2011

Coalescence in Bellman-Harris and multi-type branching processes

Jyy-i Joy Hong
Iowa State University

Coalescence in Bellman-Harris and multi-type branching processes

by

Jyy-I Hong

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Mathematics

Program of Study Committee:

Krishna B. Athreya, Major Professor

Clifford Bergman

Dan Nordman

Ananda Weerasinghe

Paul E. Sacks

Iowa State University

Ames, Iowa

2011

DEDICATION

I would like to dedicate this thesis to my parents Wan-Fu Hong and Wen-Hsiang Tseng for their unconditional love and support. Without them, the completion of this work would not have been possible.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First of all, I would like to take this opportunity to express my appreciation to my advisor Dr. Krishna B. Athreya for his excellent teaching and expert guidance. From the knowledge of mathematics to the advice of study, research and even life, I always feel so comfortable and encouraged to discuss with him and learn from him. Secondly, I would like to thank the members of my committee: Dr. Ananda Weerasinghe, Dr. Paul E. Sacks, Dr. Clifford Bergman and Dr. Dan Nordman for their precious advices and suggestions. Also, my thanks are given to Dr. Jerold Mathews for the wonderful reading time he spent in improving my English and to Dr. Leslie Hogben for her helps in many ways.

Furthermore, I would like to give my thanks to Dr. Jhishen Tsay (National Sun Yat-sen University, Taiwan) and his wife Chun-Tor Lee for their continuous encouragement when I was struggling and depressed. In addition, I would like to thank all the friends from LIFE at Campus Baptist Church for everything they have done for me.

Finally, I would like to give my deepest appreciation to my mother, my sister, Mou-Mou, My brother, Shih-Hsun, and my brother-in-low, Cohuahua, for their love, care, understanding and patience. With their endless support in my life, I was able to pursue my dreams. I also give my special thanks to my husband, Chao-Chun, for his love and company during these years and it means a lot to me.

ABSTRACT

For branching processes, there are many well-known limit theorems regarding the evolution of the population in the future time. In this dissertation, we investigate the other direction of the evolution, that is, the past of the processes. We pick some individuals at random by simple random sampling without replacement and trace their lines of descent backward in time until they meet. We study the coalescence problem of the discrete-time multi-type Galton-Watson branching process and both the continuous-time single-type and multi-type Bellman-Harris branching processes including the generation number, the death time (in the continuous-time processes) and the type (in the multi-type processes) of the last common ancestor (also called the most recent common ancestor) of the randomly chosen individuals for the different cases (supercritical, critical, subcritical and explosive).

CHAPTER 1. PRELIMINARIES

1.1 Introduction

The study of branching processes has a long history and was essentially motivated by the observation of the extinction of certain family lines of the European aristocracy in contrast to the rapid exponential growth of the whole population. Francis Galton formulated this extinction problem and originally posed it in the Educational Times in 1874 and the Reverend Henry William Watson replied with a solution (see Harris (1963)). Seneta and Heyde (1977) have pointed out that the French mathematician Bienaymé had formulated essentially the same model fifty years earlier.

The model of Galton and Watson (called the Galton-Watson branching process) appeared to have been neglected for many years after its creation. After 1940, interest in this model increased, partly because of the analogy between the growth of families and nuclear chain reactions and also partly because of the increased general interest in applications of probability theory. Since then, branching processes have been regarded as appropriate probability models for the description of the behavior of systems whose components (cells, particles, individuals in general) reproduce, are transformed, or die (see Harris [20], Athreya and Ney [5], Jagers [22], Mode [28] and Sevastyanov [35]). Nowadays, this theory is an area of active and interesting research.

There are many generalizations of the single-type Galton-Watson branching process in discrete time. Of these the multi-type branching process model in discrete time is a natural one. The multi-type branching process is important because it is constructed in a way that closely matches real-life situations and hence can be used to study a wide variety of real-life problems, including those related to differences in types of ethnicities, types of genes, types of cosmic rays, etc. Another generalization is the continuous-time single-type case known as the Bellman-Harris branching process which is widely used by many fields. This device was suggested by Scott and Uhlenbeck (1942) in their treatment of cosmic rays, where the continuous variable is energy, and was used by Bartlett (1946) and Leslie (1948) in dealing with human population, where the continuous variable is age.

In the rest of this chapter, we review basic definitions and results of single-type (Section 1.2) and multi-type (Section 1.3) discrete-time Galton-Watson branching processes and single-type (Section 1.4) and multi-type (Section 1.5) continuous-time Bellman-Harris age-dependent branching processes. We discuss results on the extinction probabilities, the growth rates of population and some other convergent properties. The results are fundamental and may be found in the books on branching processes mentioned earlier. Here, we state the results which are needed in this thesis based on the books, *Branching Processes* written by Athreya and Ney [5] and *The Theory of Branching Processes* written by T. E. Harris [20].

In Chapter 2, we state the problem of coalescence in branching processes and review the results for all cases (supercritical, critical ,subcritical and explosive) of the discrete-time single-type Galton-Watson branching processes.

In Chapter 3, we extend the results of the problem of coalescence to the discrete-time multi-type Galton-Watson branching process including supercritical (Section 3.2), critical (Section 3.3) and subcritical (Section 3.4) cases. Also, we present the Markov property on Types (Section 3.5) along the line of descent of an individual randomly chosen from the current generation by simple random sampling.

In Chapter 4, we consider the continuous-time single-type Bellman-Harris branching processes and give proofs to the problem of coalescence in the supercritical (Section 4.2) and subcritical (Section 4.3) cases. By the results on the problem of coalescence, we also are able to investigate the branching random walks (Section 4.4).

Although the research of branching processes has a long history, the study of the problem of coalescence is still in its infancy. In Chapter 5, we state some interesting open questions related to this topic.

1.2 Discrete-time Single-type Galton-Watson Branching Processes

1.2.1 Definitions and Notations

A discrete-time single-type Galton-Watson branching process is the simplest type of branching process. This process can be thought as a population evolving in time. It starts at time 0 with Z_0 individuals, each of which lives a unit of time and produces its offsprings upon death according to the probability distribution $\{p_j\}_{j\geq 0}$ independently of others. Let Z_1 be the total number of children produced by the Z_0 individuals, that is,

$$\sum_{i=1}^{Z_0} \xi_{0,i}$$

where $\{\xi_{0,i}\}_{i\geq 1}$ are i.i.d. random variables with the probability distribution $\{p_j\}_{j\geq 0}$. It constitutes the first generation and then these individuals in the first generation go on to produce the second generation of population Z_2 and so on. So, the total size of the population in the $(n+1)$st generation, $n = 0, 1, 2, \cdots$, is given by

$$Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} \xi_{n,i} & \text{if } Z_n > 0 \\ 0 & \text{if } Z_n = 0 \end{cases}$$

where $\{\xi_{n,i} : i \geq 1, n \geq 0\}$ are i.i.d. copies with the probability distribution $\{p_j\}_{j\geq 0}$.

Then $\{Z_n\}_{n\geq 0}$ is called a Galton-Watson branching process with initial population Z_0 and offspring distribution $\{p_j\}_{j\geq 0}$. Here, $\xi_{n,i}$ is the number of offspring of the ith individual of the nth generation.

Let

$$m \equiv \sum_{j=1}^{\infty} j p_j$$

be the mean of the offspring distribution $\{p_j\}_{j\geq 1}$. We shall refer to the Galton-Watson process as subcritical, critical, supercritical or explosive according as $0 < m < 1$, $m = 1$, $1 < m < \infty$ or $m = \infty$, respectively.

Moreover, if $\mathcal{T}$ denotes the full family tree generated in this way, every individual in $\mathcal{T}$ can be identified by a finite string $(i_0, i_1, \cdots, i_n)$ meaning that this individual is in the nth generation and is the i_nth child of the individual $(i_0, i_1, \cdots, i_{n-1})$ of the $(n-1)$st generation.

1.2.2 Limit Theorems

In this section, we collect some well-known results for discrete-time single-type Galton-Watson branching processes.

Theorem 1.1. *(Supercritical Case) Let $p_0 = 0$ and $1 < m < \infty$. Then*

(a) $P(Z_n \to \infty | Z_0 > 0) = 1$.

(b) (Harris, 1960) $\left\{ W_n \equiv \frac{Z_n}{m^n} : n \geq 0 \right\}$ *is a nonnegative martingale and hence*

$$\lim_{n\to\infty} W_n \equiv W \text{ exists w.p.1.}$$

(c) (Kesten and Stigum, 1966)

$$\sum_{j=1}^{\infty} (j \log j) p_j < \infty \quad \text{if and only if} \quad E(W|Z_0 = 1) = 1$$

and then W has an absolutely continuous distribution on $(0, \infty)$ *with a positive density.*

(d) (Seneta and Heyde, 1970)

$$\exists C_n \quad s.t. \quad \frac{C_{n+1}}{C_n} \to m \quad \text{and} \quad \frac{Z_n}{C_n} \to W \text{ w.p.1}$$

and $P(0 < W < \infty) = 1$. *In particular,* $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$ *if and only if* $C_n \sim m^n$.

(e) (Athreya and Schuh [4])

$$E(W : W \leq x) \equiv L(x)$$

is slowly varying at ∞, *i.e.* $\forall 0 < c < \infty$, $\frac{L(cx)}{L(x)} \to 1$ *as* $x \to \infty$.

Under the assumption $p_0 = 0$, the population size Z_n of a supercritical Galton-Watson branching process goes to infinity as $n \to \infty$ with probability 1 and it grows like m^n. This is the stochastic analogue of the so-called Malthusian law of geometric population growth.

In the next two theorems, we present the results for the critical and subcritical cases.

Theorem 1.2. *(Critical Case) Let* $m = 1$, $p_j \neq 1$ *for any* $j \geq 1$ *and* $\sigma^2 \equiv \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Then*

(a) $P(Z_n \to 0 | 0 < Z_0 < \infty) = 1$.

(b) (Kolmogrov, 1938)

$$nP(Z_n > 0) \to \frac{\sigma^2}{2} \quad \text{as } n \to \infty.$$

(c) (Yaglom, 1947)

$$P\left(\frac{Z_n}{n} > x \middle| Z_n > 0\right) \to e^{-\frac{2}{\sigma^2}x} \quad , 0 < x < \infty.$$

(d) (Athreya [12]) For $1 \le k < n$, let

$$V_{n,k} \equiv \left\{ \frac{Z^{(k)}_{n-k,i}}{n-k} I_{(Z^{(k)}_{n-k,i} > 0)} : 1 \le i \le Z_k \right\}$$

on the event $\{Z_k > 0\}$, where $\{Z^{(k)}_{j,i} : j \ge 0\}$ is the G-W process initiated by the ith individual in the kth generation.

Let $k \to \infty$, $n \to \infty$ such that $\frac{k}{n} \to u$, $0 < u < 1$.

Then the sequence of point processes $\{V_{n,k}\}_{n\ge1}$ conditioned on $\{Z_n \ge 1\}$ converges weakly to the point process

$$V \equiv \{\eta_j : j = 1, 2, \cdots, N_u\}$$

where $\{\eta_j\}_{j\ge1}$ are i.i.d. $\exp(1)$, N_u is Geo(u), i.e., $P(N_u = k) = (1-u)u^{k-1}$, $k \ge 1$ and $\{\eta_j\}_{j\ge1}$ and N_u are independent.

Theorem 1.3. *(Subcritical Case) (Yaglom, 1947) Let $0 < m \equiv \sum_{j=1}^{\infty} jp_j < 1$. Then*

(a) *For $j \ge 1$, $\lim_{n\to\infty} P(Z_n = j | Z_n > 0) \equiv b_j$ exists, $\sum_{j=0}^{\infty} b_j = 1$ and $B(s) \equiv \sum_{j=0}^{\infty} b_j s^j$, $0 \le s \le 1$ is the unique solution of the functional equation*

$$B(f(s)) = mB(s) + (1-s) \quad , 0 \le s \le 1$$

where $f(s) \equiv \sum_{j=0}^{\infty} p_j s^j$, in the class of all probability generating functions vanishing at 0.

(b) $\sum_{j=1}^{\infty} jb_j < \infty$ *iff* $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$.

(c) $\lim_{n\to\infty} \frac{P(Z_n > 0|Z_0 = 1)}{m^n} = \frac{1}{\sum_{j=1}^{\infty} jb_j}$.

(d) *If* Z_0 *is a random variable and* $EZ_0 < \infty$*, then*

$$\lim_{n\to\infty} P(Z_n = j|Z_n > 0) = b_j \quad , \forall j \geq 1$$

and if, in addition, $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$ *then*

$$\sum_{j=1}^{\infty} jb_j < \infty \quad \text{and} \quad \lim_{n\to\infty} \frac{P(Z_n > 0)}{m^n} = \frac{EZ_0}{\sum_{j=1}^{\infty} jb_j}.$$

In both of the critical and subcritical Galton-Watson branching processes, the population will die out eventually with probability 1. But, conditioned on the event of non-distinction, i.e. the set $\{Z_n > 0\}$, Z_n will go to infinity in distribution with the growth rate of n in the critical case while Z_n will converge to a proper random variable in distribution in the subcritical case as $n \to \infty$.

We present the results of P. L. Davies and D. R. Grey for the explosive Galton-Watson branching process as follows.

Theorem 1.4. *(Explosive Case) Let* $p_0 = 0$*,* $m = \infty$ *and, for some* $0 < \alpha < 1$*,*

$$\frac{\sum_{j>x} p_j}{x^\alpha L(x)} \to 1 \quad \text{as } x \to \infty$$

where $L : (1, \infty) \to (0, \infty)$ *is a function slowly varying at* ∞*. Then*

(a) *(Davies [16])*

$$\alpha^n \log Z_n \to \eta \quad \text{w.p.1}$$

and $P(0 < \eta < \infty) = 1$ *and* η *has a continuous distribution.*

(b) *(Grey [19]) Let* $\{Z_n^{(1)}\}_{n\geq 0}$ *and* $\{Z_n^{(2)}\}_{n\geq 0}$ *be two i.i.d. copies of a Galton-Watson branching process with the offspring distribution* $\{p_j\}_{j\geq 0}$ *satisfying* $p_0 = 0$, $m \equiv \sum_{j=1}^{\infty} jp_j = \infty$ *and* $Z_0^{(1)} = Z_0^{(2)} = 1$*. Then, w.p.1,*

$$\frac{Z_n^{(1)}}{Z_n^{(2)}} \to \begin{cases} 0, & \text{with prob. } \frac{1}{2} \\ \infty, & \text{with prob. } \frac{1}{2}. \end{cases}$$

It is easy to deduce b) from a) in the above theorem as

$$\alpha^n \Big(\log Z_n^{(1)} - \log Z_n^{(2)} \Big) \to \eta_1 - \eta_2 \equiv \eta, \text{ say}$$

and

$$P(\eta > 0) = P(\eta < 0) = \frac{1}{2}.$$

1.3 Discrete-time Multi-type Galton-Watson Branching Processes

1.3.1 Definitions, Assumptions and Notations

In a discrete-time single-type Galton-Watson branching process, we assume that each individual lives for a fixed unit time and then produces its children according to the same offspring distribution. In this section, we allow a number of distinguishable types of individuals having different offspring distributions.

First, we consider a finite number d of individual types. Such processes arise in a variety of applications in biology and physics and they could represent genetic or mutant types in the real populations such as animal population, bacterial population or photons, etc.

Through out this section and next chapter, we adopt the following conventions.

1. $\mathbb{N}_0$ is the set of all nonnegative integers.

2. $\mathbb{N}_0^d \equiv \Big\{ \mathbf{j} \equiv (j_1, j_2, \cdots, j_d) : j_i \in \mathbb{N}_0, i = 1, 2, \cdots, d \Big\}$

3. $\mathbf{0} = (0, 0, \cdots, 0)$ and $\mathbf{1} = (1, 1, \cdots, 1)$ in $\mathbb{N}_0^d$

4. $\mathbf{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0) \in \mathbb{N}_0^d$ with the 1 in the ith component.

5. $\mathbf{u} \leq \mathbf{v}$ means $u_i \leq v_i$ for $i = 1, 2, \cdots, d$ while $\mathbf{u} < \mathbf{v}$ means $u_i \leq v_i$ for all i and $u_i < v_i$ for at least one i.

6. The vector of absolute values is

$$|\mathbf{x}| = |x_1| + |x_2| + \cdots + |x_d|$$

7. The sup norm is

$$\|\mathbf{x}\| = \max\{|x_1|, |x_2|, \cdots, |x_d|\}$$

8. The product notation is

$$\mathbf{x}^{\mathbf{y}} = \prod_{i=1}^{d} x_i^{y_i}$$

9. For a matrix **M**, the super norm is

$$\|\mathbf{M}\| = \max\{|m_{ij}| : i, j = 1, 2, \cdots, d\}$$

Let $\mathbf{Z}_n = (Z_{n,1}, Z_{n,2}, \cdots, Z_{n,d})$ be the population vector in the nth generation, $n = 0, 1, 2, \cdots$, where $Z_{n,i}$ is the number of individuals of type i in the nth generation. We assume that each individual of type i, $i = 1, 2, \cdots, d$, lives a unit of time and, upon death, produces children of all types and according to the offspring distribution $\{p^{(i)}(\mathbf{j}) \equiv p^{(i)}(j_1, j_2, \cdots, j_d)\}_{\mathbf{j}\in\mathbb{N}^d}$ and independently of other individual, where $p^{(i)}(j_1, j_2, \cdots, j_d)$ is the probability that a type i parent produces j_1 children of type 1, j_2 children of type 2, $\cdots$, j_d children of type d. Therefore, each component of the vector of the probability generating functions $\mathbf{f} = (f^{(1)}, f^{(2)}, \cdots, f^{(d)})$ can be written as:

$$f^{(i)}(s_1, s_2, \cdots, s_d) = \sum_{j_1, j_2, \cdots, j_d \geq 0} p^{(i)}(j_1, j_2, \cdots, j_d) s_1^{j_1} s_2^{j_2} \cdots s_d^{j_d}$$

where $0 \leq s_r \leq 1$, $r = 1, 2, \cdots, d$, being the probability generating function of the number of various types produced by a type i individual,

Thus, a discrete-time multi-type Galton-Watson branching process $\{\mathbf{Z}_n\}_{n\geq 0}$ is a Markov chain on $\mathbb{N}_0^d$ with the transition function

$$P(\mathbf{i}, \mathbf{j}) = P(\mathbf{Z}_{n+1} = \mathbf{j} | \mathbf{Z}_n = \mathbf{i}) \quad \forall \mathbf{i}, \mathbf{j} \in \mathbb{N}_0^d$$

such that $\sum_{\mathbf{j}\in\mathbb{N}_0^d}^{\infty} P(\mathbf{i}, \mathbf{j}) \mathbf{s}^{\mathbf{j}} = \Big(\mathbf{f}(\mathbf{s})\Big)^{\mathbf{i}}$ (see notation (8)).

When the process is initiated in state $\mathbf{e}_i$, we will denote the process $\{\mathbf{Z}_n\}_{n\geq 0}$ by

$$\mathbf{Z}_n^{(i)} = (Z_{n,1}^{(i)}, Z_{n,2}^{(i)}, \cdots, Z_{n,d}^{(i)})$$

where $Z_{n,j}^{(i)}$ is the number of type j individuals in the nth generation for a process $\mathbf{Z}_0 = \mathbf{e}_i$. The generating function of $\mathbf{Z}_n^{(i)}$ will be denoted by $\mathbf{f}_n^{(i)}(\mathbf{s})$.

Also, we let $\xi_{n,r}^{(j)}$ be the vector of offsprings of the rth individual of type j in the nth generation then $\xi_{n,r}^{(j)} \sim \{p^{(j)}(\cdot)\}$, i.e., $P(\xi_{n,r}^{(j)} = \cdot) = p^{(j)}(\cdot)$. Then, the population in the $(n+1)$th can be expressed as

$$\mathbf{Z}_{n+1} = \sum_{j=1}^{d} \sum_{r=1}^{Z_{n,j}} \xi_{n,r}^{(j)}.$$

Let $m_{ij} = E(Z_{1,j}|\mathbf{Z}_0 = \mathbf{e}_i)$ be the expected number of type j offspring of a single type i individual in one generation for any $i, j = 1, 2, \cdots, d$. Then, we define the mean matrix

$$\mathbf{M} = \{m_{ij} : i, j = 1, 2, \cdots, d\}.$$

Clearly, we get $E(\mathbf{Z}_n|\mathbf{Z}_0) = \mathbf{Z}_0\mathbf{M}^n$. We let $m_{ij}^{(n)}$ be the (i, j)th element of $\mathbf{M}^n$.

When the higher moments exist, we can denote them by the following notations. First, we let

$$q_n^{(r)}(i, j) = E(Z_{n,i}^{(r)} Z_{n,j}^{(r)} - \delta_{i,j} Z_{n,i}^{(r)}) \quad i, j, r = 1, 2, \cdots, d$$

and define the matrix

$$\mathbf{Q}_n^{(r)} = \{q_n^{(r)}(i, j) : i, j = 1, 2, \cdots, d\},$$

the vector of matrices $\mathbb{Q}_n = (\mathbf{Q}_n^{(1)}, \mathbf{Q}_n^{(2)}, \cdots, \mathbf{Q}_n^{(d)})$, the quadratic form

$$\mathbf{Q}_n^{(r)}[\mathbf{s}] = \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} s_i q_n^{(r)}(i, j) s_j,$$

and the vectors of quadratic forms

$$\mathbb{Q}_n[\mathbf{s}] = (\mathbf{Q}_n^{(1)}[\mathbf{s}], \mathbf{Q}_n^{(2)}[\mathbf{s}], \cdots, \mathbf{Q}_n^{(d)}[\mathbf{s}]) \tag{1.1}$$

and let $\mathbb{Q}[\mathbf{s}] \equiv \mathbb{Q}_1[\mathbf{s}]$.

We also impose the following assumptions to the process $\{\mathbf{Z}_n\}_{n\geq 0}$:

1. The branching process $\{\mathbf{Z}_n\}_{n\geq 0}$ is a non-singular process, i.e., for every i, the probability that each individual has exactly one offspring of the same type is less than 1.

2. The branching process $\{\mathbf{Z}_n\}_{n\geq 0}$ is a positive regular process. That is, the mean matrix $\mathbf{M}$ is strictly positive (there exists an n such that $m_{ij}^{(n)} > 0$ for all $i, j = 1, 2, \cdots, d$).

By the Frobenius theorem, the strictly positive matrix $\mathbf{M}$ has a maximal eigenvalue ρ which is positive, simple and has associated positive right and left eigenvectors $\mathbf{u}$ and $\mathbf{v}$. Moreover, these can be normalized so that $\mathbf{u} \cdot \mathbf{v} = 1$ and $\mathbf{u} \cdot \mathbf{1} = 1$, then one can write

$$\mathbf{M}^n = \rho^n \mathbf{P} + \mathbf{R}^n$$

where $\mathbf{P}$ is the matrix whose (i, j)th entry is $u_i v_j$, and where $\mathbf{PR} = \mathbf{RP} = \mathbf{0}$ and the $r_{ij}^{(n)} \le c\rho_0^n$, $i, j = 1, 2, \cdots, d$, for some $c < \infty$ and $0 < \rho_0 < \rho$, where $r_{ij}^{(n)}$ is the (i, j)th entry of $\mathbf{R}^n$.

In a discrete-time multi-type Galton-Watson branching process, the role of the crucial criticality parameter is now played by the maximal eigenvalue ρ of the mean matrix $\mathbf{M}$. The process is called a supercritical, critical or subcritical branching process according as $\rho > 1$, $\rho = 1$ or $\rho < 1$, respectively.

1.3.2 Limit Theorems

Let $\{\mathbf{Z}_n\}_{n \ge 1}$ be a nonsingular and positive regular branching process and let $\mathbf{M}$ be its mean matrix with the maximal eigenvalue ρ.

First, we present the result of the probability of the extinction.

Theorem 1.5. *(Harris, 1963) Let*

$$\mathbf{q} = (q^{(1)}, q^{(2)}, \cdots, q^{(d)})$$

where $q^{(i)}$ is the probability of eventual extinction of the process initiated by a single individual of type i, $i = 1, 2, \cdots, d$. Then

(a) If $\rho \le 1$, then $\mathbf{q} = \mathbf{1}$.

(b) If $\rho > 1$, then $\mathbf{q} < \mathbf{1}$.

Next, we look at three limit theorems for multi-type branching processes.

In fact, the asymptotic behavior of the multi-type branching process offers no new surprises. In supercritical case, as in the single-type process, the total population $|\mathbf{Z}_n|$ grows with a geometric rate of ρ^n (need an analog of the $L \log L$ condition in Theorem 1.1 (c)) and the proportions of individuals of various types approach the corresponding ratios of the components of the left eigenvector of the mean matrix $\mathbf{M}$.

Theorem 1.6. *(Supercritical Case) Let* $\rho > 1$. *Then*

(a) (Kesten and Stigum, 1966)

$$\lim_{n\to\infty}\left(\frac{Z_n}{\rho^n}\right) = \mathbf{v}W \quad w.p.1$$

where W is a nonnegative random variable such that

$$P(W > 0) > 0 \quad \text{if and only if} \quad E\|Z_1\| \log \|Z_1\| < \infty.$$

Moreover, if $E\|Z_1\| \log \|Z_1\| < \infty$, *then*

$$E(W|\mathbf{Z}_0 = \mathbf{e}_i) = u_i \quad i = 1, 2, \cdots, d$$

and $P(W = 0|\mathbf{Z}_0 = \mathbf{e}_i) = q^{(i)}$ *for* $i = 1, 2, \cdots, d$.

(b) Let $W_n = \dfrac{\mathbf{u}\cdot\mathbf{Z}_n}{\rho^n}$ *and* $\mathbb{F}_n$ *be the* σ*-algebra generated by* $\{\mathbf{Z}_i : 1 \leq i \leq n\}$. *Then* $\{(W_n, \mathbb{F}_n) : n \geq 0\}$ *is a nonnegative martingale and hence* $\lim_{n\to\infty} W_n$ *exists w.p.1 and equals W in (a).*

Hoppe (1976) combined the functional equation approach of Seneta with the exponential martingale of Heyde to show that the analogous results of Seneta for the single-type Galton-Watson branching process also holds for the multi-type process.

Theorem 1.7. *(Hoppe [21]) Let* $1 < \rho < \infty$ *and* $\mathbf{Z}_0 = \mathbf{e}_i$, *for any* $i = 1, 2, \cdots, d$. *Then there exist positive sequence* $\{\mathbf{C}_0\}_{n\geq 1}$ *of vectors and related scalars* $\{\gamma_n\}_{n\geq 0}$ *such that*

(a) $\lim_{n\to\infty} \mathbf{C}_n \cdot \mathbf{Z}_n = W^{(i)}$ *w.p.1;*

(b) $\lim_{n\to\infty} \dfrac{\gamma_n}{\gamma_{n+1}} = \rho$*;*

(c) $\lim_{n\to\infty} \dfrac{\mathbf{C}_n}{\gamma_n} = \mathbf{u}$*;*

(d) $\lim_{n\to\infty} \gamma_n \mathbf{u} \cdot \mathbf{Z}_n = W^{(i)}$ *w.p.1;*

(e) $\lim_{n\to\infty} \gamma_n \mathbf{Z_n} = W^{(i)}\mathbf{v}$ *in probability; and*

(f) $W^{(i)}$ *is a random variable such that* $P(W^{(i)} < \infty) = 1$ *and* $P(W^{(i)} = 0) = q^{(i)}$.

(g) $\mathbf{C}_n \sim \rho^{-n}L(\rho^{-n})\mathbf{u}$ *as* $n \to \infty$, *where* $L(s)$ *varies slowly as* $s \to 0$.

In the critical case, we condition on non-extinction and normalize the process $\mathbf{Z}_n$ by dividing it the generation number n, the limit law again is exponential as it is in the single-type critical branching process.

Theorem 1.8. *(Critical Case) Let $\rho = 1$ and $E\|\mathbf{Z}_1\|^2 < \infty$. Then*

(a) (Joffe and Spitzer, 1967)

$$\lim_{n\to\infty} nP(\mathbf{Z}_n \neq \mathbf{0}|\mathbf{Z}_0 = \mathbf{i}) = \frac{\mathbf{i}\cdot\mathbf{u}}{\mathbf{v}\cdot\mathbf{Q}[\mathbf{u}]}.$$

(b) (Joffe and Spitzer, 1967) Let $\rho = 1$ and $E\|\mathbf{Z}_1\|^2 < \infty$. If $\mathbf{w}\cdot\mathbf{v} > 0$, then $\frac{\mathbf{Z}_n\cdot\mathbf{w}}{n}$, conditioned on $\mathbf{Z}_n \neq 0$, converges in distribution to the random variable Y with density

$$f(s) = \frac{1}{\gamma_1}e^{-\frac{x}{\gamma_1}}, \quad x \geq 0$$

where $\gamma_1 = \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{Q}[\mathbf{u}]}$.

(c) (Ney, 1967) Let $\rho = 1$ and $E\|\mathbf{Z}_1\|^2 < \infty$. If $\mathbf{w}\cdot\mathbf{v} = 0$, then $\frac{\mathbf{Z}_n\cdot\mathbf{w}}{\sqrt{n}}$, conditioned on $\mathbf{Z}_n \neq 0$, converges in distribution to the random variable with density

$$f_2(s) = \frac{1}{\gamma_2}e^{-\frac{|x|}{\gamma_2}}, \quad -\infty < x < \infty,$$

for some $\gamma_2 > 0$.

The same approach is applied to the subcritical case, that is, conditioned on the event of non-extinction, the process will converge to a proper random variable in distribution. Moreover, the probability of the event $\{\mathbf{Z}_n \neq \mathbf{0}\}$ of non-extinction has a geometric rate of decay ρ^n.

Theorem 1.9. *(Subcritical Case) (Joffe and Spitzer, 1967) Let $\rho < 1$. Then*

(a)

$$\frac{\mathbf{v}\cdot[\mathbf{1}-\mathbf{f}_n(\mathbf{s})]}{\rho^n} \downarrow \mathbf{Q}(\mathbf{s}) \quad \text{as } n\to\infty, \quad \mathbf{0}\leq\mathbf{s}\leq\mathbf{1},$$

where $\mathbf{Q}(\cdot)$ is non-increasing and positive if and only if $E\|\mathbf{Z}_1\|\log\|\mathbf{Z}_1\| < \infty$;

(b)

$$\lim_{n\to\infty}\frac{\mathbf{1}-\mathbf{f}_n(\mathbf{s})}{\rho^n} = \mathbf{Q}(\mathbf{s})\mathbf{u};$$

(c)

$$\lim_{n\to\infty} \rho^{-n} P(\mathbf{Z}_n \neq \mathbf{0} | \mathbf{Z}_0 = \mathbf{i}) = \mathbf{Q}(\mathbf{0})(\mathbf{i} \cdot \mathbf{u}).$$

(d)

$$\lim_{n\to\infty} P(\mathbf{Z}_n = \mathbf{j} | \mathbf{Z}_0 = \mathbf{i}, \mathbf{Z}_n \neq \mathbf{0}) = b(\mathbf{j})$$

exists, is independent of $\mathbf{i}$*, and is a probability measure on* $\mathbb{R}^+$*. Furthermore,*

$$\sum \mathbf{j} b(\mathbf{j}) < \infty \quad \text{if and only if } E\|\mathbf{Z}_1\| \log \|\mathbf{Z}_1\| < \infty.$$

1.4 Continuous-time Single-type Age-dependent Bellman-Harris Branching Processes

1.4.1 Definitions, Assumptions and Notations

In the discrete-time single-type Galton-Watson branching process, the lifetime of each individual was one unit of time. A natural generalization is to allow these lifetimes to be random variables.

Here, we first consider single-type branching process and we assume that each individual lives for a random amount of time, say L, with distribution function G and, upon its death, produces a random number ξ of children according to the offspring distribution $\{p_j\}_{j\geq 0}$. The reproduction of each individual is independent of its lifetime and of other individuals.

Let $Z(t)$ be the population at time t, i.e., the number of individuals alive at time t. Then $\{Z(t) : t \geq 0\}$ is called a continuous-time single-type Bellman-Harris branching process with the lifetime distribution $G(\cdot)$ and the offspring distribution $\{p_j\}_{j\geq 0}$.

The Galton-Watson branching process can be viewed as a special case of Bellman-Harris branching process when the lifetime $L \equiv 1$. A Bellman-Harris process is in general not Markovian, unless the lifetimes are independent exponentially distributed random variables. In such a case, i.e., the lifetimes are independently exponentially distributed, the process is called a continuous-time Markov branching process. A general Bellman-Harris process is also called a continuous-time age-dependent branching process.

As in the single-type Galton-Watson branching process, let

$$m \equiv \sum_{j=1}^{\infty} j p_j$$

and the Bellman-Harris branching process is called a supercritical, critical or subcritical process according as $m > 1$, $m = 1$ or $m < 1$.

For the lifetime distribution G, we assume throughout that $G(0+) = 0$, i.e., that there is zero probability of instantaneous death. Harris (1963) showed that, together with finite mean individual production, this guarantees the a.s. finiteness of the process for all time $t > 0$, i.e., $P(Z(t) < \infty) = 1$ for all $0 < t < \infty$.

Next, we introduce a parameter α which will describe the growth rate of the population in the supercritical case.

Definition 1.1. The Malthusian parameter for m and G is the root α in $\mathbb{R}$ (provided it exists) such that

$$m \int_0^\infty e^{-\alpha x} dG(x) = 1$$

Due to the monotonicity of the left side of the equation as a function of α, such a root, when it exists, it always unique. Also, such a Malthusian parameter always exists and is necessarily nonnegative when $m \geq 1$.

1.4.2 Limit Theorems

Let α denote the Malthusian parameter for the offspring mean m and the lifetime distribution G.

Let f be the generating function of the offspring distribution and let $F(s,t) = \sum_{j=0}^{\infty} P(Z(t) = j | Z(0) = 1) s^j$, then $F(s,t)$ is the unique bounded solution of the following integral equation

$$F(s,t) = s[1 - G(t)] + \int_0^t f(F(s, t-x)) dG(x), \quad |s| \leq 1.$$

We shall say that F is the generation function of the process determined by (f, G).

Let q be the probability of the extinction, i.e., $q = P(Z(t) = 0 \text{ for some } t)$. Then the following theorem is a direct generalization of the geometric convergence rate of the generating function of a Galton-Watson process.

Theorem 1.10. *If $m \neq 1$, $0 < \gamma = f'(q)$, G is non-lattice and the Malthusian parameter α for γ and G exists, and $\mu_\alpha = \gamma \int t e^{-\alpha t} dG(t) < \infty$, then*

$$\lim_{t \to \infty} e^{-\alpha t}(q - F(s,t)) \equiv \mathbf{Q}(s) \quad \textit{exists for } 0 \leq s \leq 1.$$

Furthermore,

$$\mathbf{Q}(s) \equiv 0 \quad \textit{if and only if} \quad m < 1 \textit{ and } \sum_{j=1}^{\infty} (j \log j) p_j = \infty.$$

If $m > 1$ *or* $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$, *then* $\mathbf{Q}(s) \neq 0$ *for* $s \neq q$.

The next theorem is for the supercritical case.

Theorem 1.11. *(Supercritical Case) Let* $1 < m < \infty$. *Then*

(a) *If* $\sum_{j=1}^{\infty} (j \log j) p_j = \infty$, *then*

$$e^{-\alpha t} Z(t) \to 0 \quad \textit{w.p.1.}$$

(b) *Let* $Z_0 = 1$ *and* $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$, *then*

$$e^{-\alpha t} Z(t) \to W \quad \textit{w.p.1}$$

where W *is a nonnegative random variable such that*

(i) $EW = 1$.

