} \gamma \tilde{A}_\nu(\gamma)$ exist, the fact that they are equal is a well-known basic result of Laplace transform theory, known as the final value theorem. Deducing the former limit from the existence of the latter requires a Tauberian theorem, see Feller [12, Section XIII.5] or Korevaar [23]. As a particular case of [12, Theorem 2, p. 445],

$$\lim_{\gamma \searrow 0} \gamma \tilde{A}_\nu(\gamma) = C \iff \lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t A_\nu(s) ds = C,$$

which is the required result.

Remark 2.3. On the k -ary tree, (2.3) and (2.4) become

$$\nu = \inf \left\{ \gamma > 0, \sum_{n=1}^{\infty} k^n \mathcal{L}_n(\gamma) < \infty \right\},$$

and

$$C = \lim_{\gamma \searrow 0} \frac{\gamma(k-1)}{\gamma + \nu} \sum_{n=1}^{\infty} k^{n-1} \mathcal{L}_n(\gamma + \nu).$$

The proof is the same as the proof of Theorem 2.1.

Without any further assumption, nothing more can be obtained than the Cesàro convergence (2.5) as the example of constant lifetimes shows. To conclude that $\lim A_\nu(t) = C$ in the i.i.d. case, Bellman and Harris [2] used Ikehara’s Tauberian theorem. For the BMC case, we will need not only a limit, but also an exponential speed of convergence. Although we have not found Lemma 2.2 in the literature, it cannot be considered as new; it is closely related to a large corpus of results going back to Haar, Wiener, and Ikehara; see [23].

Lemma 2.2. *Let f be a function and \tilde{f} its Laplace transform. Suppose that there exist two positive reals δ and ε such that:*

1. \tilde{f} is analytic in $\{z = x + iy, |x| < \delta + \varepsilon\} \setminus \{0\}$,
2. \tilde{f} has a simple pole at 0 with residue C ,
3. $\int_{-\infty}^{\infty} |\tilde{f}(\delta + iy)| dy < \infty$,
4. $\lim_{y \rightarrow \pm\infty} \tilde{f}(x + iy) = 0$, uniformly in $x \in [-\delta, \delta]$,
5. $\psi := \int_{-\infty}^{\infty} |\tilde{f}(-\delta + iy)| dy < \infty$.

Then for all $t > 0$,

$$|f(t) - C| \leq \frac{\psi}{2\pi} e^{-\delta t}.$$

Proof. By the inversion equation,

$$f(t) = \frac{1}{2\pi i} \lim_{\beta \rightarrow \infty} \int_{\delta - i\beta}^{\delta + i\beta} \tilde{f}(\gamma) e^{\gamma t} d\gamma.$$

Let \mathcal{C} be the closed rectangular contour linking the points $\delta - i\beta$, $\delta + i\beta$, $-\delta + i\beta$, and $-\delta - i\beta$. This contour encloses the simple pole at $\gamma = 0$. By the residue theorem,

$$\int_{\mathcal{C}} \tilde{f}(\gamma) e^{\gamma t} d\gamma = 2\pi i \operatorname{res}_{\gamma=0}(\tilde{f}(\gamma) e^{\gamma t}),$$

where

$$\operatorname{res}_{\gamma=0}(\tilde{f}(\gamma) e^{\gamma t}) = \lim_{\gamma \rightarrow 0} \gamma \tilde{f}(\gamma) e^{\gamma t} = C.$$

Consequently,

$$\frac{1}{2\pi i} \int_{\delta - i\beta}^{\delta + i\beta} \tilde{f}(\gamma) e^{\gamma t} d\gamma - C = I_1 + I_2 + I_3$$

with

$$\begin{aligned} I_1 &= -\frac{1}{2\pi i} \int_{\delta + i\beta}^{-\delta + i\beta} \tilde{f}(\gamma) e^{\gamma t} d\gamma, \\ I_2 &= -\frac{1}{2\pi i} \int_{-\delta + i\beta}^{-\delta - i\beta} \tilde{f}(\gamma) e^{\gamma t} d\gamma, \\ I_3 &= -\frac{1}{2\pi i} \int_{-\delta - i\beta}^{\delta - i\beta} \tilde{f}(\gamma) e^{\gamma t} d\gamma. \end{aligned}$$

By condition 4, I_1 and I_3 tend to 0 as β tends to ∞ . Therefore,

$$\begin{aligned} |f(t) - C| &\leq \frac{1}{2\pi} \left| \int_{-\infty}^{+\infty} \tilde{f}(-\delta + iy) e^{(-\delta + iy)t} dy \right| \\ &\leq \int_{-\infty}^{+\infty} |\tilde{f}(-\delta + iy)| e^{-\delta t} dy \\ &= \frac{\psi}{2\pi} e^{-\delta t}, \end{aligned}$$

by condition 5.

3. Bifurcating Markov chains

Bifurcating Markov chains were first considered in the 1970s, beginning with Spitzer [33] in the binary valued case (see [3] for further references). Bifurcating Markov chains were studied as cell lineage models by Guyon [14], see also [4]. As for ordinary Markov chains, the probability distribution of a BMC is determined by an initial measure and a transition kernel. Here is the definition, adapted to our case (as usual, \mathcal{B} denotes the Borel σ -algebra).

Definition 3.1. A transition kernel P is a mapping defined on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^+)$ such that:

- for all $B \in \mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^+)$, $t \mapsto P(t, B)$ is $\mathcal{B}(\mathbb{R}^+)$ -measurable,
- for all $t \in \mathbb{R}^+$, $B \mapsto P(t, B)$ is a probability measure on $\mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^+)$.

