Ancestral lineages and limit theorems for branching Markov chains

Vincent Bansaye

Ancestral lineages and limit theorems for branching Markov chains

Vincent Bansaye *

August 13, 2013

Abstract

We consider a branching model in discrete time where each individual has a trait in some general state space. Both the reproduction law and the trait inherited by the offsprings may depend on the trait of the mother and the environment. We study the long time behavior of the population and the ancestral lineage of typical individuals under general assumptions, which we specify for applications to some models motivated by biology. Our results focus on the growth rate, the trait distribution among the population for large time, so as local densities and the position of extremal individuals. The approach consists in comparing the branching Markov chain to a well chosen (possibly non-homogeneous) Markov chain. It relies in particular on an extension of many-to-one formula [G07, BDMT11] and spine decomposition in the vein of [LPP95, KLPP97, GB03]. The applications use properties of the underlying genealogy and sufficient conditions for the ergodic convergence of Markov chains.

Key words. Branching processes, Markov chains, Varying environment, Genealogies.

A.M.S. Classification. 60J80, 60J05, 60F05, 60F10

1 Introduction

We are interested in a branching Markov chain in a time varying environment, which means a non homogeneous multitype branching processes with possibly infinite number of types. Conditionally on the environment (quenched), this process satisfies the branching property, which means that each individual evolves independently.

Let (E, T) be a pair consisting of a set E of environments and an invertible map T on E. One can keep in mind the case when the environment is $\mathbf{e} = (e_i : i \in \mathbb{Z})$ and $T\mathbf{e} = (e_{i+1} : i \in \mathbb{Z})$.

*CMAP, Ecole Polytechnique, CNRS, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: vincent.bansaye@polytechnique.edu

Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ be a measurable space which gives the state space of the branching Markov chain. The example $\mathcal{X} \subset \mathbb{R}^d$ will be relevant for the applications.
For each $k \in \mathbb{N}$ and $\mathbf{e} \in E$, let $P^{(k)}(., \mathbf{e}, .)$ be a function from $\mathcal{X} \times \mathcal{B}_{\mathcal{X}^k}$ into $[0, 1]$ which satisfies
a) For each $x \in \mathcal{X}$, $P(x, \mathbf{e}, .)$ is a probability measure on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$.
b) For each $A \in \mathcal{B}_{\mathcal{X}}$, $P(., \mathbf{e}, A)$ is a $\mathcal{B}_{\mathcal{X}}$ measurable function.

In the whole paper, we use the classical notation $u = u_1 u_2 ... u_n$ with $u_i \in \mathbb{N}^*$ to identify an individual in the population. We denote by $|u| = n$ the generation of the individual u, by $N(u)$ the number of offsprings of the individual u and by $X(u) \in \mathcal{X}$ the trait (or position) of the individual u.

Each individual with trait $x \in \mathcal{X}$ which lives in environment $\mathbf{e} \in E$ gives birth independently to a random number of offspring, whose law both depend on x and $\mathbf{e}$. The number of offsprings is distributed as a r.v. $N(x, \mathbf{e})$ whose mean is denoted by

$$m(x, \mathbf{e}) = \mathbb{E}(N(x, \mathbf{e})).$$

In the whole paper, we assume that $m(x, \mathbf{e}) > 0$ for each $x \in \mathcal{X}, \mathbf{e} \in E$ for convenience. A natural framework for our models will be given $\mathbf{e} = (e_i : i \in \mathbb{Z})$ and $N(x, \mathbf{e})$ depends only on x and e_0, so that e_i yields the environment in generation i and and the reproduction law in environment in generation n just depends on e_n.

If the environment is $\mathbf{e}$, we denote by $\mathbb{P}_e$ the associated probability. The distribution of the traits of the offsprings of the individual u living in generation n ($|u| = n$) is given by

$$\mathbb{P}_{\mathbf{e}}(X_{u1} \in dx_1, \cdots, X_{uk} \in dx_k \mid (X(v) : |v| \leq n, N(u) = k) = P^{(k)}(X(u), T^n \mathbf{e}, dx_1 \cdots, dx_k).$$

In other words, one individual with trait x living in environment $\mathbf{e}$ gives birth to a set of individuals $(X_1, \cdots, X_{N(x,\mathbf{e})})$ whose trait are specified by $(P^{(k)}(\mathbf{e}, x.) : k \in \mathbb{N}, \mathbf{e} \in E)$.
This process is t a multitype branching process in varying environment where the types take value in $\mathcal{X}$. They have been largely studied for finite number of types, whereas much less is known or understood in the infinite case, but some results due to Seneta, Vere Jones, Moy, Kesten for countable types.
The case of branching random walk has also attracted lot of attention from the pioneering works of Biggins. Then $\mathcal{X} = \mathbb{R}^d$ and the transitions $P^{(k)}$ are invariant by translation, i.e. $P^{(k)}(x, \mathbf{e}, x + dx_1 \cdots, x + dx_k)$ does not depend on $x \in \mathcal{X}$. Recently, fine results have been obtained about the extremal individuals and their genealogy for such models, see e.g. [HS09, AS10]. Such methods and results have been extended to branching random walk in random environment. In particular the recurrence property [M08], the survival and the growth rate [GMPV10, CP07, CY11], central limit theorems [Y08, N11] and large deviations results [HL11] have been obtained.
As far as I see, the methods used for such models and in particular the martingale arguments are not easily adaptable to the general case considered here. We are motivated by applications to models for biology and ecology such as cell division models for cellular aging [G07] or parasite infection [B08] and reproduction-dispersion models in non-homogeneous environment [BL12]. For that purpose, we are inspired by the utilization of auxiliary Markov chains, branching decomposition and L^2 computations, in the vein

of the works of Athreya and Khang [AK98a, AK98b], Guyon [G07]. The applications and references will be given along the paper.

We are interested in the evolution of the measure associated to the traits of the individuals:

$$Z_n := \sum_{|u|=n} \delta_{X(u)}$$

and more specifically by $Z_n(A_n) = \#\{u : |u| = n,\ X(u) \in A_n\}$. We also define

$$Z_n(f) = \sum_{|u|=n} f(X(u)), \quad f_n.Z_n = \sum_{|u|=n} \delta_{f_n(X(u))}.$$

First, we want to know if the process may survive globally and how it would grow. Thus, Section 2 extends the growth rate characterization of [BL12] for metapopulations to infinite number of patches and varying environment. Then (Section 3), we study the repartition of the population and focus the asymptotic behavior of the proportions of individuals whose trait belongs to A, i.e. $Z_n(\mathcal{X})/Z_n(A)$. It generalizes [AK98a, G07] to both varying and space dependent reproduction. It extends [BH13] to space dependent reproduction and complete the results by giving strong law of large numbers. We add that we take into account some possible renormalization of the traits via a function f_n to cover non recurrent positive cases. Finally, in Section 4, we provide some asymptotic results about $Z_n(A_n)$, outside the range of law of large numbers. It relies on the large deviations of the auxiliary process and the trajectory associated with. As an application we can derive the position of the extreme particles in some monotone models motivated by biology, where new behaviors appear.
The probabilistic approach we follow yields a way to simulate the long time distribution of the population and will be applied to some biological models motivated by cell division or reproduction - dispersion dynamics.

We end up the introduction with recalling some classical notations. The individuals If $u = u_1 \cdots u_n$ and $v = v_1 \cdots v_m$, then $uv = u_1 \cdots u_n v_1 \cdots v_m$. For two different individuals u, v of a tree, write $u < v$ if u is an ancestor of v, and denote by $u \wedge v$ the nearest common ancestor of u and v in the means that $|w| \leq |u \wedge v|$ if $w < u$ and $w < v$.

2 Growth rate of the population

We denote by $\rho_{\mathbf{e}} = \lim_{n\to\infty} n^{-1} \log \mathbb{E}_{\mathbf{e}}(Z_n)$ the growth rate of the population in the environment $\mathbf{e}$, when it exists.
We are giving an expression of this growth rate in terms of a Markov chain associated with a random lineage. Its transition kernel is defined by

$$P(x, \mathbf{e}, dy) := \frac{1}{m(x, \mathbf{e})} \sum_{k\geq 1} \mathbb{P}(N(x, \mathbf{e}) = k) \sum_{i=0}^{k-1} P^{(k)}(x, \mathbf{e}, \mathcal{X}^{i-1} dy \mathcal{X}^{k-i})$$

so that the auxiliary Markov chain X is given by

$$\mathbb{P}_{\mathbf{e}}(X_{n+1} \in dy \ | X_n = x) = P(x, T^n \mathbf{e}, dy).$$

It means that we follow a linage by choosing uniformly at random one of the offspring at each generation.

We assume now that $\mathcal{X}$ is a locally compact polish space endowed with a complete metric and its Borel σ field. Moreover E is a Polish Space and T is an homeomorphism. We can then consider the weak topology associated to $\mathcal{M}_1(\mathcal{X} \times E)$, which means the smallest topology such that $\mu \in \mathcal{M}_1(\mathcal{X} \times E) \to \int_{\mathcal{X}\times E} f(z)\mu(dz)$ is continuous soon as f is continuous and bounded.

Definition 1. *We say that X satisfies a Large deviation principle (PGD) with good rate function $I_{\mathbf{e}}$ in environment $\mathbf{e}$ when there exists a lower semi continuous function $I : \mathcal{X} \times E \to \mathbb{R}$ with compact level subsets*[1] *for the weak topology such that*

$$L_n^{\mathbf{e}} = \frac{1}{n}\sum_{k=0}^{n} \delta_{X_k, T^k\mathbf{e}}$$

satisfies for every $x \in \mathcal{X}$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mathbf{e},x}(L_n \in F) \leq -\inf_{z\in F} I_{\mathbf{e}}(z)$$

for every closed set F of $\mathcal{M}_1(\mathcal{X} \times E)$, and

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mathbf{e},x}(L_n \in O) \geq -\inf_{z\in O} I_{\mathbf{e}}(z)$$

for every open set O of $\mathcal{M}_1(\mathcal{X} \times E)$.

The existence of such a principle is classical for fixed environment $E = \{\mathbf{e}\}$, finite $\mathcal{X}$, under irreducibility assumption. We refer to Sanov's theorem, see e.g. chapter 6.2 in [DZ98]. We note that the principle can be extended to periodic environments, taking care of the irreducibility. We are using an analogous result for stationary random environment to get forthcoming Corollary 2, under Doeblin conditions, which is due to [S94].

The first question that we tackle now is the mean growth rate of the population. The branching property yields the linearity of the operator $\mu \to m(\mu) = \mathbb{E}_{\mathbf{e},\mu}(Z_1(.))$ for some measurable set A.
In the case of fixed environment, P and N do not depend on $\mathbf{e}$, so m is also fixed and the mean growth rate of the process Z is the limit of $\log |m^n|_1/n$. In the case $\mathcal{X}$ is finite, it is Perron Frobenius eigenvalue under strong irreducibility assumption, with a min max representation due to Collatz Wielandt. Krein-Rutman theorem gives an extension to infinite dimension space requiring compactness of the operator m and strict positivity.
In the random environment case, it is the Lyapounov exponent and quenched asymptotic results can be obtained in the case $\mathcal{X}$ is finite [FK60]. Then, we get here branching processes in random environment and we refer to [AK71, K74] for extinction criteria and [T88] for the growth rate of this process.

To go beyond these assumptions and get an interpretation of the growth rate in terms of reproduction-dispersion dynamics, we provide here an other characterization.
This is a functional large deviation principle relying on Varadhan's lemma. It allows to decouple the reproduction and dispersion. It yields an extension of Theorem 5.3 in [BL12] both for varying environment and infinite state pace $\mathcal{X}$. We refer to this latter article for

[1] It means that $\{\mu \in \mathcal{M}_1(\mathcal{X} \times E) : I(\mu) \leq l\}$ is compact for the weak topology

motivations in ecology, more specifically for metapopulations. The next Corollary then puts in light the dispersion strategy followed by typical individuals of the population for large times.