(ii) W *has an absolutely continuous distribution on* $(0, \infty)$.

(iii) $P(W = 0) = q = P(Z(t) = 0$ *for some* $t)$.

As in the discrete-time Galton-Watson branching processes, there exist the Seneta-Heyde normalizing constants for the continuous-time Bellman Harris branching processes. Schuh and Cohn (1982) showed by different approaches that if $1 < m < \infty$, without the hypothesis of $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$, there exist constants C_t such that

$$\frac{Z(t)}{C_t} \to W \quad \text{w.p.1}$$

as $t \to \infty$, where W is a continuous random variable on $(0, \infty)$ such that $P(W = 0) = q = P(Z(t) = 0$ for some $t)$.

Next, when $m = 1$, we have an analog of the exponential limit law of the critical Galton-Watson branching process.

Theorem 1.12. *(Critical Case) If* $m = 1$, $\sigma^2 = f''(1) < \infty$, $\mu = \int_0^\infty tG(t) < \infty$, *and* $t^2[1 - G(t)] \to 0$ *as* $t \to \infty$, *then*

(a) $\lim_{t\to\infty} P\Big(\frac{Z(t)}{t} \le x \Big| Z(t) > 0\Big) = 1 - e^{-\frac{2\mu x}{\sigma^2}},\ x \ge 0.$

(b) $P(Z(t) > 0) \sim \frac{2\mu}{\sigma^2 t}.$

In the subcritical case, conditioned on the event of non-extinction, the process $Z(t)$ converges to a proper random variable, as $t \to \infty$. The result is stated as follows.

Theorem 1.13. *(Subcritical Case) If* $m < 1$ *and* $\sum_{j=1}^{\infty}(j\log j)p_j < \infty$. *Assume that the Malthusian parameter* α *for* m *and the lifetime distribution* G *exists and* $\int_0^\infty te^{-\alpha t}dG(t) < \infty$. *Then*

(a) $P(Z(t) > 0) \sim ce^{-\alpha t}$ *for some* $c > 0$*;*

(b) *for all* $j \ge 1$,

$$\lim_{t\to\infty} P(Z(t) = j \big| Z(t) > 0) = b_j$$

exists, $\sum_{j=1}^{\infty} b_j = 1$ *and* $\sum_{j=1}^{\infty} jb_j < \infty$.

For proofs of these, see Athreya and Ney [5].

1.4.3 Age Distribution in Bellman-Harris Processes

An important and useful aspect of age-dependent branching processes is the limit behavior of the age distribution.

Consider a Bellman-Harris branching process $\{Z(t) : t \ge 0\}$ which starts at time 0 with one individual of age 0. This individual lives for a length of time L with the lifetime distribution G and, upon its death, produces ξ children according to the offspring distribution $\{p_j\}_{j\ge 0}$ independently of other individuals alive at the same time and of the lifetime. Then each individual lives for a length of time then dies and produces its offspring in the same way and so on.

We impose the assumption that $G(0+) = 0$ and G is non-lattice.

We also adopt the following notations: For any family history ω,

1. $Z(t, \omega)$ is the number of individuals alive at time t.

2. $Z(x, t, \omega)$ is the number of individuals alive at time t whose age is less than or equal to x.

3. $A(x, t, \omega) = \dfrac{Z(x, t, \omega)}{Z(t, \omega)}$.

4. α is the Malthusian parameter for m and G

5. $A(x) = \dfrac{\int_0^x e^{-\alpha u}[1 - G(u)]du}{\int_0^\infty e^{-\alpha u}[1 - G(u)]du}$

6. $G_x(t) = \dfrac{G(x + t) - G(x)}{1 - G(x)}$

Theorem 1.14. *(Athreya and Kaplan [2]) Let* $1 < m = \sum_{j=1}^{\infty} jp_j < \infty$ *and* $p_0 = 0$*. Then*

(a) $\sup_x |A(x, t, \omega) - A(x)| \xrightarrow{P} 0$

(b) If $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$*, then*

$$\sup_x |A(x, t, \omega) - A(x)| \to 0 \quad \textit{w.p.1}$$

as $t \to \infty$*.*

(c) For any bounded continuous a.e. (w.r.t. Lebesgue measure) function $h(\cdot)$ *on the support of* G*,*

$$\int_0^\infty h(x)dA(x, t, \omega) \xrightarrow{P} \int_0^\infty h(x)dA(x)$$

as $t \to \infty$*.*

Remark 1.1. In the above theorem, (a) implies (c)

Theorem 1.15. *(Athreya and Kaplan [2]) Let* $1 < m = \sum_{j=1}^{\infty} jp_j < \infty$ *and* $p_0 = 0$*. If* $\sum_{j=1}^{\infty}(j \log j)p_j = \infty$*. Then, for any* K *in the support of* G*,*

$$\lim_{t \to \infty} Z(K, t, \omega)e^{-\alpha t} = 0 \quad \textit{w.p.1}$$

Theorem 1.16. *(Athreya and Kaplan [3])Let* $m = 1$ *and assume that* $\lim_{t \to \infty} \sup_{x \geq 0}[1 - G_x(t)] = 0$*, then, for any* $\epsilon > 0$*,*

$$\lim_{t \to \infty} P(\sup_{x \geq 0} |A(x, t, \omega) - A(x)| > \epsilon \big| Z(t) > 0) = 0$$

1.5 Continuous-time Multi-type Age-dependent Bellman-Harris Branching Processes

1.5.1 Definitions, Assumptions and Notations

The Bellman-Harris processes can be made more general by allowing individuals to be of different types. The population consists of d types of individuals whose lifetimes and reproductive behaviors are dependent on their types.

The lifetime L_i of a type i individual is a random variable with distribution $G_i(\cdot)$, $i = 1, 2, \cdots, d$. Also, a type i individual, upon its death, produces $\xi_{i,j}$ children of type j, $j = 1, 2, \cdots, d$, according to the probability distribution $\{p^{(i)}(\mathbf{j}) \equiv p^{(i)}(j_1, j_2, \cdots, j_d)\}_{\mathbf{j}\in\mathbb{N}^d}$ and independently of other individual, where $p^{(i)}(j_1, j_2, \cdots, j_d)$ is the probability that a type i parent produces j_1 children of type 1, j_2 children of type 2, $\cdots$, j_d children of type d. As in the multi-type Galton-Watson process, we still denote the generating functions of the offspring distributions by $\mathbf{f}(\mathbf{s}) = (f^{(1)}(\mathbf{s}), f^{(2)}(\mathbf{s}), \cdots, f^{(d)}(\mathbf{S}))$.

Let $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \cdots, Z_d(t))$ be the population vector of the individuals alive at time t, $t \geq 0$, where $Z_i(t)$ is the number of individuals of type i alive at time t. Then $\{\mathbf{Z}(t) : t \geq 0\}$ is called a continuous-time age-dependent branching process.

As in the discrete-time multi-type Galton-Watson branching processes, we let $m_{ij} = E(\xi_{i,j})$ be the expected number of type j offspring of a single type i individual in one generation for any $i, j = 1, 2, \cdots, d$ and define the mean matrix $\mathbf{M} = \{m_{ij} : i, j = 1, 2, \cdots, d\}$.

Assume the $\mathbf{M}$ is nonsingular and positively regular and write ρ for its Perron-Frobenius root (the maximal eigenvale). The process is called supercritical, critical or subcritical case according as $\rho > 1$, $\rho = 1$ or $\rho < 1$, respectively.

Let $\widehat{G}(\alpha) = \int_0^\infty e^{-\alpha t} G(dt)$ be the Laplace transform of any probability distribution G.

Let $\widehat{\mathbf{M}}(\alpha) = ((m_{ij}\widehat{G}_i(\alpha)))_{i,j=1}^d$.

Now, we can define an analog to the concept of a Malthusian parameter for a multi-type Bellman-Harris processes.

Definition 1.2. The Malthusian parameter α for the matrix M and the probability distributions $\{G_i : i = 1, 2, \cdots, d\}$ is defined to be the number α for which the matrix $\widehat{\mathbf{M}}(\alpha)$ has the Perron-Frobenius root (the maximal eigenvalue) 1, provided it exists.

In the critical and supercritical cases, the Malthusian parameter α always exists and is nonnegative.

1.5.2 Limit Theorems

Let $\mathbf{u}$ and $\mathbf{v}$ be the right and left eigenvector of $\mathbf{M}$ corresponding to the maximal eigenvalue ρ such that $\mathbf{1} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{v} = 1$.

Then, this Malthusian parameter α, again, related to the growth rate in the supercritical case.

Here, we also assume that the offspring mean matrix $\mathbf{M}$ is nonsingular and strictly positive.

Theorem 1.17. *(Supercritical Case) Let $\rho > 1$ and $E(\xi_{ij}^2) < \infty$ for $i, j = 1, 2, \cdots, d$. Then*

$$\lim_{t\to\infty} e^{-\alpha t}\mathbf{Z}(t) = \mathbf{v}W \quad \textit{exists w.p.1}$$

where W is a one-dimensional random variable.

Theorem 1.18. *(Critical Case) Let $\rho = 1$, $E(\xi_{ij}^2) < \infty$ for $i, j = 1, 2, \cdots, d$ and $t^2[1 - G_i(t)] \to 0$ as $t \to \infty$, for all $i = 1, 2, \cdots, d$. Then*

$$\lim_{t\to\infty} tP(\mathbf{Z}(t) \neq 0 \big| \mathbf{Z}(0) = \mathbf{e}_i) = \left[\frac{\mu \cdot (\mathbf{u} \otimes \mathbf{v})}{\mathbf{Q}[\mathbf{u}]}\right] u_i$$

where $\mu = (\mu_1, \mu_2, \cdots, \mu_d)$ is the vector of means of $(G_1(t), G_2(t), \cdots, G_d(t))$, $\mathbf{u} \otimes \mathbf{v} = (u_1 v_1, u_2 v_2, \cdots, u_d v_d)$ and $\mathbf{Q}$ is the second moment quadratic for associated with $\mathbf{f}$.

Moreover, the corresponding exponential limit law for $\left\{\frac{\mathbf{Z}(t)}{t}\Big|\mathbf{Z}(t) \neq \mathbf{0}\right\}$ has been proved by H. Weiner (1970) under the very strong assumption that all moments of $\mathbf{f}(\mathbf{s})$ exists.

The next result is the analog of the limit law in the subcritical multi-type Galton-Watson branching process.

Theorem 1.19. *(Subcritical Case) Let $\rho < 1$. Assume that $E(\xi_{ij} \log^+ \xi_{ij}) < \infty$ for all $i, j = 1, 2, \cdots, d$, the Malthusian parameter α exists and $\int_0^\infty te^{\alpha t} dG_i(t) < \infty$ for $i = 1, 2, \cdots, d$. Then, as $t \to \infty$,*

(a) $P(\mathbf{Z}(t) \neq 0 | \mathbf{Z}(0) = \mathbf{e}_i) \sim c_i e^{-\alpha t}$ *for some* $c_i > 0$.

(b) $P(\mathbf{Z}(t) = \mathbf{j} | \mathbf{Z}(0) = \mathbf{e}_i, \mathbf{Z}(t) \neq 0) \to b(\mathbf{j})$ *where* $b(\mathbf{j})$ *is a probability measure on* $\mathbb{R}_+^d - \{\mathbf{0}\}$.

1.5.3 Age Distribution in Multi-type Age-dependent Branching Processes

We also have the analogs of limits theorems of the age distribution in the multi-type age-dependent process.

Assume, for each $i = 1, 2, \cdots, d$, the lifetime distribution G_i of a type i individual is non-lattice and $G_i(0+) = 0$. Also, for any family history ω, let

1. $\mathbf{Z}(t, \omega) = (Z_1(t, \omega), Z_2(t, \omega), \cdots, Z_d(t, \omega))$. where $Z_i(t, \omega)$ is the number of type i individuals alive at time t.
2. $|\mathbf{Z}(t, \omega)| = \sum_{i=1}^{d} Z_i(t, \omega)$
3. $\mathbf{Z}(t, x, \omega) = (Z_1(t, x, \omega), Z_2(t, x, \omega), \cdots, Z_d(t, x, \omega))$. where $Z_i(t, x, \omega)$ is the number of type i individuals alive at time t whose age is less than or equal to x.
4. for $i = 1, 2, \cdots, d$, $A_i(t, x, \omega) = \dfrac{Z_i(t, x, \omega)}{Z_i(t, \omega)}$ for $Z_i(t, \omega) > 0$
5. $A_i(x) = \dfrac{\int_0^x e^{-\alpha u}[1 - G_i(u)]du}{\int_0^\infty e^{-\alpha u}[1 - G_i(u)]du}$

Recall that ξ_{ij} is the number of type j children produced by a type i parent, $i, j = 1, 2, \cdots, d$, according to the offspring distribution.

Theorem 1.20. *Assume the process is supercritical, i.e.* $\rho > 1$ *and* $E(\xi_{ij} \log^+ \xi_{ij}) < \infty$ *for all* $i, j = 1, 2, \cdots, d$, *then, conditioned on the event of non-extinction, for any* $i = 1, 2, \cdots, d$,

$$\sup_{x \geq 0} \left| A_i(t, x, \omega) - A_i(x) \right| \to 0 \quad w.p.1 \quad as\ t \to \infty.$$

For proofs, see Athreya and Ney [5].

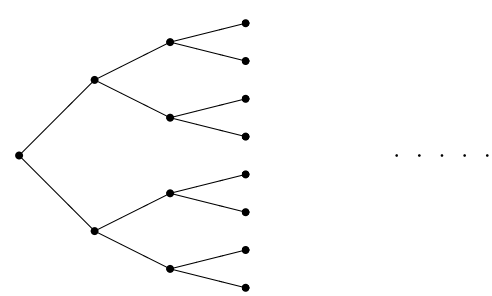
CHAPTER 2. REVIEW OF THE COALESCENCE IN DISCRETE-TIME SINGLE-TYPE GALTON-WATSON BRANCHING PROCESSES

2.1 Introduction

To investigate an old population, there are two interesting research directions. One is to predict the future behavior of this population, such as the probability of extinction, the growth rate, the stability of the composition of the population in a multi-type case and the limit distribution of the ages of individuals. On the other hand, we can also study the evolution of this old population backward in time given its probability structure. We have seen many classical theorems regarding the first direction for different branching processes in Chapter 1. In this chapter, we will review some known results for the other direction in the discrete-time single-type Galton-Watson branching processes for the other direction.

We start with the coalescence problem for the binary tree case.

Consider a binary tree $\mathcal{T}$ starting with one individual. In such a tree, each individual produces exactly two offspring upon its death. So, there are 2^n individuals in the nth generation for any $n = 0, 1, 2, \cdots$.



Now, pick two individuals from the nth generation by the simple random sampling without replacement and trace their lines of descent backward in time until they meet. Call that generation X_n. Then, for $k = 1, 2, \cdots, n$,

$$P(X_n < k) = \frac{\binom{2^k}{2} 2^{n-k} 2^{n-k}}{\binom{2^n}{2}} = \frac{2^k(2^k - 1)2^{n-k}2^{n-k}}{2^n(2^n - 1)} = \frac{1 - 2^{-k}}{1 - 2^{-n}}$$

So, $\lim_{n\to\infty} P(X_n < k) = 1 - 2^{-k}$, $k = 1, 2, \cdots$. Thus, $X_n \xrightarrow{d} Geo(\frac{1}{2})$.

Similar results holds for any regular b-nary tree, $b \geq 2$. This suggests that a similar behavior might hold for Galton-Watson trees.

Let $\{Z_n\}_{n\geq 1}$ be a discrete-time single-type Galton-Watson branching process with offspring distribution $\{p_j\}_{j\geq 0}$ and initiated size Z_0. Here, the process $\{Z_n\}_{n\geq 0}$ is generated by the way described in Section 1.2 and we will also adopt the notations introduced in the same section.

If $\mathcal{T}$ denotes the full family tree, then every individual in $\mathcal{T}$ can be identified by a finite string $(i_0, i_1, \cdots, i_n)$ meaning that this individual is in the nth generation and is the i_nth child of the individual $(i_0, i_1, \cdots, i_{n-1})$ in the $(n-1)$st generation.

Pick two individuals form the population in the nth generation (assuming $Z_n \geq 2$) by the simple random sampling without replacement and trace their lines of descent backward in time until they meet for the first time. Call this common ancestor *the last common ancestor* or *the most recent common ancestor* of these two randomly chosen individuals. Let $X_{n,2}$ be the number of the generation which this common ancestor belonged to. Then we can ask the following questions.

(1) What is the distribution of $X_{n,2}$?

(2) What happens to $X_{n,2}$ when $n \to \infty$?

Similarly, if we pick k individuals randomly from the nth generation, $k \geq 2$, and trace their lines of decent backward in time until they meet. Let $X_{n,k}$ be the generation number of the last common ancestor of these randomly chosen individuals. Moreover, let Y_n be the generation number of the last common ancestor of all the individuals in the nth generation and we are also interested in the limit behaviors of the distributions of $X_{n,k}$, $k \geq 2$, and Y_n when $n \to \infty$.

In each of the following sections, we present the results on the coalescence problem for different cases (supercritical, critical, subcritical and explosive) in the discrete-time Single-type Galton-Watson branching processes. For proofs, see Athreya [10] and [12].

2.2 The Supercritical Case

In the supercritical case, Athreya showed that the coalescence time $X_{n,k}$ will go way back to the beginning of the tree for any $k \geq 2$. That is, $X_{n,k}$ converges to a proper random variable in distribution as $n \to \infty$.

Theorem 2.1. *Let* $p_0 = 0$, $1 < m \equiv \sum_{j=1}^{\infty} jp_j < \infty$. *Then, for almost all trees* $\mathcal{T}$,

(a) for $\forall 1 \leq r < \infty$,

$$\lim_{n\to\infty} P(X_{n,2} < r|\mathcal{T}) \equiv \pi_2(r,\mathcal{T}) \quad \textit{exists}$$

and $\pi_2(r,\mathcal{T}) \uparrow 1$ *as* $r \uparrow \infty$.

(b) for $\forall k \geq 2$, $\forall 1 \leq r < \infty$,

$$\lim_{n\to\infty} P(X_{n,k} < r|\mathcal{T}) \equiv \pi_k(r,\mathcal{T}) \quad \textit{exists}$$

and $\pi_k(r,\mathcal{T}) \uparrow 1$ *as* $r \uparrow \infty$.

(c) for almost all trees $\mathcal{T}$,

$$Y_n \to N(\mathcal{T})$$

where $N(\mathcal{T}) = \max\{n \geq 1 : Z_n = 1\}$. *Also,*

$$\lim_{n\to\infty} P(Y_n = k) = (1 - p_1)p_1^k \quad , k \geq 0.$$

2.3 The Critical Case

In a discrete-time single-type critical Galton-Watson branching process, unlike the results in the supercritical case, the coalescence time $X_{n,k}$ of k randomly chosen individuals, $k \geq 2$, as well as the coalescence time Y_n of the whole population are not close to the beginning of the tree when n gets large. In fact, they are of order n. That is, $\frac{X_{n,k}}{n}$ (conditioned on the set $\{Z_n \geq k\}$), $k \geq 2$, and $\frac{Y_n}{n}$ (conditioned on $\{Z_n \geq 1\}$) converge to proper random variables, respectively, when $n \to \infty$.

Theorem 2.2. *Let* $m = 1$, $p_1 < 1$ *and* $\sigma^2 = \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$, *Then, for* $0 < u < 1$, *(a)* $\lim_{n\to\infty} P\Big(\frac{X_{n,2}}{n} < u \Big| Z_n \geq 2\Big) \equiv H_2(u)$ *exists and for* $0 < u < 1$,

$$H_2(u) \equiv 1 - E\varphi(N_u)$$

where N_u *is a geometric random variable with distribution*

$$P(N_u = k) = (1-u)u^{k-1} \quad , k \geq 1$$

and for $j \geq 1$,

$$\varphi(j) \equiv E\Big(\frac{\sum_{i=1}^{j} \eta_i^2}{(\sum_{i=1}^{j} \eta_i)^2}\Big)$$

where $\{\eta_i\}_{i\geq 1}$ *are i.i.d. exponential random variable with* $E\eta_1 = 1$.

Further, $H_2(\cdot)$ *is absolutely continuous on* $[0, 1]$ *with* $H(0+) = 0$ *and* $H(1-) = 1$.

(b) *for* $0 < u < 1$, $1 < k < \infty$,

$$\lim_{n\to\infty} P\Big(\frac{X_{n,k}}{n} < u \Big| Z_n \geq k\Big) \equiv H_k(u) \quad \textit{exists}$$

and $H_k(\cdot)$ *is an absolutely continuous distribution function with* $H_k(0+) = 0$ *and* $H_k(1-) = 1$.

(c) *for* $0 < u < 1$, $\lim_{n\to\infty} P\Big(\frac{Y_n}{n} < u \Big| Z_n \geq 1\Big) = u$.

Remark 2.1. Theorem 2.2 (c) shows that $\frac{Y_n}{n}$ is available in Zubkov [36].

2.4 The Subcritical Case

The next theorem provides a sharp contract to the results in the supercritical case and critical case. In a subcritical Galton-Watson branching process, the coalescence time $X_{n,k}$, $k \geq 2$, takes place close to the present and same is true for the coalescence time Y_n of the whole population in the nth generation.

Theorem 2.3. *Let* $0 < m \equiv \sum_{j=1}^{\infty} jp_j < 1$. *Then*

(a) *For* $k \geq 1$, $\lim_{n\to\infty} P(n - X_n > k | Z_n \geq 2) = \frac{E\phi_k(Y)}{E\psi_k(Y)} \equiv \pi_k$, *say, where*

$$\phi_k(j) = E\Big(\frac{\sum_{i_1\neq i_2=1}^{j} Z_{k,i_1} Z_{k,i_2}}{(\sum_{i=1}^{j} Z_{k,i})(\sum_{i=1}^{j} Z_{k,i} - 1)} I(\sum_{i=1}^{j} Z_{k,i} \geq 1)\Big)$$

and

$$\psi_k(j) = P\Big(\sum_{i=1}^{j} Z_{k,i} \geq 2\Big)$$

where $\{Z_{r,i} : r \geq 0\}$, $i = 1, 2, \cdots$ *are i.i.d. copies of a Galton-Watson branching process* $\{Z_r : r \geq 0\}$ *with* $Z_0 = 1$ *and the given offspring distribution* $\{p_j\}_{j\geq 0}$ *and* Y *is a random variable with distribution* $\{b_j\}_{j\geq 1}$ *where*

$$b_j \equiv \lim_{n\to\infty} P(Z_n = j | Z_n > 0, Z_0 = 1) \quad \text{which exists.}$$

Further, if $\sum_{j=1}^{\infty} j \log j p_j < \infty$, *then* $\lim_{k\uparrow\infty} \pi_k = 0$ *and hence* $n - X_n$ *conditioned on* $Z_n \geq 2$ *converges to a proper distribution on* $\{1, 2, \cdots\}$.

(b) *For* $k \geq 1$, $\lim_{n\to\infty} P(n - Y_n > k | Z_n \geq 1) \equiv \tilde{\pi}_k$ *exists and equals*

$$E\Big(\frac{1 - q_k^Y}{m^k}\Big) - E\Big(\frac{Y q^{k-1}(1 - q_k)}{m^k}\Big)$$

where Y *is a random variable with distribution*

$$P(Y = j) = b_j = \lim_{n\to\infty} P(Z_n = j | Z_n > 0, Z_0 = 1)$$

and $q_k = P(Z_k = 0 | Z_0 = 1)$.

Further, if $\sum_{j=1}^{\infty} j \log j p_j < \infty$, *then* $\lim_{k\to\infty} \tilde{\pi}_k = 0$.

That is, $n - Y_n$ *conditioned on* $\{Z_n > 0\}$ *converges in distribution as* $n \to \infty$ *to a proper distribution on* $\{1, 2, \cdots\}$.

2.5 The Explosive Case

If one considers a rapidly growing population, then two individuals chosen randomly from the nth generation are unlikely to be closely related to each other when n gets large. Theorem 2.1 says that in the supercritical case, the coalescence times do go way back to the beginning of the tree. Surprisingly, it turns out that, when $m = \infty$, this is only true for the coalescence time Y_n of the whole population in the nth generation when n gets large. The coalescence times $X_{n,k}$, $k \geq 2$, turn out to be very close to the present and, in fact, $n - X_{n,k}$, $k \geq 2$ converges to a proper random variable in distribution when $n \to \infty$.

Theorem 2.4. *Let* $p_0 = 0$, $m = \sum_{j=1}^{\infty} j p_j = \infty$, *and for some* $0 < \alpha < 1$, *and a function* $L : (1, \infty) \to (0, \infty)$ *slowly varying at* ∞. *Let*

$$\frac{\sum_{j>x} p_j}{x^\alpha L(x)} \to 1 \quad \text{as } x \to \infty.$$

Then

(a) For almost all trees $\mathcal{T}$ *and* $k = 1, 2, \cdots$, *as* $n \to \infty$,

$$P(X_{n,2} < k|\mathcal{T}) \to 0$$

and

$$P(n - X_{n,2} < k) \to \pi_2(k) \quad \text{exists}$$

and $\pi_2(k) \uparrow 1$ *as* $k \uparrow \infty$.

(b) For any $1 < j < \infty$ *and* $k = 1, 2, \cdots$

$$P(X_{n,j} < k|\mathcal{T}) \to 0 \quad \text{as } n \to \infty$$

and $P(n - X_{n,j} < k) \to \pi_j(k)$ *exists and* $\pi_j(k) \uparrow 1$ *as* $k \uparrow \infty$.

(c) $Y_n \xrightarrow{d} N(\mathcal{T}) \equiv \max\{j : Z_j = 1\} < \infty$ *and*

$$P(Y_n = k) \to (1 - p_1) p_1^{k-1} \quad , k \geq 1.$$

CHAPTER 3. COALESCENCE IN DISCRETE-TIME MULTI-TYPE GALTON-WATSON BRANCHING PROCESSES

3.1 Introduction

Throughout this chapter, we consider a $d-$type ($2 \leq d < \infty$) Galton-Watson branching process and also adopt all the definitions and notations described in Section 1.3.

Let $\{\mathbf{Z}_n\}_{n\geq 0}$ be a discrete-time multi-type Galton-Watson branching process, i.e.,

$$\mathbf{Z}_n = (Z_{n,1}, Z_{n,2}, \cdots, Z_{n,d})$$

is the population vector in the nth generation, $n = 0, 1, 2, \cdots$, where $Z_{n,i}$ is the number of individuals of type i in the nth generation, $1 \leq i \leq d$.

We impose the following assumptions to the process $\{\mathbf{Z}_n\}_{n\geq 0}$:

1. The branching process $\{\mathbf{Z}_n\}_{n\geq 0}$ is a non-singular process, i.e., for every i, the probability that each individual has exactly one offspring of the same type is less than 1.

2. The branching process $\{\mathbf{Z}_n\}_{n\geq 0}$ is a positive regular process. That is, the mean matrix $\mathbf{M}$ is strictly positive (there exists an n such that $m_{ij}^{(n)} > 0$ for all $i, j = 1, 2, \cdots, d$).

Let $\mathcal{T}$ denote the full discrete-time multi-type Galton-Watson family tree, every individual in $\mathcal{T}$ can be identified by a finite string $((r_0, i_0), (r_1, i_1), \cdots, (r_n, i_n))$ meaning that this individual is in the nth generation and is the r_nth child of type i_n of the individual $((r_0, i_0), (r_1, i_1), \cdots, (r_{n-1}, i_{n-1}))$ in the $(n-1)$st generation.

Let $k \geq 2$ be a positive integer. Now, we pick k individuals from the population in the nth generation (assuming $|\mathbf{Z}_n| \geq k$) by the simple random sampling without replacement and trace their lines of descent backward in time until they meet for the first time. Call this common ancestor *the last common ancestor* or *the most recent common ancestor* of these two randomly chosen individuals. Let $X_{n,k}$, the coalescence time, be the number of the generation which the last common ancestor belonged to. Then we are interested in the following questions.

(1) What is the distribution of $X_{n,k}$?

(2) What happens to $X_{n,k}$ when $n \to \infty$?

(3) What happens when $k \to \infty$?

(4) What happens to the generation number of the last common ancestor of the whole population in the nth generation when n gets large?

We have seen the results on the discrete-time single-type Galton-Watson branching process in Chapter 2 and we would like to extend those to the multi-type supercritical, critical and subcritical branching processes. Moreover, we are also interested in the questions involving the types:

(5) What is the joint distribution of the type and the generation number of the last common ancestor and the types of the randomly chosen individuals?

(6) What happens to this joint distribution when n gets large?

We present the results for the supercritical, critical and subcritical cases in Section 3.2, 3.3 and 3.4, respectively. In Section 3.5, we also investigate the Markov property of the limit law of the types of the ancestors of any random chosen individual from the $n$the generation along its line of the descent .

3.2 Results in The Supercritical Case

For a supercritical branching process, we assume that each individual has to produce at least one offspring w.p.1 upon death, that is, $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ for any $i = 1, 2, \cdots, d$. Also, $E(Z_{1,j}\big|\mathbf{Z}_0 = \mathbf{e}_i) \equiv m_{ij} < \infty$ for all $1 \le i, j \le d$.

Let ρ be the maximal eigenvalue of $\mathbf{M} = \{m_{ij} : i, j = 1, 2, \cdots, d\}$.

3.2.1 The Statements of Results

Theorem 3.1. *Let* $\rho > 1$*,* $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ *and* $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ *for all* $1 \leq i \leq d$*. Then, for* $k = 2, 3, \cdots$*,*

(a) for almost all trees $\mathcal{T}$ *and* $r = 1, 2, \cdots$*,*

$$P(X_{n,k} < r|\mathcal{T}) \to \phi_k(r, \mathcal{T}) \equiv 1 - \frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k}$$

as $n \to \infty$*, where* $\{W_{r,i} : i \geq 1, r \geq 1\}$ *are i.i.d. copies of* $W \equiv lim_{n\to\infty} \frac{\mathbf{u} \cdot \mathbf{Z}_n}{\rho^n}$ *in Theorem 1.6 (a).*

(b) there exist random variable $\tilde{X}_k$ *such that* $X_{n,k} \xrightarrow{d} \tilde{X}_k$ *as* $n \to \infty$*, where*

$$P(\tilde{X}_k < r) \equiv \phi_k(r) = 1 - E\left(\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k}\right)$$

for any $r = 1, 2, \cdots$*.*

Remark 3.1. It may be noted that $W \equiv lim_{n\to\infty} \frac{\mathbf{u} \cdot \mathbf{Z}_n}{\rho^n}$ has the same distribution for all $\mathbf{Z}_0$.

Now, here arises an interesting question. Since the coalescence time of $X_{n,k}$ of k randomly chosen individuals in the nth generation converges in distribution to a proper random variable $\tilde{X}_k$ as $n \to \infty$ for every positive integer $k \geq 2$, what happens to $\tilde{X}_k$ if $k \to \infty$? The following theorem tells us that $\tilde{X}_k$ will also converge in distribution as $k \to \infty$ to the generation which is the last time when the tree consists of only one individual.

Theorem 3.2. *Let* $\rho > 1$ *and* $E\|Z_1\| \log \|Z_1\| < \infty$*. Let* $U = \min\{n \geq 1 : |\mathbf{Z}_n| \geq 2\}$ *be the first time when the population exceeds* 1*. Then* $\tilde{X}_k \xrightarrow{d} U - 1$ *as* $k \to \infty$*.*

Next, we pick two individuals (i.e. consider $k = 2$) at random by simple random sampling without replacement from the nth generation and trace their lines of decent backward in time to find their last common ancestor. Let $X_{n,2}$ be the generation number of this common ancestor, η_n the type of this common ancestor and $(\zeta_{n,1}, \zeta_{n,2})$ be the typea of the chosen individuals. The following theorem asserts that the joint distribution of $(X_{n,2}, \eta_n, \zeta_{n,1}, \zeta_{n,2})$ converges as $n \to \infty$ to a proper distribution.

Theorem 3.3. *Let $\rho > 1$, $\mathbf{Z}_0 = \mathbf{e}_{i_0}$, $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ and $E\|Z_1\| \log \|Z_1\| < \infty$. Then*

$$\lim_{n\to\infty} P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2) \equiv \varphi_2(r, j, i_1, i_2) \quad \text{exists}$$

and $\sum_{(r,j,i_1,i_2)} \varphi_2(r, j, i_1, i_2) = 1.$

The following extension of Theorem 3.3 is also valid to any integer $k = 2, 3, \cdots$.

Theorem 3.4. *Let $\rho > 1$, $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ and $E\|Z_1\| \log \|Z_1\| < \infty$. Then, for any $2 \le k < \infty$,*

$$\lim_{n\to\infty} P(X_{n,k} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2, \cdots, \zeta_{n,k} = i_k) \equiv \varphi_k(r, j, i_1, i_2, \cdots, i_k)$$

exists and $\sum_{(r,j,i_1,i_2,\cdots,i_k)} \varphi_k(r, j, i_1, i_2, \cdots, i_k) = 1.$

3.2.2 The Proof of Theorem 3.1

We need the following lemma for the proofs.

Lemma 3.1. *(O'Brien, 1980) Assume $W_1, W_2, \cdots$ are pairwise independent.*

$$\frac{\max\{W_1, W_2, \cdots, W_n\}}{\sum_{i=1}^{n} W_i} \to 0 \quad \text{in probability}$$

if and only if $L(x) \equiv E(W : W \le x)$ is slowly varying at ∞.

Now, we begin to prove Theorem 3.1.

Let $\left\{\mathbf{Z}^{(l)}_{p,i,n-p}\right\}_{n\ge p}$ be the be the discrete-time multi-type Galton-Watson branching process initiated by the ith individual of type l in the pth generation.

For any $k \ge 2$, we pick k individuals by simple random sampling without replacement from the population in the nth generation and let $X_{n,k}$ be the generation number of their last common ancestor.

(a) For almost all trees $\mathcal{T}$ and $r = 1, 2, \cdots$,

$$\begin{aligned} P(X_{n,k} \ge r|\mathcal{T}) &= \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} \left|\mathbf{Z}^{(l)}_{r,i,n-r}\right|\left(\left|\mathbf{Z}^{(l)}_{r,i,n-r}\right| - 1\right)\cdots\left(\left|\mathbf{Z}^{(l)}_{r,i,n-r}\right| - k + 1\right)}{|\mathbf{Z}_n|(|\mathbf{Z}_n| - 1)\cdots(|\mathbf{Z}_n| - k + 1)} \\ &= \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} \frac{\left|\mathbf{Z}^{(l)}_{r,i,n-r}\right|}{\rho^{n-r}} \frac{\left|\mathbf{Z}^{(l)}_{r,i,n-r}\right|-1}{\rho^{n-r}} \cdots \frac{\left|\mathbf{Z}^{(l)}_{r,i,n-r}\right|-k+1}{\rho^{n-r}}}{\left(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}^{(l)}_{r,i,n-r}|}{\rho^{n-r}}\right)\left(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}^{(l)}_{r,i,n-r}|}{\rho^{n-r}} - \frac{1}{\rho^{n-r}}\right)\cdots\left(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}^{(l)}_{r,i,n-r}|}{\rho^{n-r}} - \frac{k-1}{\rho^{n-r}}\right)} \end{aligned}$$

Since $\rho > 1$ and $E\|Z_1\| \log \|Z_1\| < \infty$, by Theorem 1.6 (a), we know that $\frac{\left|\mathbf{Z}^{(l)}_{r,i,n-r}\right|}{\rho^{n-r}} \to (\mathbf{1} \cdot \mathbf{v})W_{r,i}$ w.p.1, for any $i = 1, 2, \cdots$, where $\{W_{r,i} : i \geq 1, r \geq 1\}$ are i.i.d. copies of W in Theorem 1.6 (a). So,

$$P(X_{n,k} \geq r|\mathcal{T}) \to \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} ((1 \cdot \mathbf{v})W_{r,i})^k}{\left(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} (1 \cdot \mathbf{v})W_{r,i}\right)^k} = \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} W_{r,i}^k}{\left(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} W_{r,i}\right)^k} = \frac{\sum_{i=1}^{|Z_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|Z_r|} W_{r,i}\right)^k} \equiv 1 - \phi_k(r, \mathcal{T})$$

and hence (a) is proved.