Definition 3.2. Let μ be a probability measure on \mathbb{R}^+ , and P be a transition kernel. A BMC $(T_v)_{v \in \mathbb{T}}$ with initial measure μ and transition kernel P is inductively defined as follows:

- T_0 has distribution μ ,
- for $n \geq 1$, $(T_w)_{w \in \Gamma_{n+1}}$ and $(T_u)_{u \in \Gamma_0 \cup \dots \cup \Gamma_{n-1}}$ are independent conditionally upon $(T_v)_{v \in \Gamma_n}$,
- for all $n \geq 0$, the conditional distribution of $(T_w)_{w \in \Gamma_{n+1}}$ knowing $(T_v)_{v \in \Gamma_n}$ is defined for $(B_v)_{v \in \Gamma_n} \in \mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^+)$ by

$$\mathbb{P}\{\text{for all } v \in \Gamma_n, (T_{v0}, T_{v1}) \in B_v \mid \text{for all } v \in \Gamma_n, T_v = t_v\} = \prod_{v \in \Gamma_n} P(t_v, B_v).$$

In other words, given the lifetimes of mothers in generation n , the lifetimes of couples of daughters in generation $n + 1$ are drawn independently, each according to the transition kernel P . We set P to be symmetric: for all $t \in \mathbb{R}^+$, for all $B \in \mathcal{B}(\mathbb{R}^+)$,

$$P_0(t, B) := P(t, B \times \mathbb{R}^+) = P(t, \mathbb{R}^+ \times B) =: P_1(t, B). \tag{3.1}$$

This assumption is not necessary [14]. As an example, consider the case where all lifetimes are independent, and for each v , T_{v0} is distributed as T_{00} , T_{v1} as T_{01} with $\mathbb{E}\{e^{-\nu T_{00}}\} > \mathbb{E}\{e^{-\nu T_{01}}\}$ for all positive ν : exponential growth of N_t holds, even though the kernel is not symmetric.

The initial measure μ is supposed to be invariant for both marginal kernels such that

$$\int_{\mathbb{R}^+} P_0(u, B) d\mu(u) = \mu(B) \quad \text{for all } B \in \mathcal{B}(\mathbb{R}^+). \tag{3.2}$$

Symmetry and invariance imply that the distribution of $(T_v)_{v \in \mathbb{T}}$ is automorphism invariant in the sense of [27]. In particular it is birth- and fork-stationary, in the sense of Definitions 2.2 and 2.3.

Remark 3.1. On the k -ary tree, a transition kernel is a mapping defined on $\mathbb{R}^+ \times \mathcal{B}((\mathbb{R}^+)^k)$. The generalization of Definition 3.2 is straightforward: knowing the lifetimes of mothers in generation n , the lifetimes for all k -tuples of daughters are drawn independently according to the transition kernel.

Let $(N_t)_{t \geq 0}$ be the branching process associated to $(T_v)_{v \in \mathbb{T}}$ (Definition 2.1). Our goal is to prove the extension of Theorem 1.1, i.e. the almost sure convergence of $e^{-\nu t} N_t$. Asymptotics on the Laplace transform of S_n will be needed; the expressions of ν and C given in Theorem 2.1 suggest using multiplicative ergodicity for the sums of lifetimes S_n [22]. In order to enhance the similarity with the i.i.d. case we chose to express multiplicative ergodicity in a slightly different manner.

Definition 3.3. The sums S_n are said to be multiplicatively ergodic if for all $u \in \mathbb{R}^+$, $\gamma \in \mathbb{R}^+$, $n \in \mathbb{N}^*$,

$$\mathcal{L}_n(\gamma, u) = \alpha(\gamma, u)L^n(\gamma) + r_n(\gamma, u), \tag{3.3}$$

where α , L , and r_n are such that:

- (i) the equation $2L(\gamma) = 1$ has a unique solution denoted by ν ,
- (ii) the mapping L is derivable at ν and $L'(\nu) < 0$,
- (iii) the series $\sum_n 2^n r_n(\gamma, u)$ converges uniformly in γ in a neighbourhood of ν , uniformly in u ,
- (iv) the mappings $u \mapsto \alpha(\gamma, u)$ and $u \mapsto r_n(\gamma, u)$ are μ -integrable, uniformly in the other variables,
- (v) the mapping $(y, z) \mapsto \alpha(\nu, y)\alpha(\nu, z)$ is $P(x, (y, z))$ -integrable, uniformly in x .

Observe that under Definition 3.3, for all u the sum $\sum_n 2^n \mathcal{L}_n(\gamma, u)$ converges for $\gamma > \nu$ and diverges for $\gamma \leq \nu$. Therefore, the same holds for $\sum_n 2^n \mathcal{L}_n(\gamma)$, and the definition of ν by $2L(\nu) = 1$ is coherent with (2.3).

Theorem 4.1 of Kontoyiannis and Meyn [22, p. 325] relates multiplicative ergodicity to geometric ergodicity. More precise analyticity conditions will be needed for the following function:

$$B_\nu(t, u) = e^{-\nu t} \sum_{n=1}^{\infty} 2^{n-1} \mathbb{P}\{S_n \leq t \mid T_0 = u\}. \tag{3.4}$$

Under (3.3), its Laplace transform is

$$\tilde{B}_\nu(\gamma, u) = \frac{\alpha(\nu + \gamma, u)}{\nu + \gamma} \frac{L(\nu + \gamma)}{1 - 2L(\nu + \gamma)} + \frac{1}{\nu + \gamma} \sum_{n=1}^{\infty} 2^{n-1} r_n(\nu + \gamma, u).$$

Our hypotheses will be the following.

- (C₁) For all $u > 0$, $B_\nu(t, u)$ and $\tilde{B}_\nu(\gamma, u)$ satisfy the hypotheses of Lemma 2.2, for some $\delta > 0$ (not depending on u), and

$$C(u) = \lim_{\gamma \rightarrow 0} \gamma \tilde{B}_\nu(\gamma, u) = -\frac{\alpha(\nu, u)}{4\nu L'(\nu)}.$$

Let

$$\psi(u) = \int_{-\infty}^{\infty} |\tilde{B}_\nu(-\delta + iy, u)| dy.$$

- (C₂) The mapping $(y, z) \mapsto \alpha(\nu, y)\psi(z)$ is $P(x, (y, z))$ -integrable, uniformly in x .

Admittedly, (C₁) and (C₂) are not easy to verify, unless an explicit expression of $\mathcal{L}_n(\gamma, u)$ is available. This will be the case for the model we present in the next section.

