ESAIM: Probability and Statistics www.esaim-ps.org

ESAIM: PS 19 (2015) 307–326 DOI: 10.1051/ps/2014027

WEAK LAW OF LARGE NUMBERS FOR SOME MARKOV CHAINS ALONG NON HOMOGENEOUS GENEALOGIES

VINCENT BANSAYE¹ AND CHUNMAO HUANG²

Abstract. We consider a population with non-overlapping generations, whose size goes to infinity. It is described by a discrete genealogy which may be time non-homogeneous and we pay special attention to branching trees in varying environments. A Markov chain models the dynamic of the trait of each individual along this genealogy and may also be time non-homogeneous. Such models are motivated by transmission processes in the cell division, reproduction-dispersion dynamics or sampling problems in evolution. We want to determine the evolution of the distribution of the traits among the population, namely the asymptotic behavior of the proportion of individuals with a given trait. We prove some quenched laws of large numbers which rely on the ergodicity of an auxiliary process. A central limit is also established in the transient case.

Mathematics Subject Classification. 60J10, 60K37, 60J80, 60F15.

Received September 7, 2013. Revised June 11, 2014.

1. Introduction

In this paper, we consider Markov chains which are indexed by discrete trees. This approach is motivated in particular by the study of structured populations. Thus the tree describes the genealogy of the population in discrete time, with non overlapping generations and the nodes of the tree are the individuals. We consider a trait in the population, which could be the location of the individual, its phenotype, its genotype or any biological characteristic. Letting this trait evolve as a Markov chain and be transmitted to the offspring with a random transition leads us to consider a Markov chain indexed by the genealogical tree. Such a process can also be regarded as a branching particle system where the offspring of each particle is given by the genealogy and the associated traits by the Markov chain.

The evolution of a time homogeneous Markov chain indexed by a binary tree is well known thanks to the works of [4,19]. As soon as the Markov chain along a random lineage of the binary tree is ergodic, a law of large numbers holds. It yields the convergence for the proportions of individuals in generation n whose trait has some given value. More specifically, this asymptotic proportion is given by the stationary measure of the ergodic Markov chain. This convergence holds in probability in general, and under additional assumptions on the speed of convergence of the Markov chain or uniform ergodicity, it also holds almost surely. Such results have been extended and modified to understand the (random) transmission of some biological characteristic of

Keywords and phrases. Markov chain, random environment, branching processes, law of large numbers.

¹ Ecole Polytechnique, CMAP, 91128 Palaiseau, France. bansaye@cmap.polytechnique.fr

² Harbin institute of technology at Weihai, Department of mathematics, Weihai 264209, P.R. China. cmhuang@hitwh.edu.cn.

dividing cells such as cellular aging, cell damages, parasite infection... In particular, [5] considered non ergodic Markov chains and rare events associated to the Markov chain for cell division with parasite infection. In [14], the authors considered a Markov chain indexed by a Galton–Watson tree, which is motivated by cellular aging when the cells may die. In the same vein, the almost sure convergence in the case of bifurcating autoregressive Markov chain is achieved in [8] via martingale arguments. Such results have been extended recently and one can see the works of De Saporta et al. For biological motivations in this vein, we also refer to [32,34].

In this paper, we consider similar questions stated above for the case where both the genealogical tree and the Markov chain along the branches are time non-homogeneous. In particular, we are motivated by the fact that the cell division is affected by the media. This latter is often time non-homogeneous, which may be due to the variations of the available resources or the environment, a medical treatment... Such phenomena are well known in biology from the classical studies of Gause about Paramecium or Tilman about diatoms. The cell genealogy may be modelled by a Galton-Watson process in a varying (or random) environment. It is quite straightforward to extend the weak law of large numbers to the case of non-homogeneous genealogies if the branching events are symmetric and independent (each child obtains an i.i.d. copy) and the Markov chain along the branches is time homogeneous. However, as the convergence of non-homogeneous Markov chains is a delicate problem, we need to consider new limit theorems to understand the evolution of the traits in the cell population. As stated in the next section, the asymptotic proportion can still be characterized as the stationary probability of an auxiliary Markov chain, in the same vein as [14]. It yields a natural interpretation of the repartition of the traits as a stationary probability and the description of the lineage of a typical individual, which then can be easily simulated. A large literature also exists concerning asymptotic behavior of even-odd Markov chains along time homogeneous trees (see e.q. [30]), with different motivations. We stress that in our model the trait of the cell does not influence its division, which means that the genealogical tree may be random but does not depend on the evolution of the Markov chain along its branches. When such a dependence holds (in continuous time, with fixed environments), some many to one formulas can be found in [20] and asymptotic proportions were briefly considered in [6].

Letting the trait be (replaced by) the location of the individual, the process considered here is more usually called a Branching Markov Chain. The particular case that the motion of each individual has i.i.d. increments, i.e. branching random walks, has been largely studied from the pioneering works of Biggins [9-11]. The density occupation, the law of large numbers, central limit theorems, large deviations and the positions of the extremal particles have been considered. These results, such as recurrence, transience or survival criteria, have been partially extended to random environment both in time and space, see e.g. [12, 13, 18, 22, 27, 28, 35]. Here, we consider both the case of non i.i.d. displacements and non homogeneous environments. We mainly focus on (positive) recurrent branching random walks. This provides some tractable models for reproduction-dispersion of species which evolve in a spatially and temporally non-homogeneous environment and compact state space. With our assumptions, the time environment may influence both the reproduction and the dispersion. One can figure out the effects of the humidity and the enlightenment for the reproduction of plants and the wind for pollination. The spatial environment (such as the intensity of wind, the relief...) may here influence only the dispersion. The fact that the space location does not influence the reproduction events requires space homogeneity of some environmental parameters, such as the light exposure, quality of the ground. An extension to space dependent reproduction is a challenging problem. As a motivating article in this vein in ecology, we refer for instance to [16].

Finally, such models might be a first step to consider evolution processes on larger time scale with time inhomogeneity. The trait would then correspond to a phenotype or a genotype. The fact that the branching event does not depend on the trait (neutral theory) may hold in some cases or be used as a zero hypothesis, see for example [23]. More generally, the non-homogeneity of the branching rates in the genealogies raises many difficulties but has various motivations. As an example, we refer to [33] for discussions on time non-homogeneity for extinction and speciation.

This paper is arranged as follows. We first give some basic notations in Section 1. Then we state the results of the paper and consider some applications in Sections 3 and 4. In Section 3, we give a very general statement

which ensures the convergence in probability of the proportions of individuals with a given trait as time goes to infinity. The fact that the common ancestor of two individuals is not recent yields a natural setting for law of large numbers. Here the genealogical tree may be very general but the assumptions that we need are often not satisfied and the asymptotic proportions are not explicit. That's why we focus next on time non-homogeneous tree with branching properties (Sect. 4). That allows us to get a many to one formula, in the same vein as [6,15,19,20] (see Lem. 6.1). We can then state two weak laws of large numbers, one for the population of the nth generation, and the other for the whole population. After that, we derive a central limit theorem and apply it to some branching random walks in random environments. Finally, the rest of the paper (Sects. 5 and 6) is dedicated to the proofs.

2. Notations

In the whole paper, we need the following notations. Let $(\mathcal{X}, B_{\mathcal{X}})$ be a measurable space. The process starts with an initial single individual \emptyset whose trait $X(\emptyset) \in \mathcal{X}$ has distribution ν . This initial individual produces a random number $N = N(\emptyset) < \infty$ of particles of generation 1, denoted by $1, 2, \ldots, N$, with traits determined by

$$\mathbb{P}(X(1) \in dx_1, \dots, X(k) \in dx_k | N = k, X(\emptyset) = x) = p^{(k)}(0)(x, dx_1, \dots, dx_k),$$

where for each $k, n \in \mathbb{N}$ and $x \in \mathcal{X}$, $p^{(k)}(n)(x, \cdot)$ is a probability measure on $(\mathcal{X}^k, B_{\mathcal{X}^k})$. More generally, each individual $u = u_1 \dots u_n$ of generation n whose trait is X(u) yields $N(u) < \infty$ offspring in generation n + 1, denoted by $u1, u2, \dots, uN(u)$, whose traits are determined by

$$\mathbb{P}(X(u1) \in dx_1, \dots, X(uk) \in dx_k | N(u) = k, X(u) = x) = p^{(k)}(n)(x, dx_1, \dots, dx_k).$$

The individuals of each generation evolve independently, so the process enjoys the branching property. Let $\mathbb{N} = \{1, 2, \ldots\}$ and write

$$U = \bigcup_{n=0}^{\infty} \mathbb{N}^n$$

for the set of all finite sequences $u = u_1 \dots u_n$ with $u_i \in \mathbb{N}$, where by convention $\mathbb{N}^0 = \{\emptyset\}$ contains the null sequence \emptyset . If $u = u_1 \dots u_n \in U$, we write |u| = n for the length of u; if $u = u_1 \dots u_n, v = v_1 \dots v_m \in U$, we write $uv = u = u_1 \dots u_n v_1 \dots v_m$ for the sequence obtained by juxtaposition. In particular, $u\emptyset = \emptyset u = u$. We partially order U by writing $u \leq v$ to mean that for some $u' \in U$, v = uu', and by writing u < v to mean that $u \leq v$ and $u \neq v$. Moreover, if u, v are two sequences, we write $u \wedge v$ for the maximal common sequence (common ancestor) of u and v, namely $v = u \wedge v$ if v = v is the maximal sequence such that $v \leq u$ and $v \leq v$.