(b) Since $P(X_{n,k} \geq r) = E(P(X_{n,k} \geq r)\big|\mathcal{T})$ and hence, by the bounded convergence theorem,

$$P(X_{n,k} \geq r) \to E\left(\frac{\sum_{i=1}^{|Z_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|Z_r|} W_{r,i}\right)^k}\right) \equiv 1 - \phi_k(r) \quad \text{as } n \to \infty$$

for $r = 1, 2, \cdots$.

Moreover, since $E\|Z_1\| \log \|Z_1\| < \infty$, by Kesten and Stigum's result (1966), $EW < \infty$. Hence, if

$$L(x) \equiv E(W : W \leq x)$$

then

$$\lim_{x\to\infty} L(x) = \lim_{x\to\infty} E(W : W \leq x) = EW$$

where $0 < EW < \infty$. So, for any $0 < c < \infty$,

$$\lim_{x\to\infty} \frac{L(cx)}{L(x)} = 1.$$

That is, the function $E(W_{r,1} : W_{r,1} \leq x)$ in x is slowly varying at ∞.

Therefore, by Lemma 3.1,

$$\frac{\max_{1\leq i\leq n} W_{r,i}}{\sum_{i=1}^{n} W_{r,i}} \to 0 \quad \text{in probability}$$

as $n \to \infty$. So, since $|Z_r| \to \infty$ w.p.1 as $r \to \infty$, by the bounded convergence theorem, we have

$$E\left(\frac{\sum_{i=1}^{|Z_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|Z_r|} W_{r,i}\right)^k}\right) \leq E\left(\frac{\max_{1\leq i\leq |Z_r|} W_{r,i}}{\sum_{i=1}^{|Z_r|} W_{r,i}}\right)^{k-1} \to 0 \qquad \text{as } r \to \infty.$$

Thus, ϕ_k is a proper probability distribution. So, there exists a random variable $\tilde{X}_k$ with $P(\tilde{X}_k < r) = \varphi_k(r)$ for any $r \geq 1$ such that $X_{n,k} \xrightarrow{d} \tilde{X}_k$ as $n \to \infty$ and we have completed the proof of Theorem 3.1.

Remark 3.2. Theorem 3.1 (a) and (b) should be valid just with $\rho > 1$. That is, the assumption $E\|Z_1\| \log \|Z_1\| < \infty$ can be dropped. This will need Hoppe's result [21] and the result that the function $E(W : W \leq x)$ is slowly varying at ∞. For single-type case, this was proved by Athreya and Schun [4]. It can be adapted to the multi-type case.

3.2.3 The Proof of Theorem 3.2

We prove this theorem in two steps.

Step 1.

Since $U = \min\{n \geq 1 : |\mathbf{Z}_n| \geq 2\}$, for almost all trees $\mathcal{T}$ and any $r = 1, 2, \cdots$, we have that

$$\phi_k(r, \mathcal{T}) = 1 - \frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} = \begin{cases} 1 - \dfrac{W_{r,1}^k}{W_{r,1}^k} & \text{if } r \leq U - 1 \\[2ex] 1 - \dfrac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} & \text{if } r > U \end{cases}$$

Also, the assumption that $P(\mathbf{Z}_1 = \mathbf{0} | \mathbf{Z}_0 = \mathbf{e}_i) = 0$ for all $i = 1, 2, \cdots, d$ implies that

$$P\left(0 < \frac{\max_{1 \leq i \leq N} W_{r,i}}{\sum_{i=1}^{N} W_{r,i}} < 1\right) = 1$$

for any $N \geq 2$. So, for almost all trees $\mathcal{T}$,

$$\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} \leq \left(\frac{\max_{1 \leq i \leq |\mathbf{Z}_r|} W_{r,i}}{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}}\right)^{k-1} \to 0 \quad \text{as } k \to \infty$$

and hence, for $r = 1, 2, \cdots$,

$$\lim_{k \to \infty} \phi_k(r, \mathcal{T}) = \begin{cases} 0 & \text{if } r \leq U - 1 \\ 1 & \text{if } r > U \end{cases}$$

and Step 1. is proved.

Step 2.

We have that

$$E\left(\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k}\right) = E\left(\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} I(r \le U-1)\right) + E\left(\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} I(r \ge U)\right)$$

$$= P(r \le U-1) + E\left(E\left(\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} I(r \ge U)\middle| |\mathbf{Z}_r|\right)\right)$$

Since $\{W_{r,i} : i \ge 1\}$ are i.i.d.,

$$E\left(E\left(\frac{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}^k}{\left(\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}\right)^k} I(r \ge U)\middle| |\mathbf{Z}_r|\right)\right) = E\left(|\mathbf{Z}_r| \cdot E\left(\left(\frac{W_{r,1}}{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}} I(r \ge U)\right)^k \middle| |\mathbf{Z}_r|\right)\right)$$

Also, $P\left(0 < \frac{W_{r,i}}{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}} < 1\right) = 1$ implies that $\left(\frac{W_{r,i}}{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}} I(r \ge U)\right)^k \to 0$ w.p.1 as $k \to \infty$, and hence

$$E\left(|\mathbf{Z}_r| \cdot E\left(\left(\frac{W_{r,1}}{\sum_{i=1}^{|\mathbf{Z}_r|} W_{r,i}} I(r \ge U)\right)^k \middle| |\mathbf{Z}_r|\right)\right) \to 0 \quad \text{as } k \to \infty$$

by the bounded convergence theorem.

Therefore,

$$P(\tilde{X}_k < r) = \phi_k(r) = E(\phi_k(r, \mathcal{T})) \to 1 - P(r \le U-1) = P(U-1 < r)$$

for any $r = 1, 2, \cdots$. So, $\tilde{X}_k \xrightarrow{d} U - 1$ as $k \to \infty$ and the proof is complete.

3.2.4 The Proof of Theorem 3.3

Let $\xi_{n,j}^{(i)} = (\xi_{n,j}^{(i)1}, \xi_{n,j}^{(i)2}, \cdots, \xi_{n,j}^{(i)d})$ be the vector of the offsprings of the jth individual of the type i in the nth generation. Let $\{\mathbf{Z}_{p,r,s,n}^{j(l)}\}_{n\ge0}$ be the multi-type Galton-Watson branching process initiated by the sth child of type l of the pth individual of type j in the rth generation. So,

$$\left\{\mathbf{Z}_{p,r,s,n}^{j(l)} = (Z_{p,r,s,n}^{j(l)1}, Z_{p,r,s,n}^{j(l)2}, \cdots, Z_{p,r,s,n}^{j(l)d})\right\}_{n\ge0}$$

has the same distribution as $\{\mathbf{Z}_n|\mathbf{Z}_0 = \mathbf{e}_l\}$ does.

Let $A_{n,i}$ be the type of the ancestor in the next generation after the last common ancestor of the ith chosen individual, $i = 1, 2$. Then

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2}) \\
= \;& E\Big(P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2}|\mathcal{T})\Big) \\
= \;& E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} Z_{p,r,s,n-r-1}^{j(l)i} Z_{p,r,t,n-r-1}^{j(l)i}}{|\mathbf{Z}_n|(|\mathbf{Z}_n| - 1)}\right) \\
= \;& E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} \frac{Z_{p,r,s,n-r-1}^{j(l)i}}{\rho^{n-r-1}} \frac{Z_{p,r,t,n-r-1}^{j(l)i}}{\rho^{n-r-1}}}{\frac{|\mathbf{Z}_n|}{\rho^{n-r-1}} \frac{|\mathbf{Z}_n|-1}{\rho^{n-r-1}}}\right) \\
\longrightarrow \;& E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} (v_i W_{p,r,s})(v_i W_{p,r,t})}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+n}|} W_{r+1,s}\right)^2}\right) = E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} v_i^2 W_{p,r,s} W_{p,r,t}}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+n}|} W_{r+1,s}\right)^2}\right)
\end{aligned}$$

where $\{W_{p,r,s}\}_{s\geq 1}$ and $\{W_{r+1,s}\}_{s\geq 1}$ are i.i.d. copies of the random variable W with $\lim_{n\to\infty} \frac{\mathbf{Z_n}}{\rho^n} = \mathbf{v}W$.

Similarly, we have

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} \neq A_{n,2}) \\
= \;& E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l\neq=1}^{d} \sum_{s=1}^{\xi_{r,p}^{(j)l}} \sum_{t=1}^{\xi_{r,p}^{(j)q}} \frac{Z_{p,r,s,n-r-1}^{j(l)i}}{\rho^{n-r-1}} \frac{Z_{p,r,t,n-r-1}^{j(l)i}}{\rho^{n-r-1}}}{\frac{|\mathbf{Z}_n|}{\rho^{n-r-1}} \frac{|\mathbf{Z}_n|-1}{\rho^{n-r-1}}}\right) \longrightarrow E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l\neq q=1}^{d} \sum_{s=1}^{\xi_{r,p}^{(j)l}} \sum_{t=1}^{\xi_{r,p}^{(j)q}} v_i^2 W_{p,r,s} W_{p,r,t}}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+n}|} W_{r+1,s}\right)^2}\right)
\end{aligned}$$

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, i_1 = \zeta_{n,1} \neq \zeta_{n,2} = i_2, A_{n,1} = A_{n,2}) \\
= \;& E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} \frac{Z_{p,r,s,n-r-1}^{j(l)i_1}}{\rho^{n-r-1}} \frac{Z_{p,r,t,n-r-1}^{j(l)i_2}}{\rho^{n-r-1}}}{\frac{|\mathbf{Z}_n|}{\rho^{n-r-1}} \frac{|\mathbf{Z}_n|-1}{\rho^{n-r-1}}}\right) \\
\longrightarrow \;& E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} (v_{i_1} W_{p,r,s})(v_{i_2} W_{p,r,t})}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+n}|} W_{r+1,s}\right)^2}\right) = E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^{d} \sum_{s\neq t=1}^{\xi_{r,p}^{(j)l}} v_{i_1} v_{i_2} W_{p,r,s} W_{p,r,t}}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+n}|} W_{r+1,s}\right)^2}\right)
\end{aligned}$$

and

$$P(X_{n,2}=r,\eta_n=j,i_1=\zeta_{n,1}\neq\zeta_{n,2}=i_2,A_{n,1}\neq A_{n,2}) \longrightarrow E\left(\frac{\sum_{p=1}^{Z_r^{(j)}}\sum_{l\neq q=1}^{d}\sum_{s=1}^{\xi_{r,p}^{(j)l}}\sum_{t=1}^{\xi_{r,p}^{(j)q}} v_{i_1}v_{i_2}W_{p,r,s}W_{p,r,t}}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+n}|}W_{r+1,s}\right)^2}\right)$$

Therefore, as $n\longrightarrow\infty$,

$$P(X_{n,2}=r,\eta_n=j,\zeta_{n,1}=i_1,\zeta_{n,2}=i_2) \longrightarrow v_{i_1}v_{i_2}E\left(\frac{\sum_{p=1}^{Z_r^{(j)}}\sum_{s\neq t=1}^{|\xi_{r,p}^{(j)}|}W_{p,r,s}W_{p,r,t}}{\left(\sum_{s=1}^{|\mathbf{Z}_{r+1}|}W_{r+1,s}\right)^2}\right)\equiv\varphi_2(r,j,i_1,i_2).$$

By Theorem.3.1, we know that $X_{n,2}\xrightarrow{d}\tilde{X}_2$ and then $\{X_{n,2}\}_{n\geq 0}$ is tight. Also, η_n, $\zeta_{n,1}$ and $\zeta_{n,2}$ are random variables taking values on a finite set $\{1,2,\cdots,\alpha\}$. Hence, $\{(X_{n,2},\eta_n,\zeta_{n,1},\zeta_{n,2})\}_{n\geq 0}$ is tight and the limit $\varphi_2(r,j,i_1,i_2)$ of $P(X_{n,2}=r,\eta_n=j,\zeta_{n,1}=i_1,\zeta_{n,2}=i_2)$ is a probability distribution. Thus, $\sum_{(r,j,i_1,i_2)}\varphi_2(r,j,i_1,i_2)=1$ and the proof is complete.

3.3 Results in The Critical Case

Now, we consider a discrete-time multi-type critical Galton-Watson branching process.

We begin with Theorem 3.5 which shows the convergence of some point process constructed from the original branching process and, by using it, we are able to prove the results on the coalescence problem for the multi-type critical branching process.

3.3.1 The statements of Results

For any $t<n$, let $\left\{\mathbf{Z}_{t,i,n-t}^{(l)}=(Z_{t,i,n-t}^{(l)1},Z_{t,i,n-t}^{(l)2},\cdots,Z_{t,i,n-t}^{(l)d})\right\}_{n\geq t}$ be the branching process initiated by the ith individual of type l in the tth generation and let $J_t^{(l)}$ be the set of all $i\in\{1,2,\cdots,Z_t^{(l)}\}$ such that $|\mathbf{Z}_{t,i,n-t}^{(l)}|>0$, $l=1,2,\cdots,d$.

Theorem 3.5. *Let $\rho=1$ and $E\|\mathbf{Z}_1\|^2<\infty$. On the event $A_n\equiv\{|\mathbf{Z}_n|>0\}$, for $t<n$, consider the random point process*

$$V_n\equiv\left\{\left.\frac{\mathbf{Z}_{t,i,n-t}^{(l)}}{n-t}\;\right|\; i\in J_t^{(l)},l=1,2,\cdots,d\right\}.$$

Let $n \longrightarrow \infty$, $t \longrightarrow \infty$ and $\frac{t}{n} \longrightarrow \alpha$ for $\alpha \in (0,1)$, then, conditioned on A_n, the distribution of the random point process V_n converges to a random point process $V \equiv \{ \mathbf{Y}_i \mid 1 \le i \le N_\alpha \}$ where $\{ \mathbf{Y}_i = (v_1 Y_i, v_2 Y_i, \cdots, v_d Y_i) \}_{i \ge 1}$ are i.i.d. random vectors with $Y_i \sim \exp\left(\frac{1}{\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}]}\right)$, N_α is a random variable independent of $\{\mathbf{Y}_i\}_{i\ge 1}$ with distribution $P(N_\alpha = j) = (1-\alpha)\alpha^{j-1}$ for $j \ge 1$ and $\mathbb{Q}$ is the quadratic form as defined in (1.1).

Since the two vectors $\mathbf{u}$ and $\mathbf{v}$, the left and right eigenvectors of the offspring mean matrix $\mathbf{M}$ associated with the maximal eigenvalue ρ, are normalized such that $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{1} = 1$, an analog stated as in the following corollary can be obtained along the lines of the proof of Theorem 3.5.

Corollary 3.1. *Under the same hypotheses of the Theorem3.5, consider the random point process*

$$V_n' \equiv \left\{ \frac{\left|\mathbf{Z}_{t,i,n-t}^{(l)}\right|}{n-t} \;\middle|\; i \in J_t^{(l)}, l = 1, 2, \cdots, d \right\}.$$

Let $n \longrightarrow \infty$, $t \longrightarrow \infty$ and $\frac{t}{n} \longrightarrow \alpha$ for $\alpha \in (0,1)$, then, conditioned on A_n, the distribution of the random point process V_n' converges to a random point process $V' \equiv \{ Y_i \mid 1 \le i \le N_\alpha \}$ where $\{ Y_i \}_{i\ge1}$ are i.i.d. exponential random variables with mean $\frac{1}{\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}]}$ and N_a is a random variable independent of $\{Y_i\}_{i\ge1}$ with distribution $P(N_\alpha = j) = (1-\alpha)\alpha^{j-1}$ for $j \ge 1$.

Next, we move on to the coalescence problem on critical branching process. Let $k \ge 2$ be an integer. Pick k individuals at random from the nth generation (by simple random sampling without replacement) and trace their lines of decent backward in time to find their last common ancestor. Let $X_{n,k}$ be the generation number of this common ancestor.

Theorem 3.6. *Let $\rho = 1$ and $E\|\mathbf{Z}_1\|^2 < \infty$. Then, for $k = 2, 3, \cdots$, there exists a random variable $\tilde{X}_k$ such that $\frac{X_{n,k}}{n}\Big| |\mathbf{Z}_n| \ge k \xrightarrow{d} \tilde{X}_k$ as $n \to \infty$ and, for any $\alpha \in (0,1)$,*

$$P(\tilde{X}_k < \alpha) = 1 - E(\phi_k(N_\alpha)) \equiv H_k(\alpha)$$

where $\phi_k(x) = E\left(\frac{\sum_{i=1}^{x} Y_i^k}{\left(\sum_{i=1}^{x} Y_i\right)^k}\right)$, $\{Y_i\}_{i\ge1}$ and N_α are as defined in Theorem 3.5.

Theorem 3.6 tells us that the generation number of the last common ancestor of any finite number of individuals randomly chosen from the population of the nth generation grows like n. That is, the coalescence time X_n is not close either to the beginning of the tree or the present when n gets large. This result is consistent with what we have seen in the discrete-time single-type Galton-Watson branching process.

Now, we trace the lines of descent of all the individuals in the nth generation backward in time till they meet. Let T_n be the coalescence of whole population of the nth generation (we also call T_n *the total coalescence time* of all the individuals in the nth generation). The asymptotic behavior of the total coalescence time T_n in the multi-type critical branching process offers no new surprises. As in the one-dimensional case, we condition on non-extinction and normalized by dividing it by the generation number, the limit distribution again is uniform in $(0, 1)$.

Theorem 3.7. *Let $\rho = 1$ and $E\|\mathbf{Z}_1\|^2 < \infty$. Then there exists a random variable $\tilde{T}$ such that $\frac{T_n}{n}\Big||\mathbf{Z}_n| > 0 \xrightarrow{d} \tilde{T}$ as $n \to \infty$, where $\tilde{T}$ has a uniform distribution in $(0, 1)$.*

3.3.2 The Proof of Theorem 3.5

To prove the convergene of the random point process $\{V_n\}\}_{n\geq n}$. we first consider the Laplace functional of this process V_n

$$\varphi_n(\theta_1, \theta_2, \cdots, \theta_d, f_1, f_2, \cdots, f_d) \equiv E\left(e^{-\sum\limits_{l=1}^{k}\sum\limits_{i\in J_t^{(l)}}\sum\limits_{p=1}^{d}\theta_p f_p(\frac{Z_{t,i,n-t}^{(l)p}}{n-t})}\middle||\mathbf{Z}_n| > 0, \mathbf{Z}_0 = \mathbf{e}_{i_0}\right)$$

where $\theta_1, \theta_2, \cdots, \theta_d > 0$ and $f_1, f_2, \cdots, f_d : \mathbb{R}^+ \to \mathbb{R}^+$ are bounded and continuous functions.

Let $Y_{n,t} = e^{-\sum\limits_{l=1}^{k}\sum\limits_{i\in J_t^{(l)}}\sum\limits_{p=1}^{d}\theta_p f_p(\frac{Z_{t,i,n-t}^{(l)p}}{n-t})}$.

Then

$$\begin{aligned} P(|\mathbf{Z}_n| > 0\big|\mathbf{Z}_0 = \mathbf{e}_{i_0})E(Y_{n,t}\big||\mathbf{Z}_n| > 0, \mathbf{Z}_0 = \mathbf{e}_{i_0}) &= E(Y_{n,t}I_{\{|\mathbf{Z}_n|>0\}}\big|\mathbf{Z}_0 = \mathbf{e}_{i_0}) \\ &= E(E(Y_{n,t}I_{\{|\mathbf{Z}_n|>0\}}|\mathbf{Z}_j, j \leq t)\big|\mathbf{Z}_0 = \mathbf{e}_{i_0}) \end{aligned}$$

By the Markov property,

$$\begin{aligned}
& E(Y_{n,t} I_{\{|\mathbf{Z}_n|>0\}} \big| \mathbf{Z}_j, j \le t) \\
= \; & E(Y_{n,t} I_{\{|\mathbf{Z}_n|>0\}} \big| \mathbf{Z}_t) \\
= \; & E(Y_{n,t} I_{\{|\mathbf{Z}_n|>0\}} I_{\{|\mathbf{Z}_t|>0\}} \big| \mathbf{Z}_t) \\
= \; & E(Y_{n,t} \big| \mathbf{Z}_t) I_{\{|\mathbf{Z}_n|>0\}} - E(Y_{n,t} I_{\{|\mathbf{Z}|_n=0\}} \big| \mathbf{Z}_t) I_{\{|\mathbf{Z}_t|>0\}} \\
= \; & (g_{n-t}^{(1)})^{Z_t^{(1)}} (g_{n-t}^{(2)})^{Z_t^{(2)}} \cdots (g_{n-t}^{(d)})^{Z_t^{(d)}} I_{|\mathbf{Z}_t|>0} - (q_{n-t}^{(1)})^{Z_t^{(1)}} (q_{n-t}^{(2)})^{Z_t^{(2)}} \cdots (q_{n-t}^{(d)})^{Z_t^{(d)}} I_{|\mathbf{Z}_t|>0}
\end{aligned}$$

where $g_j^{(l)}(\theta) = E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j}) I_{\{|\mathbf{Z}_j|>0\}}} \Big| \mathbf{Z}_0 = \mathbf{e}_l\Big)$ and $q_j^{(l)}(d) = P(|\mathbf{Z}_j| = 0 | \mathbf{Z}_0 = \mathbf{e}_l)$ for $j \ge 1$.

Note that

$$\begin{aligned}
& g_j^{(l)}(\theta) \\
= \; & E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j}) I_{\{|\mathbf{Z}_j|>0\}}} \Big| \mathbf{Z}_0 = \mathbf{e}_l\Big) \\
= \; & E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j}) I_{\{|\mathbf{Z}_j|>0\}}} \Big| |\mathbf{Z}_j| = 0, \mathbf{Z}_0 = \mathbf{e}_l\Big) P(|\mathbf{Z}_j| = 0 \big| \mathbf{Z}_0 = \mathbf{e}_l) \\
& \quad + E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j}) I_{\{|\mathbf{Z}_j|>0\}}} \Big| |\mathbf{Z}_j| > 0, \mathbf{Z}_0 = \mathbf{e}_l\Big) P(|\mathbf{Z}_j| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l) \\
= \; & P(|\mathbf{Z}_j| = 0 \big| \mathbf{Z}_0 = \mathbf{e}_l) + E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j}) I_{\{|\mathbf{Z}_j|>0\}}} \Big| |\mathbf{Z}_j| > 0, \mathbf{Z}_0 = \mathbf{e}_l\Big) P(|\mathbf{Z}_j| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l) \\
= \; & q_j^{(l)} + (1 - q_j^{(l)}) E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j}) I_{\{|\mathbf{Z}_j|>0\}}} \Big| |\mathbf{Z}_j| > 0, \mathbf{Z}_0 = \mathbf{e}_l\Big).
\end{aligned}$$

Let $\tilde{g}_j^{(l)}(d) = E\Big(e^{-\sum_{p=1}^{d} \theta_p f_p(\frac{Z_j^{(p)}}{j})} \Big| |\mathbf{Z}_j| > 0, \mathbf{Z}_0 = \mathbf{e}_l\Big)$.

It is known that in the critical case, i.e., $\rho = 1$, if $E\|\mathbf{Z}_1\|^2 < \infty$, then, as $j \to \infty$,

$$j(1 - q_j^{(l)}) = jP(|\mathbf{Z}_j| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l) \to \frac{u_l}{v \cdot \mathbb{Q}[\mathbf{u}]}$$

and

$$\frac{\mathbf{Z}_j}{j} \Big| |\mathbf{Z}_j| > 0, \mathbf{Z}_0 = \mathbf{e}_l \xrightarrow{d} \mathbf{v} Y$$

where $Y \sim \exp\Big(\frac{1}{\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}]}\Big)$. Since $f_1, f_2, \cdots, f_d$ are bounded and continuous, as $j \to \infty$, we have

$$\tilde{g}_j^{(l)}(\theta) \quad \to \quad E\Big(e^{-\sum_{p=1}^{\alpha} \theta_p f_p(v_p Y)}\Big) \equiv g(\theta) = \frac{1}{\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}]} \int_0^\infty e^{-\sum_{p=1}^{\alpha} \theta_p f_p(v_p y)} e^{-\frac{y}{\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}]}} dy.$$

Also, we have that

$$g_j^{(l)}(\theta) = q_j^{(l)} + (1 - q_j^{(l)})\tilde{g}_j^{(l)}(\theta) = 1 + (1 - q_j^{(l)})(\tilde{g}_j^{(l)}(\theta) - 1)$$

and so, as $j \to \infty$,

$$(g_j^{(l)}(\theta))^j = \left(1 + \frac{j(1 - q_j^{(l)})(\tilde{g}_j^{(l)}(\theta) - 1)}{j}\right)^j \to e^{\frac{u_l}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)}.$$

Now, consider the quantity

$$\begin{aligned}
& \frac{E\left((g_{n-t}^{(1)}(\theta))^{Z_t^{(1)}}(g_{n-t}^{(2)}(\theta))^{Z_t^{(2)}}\cdots(g_{n-t}^{(d)}(\theta))^{Z_t^{(d)}}\middle|\mathbf{Z}_0 = \mathbf{e}_{i_0}\right)}{P(|\mathbf{Z}_t| > 0|\mathbf{Z}_0 = \mathbf{e}_{i_0})} \\
= & \frac{E\left(\left((g_{n-t}^{(1)}(\theta))^{n-t}\right)^{\frac{t}{n-t}\frac{Z_t^{(1)}}{t}}\left((g_{n-t}^{(2)}(\theta))^{n-t}\right)^{\frac{t}{n-t}\frac{Z_t^{(2)}}{t}}\cdots\left((g_{n-t}^{(d)}(\theta))^{n-t}\right)^{\frac{t}{n-t}\frac{Z_t^{(d)}}{t}}\middle|\mathbf{Z}_0 = \mathbf{e}_{i_0}\right)}{P(|\mathbf{Z}_t| > 0|\mathbf{Z}_0 = \mathbf{e}_{i_0})}
\end{aligned}$$

and, if $n \to \infty$, $t \to \infty$ and $\frac{t}{n} \to \alpha$, $0 < \alpha < 1$, then it converges to

$$\begin{aligned}
& E\left(e^{\frac{u_1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)\frac{\alpha}{1-\alpha}v_1 Y}e^{\frac{u_2}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)\frac{\alpha}{1-\alpha}v_2 Y}\cdots e^{\frac{u_d}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)\frac{\alpha}{1-\alpha}v_d Y}\right) \\
= & E\left(e^{\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)\frac{\alpha}{1-\alpha}Y}\right) \\
= & E\left(e^{\frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)\frac{\alpha}{1-\alpha}Y}\right) \\
= & \frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\int_0^\infty e^{\frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}(g(\theta)-1)\frac{\alpha}{1-\alpha}y}e^{-\frac{y}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}}dy \\
= & \frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\int_0^\infty e^{-\frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\left(1-(g(\theta)-1)\frac{\alpha}{1-\alpha}\right)y}dy \\
= & \frac{1}{1 - (g(\theta) - 1)\frac{\alpha}{1-\alpha}} \\
= & \frac{1 - \alpha}{1 - \alpha g(\theta)}.
\end{aligned}$$

On the other hand, consider the following

$$\begin{aligned}
& \frac{E\left((q_{n-t}^{(1)}(\theta))^{Z_t^{(1)}}(q_{n-t}^{(2)}(\theta))^{Z_t^{(2)}}\cdots(q_{n-t}^{(\alpha)}(d))^{Z_t^{(d)}}\middle|\mathbf{Z}_0 = \mathbf{e}_{i_0}\right)}{P(|\mathbf{Z}_t| > 0|\mathbf{Z}_0 = \mathbf{e}_{i_0})} \\
= & \frac{E\left(\left((1 - (1 - q_{n-t}^{(1)}))^{n-t}\right)^{\frac{t}{n-t}\frac{Z_t^{(1)}}{t}}\left((1 - (1 - q_{n-t}^{(2)}))^{n-t}\right)^{\frac{t}{n-t}\frac{Z_t^{(2)}}{t}}\cdots\left((1 - (1 - q_{n-t}^{(d)}))^{n-t}\right)^{\frac{t}{n-t}\frac{Z_t^{(d)}}{t}}\middle|\mathbf{Z}_0 = \mathbf{e}_{i_0}\right)}{P(|\mathbf{Z}_t| > 0|\mathbf{Z}_0 = \mathbf{e}_{i_0})}
\end{aligned}$$

and if $n \to \infty$, $t \to \infty$ and $\dfrac{t}{n} \to \alpha$, $0 < \alpha < 1$, then it converges to

$$\begin{aligned}
& E\Big(e^{-\frac{u_1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\frac{\alpha}{1-\alpha}v_1 Y} e^{-\frac{u_2}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\frac{\alpha}{1-\alpha}v_2 Y}\cdots e^{-\frac{u_d}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\frac{\alpha}{1-\alpha}v_d Y}\Big)\\
= \; & E\Big(e^{\frac{-\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\frac{\alpha}{1-\alpha}Y}\Big)\\
= \; & E\Big(e^{-\frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\frac{\alpha}{1-\alpha}Y}\Big)\\
= \; & \frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\int_0^\infty e^{-\frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\frac{\alpha}{1-\alpha}y} e^{-\frac{y}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}}dy\\
= \; & \frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\int_0^\infty e^{-\frac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\left(\frac{\alpha}{1-\alpha}+1\right)y}dy\\
= \; & \frac{1}{\frac{\alpha}{1-\alpha}+1}\\
= \; & 1-\alpha.
\end{aligned}$$

Moreover, by Theorem 1.8 (a), we know that

$$\frac{P(|\mathbf{Z}_t|>0|\mathbf{Z}_0=\mathbf{e}_{i_0})}{P(|\mathbf{Z}_n|>0|\mathbf{Z}_0=\mathbf{e}_{i_0})} = \frac{tP(|\mathbf{Z}_t|>0|\mathbf{Z}_0=\mathbf{e}_{i_0})}{nP(|\mathbf{Z}_n|>0|\mathbf{Z}_0=\mathbf{e}_{i_0})}\frac{n}{t} \to \frac{\frac{u_{i_0}}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}}{\frac{u_{i_0}}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}}\frac{1}{\alpha} = \frac{1}{\alpha}$$

as $t, n \to \infty$.

Hence,

$$\begin{aligned}
& \varphi_n(\theta_1,\theta_2,\cdots,\theta_d,f_1,f_2,\cdots,f_d)\\
= \; & E(Y_{n,t}\big||\mathbf{Z}_n|>0,\mathbf{Z}_0=\mathbf{e}_{i_0})\\
= \; & \frac{P(|\mathbf{Z}_t|>0|\mathbf{Z}_0=\mathbf{e}_{i_0})}{P(|\mathbf{Z}_n|>0|\mathbf{Z}_0=\mathbf{e}_{i_0})}\Bigg(\frac{E\Big((g^{(1)}_{n-t}(\theta))^{Z^{(1)}_t}(g^{(2)}_{n-t}(\theta))^{Z^{(2)}_t}\cdots(g^{(d)}_{n-t}(\theta))^{Z^{(d)}_t}\Big|\mathbf{Z}_0=\mathbf{e}_{i_0}\Big)}{P(|\mathbf{Z}_t|>0\big|\mathbf{Z}_0=\mathbf{e}_{i_0})}\\
& -\frac{E\Big((q^{(1)}_{n-t}(\theta))^{Z^{(1)}_t}(q^{(2)}_{n-t}(\theta))^{Z^{(2)}_t}\cdots(q^{(d)}_{n-t}(\theta))^{Z^{(d)}_t}\Big|\mathbf{Z}_0=\mathbf{e}_{i_0}\Big)}{P(|\mathbf{Z}_t|>0\big|\mathbf{Z}_0=\mathbf{e}_{i_0})}\Bigg)\\
\to \; & \frac{1}{\alpha}\Big(\frac{1-\alpha}{1-\alpha g(\theta)}-(1-\alpha)\Big)\\
= \; & \frac{(1-\alpha)g(\theta)}{1-\alpha g(\theta)}\\
= \; & \sum_{j=0}^{\infty}(1-\alpha)\alpha^j(g(\theta))^{j+1}\\
= \; & \sum_{j=1}^{\infty}(1-\alpha)\alpha^{j-1}(g(\theta))^j.
\end{aligned}$$

Let $V \equiv \{\, \mathbf{Y}_i \mid 1 \le i \le N_\alpha \,\}$ where $\{\, \mathbf{Y}_i = (v_1 Y_i, v_2 Y_i, \cdots, v_d Y_i) \,\}_{i\ge1}$ are i.i.d. random vectors with $Y_i \sim \exp\Big(\dfrac{1}{\mathbf{v}\cdot\mathbb{Q}[\mathbf{u}]}\Big)$ and N_α is a random variable independent of $\{\mathbf{Y}_i\}_{i\ge1}$ with distribution $P(N_\alpha = j) =$

$(1-\alpha)\alpha^{j-1}$ for $j \geq 1$. Then, for any $\theta_1, \theta_2, \cdots, \theta_d > 0$ and any bounded, nonnegative and continuous functions $f_1, f_2, \cdots, f_d$, the Laplace functional of V is

$$E\Big(e^{-\sum_{i=1}^{N_\alpha}\sum_{p=1}^{d}\theta_p f_p(Y_i^{(p)})}\Big) = \sum_{j=1}^{\infty}(1-\alpha)\alpha^{j-1}(g(\theta))^j.$$

Therefore, for any $a \in (0,1)$, by the continuous mapping theorem for random measures (see Kallenberg [25]), the sequence of random point processes $V_n \equiv \left\{ \frac{\mathbf{Z}^l_{t,i,n-t}}{n-t} \;\middle|\; i \in J_t^{(l)}, l = 1, 2, \cdots, d \right\}$, $n \geq 1$, conditioned on $\{|\mathbf{Z}_n| > 0, \mathbf{Z}_0 = \mathbf{e}_{i_0}\}$ converges in distribution to the random point process $V \equiv \{ \mathbf{Y}_i \mid 1 \leq i \leq N_\alpha \}$ as $n, t \to \infty$, $\frac{t}{n} \to \alpha$. The proof is complete.

3.3.3 The Proof of Theorem 3.6

Now, we are going to prove the convergence in distribution of the generation number $X_{n,k}$ of the last common ancestor of k individuals randomly chosen from the population in the nth generation.