Using condition (iv) of Definition 3.3, let

$$\alpha(\gamma) = \int_{\mathbb{R}^+} \alpha(\gamma, u) d\mu(u) \quad \text{and} \quad r_n(\gamma) = \int_{\mathbb{R}^+} r_n(\gamma, u) d\mu(u).$$

Then

$$\mathcal{L}_n(\gamma) = \alpha(\gamma)L^n(\gamma) + r_n(\gamma). \tag{3.5}$$

The proportionality constant C naturally relates to $\alpha(\nu)$ and $L(\nu)$.

Theorem 3.1. *Assume that symmetry (3.1), invariance (3.2), and multiplicative ergodicity (3.5) hold, together with (\mathcal{C}_1) and (\mathcal{C}_2) . Then*

$$\lim_{n \rightarrow \infty} e^{-\nu t} N_t = W \quad \text{a.s.,}$$

where

- W is a random variable with expectation C and finite variance,
- the growth rate ν is such that
- the proportionality constant C is

$$2L(\nu) = 1, \tag{3.6}$$

$$C = -\frac{\alpha(\nu)}{4\nu L'(\nu)}. \tag{3.7}$$

Remark 3.2. On the k -ary tree, (3.6) and (3.7) become

$$kL(\nu) = 1 \quad \text{and} \quad C = -\frac{(k-1)\alpha(\nu)}{k^2\nu L'(\nu)}.$$

As mentioned in the introduction, we have tried to provide the weakest possible conditions to ensure convergence in L^2 on the one hand (Proposition 6.1), almost sure convergence on the other hand (Proposition 6.2). The proof of Theorem 3.1 will be completed in Section 7 by checking the hypotheses of Propositions 6.1 and 6.2 for the BMC model.

4. An explicit model

In this section an explicit example for the result of the previous section is constructed: a BMC with prescribed invariant measure μ , for which symmetry (3.1), invariance (3.2), and multiplicative ergodicity (3.5) hold. The model depends on an identifiable set of parameters, potentially adjustable to observed data.

The construction of stationary processes with prescribed marginal distributions has been the object of many studies, see Pitt *et al.* [29] and the references therein. We will follow a simple approach, first constructing a bifurcating autoregressive process, then transforming it to obtain the desired marginals. Bifurcating autoregressive (BAR) processes were introduced by Cowan and Staudte [5] precisely as cell lineage models. They have been extensively studied since, and the problem of parameter estimation has recently received a lot of attention [4], [8], [9], [14]. Our model is similar to that of [14]. The construction begins with a family of i.i.d. random variables $(\varepsilon_\nu)_{\nu \in \mathbb{T}}$, each with standard Gaussian $\mathcal{N}(0, 1)$ distribution. Let ρ_m and ρ_s be two reals in $(-1; 1)$; they will be the mother and sister correlations of our BAR process. It is defined inductively by $X_0 = \varepsilon_0$, and for all $\nu \in \mathbb{T}$,

$$X_{\nu 0} = \rho_m X_\nu + \sqrt{1 - \rho_m^2} \varepsilon_{\nu 0}, \tag{4.1a}$$

$$X_{\nu 1} = \rho_m X_\nu + \sqrt{1 - \rho_m^2} (\rho_s \varepsilon_{\nu 0} + \sqrt{1 - \rho_s^2} \varepsilon_{\nu 1}). \tag{4.1b}$$

By construction, $(X_\nu)_{\nu \in \mathbb{T}}$ is both a BMC on \mathbb{T} and a Gaussian process. It is symmetric in the sense of (3.1) and the standard Gaussian distribution is the invariant distribution of the marginal kernel in the sense of (3.2). Let f denote the composition of the quantile function of the desired

distribution μ on \mathbb{R}^+ by the distribution function of the $\mathcal{N}(0, 1)$. If X follows the $\mathcal{N}(0, 1)$ then $T = f(X)$ has distribution μ . For all $v \in \mathbb{T}$, let $T_v = f(X_v): (T_v)_{v \in \mathbb{T}}$ is a BMC on \mathbb{T} for which (3.1) and (3.2) hold. Observe, moreover, that (X_v) converges at geometric speed along the rays of \mathbb{T} , hence multiplicative ergodicity holds for the birth dates of (T_v) by Theorem 4.1 of [22, p. 325]. Of course, the mother and sister correlations are not ρ_m and ρ_s anymore. But they can be computed in terms of f , ρ_m , and ρ_s , and so the model can be adjusted to fit not only the observed distribution of lifetimes but also estimated correlations.

As remarked as early as 1932 by Rahn [31], actual lifetime data show a unimodal right-skewed shape. They have been fitted by many types of distributions such as gamma and lognormal; see [18] and the references therein. The difficulty is to exhibit a realistic example where the hypotheses of Theorem 3.1 hold with explicitly computable α and L . We propose to transform the standard Gaussian variables X_v of the BAR process defined by (4.1) by the following function, depending on three parameters:

$$f(x) = a + b(x + c)^2.$$

If X is normally distributed then $(X + c)^2$ has a noncentral chi-squared distribution and the shape can be adjusted by c ; using the location and scale parameters a and b , X can be fitted to actual lifetime data. The Laplace transforms of quadratic forms of autoregressive processes can be explicitly computed using a technique due to Kleptsyna *et al.* [20]. The expression of the Laplace transform $\mathcal{L}_n(\gamma, u)$ has been stated in [21]. From there, it is easy to check that the hypotheses of Theorem 3.1 are satisfied. Actually, it can be checked that $\alpha(\gamma, u)$ and $r_n(\gamma, u)$ are uniformly bounded in u and γ over $(\mathbb{R}^+)^2$, which considerably simplifies conditions (iv) and (v) of Definition 3.3 as well as conditions (\mathcal{C}_1) and (\mathcal{C}_2) . Details being omitted, only the explicit expressions of $L(\gamma)$ and $\alpha(\gamma)$ will be given.

