Ancestral lineages and limit theorems for branching Markov chains in varying environment

Vincent Bansaye *

Abstract

We consider branching processes in discrete time for structured population in varying environment. Each individual has a trait, which belongs to some general state space and both the reproduction law and the trait inherited by the offsprings may depend on the trait of the mother and the environment. We study in this paper the long time behavior of the population and the ancestral lineage of typical individuals under general assumptions. We focus on the mean growth rate and the trait distribution among the population. A key role is played by well chosen (possibly non-homogeneous) Markov chains and the approach relies on many-to-one formulae and the analysis of the genealogy. The applications use large deviations principles or the Harris ergodicity for these auxiliary 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, which means a multitype branching process whose number of types may be infinite. The environment may evolve in time and influence the whole population. Before the general definitions, we give an informal description of the model and explain our motivations. In population dynamics and populations genetics or in ecology, individual behaviors can be affected by genetic or phenotypic traits and the environment. Each individual will be characterized by a trait x taking values in a trait space $\mathcal{X}$, which will be typically a subspace of $\mathbb{R}^d$ giving the size, the age, the position... of the individual. The environment in the current generation can consist in abiotic parameters (temperature, humidity, wind...). Both the trait and the environment influence the reproduction law of the individual (i.e. the number of offsprings) and the distribution of the traits of the offsprings.

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

We provide here general statements on the long time behavior relying on auxiliary Markov chains. This approach allows to get powerful probabilistic tools and avoid the spectral approach to consider models where the eigenvector associated to the maximal eigenvalue of the mean operator is degenerated (see the example of Kimmel's model in the last section) and models in varying environment.
The number of offsprings of an individual u is denoted by $N(u)$ and its traits is noted $X(u) \in \mathcal{X}$. In each generation n, the set of individuals is a random set $\mathbb{G}_n$ and each individual $u \in \mathbb{G}_n$ behaves independently. One individual with trait x living in environment $\mathbf{e} \in E$ gives birth to $N(x, \mathbf{e})$ individuals. Conditionally on $N(x, \mathbf{e}) = k$, the distribution of the k traits of the offspring's is distributed as $P^{(k)}(x, \mathbf{e}, .)$.

In the whole paper, we assume that $\mathcal{X}$ and E are measurable spaces endowed with their σ-algebra $\mathcal{B}_{\mathcal{X}}$ and $\mathcal{B}_E$) respectively. Moreover $T : E \to E$ a measurable application providing the environment dynamics. For each $k \geq 1$ 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 $x \in \mathcal{X}$, $P^{(k)}(x, \mathbf{e}, .)$ is a probability measure on $(\mathcal{X}^k, (\mathcal{B}_{\mathcal{X}})k)$.
b) For each $A \in (\mathcal{B}_{\mathcal{X}})^k$, $P^{(k)}(., \mathbf{e}, A)$ is a $\mathcal{B}_{\mathcal{X}}$ measurable function.

We use the classical Ulam-Harris-Neveu notations for discrete trees. Each individual in the population is an element of

$$\mathcal{U} = \{\varnothing\} \cup \bigcup_{n \geq 1} (\mathbb{N}^*)^n$$

and is denoted by $u = u_1 u_2 \cdots u_n$ with $u_i \in \mathbb{N}^* = \{1, 2, \dots\}$. Thus $\varnothing$ is the root of the tree and $u = u_1 u_2 \cdots u_n$ is the u_n'th child of the u_{n-1}'th child of the $\dots u_1$'st child of the root. We denote by $|u| = n$ the generation of the individual u. 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, we write $u \leq v$ if u is an ancestor of v, i.e. $\exists w \in \mathcal{U}$ such that $v = uw$. Finally, we denote by $u \wedge v$ the nearest common ancestor of u and v. This latter is the element $w \in \mathcal{U}$ such that $w \leq u$ and $w \leq v$ whose generation $|w|$ is maximal.
For any generation, each individual with trait $x \in \mathcal{X}$ which lives in environment $\mathbf{e} \in E$ gives birth independently to a random number of offsprings, whose law both depend on x and $\mathbf{e}$. This number of offsprings is distributed as an integer valued random variable $N(x, \mathbf{e})$ whose mean is denoted by

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

In the whole paper, we assume that $m(x, \mathbf{e}) \in (0, \infty)$ for each $x \in \mathcal{X}, \mathbf{e} \in E$.

Let us consider a family of independent integer valued random variables $(N(u, x, \mathbf{e}) : u \in \mathcal{U}, x \in \mathcal{X}, \mathbf{e} \in E)$ such that for any $u \in \mathcal{U}$, $x \in \mathcal{X}$ and $\mathbf{e} \in E$, $N(u, x, \mathbf{e})$ is distributed as $N(x, \mathbf{e})$. The branching Markov chain $(X(u) : u \in \mathbb{G}_n)_n$ in environment $\mathbf{e}$ starting from one single individual with trait $x \in \mathcal{X}$ is defined recursively as follows under the probability $\mathbb{P}_{\mathbf{e}, \delta_x}$:
i) $\mathbb{G}_0 = \{\varnothing\}$ and $X(\varnothing) = x$.
ii) For each $n \geq 0$ and $u \in \mathbb{G}_n$, we write $N_{\mathbf{e}}(u) = N(u, X(u), T^n \mathbf{e})$ and we set

$$\mathbb{G}_{n+1} = \{ui : u \in \mathbb{G}_n, 1 \leq i \leq N_{\mathbf{e}}(u)\}$$

and for any collection $(A_v : v \in \mathcal{U})$ of measurable subsets of $\mathcal{X}$,

$$\mathbb{P}_{\mathbf{e},\delta_x}\left(\bigcap_{u\in\mathbb{G}_n, i\le N_{\mathbf{e}}(u)} \{X(ui)\in A_{ui}\} \,\Big|\, (X(u): u\in\mathbb{G}_m, m\le n)\right) = \prod_{u\in\mathbb{G}_n} P^{(N_{\mathbf{e}}(u))}\left(X(u), T^n\mathbf{e}, A_{u1}\times\ldots\times A_{uN_{\mathbf{e}}(u)}\right)$$

Under $\mathbb{P}_{\mathbf{e}}$, the branching Markov chain $X = (X(u) : u \in \mathbb{G}_n)_n$ is a multitype branching process in the varying environment $\mathbf{e}$ where the type takes values in $\mathcal{X}$ and between generation n and $n+1$ for $n \in \mathbb{N}$, the individuals live in environment $T^n\mathbf{e}$. Let us describe an example which cover our motivations and applications for modeling in varying environment. Multitype branching process in varying environment have been largely studied in fixed environment for finite number of types and we refer e.g. to [M71] for an overview. Much less is known or understood in the infinite type case and we here simply mention [M67] for a first work in this vein. The case of branching random walk has attracted lots of attention from the pioneering works of Biggins (see e.g. [B77, B90]) and inspired the induced (or auxiliary) Markov chains developed here. Then $\mathcal{X} = \mathbb{R}^d$ and the transitions $P^{(k)}$ are invariant by translation, i.e. $P^{(k)}(x,\mathbf{e},(x+A_1)\cdots(x+A_k))$ 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] and branching random walk in random environment have been investigated. In particular recurrence properties [CP07a], the survival and the growth rate [BCGH93, GMPV10, CP07b, CY11] and central limit theorems [Y08, N11] have been obtained.

Spine technics for homogeneous branching processes have been well developed from the works on size biased trees of Kallenberg [K77], Chauvin and Rouault and Wakolbinger [CR88, CRW91] and Lyons, Peres and Pemantle [LPP95, L97]. In the multitype case [KLPP97, A00, GB03], they rely on spectral tools and on the martingale associated to the maximal eigenvector of the mean operator (see Appendix). In varying environment, we mention[G99] for a spine description of the tree, while in the multitype case, spectral tools are no longer as powerful. We mention [C89] for the asymptotic behavior of the distribution of traits in branching processes in varying environment but the assumptions seem difficult to check when coming to applications. We use here auxiliary Markov chains, for which large deviations or geometric ergodicity can be obtained in varying or random environment (see Section 4.1 and 4.2), or for degenerated limiting distributions (see Section 4.3), using Doeblin and Lyapounov type conditions. 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] on which we will come back along the paper.

Notation. The space $\chi\times E$ is endowed with the product topology, which make it polish. For convenience we write $\mathcal{B}_{\mathcal{X}^k\times E^i} = (\mathcal{B}_{\mathcal{X}})^k \times (\mathcal{B}_E)^i$ the product σ-algebra for $k, i \ge 0$. Moreover we denote $\mathcal{B}(S)$ the set of measurable functions on S endowed with its σ-algebra. We denote $\mathcal{M}(S)$, resp. $\mathcal{M}_f(S)$ and $\mathcal{M}_1(S)$ the set of (non-negative) measures, resp. finite

measure and probability distribution of a measurable space S. When S is a topological space, we endow $\mathcal{M}_1(S)$ with the weak topology. We recall that it is the smallest topology such that $\mu \in \mathcal{M}_1(S) \to \int_S f(z)\mu(dz)$ is continuous as soon as f is real, continuous and bounded on S.
Finally, we write , $\nu(f) = \int_{\mathcal{X}} f(y)\nu(dy)$ when ν is a measure and f a measurable function on $\mathcal{X}$. We are interested in the evolution of the random measure $Z_n \in \mathcal{M}_f(\mathcal{X})$ associated to the traits of the individuals:

$$Z_n := \sum_{u\in\mathbb{G}_n} \delta_{X(u)}, \quad Z_n(f) = \sum_{u\in\mathbb{G}_n} f(X(u))$$

and more specifically by $Z_n(A_n) = \#\{u \in \mathbb{G}_n : X(u) \in A_n\}$. We note that $\#\mathbb{G}_n = Z_n(\mathcal{X})$ is the total size of the population in generation n and we also define the measure of the scaled traits of the individuals in generation n :

$$f_n.Z_n = \sum_{u\in\mathbb{G}_n} \delta_{f_n(X(u))}.$$

Organization of the paper. First, we provide in Section 2.1 an expression of the mean growth rate of the population :

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbf{e},\delta_x}(Z_n(\mathcal{X}))$$

in terms of large deviations of an auxiliary Markov chain. This gives an expression of the well known Perron Frobenius root (or Lyapounov exponent in random environment), with a probabilistic trajectorial interpretation in terms of reproduction of individuals and dispersion of the traits. It relies on a first usual many-to-one formula (Lemma 1) weighted with total mean number of offspring along the ancestral lineage. It is made more explicit for some ergodic stationary random environment using a variational principle due to [S94] (Section 4.1).
Then, we study in Section 2.2 the repartition of the traits among the population and focus on the asymptotic behavior of the proportions of individuals whose trait belong to A, i.e. $Z_n(A)/Z_n(\mathcal{X})$. For that purpose, we use another non-homogeneous auxiliary Markov chain arising from the many-to-one formula (Section 3.1) and adapt computations in the works of [AK98a, G07, BH13]. This extends classical law of large numbers to both varying environment and trait dependent reproduction. Let us add that we take into account some possible renormalization of the traits via a function f_n to cover non recurrent positive cases.
Section 3 is dedicated to the proofs of these results and Section 4 to examples and applications.