Theorem 1. *Assume that X satisfies a PGD with good rate function $I_{\mathbf{e}}$ in environment $\mathbf{e}$ and $\log m : \mathcal{X} \times E \to (-\infty, \infty)$ is continuous and bounded. Then, for every $x \in \mathcal{X}$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbf{e},\delta_x}(Z_n(\mathcal{X})) = \sup_{\mu\in\mathcal{M}_1(\mathcal{X}\times E)} \left\{ \int_{\mathcal{X}\times E} m(x,e)\mu(dxde) - I_{\mathbf{e}}(\mu) \right\} := \varrho_{\mathbf{e}}$$

and

$$M_{\mathbf{e}} := \left\{ \mu \in \mathcal{M}_1(\mathcal{X}\times E) : \int \log m(x,e)\mu(dxde) - I_{\mathbf{e}}(\mu) = \varrho_{\mathbf{e}} \right\}$$

is compact and non empty.

In particular, $\limsup_{n\to\infty} \frac{1}{n} \log Z_n(\mathcal{X}) \leq \varrho_{\mathbf{e}}$ a.s. The limit can hold only on the survival event. It is the case under classical $N \log N$ moment assumption for finite state space $\mathcal{X}$, see [LPP95] for fixed environment and [AK71] in random environment. But it is a rather delicate problem when the number of types is infinite.

We introduce now the event

$$\mathcal{S} := \left\{ \liminf_{n\to\infty} \frac{1}{n} \log Z_n(\mathcal{X}) \ \geq \ \varrho_{\mathbf{e}} \right\}.$$

Conditionally on $\mathcal{S}$, we let U_n be an individual uniformly chosen at random in generation n and the trait frequency up to time n and the associated environment :

$$\nu_n(A) := \frac{1}{n+1} \#\{0 \leq i \leq n : (X_i(U_n), T^i\mathbf{e}) \in A\} \qquad (A \in \mathcal{B}_{\mathcal{X}\times E}).$$

where $X_i(u)$ the position of the ancestor of u in generation i. We prove that the support of ν_n converges in probability to $M_{\mathbf{e}}$ on the event $\mathcal{S}$.

Corollary 1. *Under the assumptions of Theorem 1, we further suppose that $\varrho_{\mathbf{e}} > 0$ and $\mathcal{S}$ has positive probability. Then,*

$$\mathbb{P}_{\mathbf{e}}(\nu_n \in F | \mathcal{S}) \overset{n\to\infty}{\longrightarrow} 0$$

for every closed set F of $\mathcal{M}_1(\mathcal{X}\times E)$ which is disjoint of $M_{\mathbf{e}}$.

This result yields an information on the pedigree [JN96, NJ84] or ancestral lineage of a typical individual. It says that the trait frequency of the lineage of a typical individual converges to one of the argmax of $\varrho_{\mathbf{e}}$. We are going a bit farther in the next Section, with a description of this ancestral lineage via size biased random choice. Seeing the population from a typical individual via spine decomposition has been firstly achieved for Galton-Watson processes in [LPP95]. We refer to [KLPP97] for an extension to multitype Galton Watson processes, [GB03] for continuous time and [G99] for related results in varying environment.

Let us now specify the theorem for stationary ergodic environment $\mathcal{E} \in E$, under Doeblin assumptions. Following [S94], we let π be a T invariant ergodic probability, i.e. $\pi \circ T^{-1} = \pi$ and if $A \in \mathcal{B}_E$ satisfies $T^{-1}A = A$, then $\pi(A) \in \{0, 1\}$. Then we need :

Assumption A. There exist a positive integer b, a T invariant subset E' of E and a measurable function $M : E \to [1, \infty)$ such that $\log M \in L^1(\pi)$, $\pi(E') = 1$ and for all $x, y \in \mathcal{X}$, $A, B \in \mathcal{B}_{\mathcal{X}}$ and $\mathbf{e} \in E'$,

$$P^b(x, \mathbf{e}, A) \leq M(\mathbf{e}) P^b(y, \mathbf{e}, A).$$

We denote by $V_b(\mathcal{X} \times E)$ the set of bounded continuous functions that map $\mathcal{X} \times E$ into $[1, \infty)$ to state the result.

Corollary 2. *Under Assumption A, we have π a.s., for every $x \in \mathcal{X}$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathcal{E},\delta_x}(Z_n(\mathcal{X})) = \sup_{\mu \in \mathcal{M}(\mathcal{X}\times E)} \left\{ \int \log(m(x,e))\mu(dx,de) - I(\mu) \right\},$$

where I is defined by

$$I(\mu) := \sup \left\{ \int_{\mathcal{X}\times E} \log \left(\frac{u(x,e)}{\int_{\mathcal{X}} P(x,e,dy) u(y,Te)} \right) \mu(dx,de) : u \in V_b(\mathcal{X} \times E) \right\}.$$

Proof of Theorem 1. We first note that by Markov and branching property

$$\mathbb{E}_{\mathbf{e},\nu}(Z_{n+1}(A)) = \int_{\mathcal{X}} \mathbb{E}_{\mathbf{e},\nu}(Z_n(dy)) \mathbb{E}_{T^n\mathbf{e},\delta_y}(Z_1(A)). \tag{1}$$

As

$$\begin{aligned} \mathbb{E}_{T^n\mathbf{e},\delta_y}(Z_1(A)) &= \sum_{k\geq 1} \mathbb{P}(N(y, T^n\mathbf{e}) = k) \sum_{i=1}^{k} P^{(k)}(y, T^n\mathbf{e}, \mathcal{X}^{k-i} \times A \times \mathcal{X}^i) \\ &= m(y, T^n\mathbf{e}) P(y, T^n\mathbf{e}, A), \end{aligned}$$

we get

$$\mathbb{E}_{\mathbf{e},\nu}(Z_{n+1}(A)) = \int_{\mathcal{X}} \mathbb{E}_{\mathbf{e},\nu}(Z_n(dy)) m(y, T^n\mathbf{e}) P(y, T^n\mathbf{e}, A)$$

By induction,

$$\mathbb{E}_{\mathbf{e},\nu}(Z_n(A)) = \int_{\mathcal{X}^n \times A} \nu(dy_0) \prod_{i=1}^{n-1} m(y_i, T^i\mathbf{e}) P(y_i, T^i\mathbf{e}, dy_{i+1}) \tag{2}$$

and in particular

$$\mathbb{E}_{\mathbf{e},\nu}(Z_n(\mathcal{X})) = \mathbb{E}_{\mathbf{e},\nu} \left(\prod_{i=0}^{n} m(X_i, T^i\mathbf{e}) \right).$$

Thus

$$\mathbb{E}_{\mathbf{e},\nu}(Z_n(\mathcal{X})) = \mathbb{E}_{\mathbf{e},\nu} \left(\exp \left(n \int_{\mathcal{X}\times E} \log(m(x,e)) L_n^{\mathbf{e}}(dx,de) \right) \right)$$

As $\log m$ is bounded and continuous, so is

$$\mu \in \mathcal{M}_1(\mathcal{X} \times E) \to \phi(\mu) = \int_{\mathcal{X}\times E} \log(m(x,e))\mu(dx,de).$$

Using the LDP principle satisfied by $L_n^{\mathbf{e}}$, we can apply Varadhan's lemma (see [DZ98] Theorem 4.3.1) to the previous function to get the first part of the Theorem.

Let us now consider a sequence μ_n such that

$$\int_{\mathcal{X}\times E} \log(m(x,e))\mu_n(dxde) - I_{\mathbf{e}}(\mu_n) \overset{n\to\infty}{\longrightarrow} \varrho_{\mathbf{e}}.$$

Then $I_e(\mu_n)$ is upper bounded, which ensures that μ_n belongs to a sublevel set. By Defintion 1, such a set is compact so can extract a subsequence $\mu_{\phi(n)}$ which converges weakly in $\mathcal{M}(\mathcal{X}, E)$. As I_e is lower semicontinuous, the limit μ of this subsequence satisfies

$$\liminf_{n\to\infty} I_{\mathbf{e}}(\mu_{\phi(n)}) \geq I_{\mathbf{e}}(\mu).$$

Recalling that $\mu \to \int \log m\mu$ is continuous, we get

$$\varrho_{\mathbf{e}} = \lim_{n\to\infty} \left\{ \int_{\mathcal{X}\times E} m(x,e)\mu_{\phi(n)}(dxde) - I_{\mathbf{e}}(\mu_{\phi(n)}) \right\} \leq \int m(x,e)\mu(dxde) - I_{\mathbf{e}}(\mu)$$

and μ is a maximizer. That ensures that $M_{\mathbf{e}}$ is compact and non empty. □

Proof of Corollary 1. Let us consider a closed subset of F of $\mathcal{M}_1(\mathcal{X}\times E)$ which is disjoint of $M_{\mathbf{e}}$. Then, using again the continuity of $\mu \to \int m\mu$, we have

$$\varrho_F := \sup_{\mu\in F} \left\{ \int_{\mathcal{X}\times E} m(x,e)\mu(dxde) - I_{\mathbf{e}}(\mu) \right\} < \varrho_{\mathbf{e}}.$$

We define for any individual u in generation n

$$\nu_n(u)(A) = \frac{1}{n+1} \sum_{i\leq n} \delta_{X_i(u)}.$$

Following the computation (2),

$$\mathbb{E}_{\mathbf{e},\nu} \left(\#\{u : |u| = n, \nu_n(u) \in F\} \right) = \mathbb{E}_\nu \left(\exp\left(n \int_{\mathcal{X}\times E} m(x,e) L_n^{\mathbf{e}}(dx,de) \right) 1_{L_n^{\mathbf{e}}\in F} \right)$$

Applying again Varadhan's to the bounded continuous function $\phi : \mathcal{M}_1(\mathcal{X}\times E) \to \mathbb{R}$ defined by every $\mu \in F$,

$$\phi(\mu) = \min\left(\int_{\mathcal{X}\times E} m(x,e) L_n^{\mathbf{e}}(dx,de), \varrho_F \right),$$

we get,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbf{e},\nu} \left(\#\{u : |u| = n, F_n(u) \in F\} \right) &\leq \sup\{\phi(\mu) - I_{\mathbf{e}}(\mu) : \mu \in \mathcal{M}_1(\mathcal{X}\times E)\} \\ &\leq \varrho_F. \end{aligned}$$

Adding that if $\rho' \in (\rho_F, \rho_{\mathbf{e}})$,

$$\begin{aligned} \mathbb{P}(\nu_n(U_n) \in F | \mathcal{S}) &\leq \mathbb{E}\left(\#\{u : |u| = n, \nu_n(u) \in F\} / Z_n(\mathcal{X}) | S \right) \\ &\leq e^{-\rho' n} \mathbb{E}\left(\#\{u : |u| = n, \nu_n(u) \in F\} \right) / \mathbb{P}(\mathcal{S}) \end{aligned}$$

for n large enough by definition of $\mathcal{S}$, we get that the left hand side goes to 0. □

Proof of Corollary 2. Under Assumption A, Theorem 3.3 [S94] ensures that there exists a function I which satisfies Definition 1 π a.s. and uniformly in $x \in \mathcal{X}$. The result is then a direct application of the previous Theorem. □

We have given above an expression of the mean growth rate and specify the ancestral lineage of surviving individuals. It leaves several open questions and we are considering the following ones in the next Section, which are linked :
Does the process grows like its mean when it survives ?
How is the population spread for large times ?

3 Law of large numbers

We consider the mean measure under the environment $\mathbf{e}$:

$$m_n(x, \mathbf{e}, A) := \mathbb{E}_{\delta_x, \mathbf{e}}\left(Z_n(A)\right) = \mathbb{E}_{\delta_x, \mathbf{e}}\left(\#\{u : |u| = n, X(u) \in A\right) \qquad (A \in \mathcal{B}_{\mathcal{X}}).$$

It yields the mean number of descendant in generation n of an initial individual with trait x, whose trait belongs A. Similarly we consider its mean descendance in generation n We define a new family of Markov kernel Q_n by

$$Q_n(x, \mathbf{e}, dy) := m_1(x, \mathbf{e}, dy)\frac{m_{n-1}(y, T\mathbf{e}, \mathcal{X})}{m_n(x, \mathbf{e}, \mathcal{X})}.$$

The fact that $Q_n(x, \mathbf{e}, \mathcal{X}) = 1$ for all $n \in \mathbb{N}, x \in \mathcal{X}, \mathbf{e} \in E$ comes directly from the branching property (1). We introduce the associated semigroup, more precisely the successive composition of Q_j between the generations i and n :

$$Q_{i,n}(x, \mathbf{e}, A) = Q_{n-i}(x, T^i\mathbf{e}, .) * Q_{n-i-1}(., T^{i+1}\mathbf{e}, .) * \cdots * Q_1(., T^{n-1}\mathbf{e}, .)(A),$$

where we recall the notation $Q(x, .) * Q'(., .)(A) = \int_{\mathcal{X}} Q(x, dy)Q'(y, A)$. The next section links the semigroups m_n and $Q_{0,n}$.