First, conditioned on the set $\{|\mathbf{Z}_n| \geq 2\}$, for almost all trees $\mathcal{T}$ and any integer $r = 1, 2, \cdots$, we have that

$$P(X_{n,k} \geq r|\mathcal{T}) = \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}^{(l)}_{r,i,n-r}|(|\mathbf{Z}^{(l)}_{r,i,n-r}|-1)\cdots(|\mathbf{Z}^{(l)}_{r,i,n-r}|-k+1)}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)}$$

Hence, for any $\alpha \in (0,1)$, let $r = [n\alpha] + 1$, then

$$\begin{aligned}
& P\Big(\frac{X_{n,k}}{n} < \alpha \Big| |\mathbf{Z}_n| \geq 2\Big) \\
= \;& P(X_{n,k} < n\alpha \big| |\mathbf{Z}_n| \geq 2) \\
= \;& 1 - P(X_{n,k} \geq r \big| |\mathbf{Z}_n| \geq 2) \\
= \;& 1 - E\left(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}^{(l)}_{r,i,n-r}|(|\mathbf{Z}^{(l)}_{r,i,n-r}|-1)\cdots(|\mathbf{Z}^{(l)}_{r,i,n-r}|-k+1)}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)}\middle| |\mathbf{Z}_n| \geq 2\right) \\
= \;& 1 - \frac{1}{P(|\mathbf{Z}_n| \geq 2 | |\mathbf{Z}_n| > 0)} \\
& \cdot E\left(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^k + \sum_{s=1}^{k-1}\Big((-1)^s\big(\sum_{1\leq q_1<q_2<\cdots<q_s\leq k-1} q_1 q_2\cdots q_s\big)\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^{k-s}\Big)}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} I_{\{|\mathbf{Z}_n|\geq 2\}}\middle| |\mathbf{Z}_n| > 0\right)
\end{aligned}$$

$$= \quad 1 - \frac{1}{P(|\mathbf{Z}_n| \geq 2 \| \mathbf{Z}_n| > 0)} E\Bigg(\Bigg(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}^{(l)}|^k}{|\mathbf{Z}_n|^k} + \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^k}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} - \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}^{(l)}|^k}{|\mathbf{Z}_n|^k}$$

$$+ \frac{\sum_{s=1}^{k-1}\Big((-1)^s(\sum_{1\leq q_1<q_2<\cdots<q_s\leq k-1} q_1 q_2\cdots q_s)\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^{k-s}\Big)}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)}\Bigg) I_{\{|\mathbf{Z}_n|\geq 2\}}\Bigg| |\mathbf{Z}_n| > 0\Bigg)$$

$$= \quad 1 - \frac{1}{P(|\mathbf{Z}_n| \geq 2 \| \mathbf{Z}_n| > 0)} E\Bigg(\frac{\sum_{l=1}^{d}\sum_{i\in J_r^{(l)}} \Big(\frac{|\mathbf{Z}_{r,i,n-r}|}{n-r}\Big)^k}{\Big(\sum_{l=1}^{d}\sum_{i\in J_r^{(l)}} \frac{|\mathbf{Z}_{r,i,n-r}|}{n-r}\Big)^k} I_{\{|\mathbf{Z}_n|\geq 2\}}\Bigg| |\mathbf{Z}_n| > 0\Bigg)$$

$$+ \frac{1}{P(|\mathbf{Z}_n| \geq 2 \| \mathbf{Z}_n| > 0)} E\Bigg(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^k}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} - \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}^{(l)}|^k}{|\mathbf{Z}_n|^k} I_{\{|\mathbf{Z}_n|\geq 2\}}\Bigg| |\mathbf{Z}_n| > 0\Bigg)$$

$$+ \frac{1}{P(|\mathbf{Z}_n| \geq 2 \| \mathbf{Z}_n| > 0)} \sum_{s=1}^{k-1}\Bigg((-1)^s(\sum_{1\leq q_1<q_2<\cdots<q_s\leq k-1} q_1 q_2\cdots q_s) E\Bigg(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^{k-s}}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} I_{\{|\mathbf{Z}_n|\geq 2\}}\Bigg| |\mathbf{Z}_n| > 0\Bigg)\Bigg)$$

Since, conditioned on $\{|\mathbf{Z}_n| > 0\}$, we have that

$$0 \leq \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}^{(l)}|^k}{|\mathbf{Z}_n|^k} \leq \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^k}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} \leq \frac{|\mathbf{Z}_n|^k}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} \leq 1$$

and

$$0 \leq \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^{k-s}}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} \leq \frac{|\mathbf{Z}_n|^{k-s}}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} \leq 1, \text{ for } s = 1, 2, \cdots, k-1,$$

thus, by the bounded convergence theorem, as $n \to \infty$,

$$E\Bigg(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^k}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} - \frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}^{(l)}|^k}{|\mathbf{Z}_n|^k} I_{\{|\mathbf{Z}_n|\geq 2\}}\Bigg| |\mathbf{Z}_n| > 0\Bigg) \to 0$$

and

$$E\Bigg(\frac{\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}|^{k-s}}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)\cdots(|\mathbf{Z}_n|-k+1)} I_{\{|\mathbf{Z}_n|\geq 2\}}\Bigg| |\mathbf{Z}_n| > 0\Bigg) \to 0 \text{ for } s = 1, 2, \cdots, k.$$

It is also known that, in the critical case, $\frac{\mathbf{Z}_n}{n}\Big| |\mathbf{Z}| > 0 \xrightarrow{d} \mathbf{v}Y$ and Y is exponentially distributed with parameter $\frac{1}{\mathbf{v}\cdot\mathbf{Q}[\mathbf{u}]}$, so

$$P(|\mathbf{Z}_n| \geq 2 \| \mathbf{Z}_n| > 0) \to 1 \qquad \text{as} \qquad n \to \infty.$$

Therefore, by the continuous mapping theorem and Theorem 3.5,

$$P\Big(\frac{X_{n,k}}{n} < \alpha \Big| |\mathbf{Z}_n| > 0\Big) \to 1 - E\Bigg(\frac{\sum_{i=1}^{N_\alpha} Y_i^k}{(\sum_{i=1}^{N_\alpha} Y_i)^k}\Bigg) \equiv H_k(\alpha) \qquad \text{as } n \to \infty.$$

Let $\phi_k(x) = E\Bigg(\frac{\sum_{i=1}^{x} Y_i^k}{(\sum_{i=1}^{x} Y_i)^k}\Bigg)$. since $EY_1 < \infty$, we have that $\phi_k(x) \to 0$ as $x \to \infty$. Also,

$$\lim_{\alpha\to 1} P(N_\alpha = x) = \lim_{\alpha\to 1}(1-\alpha)\alpha^{x-1} = 0 \qquad \text{for any} x \geq 0.$$

So, $N_\alpha \to \infty$ as $\alpha \to 1$. By the Bounded Convergence Theorem again,

$$E(\phi_k(N_\alpha)) = E\Bigg(\frac{\sum_{i=1}^{N_\alpha} Y_i^k}{(\sum_{i=1}^{N_\alpha} Y_i)^k}\Bigg) \downarrow 0 \qquad \text{as } \alpha \to 1.$$

and hence $H_k(\alpha) = 1 - E(\phi_k(N_\alpha)) \uparrow 1$ as $\alpha \to 1$. Moreover, $H_k(0) = 0$. Therefore, H_k is a proper probability distribution and hence there exists a random variable $\tilde{X}_k$ with $P(\tilde{X}_k \leq \alpha) = H_k(\alpha)$ for $a \in (0, 1)$ such that

$$\frac{X_{n,k}}{n} \Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} \tilde{X}_k \qquad \text{as } n \to \infty.$$

We complete the proof of Theorem 3.6.

3.3.4 The Proof of Theorem 3.7

At the end of this section, we will prove the convergence in distribution of the total coalescence time T_n normalized by dividing by the generation number n as $n \to \infty$.

For any $\alpha \in (0, 1)$ and any $n \in \mathbb{N}$, let $r = [n\alpha] + 1$.

Let $\mathbf{Z}^{(l)}_{r,i,n-r}$ be the d-type Galton-Watson branching process initiated by the ith individual of type l in the rth generation, where $i = 1, 2, \cdots, Z^{(l)}_r$ and $l = 1, 2, \cdots, d$.

The event $\{T_n \geq r\}$ for $1 \leq r \leq n$ conditioned on $\{|\mathbf{Z}_n| > 0\}$ occurs if and only if all the individuals in the nt generation come from the $(n-r)$th generation of the tree initiated by exactly one individual in the rth generation. That is, $|\mathbf{Z}^{(l)}_{r,i,n-r}| = 0$ for all but one $l = 1, 2, \cdots, d$ and one $i = 1, 2, \cdots, Z^{(l)}_r$. Then, for almost all trees $\mathcal{T}$,

$$P(T_n \geq r|\mathcal{T}) = \sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} P(|\mathbf{Z}^{(l)}_{r,i,n-r}| > 0) \cdot \prod_{j\neq i} P(|\mathbf{Z}^{(l)}_{r,j,n-r}| = 0) \cdot \prod_{p\neq l}\prod_{j=1}^{Z_r^{(p)}} P(|\mathbf{Z}^{(p)}_{r,j,n-r}| = 0)$$

Hence,

$$\begin{aligned}
& P\Big(\frac{T_n}{n} > \alpha \Big| |\mathbf{Z}_n| > 0\Big) \\
= \;& P(T_n > n\alpha \,|\, |\mathbf{Z}_n| > 0) \\
= \;& P(T_n \geq r \,|\, |\mathbf{Z}_n| > 0) \\
= \;& E\Big(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} P(|\mathbf{Z}^{(l)}_{r,i,n-r}| > 0) \cdot \prod_{j\neq i} P(|\mathbf{Z}^{(l)}_{r,j,n-r}| = 0) \cdot \prod_{p\neq l}\prod_{j=1}^{Z_r^{(p)}} P(|\mathbf{Z}^{(p)}_{r,j,n-r}| = 0)\Big| |\mathbf{Z}_n| > 0\Big) \\
= \;& \frac{1}{P(|\mathbf{Z}_n| > 0)} E\Big(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} P(|\mathbf{Z}^{(l)}_{r,i,n-r}| > 0) \cdot \prod_{j\neq i} P(|\mathbf{Z}^{(l)}_{r,j,n-r}| = 0) \cdot \prod_{p\neq l}\prod_{j=1}^{Z_r^{(p)}} P(|\mathbf{Z}^{(p)}_{r,j,n-r}| = 0) I_{\{|\mathbf{Z}_n|>0\}}\Big) \\
= \;& \frac{1}{P(|\mathbf{Z}_n| > 0)} E\Big(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} P(|\mathbf{Z}^{(l)}_{r,i,n-r}| > 0) \cdot \prod_{j\neq i} P(|\mathbf{Z}^{(l)}_{r,j,n-r}| = 0) \cdot \prod_{p\neq l}\prod_{j=1}^{Z_r^{(p)}} P(|\mathbf{Z}^{(p)}_{r,j,n-r}| = 0) I_{\{|\mathbf{Z}_n|>0\}} I_{\{|\mathbf{Z}_r|>0\}}\Big) \\
= \;& \frac{1}{P(|\mathbf{Z}_n| > 0)} E\Big(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} P(|\mathbf{Z}^{(l)}_{r,i,n-r}| > 0) \cdot \prod_{j\neq i} P(|\mathbf{Z}^{(l)}_{r,j,n-r}| = 0) \cdot \prod_{p\neq l}\prod_{j=1}^{Z_r^{(p)}} P(|\mathbf{Z}^{(p)}_{r,j,n-r}| = 0) \\
& \qquad \cdot (1 - I_{\{|\mathbf{Z}_n|=0\}}) I_{\{|\mathbf{Z}_r|>0\}}\Big) \\
= \;& \frac{1}{P(|\mathbf{Z}_n| > 0)} E\Big(\sum_{l=1}^{d}\sum_{i=1}^{Z_r^{(l)}} P(|\mathbf{Z}^{(l)}_{r,i,n-r}| > 0) \cdot \prod_{j\neq i} P(|\mathbf{Z}^{(l)}_{r,j,n-r}| = 0) \cdot \prod_{p\neq l}\prod_{j=1}^{Z_r^{(p)}} P(|\mathbf{Z}^{(p)}_{r,j,n-r}| = 0) I_{\{|\mathbf{Z}_r|>0\}}\Big) \\
= \;& \frac{P(|\mathbf{Z}_r| > 0)}{P(|\mathbf{Z}_n| > 0)} E\Big(\sum_{l=1}^{d} Z_r^{(l)} g^{(l)}_{n-r} (1 - g^{(l)}_{n-r})^{Z_r^{(l)}-1} \cdot \prod_{p\neq l} (1 - g^{(p)}_{n-r})^{Z_r^{(p)}} \Big| |\mathbf{Z}_r| > 0\Big) \\
& \qquad \text{where } g_n^{(l)} = P(|\mathbf{Z}_n| > 0 \big| \mathbf{Z_0} = \mathbf{e}_l) \\
= \;& \frac{P(|\mathbf{Z}_r| > 0)}{P(|\mathbf{Z}_n| > 0)} E\Big(\sum_{l=1}^{d} \frac{Z_r^{(l)}}{r} \cdot (n-r) g^{(l)}_{n-r} \cdot \frac{r}{n-r} (1 - \frac{(n-r) g^{(l)}_{n-r}}{n-r})^{(n-r)\frac{Z_r^{(l)}-1}{r}\frac{r}{n-r}} \\
& \qquad \cdot \prod_{p\neq l} (1 - \frac{(n-r) g^{(p)}_{n-r}}{n-r})^{(n-r)\frac{Z_r^{(p)}}{r}\frac{r}{n-r}} \Big| |\mathbf{Z}_r| > 0\Big)
\end{aligned}$$

Let h_n be the function defined by

$$\begin{aligned}
& h_n(x_1, x_2, \cdots, x_d) \\
\equiv \;& \sum_{l=1}^{d} x_l \cdot (n-r) g^{(l)}_{n-r} \cdot \frac{r}{n-r} (1 - \frac{(n-r) g^{(l)}_{n-r}}{n-r})^{(n-r)\left(x_l - \frac{1}{r}\right)\frac{r}{n-r}} \cdot \prod_{p\neq l} (1 - \frac{(n-r) g^{(p)}_{n-r}}{n-r})^{(n-r) x_p \frac{r}{n-r}}.
\end{aligned}$$

Since, as $n \to \infty$,

$$ng_n^{(l)} = nP(|\mathbf{Z}_n| > 0 \big| \mathbf{Z_0} = \mathbf{e}_l) \to \frac{u_l}{\mathbf{v} \cdot \mathbf{Q[u]}}$$

and

$$\frac{r}{n-r} \to \frac{\alpha}{1-\alpha},$$

and hence, as $n \to \infty$,

$$h_n(x_1, x_2, \cdots, x_d) \to \sum_{l=1}^{d} x_l \cdot \frac{u_l}{\mathbf{v} \cdot \mathbf{Q[u]}} \cdot \frac{\alpha}{1-\alpha} e^{-\frac{u_l}{\mathbf{v}\cdot\mathbf{Q[u]}} \cdot x_l \cdot \frac{\alpha}{1-\alpha}} \cdot \prod_{p \neq l} e^{-\frac{u_p}{\mathbf{v}\cdot\mathbf{Q[u]}} \cdot x_p \cdot \frac{\alpha}{1-\alpha}} \equiv h(x_1, x_2, \cdots, x_d).$$

We have that $h_n \to h$ uniformly on any compact set since h_n and h are continuous and bounded. Then, as $n \to \infty$,

$$\begin{aligned}
& E\Big(h_n\big(\frac{Z_r^{(1)}}{r}, \frac{Z_r^{(2)}}{r}, \cdots, \frac{Z_r^{(d)}}{r}\big)\Big| |\mathbf{Z}_n| > 0\Big) \\
\to \quad & E(h(v_1 Y, v_2 Y, \cdots, v_d Y)) \\
= \quad & E\Big(\sum_{l=1}^{d} v_l Y \cdot \frac{u_l}{\mathbf{v} \cdot \mathbf{Q[u]}} \cdot \frac{\alpha}{1-\alpha} e^{-\frac{u_l}{\mathbf{v}\cdot\mathbf{Q[u]}} \cdot v_l Y \cdot \frac{\alpha}{1-\alpha}} \cdot \prod_{p \neq l} e^{-\frac{u_p}{\mathbf{v}\cdot\mathbf{Q[u]}} \cdot v_p Y \cdot \frac{\alpha}{1-\alpha}}\Big) \\
= \quad & E\Big(\frac{Y}{\mathbf{v} \cdot \mathbf{Q[u]}} \cdot \frac{\alpha}{1-\alpha} e^{-\frac{Y}{\mathbf{v}\cdot\mathbf{Q[u]}} \cdot \frac{\alpha}{1-\alpha}}\Big)
\end{aligned}$$

and

$$\frac{P(|\mathbf{Z}_r| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l)}{P(|\mathbf{Z}_n| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l)} = \frac{rP(|\mathbf{Z}_r| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l)}{nP(|\mathbf{Z}_n| > 0 \big| \mathbf{Z}_0 = \mathbf{e}_l)} \cdot \frac{n}{r} \to \frac{\frac{u_l}{\mathbf{v}\cdot\mathbf{Q[u]}}}{\frac{u_l}{\mathbf{v}\cdot\mathbf{Q[u]}}} \cdot \frac{1}{\alpha} = \frac{1}{\alpha}.$$

So, for $\alpha \in (0, 1)$,

$$\begin{aligned}
\lim_{n\to\infty} P\Big(\frac{T_n}{n} > \alpha \Big| |\mathbf{Z}_n| > 0\Big) &= \frac{1}{\alpha} E\Big(\frac{Y}{\mathbf{v} \cdot \mathbf{Q[u]}} \cdot \frac{\alpha}{1-\alpha} e^{-\frac{Y}{\mathbf{v}\cdot\mathbf{Q[u]}} \cdot \frac{\alpha}{1-\alpha}}\Big) \\
&= \frac{1}{\alpha} \frac{1}{\mathbf{v} \cdot \mathbf{Q[u]}} \frac{\alpha}{1-\alpha} \int_0^\infty y e^{-\frac{1}{\mathbf{v}\cdot\mathbf{Q[u]}} \frac{\alpha}{1-\alpha} y} \frac{1}{\mathbf{v} \cdot \mathbf{Q[u]}} e^{-\frac{1}{\mathbf{v}\cdot\mathbf{Q[u]}} y} dy \\
&= 1 - \alpha.
\end{aligned}$$

Hence, $\frac{T_n}{n}\Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} \tilde{T}$ as $n \to \infty$, where $\tilde{T}$ is an uniform $(0, 1)$ random variable and we prove Theorem 3.7.

3.4 Results in The Subcritical Case

3.4.1 The Statements of Results

In this section, we consider a discrete-time multi-type subcritical branching process, i.e. $0 < \rho < 1$ where ρ is the maximal eigenvalue of the mean matrix $\mathbf{M}$, and investigate the coalescence problems for this process.

For any integer $k \geq 2$, let $X_{n,k}$ be the generation number of the last common ancestor of any k random chosen individuals in the nth generation. First, we prove the result in Theorem 3.8 regarding to the limit behavior of $X_{n,k}$ as $n \to \infty$ when $k = 2$.

Theorem 3.8. *Let $0 < \rho < 1$ and $E\|Z_1\| \log \|Z_1\| < \infty$. Then there exists a random variable $\tilde{X}_2$ such that $n - X_{n,2} \Big| |Z_n| \geq 2 \xrightarrow{d} \tilde{X}_2$ as $n \to \infty$, and, for any $r = 0, 1, 2 \cdots$,*

$$P(\tilde{X}_2 \leq r) = 1 - \frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\phi(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)}, r)\Big) \equiv H_2(r)$$

where

$$\phi(t_1, t_2, \cdots, t_d, r) = E\left(\frac{\sum_{l=1}^{d} \sum_{i \neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}||\tilde{\mathbf{Z}}_{r,j}^{(l)}| + \sum_{l \neq p=1}^{d} \sum_{i=1}^{t_l} \sum_{j=1}^{t_p} |\tilde{\mathbf{Z}}_{r,i}^{(l)}||\tilde{\mathbf{Z}}_{r,j}^{(p)}|}{(\sum_{l=1}^{d} \sum_{i \neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}|)(\sum_{l=1}^{d} \sum_{i \neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}| - 1)} I_{\{\sum_{l=1}^{d} \sum_{i \neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}| \geq 2\}} \right)$$

and $\left\{\tilde{Z}_{r,i}^{(l)} : i \geq 1\right\}_{r \geq 0}$ are i.i.d copies of the branching process initiated by an individual of type l, $l = 1, 2, \cdots, d$.

Theorem 3.8 shows that the coalescence time $X_{n,2}$ does not go way back to the beginning of the tree. Instead, it is very close to the present so that the different $n - X_{n,2}$ between the generation number of the last common ancestor and the number of the current generation converges in distribution as $n \to \infty$ and we have seen this phenomenon in the single-type subcritical branching process.

We also have the similar analog for the general $k = 2, 3, \cdots$.

Corollary 3.2. *Let $0 < \rho < 1$ and $E\|Z_1\| \log \|Z_1\| < \infty$. Then, for $k = 2, 3, \cdots$, there exists a random variable $\tilde{X}_k$ such that $n - X_{n,k} \Big| |Z_n| \geq k \xrightarrow{d} \tilde{X}_k$ as $n \to \infty$.*

Now, to find the limit behavior of T_n which is the generation number of the last common ancestor of the whole population, we trace the lines of descent of all the individuals in the nth generation backward in time till they meet. The following theorem tells that the limit law $\tilde{T}$ of the total coalescence times is very close to the present.

Theorem 3.9. *Let* $0 < \rho < 1$ *and* $E\|Z_1\| \log \|Z_1\| < \infty$. *Then there exists a random variable* $\tilde{T}$ *such that* $n - T_n \Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} \tilde{T}$ *as* $n \to \infty$, *and, for any* $r = 0, 1, 2, \cdots$,

$$P(\tilde{T} \le r) = \rho^{-r} E\left(\sum_{l=1}^{d} Y^{(l)} g_r^{(l)} (1 - g_r^{(l)})^{Y^{(l)}-1} \cdot \prod_{p \ne l} (1 - g_r^{(p)})^{Y^{(p)}} \right) \equiv \pi(r).$$

where $\mathbf{Y}$ *is the random vector with distribution* $\{b(\mathbf{j})\}_{\mathbf{j} \in \mathbb{R}_+^d}$ *defined as in Theorem 1.9 (d).*

Next, we would like to look at the limit of the joint distribution of the generation number and the type of the last common ancestor and the types of the randomly chosen individuals.

Consider $k = 2$, i.e., pick two individuals at random from the nth generation (by simple random sampling without replacement) and trace their lines of decent back in time to find their last common ancestor. Let $X_{n,2}$ be the generation number of this common ancestor, η_n the type of this last common ancestor and $\zeta_{n,i}$ the type of the ith chosen individual.

Theorem 3.10. *Let* $0 < \rho < 1$ *and* $E\|Z_1\| \log \|Z_1\| < \infty$. *Then*

$$\lim_{n \to \infty} P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2 \Big| |\mathbf{Z}_n| \ge 2) \equiv \psi_2(r, j, i_1, i_2) \quad \textit{exists}$$

and $\displaystyle\sum_{(r,j,i_1,i_2)} \psi_2(r, j, i_1, i_2) = 1.$

Corollary 3.3. *Let* $0 < \rho < 1$ *and* $E\|Z_1\| \log \|Z_1\| < \infty$. *Then*

$$\lim_{n \to \infty} P(X_{n,k} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2, \cdots, \zeta_{n,k} = i_k \Big| |\mathbf{Z}_n| \ge k) \equiv \psi_k(r, j, i_1, i_2, \cdots, i_k)$$

exists and $\displaystyle\sum_{(r,j,i_1,i_2,\cdots,i_k)} \psi_k(r, j, i_1, i_2, \cdots, i_k) = 1.$

3.4.2 The Proof of Theorem 3.8

For any $r \geq 0$,

$$
\begin{aligned}
& P(n - X_{n,2} > r \Big| |\mathbf{Z}_n| \geq 2) \\
= \;& P(X_{n,2} < n - r \Big| |\mathbf{Z}_n| \geq 2) \\
= \;& E\left(\frac{\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\mathbf{Z}_{n-r,i,r}^{(l)}||\mathbf{Z}_{n-r,j,r}^{(l)}| + \sum_{l\neq p=1}^{d}\sum_{i=1}^{Z_{n-r}^{(l)}}\sum_{j=1}^{Z_{n-r}^{(p)}} |\mathbf{Z}_{n-r,i,r}^{(l)}||\mathbf{Z}_{n-r,j,r}^{(p)}|}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)} \Bigg| |\mathbf{Z}_n| \geq 2 \right) \\
= \;& \frac{1}{P(|\mathbf{Z}_n| \geq 2)} E\left(\frac{\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\mathbf{Z}_{n-r,i,r}^{(l)}||\mathbf{Z}_{n-r,j,r}^{(l)}| + \sum_{l\neq p=1}^{d}\sum_{i=1}^{Z_{n-r}^{(l)}}\sum_{j=1}^{Z_{n-r}^{(p)}} |\mathbf{Z}_{n-r,i,r}^{(l)}||\mathbf{Z}_{n-r,j,r}^{(p)}|}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)} I_{\{|\mathbf{Z}_n|\geq 2\}} \right) \\
= \;& \frac{1}{P(|\mathbf{Z}_n| \geq 2, |\mathbf{Z}_n| > 0)} E\left(\frac{\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\mathbf{Z}_{n-r,i,r}^{(l)}||\mathbf{Z}_{n-r,j,r}^{(l)}| + \sum_{l\neq p=1}^{d}\sum_{i=1}^{Z_{n-r}^{(l)}}\sum_{j=1}^{Z_{n-r}^{(p)}} |\mathbf{Z}_{n-r,i,r}^{(l)}||\mathbf{Z}_{n-r,j,r}^{(p)}|}{|\mathbf{Z}_n|(|\mathbf{Z}_n|-1)} I_{\{|\mathbf{Z}_n|\geq 2\}} I_{\{|\mathbf{Z}_{n-r}|>0\}} \right) \\
= \;& \frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| \geq 2 \Big| |\mathbf{Z}_n| > 0) P(|\mathbf{Z}_n| > 0)} E\left(\frac{\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\tilde{\mathbf{Z}}_{r,i}^{(l)}||\tilde{\mathbf{Z}}_{r,j}^{(l)}| + \sum_{l\neq p=1}^{d}\sum_{i=1}^{Z_{n-r}^{(l)}}\sum_{j=1}^{Z_{n-r}^{(p)}} |\tilde{\mathbf{Z}}_{r,i}^{(l)}||\tilde{\mathbf{Z}}_{r,j}^{(p)}|}{(\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\tilde{\mathbf{Z}}_{r,i}^{(l)}|)(\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\tilde{\mathbf{Z}}_{r,i}^{(l)}| - 1)} I_{\{\sum_{l=1}^{d}\sum_{i\neq j=1}^{Z_{n-r}^{(l)}} |\tilde{\mathbf{Z}}_{r,i}^{(l)}| \geq 2\}} \Bigg| |\mathbf{Z}_{n-r}| > 0 \right) \\
& \qquad \text{where } \tilde{\mathbf{Z}}_{r,i}^{(l)} \sim \mathbf{Z}_r^{(l)} \quad \forall i \text{ and } l = 1, 2, \cdots, d \\
= \;& \frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| \geq 2 \Big| |\mathbf{Z}_n| > 0) P(|\mathbf{Z}_n| > 0)} E\left(\phi(Z_{n-r}^{(1)}, Z_{n-r}^{(2)}, \cdots, Z_{n-r}^{(d)}, r) \Big| |\mathbf{Z}_{n-r}| > 0 \right)
\end{aligned}
$$

$$
\text{where } \phi(t_1, t_2, \cdots, t_d, r) = E\left(\frac{\sum_{l=1}^{d}\sum_{i\neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}||\tilde{\mathbf{Z}}_{r,j}^{(l)}| + \sum_{l\neq p=1}^{d}\sum_{i=1}^{t_l}\sum_{j=1}^{t_p} |\tilde{\mathbf{Z}}_{r,i}^{(l)}||\tilde{\mathbf{Z}}_{r,j}^{(p)}|}{(\sum_{l=1}^{d}\sum_{i\neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}|)(\sum_{l=1}^{d}\sum_{i\neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}| - 1)} I_{\{\sum_{l=1}^{d}\sum_{i\neq j=1}^{t_l} |\tilde{\mathbf{Z}}_{r,i}^{(l)}| \geq 2\}} \right)
$$

We know that $\mathbf{Z}_{n-r} \Big| |\mathbf{Z}_{n-r}| > 0 \xrightarrow{d} \mathbf{Y} \equiv (Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})$ as $n \to \infty$, so

$$
E\left(\phi(Z_{n-r}^{(1)}, Z_{n-r}^{(2)}, \cdots, Z_{n-r}^{(d)}, r) \Big| |\mathbf{Z}_{n-r}| > 0 \right) \to E\left(\phi(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)}, r) \right)
$$

as $n \to \infty$, since $\phi(\cdot, r)$ is continuous for any fixed $r \geq 0$.

Also, as $n \to \infty$, we have that $P(|\mathbf{Z}_n| \geq 2 \Big| |\mathbf{Z}_n| > 0) \to P(|\mathbf{Y}| \geq 2)$ and

$$
\frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| > 0)} \to \rho^{-r},
$$

hence, for any $r \geq 0$,

$$\begin{aligned}
&\lim_{n\to\infty} P(n - X_{n,2} > r \Big| |\mathbf{Z}_n| \geq 2) \\
= &\lim_{n\to\infty} \frac{1}{P(|\mathbf{Z}_n| \geq 2\Big||\mathbf{Z}_n| > 0)} \frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| > 0)} E\Big(\phi(Z_{n-r}^{(1)}, Z_{n-r}^{(2)}, \cdots, Z_{n-r}^{(d)}, r)\Big||\mathbf{Z}_{n-r}| > 0\Big) \\
= &\frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\phi(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)}, r)\Big) \\
\equiv &1 - \pi(r)
\end{aligned}$$

Now, it remains to show that $H_2(r) \to 1$ as $r \to \infty$.

Let $\mathbf{f}_n(\mathbf{s}) = (f_r^{(1)}(\mathbf{s}), f_r^{(2)}(\mathbf{s}), \cdots, f_r^{(d)}(\mathbf{s}))$ be the probability generating function of $\mathbf{Z}_n$, then

$$\begin{aligned}
&\frac{1}{\rho^r} E\Big(\phi(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})\Big) \\
= &\frac{1}{\rho^r} E\Big(1 - \prod_{l=1}^{d} (f_r^{(l)}(\mathbf{0}))^{Y^{(l)}} - \sum_{l=1}^{d} Y^{(l)}(1 - f_r^{(l)}(\mathbf{0}))(f_r^{(l)}(\mathbf{0}))^{(Y^{(l)}-1)} \prod_{p\neq l} (f_r^{(p)}(\mathbf{0}))^{Y^{(p)}}\Big) \\
= &E\Bigg(\frac{1 - \prod_{l=1}^{d} (f_r^{(l)}(\mathbf{0}))^{Y^{(l)}}}{\rho^r}\Bigg) - \sum_{l=1}^{d} \frac{1 - f_r^{(l)}(\mathbf{0})}{\rho^r} E\Big(Y^{(l)}(f_r^{(l)}(\mathbf{0}))^{(Y^{(l)}-1)} \prod_{p\neq l} (f_r^{(p)}(\mathbf{0}))^{Y^{(p)}}\Big)
\end{aligned}$$

First, since $E(Z_1^{(j)}(\log Z_1^{(j)}) < \infty$ for any $j = 1, 2, \cdots, d$, we have, for $l = 1, 2, \cdots, d$, $EY^{(l)} < \infty$ and

$$\frac{1 - f_r^{(l)}(\mathbf{0})}{\rho^r} \to \frac{u_l}{\mathbf{u} \cdot E\mathbf{Y}} \qquad \text{as } r \to \infty.$$

Also, $f_r^{(l)}(\mathbf{0}) \to 1$ as $r \to \infty$ for $l = 1, 2, \cdots, d$. By the bounded convergence theorem,

$$\sum_{l=1}^{d} \frac{1 - f_r^{(l)}(\mathbf{0})}{\rho^r} E\Big(Y^{(l)}(f_r^{(l)}(\mathbf{0}))^{(Y^{(l)}-1)} \prod_{p\neq l} (f_r^{(p)}(\mathbf{0}))^{Y^{(p)}}\Big) \to \sum_{l=1}^{d} \frac{u_l}{\mathbf{u} \cdot E\mathbf{Y}} EY^{(l)} = 1,$$

as $r \to \infty$.

Secondly, under the condition $E(Z_1^{(j)}(\log Z_1^{(j)}) < \infty$ for any $\mathbf{a} = (a_1, a_2, \cdots, a_d) \in \mathbb{N}_0^d$, we have that

$$\rho^{-r} P(|\mathbf{Z}_r| > 0 \Big| \mathbf{Z}_0 = \mathbf{a}) \to \frac{\mathbf{u} \cdot \mathbf{a}}{\mathbf{u} \cdot E\mathbf{Y}}$$

and

$$\frac{\mathbf{v} \cdot (1 - \mathbf{f}_r(\mathbf{0}))}{1 - \prod_{l=1}^{d} (f_r^{(l)}(\mathbf{0}))^{a_l}} = \sum_{l=1}^{d} \frac{v_l \cdot (1 - f_r^{(l)}(\mathbf{0}))}{1 - \prod_{l=1}^{d} (f_r^{(l)})^{a_l}(\mathbf{0})}.$$

Since, as r is increasing, $\dfrac{1-f_r^{(l)}(\mathbf{0})}{1-\prod_{l=1}^{d}(f_r^{(l)})^{a_l}(\mathbf{0})}$ is decreasing, $\dfrac{1-\prod_{l=1}^{d}(f_r^{(l)}(\mathbf{0}))^{a_l}}{\mathbf{v}\cdot(1-\mathbf{f}_r(\mathbf{0}))}$ is increasing and hence, by the monotone convergence theorem, we have

$$\begin{aligned}
E\left(\frac{1-\prod_{l=1}^{d}(f_r^{(l)}(\mathbf{0}))^{Y^{(l)}}}{\mathbf{v}\cdot(1-\mathbf{f}_r(\mathbf{0}))}\right) &= E\left(\frac{\rho^{-r}\left(1-\prod_{l=1}^{d}(f_r^{(l)}(\mathbf{0}))^{Y^{(l)}}\right)}{\rho^{-r}\left(\mathbf{v}\cdot(1-\mathbf{f}_r(\mathbf{0}))\right)}\right)\\
&= E\left(\frac{\rho^{-r}P(|\mathbf{Z}_r|>0\big|\mathbf{Z}_0=\mathbf{Y})}{\rho^{-r}\left(\mathbf{v}\cdot(1-\mathbf{f}_r(\mathbf{0}))\right)}\right)\\
&\to E\left(\frac{\frac{\mathbf{u}\cdot\mathbf{Y}}{\mathbf{u}\cdot E\mathbf{Y}}}{\frac{1}{\mathbf{u}\cdot E\mathbf{Y}}}\right)=\mathbf{u}\cdot E\mathbf{Y}.
\end{aligned}$$

So, as $r\to\infty$,

$$E\left(\frac{1-\prod_{l=1}^{d}(f_r^{(l)}(\mathbf{0}))^{Y^{(l)}}}{\rho^r}\right)=\frac{\mathbf{v}\cdot(1-\mathbf{f}_r(\mathbf{0}))}{\rho^r}E\left(\frac{1-\prod_{l=1}^{d}(f_r^{(l)}(\mathbf{0}))^{Y^{(l)}}}{\mathbf{v}\cdot(1-\mathbf{f}_r(\mathbf{0}))}\right)\to\frac{1}{\mathbf{u}\cdot E\mathbf{Y}}\cdot(\mathbf{u}\cdot E\mathbf{Y})=1.$$

So, we obtain that $\frac{1}{\rho^r}E\left(\phi(Y^{(1)},Y^{(2)},\cdots,Y^{(d)})\right)\to 1-1=0$ as $r\to\infty$ and hence $H_2(r)\to 1$ as $r\to\infty$ provided $P(|\mathbf{Y}|\geq 2)>0$. That is, $H_2(\cdot)$ is a probability distribution on $\mathbb{N}_0$. Therefore, there exists a random variable $\tilde{X}_2$ such that $n-X_{n,2}\Big||Z_n|\geq 2\xrightarrow{d}\tilde{X}_2$ as $n\to\infty$.