2 Main results

2.1 Growth rate of the population

We give an expression of the growth rate in terms of a time non-homogeneous Markov chain $\mathfrak{X}$ associated with a random lineage. We follow a lineage by choosing uniformly at random one of the offsprings at each generation, biased by the number of children, and the transition kernel P of $\mathfrak{X}$ is defined for $x \in X$, $\mathbf{e}$ and $A \in \mathcal{B}_{\mathcal{X}}$ by

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

More precisely, the law of $\mathfrak{X}$ under is given for $n \geq 0$ by

$$\mathbb{P}_{\mathbf{e}}(\mathfrak{X}_{n+1} \in A\,|\mathfrak{X}_n = x) = P(x, T^n\mathbf{e}, A).$$

In this section, we assume that $\mathcal{X}$ is a locally compact polish space endowed with a complete metric and $\mathcal{B}_{\mathcal{X}}$ is the associated Borel σ-algebra. Moreover E is a Polish space and $\mathcal{B}_E$ is the associated Borel σ-algebra. We consider now $\mathbf{e} \in E$ and introduce the following assumption, which ensures that the empirical measure asscoiated with $(\mathfrak{X}_k, T^k\mathbf{e})_{k\geq 0}$ satisfies a Large Deviation Principle (LDP) with good rate function $I_{\mathbf{e}}$.

Assumption 1. The function $I_{\mathbf{e}} : \mathcal{X}\times E \to \mathbb{R}\cup\{\infty\}$ is lower semi-continuous for the weak topology and with compact level subsets [1] and

$$L_n^{\mathbf{e}} = \frac{1}{n+1}\sum_{k=0}^{n}\delta_{\mathfrak{X}_k, T^k\mathbf{e}}$$

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 (i.e. $E = \{\mathbf{e}\}$, $T\mathbf{e} = \mathbf{e}$), finite trait space $\mathcal{X}$, under irreducibility assumption. We refer to Theorem 3.1.6 in [DZ98]. We note that the principle can be extended to periodic environments, taking care of the irreducibility. This is a challenging question in random environment and we use a LDP principle in Proposition 1 in Section 4.1 due to [S94] for some stationary ergodic environments.

In fixed environment and finite state space, the mean growth rate ρ of the population is given by the maximal eigenvalue of the operator corresponding to the mean number of offsprings, due to Perron-Frobenius theorem. Collatz-Wielandt formula provides a min-max representation of ρ, while Krein-Rutman theorem gives an extension to infinite dimension space requiring compactness of the operator m and strict positivity, see also Appendix for complements in the homogenous framework.
In the random environment case, it corresponds to the Lyapounov exponent and quenched asymptotic results can be obtained in the case when $\mathcal{X}$ is finite [FK60]. Then,

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

the process is a branching process in random environment and we refer to [AK71, K74] for extinction criteria and [C89, T88] for its growth rate.
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 follows the approach of [BCGH93] in fixed environment and is reminiscent of multiplicative ergodicity in [MT09]. It allows to decouple the reproduction and dispersion in the dynamic and we refer to [BL12] for motivations in ecology, more specifically for metapopulations, see in particular Theorem 5.3. The next corollary then puts in light the dispersion strategy followed by typical individuals of the population for large times.

Theorem 1. *Assume that Assumption 1 hold and* $\log m : \mathcal{X}\times E \to (-\infty,\infty)$ *is continuous and bounded. Then, for every* $x_0 \in \mathcal{X}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbf{e},\delta_{x_0}}(Z_n(\mathcal{X})) = \sup_{\mu\in\mathcal{M}_1(\mathcal{X}\times E)} \left\{ \int_{\mathcal{X}\times E} \log(m(x,e))\mu(dxde) - I_{\mathbf{e}}(\mu) \right\} := \rho_{\mathbf{e}}$$

and $\rho_{\mathbf{e}} \in (-\infty,\infty)$. *Moreover*

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

is compact and non empty.

In particular, $\limsup_{n\to\infty} \frac{1}{n} \log Z_n(\mathcal{X}) \le \rho_{\mathbf{e}}$ 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 e.g. [LPP95] for one type of individual and fixed environment and [AK71] in random environment. But it is a delicate problem when the number of types is infinite, see [A00] for fixed environment using spectral tools.

Under the assumptions of Theorem 1 and $\rho_{\mathbf{e}} > 0$, we introduce the event

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

Conditionally on $\mathcal{S}$, $\#G_n < \infty$ and we let U_n be an individual uniformly chosen at random in generation n. Let us then focus on its trait frequency up to time n and the associated environment :

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

where $X_i(u)$ is the trait of the ancestor of u in generation i when $i \le |u|$, i.e.

$$X_i(u_1\cdots u_n) := X_i(u_1\cdots u_i).$$

We check now 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* $\rho_{\mathbf{e}} > 0$ *and that the probability of* $\mathcal{S}$ *is positive. Then, for every* $x_0 \in \mathcal{X}$,

$$\mathbb{P}_{\mathbf{e},\delta_{x_0}}(\nu_n \in \mathcal{C} | \mathcal{S}) \stackrel{n\to\infty}{\longrightarrow} 0,$$

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

This result deals with the pedigree [NJ84, JN96, GB03] or ancestral lineage of a typical individual. It ensures that the trait frequency along the lineage of an individual chosen uniformly is close to one of the argmax of $\rho_{\mathbf{e}}$ for large times.

We aim now at checking when process indeed grows like its mean and how is the population spread for large times.

2.2 Law of large numbers and distribution of traits

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

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

It yields the mean number of descendant in generation n, whose trait belongs to A, of an initial individual with trait x. By now, we assume that for all $x \in \mathcal{X}, \mathbf{e} \in E$ and $n \geq 0$,

$$m_n(x,\mathbf{e},\mathcal{X}) < \infty.$$

We define a new family of Markov kernels $(Q_n : n \geq 1)$ by

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

for $A \in \mathcal{B}(\mathcal{X})$. 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. We introduce the associated semigroup, more precisely the successive composition of $(Q_j)_j$ between the generations i and n defined as follows :

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

for $A \in \mathcal{B}_{\mathcal{X}}$, where we use notation $Q(x,.) * Q'(.,.)(A) = \int_{\mathcal{X}} Q(x,dy)Q'(y,A)$. With a slight abuse, we also write $Q(\lambda,\mathbf{e},f)(x) = \int_{\mathcal{X}^2} \lambda(dx)Q(x,\mathbf{e},dy)f(y)$ when λ is a probability measure and f a measurable (positive or bounded) function.

We consider the empirical distribution of traits and prove that its asymptotic behavior is directly linked to the ergodic behavior of the auxiliary Markov chain associated to $(Q_{i,n} : i = 0,\ldots,n)$, being inspired and extending results of Athreya [AK98a, AK98b] or [G07]. First, we assume that the population has a positive growth rate and prove the a.s. convergence of the proportion of individuals with a given trait, under some uniform ergodic behavior of the auxiliary Markov chain. Then we relax the assumption on the growth rate and partially the uniform ergodicity and prove under L^2 assumptions a weak and strong law of large numbers on the empirical measure.

In the two next parts, we use a transformation f_n of the value of the traits in generation n. It is bound to make the process ergodic if it is not originally. We refer to the two last subsections for examples but one can have in mind the case when the auxiliary Markov chain X_n satisfies a central limit theorem, $f_n(x) = (x-a_n)/b_n$ and $f(X_n)$ converges to the same distribution whatever the initial value X_0 is. Such convergence hold for example for branching random walks. In the last part of this section, we compare the results obtained (Theorem 2 versus Theorems 3 and 4).

2.2.1 Branching decomposition and asymptotical distribution of traits

In this part, we focus on the event when the process survives and actually assume that the population has a positive growth rate :

$$\mathcal{T} := \left\{ \forall n \geq 0,\ Z_n(\mathcal{X}) > 0;\ \liminf_{n\to\infty} \frac{Z_{n+1}(\mathcal{X})}{Z_n(\mathcal{X})} > 1 \right\}.$$

We have then the following strong law of large numbers on this event.

Theorem 2. *Let $\mathbf{e} \in E$ and $f \in \mathcal{B}(\mathcal{X})$ bounded. We assume that there exists a measure ν with finite first moment such that for all $x \in \mathcal{X}, k, l \geq 0$,*

$$\mathbb{P}(N(x, T^k\mathbf{e}) \geq l) \leq \nu[l, \infty). \tag{2}$$

Assume also that there exists a sequence $(\mu_n)_n$ of $\mathcal{M}_1(\mathcal{X})$ and a sequence $(f_n)_n$ of $\mathcal{B}(\mathcal{X})$ such that

$$\sup_{\substack{\lambda\in\mathcal{M}_1(\mathcal{X})\\ n\geq 0}} \left| Q_{0,p}(\lambda, T^n\mathbf{e}, f\circ f_{n+p}) - \mu_{n+p}(f) \right| \longrightarrow 0 \tag{3}$$

as $p \to \infty$. Then, for any $x_0 \in \mathcal{X}$,

$$\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f) \stackrel{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_{x_0}} \text{ a.s. on the event } \mathcal{T}. \tag{4}$$

This provides a strong law of large numbers relying on the uniform ergodicity $(Q_{i,n} : i \leq n)$. It extends [AK98a, AK98b] with similar arguments to the non-neutral framework (the reproduction law may depend on the trait) and to time varying environment. We refer to Section 4.3 for an application.

2.2.2 L^2 convergence and asymptotical empirical measure of traits

In this section, we state weak and strong law of large numbers by combining L^2 computations, the ergodicity of the auxiliary Markov chain Y and the position of the most recent common ancestor of the individuals in generation n.

Assumption 2. Let $\mathbf{e}_n \in E$, $\mathcal{F} \subset \mathcal{B}(\mathcal{X})$, $f_n \in \mathcal{B}(\mathcal{X})$ and $\mu_n \in \mathcal{M}_1(\mathcal{X})$ for each $n \in \mathbb{N}$.

(a) For all $\lambda \in \mathcal{M}_1(\mathcal{X})$ and $i \in \mathbb{N}$,

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

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

$$\sup_{\lambda\in\mathcal{M}_1(\mathcal{X}), f\in\mathcal{F}} \left| Q_{k_n,n}(\lambda, \mathbf{e}_n, f\circ f_n) - \mu_n(f) \right| \stackrel{n\to\infty}{\longrightarrow} 0.$$

We note that in (a) and (b) one can choose $\lambda_0 \in \mathcal{M}_1(\mathcal{X})$ and take $\mu_n(f) := Q_{0,n}(\lambda_0, \mathbf{e}_n, f\circ f_n)$. Moreover (b) (uniform ergodicity) clearly implies (a). Sufficient conditions will be given in the applications. In particular, they are linked to Harris ergodic theorem and they can be obtained from Doeblin and Lyapounov type conditions.