3.1 The auxiliary process and the many to one formula

The following many to one formula links the expectation of the number of individuals whose trait belongs to A to the probability that the Markov chain associated to the kernel Q_n belongs to A. We recall the notation $Q(x, \mathbf{e}, f)(x) = \int_{\mathcal{X}} f(y)Q(x, \mathbf{e}, dy)$.

Lemma 1. *For all $n \in \mathbb{N}$, $x \in \mathcal{X}$ and $f \in \mathcal{B}(\mathcal{X})$, we have*

$$m_n(x, \mathbf{e}, f) = m_n(x, \mathbf{e}, \mathcal{X})Q_{0,n}(x, \mathbf{e}, f).$$

We note that $m_n(x, \mathbf{e}, \mathcal{X})$ is the mean number of individuals in generation n considered in the previous Section. Here, combining the branching property and the lemma above yields an other expression of the growth rate, linked with the ergodic behavior of $Q_{0,n}$ we are considering below :

$$\frac{m_{n+1}(x, \mathbf{e}, \mathcal{X})}{m_n(x, \mathbf{e}, \mathcal{X})} = \int_{\mathcal{X}} m(y, T^n\mathbf{e})Q_{0,n}(x, \mathbf{e}, dy).$$

The many-to-one formula is linked to the spine decomposition mentioned above and $Q_{0,n}$ yield the dynamic of the trait of a typical individual in a size biased tree, "along the

spine". More explicitly, it is in the same vein as the many one to one formula for binary tree [G07], Galton -Watson trees [DM10] and Galton- Watson trees in stationary random environments [BH13]. In continuous time, many- to- one formula and formula for forks can been found in [BDMT11]. But these later do not let the reproduction depend on the position.
In continuous time, we refer to [C11, HR12, HR13] for other many to one formulas and asymptotic results with reproduction law depend on the trait in some particular cases.

Proof. We recall the notation of the previous Section 2 and the Markov chain X. Following the computations of the proof of Theorem 1 to obtain (2), we get

$$\begin{aligned}\mathbb{E}_{\nu,\mathbf{e}}\left(\sum_{|u|=n} f(X(u))\right) &= \mathbb{E}_{\nu,\mathbf{e}}\left(f(X_n)\prod_{i=0}^{n-1} m(X_i, T^i\mathbf{e}, \mathcal{X})\right)\\ &= \int_{\mathcal{X}^{n+1}} f(x_n)\nu(dx_0)\prod_{i=0}^{n-1} m(x_i, T^i\mathbf{e}, dx_{i+1}).\end{aligned}$$

Moreover

$$\prod_{i=0}^{n-1} Q_{n-i}(x_i, T^i\mathbf{e}, dx_{i+1}) = \frac{m_0(x_n, \mathbf{e}, \mathcal{X})}{m_n(x_0, \mathbf{e}, \mathcal{X})}\prod_{i=0}^{n-1} m_1(x_i, T^i\mathbf{e}, dx_{i+1}).$$

As $m_0(x_n, \mathbf{e}, \mathcal{X}) = 1$, we get

$$\mathbb{E}_{\nu,\mathbf{e}}\left(\sum_{|u|=n} f_n(X(u))\right) = m_n(x, \mathbf{e}, \mathcal{X})\int_{\mathcal{X}} f(x_n)\int_{\mathcal{X}^n}\prod_{i=0}^{n-1} Q_n(x_i, T^i\mathbf{e}, dx_{i+1}),$$

which completes the proof by the definition of $Q_{0,n}$. □

Our aim is now to get ride of the expectation and obtain the repartition of the population for large times. We want to derive it from the stationary distribution of this auxiliary Markov chain with kernel Q_n and prove a law of large number on the proportions of individuals whose trait traited belongs to A. One approach would be to a find a martingale via maximal eigenvalue and eigenvector, as for finite type and fixed environment. It has been extended to branching processes with infinite number of types in [A00] but the assumptions required are not easily fulfilled, at least regarding the biological and ecological we give in motivations in this work. Moreover the generalization to varying environment seems more adapted to the technicals described here. Thus, we are here following ideas developed in [AK98a, AK98b] using the branching property and the growth of the population or that in [G07] relying on L^2 computations and control of the underlying genealogy.

3.2 Branching decomposition

In this part, we focus on the particular case when extinction does not occur and actually assume that the population has a positive growth rate. We have then the following strong law of large numbers.

Theorem 2. *Let us fix* $\mathbf{e} \in E$ *and* $\mathcal{F} \subset \mathcal{B}(\mathcal{X})$. *We assume that*

$$\mathbb{P}_{\mathbf{e}}\left(\forall n,\ Z_n(\mathcal{X}) > 0;\ \liminf_{n\to\infty} \frac{Z_{n+1}(\mathcal{X})}{Z_n(\mathcal{X})} > 1\right) = 1$$

and that the exists a measure Q *with finite first moment such that for all* $x \in \mathcal{X}, k, l \geq 0$,

$$\mathbb{P}_{\mathbf{e}}(N(x, T^k\mathbf{e}) \geq l) \leq Q[l, \infty). \tag{3}$$

Assume also that there exists a sequence of probability measure μ_n *such that*

$$\sup_{\lambda\in\mathcal{M}_1(\mathcal{X})} \left|Q_{i,n}(\lambda, T^i\mathbf{e}, f \circ f_n) - \mu_n(f)\right| \longrightarrow 0, \tag{4}$$

uniformly for $n - i \to \infty$ *and* $f \in \mathcal{F}$. *Then,*

$$\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f) \overset{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e}}\ a.s. \tag{5}$$

This result extends [AK98a, AK98b] to the case when the reproduction law may depend on the trait, the Markov kernel $P^{(k)}$ is not a direct product of the same kernel and to time varying environment. It yields a strong law of large numbers relying on the uniform ergodicity of the auxiliary Markov chain $Q_{i,n}$. The assumption of a.s. survival and positive growth rate will be relaxed in the next part using L^2 assumptions.

Proof. The branching property gives a natural decomposition of the population in generation $n + p$, as already used in [AK98b] :

$$Z_{n+p}(\mathcal{X}) = \sum_{|u|=n} Z_p^{(u)}(\mathcal{X}),$$

where $Z^{(u)}$ is the branching Markov chain whose root is the individual u and environment is $T^n\mathbf{e}$. First, we check that

$$\frac{1}{Z_{n+p}(\mathcal{X})} \sum_{|u|=n} m_p(X(u), T^n\mathbf{e}, \mathcal{X}) \to 1 \quad \text{a.s.} \tag{6}$$

as n goes to ∞. Indeed,

$$Z_{n+p}(\mathcal{X}) - \sum_{|u|=n} m_p(X(u), T^n\mathbf{e}, \mathcal{X}) = \sum_{|u|=n} \left[Z_p^{(u)}(\mathcal{X}) - m_p(X(u), T^n\mathbf{e}, \mathcal{X})\right] = Z_n(\mathcal{X})\epsilon_{n,p},$$

where

$$\epsilon_{n,p} := \frac{1}{Z_n(\mathcal{X})} \sum_{|u|=n} X_{p,u}, \qquad X_{p,u} = Z_p^{(u)}(\mathcal{X}) - m_p(X(u), T^n\mathbf{e})$$

We note that $(X_{p,u} : |u| = n)$ are independent conditionally on $\mathcal{F}_n = \sigma(X(v) : |v| \leq n)$, $\mathbb{E}(X_{p,u}) = 0$ and $|X_{p,u}| \leq |Z_p^{(u)}(\mathcal{X})| + m_p(X(u), T^n\mathbf{e})$, so that the stochastic domination assumption (3) ensures that there exists a probability distribution with finite first moment Q' such that

$$\sup_{u\in\mathbb{T}} \mathbb{P}_{\mathbf{e}}(|X_{p,u}| > t|\mathcal{F}_{|u|}) \leq Q'(t, \infty),$$

where we recall that $\mathbb{T}$ is the set of all individuals. We can then apply the law of large number of Lemma 1 in [AK98a] to get that for every $p \geq 0$, $\epsilon_{n,p} \to 0$ as $n \to \infty$ $\mathbb{P}_{\mathbf{e}}$ a.s. and obtain (6).
Let $\epsilon > 0$. The strong ergodicity assumption (4) ensures that there exists $p \in \mathbb{N}$ such that for every $n \in \mathbb{N}$,

$$\sup_{\lambda \in \mathcal{M}_1} \left| Q_{n,n+p}(\lambda, T^n \mathbf{e}, f \circ f_n) - \mu_{n+p}(f) \right| \leq \epsilon. \tag{7}$$

Then, using the same branching decomposition, we are proving

$$\begin{aligned} &\limsup_{n\to\infty} \left| \frac{Z_{n+p}(f)}{Z_{n+p}(\mathcal{X})} - \mu_{n+p}(f) \right| \\ &\quad = \limsup_{n\to\infty} \left| \frac{1}{Z_n(\mathcal{X})} \sum_{|u|=n} \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} Z_p^{(u)}(f) - \mu_{n+p}(f) \right| \leq \epsilon \quad \text{a.s} \end{aligned}$$

For that purpose, we split the last term and use the many-to-one formula (Lemma 1)

$$\begin{aligned} &\left| \frac{1}{Z_n(\mathcal{X})} \sum_{|u|=n} \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} Z_p^{(u)}(f) - \mu_{n+p}(f) \right| \\ &\quad \leq \frac{1}{Z_n(\mathcal{X})} \left| \sum_{|u|=n} \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} \left[Z_p^{(u)}(f) - m_p(X(u), T\mathbf{e}, f) \right] \right| \\ &\qquad + \frac{1}{Z_{n+p}(\mathcal{X})} \left| \sum_{|u|=n} m_p(X(u), T^n \mathbf{e}, \mathcal{X}) [Q_p(X(u), T^n \mathbf{e}, f) - \mu_{n,n+p}(f)] \right| \\ &\qquad + \mu_{n,n+p}(f) \left| \sum_{|u|=n} \frac{m_p(X(u), T^n \mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} - 1 \right|. \end{aligned}$$

To prove that the first term go to zero a.s. as $n \to \infty$, we use again the law of large numbers of Lemma 1 in [AK98a] with now

$$X_{u,n} = \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} \left[Z_p^{(u)}(f) - m_p(X(u), T\mathbf{e}, f) \right].$$

We note that $X_{u,n} \leq Z_p^{(u)}(f) + m_p(X(u), T^n \mathbf{e}, f) \leq Z_p^{(u)}(\mathcal{X}) + m_p(X(u), T^n \mathbf{e}, f)$ yields the stochastic domination of $X_{u,n}$ required as above for $X_{p,u}$. Thus, using (7),

$$\begin{aligned} &\limsup_{n\to\infty} \left| \frac{1}{Z_n(\mathcal{X})} \sum_{u \in \mathbb{G}_n} \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} Z_p^{(u)}(f) - \mu_{n+p}(f) \right| \\ &\quad \leq \epsilon \limsup_{n\to\infty} \frac{1}{Z_n(\mathcal{X})} \sum_{|u|=n} \frac{m_p(X(u), T^n \mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} + \mu(f) \limsup_{n\to\infty} \left| \sum_{|u|=n} \frac{m_p(X(u), T^n \mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} - 1 \right| \\ &\quad \leq \epsilon.1 + 0, \end{aligned}$$

recalling (6). It ends up the proof by letting $\epsilon \to 0$. □

3.3 L^2 approach

In this Section, we state weak and strong law of large numbers using L^2 computations with ergodicity of the auxiliary Markov chain and a control on the most recent common ancestor of the individuals.
We recall the notations $Q(\lambda, \mathbf{e}, f)(x) = \int_{\mathcal{X}^2} \lambda(dx)Q(x, \mathbf{e}, dy)f(y)$ and $\mathcal{B}(\mathcal{X})$ for the set of measurable functions from $\mathcal{X}$ to $\mathbb{R}$. We note $\mathcal{B}_b(\mathcal{X})$ the set of measurable functions from $\mathcal{X}$ to $\mathbb{R}$, which are bounded by b.