Let $\mathbb{T} \subset U$ be the genealogical tree rooted at \emptyset with defining elements (N(u)) $(u \in U)$ satisfying: (i) $\emptyset \in \mathbb{T}$; (ii) if $u \in \mathbb{T}$, then $uk \in \mathbb{T}$ if and only if $1 \le k \le N(u)$; (iii) if $uk \in \mathbb{T}$, then $u \in \mathbb{T}$. Denote

$$\mathbb{T}_n = \{ u \in \mathbb{T} : |u| = n \}$$

the set of all individuals in generation n. Let

$$Z_n = \sum_{u \in \mathbb{T}_n} \delta_{X(u)}$$

be the counting measure of particles of generation n. In fact, for any measurable set A of \mathcal{X} ,

$$Z_n(A) = \#\{u \in \mathbb{T}_n : X(u) \in A\}$$

denotes the number of individuals whose traits belong to A. In particular, we write $N_n = Z_n(\mathcal{X})$. Since $N(u) < \infty$ for all u, we have $N_n < \infty$ for all n. Our purpose is to investigate the asymptotic proportion of individual whose trait belongs to A, which is given by $Z_n(A)/N_n$.

3. Weak law of large numbers for non-homogeneous trees

In this section, the genealogical tree \mathbb{T} is fixed (non random). We require that the size of the population in generation n goes to infinity as $n \to \infty$.

We consider a transition kernel $(p^{(k)}(n)(x, dx_1, \dots, dx_k) : k, n \ge 0)$. Then the Markov chain X along the tree \mathbb{T} is specified recursively by

$$\mathbb{E}\left[\prod_{u\in\mathbb{T}_n} F_u(X(u1),\dots,X(uN(u))) \,\Big| (X(v):|v|\leq n)\right] = \prod_{u\in\mathbb{T}_n} \int F_u(x_1,\dots,x_{N(u)}) p^{(N(u))}(n)(x,\mathrm{d}x_1,\dots\mathrm{d}x_k).$$
(3.1)

where $(F_u : u \in \mathbb{T}) \in \mathcal{B}_b^{\mathbb{T}}$ and \mathcal{B}_b is the set of bounded measurable functions from $\bigcup_{k>0} \mathcal{X}^k$ to \mathbb{R} . The trees rooted at u are defined similarly:

$$\mathbb{T}(u) = \{uv \in \mathbb{T}\}, \qquad \mathbb{T}_n(u) = \{uv \in \mathbb{T}_n(u) : |v| = n\},$$

$$Z_n^{(u)} = \sum_{uv \in \mathbb{T}_n(u)} \delta_{X(uv)}, \qquad N_n(u) = Z_n^{(u)}(\mathcal{X}) = \#\mathbb{T}_n(u).$$

Proposition 3.1. Let $A \in B_{\mathcal{X}}$. We assume that

- (i) $N_n \to \infty$ as $n \to \infty$;
- (ii) $\limsup_{\substack{n\to\infty\\ chosen \ in \ \mathbb{T}_n;}} \mathbb{P}(|U_n \wedge V_n| \geq K) \to 0 \ as \ K \to \infty, \ where \ U_n, \ V_n \ are \ two \ individuals \ uniformly \ and \ independently$
- (iii) there exists $\mu(A) \in \mathbb{R}$ such that for all $u \in \mathbb{T}$ satisfying $N_n(u) > 0$ for all $n \geq 1$, and for all $x \in \mathcal{X}$,

$$\lim_{n\to\infty}\mathbb{P}\left(X(U_n^{(u)})\in A\bigg|X(u)=x\right)=\mu(A),$$

where $U_n^{(u)}$ denotes an individual uniformly chosen in $\mathbb{T}_n(u)$.

Then

$$\frac{Z_n(A)}{N} \to \mu(A)$$
 in \mathbb{P} -probability.

The assumption (ii) means that the common ancestor of two individuals chosen randomly is at the be beginning of the tree. Then assumption (iii) ensures that any sampling is giving the same distribution. The assumptions (i) and (ii) hold for many classical genealogies, such as supercritical branching genealogies. The next proposition relax assumption (ii) but requires strong ergodicity. The assumption (iii) is difficult to obtain in general. We give below a simple example where it holds, whereas the the next section is giving better sufficient conditions, in the branching framework.

Example 3.2 (Homogeneous kernels). The assumption (iii) is satisfied as soon as

$$p(x, dy) := \frac{1}{k} \sum_{i=1}^{k} p^{(k)}(n) \left(x, \mathcal{X}^{i-1} dy \mathcal{X}^{k-i} \right)$$

depends neither of k nor of n and $\mathbb{P}_x(Y_n \in A) \to \mu(A)$ as $n \to \infty$ for every $x \in \mathcal{X}$, where Y_n is a Markov chain with transition kernel p. This latter convergence is related to the ergodicity of the Markov chain Y, for which sufficient conditions are known, see e.q. [26]. One can mention two particular cases. First, when \mathbb{T} is the binary tree, we are in the usual framework handled in [19]. Second, one can have $p^{(k)}(n)(x, dx_1, \dots, dx_k) =$ $\prod_{i=1}^k p(x, \mathrm{d}x_i).$

In the forthcoming result, we note that the common ancestor does not need to be at the beginning of the tree, but the convergence required is stronger.

Proposition 3.3. Let $A \in B_{\mathcal{X}}$. We assume that

- (i) $N_n \to \infty$ as $n \to \infty$;
- (ii) $\limsup_{n\to\infty} \mathbb{P}(|U_n \wedge V_n| \geq n-K) \to 0$ as $K\to\infty$, where U_n , V_n are two individuals uniformly and independently chosen in \mathbb{T}_n ;
- (iii) there exists $\mu(A) \in \mathbb{R}$ such that

$$\lim_{n \to \infty} \sup_{\substack{u \in \mathbf{\Gamma}_n^+ \\ x \in \mathcal{X}}} \left| \mathbb{P}\left(X(U_n^{(u)}) \in A \middle| X(u) = x \right) - \mu(A) \right| = 0,$$

where $\Gamma_n^+ = \{u \in \mathbb{T} : |u| \le n, \ N_n(u) > 0\}$ and $U_n^{(u)}$ is an individual uniformly chosen in $\mathbb{T}_n(u)$.

Then

$$\frac{Z_n(A)}{N_n} \to \mu(A)$$
 in \mathbb{P} -probability.

Contrarily to the previous proposition, we note that under the assumption (i), the assumption (ii) is satisfied for any tree \mathbb{T} where each individual has at most q (constant) offspring. It is also satisfied for Wright Fisher (or Moran) genealogies when considering the backward genealogy $\mathbb{T}^{(n)}$ with $\#\mathbb{T}^{(n)}_n \to \infty$. Considering the homogeneous case described in Example 3.2, (iii) is satisfied when strong ergodicity holds for Y. It it the case for example in finite state space or under Doeblin type conditions. In this situation, Proposition 3.3 can be applied.

4. Quenched Law of large numbers for branching Markov chains in random environment

In this section, the genealogical tree \mathbb{T} is random. Then the population evolves following a branching process in random environment (BPRE), described as follows. Let $\xi = (\xi_0, \xi_1, \ldots)$ be a sequence of random variables taking values in some measurable space Ω , which will come in applications below from a stationary and ergodic process. Each ξ_n corresponds to a probability distribution on $\mathbb{N} = \{0, 1, 2, \ldots\}$, denoted by $p(\xi_n) = \{p_k(\xi_n) : k \geq 0\}$. This infinite vector ξ is called a random environment.

We consider now random measurable transition kernels $(p_{\xi_n}^{(k)}(x, dx_1, \dots, dx_k) : k, n \geq 0)$, which are indexed by the *n*th environment component ξ_n . The process X is a Markov chain along the random tree \mathbb{T} with transition kernels p. Conditionally on (ξ, \mathbb{T}) , the process is constructed following (3.1) with $p_{\xi_n}^{(k)}$ in place of $p^{(k)}(n)$. More specifically, the successive offspring distributions are $\{p(\xi_n)\}$, so that the number of offspring N(u) of individual u of generation n is distributed as $p(\xi_n)$ and the traits of its offspring $\{X(ui)\}$ are determined by

$$\mathbb{P}_{\xi}(X(u1) \in dx_1, \dots, X(uk) \in dx_k | N(u) = k, X(u) = x) = p_{\xi_n}^{(k)}(x, dx_1, \dots, dx_k).$$

We note that the offspring number N(u) does not depend on the parent's trait X(u), and the offspring traits $\{X(ui): i=1,\ldots,N(u)\}$ may depend on N(u),X(u) and ξ_n .

Given ξ , the conditional probability will be denoted by \mathbb{P}_{ξ} and the corresponding expectation by \mathbb{E}_{ξ} . The total probability will be denoted by \mathbb{P} and the corresponding expectation by \mathbb{E} . As usual, \mathbb{P}_{ξ} is called *quenched law*, and \mathbb{P} annealed law.

Let $\mathcal{F}_0 = \mathcal{F}(\xi) = \sigma(\xi_0, \xi_1, \ldots)$ and $\mathcal{F}_n = \mathcal{F}_n(\xi) = \sigma(\xi_0, \xi_1, \ldots, (N(u) : |u| < n))$ be the σ -field generated by the process ξ and the random variables N(u) with |u| < n, so that N_n , the size of the population in generation n, are \mathcal{F}_n -measurable. Denote

$$m_n = \sum_k k p_k(\xi_n)$$
 for $n \ge 0$,

and assume that $0 < m_n < \infty$ for all n. Set

$$P_0 = 1$$
 and $P_n = m_0 \dots m_{n-1}$ for $n \ge 1$.

Thus, for every $n \in \mathbb{N}$, $P_n = \mathbb{E}_{\xi} N_n$. It is well-known that the normalized population size $W_n = N_n/P_n$ is a non-negative martingale with respect to \mathcal{F}_n , so the limit $W = \lim_{n \to \infty} W_n$ exists a.s.

In the rest of this section, we make the following classical assumptions for the supercritical regime:

Assumption 4.1.