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

Assumption 3. (a) For every $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that for n large enough,

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_{x_0}}(\#\{u,v \in \mathbb{G}_n : \ |u \wedge v| \geq K\})}{m_n(x_0,\mathbf{e}_n,\mathcal{X})^2} \leq \epsilon. \tag{5}$$

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

$$\sup_{n\geq i} \frac{m_{n-i}(y,T^i\mathbf{e}_n,\mathcal{X})}{m_n(x_0,\mathbf{e}_n,\mathcal{X})} \leq C_i(y) \quad \text{and} \quad \mathbb{E}\Big(\max\{C_i(X(w))^2 : w \in \mathbb{G}_{i+1}\}\Big) < \infty.$$

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

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_{x_0}}(\#\{u,v \in \mathbb{G}_n : \ |u \wedge v| \geq n-K\})}{m_n(x_0,\mathbf{e}_n,\mathcal{X})^2} \overset{n\to\infty}{\longrightarrow} 0. \tag{6}$$

Moreover,

$$\sup_{n\in\mathbb{N}} \mathbb{E}_{\mathbf{e}_n,\delta_{x_0}}(Z_n(\mathcal{X})^2)/m_n(x_0,\mathbf{e}_n,\mathcal{X})^2 < \infty.$$

These expressions can be rewritten in terms of normalized variance of $Z_n(\mathcal{X})$ and more tractable sufficient assumptions can be specified, using Lemma 5, see also the applications below. We also observe that these assumptions require that each reproduction law involved in the dynamic has a finite second moment and that $m_n(x,\mathbf{e}_n,\mathcal{X}) \to +\infty$.
The statement (5) 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 finiteness (6) says that the common ancestor is not at the end of the genealogy, so it is weaker. For a simple example where (5) is fulfilled but (6) is not, one can consider the tree T_n which is composed by a single individual until the generation $n-k_n$ and the binary tree between the generations $n-k_n$ and n, with $k_n \to \infty$.

Theorem 3 (Weak LLN)**.** *Let $\mathbf{e}_n \in E$, $x_0 \in \mathcal{X}$, $f_n : \mathcal{X} \to \mathcal{X}$ and F be a subset of $\mathcal{B}(\mathcal{X})$ such that* $\sup_{f\in\mathcal{F}} \| f \|_\infty < \infty$.

We assume either that Assumptions 2(a) and 3(a) hold or that Assumptions 2(b) and 3(b) hold. Then, uniformly for $f \in \mathcal{F}$,

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x_0,\mathbf{e}_n,\mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \qquad in \quad L^2_{\mathbf{e}_n,\delta_{x_0}}. \tag{7}$$

The (uniform) boundedness of f could be relaxed using domination assumptions following [G07, DM10], but this is not in the scope of this paper. Moreover, Let us mention that the a.s. convergence may fail in the theorem above, even in the field of applications we can have in mind. One can think for example of an underlying genealogical tree growing very slowly and the trait of the individuals are given by i.i.d. random variables.

We give now a strong law of large numbers, under stronger assumptions, using the definition

$$\mathcal{V}_i := \{(wa,wb) \in \mathbb{G}_i^2 : |w| = i-1, \ a \neq b\}. \tag{8}$$

Theorem 4 (Strong LLN)**.** *Let* $\mathbf{e} \in E$*,* $x_0 \in \mathcal{X}$ *and* $f \in \mathcal{B}_b(\mathcal{X})$*.*
Assume that

$$\liminf_{n\to\infty} m_n(x, \mathbf{e}, \mathcal{X}) > 0; \quad \sum_{i\geq 1} \mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e}, u_1, u_2)\right) < \infty, \tag{9}$$

and that there exists a sequence of probability measure μ_n *on* $\mathcal{X}$ *such that*

$$\sup_{i\in\mathbb{N}} \sum_{n\geq i} \sup_{\lambda\in\mathcal{M}_1(\mathcal{X})} \left|Q_{i,n}(\lambda, T^i\mathbf{e}, f\circ f_n) - \mu_n(f)\right|^2 < \infty. \tag{10}$$

Then $Z_n(\mathcal{X})/m_n(x_0, \mathbf{e}, \mathcal{X})$ *is bounded in* $L^2_{\mathbf{e},\delta_{x_0}}$ *and*

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x_0, \mathbf{e}, \mathcal{X})} \stackrel{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_x} \ a.s.$$

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 Section 4.2 for an application.

2.2.3 Comments on the results on the distribution of traits

The two previous Sections 2.2.1 and 2.2.2 deal with the empirical distribution of traits $f_n.Z_n(f)$ and aim at describing the asymptotic behavior of

$$f_n.Z_n(1_A) = \#\{u \in \mathbb{G}_n : X_u \in f_n^{-1}(A)\}$$

which provides the number of individuals with a given trait. As briefly explained above (see also the proofs below), the assumptions and the techniques required are different. The normalization are related but different and respectively given by the total number of indivuals $Z_n(\mathcal{X})$ at generation n and the mean number $m_n(x_0, \mathbf{e}, \mathcal{X}) = \mathbb{E}(Z_n(\mathcal{X}))$ of individuals in generation n. Finally, the limiting distribution μ_n is inherited from the same auxiliary inhomogeneous Markov chain given by $(Q_j)_j$. Roughly speaking, if one could replace $Z_n(\mathcal{X})$ by its mean, the conclusion of Theorems 2 and 4 would be the same, while relying on different assumptions. But finding tractable general conditions ensuring that the limiting behavior of $Z_n(\mathcal{X})/m_n(x_0, \mathbf{e}, \mathcal{X})$ is non-degenerated is delicate, even in the homogeneous framework (see the comments in Appendix) or under L^2 assumptions in finite dimension with varying or random environment (see e.g. [BCN99, C89]). Let us also mention that for a particular class of branching structured in varying environment and infinite dimension inherited from an aged structure, we prove in [BC15] a Kesten Stigum result using the speed of convergence of the auxiliary Markov chain introduced here. We believe that generalizations in these directions (see also Section 4.2) could be achieved but it is out of the scope of this paper. We make here only simple links between the two parts.
First, the conclusion of Theorem 2 ensures the conclusion of Theorem 3 since $Z_n(\mathcal{X})/m_n(x_0, \mathbf{e}, \mathcal{X})$ is bounded in probability. Conversely, under the assumptions of Theorem 3, the probability of the event $\{Z_n(\mathcal{X})/m_n(x_0, \mathbf{e}, \mathcal{X}) \geq \epsilon\}$ is lower bounded for ϵ small enough by Paley Zygmund inequality . Then, on this event, we note that $f_n.Z_n(f)/Z_n(\mathcal{X}) - \mu_n(f) \to 0$ in probability as $n \to \infty$.

3 Proofs

3.1 Weighted many-to-one formula and auxiliary Markov chain

We recall that $X_i(u)$ is the trait of the ancestor of u in generation i and $\mathcal{B}(\mathcal{X}^k)$ is the set of measurable functions on $\mathcal{X}^k$. The first formula we provide below provides an expressions of the ancestral lineages in the size biased tree at a fixed time. It standard in many related contexts, see e.g. [G07, BH13] for discrete time branching processes (in the neutral case) and [GB03, BDMT11, C11, HR12, HR13] for continuous time branching processes. It is a version of the so called many-to-one formulas (or Feynmac-Kac formula), weighted with respect to the current generation, whose proof in our framework is provided for safe of completeness.

Lemma 1. *Let $F \in \mathcal{B}(\mathcal{X}^k)$ non-negative. Then, for every $\mathbf{e} \in E$ and $x_0 \in \mathcal{X}$,*

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\sum_{u\in\mathbb{G}_n} F(X_0(u),\ldots,X_n(u))\right) = \mathbb{E}_{\mathbf{e},x_0}\left(F(\mathfrak{X}_0,\ldots,\mathfrak{X}_n)\prod_{i=0}^{n-1} m(\mathfrak{X}_i,T^i\mathbf{e})\right).$$

Proof. For every $f_0,\ldots,f_n \in \mathcal{B}(\mathcal{X})$ non-negative, by branching property

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\sum_{u\in\mathbb{G}_n} f_0(X_0(u))\times\cdots\times f_n(X_n(u))\right)$$
$$= \int_{\mathcal{X}} m_{\mathbf{e}}(\delta_{x_0})(dx_1)\mathbb{E}_{T\mathbf{e},\delta_{x_1}}\left(\sum_{u\in\mathbb{G}_{n-1}} f_1(X_0(u))\times\cdots\times f_n(X_{n-1}(u))\right),$$

where

$$m_{\mathbf{e}}(\delta_{x_0})(dx_1) = \mathbb{E}_{\mathbf{e},\delta_{x_0}}(\#\{u\in\mathbb{G}_1 : X(u)\in dx_1\}) = m(x_0,\mathbf{e})P(x_0,\mathbf{e},dx_1).$$

So by induction

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\sum_{u\in\mathbb{G}_n} f_0(X_0(u))\times\cdots\times f_n(X_n(u))\right)$$
$$= f_0(x_0)\int_{\mathcal{X}^n}\prod_{i=0}^{n-1} m(x_i,T^i\mathbf{e})P(x_i,T^i\mathbf{e},dx_{i+1})f_{i+1}(x_{i+1}).$$

This completes the proof by a monotone class argument. □

We rewrite now the previous formula to get ride of the weights and make appear the auxiliary Markov chain giving the traits of typcial individuals in the population. It links the expectation of the number of individuals whose trait belongs to A to the probability that the Markov chain $Y^{(n)}$ belongs to A, where $(Y_i^{(n)} : i = 0,\ldots,n)$ is associated with the transition kernels $(Q_{n-i}(.,T^i\mathbf{e},.) : i = 0,\ldots,n-1)$:

$$\mathbb{P}_{\mathbf{e}}(Y_{i+1}^{(n)} \in A \,|Y_i^{(n)} = x) = Q_{n-i}(x,T^i\mathbf{e},A)$$

where $x \in \mathcal{X}$, $\mathbf{e} \in E$ and $A \in \mathcal{B}_{\mathcal{X}}$.

Lemma 2. *For all $n \in \mathbb{N}, \mathbf{e} \in E, x_0 \in \mathcal{X}$ and $F \in \mathcal{B}(\mathcal{X}^{n+1})$ non-negative, we have*

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\sum_{u\in\mathbb{G}_n} F(X_0(u),\ldots,X_n(u))\right) = m_n(x_0,\mathbf{e},\mathcal{X})\mathbb{E}_{\mathbf{e},x_0}(F(Y_0^{(n)},\ldots,Y_n^{(n)})).$$

In particular for each $f \in \mathcal{B}(\mathcal{X})$ non-negative,

$$m_n(x_0,\mathbf{e},f) = m_n(x_0,\mathbf{e},\mathcal{X})Q_{0,n}(x_0,\mathbf{e},f).$$

Additional work would be required to provide a more complete description of the tree seen from a typical individual. In this paper, we simply mention [G99] for related results for single type branching process in varying environment.