The main assumption we are using concern the ergodic behavior of the time non-homogeneous auxiliary Markov chain associated to the transitions kernels $Q_{i,n}$.

Assumption 1. Let $\mathbf{e}_n \in E$, $\mathcal{F} \subset \mathcal{B}(\mathcal{X})$, $f_n \in \mathcal{B}(\mathcal{X})$ and $\mu_n \in \mathcal{M}(\mathcal{X})$ for each $n \in \mathbb{N}$.
(a) For all $\lambda \in \mathcal{M}(\mathcal{X})$ and $i \in \mathbb{N}$,

$$\sup_{f\in\mathcal{F}} \left|Q_{i,n}(\lambda, \mathbf{e}_n, f \circ f_n) - \mu_n(f)\right| \overset{n\to\infty}{\longrightarrow} 0$$

(b) For every $k_n \leq n$ such that $n - k_n \to \infty$,

$$\sup_{\lambda\in\mathcal{M}(\mathcal{X}), f\in\mathcal{F}} \left|Q_{k_n,n}(\lambda, \mathbf{e}_n, f \circ f_n) - \mu_n(f)\right| \overset{n\to\infty}{\longrightarrow} 0.$$

The second assumption (uniform ergodicity) clearly implies the first one. Sufficient conditions will be given in the applications, such as Doeblin's conditions. The function f_n is bound to make the process ergodic if it is not originally. We have for example in mind the case when the auxiliary Markov chain X_n satisfies a central limit theorem, i.e. $f_n(x) = (x - a_n)/b_n$ when $f(X_n)$ converges to the same distribution whatever the initial value X_0 is. It will be the case for the applications to branching random walks.

We consider now the genealogies of the population and the time of the most recent common ancestor of two individuals chosen uniformly.

Assumption 2. (a) For every $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that for n large enough,

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_x}(\#\{u, v : |u| = |v| = n,\ u \wedge v \geq K\})}{m_n(x, \mathbf{e}_n, \mathcal{X})^2} \leq \epsilon. \tag{8}$$

Moreover there exists $C_i \in \mathcal{B}(\mathcal{X}^2)$ such that for all $i \in \mathbb{N}, x, y \in \mathcal{X}$,

$$\sup_{n\geq i} \frac{m_{n-i}(y, T^i\mathbf{e}_n, \mathcal{X})}{m_n(x, \mathbf{e}_n, \mathcal{X})} \leq C_i(x, y), \quad \text{with } \mathbb{E}\left(\max\{C_i(x, X(w))^2 : |w| = i + 1\}\right) < \infty.$$

(b) For every $K \in \mathbb{N}$,

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_x}(\#\{u, v : |u| = |v| = n,\ u \wedge v \geq n - K\})}{m_n(x, \mathbf{e}_n, \mathcal{X})^2} \overset{n\to\infty}{\longrightarrow} 0. \tag{9}$$

Moreover

$$\sup_{n\in\mathbb{N}} \mathbb{E}(Z_n(\mathcal{X})^2)/m_n(x, \mathbf{e}_n, \mathcal{X})^2 < \infty.$$

These expressions can be rewritten in terms of variance of $Z_n(\mathcal{X})$ and more tractable sufficient assumptions can be specified, see the applications. We observe also that these assumptions require that $Z_n(\mathcal{X})$ has a finite second moment, so each reproduction law involved in the dynamic has a finite second moment. Moreover $m_n(x, \mathbf{e}_n, \mathcal{X})$ has to go to ∞.

The assumption (8) says that the common ancestor is at the beginning of the genealogy. It is the case for Galton-Watson trees, branching processes in random environment and many others "regular trees". The assumption (9) says that the common ancestor is not at the end of the genealogy, so it is weaker. For a simple example where (8) is fulfilled but (9) is not, one can consider the tree T_n which is composed by a single individual until generation $n - k_n$ and equal to the binary tree between the generations $n - k_n$ and n, with $k_n \to \infty$. One can construct also examples of branching Markov chain with time homogeneous reproduction. It can be achieved for example by considering increasing Markov chains and increasing mean reproduction (which may be deterministic) with respect to $x \in \mathcal{X}$.

Theorem 3 (Weak LLN)**.** *Let $\mathbf{e}_n \in E^n$, $x \in \mathcal{X}$, $f_n : \mathcal{X} \to \mathcal{X}$ and $\mathcal{F} \subset \mathcal{B}_b(\mathcal{X})$.*

We assume either that Assumptions 1(a) and 2(a) hold or that Assumptions 1(b) and 2(b) hold. Then, uniformly for $f \in \mathcal{F}$,

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x, \mathbf{e}_n, \mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \tag{10}$$

in $L^2_{\mathbf{e}_n,\delta_x}$ and for all $\epsilon, \eta > 0$,

$$\mathbb{P}_{\mathbf{e}_n,\delta_x}\left(\left|\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f)\right| \geq \eta \ ; \ Z_n(\mathcal{X})/m_n(x, \mathbf{e}_n, \mathcal{X}) \geq \epsilon\right) \overset{n\to\infty}{\longrightarrow} 0.$$

We note that $f_n.Z_n(1_A)/Z_n(\mathcal{X})$ is the proportion of individuals in generation n whose trait belongs to $f_n^{-1}(A)$.

We recover the classical weak law of large numbers for Markov chains along Galton-Watson trees [DM10] and along branching processes in random environment [BH13]. Indeed, in these cases $W_n = Z_n/m_n(x, \mathbf{e}, \mathcal{X})$ is (a.s. with respect to the environment) a martingale which converges to a positive limit on the non extinction event thanks to L^2 assumptions, so that we obtain

$$\mathbb{P}\left(\left|\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f)\right| \geq \eta \ ; \forall n \in \mathbb{N}, Z_n(\mathcal{X}) > 0\right) \overset{n\to\infty}{\longrightarrow} 0.$$

We give also new law of large numbers in the forthcoming applications. Finally, we note that the Theorem holds also if $f_n : \mathcal{X} \to \mathcal{X}'$ and can be extended to unbounded f with domination assumptions following [G07].

Proof. Let us prove the first part of the Theorem under Assumptions 1(a) and 2(a). In the whole proof, x is fixed and we omit δ_x in the notation of the probability and expectation. For convenience, we also write $m(x, \mathbf{e}_n) := m(x, \mathbf{e}_n, \mathcal{X})$ and denote

$$g_n(x) := f(f_n(x)) - \mu_n(f).$$

Let us compute for $K \geq 1$,

$$\begin{aligned}
&\mathbb{E}_{\mathbf{e}_n}\left(Z_n(g_n)^2\right) \\
&\quad = \mathbb{E}_{\mathbf{e}_n}\Big(\sum_{|u|=|v|=n} g_n(X(u))g_n(X(v)) \\
&\quad = \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n \\ |u\wedge v|<K}} g_n(X(u))g_n(X(v)) \right) + \mathbb{E}\left(\sum_{\substack{|u|=|v|=n \\ |u\wedge v|\geq K}} g_n(X(u))g_n(X(v)) \right)
\end{aligned}$$

The second term of the right hand side is smaller than

$$2 \parallel f \parallel_\infty^2 \mathbb{E}(\#\{|u| = |v| = n : |u \wedge v| > K\}) \leq 2b^2 m(x, \mathbf{e}_n)^2.\epsilon_{K,n},$$

where $\limsup_{n\to\infty} \epsilon_{K,n} \to 0$ as $K \to \infty$ using the first part of Assumption 2(a). So we just deal with the first term and consider $i = 1, \cdots K$. Thanks to the branching property,

$$\begin{aligned}
&\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n \\ |u\wedge v|=i-1}} g_n(X(u))g_n(X(v)) \right) \\
&\qquad = \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|w|=i-1 \\ |wa|=|wb|=i}} \sum_{\substack{|u|=n \\ u\geq wa}} \sum_{\substack{|v|=n \\ v\geq wb}} g_n(X(u))g_n(X(v)) \right) \\
&\qquad = \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|w|=i-1 \\ |wa|=|wb|=i}} R_{i,n}(X(wa))R_{i,n}(X(wb)) \right),
\end{aligned}$$

where the many-to-one formula of Lemma 1 allows us to write

$$R_{i,n}(x) := \mathbb{E}_{T^i\mathbf{e}_n,\delta_x}\left(\sum_{|u|=n-i} g_n(X(u)) \right) = m_{n-i}(x, T^i\mathbf{e}_n)Q_{n-i}(x, T^i\mathbf{e}_n, g_n).$$

Assumption 1 (a) ensures that

$$F_{i,n}(u) := R_{i,n}(X(u))/m_{n-i}(X(wa), T^i\mathbf{e}_n)$$

goes to 0 a.s. for each $i \in \mathbb{N}, |u| = i$ uniformly for $f \in \mathcal{F}$. We also note that this quantity is bounded by b. Then,

$$\begin{aligned}
&m(x, \mathbf{e}_n, \mathcal{X})^{-2}\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n \\ |u\wedge v|\leq K}} g_n(X(u))g_n(X(v)) \right) \\
&= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{i\leq K, |w|=i-1 \\ |wa|=|wb|=i}} F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa), T^i\mathbf{e}_n)m_{n-i}(X(wb), T^i\mathbf{e}_n, \mathcal{X})}{m_n(x, \mathbf{e}_n)^2} \right).
\end{aligned}$$

As

$$F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa),T^i\mathbf{e}_n,\mathcal{X})m_{n-i}(X(wb),T^i\mathbf{e}_n)}{m_n(x,\mathbf{e}_n)^2}$$
$$\leq \sup_n \frac{m_{n-i}(X(wa),T^i\mathbf{e}_n)}{m_n(x,\mathbf{e}_n)}.\sup_n \frac{m_{n-i}(X(wb),T^i\mathbf{e}_n)}{m_n(x,\mathbf{e}_n)},$$

we can use the second part of Assumption 2(a) to conclude by bounded convergence. It yields the $L^2_{\mathbf{e}_n}$ convergence (10) under Assumptions 1(a) and 2(a), uniformly for $f \in \mathcal{F}$.

The proof of (10) under Assumptions 1(b) and 2(b) is almost the same, replacing K by $n-k_n$ with $k_n \to \infty$. Indeed, Assumption 1(b) ensures that there exists $k_n \to \infty$ such that

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_x}\left(\#\{|u|=|v|=n : u\wedge v > n-k_n\}\right)}{m_n(x,\mathbf{e}_n,\mathcal{X})^2} \overset{n\to\infty}{\longrightarrow} 0$$

whereas

$$\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{i\leq n-k_n,|w|=i-1\\|wa|=|wb|=i}} F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa),T^i\mathbf{e}_n)m_{n-i}(X(wb),T^i\mathbf{e}_n,\mathcal{X})}{m_n(x,\mathbf{e}_n)^2}\right)$$
$$\leq \left(\sup_{n-i\geq k_n,x\in\mathcal{X}} F_{i,n}(x)\right)^2 \frac{\mathbb{E}(Z_n(\mathcal{X})^2)}{m_n(x,\mathbf{e}_n)^2}.$$

Assumption 1(a) ensures that $\sup_{n-i\geq k_n,x\in\mathcal{X}} F_{i,n}(x) \to 0$ as $k_n \to \infty$ and the second part of Assumption 2(b) ensures that $\mathbb{E}(Z_n(\mathcal{X})^2)/m_n(x,\mathbf{e}_n)$ is bounded. The conclusion is thus the same.