- (i) The environment $\xi = (\xi_0, \xi_1, ...)$ is a stationary ergodic sequence.
- (ii) We assume that $\mathbb{P}(m_0 = 0) = 0$, $\mathbb{P}(p_0(\xi_0) = 1) < 1$ and $\mathbb{E}(\log m_0) < \infty$.
- (iii) We focus on the supercritical non-degenerated case

$$\mathbb{E}(\log m_0) > 0, \qquad \mathbb{E}\left(\log \frac{\mathbb{E}_{\xi} N^2}{m_0^2}\right) < \infty.$$
 (4.1)

The assumption (i) allows to get the asymptotic behavior of the size of the population. The assumption (ii) avoids some degenerated cases. Denoting by

$$q(\xi) = \mathbb{P}_{\xi}(N_n = 0 \text{ for some } n)$$

the extinction probability, it is well known (cf. [2,3]) that the non-extinction event $\{N_n \to \infty\}$ has quenched probability $1 - q(\xi)$. Moreover, the condition $\mathbb{E}(\log m_0) \le 0$ implies that $q(\xi) = 1$ a.s., whereas $\mathbb{E}(\log m_0) > 0$ (supercritical case) yields

$$q(\xi) < 1$$
 a.s

The assumption (iii) ensures that the limit W is positive on the non-extinction event. Moreover, one can see that $\lim_{n\to\infty} P_n = \infty$, since $\mathbb{E}(\log m_0) > 0$. We refer to [2,3] for the statements and proofs of these results.

4.1. Law of large numbers in generation n

We first give a forward law of large numbers in generation n for the model introduced above, with the help of an auxiliary Markov process constructed as follows. Let

$$P_{\xi_n}^{(k,i)}(x,\cdot) = p_{\xi_n}^{(k)}\left(x, \mathcal{X}^{i-1} \times \cdot \times \mathcal{X}^{k-i}\right)$$

be the *i*th marginal of $p_{\xi_n}^{(k)}$ and introduce the random transition probability

$$Q_n(x,\cdot) = Q(T^n \xi; x, \cdot) = \frac{1}{m_n} \sum_{k=0}^{\infty} p_k(\xi_n) \sum_{i=1}^k P_{\xi_n}^{(k,i)}(x, \cdot),$$

where $T\xi = (\xi_1, \xi_2, ...)$ if $\xi = (\xi_0, \xi_1, ...)$. We note that for each $\xi \in \Omega$, the Markov transition kernel $Q(\xi; \cdot, \cdot)$ is a function from $\mathcal{X} \times B_{\mathcal{X}}$ into [0, 1] satisfying:

- for each $x \in \mathcal{X}$, $Q(\xi; x, \cdot)$ is a probability measure on $(\mathcal{X}, B_{\mathcal{X}})$;
- for each $A \in B_{\mathcal{X}}$, $Q(\xi; \cdot, A)$ is a $B_{\mathcal{X}}$ -measurable function on \mathcal{X} .

Given the environment ξ , we define an auxiliary Markov chain in varying environment Y, whose transition probability in generation j is Q_j :

$$\mathbb{P}_{\xi}(Y_{j+1} \in dy | Y_j = x) = Q_j(x, dy).$$

As usual, we denote by $\mathbb{P}_{\xi,x}$ the quenched probability when the process Y starts from the initial value x, and by $\mathbb{E}_{\xi,x}$ the corresponding expectation.

As stated below, the convergence of the measure $Z_n(\cdot)$ normalized comes from the ergodic behavior of Y_n . In the same vein as [14], we have

Theorem 4.2. Let $A \in B_{\mathcal{X}}$. We assume that there exists a sequence $(\mu_{\xi,n}(A))_n$ such that for almost every ξ and for each $r \in \mathbb{N}$,

$$\lim_{n \to \infty} \left(\mathbb{P}_{T^r \xi, x} (Y_{n-r} \in A) - \mu_{\xi, n}(A) \right) = 0 \quad \text{for every } x \in \mathcal{X}.$$
 (4.2)

Then we have for almost all ξ , conditionally on the non-extinction event,

$$\frac{Z_n(A)}{N_n} - \mu_{\xi,n}(A) \to 0 \qquad in \ \mathbb{P}_{\xi}\text{-probability.}$$

$$\tag{4.3}$$

This theorem is adapted to the underlying branching genealogy. The proof is deferred to Section 6.2, where a more general result is obtained. The condition (4.2) holds if the auxiliary Markov chain is weakly ergodic, for suitable sets A. For sufficient (and necessary) conditions of weak ergodicity in the non-homogeneous case, we refer in particular to [26].

Let us now give more trackable conditions. We derive a result of (quenched) weak law of large numbers, under a stronger assumption, and then give two examples.

Corollary 4.3. Let $A \in B_{\chi}$. We assume that there exists $\mu(A) \in \mathbb{R}$ such that for almost all ξ ,

$$\lim_{n \to \infty} \mathbb{P}_{\xi, x}(Y_n \in A) = \mu(A) \qquad \text{for every } x \in \mathcal{X}.$$
 (4.4)

Then we have for almost all ξ , conditionally on the non-extinction event,

$$\frac{Z_n(A)}{N_n} \to \mu(A) \qquad in \ \mathbb{P}_{\xi}\text{-probability.} \tag{4.5}$$

We now give two examples. In the first one, both reproduction and transition kernels are time-homogeneous. In the second one, the reproduction is time non-homogeneous but each individual inherits i.i.d. positions. In both cases, Y is time homogeneous and (4.4) can be checked easily.

Example 4.4 (Homogeneous Markov chains along Galton Watson trees). We focus here on the case when the time environment is non random, *i.e.* ξ_n is constant for every $n \in \mathbb{N}$. The genealogical tree is a Galton Watson tree, whose offspring distribution is specified by $\{p_k : k \geq 0\}$. Moreover, we assume that

$$\mathbb{P}_{\xi}(X(u1) \in dx_1, \dots, X(uk) \in dx_k | N(u) = k, X(u) = x) = p^{(k)}(x, dx_1, \dots, dx_k).$$

does not depend on ξ . Then, denoting by m the mean number of offspring per individual and

$$P^{(k,i)}(x,\cdot) = p^{(k)}(x,\mathcal{X}^{i-1} \times \cdot \times \mathcal{X}^{k-i}),$$

the auxiliary process Y is a time homogeneous Markov chain whose transition kernel is given by

$$Q(x,\cdot) = \frac{1}{m} \sum_{k=0}^{\infty} p_k \sum_{i=1}^{k} P^{(k,i)}(x,\cdot).$$

Thus a law of large number in probability is obtained as soon as the ergodicity of the Markov chain Y is proved. Indeed, it ensures that there exists a probability measure μ such that for each $x \in \mathcal{X}$ and a measurable set A such that $\mu(\partial A^3) = 0$, and $Q^n(x, A) = \mathbb{P}_x(Y_n \in A) \to \mu(A)$, so that we can apply Corollary 4.3.

We recall that sufficient conditions for the ergodicity of a Markov chain are known in the literatures, see e.g. [25]. This result is a simple generalization of law of large numbers on the binary tree in [19] and that on Galton Watson trees with at most two offspring given in [14]. A continuous time analogue can be found in [7].

 $^{^3 \}partial A$ is the boundary of A.

Example 4.5 (Symmetric homogeneous Markov chains along branching trees in random environment). We consider a Branching Markov Chain on \mathcal{X} in the stationary and ergodic environment ξ . Given ξ , for each u of generation n, the number of its offspring N(u) is determined by distribution $p(\xi_n) = \{p_k(\xi_n) : k \geq 0\}$. The offspring positions $\{X(ui)\}$, independent of each other conditioned on the position of u, are determined by

$$\mathbb{P}_{\varepsilon}(X(ui) \in \mathrm{d}y | X(u) = x) = p(x, \mathrm{d}y),$$

where $p(x,\cdot)$ is a probability on \mathcal{X} . This example is actually a particular case of Example 3.2. We can see that

$$p_{\xi_n}^{(k)}(x, dx_1, \dots, dx_k) = \prod_{i=1}^k \mathbb{P}_{\xi}(X(ui) \in dx_i | X(u) = x) = \prod_{i=1}^k p(x, dx_i).$$

Therefore, $Q(x, dy) = Q_n(x, dy) = p(x, dy)$ and Y_n is a time-homogeneous Markov chain with transition probability p. As in the previous example, the problem is reduced to the ergodicity of Y. Such a class may be relevant to model the dispersion of plants in spaces where the reproduction is homogeneous and the time environment only influences the genealogy.

The convergence of the Markov chain (4.4) is difficult to get under general assumptions. Indeed, when the auxiliary kernel Q depends on ξ , the auxiliary process is time non-homogeneous and the convergence in distribution (4.4) won't hold in general⁴. Thus, we derive in the next subsection a law of large numbers on the whole tree.

4.2. Law of large numbers on the whole tree

In the framework of Markov chains with stationary and ergodic environments, quenched ergodic theorems are known (see e.g. [29,31]). They ensure the convergence

$$\frac{1}{n} \sum_{k=1}^{n} Q_0 \dots Q_{k-1}(x, \cdot) \to \mu(\cdot) \quad \text{as } n \to \infty, \quad \text{for every } x \in \mathcal{X}.$$

It leads us to consider the following limit theorems on the whole tree, where each generation of the tree has the same mean weight in the limit. Such an approach is both adapted to the branching (forward) genealogy and the convergence of the underlying auxiliary time non-homogeneous Markov chain Y, whose transition are stationary and ergodic. It differs from the usual limit theorem on the whole tree [14,19] where each cell has the same weight, but not each generation.

Theorem 4.6. Let $A \in B_{\mathcal{X}}$. We assume that there exists $\mu(A) \in \mathbb{R}$ such that for almost all ξ ,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{P}_{\xi, x}(Y_k \in A) = \mu(A) \quad \text{for every } x \in \mathcal{X}.$$
 (4.6)

Then we have for almost all ξ , conditionally on the non extinction event,

$$\frac{1}{n} \sum_{k=1}^{n} \frac{Z_k(A)}{N_k} \to \mu(A) \qquad in \ \mathbb{P}_{\xi}\text{-probability.}$$

$$\tag{4.7}$$

The proof is also deferred to Section 5.4.