Proof. By a telescopic argument :

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

Adding that $m_1(x_i,T^i\mathbf{e},dx_{i+1}) = m(x_i,T^i\mathbf{e})P(x_i,T^i\mathbf{e},dx_{i+1})$, the first part of the lemma is a consequence of Lemma 1. We then deduce the second part by applying the identity obtained to $F(x_0,\ldots,x_n) = f(x_n)$. □

3.2 Proofs of Theorem 1 and Corollary 1

Proof of Theorem 1. The previous lemma applied to $F = 1$ ensures that

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}(Z_n(\mathcal{X})) = \mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\prod_{i=0}^{n-1} m(\mathfrak{X}_i,T^i\mathbf{e})\right).$$

Thus

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}(Z_n(\mathcal{X})) = \mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\exp\left(n\int_{\mathcal{X}\times E}\log(m(x,e))L_{n-1}^{\mathbf{e}}(dx,de)\right)\right).$$

As $\log m$ is bounded and continuous by assumption, 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. The fact that $\rho_{\mathbf{e}} < \infty$ is due to the fact that m is bounded. Moreover $m(x,\mathbf{e}) > 0$ for any $x \in \mathcal{X}$ ensures that $\rho_{\mathbf{e}} > -\infty$.

Let us now consider a sequence μ_n such that

$$\phi(\mu_n) - I_{\mathbf{e}}(\mu_n) \stackrel{n\to\infty}{\longrightarrow} \rho_{\mathbf{e}}.$$

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

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

Recalling that ϕ is continuous, we get

$$\rho_{\mathbf{e}} = \lim_{k\to\infty} \left\{\phi(\mu_{n_k}) - I_{\mathbf{e}}(\mu_{n_k})\right\} \le \phi(\mu) - I_{\mathbf{e}}(\mu)$$

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

Proof of Corollary 1. We define for any individual $u \in \mathbb{G}_n$

$$\nu_n(u) = \frac{1}{n+1} \sum_{0\le i\le n} \delta_{X_i(u)}.$$

Using Lemma 1 with $F(x_0,\ldots,x_n) = \mathbb{1}(\frac{1}{n+1}\sum_{0\le i\le n}\delta_{x_i} \in \mathcal{C})$, we have

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}(\#\{u \in \mathbb{G}_n : \nu_n(u) \in \mathcal{C}\}) = \mathbb{E}_{\mathbf{e},x_0}\left(\exp\left(n\int_{\mathcal{X}\times E} \log(m(x,e))L_n^{\mathbf{e}}(dx,de)\right)1_{L_n^{\mathbf{e}}\in\mathcal{C}}\right)$$

Let $\mathcal{C}$ be a closed subset of $\mathcal{M}_1(\mathcal{X}\times E)$ which is disjoint of $M_{\mathbf{e}}$. Recalling that $\mathcal{X}\times E$ is polish, $\mathcal{M}_1(\mathcal{X}\times E)$ endowed with the weak topology can be metrizable by a distance d, which can be bounded by 1. Then

$$\mu \in \mathcal{M}_1(\mathcal{X}\times E) \to d(\mu,\mathcal{C}) = \inf_{\nu\in\mathcal{C}} d(\mu,\nu)$$

is continuous and bounded, so the function

$$\phi(\mu) = -d(\mu,\mathcal{C}) + \int_{\mathcal{X}\times E} \log(m(x,e))\mu(dx,de)$$

is continuous and bounded from $\mathcal{M}_1(\mathcal{X}\times E)$ to $\mathbb{R}$. Applying again Varadhan's lemma, we get

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{E}_{\mathbf{e},\nu}(\#\{u \in \mathbb{G}_n : \nu_n(u) \in \mathcal{C}\})\\ \le \mathbb{E}_{\mathbf{e},\delta_{x_0}}(\exp(n\phi(L_n^{\mathbf{e}}))\\ \le \sup\{\phi(\mu) - I_{\mathbf{e}}(\mu) : \mu \in \mathcal{M}_1(\mathcal{X}\times E)\}.\end{aligned}$$

We add now that

$$\phi(\mu) - I_{\mathbf{e}}(\mu) \ge \rho_{\mathbf{e}}$$

implies that $d(\mu,\mathcal{C}) = 0$ and then $\mu \in \mathcal{C}$. Thus, using as in the end of the previous an extraction argument and the fact that $\mathcal{C}\cap M_{\mathbf{e}} = \varnothing$ with $\mathcal{C}$ closed, we get

$$\sup\{\phi(\mu) - I_{\mathbf{e}}(\mu) : \mu \in \mathcal{M}_1(\mathcal{X}\times E)\} < \rho_{\mathbf{e}}$$

and we can choose ρ' such that

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{E}_{\mathbf{e},\delta_{x_0}}(\#\{u \in \mathbb{G}_n : \nu_n(u) \in \mathcal{C}\}) < \rho' < \rho_{\mathbf{e}}.$$

Adding that

$$\begin{aligned}\mathbb{P}_{\mathbf{e},\delta_{x_0}}(\nu_n(U_n) \in \mathcal{C}|\mathcal{S}) &\le \mathbb{E}_{\mathbf{e},\delta_{x_0}}(\#\{u \in \mathbb{G}_n, \nu_n(u) \in \mathcal{C}\}/Z_n(\mathcal{X})|\mathcal{S})\\ &\le e^{-\rho' n}\mathbb{E}_{\mathbf{e},\delta_{x_0}}(\#\{u \in \mathbb{G}_n, \nu_n(u) \in \mathcal{C}\})/\mathbb{P}(\mathcal{S})\end{aligned}$$

for n large enough by definition of $\mathcal{S}$ and that the right hand side goes to 0 ends up the proof. □

3.3 Proof of Theorem 2

We first state a law of large numbers, which is being used several time. It is an easy extension of Lemma 1 in [AK98a], which itself is proved using [K72].

Lemma 3. *Let $\{\mathcal{F}_n\}_0^\infty$ be a filtration. Let $\{Y_{n,i} : n, i \geq 1\}$ be real valued random variables such that for each n, conditionally on $\mathcal{F}_n$, $\{Y_{n,i} : i \geq 1\}$ are centered independent r.v. Let $\{N_n : n \geq 1\}$ be non-negative integer valued r.v. such that for each n, N_n is $\mathcal{F}_n$ measurable.*
We assume that there exists a random measure μ with finite first moment such that

$$\forall t > 0, \qquad \sup_{n,i\geq 1} \mathbb{P}(|Y_{n,i}| > t|\mathcal{F}_n) \leq \mu(t,\infty) \quad a.s.$$

Then

$$\frac{1}{N_n}\sum_{i=1}^{N_n} Y_{n,i} \overset{n\to\infty}{\longrightarrow} 0$$

a.s. on the event $\{\forall n \geq 0 : N_n > 0,\ \liminf_{n\to\infty} N_{n+1}/N_n > 1\}$.

Proof. The proof can be simply adapted from the proof of Lemma 1 in [AK98a]. For any $\delta > 0$, $n_0 \geq 2$ and $l > 1$, we define

$$A_n := \left\{ \left| \frac{1}{N_n}\sum_{i=1}^{N_n} Y_{n,i} \right| > \delta;\ \forall k = n_0, \ldots, n : \frac{N_k}{N_{k-1}} \geq l \right\}$$

and prove similarly that $\sum_{n\geq n_0} \mathbb{P}(A_n|\mathcal{F}_n) < \infty$ using the domination by μ. This yields the expected a.s. convergence on the event $\{\forall n \geq n_0,\ N_{n+1}/N_n \geq l\}$ by conditional Borel Cantelli Lemma [Chow and Teicher, 1988, p249] and the result follows by monotone limit. □

We use this lemma to prove the following result.

Lemma 4. *Under the assumptions of Theorem 2, we have*

$$\frac{1}{Z_{n+p}(\mathcal{X})}\sum_{u\in\mathbb{G}_n} m_p(X(u), T^n\mathbf{e}, \mathcal{X}) \overset{n\to\infty}{\longrightarrow} 1 \quad \mathbb{P}_{\mathbf{e},\delta_x}\ a.s\ on\ \mathcal{T}.$$

Proof. The branching property gives a natural decomposition of the population in generation $n+p$:

$$Z_{n+p}(\mathcal{X}) = \sum_{u\in\mathbb{G}_n} Z_p^{(u)}(\mathcal{X}),$$

where

$$Z_p^{(u)} = \sum_{v\in\mathbb{G}_p^{(u)}} \delta_{X^{(u)}(v)}$$

and $X^{(u)}$ is the branching Markov chain X rooted in u whose set of individuals in generation p is $\mathbb{G}_p^{(u)} = \{v \in \mathcal{U} : |v| = p, uv \in \mathbb{G}_{n+p}\}$. Moreover,

$$\begin{aligned} Z_{n+p}(\mathcal{X}) - \sum_{u\in\mathbb{G}_n} m_p(X(u), T^n\mathbf{e}, \mathcal{X}) &= \sum_{u\in\mathbb{G}_n}\left[Z_p^{(u)}(\mathcal{X}) - m_p(X(u), T^n\mathbf{e}, \mathcal{X}) \right] \\ &= Z_n(\mathcal{X})\epsilon_{n,p}, \end{aligned}$$

where

$$\epsilon_{n,p} := \frac{1}{Z_n(\mathcal{X})} \sum_{u\in\mathbb{G}_n} Y^{(n)}_{p,u}, \qquad Y^{(n)}_{p,u} := Z^{(u)}_p(\mathcal{X}) - m_p(X(u), T^n\mathbf{e}, \mathcal{X}).$$

By construction, $(Y^{(n)}_{p,u} : u \in \mathbb{G}_n)$ are independent conditionally on $\mathcal{F}_n = \sigma(X(v) : v \in \mathbb{G}_k, k \le n) = \sigma(Z_k : k \le n)$, $\mathbb{E}(Y^{(n)}_{p,u}) = 0$ and $|Y^{(n)}_{p,u}| \le |Z^{(u)}_p(\mathcal{X})| + m_p(X(u), T^n\mathbf{e}, \mathcal{X})$, so that the stochastic domination assumption (2) ensures that there exists a measure with finite first moment μ such that

$$\sup_{u\in\cup_n\mathbb{G}_n} \mathbb{P}_{\mathbf{e},\delta_x}(|Y^{(n)}_{p,u}| > t|\mathcal{F}_{|u|}) \le \mu(t,\infty).$$

We can then apply the law of large number of Lemma 4 to get that for every $p \ge 0$, $\epsilon_{n,p} \to 0$ a.s. on the event $A_{n_0,l}$, as $n \to \infty$. Recalling that $Z_{n+p}(\mathcal{X}) \ge Z_n(\mathcal{X})$ for n large enough ends up the proof. □

We can now prove the Theorem.