The proof of the last part of the Theorem comes from Cauchy Schwartz inequality :

$$\mathbb{E}_{\mathbf{e}_n}\left(1_{Z_n(\mathcal{X})/m(x,\mathbf{e}_n)\geq\epsilon}\left[\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})}-\mu_n(f)\right]\right)^2$$
$$\leq \mathbb{E}_{\mathbf{e}_n}\left(\frac{m_n(x,\mathbf{e}_n)^2}{Z_n(\mathcal{X})^2}1_{Z_n(\mathcal{X})/m_n(x,\mathbf{e}_n)\geq\epsilon}\right)\mathbb{E}_{\mathbf{e}_n}\left(\left[\frac{f_n.Z_n(f)-Z_n(\mathcal{X})\mu_n(f)}{m_n(x,\mathbf{e}_n)}\right]^2\right).$$

So the first part of the theorem yields the second one. □

We give now a strong law of large numbers. For that purpose, we define

$$V_i(x_0,x_1)=\sup_{k\geq 0}\frac{m_k(x_0,T^i\mathbf{e},\mathcal{X})m_k(x_1,T^i\mathbf{e},\mathcal{X})}{m_{i+k}(x,\mathbf{e},\mathcal{X})^2}.$$

Lemma 2. *Let $\mathbf{e}\in E$, $x\in\mathcal{X}$ and assume that*

$$\sum_{n\geq 0} m_n(x,\mathbf{e},\mathcal{X})^{-1}<\infty;\quad \sum_{i\geq 1}\mathbb{E}\left(\sum_{\substack{|w|=i-1\\|wa|=|wb|=i}} V_i(X(wa),X(wb))\right)<\infty, \tag{11}$$

then $Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X})$ is bounded in $L^2_{\mathbf{e},\delta_\mathbf{x}}$.

The proof of this Lemma is given after the following main result.

Theorem 4 (Strong LLN)**.** *Let $\mathbf{e} \in E$, $x \in \mathcal{X}$. Assume that (11) hold. Assume also that there exist $\mathcal{F} \subset \mathcal{B}_b(\mathcal{X})$ and a sequence of probability measure μ_n on $\mathcal{X}$ such that*

$$\sup_{i\in\mathbb{N}, f\in\mathcal{F}} \sum_{n\geq i} \sup_{\lambda\in\mathcal{M}_1} \left|Q_{i,n}(\lambda, T^i\mathbf{e}, f\circ f_n) - \mu_n(f)\right|^2 < \infty. \tag{12}$$

Then uniformly for $f \in \mathcal{F}$

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x, \mathbf{e}_n, \mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_x} \ a.s.$$

The first assumption is related to the genealogy of the population and the second one is linked to the ergodic property of the auxiliary Markov chain Y. Both assumptions are stronger that their counterpart of the previous theorem.
We refer to [G07] for more general conditions on the functions $f \in \mathcal{F}$ in the fixed environment case, when the reproduction law does not depend on the position.
We note that under the Assumptions of the Theorem, $Z_n(\mathcal{X})/m_n(x, \mathbf{e}, \mathcal{X})$ is bounded in $L^2_{\mathbf{e}}$ thanks to the previous Lemma. It entails that the probability of the event $\{Z_n(\mathcal{X})/m_n(x, \mathbf{e}_n, \mathcal{X}) \geq \epsilon\}$ is positive for ϵ small enough and every $n \geq 1$. On this event, we get $f_n.Z_n(f)/Z_n(\mathcal{X}) - \mu_n(f) \to 0$ a.s. as $n \to \infty$.

Proof of Lemma 2. We omit still in the notations the initial state δ_x and write $m_n(x, \mathbf{e})$ for $m_n(x, \mathbf{e}, \mathcal{X})$. Using the branching property, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})^2) &= \mathbb{E}_{\mathbf{e}}\left(\sum_{|u|=|v|=n} 1\right) \\
&= \mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})) + \mathbb{E}_{\mathbf{e}}\left(\sum_{i\leq n} \sum_{\substack{|w|=i-1 \\ |wa|=|wb|=i}} \sum_{\substack{|u|=n: u>wa \\ |v|=n: v>wb}} 1\right) \\
&= m_n(x, \mathbf{e}) + \sum_{i\leq n} \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{|w|=i-1 \\ |wa|=|wb|=i}} m_{n-i}(X(wa), T^i\mathbf{e}) m_{n-i}(X(wb), T^i\mathbf{e})\right).
\end{aligned}$$

Then,

$$\frac{\mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})^2}{m_n(x, \mathbf{e})^2} \leq \frac{1}{m_n(x, \mathbf{e})} + \sum_{i\leq n} \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{|w|=i-1 \\ |wa|=|wb|=i}} V_i(X(wa), X(wb)\right),$$

which ends up the proof. □

Proof of Theorem 4. In this proof, we also omit the uniformity with respect to f. Let us prove that

$$\mathbb{E}_{\mathbf{e}}\left(\sum_{n\geq 1}\left[\frac{f_n.Z_n(f) - \mu(f)Z_n(\mathcal{X})}{m_n(x, \mathbf{e})}\right]^2\right) < \infty$$

to get that

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x,\mathbf{e})}$$

goes to 0 a.s. and conclude. We use the notations of the proof of the previous Theorem :

$$g_n(x) := f(f_n(x)) - \mu_n(f)$$

and follow the approach of [G07]. Using Fubini inversion, the branching property and the notations of the proof of Theorem 3, we have

$$\begin{aligned}
&\sum_{n\geq 0} m_n(x,\mathbf{e})^{-2}\mathbb{E}(Z_n(g_n)^2) \\
&= \mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}}\sum_{|u|=|v|=n} m_n(x,\mathbf{e})^{-2} g_n(X(u))g_n(X(v))\right) \\
&= \mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}}\sum_{i\leq n}\sum_{\substack{|u|=|v|=n\\ |u\wedge v|=i}} m_n(x,\mathbf{e})^{-2} g_n(X(u))g_n(X(v))\right) \\
&= \mathbb{E}_{\mathbf{e}}\left(\sum_{i\leq n}\sum_{\substack{|w|=i-1\\ |wa|=|wb|=i}}\sum_{\substack{|u|=n:u>wa\\ |v|=n:v>wb}} m_n(x,\mathbf{e})^{-2} g_n(X(u))g_n(X(v))\right) \\
&\qquad +\mathbb{E}\left(\sum_{n\in\mathbb{N},|u|=n} m_n(x,\mathbf{e})^{-2} g_n(X(u))^2\right) \\
&\leq \mathbb{E}\left(\sum_{i\leq n}\sum_{\substack{|w|=i-1\\ |wa|=|wb|=i}} \frac{m_{n-i}(X(wa),T^i\mathbf{e})m_{n-i}(X(wb),T^i\mathbf{e})}{m_n(x,\mathbf{e})^2} R_{i,n}(X(wa))R_{i,n}(X(wb))\right) \\
&\qquad + \parallel g_n \parallel_\infty \mathbb{E}\left(\sum_{n\in\mathbb{N}} m_n(x,\mathbf{e})^{-2} Z_n(\mathcal{X})\right) \\
&\leq \mathbb{E}\left(\sum_{\substack{i\in\mathbb{N},|w|=i-1\\ |wa|=|wb|=i}} V_i(X(wa),X(wb))H_i\right) + 2b\sum_{n\in\mathbb{N}} m_n(x,\mathbf{e})^{-1},
\end{aligned}$$

where

$$H_i = \sup_{y,z}\sum_{n\geq i} R_{i,n}(y)R_{i,n}(z), \qquad V_i(x_0,x_1) = \sup_{n\geq i}\frac{m_{n-i}(x_0,T^i\mathbf{e})m_{n-i}(x_1,T^i\mathbf{e})}{m_n(x,\mathbf{e})^2}$$

Then, the assumptions ensure that

$$\sum_{n\geq 0} m_n(x,\mathbf{e},\mathcal{X})^{-2}\mathbb{E}(Z_n(g_n)^2) \leq 2b\sum_{n\geq 0} m_n(x,\mathbf{e})^{-1}$$
$$+\sup_{i\in\mathbb{N}} H_i.\sum_{i\in\mathbb{N}}\mathbb{E}\left(\sum_{\substack{|w|=i-1\\|wa|=|wb|=i}} V_i(X(wa),X(wb))\right) < \infty$$

It ensures that $Z_n(g_n)/m_n(x,\mathbf{e}) \to 0$ a.s. and completes the proof. □

3.4 Applications

We now provide some applications of the previous results. First we recall that the previous results extend the law of large numbers for proportions of individuals in generation n obtained in [G07, DM10] for cellular aging or [BH13] for branching process in (stationary ergodic) random environment.
Now we provide some applications of the previous Theorems, where the assumptions can be more easily satisfied. They are motivated by the assumptions arising in reproduction-dispersion or cell division models.

3.4.1 Under Doeblin's conditions

We assume Doeblin's conditions on the mean measure

Assumption 3. There exist $M : E \to [1,\infty)$ such that for all $x \in \mathcal{X}, \mathbf{e} \in E$,

$$m(x,\mathbf{e},A) \leq M(\mathbf{e})m(y,\mathbf{e},A).$$

We could relax this assumption, for example by asking such an inequality for m_b instead of m, for some $b \geq 1$. We note This assumption hold if both $m(x,\mathbf{e},\mathcal{X})$ and $P(x,\mathbf{e},.)$ satisfy Doeblin's conditions. We also refer to [M13] for more general conditions in the-non homogeneous framework.
Let us denote

$$\sigma(\mathbf{e}) := \sup_{x\in\mathcal{X}}\mathbb{E}(N(x,\mathbf{e})^2), \qquad D(\mathbf{e}) := \frac{\sigma(\mathbf{e})M(\mathbf{e})M(T\mathbf{e})^2}{m(x,T\mathbf{e})}$$

to state the result.

Corollary 3. *Let $\mathbf{e} \in E, x \in \mathcal{X}$ and $f \in \mathcal{B}_b(\mathcal{X})$. We assume that Assumption 3 holds with*

$$\sum_{n\geq 1}\frac{1+D(T^{n-1}\mathbf{e})}{m_n(x,\mathbf{e},\mathcal{X})} < \infty, \quad \sum_{n\geq 0}\prod_{k=0}^{n}(1-1/M(T^k\mathbf{e})^2) < \infty. \tag{13}$$

Then, $Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X})$ is bounded in $L^2_{\mathbf{e},\delta_x}$ and

$$\frac{Z_n(f) - Z_n(\mathcal{X})Q_{0,n}(x,\mathbf{e},f)}{m_n(x,\mathbf{e},\mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_x} \quad a.s.$$

We note that the assumptions are fulfilled soon as m_n tends fast enough to ∞ (in particular with geometric convergence) and both σ, M, m are bounded. The proof here uses Theorem 4 and some additional lemmas allowing to check the assumptions required. We may derive an application of this result to the random environment framework directly or relax the assumptions to get only convergence in probability.
We note that the assumptions above ensure that $m_n(x, \mathbf{e}, \mathcal{X})$ goes to ∞ (supercriticality). Second, we recall that the fact that $Z_n(\mathcal{X})/m_n(x, \mathbf{e}, \mathcal{X})$ is bounded ensures that $Z_n(\mathcal{X}) \to \infty$ with positive probability. Using Paley-Sigmund inequality and a stronger assumption in the first part of (13) to ensure the uniformity with respect to the initial environment, one can prove that soon as $Z_n(\mathcal{X}) \to \infty$, $\liminf_{n\to\infty} Z_n(\mathcal{X})/m_n(x, \mathbf{e}, \mathcal{X}) > 0$. Then, one $\mathbb{P}_{\mathbf{e},\delta_x}$ a.s on the event $\{Z_n(\mathcal{X}) \to \infty\}$, $Z_n(f)/Z_n(\mathcal{X}) - Q_{0,n}(x, \mathbf{e}, f) \to 0$ as $n \to \infty$. One can derive an analogous result such from Theorem 2 soon as $\liminf_{n\to\infty} Z_{n+1}(\mathcal{X})/Z_n(\mathcal{X}) > 1$. To give a simple example, we mention the case of reproduction laws which are lower bounded by a supercritical offspring distribution.