3.4.3 The Proof of Theorem 3.9

Let $\mathbf{Z}_{r,i,n-r}^{(l)}$ be the d-type Galton-Watson branching process initiated by the ith individual of type l in the rth generation, where $i=1,2,\cdots,Z_r^{(l)}$ and $l=1,2,\cdots,d$.

For any $r\geq 0$,

$$\begin{aligned}
&P(n-T_n\leq r\big||\mathbf{Z}_n|>0)\\
=\;&P(T_n>n-r\big||\mathbf{Z}_n|>0)\\
=\;&P(T_n\geq r||\mathbf{Z}_n|>0)\\
=\;&E\left(\sum_{l=1}^{d}\sum_{i=1}^{Z_{n-r}^{(l)}}P(|\mathbf{Z}_{n-r,i,r}^{(l)}|>0)\cdot\prod_{j\neq i}P(|\mathbf{Z}_{n-r,j,r}^{(l)}|=0)\cdot\prod_{p\neq l}\prod_{j=1}^{Z_{n-r}^{(p)}}P(|\mathbf{Z}_{n-r,j,r}^{(p)}|=0)\Big||\mathbf{Z}_n|>0\right)\\
=\;&\frac{P(|\mathbf{Z}_{n-r}|>0)}{P(|\mathbf{Z}_n|>0)}E\left(\sum_{l=1}^{d}Z_{n-r}^{(l)}g_r^{(l)}(1-g_r^{(l)})^{Z_{n-r}^{(l)}-1}\cdot\prod_{p\neq l}(1-g_r^{(p)})^{Z_{n-r}^{(p)}}\Big||\mathbf{Z}_{n-r}|>0\right)
\end{aligned}$$

where $g_n^{(l)} = P(|\mathbf{Z}_n| > 0 \big| \mathbf{Z_0} = \mathbf{e}_l)$.

Let $h(x_1, x_2, \cdots, x_d) = E\bigg(\sum_{l=1}^{d} x_l g_r^{(l)}(1 - g_r^{(l)})^{x_l - 1} \cdot \prod_{p \neq l} (1 - g_r^{(p)})^{x_p}\bigg)$, then h is continuous at $(x_1, x_2, \cdots, x_d)$ and

$$P(n - X_n \leq r \big| |\mathbf{Z}_n| > 0) = E\Big(h(Z_{n-r}^{(1)}, Z_{n-r}^{(2)}, \cdots, Z_{n-r}^{(d)}) \Big| |\mathbf{Z}_{n-r}| > 0\Big).$$

Since $\mathbf{Z}_{n-r} \Big| |\mathbf{Z}_{n-r}| > 0 \equiv (Z_{n-r}^{(1)}, Z_{n-r}^{(2)}, \cdots, Z_{n-r}^{(d)}) \Big| |\mathbf{Z}_{n-r}| > 0 \xrightarrow{d} \mathbf{Y} \equiv (Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})$ as $n \to \infty$,

$$E\Big(h(Z_{n-r}^{(1)}, Z_{n-r}^{(2)}, \cdots, Z_{n-r}^{(d)}) \Big| |\mathbf{Z}_{n-r}| > 0\Big) \to E(h(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})).$$

as $n \to \infty$.

Also, $\dfrac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| > 0)} = \rho^{-r}$ and then, for each $r \geq 0$,

$$P(n - X_n \leq r \big| |\mathbf{Z}_n| > 0) \to \rho^{-r} E\bigg(\sum_{l=1}^{d} Y^{(l)} g_r^{(l)} (1 - g_r^{(l)})^{Y^{(l)} - 1} \cdot \prod_{p \neq l} (1 - g_r^{(p)})^{Y^{(p)}}\bigg) \equiv \pi(r)$$

as $n \to \infty$.

Moreover, since $g_r^{(l)} \to 0$ and $\rho^{-r} g_r^{(l)} \to \dfrac{u_l}{u \cdot EY}$ as $r \to \infty$, we have

$$\begin{aligned}
\lim_{r\to\infty} \pi(r) &= \lim_{r\to\infty} E\bigg(\sum_{l=1}^{d} Y^{(l)} \rho^{-r} g_r^{(l)} (1 - g_r^{(l)})^{Y^{(l)} - 1} \cdot \prod_{p \neq l} (1 - g_r^{(p)})^{Y^{(p)}}\bigg) \equiv \pi(r) \\
&= E\bigg(\sum_{l=1}^{d} Y^{(l)} \frac{u_l}{\mathbf{u} \cdot E\mathbf{Y}}\bigg) \\
&= \frac{1}{\mathbf{u} \cdot E\mathbf{Y}} \sum_{l=1}^{d} u_l E Y^{(l)} \\
&= 1.
\end{aligned}$$

Therefore, $\pi(\cdot)$ is a proper probability distribution and hence there exists a random variable $\tilde{T}$ with

$$P(\tilde{T} \leq r) = \pi(r), \qquad r = 0, 1, 2, \cdots$$

such that $n - T_n \Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} \tilde{T}$ as $n \to \infty$.

3.4.4 The Proof of Theorem 3.10

Let $\xi_{n,j}^{i} = (\xi_{n,j}^{i(1)}, \xi_{n,j}^{i(2)}, \cdots, \xi_{n,j}^{i(d)})$ be the vector of offsprings of the jth individual of type i in the nth generation.

Let $\{\mathbf{Z}^{j,l}_{p,r,s,n} : n \geq 0\}$ be the branching process initiated by the sth child of type l of the pth individual of type j in the $r$the generation, where

$$\mathbf{Z}^{j,l}_{p,r,s,n} = (Z^{j,l,(1)}_{p,r,s,n}, Z^{j,l,(2)}_{p,r,s,n}, \cdots, Z^{j,l,(d)}_{p,r,s,n}).$$

Choose two individuals at random from the nth generation and Let $A_{n,i}$ be the type of the ancestor in the next generation of the nearest common ancestor (the last common ancestor) of the ith chosen individual, $i = 1, 2$. Then

$$\begin{aligned} & P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2} \Big| |\mathbf{Z}_n| \geq 2) \\ = \;& E\left(\frac{\sum_{p=1}^{Z^{(j)}_{n-r}} \sum_{l=1}^{d} \sum_{s \neq t=1}^{\xi^{j(l)}_{n-r,p}} Z^{j,l,(i)}_{p,n-r,s,r-1} Z^{j,l,(i)}_{p,n-r,t,r-1}}{|\mathbf{Z}_n|(|\mathbf{Z}_n| - 1)} \Bigg| |\mathbf{Z}_n| \geq 2 \right) \\ = \;& \frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| \geq 2 \Big| |\mathbf{Z}_n| > 0) P(|\mathbf{Z}_n| > 0)} E\left(\frac{\sum_{p=1}^{Z^{(j)}_{n-r}} \sum_{l=1}^{d} \sum_{s \neq t=1}^{\xi^{j(l)}_{n-r,p}} Z^{j,l,(i)}_{p,n-r,s,r-1} Z^{j,l,(i)}_{p,n-r,t,r-1}}{|\mathbf{Z}_n|(|\mathbf{Z}_n| - 1)} I_{\{|\mathbf{Z}_n| \geq 2\}} \Bigg| |\mathbf{Z}_{n-r}| > 0 \right) \end{aligned}$$

Let $\xi^j = (\xi^{j(1)}, \xi^{j(2)}, \cdots, \xi^{j(\alpha)})$ be i.i.d. copies the vector of offspring of an individual of type j, then

$$\xi^j_{n-r,p} = (\xi^{j(1)}_{n-r,p}, \xi^{j(2)}_{n-r,p}, \cdots, \xi^{j(\alpha)}_{n-r,p}) \sim \xi^j = (\xi^{j(1)}, \xi^{j(2)}, \cdots, \xi^{j(\alpha)}).$$

Let $\tilde{\mathbf{Z}}^l_{r-1,s}$ be the i.i.d. copies of $\mathbf{Z}_{r-1}$ with $\mathbf{Z}_0 = \mathbf{e}_l$, then

$$\mathbf{Z}^{jl}_{p,n-r,s,r-1} = (Z^{j,l,(1)}_{p,n-r,t,r-1}, Z^{j,l,(2)}_{p,n-r,t,r-1}, \cdots, Z^{j,l,(\alpha)}_{p,n-r,t,r-1}) \sim \tilde{\mathbf{Z}}^l_{r-1,s} = (\tilde{Z}^{l(1)}_{r-1,s}, \tilde{Z}^{l(2)}_{r-1,s}, \cdots, \tilde{Z}^{l(\alpha)}_{r-1,s}).$$

So,

$$\begin{aligned} & P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2} \Big| |\mathbf{Z}_n| \geq 2) \\ = \;& \frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| \geq 2 \Big| |\mathbf{Z}_n| > 0) P(|\mathbf{Z}_n| > 0)} E\left(\frac{\sum_{p=1}^{Z^{(j)}_{n-r}} \sum_{l=1}^{d} \sum_{s \neq t=1}^{\xi^{j(l)}} \tilde{Z}^{l(i)}_{r-1,s} \tilde{Z}^{l(i)}_{r-1,t}}{|\mathbf{Z}_n|(|\mathbf{Z}_n| - 1)} I_{\left\{ \sum_{j=1}^{d} \sum_{p=1}^{Z^{(j)}_{n-r}} \sum_{l=1}^{d} \sum_{s=1}^{\xi^{j(l)}} |\tilde{\mathbf{Z}}^l_{r-1,s}| \geq 2 \right\}} \Bigg| |\mathbf{Z}_{n-r}| > 0 \right) \\ = \;& \frac{P(|\mathbf{Z}_{n-r}| > 0)}{P(|\mathbf{Z}_n| \geq 2 \Big| |\mathbf{Z}_n| > 0) P(|\mathbf{Z}_n| > 0)} E\left(\varphi_1(Z^{(1)}_{n-r}, Z^{(2)}_{n-r}, \cdots, Z^{(d)}_{n-r}) \Big| |\mathbf{Z}_{n-r}| > 0 \right) \end{aligned}$$

where

$$\varphi_1(x_1, x_2, \cdots, x_d, r) = E\left(\frac{\sum_{p=1}^{x_j} \sum_{l=1}^{d} \sum_{s \neq t=1}^{\xi^{j(l)}} \tilde{Z}^{l(i)}_{r-1,s} \tilde{Z}^{l(i)}_{r-1,t} I_{\left\{ \sum_{j=1}^{d} \sum_{p=1}^{x_j} \sum_{l=1}^{d} \sum_{s=1}^{\xi^{j(l)}} |\tilde{\mathbf{Z}}^l_{r-1,s}| \geq 2 \right\}}}{\left(\sum_{j=1}^{d} \sum_{p=1}^{x_j} \sum_{l=1}^{d} \sum_{s=1}^{\xi^{j(l)}} |\tilde{\mathbf{Z}}^l_{r-1,s}| \right) \left(\sum_{j=1}^{d} \sum_{p=1}^{x_j} \sum_{l=1}^{d} \sum_{s=1}^{\xi^{j(l)}} |\tilde{\mathbf{Z}}^l_{r-1,s}| - 1 \right)} \right).$$

Since $\varphi_1(\cdot, r)$ is continuous and $\mathbf{Z}_{n-r}\Big||\mathbf{Z}_{n-r}| > 0 \xrightarrow{d} \mathbf{Y}$ as $n \to \infty$, we have

$$P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2}\Big||\mathbf{Z}_n| \geq 2) \\ \to \frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\varphi_1(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})\Big)$$

as $n \to \infty$.

Similarly, we also have that, as $n \to \infty$,

$$P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} \neq A_{n,2}\Big||\mathbf{Z}_n| \geq 2) \\ \to \frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\varphi_2(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})\Big),$$

$$P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = i_1 \neq \zeta_{n,2} = i_2, A_{n,1} = A_{n,2}\Big||\mathbf{Z}_n| \geq 2) \\ \to \frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\varphi_3(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})\Big)$$

and

$$P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = i_1 \neq \zeta_{n,2} = i_2, A_{n,1} \neq A_{n,2}\Big||\mathbf{Z}_n| \geq 2) \\ \to \frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\varphi_4(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})\Big)$$

where

$$\varphi_2(x_1, x_2, \cdots, x_d, r) = E\left(\frac{\sum_{p=1}^{x_j}\sum_{l\neq q=1}^{d}\sum_{s=1}^{\xi^{j(l)}}\sum_{t=1}^{\xi^{j(l)}} \tilde{Z}_{r-1,s}^{l(i)}\tilde{Z}_{r-1,t}^{q(i)} I_{\{\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}|\geq 2\}}}{\left(\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}|\right)\left(\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}| - 1\right)}\right),$$

$$\varphi_3(x_1, x_2, \cdots, x_d, r) = E\left(\frac{\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s\neq t=1}^{\xi^{j(l)}} \tilde{Z}_{r-1,s}^{l(i_1)}\tilde{Z}_{r-1,t}^{l(i_2)} I_{\{\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}|\geq 2\}}}{\left(\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}|\right)\left(\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}| - 1\right)}\right).$$

and

$$\varphi_4(x_1, x_2, \cdots, x_d, r) = E\left(\frac{\sum_{p=1}^{x_j}\sum_{l\neq q=1}^{d}\sum_{s=1}^{\xi^{j(l)}}\sum_{t=1}^{\xi^{j(l)}} \tilde{Z}_{r-1,s}^{l(i_1)}\tilde{Z}_{r-1,t}^{q(i_2)} I_{\{\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}|\geq 2\}}}{\left(\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}|\right)\left(\sum_{j=1}^{d}\sum_{p=1}^{x_j}\sum_{l=1}^{d}\sum_{s=1}^{\xi^{j(l)}}|\tilde{\mathbf{Z}}_{r-1,s}^{l}| - 1\right)}\right).$$

Let $\varphi = \varphi_1 + \varphi_2 + \varphi_3 + \varphi_4$, then, as $n \to \infty$,

$$\begin{aligned} & P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2, \Big| |\mathbf{Z}_n| \geq 2) \\ \to & \frac{1}{\rho^r P(|\mathbf{Y}| \geq 2)} E\Big(\phi(Y^{(1)}, Y^{(2)}, \cdots, Y^{(d)})\Big) \equiv \psi_2(r, j, i_1, i_2). \end{aligned}$$

Since $n - X_{n,2}\Big||Z_n| \geq 2 \xrightarrow{d} \tilde{X}_2$ as $n \to \infty$, $\{n - X_{n,2}\}_{n\geq 0}$ is tight and, also, $\{\eta_n\}_{n\geq 0}$, $\{\zeta_{n,1}\}_{n\geq 0}$ and $\{\zeta_{n,2}\}_{n\geq 0}$ only take values on the finite set $\{1, 2, \cdots, d\}$, so $\Big\{(n - X_{n,2}, \eta_n, \zeta_{n,1}, \zeta_{n,2})\Big\}_{n\geq 0}$ is tight. Therefore, the limit $\psi_2(r, j, i_1, i_2)$ of $P(n - X_{n,2} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2)$ is a probability mass function on $\mathbb{N}_0 \times \{1, 2, \cdots, d\} \times \{1, 2, \cdots, d\} \times \{1, 2, \cdots, d\}$. That is,

$$\sum_{(r,j,i_1,i_2)} \psi_2(r, j, i_1, i_2) = 1.$$

3.5 The Markov Property on Types

3.5.1 The Statements of Results

Consider a discrete-time multi-type Galton-Watson branching process $\{\mathbf{Z}_n\}_{n\geq 0}$.

We pick an individual at random from the nth generation and record its type, then trace its line of descent backward and also record the types of its ancestors along the line of descent.

Let $I_{n,0}$ be the type of this randomly chosen individual.

Let $I_{n,i}$ be the ancestor of this individual in the $(n - i)$th generation, $i = 1, 2, \cdots, n$.

The first theorem is a result on the limit behavior of the types of the last k ancestors of a randomly chosen individual in the nth generation and the Markov property on the types of the ancestors along its line of descent as $n \to \infty$ for the supercritical case.

Theorem 3.11. *Let $1 < \rho < \infty$, $|\mathbf{Z}_0| = 1$, $E||Z_1|| \log ||Z_1|| < \infty$ and assume $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ for any $i = 1, 2, \cdots, d$. Then, for any integer $k \geq 0$, there exist random variables $\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k$ such that*

$$(I_{n,0}, I_{n,1}, \cdots, I_{n,k}) \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k) \quad \text{as } n \to \infty,$$

and, for any $i_0, i_1, \cdots, i_k \in \{1, 2, \cdots, d\}$,

$$P(\tilde{I}_0 = i_0, \tilde{I}_1 = i_1, \cdots, \tilde{I}_k = i_k) = \frac{v_{i_k} m_{i_k i_{k-1}} m_{i_{k-1} i_{k-2}} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v})\rho^k}$$

where $\mathbf{v} = (v_1, v_2, \cdots, v_d)$ *is a left eigenvector of the offspring mean matrix* $\mathbf{M} = \{m_{ij} : i, j = 1, 2, \cdots, d\}$ *associated with the maximal eigenvalue* ρ*.*

Moreover, $\{\tilde{I}_n\}_{n\geq 0}$ *is a Markov chain with the state space* $\{1, 2, \cdots, d\}$ *and*

(a) the initial distribution $\lambda_0 \equiv (\lambda_0(1), \lambda_0(2), \cdots, \lambda_0(d))$ *where*

$$\lambda_0(i) = \frac{v_i}{\mathbf{1} \cdot \mathbf{v}} \quad \text{for any } i = 1, 2, \cdots, d.$$

(b) the transition probability $\mathbf{P} \equiv (p_{ij} : i, j = 1, 2, \cdots, d)$*, where*

$$p_{ij} = \frac{v_j m_{ji}}{v_i \rho} \quad \text{for any } n = 0, 1, 2, \cdots$$

(c) the stationary distribution $\pi \equiv (\pi_1, \pi_2 \cdots, \pi_d)$ *where*

$$\pi_i = \frac{u_i v_i}{\mathbf{u} \cdot \mathbf{v}} \quad \text{for any } i = 1, 2, \cdots, d.$$

We also have an analog of the result on the limit law and the Markov property of the types along the line of ancestor of any individual randomly chosen from the nth generation as $n \to \infty$ for the critical case.

Theorem 3.12. *Let* $\rho = 1$*,* $|\mathbf{Z}_0| = 1$ *and* $E\|Z_1\|^2 < \infty$*. Then, for any integer* $k \geq 0$*, there exist random variables* $\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k$ *such that*

$$(I_{n,0}, I_{n,1}, \cdots, I_{n,k})\Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k) \quad \text{as } n \to \infty,$$

and, for any $i_0, i_1, \cdots, i_k \in \{1, 2, \cdots, d\}$*,*

$$P(\tilde{I}_0 = i_0, \tilde{I}_1 = i_1, \cdots, \tilde{I}_k = i_k) = \frac{v_{i_k} m_{i_k i_{k-1}} m_{i_{k-1} i_{k-2}} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v})}$$

where $\mathbf{v} = (v_1, v_2, \cdots, v_d)$ *is the left eigenvector of the offspring mean matrix* $\mathbf{M} = \{m_{ij} : i, j = 1, 2, \cdots, d\}$ *associated with the maximal eigenvalue* 1*.*

Moreover, $\{\tilde{I}_n\}_{n\geq 0}$ *is a Markov chain with the state space* $\{1, 2, \cdots, d\}$ *and*

(a) the initial distribution $\lambda_0 \equiv (\lambda_0(1), \lambda_0(2), \cdots, \lambda_0(d))$ *where*

$$\lambda_0(i) = \frac{v_i}{\mathbf{1} \cdot \mathbf{v}} \quad \text{for any } i = 1, 2, \cdots, d.$$

(b) the transition probability $\mathbf{P} \equiv (p_{ij} : i, j = 1, 2, \cdots, d)$*, where*

$$p_{ij} = \frac{v_j m_{ji}}{v_i} \quad \textit{for any } n = 0, 1, 2, \cdots$$

(c) the stationary distribution $\pi \equiv (\pi_1, \pi_2 \cdots, \pi_d)$ *where*

$$\pi_i = \frac{u_i v_i}{\mathbf{u} \cdot \mathbf{v}} \quad \textit{for any } i = 1, 2, \cdots, d.$$

3.5.2 The Proof of Theorem 3.11

Let $\mathbf{Z}_n = (Z_{n,1}, Z_{n,2}, \cdots, Z_{n,d})$ be the population vector in the nth generation, $n = 0, 1, 2, \cdots$, where $Z_{n,i}$ is the number of individuals of type i in the nth generation.

We will prove this theorem using the principle of mathematical induction.

For $k = 0$, since, in the supercritical case, it is known that $\frac{\mathbf{Z}_n}{\rho^n} \to \mathbf{v}W$ w.p.1 as $n \to \infty$ and $P(0 < W < \infty) = 1$ (see Theorem 1.6), we have, by the bounded convergence theorem,

$$P(I_{n,0} = i_0) = E\Big(\frac{Z_{n,i_0}}{|\mathbf{Z}_n|}\Big) = E\Big(\frac{Z_{n,i_0}/\rho^n}{|\mathbf{Z}_n|/\rho^n}\Big) \to \frac{v_{i_0} W}{(\mathbf{1} \cdot \mathbf{v})W} = \frac{v_{i_0}}{\mathbf{1} \cdot \mathbf{v}} \equiv \lambda_0(i_0) \quad \text{as } n \to \infty.$$

Also, $\sum_{i=1}^{d} \lambda_0(i) = \sum_{i=1}^{d} \frac{v_i}{\mathbf{1} \cdot \mathbf{v}} = 1$, i.e., $\{\lambda_0(i) : i = 1, 2, \cdots, d\}$ is a proper probability distribution and hence there exists a random variable $\tilde{I}_0$ with $P(\tilde{I}_0 = i) = \lambda_0(i)$ for $i = 1, 2, \cdots, d$ such that $I_0 \xrightarrow{d} \tilde{I}_0$ as $n \to \infty$.

Next, we prove that the theorem holds for $k = 1$.

Let $\xi_{n,j}^{(i)} = (\xi_{n,r}^{(j)1}, \xi_{n,r}^{(j)2}, \cdots, \xi_{n,r}^{(j)d})$ be the vector of offsprings of the jth individual of type i in the nth generation. Then $\left\{\xi_{n,j}^{(i_1)i_0}\right\}_{j \geq 1, n \geq 1}$ are i.i.d random variables with $E\Big(\xi_{n,j}^{(i_1)i_0}\Big) = m_{i_1 i_0} < \infty$.

Since $Z_{n,i_0} \to \infty$ w.p.1, then by the strong law of large numbers, as $n \to \infty$,

$$\frac{1}{Z_{n,i_0}} \sum_{j=1}^{Z_{n,i_0}} \xi_{n,j}^{(i_1)i_0} \to m_{i_1 i_0} \quad \text{w.p.1.}$$

So, by the bounded convergence theorem,

$$\begin{aligned} P(I_{n,1} = i_1 \big| I_{n,0} = i_0) &= E\left(\frac{\sum_{j=1}^{Z_{n-1,i_1}} \xi_{n-1,j}^{(i_1)i_0}}{Z_{n,i_0}}\right) = E\left(\frac{1}{Z_{n-1,i_1}} \sum_{j=1}^{Z_{n-1,i_1}} \xi_{n-1,j}^{(i_1)i_0} \cdot \frac{Z_{n-1,i_1}/\rho^{n-1}}{Z_{n,i_0}/\rho^n} \cdot \frac{1}{\rho}\right) \\ &\to m_{i_1 i_0} \cdot \frac{v_{i_1} W}{v_{i_0} W} \cdot \frac{1}{\rho} = \frac{v_{i_1} m_{i_1 i_0}}{\rho v_{i_0}} \quad \text{as } n \to \infty \end{aligned}$$

Hence,

$$P(I_{n,0} = i_0, I_{n,1} = i_1) = P(I_{n,1} = i_1 \big| I_{n,0} = i_0)P(I_{n,0} = i_0) \to \frac{v_{i_1} m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho} \equiv \lambda_1(i_0, i_1) \quad \text{as } n \to \infty$$

and

$$\sum_{i=1}^{d}\sum_{j=1}^{d} \lambda_1(i,j) = \sum_{i=1}^{d} \frac{1}{(\mathbf{1}\cdot\mathbf{v})\rho}\Big(\sum_{j=1}^{d} v_j m_{ji}\Big) = \sum_{i=1}^{d} \frac{\rho v_i}{(\mathbf{1}\cdot\mathbf{v})\rho} = \sum_{i=1}^{d} \lambda_0(i) = 1$$

since $\mathbf{v}$ is the left eigenvector of $\mathbf{M}$ associated with the eigenvalue ρ.

So, $\{\lambda_1(i,j) : i, j = 1, 2, \cdots, d\}$ is a proper probability distribution with one marginal distribution λ_0. Thus, there exists a random variable $\tilde{I}_1$ such that $P(\tilde{I}_0 = i, \tilde{I}_1 = j) = \lambda_1(i,j)$ for $i, j = 1, 2, \cdots, d$ and $(I_0, I_1) \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1)$ as $n \to \infty$.

Assume that there exist random variables $\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k$ such that

$$P(\tilde{I}_{n,0} = i_0, \tilde{I}_{n,1} = i_1, \cdots, \tilde{I}_k = i_k) = \frac{v_{i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho^k} \equiv \lambda_k(i_0, i_1, \cdots, i_k)$$

and, as $n \to \infty$,

$$(I_0, I_1, \cdots, I_k) \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k).$$

Then

$$\begin{aligned}
& P(I_{k+1} = i_{k+1}, I_k = i_k, \cdots, I_i = i_1 \big| I_0 = i_0) \\
= \;& E\left(\frac{\sum_{j_{k+1}=1}^{Z_{n-(k+1),i_{k+1}}} \sum_{j_k=1}^{\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}} \cdots \sum_{j_1}^{\xi^{(i_2)i_1}_{n-2,j_2}} \xi^{(i_1)i_0}_{n-1,j_1}}{Z_{n,i_0}} \right) \\
= \;& E\Bigg(\frac{1}{Z_{n-(k+1),i_{k+1}}} \frac{1}{\xi_{n-(k+1),j^{(i_{k+1})i_k}_{k+1}}} \cdots \frac{1}{\xi^{(i_2)i_1}_{n-2,j_2}} \sum_{j_{k+1}=1}^{Z_{n-(k+1),i_{k+1}}} \sum_{j_k=1}^{\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}} \cdots \sum_{j_1}^{\xi^{(i_2)i_1}_{n-2,j_2}} \left(\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}} \cdots \xi^{(i_2)i_1}_{n-2,j_2} \xi^{(i_1)i_0}_{n-1,j_1} \right) \\
& \qquad \cdot \frac{Z_{n-(k+1),i_{k+1}}/\rho^{n-(k+1)}}{Z_{n,i_0}/\rho^n} \cdot \frac{1}{\rho^{k+1}} \Bigg)
\end{aligned}$$

and, again by Theorem 1.6, the strong law of large numbers and the bounded convergence theorem, we have that, as $n \to \infty$,

$$P(I_{k+1} = i_{k+1}, I_k = i_k, \cdots, I_i = i_1 \big| I_0 = i_0) \to \frac{v_{i_{k+1}} m_{i_{k+1} i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{v_{i_0} \rho^{k+1}}$$

Hence, as $n \to \infty$,

$$\begin{aligned} & P(I_{n,0} = i_0, I_{n,1} = i_1, \cdots, I_{k+1} = i_{k+1}) \\ = \; & P(I_{k+1} = i_{k+1}, I_k = i_k, \cdots, I_i = i_1 \big| I_0 = i_0) P(I_0 = i_0) \\ \to \; & \frac{v_{i_{k+1}} m_{i_{k+1} i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v}) \rho^{k+1}} \equiv \lambda_{k+1}(i_0, i_1, \cdots, i_k) \end{aligned}$$

and

$$\sum_{i_0=1}^{d} \sum_{i_1=1}^{d} \cdots \sum_{i_{k+1}=1}^{d} \lambda_{k+1}(i_1, i_1, \cdots, i_{k+1}) = \sum_{i_0=1}^{d} \sum_{i_1=1}^{d} \cdots \sum_{i_k=1}^{d} \lambda_k(i_1, i_1, \cdots, i_k) = 1.$$

So, there exists a random variable $\tilde{I}_{k+1}$ such that

$$P(\tilde{I}_{n,0} = i_0, \tilde{I}_{n,1} = i_1, \cdots, \tilde{I}_k = i_k, \tilde{I}_{k+1} = i_{k+1}) = \lambda_{k+1}(i_0, i_1, \cdots, i_k, i_{k+1}) = \frac{v_{i_{k+1}} m_{i_{k+1} i_k} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v}) \rho^{k+1}}$$

and, as $n \to \infty$, $(I_0, I_1, \cdots, I_k, I_{k+1}) \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k, \tilde{I}_{k+1})$.

By the principle of the mathematical induction, we prove that, for any integer $k \geq 0$, there exist random variables $(\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k)$ such that

$$(I_{n,0}, I_{n,1}, \cdots, I_{n,k}) \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k) \quad \text{as } n \to \infty,$$

and, for any $i_0, i_1, \cdots, i_k \in \{1, 2, \cdots, d\}$,

$$P(\tilde{I}_0 = i_0, \tilde{I}_1 = i_1, \cdots, \tilde{I}_k = i_k) = \frac{v_{i_k} m_{i_k i_{k-1}} m_{i_{k-1} i_{k-2}} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v}) \rho^k}.$$

Now, we prove the Markov property of $\{\tilde{I}_n\}_{n \geq 0}$.

For any $n \geq 1$ and any $i, j, i_0, \cdots, i_{n-1} \in \{1, 2, \cdots, d\}$, we have

$$\begin{aligned} P(\tilde{I}_{n+1} = j \big| \tilde{I}_n = i, \tilde{I}_{n-1} = i_{n-1}, \cdots, \tilde{I}_0 = i_0) &= \frac{P(\tilde{I}_{n+1} = j, \tilde{I}_n = i, \tilde{I}_{n-1} = i_{n-1}, \cdots, \tilde{I}_0 = i_0)}{P(\tilde{I}_n = i, \tilde{I}_{n-1} = i_{n-1}, \cdots, \tilde{I}_0 = i_0)} \\ &= \frac{v_j m_{ji} m_{i i_{n-1}} \cdots m_{i_1 i_0} / (\mathbf{1} \cdot \mathbf{v}) \rho^{n+1}}{v_i m_{i i_{n-1}} \cdots m_{i_1 i_0} / (\mathbf{1} \cdot \mathbf{v}) \rho^n} \\ &= \frac{v_j m_{ji}}{v_i \rho} \equiv p_{ij}. \end{aligned}$$

So, the conditional probability distribution of the future state of the chain $\{\tilde{I}_n\}_{n \geq 0}$, given the present state and the past states, only depends on the present state. Therefore, $\{\tilde{I}_n\}_{n \geq 0}$ is a Markov chain with the state space $\{1, 2, \cdots, d\}$ such that

(a) the initial distribution $\lambda_0 \equiv (\lambda_0(1), \lambda_0(2), \cdots, \lambda_0(d))$ where

$$\lambda_0(i) = \frac{v_i}{\mathbf{1} \cdot \mathbf{v}} \quad \text{for any } i = 1, 2, \cdots, d.$$

(b) the transition probability $\mathbf{P} \equiv (p_{ij} : i, j = 1, 2, \cdots, d)$, where

$$p_{ij} = \frac{v_j m_{ji}}{v_i \rho} \quad \text{for any } n = 0, 1, 2, \cdots$$

It remains to show that the Markov chain $\{\tilde{I}_n\}_{n\geq 0}$ has a stationary distribution $\pi \equiv (\pi_1, \pi_2 \cdots, \pi_d)$ where

$$\pi_i = \frac{u_i v_i}{\mathbf{u} \cdot \mathbf{v}} \quad \text{for any } i = 1, 2, \cdots, d.$$

Since $\mathbf{u} > \mathbf{0}$ and $\mathbf{v} > \mathbf{0}$, $\pi_i = \dfrac{u_i v_i}{\mathbf{u} \cdot \mathbf{v}} > 0$. Also,

$$\sum_{i=1}^{d} \pi_i = \sum_{i=1}^{d} \frac{u_i v_i}{\mathbf{u} \cdot \mathbf{v}} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{v}} = 1.$$

So, $\pi \equiv (\pi_1, \pi_2 \cdots, \pi_d)$ is a probability distribution.

Moreover, since $\mathbf{u}$ is a right eigenvector of $\mathbf{M}$ associated with the eigenvalue ρ, for any $j = 1, 2, \cdots, d$,

$$\sum_{i=1}^{d} \pi_i p_{ij} = \sum_{i=1}^{d} \frac{u_i v_i}{\mathbf{u} \cdot \mathbf{v}} \cdot \frac{v_j m_{ji}}{v_i \rho} = \frac{v_j}{\rho(\mathbf{u} \cdot \mathbf{v})} \sum_{i=1}^{d} m_{ji} u_i = \frac{v_j}{\rho(\mathbf{u} \cdot \mathbf{v})} \cdot \rho u_j = \frac{v_j u_j}{\mathbf{u} \cdot \mathbf{v}} = \pi_j$$

and hence π is a stationary distribution of the transition probability $\mathbf{P}$.

Therefore, the proof is complete.

3.5.3 The Proof of Theorem 3.12

Before we prove Theorem 3.12, we need the following lemmas.

Let $\mathbf{u}$ and $\mathbf{v} = (v_1, v_2, \cdots, v_d)$ be the right and left eigenvector, respectively, of the offspring mean matrix $\mathbf{M}$ associated with the maximal eigenvalue ρ.

Lemma 3.2. *(Mode, 1971) Let $\rho = 1$, $|\mathbf{Z}_0| = 1$ and $E\|Z_1\|^2 < \infty$. Then, for any $i = 1, 2, \cdots, d$,*

$$\lim_{n\to\infty} E\left(\frac{Z_{n,i}}{n}\middle| |\mathbf{Z}_n| > 0\right) = v_i(\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}])$$

Remark 3.3. From Lemma 3.2, we know that, as $n \to \infty$, $\frac{Z_{n,i}}{n}\Big||\mathbf{Z}_n| > 0$ converges to $v_i(\mathbf{v} \cdot \mathbb{Q}[\mathbf{u}])$ in L^1 and hence in probability.

Lemma 3.3. *(Karlin, 1966) Let* $\mathbf{K} = \{\mathbf{x} = (x_1, x_2, \cdots, x_d) : x_i > 0, \mathbf{x} \cdot \mathbf{u} = 1\}$. *Then*

$$\lim_{n\to\infty} \sup_{\mathbf{x}\in\mathbf{K}} \|\mathbf{x}\mathbf{M}^n - \rho^n - \mathbf{v}\| = 0.$$

Lemma 3.3 is in the same spirit as the Frobeninus theorem and we will make use of it to prove the next lemma.