⁴ The reader could consider for example the case of an environment containing only two components whose associated transition matrices have different stationary probability.

Example 4.7 (Doeblin conditions for the auxiliary Markov chain Y). Assume that there exist a positive integer b and a measurable function $M(\xi): \Omega \to [1, \infty)$ such that $\mathbb{E}|\log M(\xi)| < \infty$, and for almost all ξ ,

$$\mathbb{P}_{\xi,x}(Y_b \in A) \le M(\xi)\mathbb{P}_{\xi,y}(Y_b \in A)$$
 for all $x, y \in \mathcal{X}$,

which means that assumption (A) of Seppäläinen [31] is satisfied. According to Theorem 2.8 (i) and (iii) of [31] (with $f = \mathbf{1}_{A \times \Omega}$), there exists a probability Φ on $\mathcal{X} \times \Omega$ such that for almost all ξ ,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\mathbf{1}_A(Y_k)=\varPhi(A)\qquad \mathbb{P}_{\xi,x^-}\text{ a.s. for every }x\in\mathcal{X}.$$

By the dominate convergence theorem, we have for almost all ξ ,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{P}_{\xi, x}(Y_k \in A) = \Phi(A) \quad \text{for every } x \in \mathcal{X}.$$

Thus (4.6) holds with $\mu(A) = \Phi(A)$, so we can use Theorem 4.6 to get (6.7) and (4.7).

4.3. Central limit theorem

When the auxiliary Markov chain Y is a classical random walk on $\mathcal{X} \subset \mathbb{R}$, we know that Y_n satisfies a central limit theorem. Such results have been extended to the framework of random walk in random environment (see e.g. [1]) and some more general Markov chains (see e.g. [17]). It leads us to state the convergence of proportions in the case when Y_n satisfies a central limit theorem.

Theorem 4.8. Let $\mathcal{X} \subset \mathbb{R}$. We assume that for almost all ξ , Y_n satisfies a central limit theorem: there exits a sequence of random variables $\{(a_n(\xi), b_n(\xi)\}\$ satisfying $b_n(\xi) > 0$ such that

$$\lim_{n \to \infty} \mathbb{P}_{\xi, x} \left(\frac{Y_n - a_n(\xi)}{b_n(\xi)} \le y \right) = \Phi(y) \quad \text{for every } x \in \mathcal{X},$$
(4.8)

where Φ is a continuous function on \mathbb{R} . If for each $r \in \mathbb{N}$ fixed,

$$\lim_{n \to \infty} \frac{b_n(\xi)}{b_{n-r}(T^r \xi)} = 1 \qquad and \qquad \lim_{n \to \infty} \frac{a_n(\xi) - a_{n-r}(T^r \xi)}{b_{n-r}(T^r \xi)} = 0 \quad a.s., \tag{4.9}$$

then we have for almost all ξ , conditionally on the non-extinction event,

$$\frac{Z_n(-\infty, b_n(\xi)y + a_n(\xi)]}{N_n} \to \Phi(y) \qquad in \ \mathbb{P}_{\xi}\text{-probability.}$$
(4.10)

The proof is deferred to Section 5.5.

Example 4.9 (Branching random walk on \mathbb{R} in time random environment). This model is considered in Huang and Liu [22]. The environment $\xi = (\xi_n)_{n \in \mathbb{N}}$ is a stationary and ergodic process indexed by time $n \in \mathbb{N}$. Each realization of ξ_n corresponds to a distribution $\eta_n = \eta_{\xi_n}$ on $\mathbb{N} \otimes \mathbb{R}^{\mathbb{N}}$. Given the environment ξ , the process is formed as follows: at time n, each particle u of generation n, located at $X(u) \in \mathbb{R}$, is replaced by N(u) new particles of generation n+1 which scattered on \mathbb{R} with positions determined by $X(ui) = X(u) + L_i(u)$, where the point process $(N(u); L_1(u), L_2(u), \ldots)$ has distribution η_n . To fit with the notations of this paper, we can see that

$$p_{k}(\xi_{n}) = \eta_{n}(k, \mathbb{R} \times \mathbb{R} \times \dots),$$

$$p_{\xi_{n}}^{(k,i)}(x, dy) = \eta_{n}(k, \mathbb{R}^{i-1} \times \{dy - x\} \times \mathbb{R}^{k-i}) =: q_{\xi_{n}}^{(k,i)}(dy - x),$$

$$Q_{n}(x, dy) = \frac{1}{m_{n}} \sum_{k=0}^{\infty} p_{k}(\xi_{n}) \sum_{i=1}^{k} q_{\xi_{n}}^{(k,i)}(dy - x) =: q_{n}(dy - x).$$

We note that for any measurable function f on \mathbb{R} ,

$$\int f(t)q_n(\mathrm{d}t) = \frac{1}{m_n} \mathbb{E}_{\xi} \sum_{i=1}^{N(u)} f(L_i(u)) \quad (\forall u \in \mathbb{T}_n).$$

Hence q_n is the normalized intensity measure of the point process $(N(u); L_1(u), L_2(u), \ldots)$ for $u \in \mathbb{T}_n$. We define

$$Y_n = \zeta_0 + \zeta_1 + \ldots + \zeta_n,$$

where ζ_j is independent of each other under \mathbb{P}_{ξ} and the distribution of ζ_j for $j \geq 1$ is q_{j-1} . Then Y_n is a non-homogeneous Markov chain, whose transition kernel satisfies

$$\mathbb{P}_{\varepsilon}(Y_{n+1} \in dy | Y_n = x) = q_n(dy - x) = Q_n(x, dy).$$

Let $\mu_n = \int_{\mathbb{R}} tq_n(dt)$ and $\sigma_n^2 = \int_{\mathbb{R}} (t - \mu_n)^2 q_n(dt)$. If $|\mu_0| < \infty$ a.s. and $\mathbb{E}(\sigma_0^2) \in (0, \infty)$, according to [22], the sequence (q_n) satisfies a central limit theorem:

$$q_0 * \dots * q_{n-1}(b_n(\xi)y + a_n(\xi)) \to \Phi(y)$$
 a.s.,

where

$$a_n(\xi) = \sum_{i=0}^{n-1} \mu_n, \qquad b_n(\xi) = \left(\sum_{i=0}^{n-1} \sigma_n^2\right)^{1/2}$$

and Φ is the distribution function of the standard normal distribution. It follows that (4.8) holds for almost all ξ . Moreover, by the ergodic theorem, (4.9) can be verified. Thus we can apply Theorem 4.8 to this model and obtain (4.10) under the hypothesis given above. This result can also be deduced from [22], where the almost sure convergence of (4.10) is shown though some tedious calculations.

Example 4.10 (Branching random walk on \mathbb{Z} in time and space random environment). This model is considered in Liu [24]. Let $\xi = (\xi_n)_{n \in \mathbb{N}}$ be a stationary and ergodic process denoting the environment in time, and $\omega = (\omega_x)_{x \in \mathbb{Z}}$, which denotes the environment in locations, be another stationary and ergodic process taking values in [0, 1]. The two sequences ξ , ω are supposed to be independent of each other. Given the environment (ξ, ω) , each u of generation n, located at $X(u) \in \mathbb{Z}$, is replaced at time n+1 by k new particles with probability $p_k(\xi_n)$, which move immediately and independently to x+1 with probability ω_x and to x-1 with probability $1-\omega_x$. Namely, the position of ui is determined by

$$\mathbb{P}_{(\xi,\omega)}(X(ui) = y | X(u) = x) = Q(x,y) := \begin{cases} \omega_x & \text{if } y = x+1; \\ 1 - \omega_x & \text{if } y = x-1, \end{cases}$$

where $\mathbb{P}_{(\xi,\omega)}$ denotes the conditional probability given the environment (ξ,ω) . Notice that when the environment in locations ω is fixed, this process is the just one considered in Example 3.2 with the state space $\mathcal{X}=\mathbb{Z}$ and p(x,y)=Q(x,y). So the transition probability of the Markov chain Y_n is Q, which only depends on the environment in locations ω and is independent of the environment in time ξ . We can regard Y_n as a random walk on \mathbb{Z} in random environment which is studied in Alili [1]. By Theorem 6.3 of [1] and the continuity of Φ , under some hypothesis, we have for every ω ,

$$\lim_{n\to\infty}\mathbb{P}_{\omega,x}\left(\frac{Y_n-n\gamma}{\sqrt{n}}\leq y\right)=\varPhi(y)\qquad\text{for every }x\in\mathbb{Z},$$

where Φ is the distribution function of the normal distribution $\mathcal{N}(0,D)$, and γ,D are two explicit constants (see [1] for details). Therefore, we can apply Theorem 4.8 and obtain (4.10) under the probability $\mathbb{P}_{(\xi,\omega)}$.

5. Proof of Propositions 3.1 and 3.3

Proposition 3.1 is a consequence of the following result with $f(x) = \mathbf{1}_A(x) - \mu(A)$ since the L^2 convergence implies the convergence in probability.