Proposition 4.1. *Let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence of i.i.d. random variables, with common distribution $\mathcal{N}(0, 1)$. Let $\rho = \rho_m \in (-1, 1)$. Let $(X_n)_{n \in \mathbb{N}}$ be the stationary autoregressive chain defined by $X_0 = \varepsilon_0$ and for $n \geq 0$,*

$$X_{n+1} = \rho X_n + \sqrt{1 - \rho^2} \varepsilon_{n+1}.$$

Let

$$S_n = \sum_{k=0}^n f(X_k) = \sum_{k=0}^n (a + b(X_k + c)^2).$$

The sums S_n are multiplicatively ergodic in the sense of Definition 3.3. Define

$$\begin{aligned} \gamma_1 &= 2\gamma b(1 - \rho^2), & \gamma_2 &= \frac{1 - \rho}{\gamma_1 + (1 - \rho)^2}, & B &= -2\rho\gamma_2^2, & C &= \frac{2\rho\gamma_1}{1 - \rho^2}\gamma_2^2, \\ \lambda_{\pm} &= \frac{\gamma_1 + 1 + \rho^2 \pm \sqrt{(\gamma_1 + (\rho + 1)^2)(\gamma_1 + (\rho - 1)^2)}}{2}, \\ \beta_+ &= \frac{1 - \lambda_- + \gamma_1/(1 - \rho^2)}{\lambda_+ - \lambda_-}, & \beta_- &= \frac{\lambda_+ - 1 + \gamma_1/(1 - \rho^2)}{\lambda_+ - \lambda_-}, & \pi_0 &= \beta_+\lambda_+ + \beta_-\lambda_-. \end{aligned}$$

Then

$$L(\gamma) = e^{-a\gamma} (\lambda_+)^{-1/2} \exp\left(-\frac{\gamma bc^2(1 - \rho)}{2\gamma b(1 + \rho) + (1 - \rho)}\right), \tag{4.2}$$

and

$$\alpha(\gamma) = (\beta_+\lambda_+)^{-1/2} \exp\left(-\frac{c^2\gamma_1}{2(1-\rho^2)}\left(\frac{1}{\pi_0} + B\left(\frac{\rho}{\pi_0} - \frac{\rho}{\lambda_+}\right) + C\left(\frac{\rho}{\pi_0}\right)\right)\right). \tag{4.3}$$

The growth rate ν and the proportionality constant C can be derived from (4.2) and (4.3), through (3.6) and (3.7). The growth rate ν is the solution of $2L(\nu) = 1$. It has no general explicit expression, but it can be numerically computed. In the particular case $a = c = 0$, $\nu = \nu_0$ is found to be

$$\nu_0 = \frac{3}{2b} \frac{1 - \rho^2/4}{1 - \rho^2}.$$

In the general case, it can be checked that $\nu \leq \nu_0$.

5. Correlations and growth rate

As mentioned in the introduction, the influence of lifetime correlations on the exponential growth of the colony was discussed long ago [6], [16], [30]. That sister correlation does not change the exponential growth rate was noted by all the authors, and is confirmed by Theorem 2.1. The influence of mother correlation is examined here.

The hypotheses in this section are those of Theorem 2.1: birth-stationarity and definition of ν and C by (2.3) and (2.4). A general comparison result will be first obtained under association hypotheses. Recall that a sequence of random variables $(X_n)_{n \geq 0}$ is *associated* if for all n , the vector $\mathbf{X}^{(n)} = (X_1, X_2, \dots, X_n)$ satisfies the following condition: for any coordinatewise bounded and nondecreasing functions f, g on \mathbb{R}^n , $\text{cov}(f(\mathbf{X}^{(n)}), g(\mathbf{X}^{(n)})) \geq 0$. We refer the reader to [11] for more about this notion. The sequence $(X_n)_{n \geq 1}$ is *negatively associated* if for any coordinatewise bounded and nondecreasing functions f, g defined respectively on $\mathbb{R}^{|I|}$, $\mathbb{R}^{|J|}$ where I and J are disjoint subsets of \mathbb{N} , $\text{cov}(f((X_i)_{i \in I}), g((X_i)_{i \in J})) \leq 0$. This definition was introduced by Joag-Dev and Proschan [17].

Proposition 5.1. *Suppose that the sequence $(T_{0^n})_{n \geq 1}$ is associated (respectively, negatively associated) and that the hypotheses of Theorem 2.1 are satisfied. Let $(T_{0^n}^*)_{n \geq 1}$ be a sequence of independent random variables also satisfying the hypotheses of Theorem 2.1, and such that $T_{0^n}^*$ and T_{0^n} have the same distribution. Let ν, ν^* and C, C^* be the respective growth rates and proportionality constants corresponding to $(T_{0^n})_{n \geq 1}$ and $(T_{0^n}^*)_{n \geq 1}$ through (2.3) and (2.4). Then $\nu^* \leq \nu$ and $C \leq C^*$ (respectively, $\nu \leq \nu^*$ and $C^* \leq C$).*

Proof. We provide the proof only for the case of association, the case of negative association is symmetric. Let $S_n^* = T_{0^1}^* + T_{0^2}^* + \dots + T_{0^n}^*$. If $(T_{0^n})_{n \geq 1}$ is an associated sequence, then for any positive real γ ,

$$\mathbb{E}\{e^{-\gamma S_n^*}\} \leq \mathbb{E}\{e^{-\gamma S_n}\}. \tag{5.1}$$

It follows immediately that $\nu^* \leq \nu$, by (2.3). The inequality $C \leq C^*$ then follows from the fact that S_n stochastically dominates S_n^* : for all $t \geq 0$,

$$\mathbb{P}\{S_n \leq t\} \leq \mathbb{P}\{S_n^* \leq t\}. \tag{5.2}$$

Stochastic comparison results such as (5.1) and (5.2) are well-known decoupling inequalities, and we will omit their proofs, see de la Peña and Lai [7, p. 118] and Shao [32].