Lemma 3.4. *Let* $\rho = 1$, $|\mathbf{Z}_0| = 1$ *and* $E\|Z_1\|^2 < \infty$. *Then, for any* $i = 1, 2, \cdots, d$ *and any* $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\omega : \left|\frac{Z_{n,i}(\omega)}{\mathbf{u}\cdot\mathbf{Z}_n(\omega)} - v_i\right| > \epsilon \Big| |\mathbf{Z}_n| > 0\right) = 0$$

Proof. Let $\mathbf{Z}_m^{(j)l}(n) = \left(Z_{m,1}^{(j)l}(n), Z_{m,2}^{(j)l}(n), \cdots, Z_{m,d}^{(j)l}(n)\right)$ where $Z_{m,i}^{(j)l}(n)$ is the number of type i offspring in the $(m+n)$th generation of the lth individual of type j in the nth generation. Then, by the additive property,

$$\begin{aligned} Z_{n+m,i}^{(k)} &= \text{the number of the individuals of type } i \text{ in the } (m+n)\text{th generation of the process} \\ &\quad\ \text{initiated with an individual of type } k \\ &= \sum_{j=1}^{d} \sum_{l=1}^{Z_{n,j}^{(k)}} Z_{m,i}^{(j)l}(n). \end{aligned}$$

Let $\mathbf{X}_n = (X_{n,1}, X_{n,2}, \cdots, X_{n,d}) \equiv \frac{\mathbf{Z}_n}{\mathbf{u}\cdot\mathbf{Z}_n}$, that is,

$$\frac{X_{n,j}}{Z_{n,j}} = \frac{1}{\mathbf{u}\cdot\mathbf{Z}_n} \quad \text{for all } j = 1, 2, \cdots, d.$$

Then

$$\begin{aligned} \mathbf{Z}_{n+m}^{(k)} &= \sum_{j=1}^{d} \sum_{l=1}^{Z_{n,j}^{(k)}} \mathbf{Z}_m^{(j)l}(n) \\ &= \sum_{j=1}^{d} \sum_{l=1}^{Z_{n,j}^{(k)}} \left(\mathbf{Z}_m^{(j)l}(n) - \mathbf{e}_j\mathbf{M}^m\right) + \sum_{j=1}^{d} \sum_{l=1}^{Z_{n,j}^{(k)}} \mathbf{e}_j\mathbf{M}^m \\ &= \sum_{j=1}^{d} \sum_{l=1}^{Z_{n,j}^{(k)}} \left(\mathbf{Z}_m^{(j)l}(n) - \mathbf{e}_j\mathbf{M}^m\right) + \sum_{j=1}^{d} Z_{n,j}^{(k)}\mathbf{e}_j\mathbf{M}^m \end{aligned}$$

and hence

$$\begin{aligned}
\mathbf{u}\cdot\mathbf{Z}_{n+m}^{(k)} &= \sum_{j=1}^{d}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-u_j\mathbf{M}^m\Big)+\sum_{j=1}^{d}\mathbf{u}\cdot Z_{n,j}^{(k)}\mathbf{e}_j\mathbf{M}^m \\
&= \sum_{j=1}^{d}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-\rho^m u_j\Big)+\sum_{j=1}^{d}\rho^m u_j Z_{n,j}^{(k)} \\
&= \sum_{j=1}^{d}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-\rho^m u_j\Big)+\rho^m(\mathbf{u}\cdot\mathbf{Z}_n^{(k)}).
\end{aligned}$$

So,

$$\begin{aligned}
\mathbf{X}_{n+m}^{(k)} &= \frac{\mathbf{Z}_{n+m}^{(k)}}{\mathbf{u}\cdot\mathbf{Z}_{n+m}^{(k)}} \\
&= \frac{\sum_{j=1}^{d} Z_{n,j}^{(k)}\mathbf{e}_j\mathbf{M}^m+\sum_{j=1}^{d}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{Z}_m^{(j)l}(n)-\mathbf{e}_j\mathbf{M}^m\Big)}{\rho^m(\mathbf{u}\cdot\mathbf{Z}_n^{(k)})+\sum_{j=1}^{d}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-\rho^m u_j\Big)} \\
&= \frac{\sum_{j=1}^{d}\frac{1}{\rho^m\left(\mathbf{u}\cdot\mathbf{Z}_n^{(k)}\right)}Z_{n,j}^{(k)}\mathbf{e}_j\mathbf{M}^m+\sum_{j=1}^{d}\frac{1}{\rho^m\left(\mathbf{u}\cdot\mathbf{Z}_n^{(k)}\right)}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{Z}_m^{(j)l}(n)-\mathbf{e}_j\mathbf{M}^m\Big)}{1+\sum_{j=1}^{d}\frac{1}{\rho^m\left(\mathbf{u}\cdot\mathbf{Z}_n^{(k)}\right)}\sum_{l=1}^{Z_{n,j}^{(k)}}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-\rho^m u_j\Big)} \\
&= \frac{\rho^{-m}\sum_{j=1}^{d}\frac{Z_{n,j}^{(k)}}{\left(\mathbf{u}\cdot\mathbf{Z}_n^{(k)}\right)}\mathbf{e}_j\mathbf{M}^m+\sum_{j=1}^{d}\left(\frac{X_{n,j}}{Z_{n,j}}\right)\sum_{l=1}^{Z_{n,j}^{(k)}}\rho^{-m}\Big(\mathbf{Z}_m^{(j)l}(n)-\mathbf{e}_j\mathbf{M}^m\Big)}{1+\sum_{j=1}^{d}\left(\frac{X_{n,j}}{Z_{n,j}}\right)\sum_{l=1}^{Z_{n,j}^{(k)}}\rho^{-m}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-\rho^m u_j\Big)}
\end{aligned}$$

Suppressing the superscript k and letting

$$r_{nm}=\sum_{j=1}^{d}\left(\frac{X_{n,j}}{Z_{n,j}}\right)\sum_{l=1}^{Z_{n,j}}\rho^{-m}\Big(\mathbf{u}\cdot\mathbf{Z}_m^{(j)l}(n)-\rho^m u_j\Big)$$

and

$$\alpha_{nm}=\sum_{j=1}^{d}\left(\frac{X_{n,j}}{Z_{n,j}}\right)\sum_{l=1}^{Z_{n,j}}\rho^{-m}\Big(\mathbf{Z}_m^{(j)l}(n)-\mathbf{e}_j\mathbf{M}^m\Big)$$

then we can have

$$\mathbf{X}_{n+m}=\frac{\rho^{-m}\sum_{j=1}^{d}X_{n,j}\mathbf{e}_j\mathbf{M}^m+\alpha_{nm}}{1+r_{nm}}=\frac{\rho^{-m}\mathbf{X}_n\mathbf{M}^m+\alpha_{nm}}{1+r_{nm}}$$

and hence

$$\mathbf{X}_{n+m} - \mathbf{v} = \frac{\left(\mathbf{X}_n \mathbf{M}^m \rho^{-m} - \mathbf{v}\right) - \mathbf{v} r_{nm} + \alpha_{nm}}{1 + r_{nm}}. \tag{3.1}$$

First, since $\left\{\mathbf{u} \cdot \mathbf{Z}_m^{(j)l}\right\}_{n\geq 0}$ are i.i.d random variables with $E\left(\mathbf{u} \cdot \mathbf{Z}_m^{(j)l} - \rho^m u_j\right) = 0$ and $\left\{\mathbf{Z}_m^{(j)l}\right\}_{n\geq 0}$ are i.i.d random vectors with $E\left(\mathbf{Z}_m^{(j)l}(n) - \mathbf{e}_j \mathbf{M}^m\right) = \mathbf{0}$, by the strong law of large numbers, on the set $\{|\mathbf{Z}_n| > 0, \mathbf{Z}_n \to \infty\}$,

$$\frac{1}{Z_{n,j}} \sum_{l=1}^{Z_{n,j}} \left(\mathbf{u} \cdot \mathbf{Z}_m^{(j)l}(n) - \rho^m u_j\right) \to 0 \quad \text{w.p.1}$$

and

$$\frac{1}{Z_{n,j}} \sum_{l=1}^{Z_{n,j}} \left(\mathbf{Z}_m^{(j)l}(n) - \mathbf{e}_j \mathbf{M}^m\right) \to \mathbf{0} \quad \text{w.p.1}$$

and hence $\lim_{n\to\infty} |r_{nm}| = 0$ and $\lim_{n\to\infty} \|\alpha_{nm}\| = 0$. Therefore, for any $\eta > 0$,

$$\lim_{n\to\infty} P\left(|r_{nm}| > \eta, \mathbf{Z}_n \to \infty \middle| |\mathbf{Z}_n| > 0\right) = \lim_{n\to\infty} P\left(\|\alpha_{nm}\| > \eta, \mathbf{Z}_n \to \infty \middle| |\mathbf{Z}_n| > 0\right) = 0.$$

Let $\epsilon > 0$ be arbitrary, then, by Lemma 3.3, there exists an m_0 such that for all $m \geq m_0$,

$$\sup_{\mathbf{x}\in\mathbf{K}} \|\mathbf{x}\mathbf{M}^n - \rho^n - \mathbf{v}\| \leq \epsilon.$$

From (3.1), we have, for any $\epsilon > 0$, $\eta > 0$, that

$$\begin{aligned} & P\left(\mathbf{Z}_n \to \infty, \|\mathbf{X}_{n+m} - \mathbf{v}\| \leq \frac{\epsilon + \eta + \|\mathbf{v}\|\eta}{1-\eta} \middle| |\mathbf{Z}_n| > 0\right) \\ \geq \;& 1 - P\left(|r_{nm}| > \eta, \mathbf{Z}_n \to \infty \middle| |\mathbf{Z}_n| > 0\right) - P\left(\|\alpha_{nm}\| > \eta, \mathbf{Z}_n \to \infty \middle| |\mathbf{Z}_n| > 0\right). \end{aligned}$$

This implies that for any $\epsilon > 0$, $\eta > 0$,

$$\limsup_{N\to\infty} P\left(\|\mathbf{X}_{n+m} - \mathbf{v}\| > \frac{\epsilon + \eta + \|\mathbf{v}\|\eta}{1-\eta} \middle| |\mathbf{Z}_n| > 0\right) = 0,$$

which proves Lemma 3.4.

□

Remark 3.4. Lemma 3.4 demonstrates the proportions of individual of various types approach the corresponding ratios of the components of the left eigenvector $\mathbf{v}$ of the mean matrix $\mathbf{M}$ associated with the maximal eigenvalue ρ. That is, for any $i = 1, 2, \cdots, d$, the random quantity $\frac{Z_{n,i}}{|\mathbf{Z}|_n}$, conditioned on $\{|\mathbf{Z}_n| > 0\}$, converges to $\frac{v_i}{\mathbf{1} \cdot \mathbf{v}}$ (which is a non-random quantities) in probability as $n \to \infty$.

Now, we are ready to prove Theorem 3.12 using the principle of mathematical induction.

First, consider $k = 0$, for any $i_0 = 1, 2, \cdots, d$, by the bounded convergence theorem,

$$P\Big(I_0 = i_0 \Big| |\mathbf{Z}_n| > 0\Big) = E\Big(\frac{Z_{n,i_0}}{|\mathbf{Z}_n|}\Big| |\mathbf{Z}_n| > 0\Big) \to \frac{v_{i_0}}{\mathbf{1}\cdot\mathbf{v}} \equiv \lambda_0(i_0).$$

and $\sum_{i=1}^{d} \lambda_0(i) = 1$. So, there exists a random variable $\tilde{I}_0$ on $\{1, 2, \cdots, d\}$ such that $P(\tilde{I}_0 = i) = \lambda_0(i)$ for all $i = 1, 2, \cdots, d$ and

$$I_0 \Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} \tilde{I}_0 \quad \text{as } n \to \infty.$$

Next, for $k = 1$, let $\xi^{(i)}_{n,j} = (\xi^{(j)1}_{n,r}, \xi^{(j)2}_{n,r}, \cdots, \xi^{(j)d}_{n,r})$ be the vector of offsprings of the jth individual of type i in the nth generation. Then $\big\{\xi^{(i_1)i_0}_{n,j}\big\}_{j\geq 1, n\geq 1}$ are i.i.d random variables with $E\big(\xi^{(i_1)i_0}_{n,j}\big) = m_{i_1 i_0} < \infty$.

Since, on the set $\{|\mathbf{Z}_n| > 0\}$, $Z_{n,i_0} \to \infty$ w.p.1, then by the strong law of large numbers, as $n \to \infty$,

$$\frac{1}{Z_{n,i_0}} \sum_{j=1}^{Z_{n,i_0}} \xi^{(i_1)i_0}_{n,j} \Big| |\mathbf{Z}_n| > 0 \to m_{i_1 i_0} \quad \text{w.p.1.}$$

Also, it is known from Lemma 3.2 that $\frac{Z_{n,i}}{n}\Big| |\mathbf{Z}_n| > 0$ converges to $v_i Y$ in L^1 and hence in probability as $n \to \infty$, where Y is the exponential random variable defined as in Theorem 1.8 (b). So, by the bounded convergence theorem,

$$\begin{aligned} P\Big(I_{n,1} = i_1 \Big| I_{n,0} = i_0, |\mathbf{Z}_n| > 0\Big) &= E\left(\frac{\sum_{j=1}^{Z_{n-1,i_1}} \xi^{(i_1)i_0}_{n-1,j}}{Z_{n,i_0}}\Bigg| |\mathbf{Z}_n| > 0\right) \\ &= E\left(\frac{1}{Z_{n-1,i_1}} \sum_{j=1}^{Z_{n-1,i_1}} \xi^{(i_1)i_0}_{n-1,j} \cdot \frac{Z_{n-1,i_1}/(n-1)}{Z_{n,i_0}/n} \cdot \frac{n-1}{n}\right) \\ &\to m_{i_1 i_0} \cdot \frac{v_{i_1} Y}{v_{i_0} Y} = \frac{v_{i_1} m_{i_1 i_0}}{v_{i_0}} \quad \text{as } n \to \infty \end{aligned}$$

Hence, we have that, as $n \to \infty$,

$$P\Big(I_{n,0} = i_0, I_{n,1} = i_1 \Big| \mathbf{Z}_n| > 0\Big) = P\Big(I_{n,1} = i_1 \Big| I_{n,0} = i_0, |\mathbf{Z}_n| > 0\Big) P\Big(I_{n,0} = i_0 \Big| |\mathbf{Z}_n| > 0\Big) \to \frac{v_{i_1} m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})} \equiv \lambda_1(i_0, i_1)$$

and

$$\sum_{i=1}^{d}\sum_{j=1}^{d} \lambda_1(i,j) = \sum_{i=1}^{d} \frac{1}{(\mathbf{1}\cdot\mathbf{v})}\Big(\sum_{j=1}^{d} v_j m_{ji}\Big) = \sum_{i=1}^{d} \frac{v_i}{(\mathbf{1}\cdot\mathbf{v})} = \sum_{i=1}^{d} \lambda_0(i) = 1$$

since $\mathbf{v}$ is the left eigenvector of $\mathbf{M}$ associated with the eigenvalue $\rho = 1$.

So, $\{\lambda_1(i,j) : i,j = 1,2,\cdots,d\}$ is a proper probability distribution with one marginal distribution λ_0. Thus, there exists a random variable $\tilde{I}_1$ such that $P(\tilde{I}_0 = i, \tilde{I}_1 = j) = \lambda_1(i,j)$ for $i,j = 1,2,\cdots,d$ and $(I_0, I_1)\Big||\mathbf{Z}_n| > 0 \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1)$ as $n \to \infty$.

Now, assume that there exist random variables $\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k$ such that

$$P(\tilde{I}_{n,0} = i_0, \tilde{I}_{n,1} = i_1, \cdots, \tilde{I}_k = i_k) = \frac{v_{i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho^k} \equiv \lambda_k(i_0, i_1, \cdots, i_k)$$

and, as $n \to \infty$,

$$(I_0, I_1, \cdots, I_k)\Big||\mathbf{Z}_n| > 0 \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k).$$

Then

$$\begin{aligned}
&P\Big(I_{k+1} = i_{k+1}, I_k = i_k, \cdots, I_i = i_1 \Big| I_0 = i_0, |\mathbf{Z}_n| > 0\Big) \\
= \; & E\left(\frac{\sum_{j_{k+1}=1}^{Z_{n-(k+1),i_{k+1}}} \sum_{j_k=1}^{\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}} \cdots \sum_{j_1}^{\xi^{(i_2)i_1}_{n-2,j_2}} \xi^{(i_1)i_0}_{n-1,j_1}}{Z_{n,i_0}} \Bigg| |\mathbf{Z}_n| > 0 \right) \\
= \; & E\Bigg(\frac{1}{Z_{n-(k+1),i_{k+1}}} \frac{1}{\xi_{n-(k+1),j^{(i_{k+1})i_k}_{k+1}}} \cdots \frac{1}{\xi^{(i_2)i_1}_{n-2,j_2}} \sum_{j_{k+1}=1}^{Z_{n-(k+1),i_{k+1}}} \sum_{j_k=1}^{\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}} \cdots \sum_{j_1}^{\xi^{(i_2)i_1}_{n-2,j_2}} \left(\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}} \cdots \xi^{(i_2)i_1}_{n-2,j_2} \xi^{(i_1)i_0}_{n-1,j_1} \right) \\
& \qquad \cdot \frac{Z_{n-(k+1),i_{k+1}}/(n-(k+1))}{Z_{n,i_0}/n} \cdot \frac{n-(k+1)}{n} \Bigg| |\mathbf{Z}_n| > 0 \Bigg)
\end{aligned}$$

and, again by Theorem 1.6, the strong law of large numbers and the bounded convergence theorem, we have that, as $n \to \infty$,

$$P\Big(I_{k+1} = i_{k+1}, I_k = i_k, \cdots, I_i = i_1 \Big| I_0 = i_0, |\mathbf{Z}_n| > 0\Big) \to \frac{v_{i_{k+1}} m_{i_{k+1} i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{v_{i_0} \rho^{k+1}}$$

Hence, as $n \to \infty$,

$$\begin{aligned}
&P\Big(I_{n,0} = i_0, I_{n,1} = i_1, \cdots, I_{k+1} = i_{k+1} \Big| |\mathbf{Z}_n| > 0\Big) \\
= \; & P\Big(I_{k+1} = i_{k+1}, I_k = i_k, \cdots, I_i = i_1 \Big| I_0 = i_0, |\mathbf{Z}_n| > 0\Big) P\Big(I_0 = i_0 \Big| |\mathbf{Z}_n| > 0\Big) \\
\to \; & \frac{v_{i_{k+1}} m_{i_{k+1} i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})} \equiv \lambda_{k+1}(i_0, i_1, \cdots, i_k)
\end{aligned}$$

and

$$\sum_{i_0=1}^{d} \sum_{i_1=1}^{d} \cdots \sum_{i_{k+1}=1}^{d} \lambda_{k+1}(i_1, i_1, \cdots, i_{k+1}) = \sum_{i_0=1}^{d} \sum_{i_1=1}^{d} \cdots \sum_{i_k=1}^{d} \lambda_k(i_1, i_1, \cdots, i_k) = 1.$$

So, there exists a random variable $\tilde{I}_{k+1}$ such that

$$P(\tilde{I}_{n,0} = i_0, \tilde{I}_{n,1} = i_1, \cdots, \tilde{I}_k = i_k, \tilde{I}_{k+1} = i_{k+1}) = \lambda_{k+1}(i_0, i_1, \cdots, i_k, i_{k+1}) = \frac{v_{i_{k+1}} m_{i_{k+1} i_k} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v})}$$

and, as $n \to \infty$, $(I_0, I_1, \cdots, I_k, I_{k+1}) \Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k, \tilde{I}_{k+1})$.

By the principle of the mathematical induction, we prove that, for any integer $k \geq 0$, there exist random variables $(\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k)$ such that

$$(I_{n,0}, I_{n,1}, \cdots, I_{n,k}) \Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} (\tilde{I}_0, \tilde{I}_1, \cdots, \tilde{I}_k) \quad \text{as } n \to \infty,$$

and, for any $i_0, i_1, \cdots, i_k \in \{1, 2, \cdots, d\}$,

$$P(\tilde{I}_0 = i_0, \tilde{I}_1 = i_1, \cdots, \tilde{I}_k = i_k) = \frac{v_{i_k} m_{i_k i_{k-1}} m_{i_{k-1} i_{k-2}} \cdots m_{i_1 i_0}}{(\mathbf{1} \cdot \mathbf{v})}.$$

By the proof similar to the lines in the proof for the supercritical case, one can show the Markov property of $\{\tilde{I}_n\}_{n \geq 0}$ for the critical case and thus the proof is complete.

CHAPTER 4. COALESCENCE IN CONTINUOUS-TIME SINGLE-TYPE AGE-DEPENDENT BELLMAN-HARRIS BRANCHING PROCESSES

4.1 Introduction

Now, we consider a continuous-time single-type age-dependent Bellman-Harris branching process $\{Z(t) : t \geq 0\}$ with offspring distribution $\{p_j\}_{j\geq 0}$ and lifetime distribution G. Assume that this process is initiated with one individual of age 0. That is, $Z(0) = 1$.

For any family tree $\mathcal{T}$, since every individual lives a random length of time according to G, when we look at the population at time t, for any $t > 0$, those who are alive at this time may belong to different generations. But if we ignore the lifetime structure of this process, there is a corresponding discrete-time single-type Galton Watson branching process $\{Y_n\}_{n\geq 0}$ with offspring distribution $\{p_j\}_{j\geq 0}$, where Y_n is the number of individuals in the continuous-time process $Z(t)$ who were born as an nth-generation offspring, that means that each of these Y_n individuals in has exactly n ancestors along its line of descent. We call $\{Y_n\}_{n\geq 0}$ *the embedded generation process* of the process $\{Z(t) : t \geq 0\}$. Therefore, the results (presented in Section 1.2) on the discrete-time Galton-Watson branching process can be applied to this embedded process $\{Y_n\}_{n\geq 0}$ of the continuous-time Bellman-Harris branching process.

In this chapter, we will adopt all the definitions and notations described in Section 1.4.

Let's consider the coalescent problem on the continuous-time processes.

Let $k \geq 2$ be a positive integer. We pick k individuals from those who are alive at time t, (assuming $Z(t) \geq k$) by the simple random sampling without replacement and trace their lines of descent backward in time until they meet for the first time.

Let $X_k(t)$ be the generation number of the last common ancestor of these k random chosen individuals and Then we can ask the same questions as we do for the discrete-time Galton-Watson branching process. That is,

(1) What is the distribution of $X_k(t)$?

(2) What happens to $X_k(t)$ when $n \to \infty$?

Moreover, for a continuous-time branching process, we can ask more questions about the ”time”.

Let $D_k(t)$ be the coalescence time of the lines of descent of any k individuals randomly chosen from the population alive at time t. Note that $D_k(t)$ also means the death time of the last common ancestor and then the following questions are of our interest:

(3) What is the distribution of $D_k(t)$?

(4) What happens to $D_k(t)$ when $n \to \infty$?

In Sections 4.2 and 4.3, we present the results on the limits behaviors of the generation number and the death time of the last common ancestor for the supercritical and subcritical cases in the continuous-time single-type age-dependent Bellman-Harris branching process.

4.2 Results in The Supercritical Case

4.2.1 The statement of Results

The first theorem is regarding the generation number of the last common ancestor of k individuals randomly chosen from the population alive at time t.

Let $L_{n,i,k}$ be the lifetime of the ancestor in the kth generation of the ith individual in the rth generation, then $\{L_{r,i,k} : n \geq 0, i \geq 1, k = 1, 2, \cdots, n-1\}$ are i.i.d copies with the lifetime distribution G.

Let $S_{n,i} = \sum_{k=0}^{n-1} L_{n,i,k}$, then $S_{n,i}$ is the birth time of the ith individual in the nth generation.

Theorem 4.1. *Let $1 < m =< \infty$, $p_0 = 0$ and the life time distribution G is non-lattice with $G(0+) = 0$. If $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$, then, for any integer $k \geq 2$,*

(a) for almost all trees $\mathcal{T}$ and $r = 0, 1, 2, \cdots$,

$$P(X_k(t) < r \big| \mathcal{T}) \to \phi_k(r, \mathcal{T}) \equiv 1 - \frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k}{W^k}$$

as $t \to \infty$, where $\left\{W_{r,i}\right\}_{i\geq 1}$ are the i.i.d copies of the random variable W in Theorem 1.11 (b).

(b) *there exists a random variable* $\tilde{X}_k$ *on* $\{0, 1, 2, \cdots\}$ *such that* $X_k(t) \xrightarrow{d} \tilde{X}_k$ *as* $t \to \infty$ *and*

$$P(\tilde{X}_k < r) = 1 - E\left(\frac{\sum_{i=1}^{Y_r} \left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k}{W^k}\right) \equiv \phi_k(r).$$

for any $r = 0, 1, 2, \cdots$.

Similar to the result in the discrete-time process, when $k \to \infty$, the random variable $\tilde{D}_k$ also converges in distribution to a proper random variable which is the last generation consisting of only one individual. It is stated in the next theorem.

Theorem 4.2. *Let* $1 < m < \infty$ *and* $U = \min\{n \geq 1 : Y_n \geq 2\}$. *Under the same hypotheses of Theorem 4.1, then* $\tilde{X}_k \xrightarrow{d} U - 1 \quad$ *as* $k \to \infty$.

Now, we switch our focus to the death time of the last common ancestor.

Let $L_{s,i}$ be the total lifetime of the ith individual alive at time s. Then $\left\{L_{s,i}\right\}_{i \geq 1}$ are i.i.d. copies of the lifetime random variable with distribution G.

Let $a_{s,i}$ be the corresponding age and $R_{s,i}$ be the corresponding residual lifetime at time t. That is, $R_{s,i} = L_{s,i} - a_{s,i}$ for any $i \geq 1$ and any $s \geq 0$.

Theorem 4.3. *Let* $1 < m =< \infty$, $p_0 = 0$ *and the life time distribution* G *is non-lattice with* $G(0+) = 0$. *If* $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$, *then, for any integer* $k \geq 2$,

(a) *for almost all trees* $\mathcal{T}$ *and any* $s \geq 0$,

$$P(\tilde{D}_k \leq s \big| \mathcal{T}) \equiv H_k(s, \mathcal{T}) = 1 - \frac{\sum_{i=1}^{Z(s)} \left(e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\right)^k}$$

where $\{\tilde{W}_{r,i}\}_{i \geq 1}$ *are the i.i.d copies of the sum* $\sum_{j=i}^{\xi} W_i$, ξ *is the random variable with the offspring distribution* $\{p_j\}_{j \geq 0}$ *and* $\{W_i\}_{i \geq 0}$ *are i.i.d. copies of* W *as defined in Theorem 1.11 (b).*

(b) *there exists a random variable* $\tilde{D}_k$ *on the set of non-negative real numbers such that* $D_k(t) \xrightarrow{d} \tilde{D}_k$ *as* $t \to \infty$ *and*

$$P(\tilde{D}_k \leq s) = 1 - E\left(\frac{\sum_{i=1}^{Z(s)} \left(e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\right)^k}\right) \equiv H_k(s).$$

for any $s \geq 0$.

The next theorem shows that the limit law of $\tilde{D}_k$, the limit of the death time of the last common ancestor of any k randomly chosen individuals, converges to the first moment when the process slits into more than one as $k \to \infty$.

Theorem 4.4. *Let* $1 < m < \infty$ *and* $U = \min\{n \geq 1 : Y_n \geq 2\}$. *Under the same hypotheses of Theorem 4.3, then there exist a random variable* $\tilde{D}$ *such that* $\tilde{D}_k \xrightarrow{d} \tilde{D}$ *as* $k \to \infty$ *and, for any* $s \geq 0$,

$$P(\tilde{D} \leq s) = P(L_0 + L_1 + \cdots + L_{U-1} \leq s)$$

where $\{L_i\}_{i\geq 0}$ *are i.i.d. copies of the lifetime random variable L with distribution G and U is as defined in Theorem 4.2.*

4.2.2 The proof of Theorem 4.1

Let $\{Y_n\}_{n\geq 0}$ be the embedded generation process of the continuous-time Bellman-Harris process $\{Z(t) : t \geq 0\}$.

Let $\{Z_{r,i}(t) : t > 0\}$ be the the continuous-time single-type age-dependent Bellman-Harris branching process initiated with the ith individual in the rth generation when it is of age 0.

Let $L_{n,i,k}$ be the lifetime of the ancestor in the kth generation of the ith individual in the rth generation, then $\{L_{r,i,k} : n \geq 0, i \geq 1, k = 1, 2, \cdots, n-1\}$ are i.i.d copies with the lifetime distribution G.

Let $S_{n,i} = \sum_{k=0}^{n-1} L_{n,i,k}$, then $S_{n,i}$ is the birth time of the ith individual in the nth generation.

(a) For almost all trees $\mathcal{T}$ and any $r = 0, 1, 2, \cdots$,

$$\begin{aligned}
& P\Big(X_k(t) \geq r \Big| \mathcal{T}\Big) \\
= \quad & \frac{\sum_{i=1}^{Y_r} \binom{Z_{r,i}(t-S_{r,i})}{k}}{\binom{Z(t)}{k}} \\
= \quad & \frac{\sum_{i=1}^{Y_r} Z_{r,i}(t-S_{r,i})[Z_{r,i}(t-S_{r,i})-1]\cdots[Z_{r,i}(t-S_{r,i})-k+1]}{Z(t)[Z(t)-1]\cdots[Z(t)-k+1]} \qquad (4.1) \\
= \quad & \frac{\sum_{i=1}^{Y_r} e^{-\alpha(t-S_{r,i})}Z_{r,i}(t-S_{r,i}) \cdot e^{-\alpha(t-S_{r,i})}[Z_{r,i}(t-S_{r,i})-1]\cdots e^{-\alpha(t-S_{r,i})}[Z_{r,i}(t-S_{r,i})-k+1] \cdot e^{-k\alpha S_{r,i}}}{e^{-\alpha t}Z(t) \cdot e^{-\alpha t}[Z(t)-1]\cdots e^{-\alpha t}[Z(t)-k+1]}
\end{aligned}$$

where α is the Malthusian parameter for the offspring mean m and the lifetime distribution G.

It known from Theorem 1.11 that if $Z_0 = 1$, $p_0 = 0$ and $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$, then

$$e^{-\alpha t}Z(t) \to W \quad \text{w.p.1} \quad \text{as } t \to \infty$$

where W is a random variable such that $P(W > 0) = 1$. So, as $t \to \infty$,

$$P(X_k \geq r \big| \mathcal{T}) \to \frac{\sum_{i=1}^{Y_r} \left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k}{W^k} \equiv 1 - \phi_k(r, \mathcal{T})$$

as $t \to \infty$, where $\left\{W_{r,i}\right\}_{i\geq 1}$ are the i.i.d copies of W.

(b) Since $P(X_k(t) \geq r) = E(P(X_k(t) \geq r)\big|\mathcal{T})$ and hence, by the bounded convergence theorem,

$$P(X_k(t) \geq r) \to E\Bigg(\frac{\sum_{i=1}^{Y_r} \left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k}{W^k}\Bigg) \equiv 1 - \phi_k(r) \quad \text{as } t \to \infty$$

for $r = 1, 2, \cdots$.

To finish the proof, we need to show that ϕ_k is a proper probability distribution, i.e., $\phi_k(r) \to 1$ as $r \to \infty$, and it is sufficient to prove that

$$\sum_{i=1}^{Y_r} \left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k \to 0 \quad \text{in probability} \quad \text{as } r \to \infty.$$

(Then, by the bounded convergence theorem, we can have that $E\Bigg(\frac{\sum_{i=1}^{Y_r} \left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k}{W^k}\Bigg) \to 0$ as $r \to \infty$ and hence complete the proof.)

First, we have that

$$\left(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i}} W_{r,i}\right)^k \leq \sum_{i=1}^{Y_r} \left(e^{-\alpha S_{r,i}} W_{r,i}\right)^k \leq \left(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i} W_{r,i}}\right)^{k-1} \sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} \tag{4.2}$$

and

$$\begin{aligned} E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}\Big) &= E\Big(E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}\Big| L_0, L_1, \cdots, L_{r-1}, Y_0, Y_1, \cdots, Y_r\Big)\Big) \\ &= E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} E\Big(W_{r,i}\Big| L_0, L_1, \cdots, L_{r-1}, Y_0, Y_1, \cdots, Y_r\Big)\Big) \end{aligned}$$

Note that $\{W_{r,i}\}_{i\geq 1}$ are i.i.d. copies of W and are independent of $\{L_0, L_1, \cdots, L_{r-1}, Y_0, Y_1, \cdots, Y_r\}$, so

$$\begin{aligned} & E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}\Big) \\ = \; & E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} E(W)\Big) = EW \cdot E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}\Big) \\ = \; & EW \cdot E\Big(E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}\Big|Y_r\Big)\Big) = EW \cdot E\Big(Y_r E\Big(e^{-\alpha S_{r,1}}\Big|Y_r\Big)\Big) \\ = \; & EW \cdot E\Big(Y_r E\Big(e^{-\alpha S_{r,1}}\Big)\Big) = EW \cdot EY_r \cdot E\Big(e^{-\alpha S_{r,1}}\Big) = EW \cdot EY_r \cdot \Big(Ee^{-\alpha L}\Big)^r \end{aligned}$$

since $\Big\{S_{r,i} \equiv \sum_{k=0}^{r-1} L_{r,i,k}\Big\}_{i\geq 1}$ are identically distributed and $\{L_{r,i,k} : 0 \leq k \leq r-1\}$ are i.i.d copies of the lifetime random variable L for each $i \geq 1$.

From Theorem 1.11 (b), it is known that $EW = 1$. Then,

$$E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}\Big) = EW \cdot m^r \cdot (\varphi_L(\alpha))^r = EW = 1 < \infty \tag{4.3}$$

where $\varphi_L(\alpha) \equiv \int_0^\infty e^{-\alpha u} dG(u)$ and hence $m\varphi_L(\alpha) = 1$ since α is the Mathusian parameter for m and G.

For any $\eta > 0$, by Chebyshev's inequality,

$$P\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \eta\Big) \leq \frac{1}{\eta} E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}\Big) = \frac{1}{\eta}.$$

For any $\epsilon > 0$,

$$\begin{aligned} & P\Big(\Big(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i} W_{r,i}}\Big)^{k-1} \sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon\Big) \\ = \; & P\Big(\Big(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i} W_{r,i}}\Big)^{k-1} \sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon, \sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \eta\Big) \\ & \quad + P\Big(\Big(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i} W_{r,i}}\Big)^{k-1} \sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon, \sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i} \leq \eta\Big) \\ \leq \; & \frac{1}{\eta} + P\Big(\Big(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i} W_{r,i}}\Big)^{k-1} > \frac{\epsilon}{\eta}\Big) \end{aligned} \tag{4.4}$$

So, to prove that

$$\sum_{i=1}^{Y_r} \Big(e^{-\alpha S_{r,i}} W_{r,i}\Big)^k \to 0 \quad \text{in probability} \quad \text{as } r \to \infty,$$

it suffices, from (4.2) and (4.4), to prove that

$$\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} \to 0 \quad \text{in probability} \quad \text{as } r \to \infty.$$

Let $\mathfrak{F}_r$ be the $\sigma-$algebra generated by all the information up to the rth generation in the embedded tree. Then, for any $\epsilon > 0$,

$$\begin{aligned} P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon \Big| \mathfrak{F}_r\Big) &= P\Big(\exists i = 1, 2, \cdots, Y_r \text{ s.t. } e^{-\alpha S_{r,i}} W_{r,i} > \epsilon \Big| \mathfrak{F}_r\Big) \\ &\le \sum_{i=1}^{Y_r} P\Big(e^{-\alpha S_{r,i}} W_{r,i} > \epsilon \Big| \mathfrak{F}_r\Big) \\ &= \sum_{i=1}^{Y_r} P\Big(W_{r,i} > \epsilon e^{\alpha S_{r,i}} \Big| \mathfrak{F}_r\Big) \end{aligned}$$

Let $\eta(y) = \sup_{x\ge y} xP(W > x)$. Since $EW < \infty$, $xP(W > x) \to 0$ as $x \to \infty$. So, for any $l > 0$, there exists $a > 0$ s.t. $yP(W > y) < l\epsilon$ for all $y \ge a$ and hence $\eta(a) \le l\epsilon$.