Lemma 3. *Under Assumption 3, for all $x, y \in \mathcal{X}, \mathbf{e} \in E, n \geq 0, A \in \mathcal{B}_{\mathcal{X}}$,*

$$\frac{m_n(x, \mathbf{e})}{m_n(y, \mathbf{e})} \in [M(\mathbf{e})^{-1}, M(\mathbf{e})], \qquad Q_n(x, \mathbf{e}, A) \leq M(\mathbf{e})^2 Q_n(y, \mathbf{e}, A).$$

Proof. We have

$$\begin{aligned} m_n(x, \mathbf{e}) = \int_{\mathcal{X}} m_1(x, \mathbf{e}, dz) m_{n-1}(z, T\mathbf{e}) &\leq M(\mathbf{e}) \int_{\mathcal{X}} m_1(y, \mathbf{e}, dz) m_{n-1}(z, T\mathbf{e}) \\ &\leq M(\mathbf{e}) m_n(y, \mathbf{e}). \end{aligned}$$

Thus for all n, x, y,

$$m_n(x, \mathbf{e}) \leq M(\mathbf{e}) m_n(y, \mathbf{e})$$

and we obtain the first part of the Lemma. We then note that

$$\begin{aligned} Q_n(x, \mathbf{e}, A) &= \int_{\mathcal{X}} \frac{m_1(x, \mathbf{e}, dz)}{m_n(x, \mathbf{e})} m_{n-1}(z, T\mathbf{e}, A) \\ &\leq M(\mathbf{e})^2 \int_{\mathcal{X}} \frac{m_1(y, \mathbf{e}, dz)}{m_n(y, \mathbf{e})} m_{n-1}(z, T\mathbf{e}, A) \leq M(\mathbf{e})^2 Q_n(y, \mathbf{e}, A) \end{aligned}$$

to get the second part of the Lemma. □

Proof of Theorem 3. Using the branching property in generation i and the first part of Lemma 3, we have for all $x, y \in \mathcal{X}$,

$$m_{i+k}(x, \mathbf{e}) \geq m_i(x, \mathbf{e}) M(T^i e)^{-1} m_k(y, T^i \mathbf{e}).$$

Then

$$V_i(x_0, x_1) = \sup_{k\geq 0} \frac{m_k(x_0, T^i\mathbf{e}) m_k(x_1, T^i\mathbf{e})}{m_{k+i}(x, \mathbf{e})^2} \leq \frac{M(T^i\mathbf{e})^2}{m_i(x, \mathbf{e})^2}$$

and

$$\begin{aligned} \sum_{i\geq 1} \mathbb{E}_{\mathbf{e}} \left(\sum_{|w|=i-1} \sum_{|wa|=|wb|=i} V_i(X(wa), X(wb)) \right) &\leq \sum_{i\geq 1} \mathbb{E}_{\mathbf{e}}(Z_{i-1}(\mathcal{X})) \sigma(T^{i-1}\mathbf{e}) \frac{M(T^i\mathbf{e})^2}{m_i(x, \mathbf{e})^2} \\ &\leq \sum_{i\geq 1} \sigma(T^{i-1}\mathbf{e}) \frac{M(T^{i-1}(\mathbf{e}) M(T^i\mathbf{e})^2}{m(x, T^{i-1}\mathbf{e}) m_i(x, \mathbf{e})}. \end{aligned}$$

So (13) ensures that (11) is fulfilled and Lemma 2 ensures that $Z_n(\mathcal{X})/m_n(x,\mathbf{e})$ is bounded in $L^2_{\mathbf{e}}$. The first part of the Theorem is proved and we tackle now the a.s. convergence.

Using the second part of Lemma 3, we get the geometric ergodicity of $Q_{i,n}$:

$$|Q_{i,n}(\lambda,\mathbf{e},f)-Q_{i,n}(\mu,\mathbf{e},f)|\leq \| f \|_\infty \prod_{j=i}^{n}(1-1/M(T^j\mathbf{e})^2)$$

so that the second part of (13) ensures that (12) hold. Then Theorem 4 yields the expected a.s. convergence. □

3.4.2 Under Lyapounov's conditions

In the vein of Section 3.4.1, we can use Lyapounov functions to ensure geometric ergodicity and derive law of large numbers. One need to prove that $Q_{i,n}$ is geometric ergodic combining a Lyapounov function and Doeblin's assumption. It is the key point to apply Theorem 2 when the size of the population has a positive growth rate and the reproduction law is dominated. Some additional work is required to check the Assumptions of Theorem 4 to get ride of the positive growth rate assumption, as made in Section 3.4.1.

In the two next applications, we consider the case when the reproduction law does not depend on the trait, so $Q_{i,n}$ just depends on i and the auxiliary Markov chain with kernel Q_i is denoted by Y. We focus in these two examples the discussion on the functions f_n one can use to derive results.

3.4.3 Branching random walks

For branching random walks (possibly in varying environment in time and space), the auxiliary Markov chain Y is a random walk (possibly in varying environment in time and space). One way to get law of large numbers is to check some central limit theorem, i.e. the convergence in law

$$(Y_n-a_n)/b_n \Rightarrow W$$

for every initial state $x\in\mathcal{X}$. Then we can use Theorem 3 with $f_n(x)=(x-a_n)/b_n$ to obtain the asymptotic proportion of individuals whose trait x satisfies $f_n(x)\in[a,b]$. It is given by $\mathbb{P}(W\in[a,b])$ soon as $\mathbb{P}(W\in\{a,b\})=0$. .
We refer to [BH13] Section 3.4 for more details in this direction in the case when the reproduction law does not depend on the trait $x\in\mathcal{X}$ and the environment is stationary ergodic in time.

3.4.4 Kimmel's cell infection model and non ergodicity

In the Kimmel's branching model [B08] for cell division with parasite infection, the auxiliary Markov chain Y_n is a Galton-Waston in (stationary ergodic) random environment. For example, in the case when no extinction is possible, i.e. $\mathbb{P}_1(Y_1>0)$, under the usual integrability assumption we have

$$Y_n/\Pi_{i=0}^{n-1}m_i \overset{n\to\infty}{\longrightarrow} W\in(0,\infty) \qquad \text{a.s.}$$

where m_i are the successive mean number of offsprings and the distribution of W depends on the inital value Y_0. But

$$\log(Y_n)/n \overset{n\to\infty}{\longrightarrow} \mathbb{E}(\log m_0) \qquad \text{a.s.}$$

and the limit here does not depend on Y_0. So we get ergodic property and can use use Theorem 3 with $f_n(x) = \log(x)/n$. We obtain that the proportion of cells in generation whose number of parasites in between $[\exp([\mathbb{E}(\log m_0) - \epsilon]n), \exp([\mathbb{E}(\log m_0) + \epsilon]n)$ goes to 1 in probability.

Soon as the number of parasites in a cell can be equal to zero, i.e. $\mathbb{P}_1(Y_1 = 0) > 0$, ergodicity is failing and some additional work is needed. Using monotonicity argument, one may still conclude, see [B08] for an example.

4 Local densities and extremal particles.

We deal now with local densities, so we focus on the number of individuals whose trait belongs to some set A_n in generation n and their ancestral lineage.
We have proved the many-to-one formula

$$\mathbb{E}(Z_n(A_n)) = m_n(x, \mathbf{e}, \mathcal{X})Q_{0,n}(x, \mathbf{e}, A_n)$$

in the previous section. We have then checked that the ergodicity of $Q_{0,n}$ ensures that $Z_n(A)/m_n(x, \mathbf{e}, \mathcal{X}) - Q_{0,n}(x, \mathbf{e}, A)$ goes to zero under some conditions.
Now we wish to compare the asymptotic behaviors of $Z_n(A_n)$ and $m_n(x, \mathbf{e}, \mathcal{X})Q_{0,n}(x, \mathbf{e}, A_n)$, when $Q_{0,n}(x, \mathbf{e}, A_n) \to 0$ as $n \to \infty$. In particular, we are studying the links with the large deviations events of $Q_{0,n}$, i.e. exponential decrease of this quantity.
Such questions have been well studied for branching random walks from the pioneering work of Biggins [B77], and we refer to [R93] and to [R00, S08] for reviews on the topic. We mention also to [CP07, N11] for the random environment framework and to [DMS05] for large deviations Markov chain along tree with n vertices.
The upper bound for such results comes directly from Markov inequality and we are working on the lower bound. As usual, we could then derive the asymptotic behavior of the left most particle. It covers classical results for branching random walks on the speed of the extremal individual at the log scale. We provide some example motivated by cell's infection model, where the associated deviation strategy is more subtle. We mention also that $Z_n(A_n)$ may be negligible compared to $m_n(x, \mathbf{e}, \mathcal{X})Q_{0,n}(x, \mathbf{e}, A_n)$.

Definition 2. *For all $0 \le i \le n$, $A, B \in \mathcal{B}_{\mathcal{X}}$, we define the measure*

$$\mu_{i,n}(A, \mathbf{e}, B)[l, \infty) := \inf_{x \in A} \mathbb{P}_{\delta_x, T^i\mathbf{e}}(Z_{n-i}(B) \ge l)$$

We note $\bar{\mu}$ the mean of μ and $\hat{\mu}$ the variance of $\mu/\bar{\mu}$, so that

$$\bar{\mu}_{i,n}(A, \mathbf{e}, B) = \sum_{l \ge 1} \mu_{i,n}(A, \mathbf{e}, B)[l, \infty), \qquad \hat{\mu}_{i,n}(A, \mathbf{e}, B) = \frac{\sum_{l \ge 1} l^2 \mu_{i,n}(A, \mathbf{e}, B)\{l\}}{\bar{\mu}_{i,n}(A, \mathbf{e}, B)^2} - 1.$$

We start by coupling our process in the first stages by a particular branching process in varying environment to use both the convergence of the associated martingale during these first steps and a law of large number argument on the remaining time.

Lemma 4. *We assume that there exist $\phi, \psi : \mathbb{N} \to \infty$ with $\phi(n)$ non decreasing going to ∞ and*

$$0 = k_0 < k_1 < \cdots < k_{\phi(n)} = k_{0,n} < k_{1,n} < \cdots < k_{\psi(n),n} = n, \qquad B_i, B_{j,n} \subset \mathcal{X}$$

such that

$$\liminf_{i\to\infty} \bar{\mu}_{k_i,k_{i+1}}(B_i,\mathbf{e},B_{i+1}) > 1$$

and

$$\sum_{i\geq 0} \frac{\hat{\mu}_{k_i,k_{i+1}}(B_i,\mathbf{e},B_{i+1})}{\Pi_{i=0}^{n-1}\bar{\mu}_{k_i,k_{i+1}}(B_i,\mathbf{e},B_{i+1})} < \infty; \quad \sup_n \sum_{i=0}^{\psi(n)-1} \frac{\hat{\mu}_{k_{i,n},k_{i+1,n}}(B_{i,n},\mathbf{e},B_{i+1,n})}{\Pi_{i=0}^{n-1}\bar{\mu}_{k_{i,n},k_{i+1,n}}(B_{i,n},\mathbf{e},B_{i+1,n})} < \infty.$$

Then, there exists a set A, whose probability is positive and which does not depend on $(k_{i,n}, B_{i,n} : i, n \geq 0)$, such that

$$A \subset \left\{ \liminf_{n\to\infty} \frac{Z_n(B_{n,\psi(n)})}{P_n} > 0 \right\},$$

where

$$P_n := \prod_{i=0}^{\phi(n)-1} \bar{\mu}_{k_i,k_{i+1}}(B_i,\mathbf{e},B_{i+1}). \prod_{i=0}^{\psi(n)-1} \bar{\mu}_{k_{i,n},k_{i+1,n}}(B_{i,n},\mathbf{e},B_{i+1,n})$$