From Proposition 5.1, it is indicated that for a fixed marginal distribution of lifetimes, the growth rate ν should increase as the mother correlation increases from 0 to 1. This is indeed

what can be observed on the explicit model from Section 4. In that model, ν is defined by $2L(\nu) = 1$ where L is given by (4.2). In (4.2) ρ is the correlation between successive steps of the BAR process, which differs from the correlation between the lifetimes of a mother and its daughter. The latter will be defined by ϱ . The expression of ϱ as a function of the parameters a, b, c , and ρ is easily calculated. It depends only on c and ρ . Thus,

$$\varrho = \text{cor}(T_v, T_{v0}) = \frac{\rho^2 + 2c^2\rho}{1 + 2c^2}.$$

As ρ increases from 0 to 1, so does ϱ . As $\varrho \rightarrow +1$,

$$\nu \rightarrow \begin{cases} +\infty & \text{if } a = 0, \\ \frac{\log(2)}{a} & \text{otherwise.} \end{cases}$$

The limit value $\log(2)/a$ is the growth rate that would be achieved if all lifetimes were equal to a , which is the minimal value that a lifetime can take in the model.

6. Convergence in quadratic mean and almost sure

The conditions for the convergence of $e^{-\nu t} N_t$ are stated in this section. Under the hypothesis of fork-stationarity of Definition 2.3, Proposition 6.1 below provides a general condition under which $e^{-\nu t} N_t$ converges in L^2 .

Proposition 6.1. *Let $(T_v)_{v \in \mathbb{T}}$ be a fork-stationary bifurcating process. Assuming that the hypotheses of Theorem 2.1 hold, let ν and C be defined by (2.3) and (2.4). For all $t, \tau \geq 0$, let*

$$\begin{aligned} \Sigma_1(t) &= \sum_{n=0}^{\infty} (n + 1)2^n \mathbb{P}\{S_n \leq t\}, \\ \Sigma_2(t, \tau) &= \sum_{n=0}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} 2^{n+i+j} \mathbb{P}\{S_{n,i}^{(0)} \leq t, S_{n,j}^{(1)} \leq t + \tau\}. \end{aligned}$$

Assume that for all $t, \tau \geq 0$, $\Sigma_1(t)$ and $\Sigma_2(t, \tau)$ are finite, that the following limits exist, and the second one does not depend on τ . Thus,

$$\lim_{t \rightarrow +\infty} e^{-2\nu t} \Sigma_1(t) = 0, \tag{6.1}$$

$$\lim_{t \rightarrow +\infty} e^{-\nu(2t+\tau)} \Sigma_2(t, \tau) = C_2 < +\infty. \tag{6.2}$$

As $t \rightarrow \infty$, $e^{-\nu t} N_t$ converges in quadratic mean to a random variable W with expectation C .

Observe that since L^2 -convergence implies L^1 -convergence,

$$\lim_{t \rightarrow \infty} \mathbb{E}\{e^{-\nu t} N_t\} = \mathbb{E}\{W\} = C > 0.$$

Proof of Proposition 6.1. We first express the product $N_t N_{t+\tau}$ as a function of birth dates. For this, recall (2.2),

$$N_t = \frac{1}{2} + \frac{1}{2} \sum_{v \in \mathbb{T}} \mathbf{1}_{\{S_v \leq t\}}.$$

Hence,

$$(2N_t - 1)(2N_{t+\tau} - 1) = \sum_{(v,w) \in \mathbb{T}^2} \mathbf{1}_{\{S_v \leq t\}} \mathbf{1}_{\{S_w \leq t+\tau\}}.$$

For any couple $(v, w) \in \mathbb{T}^2$, one and only one of the following three cases occurs:

1. $w \preceq v$, in which case $\mathbf{1}_{\{S_v \leq t\}} \mathbf{1}_{\{S_w \leq t+\tau\}} = \mathbf{1}_{\{S_v \leq t\}}$,
2. $|v \wedge w| < \min\{|v|, |w|\}$,
3. $v \preceq w$ and $v \neq w$.

Decomposing the sum over the three cases, taking expectations on both sides, and using fork-stationarity, we obtain

$$\mathbb{E}\{(2N_t - 1)(2N_{t+\tau} - 1)\} = \Sigma_1(t) + \Sigma_2(t, \tau) + \Sigma_3(t, \tau), \tag{6.3}$$

with

$$\begin{aligned} \Sigma_1(t) &= \sum_{n=0}^{\infty} (n+1)2^n \mathbb{P}\{S_n \leq t\}, \\ \Sigma_2(t, \tau) &= \sum_{n=0}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} 2^{n+i+j} \mathbb{P}\{S_{n,i}^{(0)} \leq t, S_{n,j}^{(1)} \leq t + \tau\}, \\ \Sigma_3(t, \tau) &= \sum_{n=0}^{\infty} \sum_{j=1}^{\infty} 2^{n+j} \mathbb{P}\{S_n \leq t, S_{n+j} \leq t + \tau\}. \end{aligned}$$

From the hypotheses for all $t, \tau \geq 0$, $\Sigma_1(t)$ and $\Sigma_2(t, \tau)$ are finite. We remark that

$$\Sigma_3(t, \tau) \leq \sum_{n=0}^{\infty} \sum_{j=1}^{\infty} 2^{n+j} \mathbb{P}\{S_{n+j} \leq t + \tau\} = \sum_{m=1}^{\infty} m 2^m \mathbb{P}\{S_m \leq t + \tau\} \leq \Sigma_1(t + \tau).$$

Therefore, $\Sigma_3(t, \tau)$ is also finite. In particular, $\mathbb{E}\{N_t^2\} < \infty$ for all t . Since $\Sigma_3(t, \tau) \leq \Sigma_1(t + \tau)$,

$$\lim_{t \rightarrow \infty} e^{-v(2t+\tau)} \Sigma_3(t, \tau) = 0. \tag{6.4}$$

Collecting (6.1)–(6.4) and using the fact that

$$\lim_{t \rightarrow \infty} \mathbb{E}\{e^{-v(2t+\tau)} N_t\} = \lim_{t \rightarrow \infty} \mathbb{E}\{e^{-v(2t+\tau)} N_{t+\tau}\} = 0,$$

we obtain

$$\lim_{t \rightarrow \infty} \mathbb{E}\{e^{-v(2t+\tau)} N_t N_{t+\tau}\} = \frac{C_2}{4}.$$

Hence,

$$\begin{aligned} &\lim_{t \rightarrow \infty} \mathbb{E}\{(e^{-vt} N_t - e^{-v(t+\tau)} N_{t+\tau})^2\} \\ &= \lim_{t \rightarrow \infty} \mathbb{E}\{e^{-2vt} N_t^2 - 2e^{-v(2t+\tau)} N_t N_{t+\tau} + e^{-2v(t+\tau)} N_{t+\tau}^2\} \\ &= \frac{C_2}{4} - 2 \frac{C_2}{4} + \frac{C_2}{4} \\ &= 0, \end{aligned}$$

which completes the proof.