Let $n > \dfrac{1}{\alpha} \ln \dfrac{a}{\epsilon}$, then $\epsilon e^{\alpha n} > a$. Hence,

$$\begin{aligned} & P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon\Big) \\ = \; & P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon, \min_{1\le i\le Y_r} S_{r,i} \le n\Big) + P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon, \min_{1\le i\le Y_r} S_{r,i} > n\Big) \\ \le \; & P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + E\Big(P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon, \min_{1\le i\le Y_r} S_{r,i} > n \Big| \mathfrak{F}_r\Big)\Big) \\ \le \; & P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + E\Big(\sum_{i=1}^{Y_r} P\Big(W_{r,i} > \epsilon e^{\alpha S_{r,i}}, \min_{1\le i\le Y_r} S_{r,i} > n \Big| \mathfrak{F}_r\Big)\Big) \\ = \; & P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + \frac{1}{\epsilon} E\Big(\sum_{i=1}^{Y_r} \epsilon e^{\alpha S_{r,i}} P\Big(W_{r,i} > \epsilon e^{\alpha S_{r,i}}, \min_{1\le i\le Y_r} S_{r,i} > n \Big| \mathfrak{F}_r\Big) \cdot e^{-\alpha S_{r,i}}\Big) \\ \le \; & P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + \frac{1}{\epsilon} E\Big(\sum_{i=1}^{Y_r} \eta(a) e^{-\alpha S_{r,i}}\Big) \\ = \; & P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + \frac{1}{\epsilon}\eta(a) E\Big(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}\Big) \end{aligned}$$

and, by (4.3), we have that

$$P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon\Big) \le P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + \frac{1}{\epsilon}\eta(a) < P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) + l. \tag{4.5}$$

Moreover,

$$\begin{aligned} P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) &= \sum_{x=0}^{\infty} P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n \Big| Y_r = x\Big) P(Y_r = x) \\ &\le \sum_{x=0}^{\infty} x P(S_{r,1} \le n) P(Y_r = x) \\ &= P(S_{r,1} \le n) EY_r \\ &= P(e^{-\theta S_{r,1}} \le e^{-\theta n}) EY_r \end{aligned}$$

where $\theta > \alpha$ such that $m\varphi_L(\theta) < 1$. Then, by Markov inequality,

$$P\Big(\min_{1\le i\le Y_r} S_{r,i} \le n\Big) \le \frac{E\Big(e^{-\theta S_{r,1}}\Big)}{e^{-\theta n}} m^r = e^{\theta n}\Big(Ee^{-\theta L}\Big)^r m^r = e^{\theta n}\Big(m\varphi_L(\theta)\Big)^r \to 0 \tag{4.6}$$

as $r \to \infty$.

Since (4.5) and (4.6) together imply that $P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon\Big) < l$ as $r \to \infty$ for any $l > 0$. Hence, for any $\epsilon > 0$,

$$P\Big(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon\Big) \to 0 \quad \text{as } r \to \infty,$$

and so ϕ_k is a proper probability distribution and hence there exists a random variable $\tilde{X}_k$ on $\{0, 1, 2, \cdots\}$ such that $X_k(t) \xrightarrow{d} \tilde{X}_k$ as $t \to \infty$ and

$$P(\tilde{X}_k < r) = 1 - E\left(\frac{\sum_{i=1}^{Y_r} \Big(e^{-\alpha S_{r,i}} W_{r,i}\Big)^k}{W^k}\right) \equiv \phi_k(r).$$

for any $r = 0, 1, 2, \cdots$.

The proof of Theorem 4.1 is complete.

4.2.3 The proof of Theorem 4.2

Since the number of individuals alive at time t can be expressed as the sum of all the offsprings alive at time t of all individuals in the rth generation. That is, for $t > 0$,

$$Z(t) = \sum_{i=1}^{Y_r} Z_{r,i}(t - S_{r,i})$$

and then (5.10) can be written as

$$
\begin{aligned}
&\frac{\sum_{i=1}^{Y_r} Z_{r,i}(t-S_{r,i})[Z_{r,i}(t-S_{r,i})-1]\cdots[Z_{r,i}(t-S_{r,i}-k+1)]}{Z(t)[Z(t)-1]\cdots[Z(t)-k+1]} \\
= &\frac{\sum_{i=1}^{Y_r} Z_{r,i}(t-S_{r,i})[Z_{r,i}(t-S_{r,i})-1]\cdots[Z_{r,i}(t-S_{r,i}-k+1)]}{\left[\sum_{i=1}^{Y_r} Z_{r,i}(t-S_{r,i})\right]\left[\sum_{i=1}^{Y_r} Z_{r,i}(t-S_{r,i})-1\right]\cdots\left[\sum_{i=1}^{Y_r} Z_{r,i}(t-S_{r,i})-k+1\right]} \\
= &\frac{\sum_{i=1}^{Y_r} e^{-\alpha(t-S_{r,i})}Z_{r,i}(t-S_{r,i})\cdot e^{-\alpha(t-S_{r,i})}[Z_{r,i}(t-S_{r,i})-1]\cdots e^{-\alpha(t-S_{r,i})}[Z_{r,i}(t-S_{r,i}-k+1)]}{\left[\sum_{i=1}^{Y_r} e^{-\alpha(t-S_{r,i})}Z_{r,i}(t-S_{r,i})\right]\cdot\left[\sum_{i=1}^{Y_r} e^{-\alpha(t-S_{r,i})}Z_{r,i}(t-S_{r,i})\right]\cdots\left[\sum_{i=1}^{Y_r} e^{-\alpha(t-S_{r,i})}Z_{r,i}(t-S_{r,i})\right]}
\end{aligned}
$$

So, for almost all trees $\mathcal{T}$,

$$
P\Big(X_k(t)\geq r\Big|\mathcal{T}\Big)\to\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}}W_{r,i}\right)^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right)^k}\qquad\text{as } t\to\infty,
$$

where $\left\{W_{r,i}\right\}_{i\geq 1}$ are the i.i.d copies of W defined in Theorem 1.11 (b).

Then, by the bounded convergence theorem, for any $r=0,1,2,\cdots$,

$$
P\Big(X_k(t)\geq r\Big)\to E\left(\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}}W_{r,i}\right)^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right)^k}\right)\qquad\text{as } t\to\infty
$$

and hence, for any $r=0,1,2,\cdots$,

$$
\begin{aligned}
P\Big(\tilde{X}_k(t)\geq r\Big) &= E\left(\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}}W_{r,i}\right)^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right)^k}\right) \\
&= E\left(\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}}W_{r,i}\right)^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right)^k}I_{(r\leq U-1)}\right)+E\left(\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}}W_{r,i}\right)^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right)^k}I_{(r\geq U)}\right) \\
&= P(r\leq U-1)+E\left(E\left(\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha S_{r,i}}W_{r,i}\right)^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right)^k}I_{(r\geq U)}\middle|Y_r\right)\right) \\
&= P(r\leq U-1)+E\left(Y_rE\left(\left(\frac{e^{-\alpha S_{r,i}}W_{r,i}}{\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}}\right)^k I_{(r\geq U)}\middle|Y_r\right)\right)
\end{aligned}
$$

where $U = \min\{n \geq 1 : Y_n \geq 2\}$.

Since $p_0 = 0$ and $\sum_{j=1}^{\infty}(j\log j)p_j < \infty$, $P(0 < W_{r,i} < \infty) = 1$ for all $i \geq 0$ and all $r = 0, 1, 2, \cdots$. So, on the set $\{r \geq U\}$,

$$0 < \frac{e^{-\alpha S_{r,i}} W_{r,i}}{\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}} < 1 \quad \text{w.p.1}$$

and hence, for any $r = 0, 1, 2, \cdots$, as $k \to \infty$,

$$\left(\frac{e^{-\alpha S_{r,i}} W_{r,i}}{\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}}\right)^k \to 0 \quad \text{w.p.1.}$$

Therefore, by the bounded convergence theorem again, we have that

$$P\Big(\tilde{X}_k(t) \geq r\Big) \to P\Big(U - 1 \geq r\Big) \quad \text{as } k \to \infty$$

for any $r = 0, 1, 2, \cdots$. So, The proof is complete.

4.2.4 The proof of Theorem 4.3

We need the following lemmas to prove Theorem 4.3.

Let $\xi_{s,i}$ be the number of offsprings of the ith individual alive at time s.

Lemma 4.1. *For any $s \geq 0$, let $\left\{W_{s,i,j} : j \geq 1, i \geq 1\right\}$ be i.i.d. copies of W defined in Theorem 1.11 (b) and be indiependent of $\left\{\xi_{s,i}\right\}_{i\geq 1}$. Let $\tilde{W}_{s,i} = \sum_{j=1}^{\xi_{s,i}} W_{s,i,j}$. Then, under the hypotheses of Theorem 4.3, as $s \to \infty$,*

$$\frac{1}{Z(s)} \max_{1\leq i\leq Z(s)} \tilde{W}_{s,i} \to 0 \quad \textit{in probability.}$$

Proof. In a continuous-time single-type age-dependent Bellman-Harris branching process, $\left\{\xi_{s,i}\right\}_{i\geq 0}$ are i.i.d. copies of the offspring random variable. Also, $\{W_{s,i,j}\}$ are i.i.d. and independet of $\left\{\xi_{s,i}\right\}_{i\geq 0}$, so $\left\{\tilde{W}_{s,i}\right\}_{i\geq 1}$ are i.i.d random variables and

$$E\tilde{W}_{s,i} = E\Bigg(\sum_{j=1}^{\xi_{s,i}} W_{s,i,j}\Bigg) = E\xi_{s,i} \cdot EW_{s,i,1} = m \in (1, \infty). \tag{4.7}$$

Thus, since $E\tilde{W}_{s,i} < \infty$, for any $\epsilon > 0$,

$$nP(\tilde{W}_{s,i} > n\epsilon) \to 0 \quad \text{as } n \to \infty \tag{4.8}$$

and then,

$$\begin{aligned} P\Big(\frac{1}{n}\max_{1\le i\le n}\tilde{W}_{s,i} > \epsilon\Big) &= 1 - P\Big(\max_{1\le i\le n}\tilde{W}_{s,i} \le n\epsilon\Big) \\ &= 1 - P\Big(\tilde{W}_{s,i} \le n\epsilon \text{ for all } i = 1, 2, \cdots, n\Big) \\ &= 1 - \prod_{i=1}^{n} P\Big(\tilde{W}_{s,i} \le n\epsilon\Big) \\ &= 1 - \Big[P\Big(\tilde{W}_{s,1} \le n\epsilon\Big)\Big]^n \\ &= 1 - \Big[1 - \frac{nP\Big(\tilde{W}_{s,1} > n\epsilon\Big)}{n}\Big]^n \to 1 - e^0 = 1 \quad \text{by (4.8)} \end{aligned}$$

as $n \to \infty$.

Therefore, by the bounded convergence theorem, as $s \to \infty$,

$$P\Big(\frac{1}{Z(s)}\max_{1\le i\le Z(s)}\tilde{W}_{s,i} > \epsilon\Big) = E\Big(P\Big(\frac{1}{Z(s)}\max_{1\le i\le Z(s)}\tilde{W}_{s,i} > \epsilon\Big|Z(s)\Big)\Big) \to 0$$

since $P\Big(Z(s) \to \infty \text{ as } s \to \infty\Big) = 1$ under the assumption $p_0 = 0$. Then, Lemma 4.1 is proved.

□

Lemma 4.2. *For any $k > 0$, let $Z(s,k)$ be the number of individuals alive at time s with the residual lifetime less than or equal to k. Then, under the hypotheses of Theorem 4.3, as $s \to \infty$,*

$$\frac{Z(s,k)}{Z(s)} \to B(k) \quad \textit{in probability}$$

where

$$B(k) = \frac{\int_{[0,\infty)} e^{-\alpha x}[G(x+k) - G(x)]dx}{\int_{[0,\infty)} e^{-\alpha x}[1 - G(x)]dx}.$$

Proof. For any fixed $k > 0$, consider a function g such that

$$\begin{aligned} g(a) &\equiv P\Big(R_{s,i} \le k\Big|a_{s,i} = a\Big) = P\Big(L_{s,i} - a_{a,i} \le k\Big|L_{s,i} > a\Big) = P\Big(a < L_{s,i} \le a + k\Big|L_{s,i} > a\Big) \\ &= \frac{P\Big(a < L_{s,i} \le a + k\Big)}{P\Big(L_{s,i} > a\Big)} = \frac{G(a+k) - G(a)}{1 - G(a)}. \end{aligned}$$

Let $\mathfrak{F}_s$ be the $\sigma-$algebra generated all the history of this branching process upto time s.

Then, for any $\epsilon > 0$,

$$
\begin{aligned}
& P\left(\left|\frac{Z(s,k)}{Z(s)} - B(k)\right| > \epsilon\right) \\
= \; & E\left(P\left(\left|\frac{Z(s,k)}{Z(s)} - B(k)\right| > \epsilon \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right)\right) \\
= \; & E\left(P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} I_{(R_{s,i} \le k)} - B(k)\right| > \epsilon \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right)\right) \\
= \; & E\left(P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} \left(I_{(R_{s,i} \le k)} - g(a_{s,i})\right) + \frac{1}{Z(s)} \sum_{i=1}^{Z(s)} \left(g(a_{s,i}) - B(k)\right)\right| > \epsilon \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right)\right) \\
\le \; & E\left(P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} \left(I_{(R_{s,i} \le k)} - g(a_{s,i})\right)\right| > \frac{\epsilon}{2} \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right)\right) \\
& \qquad + E\left(P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} \left(g(a_{s,i}) - B(k)\right)\right| > \frac{\epsilon}{2} \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right)\right) \\
= \; & E\left(P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} \left(I_{(R_{s,i} \le k)} - g(a_{s,i})\right)\right| > \frac{\epsilon}{2} \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right)\right) \\
& \qquad + P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} \left(g(a_{s,i}) - B(k)\right)\right| > \frac{\epsilon}{2}\right) \qquad (4.9)
\end{aligned}
$$

Note that, for any $i \ge 1$,

$$
E\left(I_{(R_{s,i} \le k)} - g(a_{s,i}) \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right) = 0
$$

and

$$
\begin{aligned}
& Var\left(I_{(R_{s,i} \le k)} - g(a_{s,i}) \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right) \\
= \; & E\left(\left|I_{(R_{s,i} \le k)} \le k) - g(a_{s,i})\right|^2 \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right) \\
= \; & E\left(I_{(R_{s,i} \le k)} \le k) \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right) \\
& \qquad - 2g(a_{s,i}) E\left(I_{(R_{s,i} \le k)} \le k) \middle| \mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\right) + \left(g(a_{s,i})\right)^2 \\
= \; & g(a_{s,i}) - 2\left(g(a_{s,i})\right)^2 + \left(g(a_{s,i})\right)^2 = g(a_{s,i}) - \left(g(a_{s,i})\right)^2 \le \frac{1}{4}.
\end{aligned}
$$

Conditioned on $\mathfrak{F}_s$ and $(a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})$, we have that $\{I_{(R_{s,i} \le k) - g(a_{s,i})} : i = 1, 2, \cdots, Z(s)\}$ are independent. So, by Chebycheve's inequality, for any $\epsilon > 0$,

$$
\begin{aligned}
& P\Bigg(\bigg|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)}\Big(I_{(R_{s,i}\le k)} - g(a_{s,i})\Big)\bigg| > \frac{\epsilon}{2}\bigg|\mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\Bigg) \\
\le\;& \frac{4}{\epsilon^2} Var\Bigg(\frac{1}{Z(s)}\sum_{i=1}^{Z(s)}\Big(I_{(R_{s,i}\le k)} - g(a_{s,i})\Big)\bigg|\mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\Bigg) \\
=\;& \frac{4}{\epsilon^2}\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} Var\Big(I_{(R_{s,i}\le k)} - g(a_{s,i})\Big|\mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\Big) \\
\le\;& \frac{4}{\epsilon^2}\frac{1}{Z(s)}\sum_{i=1}^{Z(s)}\frac{1}{4} \\
=\;& \frac{1}{\epsilon^2} Z(s) \to 0 \quad \text{as } s \to \infty
\end{aligned}
$$

and hence

$$
P\Bigg(\bigg|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)}\Big(I_{(R_{s,i}\le k)} - g(a_{s,i})\Big)\bigg| > \frac{\epsilon}{2}\bigg|\mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\Bigg) \to 0 \quad \text{w.p.1}
$$

as $s \to \infty$. Then, by the bounded convergence theorem,

$$
E\Bigg(P\Bigg(\bigg|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)}\Big(I_{(R_{s,i}\le k)} - g(a_{s,i})\Big)\bigg| > \frac{\epsilon}{2}\bigg|\mathfrak{F}_s, (a_{s,1}, a_{s,2}, \cdots, a_{s,Z(s)})\Bigg)\Bigg) \to 0 \quad \text{as } s \to \infty. \tag{4.10}
$$

It remains to prove that

$$
P\Bigg(\bigg|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)}\Big(g(a_{s,i}) - B(k)\Big)\bigg| > \frac{\epsilon}{2}\Bigg) \to 0 \quad \text{as } s \to \infty.
$$

Let $A(x, s) = \frac{1}{Z(s)}\sum_{i=1}^{Z(s)} I_{(a_{s,i}\le x)}$. Then, by Theorem 1.14, as $s \to \infty$

$$
\sup_x \Big|A(x, s) - A(x)\Big| \to 0 \quad \text{w.p.1.}
$$

where A is as defined in Section 1.4.

Since $g(x) = \dfrac{G(x+k) - G(x)}{1 - G(x)}$ is a bounded and continuous function, by Theorem 1.14 again, we have, as $s \to \infty$

$$
\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} g(a_{s,i}) \equiv \int_{[0,\infty)} g(x)dA(x, s) \to \int_{[0,\infty)} g(x)dA(x) \quad \text{w.p.1}
$$

where

$$
\int_{[0,\infty)} g(x)dA(x) = \frac{\int_0^\infty \frac{G(x+k)-G(x)}{1-G(x)} e^{-\alpha x}(1 - G(x))dx}{\int_0^\infty e^{-\alpha x}(1 - G(x))dx} = \frac{\int_0^\infty e^{-\alpha x}(G(x + k) - G(x))dx}{\int_0^\infty e^{-\alpha x}(1 - G(x))dx} = B(k).
$$

So, as $s \to \infty$, $\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} g(a_{s,i}) \to B(k)$ w.p.1 and hence in probability. Thereofore, for any $\epsilon > 0$,

$$P\left(\left|\frac{1}{Z(s)} \sum_{i=1}^{Z(s)} g(a_{s,i}) - B(k)\right| > \frac{\epsilon}{2}\right) \to 0 \quad \text{as } s \to \infty. \tag{4.11}$$

From (4.9), (4.10) and (4.11), we have that, for any $\epsilon > 0$,

$$P\left(\left|\frac{Z(s,k)}{Z(s)} - B(k)\right| > \epsilon\right) \to 0 \quad \text{as } s \to \infty$$

and the proof is complete.

□

Lemma 4.3. *Let $\tilde{W}_{s,i}$ and $Z(s,k)$ be the random variables defined in Lemma 4.1 and Lemma 4.2, respectively. Then, under the hypotheses of Theorem 4.3, these exists a $\theta > 0$ such that, as $s \to \infty$,*

$$P\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \tilde{W}_{s,k} I_{(R_{s,i} \le k)} \ge \theta\right) \to 1.$$

Proof. Let $n_{s,1} = \min\left\{1 \le j \le Z(s) : R_{s,j} \le k\right\}$ and $n_{s,i} = \min\left\{n_{s,i-1} < j \le Z(s) : R_{s,j} \le k\right\}$ for $i \ge 2$. Then

$$\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i} \le k)} = \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \tilde{W}_{s,n_{s,i}} I_{(R_{s,n_{s,i}} \le k)} = \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \tilde{W}_{s,n_{s,i}}. \tag{4.12}$$

It is known from (4.7) that $E\tilde{W}_{s,1} > 0$ and hence there exists an $\eta > 0$ such that $P\left(\tilde{W}_{s,1} \ge \eta\right) > 0$.

Let $\mathfrak{F}_s$ be the $\sigma-$algebra generated by all the information of this Bellman-Harris branching process upto time s. Then

$$\begin{aligned}
P\left(\tilde{W}_{s,n_{s,i}} \ge \eta\right) &= E\left(P\left(\tilde{W}_{s,n_{s,i}} \ge \eta \middle| \mathfrak{F}_s\right)\right) = E\left(\sum_{j=1}^{Z(s,k)} P\left(\tilde{W}_{s,n_{s,i}} \ge \eta, n_{s,i} = j \middle| \mathfrak{F}_s\right)\right) \\
&= E\left(\sum_{j=1}^{Z(s,k)} P\left(\tilde{W}_{s,j} \ge \eta, n_{s,i} = j \middle| \mathfrak{F}_s\right)\right) \\
&= E\left(\sum_{j=1}^{Z(s,k)} P\left(\tilde{W}_{s,j} \ge \eta \middle| \mathfrak{F}_s\right) P\left(n_{s,i} = j \middle| \mathfrak{F}_s\right)\right) \\
&= E\left(\sum_{j=1}^{Z(s,k)} P\left(\tilde{W}_{s,j} \ge \eta \middle| \mathfrak{F}_s\right) P\left(n_{s,1} = j \middle| \mathfrak{F}_s\right)\right) \\
&= P\left(\tilde{W}_{s,1} \ge \eta\right).
\end{aligned}$$

Let

$$X_{s,i} = \begin{cases} 1 & , \text{if } \tilde{W}_{s,n_{s,i}} \geq \eta \\ 0 & . \text{ if } \tilde{W}_{s,n_{s,i}} < \eta \end{cases}$$

then

$$\begin{aligned} \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \tilde{W}_{s,n_{s,i}} &\geq \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \eta X_{s,i} \\ &= \eta\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\right) + \eta\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right) \\ &= \eta\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\right) + \eta P(\tilde{W}_{s,n_{s,i}} \geq \eta). \end{aligned} \tag{4.13}$$

Conditioned on $\mathfrak{F}_s$, $\left\{X_{s,i}\right\}_{i\geq 1}$ are indipendent with $E\left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\middle|\mathfrak{F}_s\right) = 0$ and

$$\begin{aligned} Var\left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\middle|\mathfrak{F}_s\right) &= P(\tilde{W}_{s,n_{s,i}} \geq \eta|\mathfrak{F}_s)\left(1 - P(\tilde{W}_{s,n_{s,i}} \geq \eta|\mathfrak{F}_s)\right) \\ &= P(\tilde{W}_{s,1} \geq \eta|\mathfrak{F}_s)\left(1 - P(\tilde{W}_{s,1} \geq \eta|\mathfrak{F}_s)\right) \\ &\leq \frac{1}{4}, \end{aligned}$$

then, by Chebychev's inequality and Lemma 4.2, for any $\epsilon > 0$,

$$\begin{aligned} & P\left(\left|\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\right| > \epsilon\middle|\mathfrak{F}_s\right) \\ \leq\ & \frac{1}{\epsilon^2} Var\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\middle|\mathfrak{F}_s\right) \\ =\ & \frac{1}{\epsilon^2} \frac{1}{Z(s,k)^2} Var\left(\left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\middle|\mathfrak{F}_s\right) \\ \leq\ & \frac{1}{4\epsilon^2 Z(s,k)} \\ =\ & \frac{1}{4\epsilon^2 Z(s)} \frac{Z(s)}{Z(s,k)} \to 0 \quad \text{in probability} \quad \text{as } s \to \infty. \end{aligned}$$

Therefore, by the bounded convergence theorem,

$$\begin{aligned} & P\left(\left|\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\right| > \epsilon\right) \\ =\ & E\left(P\left(\left|\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right)\right| > \epsilon\middle|\mathfrak{F}_s\right)\right) \to 0 \quad \text{as } s \to \infty. \end{aligned}$$

Hence,

$$P\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\right) < -\epsilon\right) \to 0 \quad \text{as } s \to \infty. \tag{4.14}$$

Let $\theta = \frac{1}{2}\eta P(\tilde{W}_{s,1} \geq \eta)$, then $\theta > 0$. Also, (4.12), (4.13) and (4.14) together imply that

$$\begin{aligned}
& P\Big(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \tilde{W}_{s,i} I_{(R_{s,i} \leq k)} \geq \theta\Big) \\
\geq \quad & P\Big(\eta\Big(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \big(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\big)\Big) + \eta P(\tilde{W}_{s,n_{s,i}} \geq \eta) \geq \frac{1}{2}\eta P(\tilde{W}_{s,1} \geq \eta)\Big) \\
= \quad & P\Big(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \big(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\big) \geq -\frac{1}{2} P(\tilde{W}_{s,1} \geq \eta)\Big) \\
= \quad & 1 - P\Big(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \big(X_{s,i} - P(\tilde{W}_{s,n_{s,i}} \geq \eta)\big) < -\frac{1}{2} P(\tilde{W}_{s,1} \geq \eta)\Big) \to 1 \quad \text{as } s \to \infty.
\end{aligned}$$

So, we have that

$$P\Big(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \tilde{W}_{s,i} I_{(R_{s,i} \leq k)} \geq \theta\Big) \to 1 \text{ as } s \to \infty$$

and hence Lemma 4.3 is proved.

□

Now, we are ready to prove Theorem 4.3.

Consider a discrete-time single-type Bellman-Harris branching process $\{Z(t) : t \geq 0\}$ with $Z(0) = 1$.

Recall the following notations: For any $i = 1, 2, \cdots, Z(s)$,

(1) $\xi_{s,i}$ is the number of offsprings of the ith individual alive at time s;

(2) $L_{s,i}$ is the corresponding total lifetime of the ith individual alive at time s.

(3) $a_{s,i}$ be the corresponding age;

(4) $R_{s,i}$ be the corresponding residual lifetime at time t.

Let $\tilde{Z}_{t-s-R_{s,i},j}$ be the branching process initiated by the jth offspring of the ith individual alive at time s.

Pick k individuals randomly from those alive at time t and trace their lines of decent backward in time until they meet. Denote the coalescence time $D_k(t)$ which also means the death time of the last common ancestor of these randomly chosen individuals.

For almost all trees $\mathcal{T}$ and $s \geq 0$,

$$\begin{aligned}
& P\Big(D_k(t) \leq s \Big| \mathcal{T}\Big) \\
= \quad & 1 - P\Big(D_k(t) > s \Big| \mathcal{T}\Big) \\
= \quad & 1 - \frac{\sum_{i=1}^{Z(s)} \Big(\sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j}\Big)\Big(\sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j} - 1\Big)\cdots\Big(\sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j} - k + 1\Big)}{\Big(\sum_{i=1}^{Z(s)}\sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j}\Big)\Big(\sum_{i=1}^{Z(s)}\sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j} - 1\Big)\cdots\Big(\sum_{i=1}^{Z(s)}\sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j} - k + 1\Big)} \\
= \quad & 1 - \frac{\sum_{i=1}^{Z(s)} \prod_{l=1}^{k} \Big(e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j} e^{-\alpha(t-s-R_{s,i})} - (l-1)e^{-\alpha(t-s-R_{s,i})}\Big)}{\prod_{l=1}^{k} \Big(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} \tilde{Z}_{t-s-R_{s,i},j} e^{-\alpha(t-s-R_{s,i})} - (l-1)e^{-\alpha(t-s-R_{s,i})}\Big)}
\end{aligned}$$

and then, by Theroem 1.11,

$$\begin{aligned}
P\Big(D_k(t) \leq s \Big| \mathcal{T}\Big) \quad \to \quad & 1 - \frac{\sum_{i=1}^{Z(s)} \Big(e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} W_{s,i,j}\Big)^k}{\Big(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} W_{s,i,j}\Big)^k} \quad \text{as } t \to \infty \\
= \quad & 1 - \frac{\sum_{i=1}^{Z(s)} \Big(e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k}{\Big(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k} \equiv H_k(s, \mathcal{T})
\end{aligned}$$

where $\{W_{s,i,j}\}_{j \geq 1}$ are i.i.d copies of W in Theorem 1.11 and $\tilde{W}_{s,i} \equiv \sum_{i=1}^{\xi_{s,i} W_{s,i,j}}$ for $i \geq 1$ and $s \geq 0$.

So, by the bounded convergence theorem, as $t \to \infty$,

$$P\Big(D_k(t) \leq s\Big) = E\Big(P\Big(D_k(t) \leq s \Big| \mathcal{T}\Big)\Big) \to 1 - E\left(\frac{\sum_{i=1}^{Z(s)} \Big(e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k}{\Big(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k}\right) \equiv H_k(s).$$

Next, we need to show that H_k is a proper probability distribution, i.e. show that $H_k(s) \to 1$ as $s \to \infty$ and it is the same as showing that

$$E\left(\frac{\sum_{i=1}^{Z(s)} \Big(e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k}{\Big(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k}\right) \to 0 \quad \text{as } s \to \infty.$$

It is suffices to prove that, as $s \to \infty$,

$$\frac{\sum_{i=1}^{Z(s)} \Big(e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k}{\Big(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \tilde{W}_{s,i}\Big)^k} \to 0 \quad \text{in probability.}$$

Moreover, since

$$\left(\frac{\max\limits_{1\le i\le Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}\right)^k \le \frac{\sum\limits_{i=1}^{Z(s)}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k} \le \left(\frac{\max\limits_{1\le i\le Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}\right)^{k-1},$$

it is enough to show that, as $s \to \infty$,

$$\frac{\max\limits_{1\le i\le Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}} \to 0 \quad \text{in probability.}$$

For any fixed $k > 0$,

$$\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i} \ge e^{-\alpha k}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\le k)}$$

and then

$$\frac{\max\limits_{1\le i\le Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}} \le \frac{\max\limits_{1\le i\le Z(s)} \tilde{W}_{s,i}}{e^{-\alpha k}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\le k)}} = \frac{e^{\alpha k}\frac{1}{Z(s)}\max\limits_{1\le i\le Z(s)} \tilde{W}_{s,i}}{\frac{Z(s,k)}{Z(s)}\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\le k)}} \tag{4.15}$$

where $Z(s,k)$ is the number of individuals alive at time s with the residual lifetime less than or equal to k.

From Lemma 4.2, we know that, as $s \to \infty$,

$$\frac{Z(s,k)}{Z(s)} \to B(k) \quad \text{in probability}$$

and hence

$$P\left(\frac{Z(s,k)}{Z(s)} < \frac{1}{2}B(k)\right) \to 0 \quad \text{as } s \to \infty.$$

Also, from Lemma 4.3, we have that, for some $\theta > 0$, as $s \to \infty$,

$$P\left(\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s,k)} \tilde{W}_{s,k} I_{(R_{s,i}\le k)} \ge \theta\right) \to 1.$$

So, for any $\delta > 0$, there exists an $M > 0$ such that for every $s > M$,

$$P\left(\frac{Z(s,k)}{Z(s)} < \frac{1}{2}B(k)\right) < \frac{\delta}{2}$$

and

$$P\left(\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s,k)} \tilde{W}_{s,k} I_{(R_{s,i}\le k)} < \theta\right) < \frac{\delta}{2}.$$

Let $A = \left\{\frac{Z(s,k)}{Z(s)} \geq \frac{1}{2}B(k)\right\}$ and $B = \left\{\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s,k)} \tilde{W}_{s,k} I_{(R_{s,i}\leq k)} \geq \theta\right\}$. Then, for any $\epsilon > 0$,

$$\begin{aligned}
& P\Bigg(\frac{e^{\alpha k}\frac{1}{Z(s)}\max\limits_{1\leq i\leq Z(s)} \tilde{W}_{s,i}}{\frac{Z(s,k)}{Z(s)}\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\leq k)}} > \epsilon\Bigg) \\
= \; & P\Bigg(\frac{1}{Z(s)}\max_{1\leq i\leq Z(s)} \tilde{W}_{s,i} > \epsilon e^{-\alpha k}\frac{Z(s,k)}{Z(s)}\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\leq k)}\Bigg) \\
\leq \; & P\Bigg(\frac{1}{Z(s)}\max_{1\leq i\leq Z(s)} \tilde{W}_{s,i} > \epsilon e^{-\alpha k}\frac{1}{2}B(k)\theta : A\cap B\Bigg) + P(A^C) + P(B^C) \\
\leq \; & P\Bigg(\frac{1}{Z(s)}\max_{1\leq i\leq Z(s)} \tilde{W}_{s,i} > \frac{1}{2}\epsilon\theta e^{-\alpha k}B(k)\Bigg) + \delta
\end{aligned}$$

for every $s > M$. Thus, for any $\delta > 0$,

$$\begin{aligned}
& \limsup_{s\to\infty} P\Bigg(\frac{e^{\alpha k}\frac{1}{Z(s)}\max\limits_{1\leq i\leq Z(s)} \tilde{W}_{s,i}}{\frac{Z(s,k)}{Z(s)}\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\leq k)}} > \epsilon\Bigg) \\
\leq \; & \limsup_{s\to\infty} P\Bigg(\frac{1}{Z(s)}\max_{1\leq i\leq Z(s)} \tilde{W}_{s,i} > \frac{1}{2}\epsilon\theta e^{-\alpha k}B(k)\Bigg) + \delta = \delta
\end{aligned}$$

i.e.,

$$\lim_{s\to\infty} P\Bigg(\frac{e^{\alpha k}\frac{1}{Z(s)}\max\limits_{1\leq i\leq Z(s)} \tilde{W}_{s,i}}{\frac{Z(s,k)}{Z(s)}\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s)} \tilde{W}_{s,i} I_{(R_{s,i}\leq k)}} > \epsilon\Bigg) = 0 \quad \text{for any } \epsilon > 0$$

and hence, from (4.15), we have that

$$\lim_{s\to\infty} P\Bigg(\frac{\max\limits_{1\leq i\leq Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}} > \epsilon\Bigg) = 0 \quad \text{for any } \epsilon > 0,$$

i.e., as $s \to \infty$,

$$\frac{\max\limits_{1\leq i\leq Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}} \to 0 \quad \text{in probability.}$$

By the bounded convergence theorem,

$$E\Bigg(\frac{\sum\limits_{i=1}^{Z(s)}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}\Bigg) \to 0 \quad \text{as } s \to \infty.$$

and thus H_k is a proper probability distribution on $\mathbb{R}_+ \equiv \{x \in \mathbb{R} : x \geq 0\}$. So, there exists a random variable $\tilde{D}_k$ on $\mathbb{R}_+$ such that $D_k(t) \xrightarrow{d} \tilde{D}_k$ as $t \to \infty$ and

$$P(\tilde{D}_k \leq s) = 1 - E\Bigg(\frac{\sum\limits_{i=1}^{Z(s)}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum\limits_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}\Bigg) \equiv H_k(s).$$

for any $s \geq 0$. The proof of Theorem 4.3 is complete.