Lemma 5.1. Let f be a bounded measurable function on \mathcal{X} . We assume that

- (i) $N_n \to \infty$ as $n \to \infty$; (ii) $\limsup_{n \to \infty} \mathbb{P}(|U_n \wedge V_n| \ge K) \to 0$ as $K \to \infty$;
- (iii) $for \ all \ u \in \mathbb{T} \ and \ all \ x \in \mathcal{X},$

$$\lim_{n \to \infty} R_u(n, x) = 0,$$

where

$$R_u(n,x) = \begin{cases} \mathbb{E}\left[\left.f(X(U_n^{(u)}))\right| X(u) = x\right], & \text{if } N_n(u) > 0, \\ 0, & \text{if } N_n(u) = 0. \end{cases}$$

Then

$$\frac{\sum_{u \in \mathbb{T}_n} f(X(u))}{N_n} \to 0 \qquad in \ L^2.$$

Proof. We first notice that

$$\mathbb{E}\left(\frac{\sum_{u\in\mathbb{T}_n}f(X(u))}{N_n}\right)^2 = \frac{1}{N_n^2}\mathbb{E}\left[\sum_{u\in\mathbb{T}_n}f^2(X(u))\right] + \frac{1}{N_n^2}\mathbb{E}\left[\sum_{\substack{u,v\in\mathbb{T}_n\\u\neq v}}f(X(u))f(X(v))\right].$$

We need to prove that both terms in the right side go to 0 as $n \to \infty$. For the first term, since f is bounded, there exists a constant C such that $|f| \leq C$. By (i),

$$\frac{1}{N_n^2} \mathbb{E}\left[\sum_{u \in \mathbb{T}} f^2(X(u))\right] = \frac{\mathbb{E}f^2(U_n)}{N_n} \le \frac{C^2}{N_n} \to 0 \quad \text{as } n \to \infty.$$

The second term can be decomposed as

$$\frac{1}{N_n^2} \mathbb{E} \left[\sum_{\substack{u,v \in \mathbb{T}_n \\ u \neq v}} f(X(u)) f(X(v)) \right] = \sum_{r=0}^K \mathbb{E} A_{n,r} + \sum_{r=K+1}^{n-1} \mathbb{E} A_{n,r},$$

where K is a fixed integer suitable large, and

$$A_{n,r} = \frac{1}{N_n^2} \sum_{w \in \mathbb{T}_r} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} \sum_{\substack{wi\bar{u} \in \mathbb{T}_{n_r}(wi) \\ wj\bar{v} \in \mathbb{T}_{n_r}(wj)}} f(X(wi\tilde{u})) f(X(wj\tilde{v})),$$

where $n_r := n - r - 1$. It is clear that

$$\mathbb{E}A_{n,r} = \frac{1}{N_n^2} \sum_{w \in \mathbb{T}_r} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} \sum_{\substack{wi\tilde{u} \in \mathbb{T}_{n_r}(wi) \\ wj\tilde{v} \in \mathbb{T}_{n_r}(wj)}} \mathbb{E}f(X(wi\tilde{u}))f(X(wj\tilde{v}))$$

$$= \sum_{w \in \mathbb{T}_r} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} a_{n,r}(wi, wj) R_{n,r}(wi, wj),$$

where $a_{n,r}(wi, wj) = N_{n_r}(wi)N_{n_r}(wj)/N_n^2$ and

$$R_{n,r}(wi, wj) = \mathbb{E}\left[R_{wi}(n_r, X(wi))R_{wj}(n_r, X(wj))\right].$$

As |f| < C, we have

$$\sum_{r=K+1}^{n-1} \mathbb{E} A_{n,r} \le C^2 \sum_{r=K+1}^{n-1} \sum_{\substack{w \in \mathbb{T}_r \ wi, wj \in \mathbb{T}_1(w) \\ i \ne j}} a_{n,r}(wi, wj) \le C^2 \mathbb{P}(|U_n \wedge V_n| \ge K+1).$$

By (ii), $\limsup_{n\to\infty} \mathbb{P}(|U_n \wedge V_n| \geq K+1) \to 0$ as $K\to\infty$. Thus $\limsup_{n\to\infty} \sum_{r=K+1}^{n-1} \mathbb{E} A_{n,r}$ is negligible for K large enough. For $0\leq r\leq K$, the fact that $R_u(n_r,x)$ goes to zero for all x and is bounded by C with respect to xenables us to apply the dominate convergence theorem and get

$$R_{n,r}(wi, wj) \to 0$$
 as $n \to \infty$.

Adding that $a_{n,r}(wi,wj)$ is bounded by 1 yields $\sum_{r=0}^K \mathbb{E} A_{n,r} \to 0$ as $n \to \infty$. This completes the proof.

Proposition 3.3 comes directly from Lemma 5.2 below, with $f(x) = \mathbf{1}_A(x) - \mu(A)$.

Lemma 5.2. Let f be a bounded measurable function on \mathcal{X} . We assume that

- (i) $N_n \to \infty$ as $n \to \infty$; (ii) $\limsup_{n \to \infty} \mathbb{P}(|U_n \wedge V_n| \ge n K) \to 0$ as $K \to \infty$;
- (iii) $\lim_{n\to\infty} \sup_{u\in\Gamma_n} |R_u(n,X(u))| = 0$, where $\Gamma_n = \bigcup_{n=1}^n \mathbb{T}_k$.

Then

$$\frac{\sum_{u \in \mathbb{T}_n} f(X(u))}{N_n} \to 0 \quad in \ L^2.$$

Proof. Similar to the proof of Lemma 5.1, but we split

$$\sum_{r=0}^{n-1} \mathbb{E} A_{n,r} = \sum_{r=0}^{n-K-1} \mathbb{E} A_{n,r} + \sum_{r=n-K}^{n-1} \mathbb{E} A_{n,r},$$

and show the negligibility of the two terms respectively. By (ii), we first have

$$\sum_{r=n-K}^{n-1} \mathbb{E} A_{n,r} \le C^2 \limsup_{n \to \infty} \mathbb{P}(|U_n \wedge V_n| \ge n - K) \to 0 \text{ as } K \to \infty.$$

Now we consider $\sum_{r=0}^{n-K-1} \mathbb{E} A_{n,r}$. For $r \leq n-K-1$ and $wi, wj \in \mathbb{T}_1(w)$ $(w \in \mathbb{T}_r, i \neq j)$, with $n_r = n-r-1$,

$$|R_{n,r}(wi,wj)| \leq \mathbb{E} |R_{wi}(n_r,X(wi))R_{wj}(n_r,X(wj))| \leq \sup_{k>K} \mathbb{E} \left(\sup_{u\in\Gamma_n} |R_u(k,X(u))|^2 \right).$$

It follows that

$$\sum_{r=0}^{n-K-1} \mathbb{E} A_{n,r} \le \sup_{k \ge K} \mathbb{E} \left(\sup_{u \in \Gamma_n} \left| R_u(k, X(u)) \right|^2 \right) \to 0 \quad \text{as } K \to \infty,$$

by (iii) and the dominate convergence theorem.

6. Proofs for branching Markov chains in random environments

In this section, we focus on the process in random environments and present proofs for the theorems stated in Section 4.

6.1. Proof of the law of large numbers in generation n

Following the definition of Y in Section 4.1, for each $u \in \mathbb{T}$, we define the Markov chain Y(u) associated to u by

$$\mathbb{P}_{\xi}(Y_{i+1}(u) \in \mathrm{d}y | Y_i(u) = x) = Q(T^{|u|+j}\xi; x, \mathrm{d}y).$$

In particular, $Y_n = Y_n(\emptyset)$. Moreover, for any bounded measurable function f on \mathcal{X} , we denote throughout

$$P_{\xi_n}^{(k,i)}f(x):=\int f(y)P_{\xi_n}^{(k,i)}(x,\mathrm{d}y)\qquad\text{and}\qquad Q_jf(x):=\int f(y)Q_j(x,\mathrm{d}y).$$

Firstly, following [14, 19, 20], we can get Lemma 6.1 below which gives the mean relation between $Y_n(u)$ and the traits of individuals in $\mathbb{T}_n(u)$. The proof is omitted

Lemma 6.1 (Many-to-one formula). For each $x \in \mathcal{X}$, $u \in \mathbb{T}$ and any bounded measurable function $f_{\xi,n}$ on \mathcal{X} ,

$$\mathbb{E}_{\xi} \left[f_{\xi,n}(Y_n(u)) \mid Y_0(u) = x \right] = (\mathbb{E}_{\xi} N_n(u))^{-1} \mathbb{E}_{\xi} \left[\sum_{uv \in \mathbb{T}_n(u)} f_{\xi,n}(X(uv)) \middle| X(u) = x \right]. \tag{6.1}$$

We now give a general statement from which we are deriving both Theorems 4.2 and 4.8. Denote

$$F_{\mathcal{E},n} = \nu_{\mathcal{E}} Q_0 \dots Q_{n-1}. \tag{6.2}$$

Proposition 6.2. Let ν_{ξ} be the distribution of $X(\emptyset)$. We assume that for almost all ξ , there exist a function g, an integer $n_0 = n_0(\xi)$ and non negative numbers $(\alpha_n, \beta_n) = (\alpha_n(\xi), \beta_n(\xi))$ such that

- (H1) for all $n \geq n_0$ and $x \in \mathcal{X}$, $\sup_{0 \leq r < n} Q_r \dots Q_{n-1} | f_{\xi,n} | (x) \leq g(x);$ (H2) for every $n \geq n_0$, $F_{\xi,n}(f_{\xi,n}^2) \leq \alpha_n$, and $\alpha_n/P_n \to 0$ as $n \to \infty$;
- (H3) for every $n \in \mathbb{N}$, $F_{\xi,n}\left(J_{T^n\xi}(g \otimes g)\right) \leq \beta_n$ and $\sum_n \frac{\beta_n}{P_n m_n^2} < \infty$, where

$$J_{\xi}(f_{\xi} \otimes g_{\xi})(x) := \mathbb{E}_{\xi,x} \sum_{\substack{i,j \in \mathbb{T}_1 \\ i \neq j}} f_{\xi}(X(i))g_{\xi}(X(j));$$

(H4) for each r fixed, $Q_r \dots Q_{n-1} f_{\xi,n}(x) \to 0$ as $n \to \infty$ for every $x \in \mathcal{X}$.