A reinforcement of (6.1) and (6.2) ensures almost sure convergence.

Proposition 6.2. *Under the hypotheses of Proposition 6.1, assume that W is a.s. positive and that*

$$\int_0^\infty e^{-2\nu t} \Sigma_1(t) dt < \infty, \tag{6.5}$$

$$\int_0^\infty |e^{-2\nu t} \Sigma_2(t, 0) - C_2| dt < \infty. \tag{6.6}$$

Then as $t \rightarrow \infty$, $e^{-\nu t} N_t$ converges a.s. to W .

Proof. From the proof of Proposition 6.1, the additional hypothesis yields

$$\int_0^\infty \mathbb{E}\{(e^{-\nu t} N_t - W)^2\} dt < \infty.$$

Almost sure convergence is deduced exactly as in the proof of Theorem 21.1 of [15, p. 148]. That W is a.s. positive cannot be obtained without stronger hypotheses. It will be proved for the BMC model in Section 7.

Remark 6.1. The only change for the k -ary tree consists of replacing 2 by k in the definitions of Σ_1 and Σ_2 from Proposition 6.1.

7. Proof of Theorem 3.1

As already remarked, symmetry (3.1) and invariance (3.2) imply birth- and fork-stationarity. We also observed that the solution of $2L(\nu) = 1$ is such that

$$\nu = \inf \left\{ \gamma > 0, \sum_{n=1}^\infty 2^n \mathcal{L}_n(\gamma) < \infty \right\}.$$

The main ingredient in the proof consists of applying Lemma 2.2 to $B_\nu(t, u)$ defined by (3.4), due to condition (\mathcal{C}_1) . This yields

$$\left| e^{-\nu t} \sum_{n=1}^\infty 2^{n-1} \mathbb{P}\{S_n \leq t \mid T_0 = u\} + \frac{\alpha(\nu, u)}{4\nu L'(\nu)} \right| \leq \frac{\psi(u)}{2\pi} e^{-\delta t}. \tag{7.1}$$

Recall from (2.6) and (3.4) that

$$\mathbb{E}\{N_t\} = 1 + \sum_{n=1}^\infty 2^{n-1} \mathbb{P}\{S_n \leq t\} = 1 + A_\nu(t) = 1 + \int_{\mathbb{R}^+} B_\nu(t, u) d\mu(u).$$

Integrating against μ (condition (iv) of Definition 3.3), we obtain

$$\lim_{t \rightarrow \infty} e^{-\nu t} \mathbb{E}\{N_t\} = \int_{\mathbb{R}^+} C(u) d\mu(u) = -\frac{\alpha(\nu)}{4\nu L'(\nu)} = C.$$

Now consider

$$\Sigma_1(t) = \sum_{n=0}^\infty (n+1) 2^n \mathbb{P}\{S_n \leq t\}.$$

The series $\sum (n + 1)2^n L^n(\gamma)$ converges for $\gamma > \nu$ and diverges for $\gamma \leq \nu$. Choose $\gamma > \nu$. By Markov's inequality

$$\Sigma_1(t) = \sum_{n=0}^{\infty} (n + 1)2^n \mathbb{P}\{S_n \leq t\} \leq e^{\gamma t} \sum_{n=0}^{\infty} (n + 1)2^n \mathcal{L}_n(\gamma).$$

Therefore, $\Sigma_1(t)$ is finite for all t . Take γ such that $\nu < \gamma < 2\nu$, thus,

$$e^{-2\nu t} \Sigma_1(t) \leq e^{(\gamma-2\nu)t} \sum_{n=0}^{\infty} (n + 1)2^n \mathcal{L}_n(\gamma).$$

There exists a constant K_1 such that for all $t \geq 0$, $e^{-2\nu t} \Sigma_1(t) \leq K_1 e^{(\gamma-2\nu)t}$, hence, (6.1) and (6.5) hold.

The convergence of $e^{-\nu(2t+\tau)} \Sigma_2(t, \tau)$ remains to be proved. Consider

$$e^{-\nu(2t+\tau)} \Sigma_2(t, \tau) = e^{-\nu(2t+\tau)} \sum_{n=0}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} 2^{n+i+j} \mathbb{P}\{S_{n,i}^{(0)} \leq t, S_{n,j}^{(1)} \leq t + \tau\}.$$

We use the Markov property after conditioning on the following event:

$$B_n := \{S_{n-1} = u, T_0^n = x, T_0^{n+1} = y, T_0^n = z\}.$$

By Definition 3.2,

$$\begin{aligned} &\mathbb{P}\{S_{n,i}^{(0)} \leq t, S_{n,j}^{(1)} \leq t + \tau \mid B_n\} \\ &= \mathbb{P}\{S_i \leq t - u - x \mid T_0 = y\} \mathbb{P}\{S_j \leq t + \tau - u - x \mid T_0 = z\}. \end{aligned}$$

Therefore,

$$\begin{aligned} &e^{-\nu(2t+\tau)} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} 2^{i+j} \mathbb{P}\{S_{n,i}^{(0)} \leq t, S_{n,j}^{(1)} \leq t + \tau \mid B_n\} \\ &= e^{-2\nu(u+x)} \left(e^{-\nu(t-u-x)} \sum_{i=1}^{\infty} 2^i \mathbb{P}\{S_i \leq t - u - x \mid T_0 = y\} \right) \\ &\quad \times \left(e^{-\nu(t+\tau-u-x)} \sum_{j=1}^{\infty} 2^j \mathbb{P}\{S_j \leq t + \tau - u - x \mid T_0 = z\} \right). \end{aligned}$$