Proof of Theorem 2. We use again the branching decomposition in generation n to write

$$\begin{aligned}
&\left|\frac{f_{n+p}.Z_{n+p}(f)}{Z_{n+p}(\mathcal{X})} - \mu_{n+p}(f)\right| \\
&= \left|\frac{1}{Z_n(\mathcal{X})} \sum_{u\in\mathbb{G}_n} \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} f_{n+p}.Z^{(u)}_p(f) - \mu_{n+p}(f)\right| \\
&\le \left|\sum_{u\in\mathbb{G}_n} \frac{X_{u,n,p}}{Z_n(\mathcal{X})}\right| + \left|\sum_{|u|=n} \frac{Y_{u,n,p}}{Z_{n+p}(\mathcal{X})}\right| + \mu_{n+p}(f)\left|\sum_{u\in\mathbb{G}_n} \frac{m_p(X(u), T^n\mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} - 1\right|,
\end{aligned}$$

where

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

and

$$Y_{u,n,p} = m_p(X(u), T^n\mathbf{e}, \mathcal{X})[Q_{0,p}(X(u), T^n\mathbf{e}, f\circ f_{n+p}) - \mu_{n+p}(f)].$$

We want to prove that these quantities go to zero. First we note that

$$\begin{aligned}
X_{u,n,p} &\le f_{n+p}.Z^{(u)}_p(f) + m_p(X(u), T^n\mathbf{e}, f\circ f_{n+p}) \\
&\le \| f \|_\infty [Z^{(u)}_p(\mathcal{X}) + m_p(X(u), T^n\mathbf{e}, \mathcal{X})],
\end{aligned}$$

so that Assumption (2) ensures that the r.v. $X_{u,n,p}$ are stochastically dominated. Then, we can apply the law of large numbers of Lemma 4 and get that for each $p \ge 0$, $\sum_{u\in\mathbb{G}_n} X_{u,n,p}/Z_n(\mathcal{X}) \to 0$ as $n \to \infty$.
Moreover the many-to-one formula (Lemma 2) ensures that

$$\begin{aligned}
Y_{u,n,p} &= m_p(X(u), T^n\mathbf{e}, \mathcal{X})[Q_{0,p}(X(u), T^n\mathbf{e}, f\circ f_{n+p}) - \mu_{n+p}(f)] \\
&\le m_p(X(u), T^n\mathbf{e}, \mathcal{X})M_p,
\end{aligned}$$

where

$$M_p := \sup_{n\ge 0} |Q_{0,p}(X(u), T^n\mathbf{e}, f\circ f_{n+p}) - \mu_{n+p}(f)|.$$

Combining these results, we get

$$\limsup_{n\to\infty}\left|\frac{f_{n+p}.Z_{n+p}(f)}{Z_{n+p}(\mathcal{X})}-\mu_{n+p}(f)\right| \\ \leq M_p \limsup_{n\to\infty}\sum_{u\in\mathbb{G}_n}\frac{m_p(X(u),T^n\mathbf{e},\mathcal{X})}{Z_{n+p}(\mathcal{X})} \\ +\parallel f\parallel_\infty \limsup_{n\to\infty}\left|\sum_{u\in\mathbb{G}_n}\frac{m_p(X(u),T^n\mathbf{e},\mathcal{X})}{Z_{n+p}(\mathcal{X})}-1\right|\leq M_p.$$

by means of Lemma 4. Using now that $M_p\to 0$ as $p\to\infty$ by (3), we have

$$\limsup_{p\to\infty}\limsup_{n\to\infty}\left|\frac{Z_{n+p}(f)}{f_{n+p}.Z_{n+p}(\mathcal{X})}-\mu_{n+p}(f)\right|=0$$

and the proof is complete. □

3.4 L^2 estimates

We consider the variance of the size of the population and for that purpose we recall notation (8).

Lemma 5. *Let* $\mathbf{e}\in E$, $x_0\in\mathcal{X}$ *and* $0\leq k\leq n$. *We have*

$$\mathbb{E}_{\mathbf{e},\delta_{x_0}}(\#\{(u,v)\in\mathbb{G}_n^2\ :\ |u\wedge v|\geq k\}) \\ = m_n(x_0,\mathbf{e})+\sum_{i=k+1}^{n}\mathbb{E}_{\mathbf{e},\delta_{x_0}}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} m_{n-i}(X(u_1),T^i\mathbf{e})m_{n-i}(X(u_2),T^i\mathbf{e})\right).$$

In particular, defining

$$V_i(\mathbf{e},u_1,u_2)=\sup_{k\geq 0}\frac{m_k(X(u_1),T^i\mathbf{e},\mathcal{X})m_k(X(u_2),T^i\mathbf{e},\mathcal{X})}{m_{i+k}(x_0,\mathbf{e},\mathcal{X})^2}$$

and assuming that for some sequence $\mathbf{e}_n\in E$,

$$\liminf_{n\to\infty} m_n(x,\mathbf{e}_n,\mathcal{X})>0;\quad \sup_{n\geq 0}\sum_{i\geq 1}\mathbb{E}_{\mathbf{e}_n,\delta_{x_0}}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i}V_i(\mathbf{e}_n,u_1,u_2)\right)<\infty, \tag{11}$$

then $Z_n(\mathcal{X})/m_n(x,\mathbf{e}_n,\mathcal{X})$ *is bounded in* $L^2_{\mathbf{e}_n,\delta_{x_0}}$.

Proof of Lemma 5. We omit the initial state δ_{x_0} in the notations and write $m_n(x,\mathbf{e})$ for $m_n(x,\mathbf{e},\mathcal{X})$. Using the branching property and distinguishing if the common ancestor of two individuals lives before generation n or in generation n, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbf{e}}(\#\{(u,v)\in\mathbb{G}_n^2 : |u\wedge v|\geq k\}) \\
&= \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{(u,v)\in\mathbb{G}_n^2\\ |u\wedge v|\geq k}} 1\right) \\
&= \mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})) + \mathbb{E}_{\mathbf{e}}\left(\sum_{k+1\leq i\leq n}\ \sum_{\substack{w\in\mathbb{G}_{i-1}\\ a\neq b\\ (wa,wb)\in\mathbb{G}_i}}\ \sum_{\substack{u\in\mathbb{G}_n: u\geq wa\\ v\in\mathbb{G}_n: v\geq wb}} 1\right) \\
&= m_n(x_0,\mathbf{e}) + \sum_{k+1\leq i\leq n}\mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{w\in\mathbb{G}_{i-1},\ a\neq b\\ (wa,wb)\in\mathbb{G}_i^2}} m_{n-i}(X(wa),T^i\mathbf{e})m_{n-i}(X(wb),T^i\mathbf{e})\right).
\end{aligned}$$

This yields the first part of the Lemma. Then, letting $k=0$ and dividing by $m_n(x,\mathbf{e})^2$ ensures that

$$\frac{\mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})^2)}{m_n(x_0,\mathbf{e})^2} \leq \frac{1}{m_n(x_0,\mathbf{e})} + \sum_{i\leq n}\mathbb{E}_{\mathbf{e}}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e},X(u_1),X(u_2))\right),$$

which ends up the proof. □

Let us end up this part with a lemma, which provides a convenient sufficient condition for (11). It will be used for applications in Section 4.2. We recall that it ensures that the common ancestor of two individuals chosen independently lives at the beginning of the tree.

Lemma 6. *Assume that there exist positive real numbers $(C(\mathbf{e}) : \mathbf{e}\in E)$ such that for all $x,y\in\mathcal{X}$ and $n\geq 0$,*

$$m_n(y,\mathbf{e},\mathcal{X}) \leq C(\mathbf{e})m_n(x,\mathbf{e},\mathcal{X}) \tag{12}$$

and consider a sequence $\mathbf{e}_n\in E$ such that

$$\sup_{n\geq 0}\sum_{i=1}^{n}\frac{1\wedge D(x,T^{i-1}\mathbf{e}_n)}{m_i(x,\mathbf{e}_n,\mathcal{X})} < \infty, \tag{13}$$

where

$$\sigma(\mathbf{e}) := \sup_{y\in\mathcal{X}}\mathbb{E}(N(y,\mathbf{e})^2), \qquad D(x,\mathbf{e}) := \frac{\sigma(\mathbf{e})C(\mathbf{e})}{m(x,\mathbf{e})}C(T\mathbf{e})^2. \tag{14}$$

Then (11) holds and $Z_n(\mathcal{X})/m_n(x,\mathbf{e}_n,\mathcal{X})$ is bounded in $L^2_{\mathbf{e}_n,\delta_x}$.
Moreover, if $\mathbf{e}_n=\mathbf{e}$ is constant, then (9) also holds.

Proof. Using the branching property in generation i and (12), we have for all $y\in\mathcal{X}$,

$$m_{i+k}(x,\mathbf{e}) \geq m_i(x,\mathbf{e})C(T^i\mathbf{e})^{-1}m_k(y,T^i\mathbf{e}). \tag{15}$$

Then

$$V_i(\mathbf{e},u_1,u_2) \le \frac{C(T^i\mathbf{e})^2}{m_i(x,\mathbf{e},\mathcal{X})^2}$$

and

$$\mathbb{E}_{\mathbf{e},\delta_x}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e}_n,u_1,u_2)\right) \quad\le\quad \mathbb{E}_{\mathbf{e}_n}(\mathcal{V}_i)\,\frac{C(T^i\mathbf{e})^2}{m_i(x,\mathbf{e},\mathcal{X})^2}.$$

Adding that

$$\mathbb{E}_{\mathbf{e}}(\mathcal{V}_i) \le m_{i-1}(x,\mathbf{e},\mathcal{X})\sigma(T^{i-1}\mathbf{e})$$

and using again (15) with $k=1$, we get

$$\mathbb{E}_{\mathbf{e},\delta_x}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e},u_1,u_2)\right) \le \frac{\sigma(T^{i-1}\mathbf{e})C(T^{i-1}\mathbf{e})}{m(x,T^{i-1}\mathbf{e})}\,\frac{C(T^i\mathbf{e})^2}{m_i(x,\mathbf{e},\mathcal{X})}.$$

Thus, (11) and (9) hold. Applying Lemma 5 and considering a sequence $\mathbf{e}_n$ in the previous estimates ends up the proof. □

3.5 Proofs of Theorem 3 and Theorem 4

Proof of Theorem 3. 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 of the expectation. For convenience, we also write $m(x,\mathbf{e}_n) := m(x,\mathbf{e}_n,\mathcal{X})$, $b := \sup_{f\in\mathcal{F}} \| f \|_\infty$ and denote

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

We compute for $K \ge 1$,

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

The second term of the right hand side is smaller than

$$2 \| f \|_\infty^2 \,\mathbb{E}(\#\{u,v\in\mathbb{G}_n : |u\wedge v| > K\}) \le 2b^2 m(x,\mathbf{e}_n)^2.\epsilon_{K,n},$$

where $\limsup_{n\to\infty}\epsilon_{K,n} \to 0$ as $K\to\infty$ using the first part of Assumption 3(a). So we just deal with the first term and the most recent common ancestor is labeled by $i-1$ for