Proof. We use a coupling of the branching Markov chain Z with a supercritical Branching Process in Varying Environment (BPVE). It is obtained by selecting the individuals whose lineage lives in the tube $(B_i : i \leq \phi(n), B_{j,n} j \leq \psi(n))$. More prisely we consider the subpopulation of Z constructed recursively by keeping the descendance of the population in generation k_i whose trait belongs to B_i for $i \leq \phi(n)$ and then whose trait belongs to $B_{j,n}$ for $j \leq \psi(n)$. The size of the population obtained by this construction in generation k_i is a.s. larger than a branching process N_i whose reproduction law in generation i is $\mu_i := \mu_{k_i,k_{i+1}}(B_i,\mathbf{e},B_{i+1})$. Similarly, the size of the population in generation $k_{\psi(n),n}$ is larger than a branching process in varying environment, with initial value equal to $N_{\phi(n)}$ and successive reproduction law $\mu_{j,n} := \mu_{k_{j,n},k_{i+1,n}}(B_{j,n},\mathbf{e},B_{j+1,n})$ jor $j = 0,\cdots,\psi(n)$. Thus,

$$Z_n(B_{n,\psi(n)}) \geq \sum_{j=1}^{N_{\phi(n)}} U_{j,n}$$

where $U_{j,n}$ is distributed as a BPVE in generation $\psi(n)$, denoted by U_n, whose successive reproduction laws are $\mu_{k_{i,n},k_{i+1,n}}(B_{i,n},\mathbf{e},B_{i+1,n})$ for $i = 0,\cdots,\psi(n)-1$. Moreover $(U_{j,n} : j = 0,\dots\psi(n),)$ are independent by branching property.
By orthogonality ansd using the assumption

$$\sum_{i\geq 0} \frac{Var(\mu_i/\bar{\mu_i})}{\Pi_{j=0}^{i-1}\bar{\mu_i}} < \infty,$$

the martingale

$$\frac{N_i}{\prod_{j=0}^{i-1}\bar{\mu}_j}$$

has a finite positive limit W on the survival event $A = \{\forall n, N_n > 0\}$. Recalling that $\liminf_{i\to\infty}\bar{\mu}_i > 1$ by assumption, A has positive probability and conditionally on that event

$$\liminf_{i\to\infty} \frac{N_{i+1}}{N_i} > 1.$$

Similarly $U_{j,n}/\mathbb{E}(U_{j,n})$ is bounded in $L^2_{\mathbf{e}}$ using the second part of the moment assumptions of the Lemma. Thus

$$X_{j,n} \stackrel{d}{=} \frac{U_{j,n} - \mathbb{E}(U_n)}{\mathbb{E}(U_n)}$$

are independent random variable independent of $(N_i : i = 0, \cdots, \phi(n))$ and are bounded in $L^2_{\mathbf{e}}$. We can then apply Lemma 1 in [AK98a] to

$$\frac{1}{N_{k_{\phi(n)}}} \sum_{j=1}^{N_{k_{\phi(n)}}} X_{j,n}$$

to get that this quantity goes a.s. to 0 as $n \to \infty$. Then

$$Z_n(B_{n,\psi(n)}) \geq N_{k_{\phi(n)}} \mathbb{E}(U_n).[1 + \epsilon_n]$$

where $\epsilon_n \to 0$. Finally, we use

$$\mathbb{E}(N_{k_{\phi(n)}}) = \prod_{i=0}^{\phi(n)-1} \bar{\mu}_{k_i,k_{i+1}}(B_i, \mathbf{e}, B_{i+1}), \quad \mathbb{E}(U_n) = \prod_{i=0}^{\psi(n)-1} \bar{\mu}_{k_{i,n},k_{i+1,n}}(B_{i,n}, \mathbf{e}, B_{i+1,n})$$

to get

$$\liminf_{n\to\infty} \frac{Z_n(B_{n,\psi(n)})}{P_n} \geq \liminf_{n\to\infty} \frac{N_{\phi(n)}}{\mathbb{E}(N_{\phi(n)})}(1 + \epsilon_n) \geq W.$$

Recalling that $A = \{W > 0\}$ has positive probability ends up the proof. □

4.1 Monotone Branching Markov chain

Our aim is to see the local densities in terms of the large deviations of the auxiliary process and the way this large deviation event is achieved. First, let us derive from the previous Lemma a result in the monotone case for the event $[a_n, \infty)$, which yield the applications for the cell models and branching random walks which initially motivated these questions.
Thus, by now, we assume that $\mathcal{X}$ is totally ordered by $\leq$ and

Assumption 4 (Monotonicity)**.** For all $x \leq y$, $\mathbf{e} \in E$ and $a \in \mathcal{X}$, we have

$$\mathbb{P}_{\delta_x,\mathbf{e}}(Z_1([a,\infty)) \geq l) \leq \mathbb{P}_{\delta_y,\mathbf{e}}(Z_1([a,\infty)) \geq l) \qquad (l \geq 0).$$

Assumption 5 (Mean growth rate)**.** Let $\rho > 0$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log m_n(x, \mathbf{e}, [a_n, \infty)) = \rho.$$

Moreover, there exist $p \geq 1$ and $b_i \in \mathcal{X}$ such that $x \geq b_0$ and

$$\liminf_{i\to\infty} m_p(b_i, T^{ip}\mathbf{e}, [b_{i+1}, \infty)) > 1.$$

Finally, for every $\epsilon > 0$, there exist $q = q(\epsilon)$, $\phi(n) \to \infty$ and $(b_{j,n} : j, n \geq 0)$ such that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{j<(n-\phi(n)p)/q} \log m_q(b_{j,n}, T^{i\phi(n)+jq}\mathbf{e}, [b_{j+1,n}, \infty)) \geq \rho - \epsilon.$$

The values $(b_i : i \leq \phi(n), b_{j,n} : j \leq \psi(n))$ correspond to the (lower) line of the trait of the subpopulation which realize the population $Z_n([a_n, \infty))$ in generation n. This line is a straight line for branching random walk or for $a_n = 1$ in the Kimmel's branching model [B08], see below. But the line is not straightforward for the other quantities of interest in Kimmel's branching model, such as the large deviations associated to $a_n \to \infty$. An other motivating example when the line is not straightforward is given by large deviations event realized in one step of the process. It can be the case for random walks with heavy tails or autoregressive processes. For that reasons, we have split the line in two parts in the previous results.

Theorem 5. *Let $\mathbf{e} \in E$ and $x \in \mathcal{X}$. Under the Assumptions 4, 5 and*

$$\sup\left\{\mathbb{E}(N(z, T^k\mathbf{e})^2) : z \in \mathcal{X}, k \geq 0\right\} < \infty,$$

we have

$$\frac{1}{n}\log Z_n([a_n, \infty)) \overset{n\to\infty}{\longrightarrow} \rho$$

with positive probability under $\mathbb{P}_{\mathbf{e},\delta_x}$.

Letting the initial population goes to infinity in this statement allows to get the result a.s. by branching property. Getting the result a.s. requires additional assumptions.

Proof. As for branching random walks, the upper bound comes directly from Markov inequality. For every $\eta > 0$,

$$\mathbb{P}_{x,\mathbf{e}}(Z_n([a_n, \infty)) \geq \exp((\rho + \eta)n)) \quad \leq \quad \exp(-(\rho + \eta)n) m_n(x, \mathbf{e}, [a_n, \infty)),$$

so that the first part of the Assumption 5 ensures that the right hand side is summable. Then Borel-Cantelli lemma yields the a.s. upper bound.

The lower bound from the previous Lemma with

$$k_i = ip, \quad k_{n,j} = \phi(n)p + jq, \qquad B_j = [b_j, \infty), \quad B_{j,n} = [b_{j,n}, \infty), \quad \psi(n) = [(n - \phi(n)p)/q],$$

where $[x]$ is the smallest integer larger or equal to x. By the monotonicity Assumption 4

$$\mu_{k_j,k_{j+1}}(B_j, \mathbf{e}, B_{j+1})(.) := \mathbb{P}_{\delta_{b_j}, T^{k_j}\mathbf{e}}\big(Z_{k_{j+1}-k_j}([b_{j+1}, \infty)) = .\big)$$

and the definition of $\mu_{k_{j,n},k_{j+1,n}}(B_{j,n}, \mathbf{e}, B_{j+1,n})(.)$ is analoguous. So

$$\bar{\mu}_{k_j,k_{j+1}}(B_j, \mathbf{e}, B_{j+1}) = m_{k_{j+1}-k_j}(b_j, T^{k_j}\mathbf{e}, [b_{j+1}, \infty)) = m_p(b_j, T^{jp}\mathbf{e}, [b_{j+1}, \infty)).$$

and the analoguous identitie hold for $\bar{\mu}_{k_{j,n},k_{j+1,n}}(B_{j,n}, \mathbf{e}, B_{j+1,n})$ By Assumption 5, we have for $\epsilon \in (0, \rho)$,

$$\liminf_{j\to\infty} \bar{\mu}_{k_j,k_{j+1}}(B_j, \mathbf{e}, B_{j+1}) > 1, \ \liminf_{n\to\infty} \frac{1}{n}\log(\Pi_{j=0}^{\psi(n)-1} \bar{\mu}_{k_{j,n},k_{j+1,n}}(B_{j,n}, \mathbf{e}, B_{j+1,n})) \geq \rho - \epsilon > 0.$$

Recalling that $\sup\{\mathbb{E}(N(x, T^k\mathbf{e})^2) : x \in \mathcal{X}, k \geq 0\} < \infty$ is assumed, we get

$$\sum_{i\geq 0} \frac{\hat{\mu}_{k_i,k_{i+1}}(B_i, \mathbf{e}, B_{i+1})}{\Pi_{j=0}^{i-1}\bar{\mu}_{k_j,k_{j+1}}(B_i, \mathbf{e}, B_{j+1})} < \infty; \quad \sup_n \sum_{i=0}^{\psi(n)-1} \frac{\hat{\mu}_{k_{i,n},k_{i+1,n}}(B_{i,n}, \mathbf{e}, B_{i+1,n})}{\Pi_{j=0}^{i-1}\bar{\mu}_{k_{j,n},k_{j+1,n}}(B_{j,n}, \mathbf{e}, B_{j+1,n})} < \infty.$$

Thus, we can apply Lemma 4 and get

$$A \subset \left\{ \liminf_{n\to\infty} \frac{Z_n([a_n,\infty))}{\prod_{i=1}^{\psi(n)} \bar{\mu}_{k_{i,n},k_{i+1,n}}(B_{i,n},\mathbf{e},B_{i+1,n})} > 0 \right\} \subset \left\{ \liminf_{n\to\infty} \frac{1}{n} \log Z_n([a_n,\infty)) \geq \rho - \epsilon \right\} \geq \eta$$

Noting that A is fixed when $\epsilon \to 0$ ends up the proof. □

As expected, we can now precise the asymptotic behavior of the extremal individuals. If $a_n(x)$ satisfies the assumptions of Theorem 5 with some rate $\rho(x)$, then, for every x such that $\rho(x) > \log m$,

$$\limsup_{n\to\infty} \frac{\max\{X(u) : |u| = n\}}{a_n(x)} \leq 1 \qquad \mathbb{P}_{\delta_x,\mathbf{e}} \text{ a.s.}$$

and for every x such that $\rho(x) < \log m$,

$$\liminf_{n\to\infty} \frac{\max\{X(u) : |u| = n\}}{a_n(x)} \geq 1 \qquad \mathbb{P}_{\delta_x,\mathbf{e}} \text{ a.s.}$$

on some event with positive probability.
The proof is standard. The first part comes directly from Borel-Cantelli Lemma, recalling that

$$\mathbb{P}_{\delta_x,\mathbf{e}}(R_n \geq a_n(x)) \leq \mathbb{E}_{\delta_x,\mathbf{e}}(Z_n([a_n(x),\infty))) = m(x,\mathbf{e},[a_n(x),\infty))$$

decreases exponentially with rate $\rho - \rho(x)$. The second part comes from the Theorem 5 which ensures that there are many particles beyond $a_n(x)$.

4.2 Monotone Markov chain indexed by branching trees

Let us specify in a simpler framework the results above, more precisely the link between the local densities and the large deviations of the auxiliary chain. We assume here that the reproduction law does not depend on the trait of the individual, so that

$$N(\mathbf{e}) := N(x,\mathbf{e}), \quad m(\mathbf{e}) := m(x,\mathbf{e}), \qquad m_n(\mathbf{e}) := m_n(x,\mathbf{e}) = \prod_{i=0}^{n-1} m(T^i\mathbf{e}). \tag{14}$$

As above, we require the monotonocity of the trait distribution : assume :

Assumption 6 (Monotonicity of P)**.** For all $x \leq y$, $\mathbf{e} \in E$ and $a \in \mathcal{X}$, we have

$$P(x,\mathbf{e},[a,\infty)) \leq P(y,\mathbf{e},[a,\infty)).$$

We assume also that the large deviations of $Q_{i,n}$ beyond a_n occur with rate $\alpha > 0$ and that the beginning of the associated trajectory is supercritical, i.e.

Assumption 7 (Large deviations of the auxiliary process Q)**.** We have

$$\lim_{n\to\infty} \frac{1}{n} \log Q_{0,n}(x,\mathbf{e},[a_n,\infty)) = -\alpha$$

Moreover, we assume that exists $p \geq 1$ and $b_i \in \mathcal{X}$ such that

$$\liminf_{i\to\infty} m_p(T^{ip}\mathbf{e}) Q_p(b_i, T^{ip}\mathbf{e}, [b_{i+1},\infty)) > 1$$

and that for every $\epsilon > 0$, there exist $q = q(\epsilon)$, $\phi(n) \to \infty$ and $(b_{j,n} : j, n \geq 0)$ such that

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{j<(n-\phi(n)p)/q} \log Q_q(b_{j,n}, T^{i\phi(n)+jq}\mathbf{e}, [b_{j+1,n},\infty)) \geq -\alpha - \epsilon.$$

These assumptions are satisfied for the applications we have in mind. For an example of large deviations following Assumption 7, a sufficient condition is $P_{a_n}(Y_n \geq a_n + b_n) \sim \mathbb{P}_0(Y_n \geq b_n)$. The trajectory associated to the large deviation event is then straight line and $k_{i,n} = k_i$ works. It holds for random walks and more generally for random walks in random environment under general moment assumptions.