By (7.1), we obtain

$$\left| \left(e^{-\nu(t-u-x)} \sum_{i=1}^{\infty} 2^i \mathbb{P}\{S_i \leq t - u - x \mid T_0 = y\} \right) + \frac{\alpha(\nu, y)}{2\nu L'(\nu)} \right| \leq \frac{\psi(y)}{\pi} e^{-\delta(t-u-x)},$$

and

$$\left| \left(e^{-\nu(t+\tau-u-x)} \sum_{j=1}^{\infty} 2^j \mathbb{P}\{S_j \leq t + \tau - u - x \mid T_0 = z\} \right) + \frac{\alpha(\nu, z)}{2\nu L'(\nu)} \right| \leq \frac{\psi(z)}{\pi} e^{-\delta(t+\tau-u-x)}.$$

For large enough t ,

$$\begin{aligned} & \left| \left(e^{-v(t-u-x)} \sum_{i=1}^{\infty} 2^i \mathbb{P}\{S_i \leq t - u - x \mid T_0 = y\} \right) \right. \\ & \times \left. \left(e^{-v(t+\tau-u-x)} \sum_{j=1}^{\infty} 2^j \mathbb{P}\{S_j \leq t + \tau - u - x \mid T_0 = z\} \right) - \frac{\alpha(v, y)\alpha(v, z)}{(2vL'(v))^2} \right| \\ & \leq -\frac{2}{\pi vL'(v)} (\alpha(v, z)\psi(y) + \alpha(v, y)\psi(z))e^{-\delta(t-u-x)}. \end{aligned}$$

Denoting by Q_n the joint distribution of (S_{n-1}, T_0^n) , define

$$C_2 = \frac{1}{(2vL'(v))^2} \sum_{n=0}^{\infty} 2^n \int_{u,x} e^{-2v(u+x)} \left(\int_{y,z} \alpha(v, y)\alpha(v, z) dP(x, (y, z)) \right) dQ_n(u, x).$$

By condition (v) of Definition 3.3, there exists K_2 such that for all x ,

$$\int_{y,z} \alpha(v, y)\alpha(v, z) dP(x, (y, z)) \leq K_2.$$

Hence,

$$\begin{aligned} C_2 & \leq \frac{K_2}{(2vL'(v))^2} \sum_{n=0}^{\infty} 2^n \int_{u,x} e^{-2v(u+x)} dQ_n(u, x) \\ & = \frac{K_2}{(2vL'(v))^2} \sum_{n=0}^{\infty} 2^n \mathbb{E}\{e^{-2vS_{n-1}+T_0^n}\} \\ & = \frac{K_2}{(2vL'(v))^2} \sum_{n=0}^{\infty} 2^n \mathcal{L}_n(2v) \\ & < \infty. \end{aligned}$$

We obtain

$$\begin{aligned} & |e^{-v(2t+\tau)} \Sigma_2(t, \tau) - C_2| \\ & \leq -\frac{2}{\pi vL'(v)} e^{-\delta t} \sum_{n=0}^{\infty} 2^n \int_{u,x} e^{-(2v-\delta)(u+x)} \left(\int_{y,z} (\alpha(v, z)\psi(y) + \alpha(v, y)\psi(z)) dP \right. \\ & \qquad \qquad \qquad \left. \times (x, (y, z)) \right) dQ_n(u, x). \end{aligned}$$

From condition (C₂) and symmetry (3.1), there exists K_3 such that for all x ,

$$-\frac{2}{\pi vL'(v)} \left(\int_{y,z} (\alpha(v, z)\psi(y) + \alpha(v, y)\psi(z)) dP(x, (y, z)) \right) \leq K_3.$$

Therefore,

$$\begin{aligned}
 |e^{-\nu(2t+\tau)} \Sigma_2(t, \tau) - C_2| &\leq K_3 e^{-\delta t} \sum_{n=0}^{\infty} 2^n \int_{u,x} e^{-(2\nu-\delta)(u+x)} dQ_n(u, x) \\
 &= K_3 e^{-\delta t} \sum_{n=0}^{\infty} 2^n \mathbb{E}\{e^{-(2\nu-\delta)(S_{n-1}+T_0^n)}\} \\
 &= K_3 e^{-\delta t} \sum_{n=0}^{\infty} 2^n \mathcal{L}_n(2\nu - \delta).
 \end{aligned}$$

For $\delta < \nu$, the series converges, hence (6.6) holds. What has been proved implies that $\Sigma_2(t, \tau)$ is finite for all τ and for t large enough. But $\Sigma_2(t, \tau)$ is a nondecreasing function of t , hence it is finite for all t and τ .

Only one point remains to be proved, that the limit of $e^{-\nu t} N_t$ is a.s. positive. Assume that $T_0 = u$ and take $t > u$. The cells alive at time t descend either from 00 or from 01. Therefore,

$$N_t = N_{t-u}^{(0)} + N_{t-u}^{(1)},$$

where $(N_s^{(0)})_{s \geq 0}$ and $(N_s^{(1)})_{s \geq 0}$ have the same distribution as $(N_s)_{s \geq 0}$. Multiply by $e^{-\nu t}$ to obtain

$$e^{-\nu t} N_t = e^{-\nu u} (e^{-\nu(t-u)} N_{t-u}^{(0)} + e^{-\nu(t-u)} N_{t-u}^{(1)}).$$

Taking the limit in L^2 as t tends to infinity, the conditional distribution of W on $T_0 = u$ is the same as the distribution of $e^{-\nu u} (W^{(0)} + W^{(1)})$ where $W^{(0)}$ and $W^{(1)}$ have the same distribution as W . In particular, for all $u > 0$,

$$\mathbb{P}\{W = 0 \mid T_0 = u\} = \mathbb{P}\{W^{(0)} = 0, W^{(1)} = 0\} \leq \mathbb{P}\{W = 0\}.$$