$i = 1, \dots, K$. Thanks to the branching property,

$$\begin{aligned}\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{u,v\in\mathbb{G}_n\\|u\wedge v|=i-1}} g_n(X(u))g_n(X(v))\right) &= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{w\in\mathbb{G}_{i-1}\\(wa,wb)\in\mathbb{G}_i^2}}\ \sum_{\substack{u\in\mathbb{G}_n\\u\geq wa}}\sum_{\substack{v\in\mathbb{G}_n\\v\geq wb}} g_n(X(u))g_n(X(v))\right)\\ &= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{w\in\mathbb{G}_{i-1}\\(wa,wb)\in\mathbb{G}_i^2}} R_{i,n}(X(wa))R_{i,n}(X(wb))\right),\end{aligned}$$

where the many-to-one formula of Lemma 2 allows us to write

$$R_{i,n}(x) := \mathbb{E}_{T^i\mathbf{e}_n}\left(\sum_{u\in\mathbb{G}_{n-i}} g_n(X(u))\right) = m_{n-i}(x, T^i\mathbf{e}_n)Q_{0,n-i}(x, T^i\mathbf{e}_n, g_n). \tag{16}$$

Then

$$\begin{aligned}&m(x_0,\mathbf{e}_n,\mathcal{X})^{-2}\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{(u,v)\in\mathbb{G}_n^2\\|u\wedge v|\leq K}} g_n(X(u))g_n(X(v))\right)\\ &= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{i<K,w\in\mathbb{G}_{i-1}\\(wa,wb)\in\mathbb{G}_i}} F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa),T^i\mathbf{e}_n)m_{n-i}(X(wb),T^i\mathbf{e}_n)}{m_n(x_0,\mathbf{e}_n)^2}\right).\end{aligned}$$

and Assumption 2(a) ensures that

$$F_{i,n}(u) := \frac{R_{i,n}(X(u))}{m_{n-i}(X(u),T^i\mathbf{e}_n)} = Q_{0,n-i}(X(u), T^i\mathbf{e}_n, f\circ f_n) - \mu_n(f) \tag{17}$$

goes to 0 a.s. for each $i \in \mathbb{N}$ and $u \in \mathbb{G}_i$. Moreover this convergence is uniform for $f \in \mathcal{F}$. Adding that $F_{i,n}$ is bounded by b, we have

$$\begin{aligned}&F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa),T^i\mathbf{e}_n)m_{n-i}(X(wb),T^i\mathbf{e}_n)}{m_n(x_0,\mathbf{e}_n)^2}\\ &\qquad\leq b^2\sup_n\frac{m_{n-i}(X(wa),T^i\mathbf{e}_n)}{m_n(x_0,\mathbf{e}_n)}.\sup_n\frac{m_{n-i}(X(wb),T^i\mathbf{e}_n)}{m_n(x_0,\mathbf{e}_n)}.\end{aligned}$$

By bounded convergence, the second part of Assumption 3(a) ensures that $Z_n(g_n)/m_n(x_0,\mathbf{e}_n) \to 0$ in $L^2_{\mathbf{e}_n}$ uniformly for $f \in \mathcal{F}$. It ends up the proof of (7) under Assumptions 2(a) and 3(a).

The proof of (7) under Assumptions 2(b) and 3(b) is almost the same, replacing K by $n-k_n$ with $k_n \to \infty$. Indeed, Assumption 3(b) ensures that there exists $k_n \to \infty$ such that

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_{x_0}}\left(\#\{(u,v)\in\mathbb{G}_n^2 : |u\wedge v| > n-k_n\}\right)}{m_n(x_0,\mathbf{e}_n)^2} \overset{n\to\infty}{\longrightarrow} 0,$$

whereas

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

$$\leq \left(\sup_{\substack{i\leq n-k_n\\x\in\mathcal{X}}} F_{i,n}(x)\right)^2 \frac{\mathbb{E}(Z_n(\mathcal{X})^2)}{m_n(x,\mathbf{e}_n)^2}.$$

Assumption 2(b) ensures that $\sup_{i\leq n-k_n, x\in\mathcal{X}} F_{i,n}(x)\to 0$ as $k_n\to\infty$ and the second part of Assumption 3(b) ensures that $\mathbb{E}_{\mathbf{e}_n}(Z_n(\mathcal{X})^2)/m_n(x,\mathbf{e}_n)^2$ is bounded. The conclusion is thus the same and the proof is complete. □

Proof of Theorem 4. The fact that $Z_n(\mathcal{X})/m_n(x_0,\mathbf{e},\mathcal{X})$ is bounded in $L^2_{\mathbf{e},\delta_x}$ comes directly from the second part of Lemma 5. To get the a.s. convergence, we prove that

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

For that purpose, we use the notations of the proof of the previous theorem, in particular $g_n(.):=f(f_n(.))-\mu_n(f)$, and we are inspired by L^2 computations for Markov chain indexed by trees, see the proof of Theorem 14 in [G07]. Using Fubini inversion, the branching property and (16), we have

$$\sum_{n\geq 0} m_n(x,\mathbf{e})^{-2}\mathbb{E}_{\mathbf{e}}(Z_n(g_n)^2)$$

$$=\mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}}\sum_{(u,v)\in\mathbb{G}_n^2} m_n(x,\mathbf{e})^{-2}g_n(X(u))g_n(X(v))\right)$$

$$=\mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}}\sum_{i\leq n}\sum_{\substack{(u,v)\in\mathbb{G}_n^2\\|u\wedge v|=i}} m_n(x,\mathbf{e})^{-2}g_n(X(u))g_n(X(v))\right)$$

$$=\mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{n\in\mathbb{N}\\i\leq n}}\sum_{\substack{w\in\mathbb{G}_{i-1}\\(wa,wb)\in\mathbb{G}_i^2}}\sum_{\substack{u\in\mathbb{G}_n:u\geq wa\\v\in\mathbb{G}_n:v\geq wb}} m_n(x,\mathbf{e})^{-2}g_n(X(u))g_n(X(v))\right)$$

$$+\mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N},u\in\mathbb{G}_n} m_n(x,\mathbf{e})^{-2}g_n(X(u))^2\right).$$

Then

$$\sum_{n\geq 0} m_n(x,\mathbf{e})^{-2}\mathbb{E}_{\mathbf{e}}(Z_n(g_n)^2)$$

$$\leq \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{n\in\mathbb{N}\\ i\leq n}}\ \sum_{\substack{w\in\mathbb{G}_{i-1}\\ (wa,wb)\in\mathbb{G}_i}} \frac{m_{n-i}(X(wa),T^i\mathbf{e})m_{n-i}(X(wb),T^i\mathbf{e})}{m_n(x,\mathbf{e})^2}F_{i,n}(X(wa))F_{i,n}(X(wb))\right)$$

$$+2\parallel g_n\parallel_\infty \mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}} m_n(x,\mathbf{e})^{-2}Z_n(\mathcal{X})\right)$$

$$\leq \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{i\in\mathbb{N},w\in\mathbb{G}_{i-1}\\ (wa,wb)\in\mathbb{G}_i}} V_i(\mathbf{e},wa,wb)H_i\right)+b\sum_{n\in\mathbb{N}} m_n(x,\mathbf{e})^{-1},$$

where $b:=(2\parallel f\parallel_\infty)^2$ and recalling definition (17),

$$H_i=\sup_{y,z}\sum_{n\geq i}F_{i,n}(y)F_{i,n}(z), \qquad V_i(\mathbf{e},u_1,u_2)=\sup_{n\geq i}\frac{m_{n-i}(X(u_1),T^i\mathbf{e})m_{n-i}(X(u_2),T^i\mathbf{e})}{m_n(x,\mathbf{e})^2}.$$

The assumptions (9) and (10) ensure that

$$\sum_{n\geq 0} m_n(x,\mathbf{e},\mathcal{X})^{-2}\mathbb{E}_{\mathbf{e}}(Z_n(g_n)^2)\leq$$

$$b\sum_{n\geq 0} m_n(x,\mathbf{e})^{-1}+\sup_{i\in\mathbb{N}} H_i.\sum_{i\in\mathbb{N}}\mathbb{E}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e},u_1,u_2)\right)<\infty.$$

Then, $Z_n(g_n)/m_n(x,\mathbf{e})\to 0$ $\mathbb{P}_{\mathbf{e}}$ a.s., which completes the proof. □

4 Applications and examples

We illustrate now the previous theorems and derive asymptotic results for branching processes in varying or random environment.

4.1 Quenched mean growth rate in some ergodic random environment

In this section, we consider the mean growth rate of the population in a random environment, under a strong Doeblin assumption of the transition kernel P defined in Section 2.1. This assumption allows us to use the large deviation principle obtained in [S94] and apply Theorem 1. Following [S94], we assume in this section that $T:E\to E$ is an homeomorphism. 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 $\pi(A)\in\{0,1\}$. We require :

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 $x,y\in\mathcal{X}$, $A\in\mathcal{B}_{\mathcal{X}}$ 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 from $\mathcal{X}\times E$ into $[1,\infty)$ and $\mathcal{M}_1^\pi(\mathcal{X}\times E)$ the set of probabilities on $\mathcal{X}\times E$ whose $\mathcal{X}$ marginal is equal to π.

Proposition 1. *Under Assumption A, we further suppose that $\log m : \mathcal{X}\times E \to (-\infty,\infty)$ is continuous and bounded. Then there exists $E'' \subset E$ such that $\pi(E'')=1$ and for all $\mathbf{e}\in E''$ and $x_0\in\mathcal{X}$,*

$$\lim_{n\to\infty}\frac{1}{n}\log \mathbb{E}_{\mathbf{e},\delta_{x_0}}(Z_n(\mathcal{X})) = \sup_{\mu\in\mathcal{M}_1(\mathcal{X}\times E)}\left\{\int \log(m(x,e))\mu(dx,de) - I(\mu)\right\},$$

where I is defined by

$$I(\mu) := \sup\left\{\int_{\mathcal{X}\times E}\log\left(\frac{u(x,e)}{\int_{\mathcal{X}}P(x,e,dy)u(y,Te)}\right)\mu(dx,de) : u\in V_b(\mathcal{X}\times E)\right\} \tag{18}$$

for $\mu\in\mathcal{M}_1^\pi(\mathcal{X}\times E)$ and $I(\mu)=+\infty$ otherwise.

Proof. Under Assumption A, Theorem 3.3 [S94] ensures that the function I defined by (18) satisfies Assumption 1, π a.e. uniformly with respect to $x\in\mathcal{X}$. The result is then a direct application of Theorem 1. □

4.2 Asymptotical empirical measure of traits in varying environment

We need here an additional assumption on the transition semi-groups, namely we require a contraction property. Such contractions can be derived from classical technics relying on Doeblin and Lyapounov type assumptions, see in particular [MT09, HM08, M13]. The key point is that they can be composed in the non-homogenous framework. We focus on the case of Doeblin type conditions to obtain the results below, while the use of Lyapounov function in our context is more involved and left for future works. We recall from Section 3.4 the definition $\sigma(\mathbf{e}) := \sup_{y\in\mathcal{X}}\mathbb{E}(N(y,\mathbf{e})^2)$.