Then we have for almost all ξ ,

$$\frac{\sum_{u \in \mathbb{T}_n} f_{\xi,n}(X(u))}{P_n} \to 0 \qquad in \ \mathbb{P}_{\xi}\text{-}L^2.$$

Proof. We need to show that

$$\mathbb{E}_{\xi} \left(\frac{\sum_{u \in \mathbb{T}_n} f_{\xi, n}(X(u))}{P_n} \right)^2 \to 0 \quad a.s. \quad \text{as } n \to \infty.$$

Similar to the proof of Proposition 5.1, we write

$$\mathbb{E}_{\xi} \left(\frac{\sum_{u \in \mathbb{T}_n} f_{\xi, n}(X(u))}{P_n} \right)^2 = \frac{1}{P_n^2} \mathbb{E}_{\xi} \left[\sum_{u \in \mathbb{T}_n} f_{\xi, n}^2(X(u)) \right] + \sum_{r=0}^K \mathbb{E}_{\xi} A_{n,r} + \sum_{r=K+1}^{n-1} \mathbb{E}_{\xi} A_{n,r},$$
(6.3)

where

$$A_{n,r} := \frac{1}{P_n^2} \sum_{w \in \mathbb{T}_r} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} \sum_{\substack{wi\tilde{u} \in \mathbb{T}_{n_r}(wi) \\ wj\tilde{v} \in \mathbb{T}_{n_r}(wj)}} f_{\xi,n}(X(wi\tilde{u})) f_{\xi,n}(X(wj\tilde{v})),$$

and $K = K(\xi) (\geq n_0)$ is suitable large. By Lemma 6.1 and condition (H2), for $n \geq n_0$,

$$\frac{1}{P_n^2} \mathbb{E}_{\xi} \left[\sum_{u \in \mathbb{T}_n} f_{\xi,n}^2(X(u)) \right] = \frac{1}{P_n} F_{\xi,n}(f_{\xi,n}^2) \le \frac{\alpha_n}{P_n} \to 0 \ a.s. \ as \ n \to \infty.$$

Again by Lemma 6.1, for every $w \in \mathbb{T}_r$, $i \neq j \in \mathbb{N}$, denoting $n_r = n - r - 1$ in the computation below,

$$B(wi,wj) = \mathbb{E}_{\xi} \left[\sum_{wi\tilde{u} \in \mathbb{T}_{n_r}(wi)} f_{\xi,n}(X(wi\tilde{u})) \sum_{wj\tilde{v} \in \mathbb{T}_{n_r}(wj)} f_{\xi,n}(X(wj\tilde{v})) \middle| \mathcal{F}_{r+1}, X(wi), X(wj) \right]$$

$$= \mathbb{E}_{\xi,X(wi)} \left[\sum_{wi\tilde{u} \in \mathbb{T}_{n_r}(wi)} f_{\xi,n}(X(wi\tilde{u})) \right] \mathbb{E}_{\xi,X(wj)} \left[\sum_{wj\tilde{v} \in \mathbb{T}_{n_r}(wj)} f_{\xi,n}(X(wj\tilde{v})) \right]$$

$$= m_{r+1}^2 \dots m_{n-1}^2 Q_{r+1,n} f_{\xi,n}(X(wi)) Q_{r+1,n} f_{\xi,n}(X(wj)),$$

with the notation $Q_{r,n} = Q_r \dots Q_{n-1}$. Thus,

$$\mathbb{E}_{\xi} A_{n,r} = \frac{1}{P_n^2} \mathbb{E}_{\xi} \left[\sum_{w \in \mathbb{T}_r} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} B(wi, wj) \right] = \frac{1}{P_r m_r^2} F_{\xi,r} \left[J_{T^r \xi} (Q_{r+1,n} f_{\xi,n} \otimes Q_{r+1,n} f_{\xi,n}) \right].$$

For $r \geq K + 1$, by conditions (H1) and (H3),

$$\sum_{r=K+1}^{n-1} \mathbb{E}_{\xi} A_{n,r} \le \sum_{r=K+1}^{\infty} \frac{1}{P_r m_r^2} F_{\xi,r} (J_{T^r \xi}(g \otimes g)) \le \sum_{r=K+1}^{\infty} \frac{\beta_r}{P_r m_r^2} \to 0 \ a.s. \ as \ K \to \infty.$$

It remains to consider $0 \le r \le K$. For almost all ξ , for each r fixed, by (H4),

$$Q_{r+1,n}f_{\xi,n}\otimes Q_{r+1,n}f_{\xi,n}(y,z)\stackrel{n\to\infty}{\longrightarrow} 0$$
 for each $(y,z)\in\mathcal{X}^2$.

Notice that by (H1) and (H3), for $n \ge n_0(\xi)$,

$$F_{\xi,r}\left[J_{T^r\xi}(Q_{r+1,n}f_{\xi,n}\otimes Q_{r+1,n}f_{\xi,n})\right] \le F_{\xi,r}(J_{T^r\xi}(g\otimes g)) \le \beta_r \qquad a.s.$$

By the dominated convergence theorem, $\sum_{r=0}^K \mathbb{E}_{\xi} A_{n,r} \to 0$ a.s. as $n \to \infty$. The proof is completed.

In particular, applying Proposition 6.2 with $f_{\xi,n} = f_{\xi} - \pi(f_{\xi})$, we obtain the following result.

Proposition 6.3. Let ν_{ξ} be the distribution of $X(\emptyset)$. We assume that for almost all ξ , there exist a function g, an integer $n_0 = n_0(\xi)$ and non negative numbers $(\alpha_n, \beta_n) = (\alpha_n(\xi), \beta_n(\xi))$ such that

- (H1) for all $n \geq n_0$ and $x \in \mathcal{X}$, $\sup_{0 \leq r < n} Q_r \dots Q_{n-1} | f_{\xi} | (x) \leq g(x)$; (H2) for every $n \geq n_0$, $F_{\xi,n}(f_{\xi}^2) \leq \alpha_n$, and $\alpha_n/P_n \to 0$ as $n \to \infty$;
- (H3) for every $n \in \mathbb{N}$, $\max\{F_{\xi,n}(J_{T^n\xi}(g \otimes g)), F_{\xi,n}(J_{T^n\xi}(g \otimes \mathbf{1})), F_{\xi,n}(J_{T^n\xi}(\mathbf{1} \otimes \mathbf{1}))\} \leq \beta_n$ and $\sum_n \frac{\beta_n}{P_n m_n^2} < \infty$;

(H4) there exists $\pi(f_{\xi}) \in \mathbb{R}$ bounded by some constant M such that for each r fixed,

$$\lim_{n \to \infty} Q_r \dots Q_{n-1} f_{\xi}(x) = \pi(f_{\xi}) \quad \text{ for every } x \in \mathcal{X}.$$

Under (H1-(H4), if additionally $W_n \to W$ in \mathbb{P}_{ξ} -L², then we have for almost all ξ ,

$$\frac{\sum_{u \in \mathbb{T}_n} f_{\xi}(X(u))}{P_n} \to \pi(f_{\xi})W \qquad in \ \mathbb{P}_{\xi}\text{-}L^2.$$

Proof of Theorem 4.2. We apply Proposition 6.2 with $f_{\xi,n}(x) = \mathbf{1}_A(x) - \mu_{\xi,n}(A)$. Indeed, for n large enough, $f_{\xi,n}$ is bounded by some constant M. We can let g = M, $\alpha_n = M^2$ and $\beta_n = \mathbb{E}_{T^n \xi} N^2$. The assumptions (4.1) ensure that a.s.,

$$\frac{\alpha_n}{P_n} \to 0, \quad \sum_n \frac{\beta_n}{P_n m_n^2} < \infty \quad \text{and} \quad W_n \to W \quad \text{in \mathbb{P}_ξ-L}^2.$$

Then Proposition 6.2 yields for almost all ξ ,

$$\frac{Z_n(A)}{P_n} - \mu_{\xi,n}(A)W \to 0 \quad \text{in } \mathbb{P}_{\xi}\text{-probability.}$$
(6.4)

Moreover, $\mathbb{E}_{\xi}W = 1$ implies that $q(\xi) = \mathbb{P}_{\xi}(W = 0) < 1$, so $\{W > 0\} = \{N_n \to \infty\}$ \mathbb{P}_{ξ} - a.s. Finally, (4.3) comes from (6.4) and the fact that $W_n \to W > 0$ a.s. on the non-extinction event $\{N_n \to \infty\}$.

Proof of Corollary 4.3. We apply Proposition 6.3 with $f(x) = \mathbf{1}_A(x)$. Take g = 1, $\alpha_n = 1$ and $\beta_n = \mathbb{E}_{T^n \xi} N^2$.

6.2. Proof of the law of large numbers on the whole tree

Proposition 6.4 below is in the same vein as Proposition 6.3. We recall the notation $F_{\xi,n} = \nu_{\xi} Q_0 \dots Q_{n-1}$.

Proposition 6.4. Let ν_{ξ} be the distribution of $X(\emptyset)$. We assume that for almost all ξ , there exist a function g and non-negative numbers $(\alpha_n, \beta_n) = (\alpha_n(\xi), \beta_n(\xi))$ such that

- (H1) for all $n \in \mathbb{N}$ and $x \in \mathcal{X}$, $\sup_{0 \le r \le n} Q_r \dots Q_{n-1} |f_{\xi}|(x) \le g(x)$;
- (H2) for every $n \in \mathbb{N}$, $\max\{F_{\xi,n}(f_{\xi}^2), F_{\xi,n}(|f_{\xi}|g), F_{\xi,n}(g)\} \leq \alpha_n$ and $\sum_n \frac{\alpha_n}{P_n} < \infty$;
- (H3) for every $n \in \mathbb{N}$, $\max\{F_{\xi,n}(J_{T^n\xi}(g \otimes g)), F_{\xi,n}(J_{T^n\xi}(g \otimes \mathbf{1})), F_{\xi,n}(J_{T^n\xi}(\mathbf{1} \otimes \mathbf{1}))\} \leq \beta_n$ and $\sum_n \frac{\beta_n}{P_n m_n^2} < \infty$;
- (H4) there exists $\pi(f_{\xi}) \in \mathbb{R}$ bounded by some constant M such that for each r fixed,

$$\lim_{n \to \infty} \frac{1}{n-r} \sum_{k=r+1}^{n} Q_r \dots Q_{k-1} f_{\xi}(x) = \pi(f_{\xi}) \quad \text{for every } x \in \mathcal{X}.$$

Under (H1)-(H4), if additionally $\frac{1}{n}\sum_{k=1}^{n}W_{k}\to W$ in \mathbb{P}_{ξ} - L^{2} , then we have for almost all ξ ,

$$\frac{1}{n} \sum_{u \in \Gamma} \frac{f_{\xi}(X(u))}{P_{|u|}} \to \pi(f_{\xi})W \qquad \text{in } \mathbb{P}_{\xi}\text{-}L^2,$$

where $\Gamma_n = \bigcup_{k=1}^n \mathbb{T}_k$.