Hence, $\mathbb{P}\{W = 0 \mid T_0 = u\} = \mathbb{P}\{W = 0\}$ μ -almost everywhere. Let

$$p = \mathbb{P}\{W = 0\} = \mathbb{P}\{W^{(0)} = 0, W^{(1)} = 0\}.$$

It follows that

$$\begin{aligned}
 p &= \int_{(\mathbb{R}^+)^3} \mathbb{P}\{W^{(0)} = 0, W^{(1)} = 0 \mid (T_0, T_{00}, T_{01}) = (x, y, z)\} dP(x, (y, z)) d\mu(x) \\
 &= \int_{(\mathbb{R}^+)^3} \mathbb{P}\{W^{(0)} = 0 \mid T_{00} = y\} \mathbb{P}\{W^{(1)} = 0 \mid T_{01} = z\} dP(x, (y, z)) d\mu(x),
 \end{aligned}$$

by Definition 3.2. Since $\mathbb{P}\{W^{(0)} = 0 \mid T_{00} = y\} = \mathbb{P}\{W^{(1)} = 0 \mid T_{01} = z\} = p$, it follows that $p = p^2$. But $p = 1$ is excluded since $\mathbb{E}\{W\} > 0$. Hence, $p = 0$.

Acknowledgements

This research was supported by Laboratoire d'Excellence TOUCAN (Toulouse Cancer). The authors wish to thank Christine Laurent-Thiébaud for her expertise on Tauberian lemmas and the anonymous referee for carefully reading the manuscript.

References

- [1] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
- [2] BELLMAN, R. AND HARRIS, T. (1952). On age-dependent binary branching processes. *Ann. Math. (2)* **55**, 280–295.
- [3] BENJAMINI, I. AND PERES, Y. (1994). Markov chains indexed by trees. *Ann. Probab.* **22**, 219–243.
- [4] BITSEKI PENDA, S. V., DJELLOUT, H. AND GUILLIN, A. (2014). Deviation inequalities, moderate deviations and some limit theorems for bifurcating Markov chains with application. *Ann. Appl. Probab.* **24**, 235–291.
- [5] COWAN, R. AND STAUDTE, R. (1986). The bifurcating autoregression model in cell lineage studies. *Biometrics* **42**, 769–783.
- [6] CRUMP, K. S. AND MODE, C. J. (1969). An age-dependent branching process with correlations among sister cells. *J. Appl. Probab.* **6**, 205–210.
- [7] DE LA PEÑA, V. H. AND LAI, T. L. (2001). Theory and applications of decoupling. In *Probability and Statistical Models with Applications*, Chapman & Hall/CRC, Boca Raton, FL, pp. 117–145.
- [8] DE SAPORTA, B., GÉGOUT-PETIT, A. AND MARSALLE, L. (2011). Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. *Electron. J. Statist.* **5**, 1313–1353.
- [9] DELMAS, J.-F. AND MARSALLE, L. (2010). Detection of cellular aging in a Galton–Watson process. *Stoch. Process. Appl.* **120**, 2495–2519.
- [10] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd edn. Springer, New York.
- [11] ESARY, J. D., PROSCHAN, F. AND WALKUP, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38**, 1466–1474.
- [12] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.
- [13] GLYNN, P. W. AND WHITT, W. (1994). Large deviations behavior of counting processes and their inverses. *Queueing Systems Theory Appl.* **17**, 107–128.
- [14] GUYON, J. (2007). Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.* **17**, 1538–1569.
- [15] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.
- [16] HARVEY, J. D. (1972). Synchronous growth of cells and the generation time distribution. *J. General Microbiol.* **70**, 99–107.
- [17] JOAG-DEV, K. AND PROSCHAN, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11**, 286–295.
- [18] JOHN, P. C. L. (ed.) (1981). *The Cell Cycle*. Cambridge University Press.
- [19] KENDALL, D. G. (1952). On the choice of a mathematical model to represent normal bacterial growth. *J. R. Statist. Soc. B* **14**, 41–44.
- [20] KLEPTSZYNA, M. L., LE BRETON, A. AND VIOT, M. (2002). New formulas concerning Laplace transforms of quadratic forms for general Gaussian sequences. *J. Appl. Math. Stoch. Anal.* **15**, 323–339.
- [21] KLEPTSZYNA, M., LE BRETON, A. AND YCART, B. (2014). Exponential transform of quadratic functional and multiplicative ergodicity of a Gauss–Markov process. *Statist. Probab. Lett.* **87**, 70–75.
- [22] KONTOYIANNIS, I. AND MEYN, S. P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.* **13**, 304–362.
- [23] KOREVAAR, J. (2004). *Tauberian Theory. A Century of Developments*. Springer, Berlin.
- [24] MARKHAM, J. F. *et al.* (2010). A minimum of two distinct heritable factors are required to explain correlation structures in proliferating lymphocytes. *J. R. Soc. Interface* **7**, 1049–1059.
- [25] MEYN, S. AND TWEEDIE, R. L. (2009). *Markov Chains and Stochastic Stability*, 2nd edn. Cambridge University Press.
- [26] NORDON, R. E., KO, K.-H., ODELL, R. AND SCHROEDER, T. (2011). Multi-type branching models to describe cell differentiation programs. *J. Theoret. Biol.* **277**, 7–18.
- [27] PEMANTLE, R. (1992). Automorphism invariant measures on trees. *Ann. Probab.* **20**, 1549–1566.
- [28] PEMANTLE, R. (1995). Tree-indexed processes. *Statist. Sci.* **10**, 200–213.
- [29] PITT, M. K., CHATFIELD, C. AND WALKER, S. G. (2002). Constructing first order stationary autoregressive models via latent processes. *Scand. J. Statist.* **29**, 657–663.
- [30] POWELL, E. O. (1956). Growth rate and generation time of bacteria with special reference to continuous culture. *Microbiol.* **15**, 492–511.
- [31] RAHN, O. (1932). A chemical explanation of the variability of the growth rate. *J. Gen. Physiol.* **15**, 257–277.
- [32] SHAO, Q.-M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theoret. Probab.* **13**, 343–356.
- [33] SPITZER, F. (1975). Markov random fields on an infinite tree. *Ann. Probab.* **3**, 387–398.
- [34] WANG, P. *et al.* (2010). Robust growth of *Escherichia coli*. *Curr. Biol.* **20**, 1099–1103.