Corollary 4. *Let $\mathbf{e} \in E$ and $x \in \mathcal{X}$. Assuming (14), 6 and 7, we have*

$$\frac{1}{n} \log \left(Z_n([a_n, \infty)) / m_n(\mathbf{e}) \right) \overset{n\to\infty}{\longrightarrow} -\alpha$$

with positive probability under $\mathbb{P}_{\mathbf{e},\delta_x}$.

As expected, the large deviation of the auxiliary Markov chain quantifies the lost of growth α of the size of the population beyond a_n, $Z_n([a_n, \infty)$, compared to the whole growth of the population given by $m_n(\mathbf{e})$.

4.3 Applications

4.3.1 Kimmel's branching model

We refer to [B08] for a complete description of the model and the motivations. The population of cells is a binary tree and the number of parasites is the trait. The auxiliary Markov process Y is then a branching process in random environment. Monotonicity (Assumption 6) is a direct consequence of the branching property of Y and the problem is reduced to check Assumption 7.

One of the motivating question in [B08] is to count the number of infected cells in the subcritical case, which means that Y is a.s. absorbed in finite time. Three regimes appear in the subcritical case [1] and in particular in the weak subcritical case

$$\mathbb{P}(Y_n > 0) \sim c n^{-3/2} \gamma^n$$

where $\gamma < \mathbb{E}_1(Y_1)$. The mean number of infected cells is equal to $2^n \mathbb{P}(Y_n > 0)$ and obtaining a.s. results was left open is this regime. Corollary 4 ensures that the number of infected cells in generation n, N_n^* satisfies

$$\frac{1}{n} \log(N_n^*) \overset{n\to\infty}{\longrightarrow} \log(2\gamma) \qquad \text{p.s.}$$

soon as $2\gamma > 0$. Indeed, the corollary is applied for p large enough such that $2^p \mathbb{P}_1(Y_p > 0) > 1$, $a_n = [1, \infty]$, $b_i = 1, b_{j,n} = 1$, $\phi(n) = o(n)$ and q is chosen such that

$$\log \mathbb{P}_1(Y_q > 0) \geq q \log(\gamma) - \epsilon.$$

Second, when counting the number of cells infected less than the typical cell in the supercritical regime, the problem is now linked to the lower large deviation of branching processes in random environment Y_n, i.e. to

$$\mathbb{P}(1 \leq Y_n \leq \exp(n\theta)), \qquad \text{where } \theta < \mathbb{E}(\log m(\mathcal{E})))$$

and the way this large deviation event is realized. We refer to [BB12] for the results. Here again Corollary 4 allows to determine the a.s. behavior of the number of cells whose number of parasites is between 1 and $\exp(\theta n)$. It is worth noting that for this question the associated trajectory is not straight and $k_{i,n}$ depends on n.

4.3.2 Branching random walks

We can recover here the classical result on the asymptotic behavior of

$$\frac{1}{n}\log Z_n[an,\infty)$$

for a branching random walk with random increment X. It converges a.s. to

$$\log(m)-\Lambda(a)$$

soon as $a \geq \mathbb{E}(X)$ and $\log(m) > \Lambda(a)$, where Λ is the rate function associated to the random walk $S=\sum_{i=0}^{n-1} X_i$, see e.g. [R00, S08]. One can extend this result to offsprings distribution in time varying environment and random walks in varying environment using the last Corollary and large deviations of random walks in varying environment. Here $b_i = aip$, $b_{j,n} = ajp + b_{\phi(n)}$, $\phi(n) = o(n)$. We refer in particular to [Z04] for results on quenched and annealed large deviations of random walk in random envrionment.

4.3.3 Perspectives

A main motivation for this work is the control of local densities in cell division models for aging [G07, DM10], for damages [ES07] or infection such as Kimmel's branching model already mentioned. An other one coming from ecology is the role of time and space inhomogenity. We aim at investigating further these questions and determine the behavior of extremal particles in these models, which present different large deviation's trajectories.

Acknowledgement. This work was partially funded by Chair Modelisation Mathematique et Biodiversite VEOLIA-Ecole Polytechnique-MNHN-F.X., the professorial chair Jean Marjoulet, the project MANEGE 'Modèles Aléatoires en Écologie, Génétique et Évolution' 09-BLAN-0215 of ANR (French national research agency).

References

[AS10] E. Aïdékon, Z. Shi (2010). Weak convergence for the minimal position in a branching random walk: a simple proof. *Periodica Mathematica Hungarica* 61 (2010) 43-54.

[AK98a] K. Athreya, H.J. Kang (1998). Some limit theorems for positive recurrent branching Markov chains I. *Adv. Appl. Prob.* 30(3). 693-710.

[AK98b] K. Athreya, H.J. Kang (1998). Some limit theorems for positive recurrent branching Markov chains II. *Adv. Appl. Prob.* 30(3). 711-722.

[AK71] K. B. Athreya, S. Karlin (1971). On branching processes with random environments II : Limit theorems. *Ann. Math. Stat.* 42. 1843-1858.

[A00] K. B. Athreya (2000). Change of measure of Markov chains and the $L \log L$ theorem for branching processes. *Bernoulli* Vol. 6, No 2, 323-338.

[B08] V. Bansaye. Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Annals of Applied Probability, Vol 18, Number 3, 967-996* (2008).

[BB12] Lower large deviations for supercritical branching processes in random environment (2012), with Christian Boeinghoff. To appear in the *Proc. of Steklov Institute of Mathematics*.

[BDMT11] V. Bansaye, J.-F. Delmas, L. Marsalle and V.C. Tran (2011). Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *Ann. Appl. Probab.* Vol. 21, No. 6, 2263-2314.

[BL12] V. Bansaye, A. Lambert (2012). New approaches of source-sink metapopulations decoupling the roles of demography and dispersal. To appear in *Theor. Pop. Biology*

[BH13] V. Bansaye, C. Huang (2013). Weak law of large numbers for some Markov chains along non homogeneous genealogies. *Preprint* avialable via Arxiv.

[B77] J.D. Biggins (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14**, 25-37.

[C11] B. Cloez (2011). Limit theorems for some branching measure-valued processes. *Avialable via* http://arxiv.org/abs/1106.0660.

[CRW91] B. Chauvin, A. Rouault and A. Wakolbinger (1991). Growing conditioned trees. *Stochastic Processes and their Applications* **39**, 117–130.

[CP07] F. Comets, S. Popov (2007). Shape and local growth for multidimensional branching random walks in random environment. *ALEA* **3**, 273-299.

[CY11] F. Comets, N. Yoshida (2011). Branching random walks in space-time random environment: survival probability, global and local growth rates. *Journal of Theor. Probab.*

[DZ98] A. Dembo, O. Zeitouni (1998). *Large deviations techniques and applications*. Applications of Mathematics (New York) 38 (Second edition ed.)

[DMS05] A. Dembo, P. Mörters, S. Sheffield (2005). Large deviations of Markov chains indexed by random trees. *Ann. Inst. H. Poincaré Probab. Statist.* 41, no. 6, 971?996.

[DM10] J.-F. Delmas, L. Marsalle (2010). Detection of cellular aging in a Galton-Watson process. *Stochastic Processes and their Applications* **120**, 2495–2519.

[E07] J. Engländer (2007). Branching diffusions, superdiffusions and random media. *Prob. Surveys.* **4**, 303-364.

[EK86] S.N. Ethier, T.G. Kurtz (1986). *Markov Processus, Characterization and Convergence*. John Wiley & Sons, New York.

[ES07] S.N. Evans, D. Steinsaltz (2007). Damage segregation at fissioning may increase growth rates: A superprocess model. *Theoretical Population Biology* **71**, 473-490.

[F71] W. Feller (1971). *An introduction to probability theory and its applications*, volume 1 and 2. Wiley.

[FK60] H. Furstenberg, H. Kesten (1960). Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457-469.

[GMPV10] N. Gantert, S. Müller, S. Popov, M. Vachkovskaia (2010). Survival of branching random walks in random environment. *Journal of Theor. Probab.*, **23**, 1002-1014.

[1] J. Geiger, G. Kersting, V. A. Vatutin (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. Henri Poincaré (B).* **39**, pp. 593–620.

[G99] J. Geiger (1999). Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Prob. 36*, 301-309.

[GB03] H.O. Georgii, E. Baake (2003). Supercritical multitype branching processes: the ancestral types of typical individuals. *Adv. in Appl. Probab.*, Vol. 35, No 4, 1090-1110.

[GRW92] L. G. Gorostiza, S. Roelly, A. Wakolbinger (1992). Persistence of critical multitype particle and measure branching processes. *Probability Theory and Related Fields*, Vol. 92, No 3, 313-335.

[G07] J. Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.* **17**, 1538–1569 (2007).

[HR12] S. C. Harris, M. I. Roberts (2012). The many-to-few lemma and multiple spines. *Avialable via* http://arxiv.org/abs/1106.4761.

[HR13] S. C. Harris, M. I. Roberts (2013). A strong law of large numbers for branching processes: almost sure spine events *Avialable via* http://arxiv.org/abs/1302.7199.

[HL11] C. Huang, Q. Liu (2011). Branching random walk with a random environment in time. *Preprint.*

[HS09] Y. Hu, Z. Shi (2009). Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.* 37 742-781.

[JN96] P. Jagers, O. Nerman (1996). The asymptotic composition of supercritical multitype branch- ing populations. In *Séminaire de Probabilités*, XXX, volume 1626 of Lecture Notes in Math., pages 40-54. Springer, Berlin, 1996.

[K74] N. Kaplan (1974). Some Results about Multidimensional Branching Processes with Random Environments. *Ann. Probab.*, Vol. 2, No. 3., 441–455.

[KLPP97] T. Kurtz, R. Lyons, R. Pemantle, Y. Peres (1997). A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes*, ed. K. B. Athreya and P. Jagers. Springer, New York. 181-185.

[LPP95] R. Lyons, R. Pemantle, Y. Peres (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.*, Vol. 23, No 3, 1125-1138.

[M13] F. Mukhamedov (2013). Weak ergodicity of nonhomogeneous Markov chains on noncommutative L1-spaces. *Banach J. Math. Anal.* Vol. 7, No. 2, 53 -73.

[M08] S. Müller (2008). A criterion for transience of multidimensional branching random walk in random environment. *Electr. Jour. Probab.* **13**, 1189-1202.

[N11] M. Nakashima (2011). Almost sure central limit theorem for branching random walks in random environment. *Ann. Appl. Probab.*, **21**(1), 351-373.

[NJ84] O. Nerman, P. Jagers (1984). The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete* 65 , no. 3, 445-460.

[R93] A. Rouault (1993). Precise estimates of presence probabilities in the branching random walk. *Stoch. Process. Appl.* 44, no. 1, 27-39.

[R00] A. Rouault (2000). Large deviations and branching processes. *Proceedings of the 9th International Summer School on Probability Theory and Mathematical Statistics* (Sozopol, 1997). Pliska Stud. Math. Bulgar. 13, 15-38.

[S94] T. Seppäläinen (1994). Large deviations for Markov chains with Random Transitions. *Ann. Prob.* **22** (2), 713-748.

[S08] Z. Shi (2008). *Random walks and trees.* Lecture notes, Guanajuato, Mexico, November 3-7.

[T88] D. Tanny (1988). A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means. *Stoch. Process. Appl.* 28, no. 1, 123-139.

[Y08] N. Yoshida (2008). Central limit theorem for random walk in random environment. *Ann. Appl. Probab.* **18** (4), 1619-1635.

[Z04] O. Zeitouni (2004). Random walks in random environment. *XXXI Summer school in probability, St Flour* (2001). Lecture notes in Math. 1837 (Springer) (2004), pp. 193?312.