Proof. We only prove the case where $\pi(f_{\xi}) = 0$. For general case, it suffices to consider $f_{\xi} - \pi(f_{\xi})$ in place of f_{ξ} . We shall prove that under the hypothesis (H1)–(H4), a.s.,

$$\mathbb{E}_{\xi} \left(\frac{1}{n} \sum_{u \in \Gamma_n} \frac{f_{\xi}(X(u))}{P_{|u|}} \right)^2 \to 0 \text{ as } n \to \infty.$$
 (6.5)

Notice that $\Gamma_n = \bigcup_{k=1}^n \mathbb{T}_k$ is the set of all individuals in the first n generation. For $u, v \in \Gamma_n$, we discuss for two cases: (i) u and v in the same life line, which means that one is an ancestor of the other, i.e. $u \wedge v = u$ or v; (ii) the contrary case, i.e. $u \wedge v \neq u$, v. So we can write

$$\mathbb{E}_{\xi} \left(\frac{1}{n} \sum_{u \in \Gamma_n} \frac{f_{\xi}(X(u))}{P_{|u|}} \right)^2 = \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{u,v \in \Gamma_n} \frac{f_{\xi}(X(u)) f_{\xi}(X(v))}{P_{|u|} P_{|v|}} \right] = S_{n,\xi} + T_{n,\xi},$$

where

$$S_{n,\xi} = \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v = u \text{ or } v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \quad \text{and} \quad T_{n,\xi} = \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{|u|}P_{|v|}} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \wedge v \neq u,v}} \frac{f_{\xi}(X(u))f_{\xi}(X(v))}{P_{\xi}(X(u))} \right] \cdot \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v$$

We need to prove that $S_{n,\xi}$, $T_{n,\xi}$ tend to 0 a.s. as $n \to \infty$. At first, $S_{n,\xi}$ can be decomposed as

$$S_{n,\xi} = \frac{2}{n^2} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_n \\ u \le v}} \frac{f_{\xi}(X(u)) f_{\xi}(X(v))}{P_{|u|} P_{|v|}} \right] + \frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{u \in \Gamma_n} \frac{f_{\xi}^2(X(u))}{P_{|u|}^2} \right].$$
 (6.6)

For the second term in the right side of the equality above, we use (H2) and get

$$\frac{1}{n^2} \mathbb{E}_{\xi} \left[\sum_{u \in \Gamma_n} \frac{f_{\xi}^2(X(u))}{P_{|u|}^2} \right] = \frac{1}{n^2} \sum_{r=1}^n \mathbb{E}_{\xi} \left[\sum_{u \in \mathbb{T}_r} \frac{f_{\xi}^2(X(u))}{P_{|r|}^2} \right] = \frac{1}{n^2} \sum_{r=1}^n P_r^{-1} F_{\xi,r}(f_{\xi}^2) \le \frac{1}{n^2} \sum_{r=1}^{\infty} P_r^{-1} \alpha_r \to 0$$

as $n \to \infty$, since $\sum_n P_n^{-1} \alpha_n < \infty$ a.s.. For $u, v \in \Gamma_n$ with u < v, we write $v = u\tilde{v}$ with $|\tilde{v}| = k - r$ if $u \in \mathbb{T}_r$ and $v \in \mathbb{T}_k$ (k > r). Then, using (H2) and the a.s. convergence of $\sum_n P_n^{-1} \alpha_n$ again, we have a.s.,

$$\frac{2}{n^{2}} \mathbb{E}_{\xi} \left[\sum_{\substack{u,v \in \Gamma_{n} \\ u < v}} \frac{f_{\xi}(X(u)) f_{\xi}(X(v))}{P_{|u|} P_{|v|}} \right] = \frac{2}{n^{2}} \sum_{r=1}^{n} \sum_{k=r+1}^{n} \mathbb{E}_{\xi} \left[\sum_{u \in \mathbb{T}_{r}} \sum_{u\tilde{v} \in \mathbb{T}_{k-r}(u)} \frac{f_{\xi}(X(u)) f_{\xi}(X(u\tilde{v}))}{P_{|r|} P_{|k|}} \right] \\
= \frac{2}{n^{2}} \sum_{r=1}^{n} \sum_{k=r+1}^{n} P_{r}^{-2} \mathbb{E}_{\xi} \left[\sum_{u \in \mathbb{T}_{r}} f_{\xi}(X(u)) Q_{r} \dots Q_{k-1} f_{\xi}(X(u)) \right] \\
\leq \frac{2(n-r)}{n^{2}} \sum_{r=1}^{n} P_{r}^{-1} F_{\xi,r}(|f_{\xi}|g) \\
\leq \frac{2}{n} \sum_{r=1}^{\infty} P_{r}^{-1} \alpha_{r} \to 0 \text{ as } n \to \infty.$$

Hence we have $S_{n,\xi} \to 0$ a.s. as $n \to \infty$. Now we consider $T_{n,\xi}$. For $K = K(\xi)$ fixed suitable large,

$$T_{n,\xi} = \sum_{r=1}^{K} \mathbb{E}_{\xi} A_{n,r} + \sum_{r=K+1}^{n-1} \mathbb{E}_{\xi} A_{n,r},$$

where

$$A_{n,r} = \frac{1}{n^2} \sum_{k,l=r+1}^n \sum_{w \in \mathbb{T}_r} \sum_{\substack{w \in \mathbb{T}_r \text{ } wi, wj \in \mathbb{T}_1(w) \text{ } wi\tilde{u} \in \mathbb{T}_{k-r-1}(wi) \\ i \neq j}} \sum_{\substack{wi\tilde{u} \in \mathbb{T}_{k-r-1}(wi) \\ wj\tilde{v} \in \mathbb{T}_{k-r-1}(wj)}} \frac{f_{\xi}(X(wi\tilde{u}))f_{\xi}(X(wi\tilde{v}))}{P_k P_l}.$$

With the notation $Q_{r,n} = Q_r \dots Q_{n-1}$, we compute

$$\mathbb{E}_{\xi} A_{n,r} = \frac{1}{n^2} \sum_{k,l=r+1}^{n} (P_k P_l)^{-1} \mathbb{E}_{\xi} \left[\sum_{w \in \mathbb{T}_r} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} \mathbb{E}_{\xi,X(wi)} \left[\sum_{wi\tilde{u} \in \mathbb{T}_{k-r-1}(wi)} f_{\xi}(X(wi\tilde{u})) \right] \right] \\
= \frac{1}{n^2} \sum_{k,l=r+1}^{n} P_{r+1}^{-2} \mathbb{E}_{\xi} \left[\sum_{\substack{w \in \mathbb{T}_r \\ i \neq j}} \sum_{\substack{wi, wj \in \mathbb{T}_1(w) \\ i \neq j}} Q_{r+1,k} f_{\xi}(X(wi)) Q_{r+1,l} f_{\xi}(X(wj)) \right] \\
= \frac{(n-r)^2}{n^2} \frac{1}{P_r m_r^2} F_{\xi,r} (J_{T^{r+1}\xi}(R_{n,r} \otimes R_{n,r})),$$

where

$$R_{n,r}(x) = \frac{1}{n-r} \sum_{k=r+1}^{n} Q_{r+1,k} f_{\xi}(x).$$

We oberve that thanks to (H1), $\sup_{0 \le r < n} |R_{r,n}| \le g$ for every n. And by (H4), for almost all ξ , for each r fixed,

$$R_{n,r} \otimes R_{n,r}(y,z) \stackrel{n \to \infty}{\longrightarrow} 0$$
 for each $(y,z) \in \mathcal{X}^2$.

Following similar arguments in the proof of Proposition 6.2, we have $T_{n,\xi} \to 0$ a.s. as $n \to \infty$. The proof is completed.