Proposition 2. *We assume that there exists $M : E\to[1,\infty)$ such that for all $x,y\in\mathcal{X},\mathbf{e}\in E$,*

$$m_1(x,\mathbf{e},.) \le M(\mathbf{e})m_1(y,\mathbf{e},.),$$

Let $\mathbf{e_0}\in E$ and $x_0\in\mathcal{X}$ be such that

$$\sum_{n\ge0}\frac{1}{m_{n+1}(x_0,\mathbf{e}_0,\mathcal{X})}\left(1+\frac{\sigma(T^n\mathbf{e}_0)M(T^n\mathbf{e}_0)}{m(x_0,T^n\mathbf{e}_0)}M(T^{n+1}\mathbf{e}_0)^2\right)<\infty$$

and

$$\sum_{n\ge0}\prod_{k=0}^{n}\left(1-1/M(T^k\mathbf{e}_0)^2\right)^2<\infty.$$

Then, $Z_n(\mathcal{X})/m_n(x_0,\mathbf{e}_0,\mathcal{X})$ is bounded in $L^2_{\mathbf{e}_0,\delta_{x_0}}$ and for every $f\in\mathcal{B}_b(\mathcal{X})$,

$$\frac{Z_n(f)-\mu_n(f)Z_n(\mathcal{X})}{m_n(x_0,\mathbf{e}_0,\mathcal{X})}\overset{n\to\infty}{\longrightarrow}0 \qquad \mathbb{P}_{\mathbf{e}_0,\delta_{x_0}}\ a.s.$$

This result can be applied to reproduction-dispersion branching processes on compact sets, where both the reproduction and the dispersion may be affected by the environment. To give a simple example when this result can be applied, to apply this proposition, one can consider the case when when M and σ are bounded and the exists $m_i > 0$ such that for any $x \in \mathcal{X}$, $m(x, T^i \mathbf{e}_0) \geq m_i$ and

$$\sum_{i \geq 0} \frac{1}{m_i^2} < \infty.$$

Let us finally observe that in the particular case when there exists $m > 0$ such that for any $i \in \mathbb{N}$, $m_i \geq m$, Theorem 2 could also be applied to obtain the a.s. behavior of the distribution of traits among the population $Z_n(f)/Z_n(\mathcal{X}) - \mu_n(f) \to 0$.
For the proof, we use the following lemma and recall definition (14). The space of probabilities $\mathcal{M}_1(\mathcal{X})$ is now endowed with a distance d bounded by 1 such that for every bounded measurable function f, there exists $c_0 > 0$ such that $|\mu(f) - \nu(f)| \leq c_0 d(\mu, \nu)$.

Lemma 7. *Assume that there exist real numbers $(C(\mathbf{e}), A_n(\mathbf{e}) : \mathbf{e} \in E)$ such that for any $\mathbf{e} \in E$, $x, y \in \mathcal{X}$, $\lambda, \mu \in \mathcal{M}_1(\mathcal{X})$ and $n \geq 0$,*

$$m_n(y, \mathbf{e}, \mathcal{X}) \leq C(\mathbf{e}) m_n(x, \mathbf{e}, \mathcal{X}), \qquad d(Q_n(\lambda, \mathbf{e}, .), Q_n(\mu, \mathbf{e}, .)) \leq A_n(\mathbf{e}) d(\lambda, \mu). \tag{19}$$

Let $\mathbf{e_0} \in E$ and $x_0 \in \mathcal{X}$ be such that

$$\sum_{n \geq 1} \frac{1 \wedge D(x_0, T^{n-1}\mathbf{e}_0)}{m_n(x_0, \mathbf{e}_0, \mathcal{X})} < \infty, \qquad \sum_{n \geq 0} \prod_{k=0}^{n} A_k(T^{n-k}\mathbf{e}_0)^2 < \infty.$$

Then, $Z_n(\mathcal{X})/m_n(x_0, \mathbf{e}_0, \mathcal{X})$ is bounded in $L^2_{\mathbf{e}_0, \delta_{x_0}}$ and for every $f \in \mathcal{B}_b(\mathcal{X})$,

$$\frac{Z_n(f) - \mu_n(f) Z_n(\mathcal{X})}{m_n(x_0, \mathbf{e}_0, \mathcal{X})} \overset{n \to \infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e}_0, \delta_{x_0}} \ a.s.$$

Proof. The proof is an application of Lemma 6 and Theorem 4. Indeed, we first use Lemma 6 to check that (9) hold and $Z_n(\mathcal{X})/m_n(x_0, \mathbf{e}, \mathcal{X})$ is bounded in $L^2_{\mathbf{e}, \delta_{x_0}}$. Moreover, by induction, the second part of (19) yields

$$d(Q_{i,n}(\lambda, \mathbf{e}, .) - Q_{i,n}(\mu, \mathbf{e}, .)) \leq \Pi_{k=i}^{n-1} A_{n-k}(T^k \mathbf{e}) d(\lambda, \mu) \leq \Pi_{k=i}^{n-1} A_{n-k}(T^k \mathbf{e}),$$

since d is bounded by 1. Then

$$|Q_{i,n}(\lambda, \mathbf{e}, f) - Q_{i,n}(\mu, \mathbf{e}, f)| \leq c_0 \Pi_{k=i}^{n-1} A_{n-k}(T^k \mathbf{e}).$$

Adding that the right-hand-side is summable allows us to get (10) and apply Theorem 4 with $f_n = Id$ and conclude. □

Proof of Proposition 2. Let us check that (19) hold and apply the previous lemma. Indeed, we obtain the first part of (19) with $C = M$ since

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

The second part of (19) comes from

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

Choosing for d the total variation distance $d(\lambda,\mu) = \sup_{\|f\|_\infty \le 1} \left| \int_{\mathcal{X}} f(x)\lambda(dx) - \int_{\mathcal{X}} f(x)\mu(dx) \right|$ we get

$$d(Q_n(\lambda,\mathbf{e},.),Q_n(\mu,\mathbf{e},.)) \le \frac{1}{M(\mathbf{e})^2} d(\lambda,\mu).$$

Setting

$$D(\mathbf{e}) = \frac{\sigma(\mathbf{e})M(\mathbf{e})}{m(x,\mathbf{e})} M(T\mathbf{e})^2, \qquad A(\mathbf{e}) = \frac{1}{M(\mathbf{e})^2},$$

we can apply Lemma 7, which ends up the proof. □

4.3 Supercritical regime in Kimmel's branching model

Let us illustrate the choice of a relevant function f_n and apply Theorem 3 to Kimmel's branching model [B08] for cell division with parasite infection. We consider two random variables $(Z^{(1)},Z^{(2)})$ taking integer values and for safe of simplicity, we assume that

$$\mathbb{P}(Z^{(1)} \ge 1) = \mathbb{P}(Z^{(2)} \ge 1) = 1, \qquad m_1 = \mathbb{E}(Z^{(1)}) > 1, \quad m_2 = \mathbb{E}(Z^{(2)}) \ge 1$$

We also assume that

$$\mathbb{E}(Z^{(1)} \log(Z^{(1)})) < \infty, \quad \mathbb{E}(Z^{(2)} \log(Z^{(2)})) < \infty. \tag{20}$$

Kimmel's branching model describes the division of a cell where in each generation, each parasite reproduces and gives birth independently to a random number of parasites, distributed as $Z^{(1)} + Z^{(2)}$, $Z^{(1)}$ of which go in the first daughter cell and $Z^{(2)}$ of which go in the second daughter cell. Here $\mathcal{X} = \mathbb{N}^*$ and the cell population is the binary tree, so

$$\mathbb{G}_n = \{1,2\}^n.$$

We start with one sigle cell with one single parasite $Z_\varnothing = \delta_1$ and for any $u \in \mathbb{G}_n$,

$$(X(u1),X(u2)) = \sum_{i=1}^{X(u)} Z_i(u),$$

where $(Z_i(u) : i \ge 1)$ are i.i.d. and distributed as $(Z^{(1)},Z^{(2)})$.

Proposition 3. *For any $\epsilon > 0$,*

$$\lim_{n\to\infty} \frac{\#\{u \in \mathbb{G}_n : |\log(X(u))/n - L| \le \epsilon\}}{2^n} = 1$$

in probability, with

$$L = \frac{1}{2}\log(m_1) + \frac{1}{2}\log(m_2).$$

This result ensures that in most of the cells $u \in \mathbb{G}_n$, the infection $X(u)$ grows exponentially with rate L.

Proof. Here $m_n(x,\mathbf{e}) = 2^n$ and (1) becomes

$$Q_n(x,\mathbf{e},dy) = \frac{1}{2}m_1(x,\mathbf{e},dy) = \frac{1}{2}(\mathbb{P}(X(u1) \in dy|X(u) = x) + \mathbb{P}(X(u2) \in dy|X(u) = x)$$

and the auxiliary Markov chain $(Y_i^{(n)} : i = 0,\dots,n) = (Y_i : u = 0,\dots,n)$ is a branching process in random environment, whose reproduction law is distributed as $Z^{(1)}$ with probability $1/2$ and as $Z^{(2)}$ with probability $1/2$. Then, under assumption (20), [AK71] ensures that for any initial state $Y_0 = i \in \mathbb{N}^*$,

$$\frac{Y_n}{\Pi_{i=0}^{n-1} M_i} \to W,$$

where W is a finite positive random variable and $(M_i : i \geq 0)$ is a sequence of i.i.d. random variable such that $\mathbb{P}(M_0 = m_1) = \mathbb{P}(M_0 = m_2) = 1/2$. In particular, $\log(Y_n)/n$ converges a.s. to L and for any continuous and bounded function f, writing $f_n(x) = \log(x)/n$ and $\mu = \delta_L$,

$$Q_{i,n}(x,\mathbf{e}, f \circ f_n) \stackrel{n\to\infty}{\longrightarrow} \mu(f)$$

for any $x \in \mathbb{N}^*$ and $i \in \mathbb{N}$. Thus Assumption 2 (a) is fulfilled. Moreover Assumption 3 (a) is easily checked since $\mathbb{G}_n$ is the binary tree and $m_n(.,.,.) = 2^n$.
Then we can apply Theorem 3 to get

$$\frac{\sum_{u\in\mathbb{G}_n} f(\log(X_u)/n)}{2^n} \stackrel{n\to\infty}{\longrightarrow} f(L)$$

in probability, where we use again $\#G_n = m_n(.,.,.) = 2^n$. □

4.4 Further comments

We finally mention some links with classical branching models.

About multitype branching processes. When the state space $\mathcal{X}$ is finite, the process Z is a multitype branching process and much finer results can be obtained. In particular, the limit behavior of $Z_n/m_n(x,\mathbf{e},\mathcal{X})$ is known, see e.g. [KLPP97] in fixed environment and [C89] in random environment. Let us yet stress that we provide in Lemma 2 a slightly different spine decomposition than [KLPP97, GB03], without projection with respect to the eigenvector associated to the mean operator or weighted paths. In Finally, we note that the previous results (and in particular Corollary 7) can be applied to branching processes in varying environment, when for example the mean growth rate decreases to 1 to mimic the effect of resources limitation.