Lemma 6.5 ([21], Cor. 1.6). Let p > 1. If $\mathbb{E}(\log m_0) > 0$ and $\mathbb{E}\left(\log \mathbb{E}_{\xi}\left(\frac{N}{m_0}\right)^p\right) < \infty$, then

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} W_k = W \qquad in \ \mathbb{P}_{\xi}\text{-}L^p.$$

Proof of Theorem 4.6. By Lemma 6.5, $\frac{1}{n} \sum_{k=1}^{n} W_k \to W$ in \mathbb{P}_{ξ} - L^2 . Applying Proposition 6.4 with $f(x) = \mathbf{1}_A(x)$, g = 1, $\alpha_n = 1$ and $\beta_n = \mathbb{E}_{T^n \xi} N^2$, we obtain for almost all ξ ,

$$\frac{1}{n} \sum_{k=1}^{n} \frac{Z_k(A)}{P_k} \to \mu(A)W \qquad \text{in } \mathbb{P}_{\xi}\text{-probability.}$$
 (6.7)

Since $W_n \to W$ a.s., for any $\delta > 0$, there exists n_0 such that $\forall n \geq n_0$,

$$|W_n - W| < \delta$$
 and $\left| \frac{1}{W_n} - \frac{1}{W} \right| < \delta$ (6.8)

a.s. on the non-extinction event. We write

$$\frac{1}{n} \sum_{k=1}^{n} \frac{Z_k(A)}{N_k} - \mu(A) = \frac{1}{n} \sum_{k=1}^{n_0} \frac{Z_k(A)}{P_k} \left(\frac{1}{W_k} - \frac{1}{W} \right) + \frac{1}{n} \sum_{k=n_0+1}^{n} \frac{Z_k(A)}{P_k} \left(\frac{1}{W_k} - \frac{1}{W} \right) + \frac{1}{W} \left(\frac{1}{N} \sum_{k=1}^{n} \frac{Z_k(A)}{P_k} - \mu(A)W \right).$$
(6.9)

Obviously, the first term in the right hand side of (6.9) tends to 0 a.s. on the non-extinction event as n goes to infinity. And the convergence of the third term is from (6.7). Therefore, to prove (4.7), we only need to show that the second term

$$\frac{1}{n} \sum_{k=n_0+1}^{n} \frac{Z_k(A)}{P_k} \left(\frac{1}{W_k} - \frac{1}{W} \right) \stackrel{n \to \infty}{\longrightarrow} 0 \quad \text{in } \mathbb{P}_{\xi}\text{-probability}$$
 (6.10)

on the non-extinction event for almost all ξ . In fact, by (6.8), we have

$$\left| \frac{1}{n} \sum_{k=n_0+1}^{n} \frac{Z_k(A)}{P_k} \left(\frac{1}{W_k} - \frac{1}{W} \right) \right| < \delta(W + \delta)$$

a.s. on the non-extinction event. The arbitrariness of δ yields (6.10).

6.3. Proof of the central limit theorem

Proof of Theorem 4.8. We shall apply Proposition 6.2 with $f_{\xi,n}(x) = \mathbf{1}_{A_n}(x) - \Phi(y)$, where $A_n = (-\infty, b_n(\xi)y + a_n(\xi)]$. By (4.8) and Dini's Theorem, we have a.s.,

$$\lim_{n \to \infty} \sup_{y \in \mathbb{R}} \left| \mathbb{P}_{\xi, x} \left(\frac{Y_n - a_n(\xi)}{b_n(\xi)} \le y \right) - \Phi(y) \right| = 0 \quad \text{for every } x \in \mathcal{X}.$$
 (6.11)

Notice that $|f_{\xi,n}| \leq 1$. Take g = 1, $\alpha_n = 1$ and $\beta_n = \mathbb{E}_{T^n \xi} N^2$. It is easy to verify that (H1)–(H3) are satisfied. For (H4), by (6.11) and the continuity of Φ , for each r fixed,

$$Q_{r} \dots Q_{n-1} f_{\xi,n}(x) = \mathbb{P}_{T^{r}\xi,x} (Y_{n-r} \leq b_{n}(\xi)y + a_{n}(\xi)) - \Phi(y)$$

$$\leq \sup_{y \in \mathbb{R}} \left| \mathbb{P}_{T^{r}\xi,x} \left(\frac{Y_{n-r} - a_{n-r}(T^{r}\xi)}{b_{n-r}(T^{r}\xi)} \leq y \right) - \Phi(y) \right| + \left| \Phi\left(\frac{b_{n}(\xi)y + a_{n}(\xi) - a_{n-r}(T^{r}\xi)}{b_{n-r}(T^{r}\xi)} \right) - \Phi(y) \right|,$$

which goes to 0 as $n \to \infty$ for every $x \in \mathcal{X}$. By Proposition 6.2, a.s.,

$$\frac{Z_n(A_n)}{P_n} - \Phi(y) \frac{N_n}{P_n} \to 0 \quad \text{in } \mathbb{P}_{\xi}\text{-}L^2.$$

Since $W_n \to W$ a.s. and in \mathbb{P}_{ξ} - L^2 , we have for almost all ξ ,

$$P_n^{-1}Z_n(-\infty, b_n(\xi)y + a_n(\xi)] \to \Phi(y)W$$
 in \mathbb{P}_{ξ} -probability,

which implies (4.10).

Acknowledgements. The authors would like to thank anonymous referees for valuable comments and remarks. This work was partially funded by Chair Modélisation Mathématique et Biodiversité VEOLIA-École 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), the National Natural Science Foundation of China, Grant No. 11101039 and the Natural Scientific Research Innovation Foundation in Harbin Institute of Technology (HIT. NSRIF. 2015102).

References

- [1] S. Alili, Asymptotic behaviour for random walk in random environments. J. Appl. Probab. 36 (1999) 334–349.
- [2] K.B. Athreya and S. Karlin, On branching processes with random environments, I: extinction probability. *Ann. Math. Stat.* **42** (1971) 1499–1520.
- [3] K.B. Athreya and S. Karlin, On branching processes with random environments, II: limit theorems. Ann. Math. Stat. 42 (1971) 1843–1858.
- [4] K.B. Athreya and H.J. Kang, Some limit theorems for positive recurrent branching Markov chains I, II. Adv. Appl. Probab. 30 (1998) 693–722.
- [5] V. Bansaye, Proliferating parasites in dividing cells: Kimmel's branching model revisited. Ann. Appl. Probab. 18 (2008) 967–996.
- [6] V. Bansaye and V.C. Tran, Branching feller diffusion for cell division with parasite infection. ALEA 8 (2011) 95–127.
- [7] V. Bansaye, J-F. Delmas, L. Marsalle and V.C. Tran, Limit theorems for Markov processes indexed by supercritical Galton Watson tree. Ann. Appl. Probab. 21 (2011) 2263–2314.
- [8] B. Bercu, B. De Saporta and A. Gégout-Petit, Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electr. J. Probab.* **14** (2009) 2492–2526.
- [9] J.D. Biggins, Martingale convergence in the branching random walk. J. Appl. Probab. 14 (1977) 25–37.
- [10] J.D. Biggins, Chernoff's theorem in the branching random walk. J. Appl. Probab. 14 (1977) 630–636.
- [11] J.D. Biggins, The central limit theorem for the supercritical branching random walk, and related results. Stoch. Proc. Appl. 34 (1990) 255–274.
- [12] F. Comets and S. Popov, Shape and local growth for multidimensional branching random walks in random environment. *ALEA* 3 (2007) 273–299.
- [13] F. Comets and N. Yoshida, Branching random walks in space-time random environment: survival probability, global and local growth rates. J. Theor. Probab. 24 (2011) 657–687.
- [14] J-F. Delmas and L. Marsalle, Detection of cellular aging in a Galton-Watson process. Stoch. Proc. Appl. 120 (2010) 2495–2519.
- [15] M. Doumic, M. Hoffmann, N. Krell and L. Robert, Statistical estimation of a growth-fragmentation model observed on a genealogical tree. Bernoulli 21 (2015) 1760–1799.
- [16] H. Fréville, B. Colas, M. Riba, H. Caswell, A. Mignot, E. Imbert and I. Olivieri, Spatial and demographic variability in the endemic plant species Centaurea Corymbosa (Asteraceae). Ecology 85 (2004) 694–703.
- [17] L. Gallardo, A central limit theorem for Markov chains and applications to hypergroups. Proc. Am. Math. Soc. 127 (1999) 1837–1845.
- [18] N. Gantert, S. Müller, S. Popov and M. Vachkovskaia, Survival of branching random walks in random environment. J. Theor. Probab. 23 (2010) 1002–1014.
- [19] J. Guyon, Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. Ann. Appl. Probab. 17 (2007) 1538–1569.
- [20] S.C. Harris and M. Roberts, The many-to-few lemma and multiple spines. Preprint available via http://arxiv.org/abs/1106. 4761 (2014).
- [21] C. Huang, Q. Liu, Convergence in L^p and its exponential rate for a branching process in a random environment. *Electron. J. Probab.* 19 (2014) 1–22.
- [22] C. Huang and Q. Liu, Branching random walk with a random environment in time. Preprint available via http://arxiv.org/abs/1407.7623 (2014).
- [23] N. Lartillot and R. Poujol, A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. Mol. Biol. Evol. 28 (2010) 729-744.
- [24] Q. Liu, Branching random walks in random environment, in Proc. of the 4th International Congress of Chinese Mathematicians. Edited by L. Ji, K. Liu, L. Yang and S.-T. Yau, Vol. II (2007) 702-719.

- [25] S.P. Meyn and R.L. Tweedie, Stability of Markovian processes I: criteria for discrete-time chains. Adv. Appl. Probab. 24 (1992) 542–574.
- [26] F. Mukhamedov, On L₁-weak ergodicity of nonhomogeneous discrete Markov processes and its applications. Available via http://arxiv.org/abs/1105.0478 (2012).
- [27] S. Müller, A criterion for transience of multidimensional branching random walk in random environment. Electr. J. Probab. 13 (2008) 1189–1202.
- [28] M. Nakashima, Almost sure central limit theorem for branching random walks in random environment. Ann. Appl. Probab. 21 (2011) 351–373.
- [29] S. Orey, Markov chains with stationary Markov transition probabilities. Ann. Probab. 19 (1991) 907–928.
- [30] H. Pan, Law of large number for even-odd Markov chain fields and a three times circulation Markov chain indexed by a tree. Ph.D. thesis (2010). http://globethesis.com/?t=2120360302966543.
- [31] T. Seppäläinen, Large deviations for Markov chains with random transitions. Ann. Probab. 22 (1994) 713–748.
- [32] D.-A. Sinclair and L. Guarente, Extrachromosomal rDNA circles—a cause of aging in yeast. Cell 91 (1997) 1033-1042.
- [33] T. Stadler, Inferring speciation and extinction processes from extant species data. Proc. Natl. Acad. Sci. 108 (2011) 16145– 16146.
- [34] E.J. Stewart, R. Madden, G. Paul and F. Taddei, Aging and death in a organism that reproduces by morphologically symmetric division. PLoS Biol. 3 (2005) e45.
- [35] N. Yoshida, Central limit theorem for random walk in random environment. Ann. Appl. Probab. 18 (2008) 1619–1635.