About branching random walks and random environment. Branching random walks have been largely studied from the pioneering works of Biggins (see e.g. [B77]) and central limit theorems have been obtained to describe the repartition of the population for large times [B90].
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). To use results of the previous Section, we first check that a central limi theorem hold, i.e. the weak convergence

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

where the limit W does not depend on the initial state $x \in \mathcal{X}$. Indeed, we can then 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 $\mathbb{P}(W \in [a,b])$ as soon as $\mathbb{P}(W \in \{a,b\}) = 0$. In other words, one need to check that the auxiliary process satisfies a central limit theorem. We refer to [BH13] Section 3.4 for some examples 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.
One can actually directly derive some (rougher) law of large numbers from the speed of random walks (in environment), i.e. using $Y_n/a_n \Rightarrow v$. As an example, we recall that in dimension 1, the random walk in random environment Y may be sub-ballistic and $b_n = n^\gamma$ with $\gamma \in (0,1)$. We finally mention [N11] when the offspring distribution is chosen in an i.i.d. manner for each time n and location $x \in \mathbb{Z}$.

5 Appendix : fixed environment and spectral gap of the mean operator

In this section, we consider a fixed environment and we compare our statements with classical results. Thus, we set $\mathbb{P} := \mathbb{P}_{\mathbf{e}}$ and

$$m_n(x,.) := m_n(x,\mathbf{e},.), \qquad m_n(\mu,.) = \int_{\mathcal{X}} \mu(dx) m_n(x,.)$$

for any $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathcal{X})$. We consider a subspace $\mathfrak{M}$ of $\mathcal{M}(\mathcal{X})$ stable by addition which contains $\mathcal{M}_1(\mathcal{X})$. We endow $\mathfrak{M}$ with a norm $\| \, . \, \|_{\mathfrak{M}}$ and assume that there exists $c > 0$ such that $\| \mu \|_{\mathfrak{M}} \leq c\mu(\mathcal{X})$ for any $\mu \in \mathfrak{M}$ and that $\mu \to m_1(\mu,.)$ is a bounded endomorphism on $(\mathfrak{M}, \| \, . \, \|_{\mathfrak{M}})$. We denote by $\mathfrak{M}'$ the topological dual of $\mathfrak{M}$ and require the following spectral properties.

Assumption 4. There exists $(\lambda, \mu_0, f_0) \in (1,\infty] \times \mathcal{M}_1(\mathcal{X}) \times \mathfrak{M}'$ such that $f_0(\mu) > 0$ for any non zero measure μ of $\mathfrak{M}$ and

$$m_1(\mu_0,.) = \lambda \mu_0(.), \qquad f_0(m_1(.,dx)) = \lambda f_0(.).$$

Moreover, there exists $a < \lambda$ and $c > 0$ such that

$$\| m_n(\mu,.) - \lambda^n f_0(\mu)\mu_0(.) \|_{\mathfrak{M}} \leq c a^n \| \mu - f_0(\mu)\mu_0(.) \|_{\mathfrak{M}} .$$

When $\mathcal{X}$ is finite, Perron Frobenius theory ensures that the previous Assumptions hold if the matrix given by the mean operator m_1 is aperiodic and irreducible. We refer to [S01] for details and extension to a denumerable state space $\mathcal{X}$. Moreover Krein Rutman Theorem allows to tackle the non-denumerable framework when the mean operator is compact and positive. Let us finally note that several technics in analysis allow to go beyond these assumptions via the decompositions of the operator, see [MS14], where an

overview of the results in the continuous time framework is given.
The previous assumption ensures that for any non-negative function f such that $\mu_0(f) \in (0,\infty)$ and any $\mu \in \mathfrak{M}$,

$$m_n(\mu, f) \sim \lambda^n f_0(\mu)\mu_0(f) \qquad \text{and} \qquad m_n(\mu, \mathcal{X}) \sim f_0(\mu)\lambda^n \qquad (n \to \infty).$$

Proposition 4. *Let $f \in \mathcal{B}(\mathcal{X})$ bounded and $x_0 \in \mathcal{X}$. If Assumption 4 hold and $\sup_{x\in\mathcal{X}} \mathbb{E}(N(x)^2) < \infty$, then*
(i) $Z_n(\mathcal{X})/m_n(x_0, \mathcal{X})$ is bounded in $L^2_{\delta_{x_0}}$;
(ii) $f_0(Z_n)/\lambda^n$ converges $\mathbb{P}_{\delta_{x_0}}$ a.s. to $W \in [0,\infty)$ and $\mathbb{P}_{\delta_{x_0}}(W > 0) > 0$;
(iii) $Z_n(f)/Z_n(\mathcal{X}) \longrightarrow \mu_0(f)$ as $n \to \infty$, $\mathbb{P}_{\delta_{x_0}}$ a.s. on the event $\{W > 0\}$.

Proof. First, we recall that $f_0(Z_n)/\lambda^n$ is martingale since f_0 is the eigenvector of the adjoint of the mean operator. Indeed, denoting by

$$Z_{1,u} = \sum_{i=1}^{N(u)} \delta_{X(ui)}$$

for $u \in \mathbb{G}_n$ the punctual measures associated to the offsprings of each individual in generation n and combining the linearity of f_0 and $\mathbb{E}$, we get :

$$\begin{aligned} \mathbb{E}(f_0(Z_{n+1}) \mid \mathcal{F}_n) &= \mathbb{E}(\sum_{u\in\mathbb{G}_n} f_0(Z_{1,u}) \mid \mathcal{F}_n) = \sum_{u\in\mathbb{G}_n} f_0(\mathbb{E}(Z_{1,u} \mid \mathcal{F}_n)) \\ &= \sum_{u\in\mathbb{G}_n} \lambda f_0(\delta_{X(u)}) = \lambda f_0(Z_n). \end{aligned}$$

Using Assumption 4, we obtain (12). Recalling $\lambda > 1$, (13) is satisfied. Then we can apply Lemma 6 and (11) hold. It ensures that $Z_n(\mathcal{X})/m_n(x,\mathcal{X})$ is bounded in L^2 and so does $f_0(Z_n)/\lambda^n$ since f_0 is bounded and $\| Z_n \|_X \leq cZ_n(\mathcal{X})$. We deduce that the martingale limit of $f_0(Z_n)/\lambda^n$ is non-degenerated and $(i)-(ii)$ are proved. Let us now focus on

$$Q_n(x, f) = m_n(x, f)/m_n(x, \mathcal{X}).$$

Using the second part of Assumption 4 and f_0 bounded, there exist constant c', c'' such that for every $y \in \mathcal{X}$

$$|Q_n(y, f) - \mu_0(f)| \leq c'(a/\lambda)^n \| \delta_y - f_0(\delta_y)\mu_0 \|_{\mathfrak{M}} \leq c''(a/\lambda)^n$$

and

$$|Q_n(y, f) - Q_n(z, f)| \leq 2c''(a/\lambda)^n.$$

It ensures that condition (10) hold since $a < \lambda$ and we can apply Theorem 4 to get

$$\frac{Z_n(f) - Z_n(\mathcal{X})\mu_0(f)}{m_n(x, \mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\delta_x} \quad \text{a.s.}$$

Adding that $\liminf_{n\to\infty} Z_n(\mathcal{X})/m_n(x,\mathcal{X}) > 0$ a.s. on the event $\{W > 0\}$ since f_0 is bounded ends up the proof. □

Let us give some additional comments. Checking that the martingale limit W is non-degenerated is delicate in general. Classical $N \log N$ moment assumptions for single type population in Kesten Stigum theorem (see [LPP95, KLPP97]) can be extended to the case $\#\mathcal{X} < \infty$ [KLPP97] and less explicit but more general criteria can be found in [A00] for $\#\mathcal{X} = \infty$. Here, we have used the L^2 computations of the previous section to get a strong law of large numbers for a general state space $\mathcal{X}$.

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.

[BCGH93] J.-B. Baillon, P. Clément, A. Greven, F.A. den Hollander (1993). Variational approach to branching random walk in random environment. *Ann. Probab.* 21, no. 1, 290-317.

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

[BDMT11] V. Bansaye, J.-F. Delmas, L. Marsalle, 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. *ESAIM* 2015.

[BC15] V. Bansaye, A. Camanès (2016). Aging branching process and queuing for an infinite bus line. *Available on Arxiv* via http://arxiv.org/abs/1506.04168 .

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

[BCN99] J. D. Biggins, H. Cohn, O. Nerman (1999). Multi-type branching in varying environment. *Stoch. Proc. Appl.* 83, 357-400.

[B90] J. D. Biggins (1990). The central limit theorem for the supercritical branching random walk and related results. *Stoch. Proc. Appl.* **34**, 255-274.

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

[CR88] B. Chauvin, A. Rouault (1988). KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probab. Theory Related Fields*, 80(2) : 299-314.

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

[C89] H. Cohn (1989). On the growth of the supercritical multitype branching processes in random environment. *Ann. Probab.* Vol. 17, No 3. 1118-1123.

[CP07a] F. Comets, S. Popov (2007). On multidimensional branching random walks in random environment. *Ann. Probab.* 35, no. 1, 68-114.

[CP07b] 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.)

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

[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.

[K77] O. Kallenberg (1977). Stability of critical cluster fields. *Math. Nachr.* 77, 7-43.

[FK60] H. Furstenberg, H. Kesten (1960). Products of random matrices. *Ann. Math. Stat.*, 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.

[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.

[G76] R. T. Goettge (1976). Limit theorems for the supercritical Galton-Watson process in varying environments. *Math. Biosci.* 28, no. 1-2, 171-190.

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

[HM08] M. Hairer, J. C. Mattingly (2008). Yet another look at Harris' ergodic theorem for Markov chains. *Avialable via* arXiv:0810.2777.

[HR12] S. C. Harris, M. I. Roberts (2012). The many-to-few lemma and multiple spines. in *Ann. Inst. Henri Poincaré*, 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 *Electr. Comm. Probab.*, 19, 28.

[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 multi-type 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.

[K72] T. Kurtz (1972). Inequalities for law of large numbers. *Ann. Math. Statist.* 43,1874-1883.

[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.

[L97] Lyons. A simple path to Biggins' martingale convergence for branching random walk. In *Classical and modern branching processes* (Minneapolis, MN, 1994), volume 84 of IMA Vol. Math. Appl., pages 217-221. Springer, New York, 1997.

[MT09] S. Meyn, L. Tweedie (2009). *Markov Chains and Stochastic Stability*. Broché.

[M71] C. J. Mode (1971). *Multi-Type Branching Processes-Theory and Application*, American Elsevier Publishing Company, Inc., New York, 300 pages, 1971.

[M67] S.-T.C. Moy (1967). Extensions of a limit theorem of Everett Ulam and Harrison multi-type branching processes to a branching process with countably many types. *Ann. Math. Statist.* 38. 992-999.

[MS14] S. Mischler, J. Scher (2013). Spectral analysis of semigroups and growth-fragmentation equations To appear in *Ann. Inst. H. Poincaré Anal. Non Linéaire* and *available on Arxiv* http://arxiv.org/abs/1310.7773.

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

[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.

[S01] E. Seneta (2006). *Non-negative matrices and Markov chains*. Springer Series in Statistics. No. 21

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

[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.