4.2.5 The proof of Theorem 4.4

Recall that $\{Y_n\}_{n\geq 0}$ is the embedded generation process of this Bellman-Harris process and $U = \min\left\{n \geq 1 : Y_j \geq 2\right\}$ which is the first generation with more than one individuals.

Let Ls, i, j be the lifetime of the ancestor in the jth generation of the ith individual alive at time s. Then $\left\{L_{s,i,j}\right\}_{j\geq 1}$ are i.i.d. random variables with the lifetime distribution G for any $i = 1, 2, \cdots, Z(s)$ and $s > 0$.

From Theorem 4.3, for any $s > 0$, we have that

$$
\begin{aligned}
P\Big(\tilde{D}_k(t) > s\Big) &= E\left(\frac{\sum_{i=1}^{Z(s)}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}\right) \\
&= E\left(\frac{\sum_{i=1}^{Z(s)}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k} I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U-1}>s)}\right) + E\left(\frac{\sum_{i=1}^{Z(s)}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k} I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U-1}\leq s)}\right) \\
&= P\Big(L_{s,i,0} + L_{s,i,1} + \cdots + L_{s,i,U-1} > s\Big) + E\left(E\left(\frac{\sum_{i=1}^{Y_r}\left(e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}\right)^k} I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U-1}\leq s)}\middle| Z(s)\right)\right) \\
&= P\Big(L_0 + L_1 + \cdots + L_2 > s\Big) + E\left(\sum_{i=1}^{Z(s)} E\left(\left(\frac{e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}\right)^k I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U-1}\leq s)}\middle| Z(s)\right)\right)
\end{aligned}
$$

where $\left\{L_i\right\}_{i\geq 0}$ are i.i.d random variables with the lifetime distribution.

Since $p_0 = 0$, $1 < m < \infty$, $\sum_{j=1}^{\infty}(j\log j)p_j < \infty$ and $\tilde{W}_{s,i} = \sum_{j=1}^{\xi_{s,i}} W_{s,i,j}$, we have that $P(0 < \tilde{W}_{s,i} < \infty) = 1$ for all $i \geq 0$ and $s > 0$. So, on the set $\{L_{s,i,0} + L_{s,i,1} + \cdots + L_{s,i,U-1} \leq s\}$,

$$0 < \frac{e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}} < 1 \quad \text{w.p.1}$$

and hence, on the set $\{L_{s,i,0} + L_{s,i,1} + \cdots + L_{s,i,U-1} \leq s\}$, for any $s > 0$, as $k \to \infty$,

$$\left(\frac{e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\tilde{W}_{s,i}}\right)^k \to 0 \quad \text{w.p.1}.$$

Therefore, by the bounded convergence theorem again, we have that

$$P\Big(\tilde{D}_k(t) > r\Big) \to P\Big(L_0 + L_1 + \cdots + L_{U-1} > s\Big) \quad \text{as } k \to \infty$$

for any $s > 0$. So, The proof is complete.

4.3 Results in The Subcritical Case

4.3.1 The statements of Results

The first result we establish for the subcritical case is the convergence of the age chart of the population.

Let $a_{t,i}$ be the age of the ith individual alive at time t, $i = 1, 2, \cdots, Z(t)$.

Recall that $f(s) = \sum_{j=0}^{\infty} p_j s^j$, $0 \le s \le 1$, is the probability generating function of the offspring distribution $\{p_j\}_{j\ge 0}$ of the process $\{Z(t)\}$.

For any continuous and bounded function $h : \mathbb{R}_+ \to \mathbb{R}_+$, let

$$H_h(s,t) = E\Big(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\Big).$$

Theorem 4.5. *Let $0 < m < 1$ and $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$. Assume that the lifetime distribution G is non-lattice, $G(0+) = 0$ and such that the Malthusian parameter α exists and $\int_0^{\infty} te^{-\alpha t} dG(t) < \infty$. If*

$$\sup_n \sup_{\substack{n \le t < n+1 \\ 0 \le u \le n}} \left| \frac{1 - f(H_h(s,t-u)) - m(1 - H_h(s,t-u))}{1 - f(H_h(s,n-u)) - m(1 - H_h(s,n-u))} \right| < \infty$$

for any $s \ge 0$ and any nonnegative, bounded and continuous function h on $\mathbb{R}_+$. Then, there exists random variables $\{\tilde{a}_i\}_{i\ge 1}$ such that, conditioned on the event $\{Z(t) > 0\}$, the point process

$$A(t) \equiv \Big\{a_{t,i} : 1 \le i \le Z(t)\Big\}$$

converges in distribution, as $t \to \infty$, to the point process

$$\tilde{A} \equiv \Big\{\tilde{a}_i : 1 \le i \le Y\Big\}$$

where Y is the random variable with the distribution $\{b_j\}_{j\ge 0}$ as defined in Theorem 1.13.

Theorem 4.6. *Let* $0 < m < 1$ *and* $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$. *Assume that the lifetime distribution G is non-lattice,* $G(o+) = 0$ *and such that the Malthusian parameter* α *exists and* $\int_0^{\infty} te^{-\alpha t} dG(t) < \infty$. *If*

$$\sup_n \sup_{\substack{n \le t < n+1 \\ 0 \le u \le n}} \left| \frac{1 - f(H_h(s, t-u)) - m(1 - H_h(s, t-u))}{1 - f(H_h(s, n-u)) - m(1 - H_h(s, n-u))} \right| < \infty$$

for any $s \ge 0$ *and any nonnegative, bounded and continuous function h on* $\mathbb{R}_+$. *Then, there exists* $\tilde{D}_2$ *on the set of non-negative real numbers such that*

$$t - D_2(t) \xrightarrow{d} \tilde{D}_2 \quad \text{as } t \to \infty,$$

and, for any $u \ge 0$,

$$P(\tilde{D}_2 \le s) = 1 - \frac{1}{e^{\alpha u} P(Y \ge 2)} E\Big(\phi(\tilde{A}, u)\Big) \equiv H_2(u)$$

where Y is the random variable with the distribution $\{b_j\}_{j \ge 0}$ *as defined in Theorem 1.13,*

$$\phi((a_1, a_2, \cdots, a_k), u) = E\left(\frac{\sum_{i \ne j = 1}^{k} \tilde{Z}_i(a_i + u)\tilde{Z}_j(a_j + u)}{\Big(\sum_{i=1}^{k} \tilde{Z}_i(a_i + u)\Big)\Big(\sum_{i=1}^{k} \tilde{Z}_i(a_i + u) - 1\Big)} I_{\left(\sum_{i=1}^{k} \tilde{Z}_i(a_i+u) \ge 2\right)} \right)$$

for any positive integer k and any positive real numbers $a_1, a_2, \cdots, a_k$ *and* $\{\tilde{Z}_i(t)\}_{i \ge 1}$ *are i.i.d. copies of* $Z(t)$.

By the same lines of the proof of Theorem 4.6, we can extend the result to any integer $k \ge 2$.

Corollary 4.1. *Let* $0 < m < 1$ *and* $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$. *Then under the same hypotheses in Theorem 4.6, for any* $k \ge 2$, *there exists* $\tilde{D}_k$ *on the set of non-negative real numbers such that* $t - D_k(t) \xrightarrow{d} \tilde{D}_k$ *as* $t \to \infty$.

4.3.2 The proof of Theorem 4.5

Let $Z(t)$ be the continuous-time single-type age-dependent Bellman-Harris braching process with $Z(0) = 1$.

Let $a_{t,i}$ be the age of the ith individual alive at time t, $i = 1, 2, \cdots, Z(t)$.

Let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be any bounded and continuous function.

For any $s \geq 0$, we consider the Laplace functional of $(a_{t,i}, a_{t,2}, \cdots, a_{t,Z(t)})$ conditioned on the set $\{Z(t) > 0\}$, then

$$\begin{aligned}
E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\Big|Z(t) > 0\right) &= \frac{1}{P(Z(t) > 0)} E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})} I_{(Z(t)>0)}\right) \\
&= \frac{1}{P(Z(t) > 0)}\left[E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right) - E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})} I_{(Z(t)=0)}\right)\right] \\
&= \frac{1}{P(Z(t) > 0)}\left[E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right) - E\left(I_{(Z(t)=0)}\right)\right] \\
&= \frac{1}{P(Z(t) > 0)}\left[E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right) - 1 + 1 - P\left(Z(t) = 0\right)\right] \\
&= \frac{1}{P(Z(t) > 0)}\left[E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right) - 1 + P\left(Z(t) > 0\right)\right] \\
&= \frac{1}{P(Z(t) > 0)}\left[E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right) - 1\right] + 1. \qquad (4.16)
\end{aligned}$$

Let $H(s,t) = E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right)$.

Recall that $f(s) = \sum_{j=0}^{\infty} p_j s^j$ is the probability generating function of the offspring distribution $\{p_j\}_{j\geq 0}$.

Let ξ is the number of offspring of an individual in the process. Note that $\xi \sim \{p_j\}_{j\geq 0}$.

Let L_0 be the total lifetime of the first ancestor in this process. So, $L_0 \sim G$.

Let $\{Z_j(t) : t \geq 0\}$ be the Bellman-Harris branching process initiated by the jth individual in the first generation.

Since $Z(t) = \sum_{j=1}^{\xi} Z_j(t - L_0)$, we have

$$\begin{aligned}
E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\Big|L_0 \leq t, \xi\right) &= E\left(e^{-s\sum_{j=1}^{\xi}\sum_{i=1}^{Z_j(t-L_0)} h(a_{t,j,i})}\Big|L_0 \leq t, \xi\right) \\
&= \left(E\left(e^{-s\sum_{i=1}^{Z(t-L_0)} h(a_{t,i})}\Big|L_0 \leq t\right)\right)^{\xi} \\
&= \left(H(s, t - L_0)\right)^{\xi}
\end{aligned}$$

where $a_{t,j,i}$ is the age (at time t) of the ith individual from the tree initiated by the jth individual in the first generation.

Hence,

$$E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\right) = E\left(\left(H(s, t - L_0)\right)^{\xi}\right) = f(H(s, t - L_0)).$$

Therefore, we have that

$$\begin{aligned} H(s,t) &= E\Big(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\Big) \\ &= E\Big(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})} : L_0 > t\Big) + E\Big(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})} : L_o \le t\Big) \\ &= e^{-sh(t)} P\big(L_0 > t\big) + \int_{[0,t]} f(H(s,t-u))dG(u) \\ &= e^{-sh(t)}\big(1 - G(t)\big) + \int_{[0,t]} f(H(s,t-u))dG(u). \end{aligned} \tag{4.17}$$

and it implies that, for any $s \ge 0$, $H(s,t)$ satisfies the integral equation:

$$(*)\begin{cases} H(s,t) = e^{-sh(t)}\big(1 - G(t)\big) + \int_{[0,t]} f(H(s,t-u))dG(u) \\ H(s,0) = e^{-sh(0)}. \end{cases}$$

Moreover,

$$H(\infty,t) \equiv \lim_{s\to\infty} H(s,t) = P\big(Z(t) = 0\big).$$

Then, by (4.16) and (4.17),

$$\begin{aligned} & E\Big(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\Big| Z(t) > 0\Big) = 1 - \frac{1}{P(Z(t)>0)}\Big[1 - H(s,t)\Big] \\ =\ & 1 - \frac{1}{P(Z(t)>0)}\Big[1 - e^{-sh(t)}\big(1 - G(t)\big) + \int_{[0,t]} f(H(s,t-u))dG(u)\Big] \\ =\ & 1 - \frac{1}{P(Z(t)>0)}\Big[\big(1 - e^{-sh(t)}\big)\big(1 - G(t)\big) + \int_{[0,t]} \Big[1 - f(H(s,t-u))\Big]dG(u)\Big] \end{aligned}$$

For any fixed $s \ge 0$, let

$$H(t) = 1 - H(s,t) \tag{4.18}$$

$$\xi_1(t) = \big(1 - e^{-sh(t)}\big)\big(1 - G(t)\big) \tag{4.19}$$

$$\xi_2(t) = \int_0^t \Big[1 - f(H(s,t-u)) - mH(t-u)\Big]dG(u) \tag{4.20}$$

$$\xi_3(t) = \xi_1(t) + \xi_2(t) \tag{4.21}$$

and then

$$H(t) = \xi_3(t) + m\int_{[0,t]} H(t-u)dG(u). \tag{4.22}$$

To complete the proof of Theorem 4.5, we need the following definition and lemmas.

Definition 4.1. Afunction ξ is directly Riemann integrable if

(a) $\sum_{n=0}^{\infty}\delta\Big(\sup_{n\delta\le t<(n+1)\delta}\xi(t)\Big)$ and $\sum_{n=0}^{\infty}\delta\Big(\inf_{n\delta\le t<(n+1)\delta}\xi(t)\Big)$ converge absolutely for sufficient small $\delta>0$; and

(b) $\delta\bigg(\sum_{n=0}^{\infty}\delta\Big(\sup_{n\delta\le t<(n+1)\delta}\xi(t)\Big)-\sum_{n=0}^{\infty}\delta\Big(\inf_{n\delta\le t<(n+1)\delta}\xi(t)\Big)\bigg)\to 0$ as $\delta\to 0$.

Remark 4.1. Some sufficient conditions for direct Riemann integrability of ξ are

(1) $\xi\ge 0$, bounded, continuous and $\sum_{n=0}^{\infty}\Big(\sup_{n\le t<n+1}\xi(t)\Big)<\infty$;

(2) $\xi\ge 0$, non-increasing and Riemann integrable in the ordinary sense;

(3) ξ is bounded by a directly Riemann integrable function;

(4) ξ is constant on the intervals $(n, n+1)$ and absolutely integrable.

Lemma 4.4 is a well-known result in the renewal theory. See Feller [17].

Lemma 4.4. *Let G be a probability distribution function and G^{*n} denote its $n-$fold convolution. Let $U=\sum_{n=0}^{\infty}G^{*n}$. If ξ is directly Riemann integrable and G is non-lattice, then*

$$\lim_{t\to\infty}(\xi * U)(t)=\frac{\int_0^{\infty}\xi(u)du}{\int_0^{\infty}udG(u)}.$$

Lemma 4.5. *If the Mathusian parameter α of m and G exists, if $e^{-\alpha t}\xi(t)$ is directly Riemann integrable, and if G us non-lattice, then the solution H of the integral equation*

$$H(t)=\xi(t)+m\int_0^t H(t-u)dG(u),\quad t\ge 0$$

satisfies

$$H(t)\sim\frac{\int_0^{\infty}e^{-\alpha u}\xi(t)du}{m\int_0^{\infty}ue^{-\alpha u}dG(u)}.$$

The proof of Lemma 4.5 can be found in Athreya and Ney [5].

Lemma 4.6. *Let H be the function defined in (4.18). Then, under the hypotheses of Theorem 4.5,*

$$\sup_{s,t\ge 0}e^{-\alpha t}H(t)<\infty.$$

Proof. For any fixed $s \geq 0$ and for any $t \geq 0$, we have

$$\begin{aligned}
\left|H(t)\right| &= \left|1 - H(s,t)\right| \\
&= \left|\left(1 - e^{-sh(t)}\right)\left(1 - G(t)\right) + \int_{[0,t]} \left[1 - f(H(s,t-u))\right] dG(u)\right| \\
&\leq \left|1 - e^{-sh(t)}\right|\left|1 - G(t)\right| + \left|\int_{[0,t]} \left[1 - f(H(s,t-u))\right] dG(u)\right|.
\end{aligned}$$

Note that $f(1) = 1$, $0 < H(s,t-u) < 1$ and f is a continuous function. Then, by the mean value theorem, there exists c such that $H(s,t-u) < c < 1$ and

$$f'(c) = \frac{f(1) - f(H(s,t-u))}{1 - H(s,t-u)}. \tag{4.23}$$

Therefore,

$$\begin{aligned}
\left|H(t)\right| &\leq \left|1 - e^{-sh(t)}\right|\left|1 - G(t)\right| + \left|\int_{[0,t]} f'(c)\left(1 - H(s,t-u)\right) dG(u)\right| \\
&\leq \left|1 - G(t)\right| + \int_{[0,t]} \left|f'(c)\right|\left|1 - H(s,t-u)\right| dG(u) \\
&\leq \left(1 - G(t)\right) + m \int_{[0,t]} \left|H(t-u)\right| dG(u). \qquad (4.24)
\end{aligned}$$

since f' is non-decreasing.

Let $me^{-\alpha t} dG(t) = dG_\alpha(t)$ and $g_\alpha(t) \equiv e^{-\alpha t}\left(1 - G(t)\right)$.

Note that $\int_0^\infty te^{-\alpha t} dG(t) < \infty$ imples the Riemann integrability of g_α.

So, $g_\alpha \geq 0$ is non-increasing and Riemann-integrable, and hence g_α is directly Riemann integrable by condition (2) in Remark 4.1.

Moreover, that G is non-lattice implies that G_α is also non-lattice.

Let G_α^{*n} be the $n-$fold convolution of G_α and $U_\alpha = \sum_{n=0}^{\infty} G_\alpha^{*n}$. Then, by Lemma 4.4, we have that

$$\lim_{t \to \infty} g_\alpha * U_\alpha(t) = \frac{\int_0^\infty g_\alpha(u) du}{\int_0^\infty u dG_\alpha(u)} < \infty.$$

Multiply both sides of (4.24) by $e^{-\alpha t}$ and , then

$$\begin{aligned} e^{-\alpha t}\big|H(t)\big| &\leq e^{-\alpha t}\Big(1 - G(t)\Big) + m\int_{[0,t]} e^{-\alpha t}\big|H(t-u)\big|dG(u) \\ &= g_\alpha(t) + \int_{[0,t]} e^{-\alpha(t-u)}\big|H(t-u)\big|dG_\alpha(u) \\ &= g_\alpha(t) + H_\alpha * G_\alpha(t) \\ &\leq g_\alpha(t) + (g_\alpha + H_\alpha * G_\alpha) * G_\alpha(t) \\ &= \cdots \\ &= g_\alpha(t) + g_\alpha * G_\alpha(t) + g_\alpha * G_\alpha^{*2}(t) + g_\alpha * G_\alpha^{*3}(t) + \cdots \\ &= g_\alpha * U_\alpha(t) \end{aligned}$$

and hence $\lim_{t\to\infty} e^{-\alpha t}\big|H(t)\big|$ is bounded by a constant for any $s \geq 0$. So,

$$\sup_{s,t\geq 0} e^{-\alpha t}H(t) < \infty.$$

□

Lemma 4.7. *Let ξ_1 be the function defined in (4.19). Then, under the hypotheses of Theorem 4.5, $e^{-\alpha t}\xi_1(t)$ is directly Riemann integrable.*

Proof. Note that

$$\left|e^{-\alpha t}\xi_1\right| = \left|e^{-\alpha t}\Big(1 - e^{-sh(t)}\Big)\Big(1 - G(t)\Big)\right| \leq e^{-\alpha t}(1 - G(t)) \equiv g_\alpha(t)$$

where g_α is known as a directly Riemann integrable function from the proof of Lemma 4.6.

So, $e^{-\alpha t}\xi_1$ is directly Riemann integrable by condition (3) in Remark 4.1.

□

Lemma 4.8. *Let ξ_2 be the function defined in (4.20). Then, under the hypotheses of Theorem 4.5,*

$$\int_0^\infty e^{-\alpha t}\big|\xi_2(t)\big|dt < \infty.$$

Proof. Recall that

$$\begin{aligned} & H(t) = \xi_1(t) + \xi_2(t) + m\int_{[0,t]} H(t-u)dG(u) \\ \Rightarrow\quad & e^{-\alpha t}H(t) = e^{-\alpha t}\xi_1(t) + e^{-\alpha t}\xi_2(t) + m\int_{[0,t]} e^{-\alpha t}H(t-u)dG(u). \end{aligned}$$

Let $H_\alpha(t) = e^{-\alpha t}H(t)$, $\xi_{1\alpha}(t) = e^{-\alpha t}\xi_1(t)$ and $\xi_{2\alpha}(t) = e^{-\alpha t}\xi_2(t)$, then

$$\begin{aligned} H_\alpha(t) &= \xi_{1\alpha}(t) + \xi_{2\alpha}(t) + \int_{[0,t]} H_\alpha(t-u)dG_\alpha(u) \\ &= \xi_{1\alpha}(t) + \xi_{2\alpha}(t) + H_\alpha * G_\alpha(t). \end{aligned} \tag{4.25}$$

We know that $\xi_{1\alpha}$ is bounded by 1 and, by Lemma 4.6, H_α is also bounded, so $\xi_{2,\alpha}$ is bounded. Take Laplace transforms on both sides of (4.25), we have that

$$\begin{aligned} & \hat{H}_\alpha(\theta) = \hat{\xi}_{1\alpha}(\theta) + \hat{\xi}_{2\alpha}(\theta) + \hat{H}_\alpha \cdot \hat{G}_\alpha(\theta) \\ \Rightarrow \quad & \tilde{H}_\alpha(\theta)\Big(1 - \hat{G}_\alpha(\theta)\Big) + \Big(-\tilde{\xi}_{2\alpha}(\theta)\Big) = \hat{\xi}_{1\alpha}(\theta). \end{aligned}$$

Note that, by (4.26),

$$\frac{f(1) - f(H(s,t-u))}{1 - H(s,t-u)} = f'(c) < f'(1) = m \tag{4.26}$$

and hence $\xi_2(t) = \int_0^t \Big[1 - f(H(s,t-u)) - mH(t-u)\Big]dG(u) < 0.$

So, we have that $H_\alpha \geq 0$, $\xi_{1\alpha} \geq 0$, $\xi_{2\alpha} \leq 0$ and $G_\alpha \leq 1$. Thus, $\hat{H}_\alpha(\theta)\Big(1 - \hat{G}_\alpha(\theta)\Big) \geq 0$, $-\hat{\xi}_{2\alpha}(\theta) \geq 0$ and $\hat{\xi}_{1\alpha}(\theta) \geq 0$.

Also, by the monotone convergence thoerem,

$$\lim_{\theta \downarrow 0} \hat{\xi}_{1\alpha}(\theta) = \lim_{\theta \downarrow 0} \int_0^\infty e^{-\theta t}\xi_{1\alpha}(t)dt = \int_0^\infty \xi_{1\alpha}(t)dt = \int_0^\infty e^{-\alpha t}\xi_1(t)dt < \infty$$

and hence $\lim_{\theta \downarrow 0}\Big(-\hat{\xi}_{2\alpha}(\theta)\Big) < \infty$ since $\hat{H}_\alpha(\theta)\Big(1 - \hat{G}_\alpha(\theta)\Big)$, $-\hat{\xi}_{2\alpha}(\theta)$ and $\hat{\xi}_{1\alpha}(\theta)$ are of the same sign.

Therefore, by the monotone convergence theorem again,

$$\begin{aligned} \int_0^\infty e^{-\alpha t}\Big|\xi_2(t)\Big|dt &= \int_0^\infty e^{-\alpha t}\Big(-\xi_2(t)\Big)dt = \int_0^\infty \Big(-\xi_{2\alpha}(t)\Big)dt = \lim_{\theta \downarrow 0} \int_0^\infty e^{-\theta t}\Big(-\xi_{2\alpha}(t)\Big)dt \\ &= \lim_{\theta \downarrow 0}\Big(-\hat{\xi}_{2\alpha}(\theta)\Big) < \infty \end{aligned}$$

□

Lemma 4.9. *Let ξ_2 be the function defined in (4.20). Then, under the hypotheses of Theorem 4.5, $e^{-\alpha t}\xi_2(t)$ is directly Riemann integrable.*

Proof. By the assumption, we have that

$$M \equiv \sup_n \sup_{\substack{n \leq t < n+1 \\ 0 \leq u \leq n}} \left| \frac{1 - f(H_h(s,t-u)) - m(1 - H_h(s,t-u))}{1 - f(H_h(s,n-u)) - m(1 - H_h(s,n-u))} \right| < \infty.$$

So, $n \le t < n+1$, we have

$$\begin{aligned}
& e^{-\alpha t}\left|\xi_2(t)\right| \\
\le\ & e^{-\alpha(n+1)}\left|\int_0^t \left[1 - f(H(s,t-u)) - mH(t-u)\right]dG(u)\right| \\
=\ & e^{-\alpha(n+1)}\left|\int_0^n \left[1 - f(H(s,t-u)) - mH(t-u)\right]dG(u) + \int_n^t \left[1 - f(H(s,t-u)) - mH(t-u)\right]dG(u)\right| \\
\le\ & e^{-\alpha(n+1)}\left|\int_0^n \left[1 - f(H(s,n-u)) - mH(n-u)\right]dG(u)\right| + e^{-\alpha(n+1)}\int_n^t \left|1 - f(H(s,t-u)) - mH(t-u)\right|dG(u) \\
=\ & e^{-\alpha(n+1)}\int_0^n \left|1 - f(H(s,t-u)) - mH(t-u)\right|dG(u) + e^{-\alpha(n+1)}\int_n^t \left|1 - f(H(s,t-u)) - mH(t-u)\right|dG(u) \\
\le\ & e^{-\alpha(n+1)}\int_0^n M\left|1 - f(H(s,n-u)) - mH(n-u)\right|dG(u) + e^{-\alpha(n+1)}\int_n^t \left|1 - f(H(s,t-u)) - mH(t-u)\right|dG(u) \\
=\ & Me^{-\alpha(n+1)}\left|\int_0^n 1 - f(H(s,n-u)) - mH(n-u)dG(u)\right| + e^{-\alpha(n+1)}\int_n^t \left|1 - f(H(s,t-u)) - mH(t-u)\right|dG(u)
\end{aligned}$$

Since $|f| \le 1$, $|H| \le 1$ and $\left|1 - f(H(s,t-u)) - mH(t-u)\right| \le 1$, we have

$$e^{-\alpha t}\left|\xi_2(t)\right| \quad \le \quad Me^{-\alpha(n+1)}\left|\xi_2(n)\right| + e^{-\alpha(n+1)}(G(t) - G(n)).$$

and then

$$\sup_{n \le t < n+1} e^{-\alpha t}\left|\xi_2(t)\right| \le e^{-\alpha}\Big(Me^{-\alpha n}\left|\xi_2(n)\right| + e^{-\alpha n}(1 - G(n))\Big).$$

Moreover, by Lemma 4.7, we know that $\int_0^\infty e^{-\alpha t}\left|\xi_2(t)\right|dt < \infty$ and, by the assumption, we also have that $\int_0^\infty e^{-\alpha t}(1 - G(t))dt < \infty$. So,

$$\sum_{n=0}^{\infty} e^{-\alpha n}\left|\xi_2(n)\right| < \infty \quad \text{and} \quad \sum_{n=0}^{\infty} e^{-\alpha n}(1 - G(n)) < \infty$$

and hence

$$\sum_{n=0}^{\infty} \sup_{n \le t < n+1} e^{-\alpha t}\left|\xi_2(t)\right| < \infty.$$

Since $e^{-\alpha t}\left|\xi_2(t)\right|$ is continuous and bounded, by condition (1) in Remark 4.1, $e^{-\alpha t}\left|\xi_2(t)\right|$ is directly Riemann integrable. Therefore, by condition (3), $e^{-\alpha t}\xi_2(t)$ is also directly Riemann integrable.

□

Now, we continue to prove Theorem 4.5.

By Lemma 4.7 and Lemma 4.9, we have that

$$e^{-\alpha t}\xi_3(t) = e^{-\alpha t}\xi_2(t) + e^{-\alpha t}\xi_3(t) \text{ is directly Riemann integrable.}$$

Then, by Lemma 4.5, we know that the solution H of the integral equation

$$H(t) = \xi_3(t) + m\int_{[0,t]} H(t-u)dG(u)$$

satisfies

$$H(t) \sim ce^{\alpha t} \quad \text{as } t \to \infty$$

where $c = \dfrac{\int_0^\infty e^{-\alpha u}\xi_3(u)du}{m\int_0^\infty ue^{-\alpha u}dG(u)}$.

Note that $H(t) = 1 - H(s,t)$ and hence $c \equiv c(s)$ depends on s.

Then

$$\begin{aligned}
\lim_{t\to\infty} E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\middle| Z(t) > 0\right) &= \lim_{t\to\infty}\left(1 - \frac{1}{P(Z(t)>0)}\Big(1 - H(s,t)\Big)\right) \\
&= 1 - \lim_{t\to\infty} \frac{1}{e^{-\alpha t}P(Z(t)>0)}\Big(1 - H(s,t)\Big)e^{-\alpha t} \\
&= 1 - \lim_{t\to\infty} \frac{1}{e^{-\alpha t}P(Z(t)>0)} H(s,t)e^{-\alpha t} \\
&\to 1 - \frac{c(s)}{\mathbf{Q}(0)} \equiv \phi(s).
\end{aligned}$$

Moreover, since, by the bounded convergence theorem,

$$\lim_{s\to 0+} H(s,t) = \lim_{s\to 0+} E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\middle| Z(t) > 0\right) = 1,$$

we have that

$$\lim_{s\to 0+} H(t) = \lim_{s\to 0+} 1 - H(s,t) = 0$$

$$\lim_{s\to 0+} \xi_1(t) = \lim_{s\to 0+} (1 - e^{-sh(t)})(1 - G(t)) = 0,$$

and, by the bounded convergence theorem again,

$$\lim_{s\to 0+} \xi_2(t) = \lim_{s\to 0+} \int_0^t (1 - f(H(s,t-u))) - mH(t-u)dG(u) = 0.$$

Hence, $\lim_{s\to 0+} \xi_3(t) = \lim_{s\to 0+} (\xi_1(t) + \xi_2(t)) = 0$.

Also, for any $s \geq 0$,

$$\left|e^{-\alpha t}\xi_3(t)\right| \leq e^{-\alpha t}\left|\xi_1(t)\right| + e^{-\alpha t}\left|\xi_2(t)\right|$$

where $e^{-\alpha t}\left|\xi_1(t)\right|$ and $e^{-\alpha t}\left|\xi_2(t)\right|$ are integrable.

Then, by the dominated convergence theorem,

$$\lim_{s\to 0+}\int_0^\infty e^{-\alpha t}\xi_3(t)dt = \int_0^\infty \lim_{s\to 0+} e^{-\alpha u}\xi_3(t)st = 0$$

and hence

$$\lim_{s\to 0+}\phi(s) = \lim_{s\to 0+} 1 - \frac{c(s)}{\mathbf{Q}(0)} = 1 - \lim_{s\to 0+}\frac{1}{\mathbf{Q}(0)}\frac{\int_0^\infty e^{-\alpha u}\xi_3(u)du}{m\int_0^\infty ue^{-\alpha u}dG(u)} = 1 - 0 = 1.$$

Therefore, ϕ is a Laplace functional of a point process.

Since, for any $s \geq 0$

$$\phi(s) = \lim_{t\to\infty} E\left(e^{-s\sum_{i=1}^{Z(t)} h(a_{t,i})}\middle|Z(t) > 0\right)$$

and

$$Z(t)\Big|Z(t) > 0 \xrightarrow{d} Y \quad \text{as } t \to \infty,$$

there exists a point process $\tilde{A} \equiv \left\{\tilde{a}_i : 1 \leq i \leq Y\right\}$ such that

$$\phi(s) = E\left(e^{-s\sum_{i=1}^{Y} h(\tilde{a}_i)}\right)$$

for any $s \geq 0$, and, as $t \to \infty$,

$$A(t)\Big|Z(t) > 0 \xrightarrow{d} \tilde{A}.$$

The proof is complete.

4.3.3 The proof of Theorem 4.6

Let $Z_{t,i}(u)$ be the branching process initiated by the ith individual alive at time t. So,

$$Z(t) = \sum_{i=1}^{Z(t-u)} Z_{t-u,i}(a_{t,i} + u). \tag{4.27}$$

For any $u \leq t$,

$$\begin{aligned}
& P\Big(t - D_2(t) \geq u \Big| Z(t) \geq 2\Big) \\
= \;& P\Big(D_2(t) \leq t - u \Big| Z(t) \geq 2\Big) \\
= \;& E\left(\frac{\sum_{i \neq j = 1}^{Z(t-u)} Z_{t-u,i}(a_{t-u,i} + u) Z_{t-u,j}(a_{t-u,j} + u)}{Z(t)(Z(t) - 1)} \Bigg| Z(t) \geq 2 \right) \\
= \;& \frac{1}{P(Z(t) \geq 2)} E\left(\frac{\sum_{i \neq j = 1}^{Z(t-u)} Z_{t-u,i}(a_{t-u,i} + u) Z_{t-u,j}(a_{t-u,j} + u)}{Z(t)(Z(t) - 1)} I_{(Z(t) \geq 2)} \right) \\
= \;& \frac{1}{P(Z(t) \geq 2, Z(t) > 0)} E\left(\frac{\sum_{i \neq j = 1}^{Z(t-u)} Z_{t-u,i}(a_{t-u,i} + u) Z_{t-u,j}(a_{t-u,j} + u)}{Z(t)(Z(t) - 1)} I_{(Z(t) \geq 2)} I_{(Z(t-u) > 0)} \right)
\end{aligned}$$

By (4.27) and the definition of the conditional probability, we have

$$\begin{aligned}
& P\Big(t - D_2(t) \geq u \Big| Z(t) \geq 2\Big) \\
= \;& \frac{P(Z(t-u) > 0)}{P(Z(t) \geq 2 \big| Z(t) > 0) P(Z(t) > 0)} E\left(\frac{\sum_{i \neq j = 1}^{Z(t-u)} Z_{t-u,i}(a_{t-u,i} + u) Z_{t-u,j}(a_{t-u,j} + u)}{\left(\sum_{i=1}^{Z(t-u)} Z_{t-u,i}(a_{t,i} + u)\right)\left(\sum_{i=1}^{Z(t-u)} Z_{t-u,i}(a_{t,i} + u) - 1\right)} \right. \\
& \qquad \left. \cdot I_{\left(\sum_{i=1}^{Z(t-u)} Z_{t-u,i}(a_{t,i}+u) \geq 2\right)} \Bigg| Z(t-u) > 0 \right) \\
= \;& \frac{P(Z(t-u) > 0)}{P(Z(t) \geq 2 \big| Z(t) > 0) P(Z(t) > 0)} E\left(\frac{\sum_{i \neq j = 1}^{Z(t-u)} \tilde{Z}_i(a_{t-u,i} + u) \tilde{Z}_j(a_{t-u,j} + u)}{\left(\sum_{i=1}^{Z(t-u)} \tilde{Z}_i(a_{t-u,i} + u)\right)\left(\sum_{i=1}^{Z(t-u)} \tilde{Z}_i(a_{t-u,i} + u) - 1\right)} \right. \\
& \qquad \left. \cdot I_{\left(\sum_{i=1}^{Z(t-u)} \tilde{Z}_i(a_{t,i}+u) \geq 2\right)} \Bigg| Z(t-u) > 0 \right) \\
= \;& \frac{1}{P(Z(t) \geq 2 \big| Z(t) > 0)} \frac{P(Z(t-u) > 0)}{P(Z(t) > 0)} E\Big(\phi(A(t-u), u) \Big| A(t-u) = (a_{t-u,1}, a_{t-u,2}, \cdots, a_{t-u,Z(t-u)})\Big)
\end{aligned}$$

where $\left\{\tilde{Z}_i(t)\right\}_{i \geq 1}$ are i.i.d. copies of $Z(t)$ and

$$\phi((a_1, a_2, \cdots, a_k), u) = E\left(\frac{\sum_{i \neq j = 1}^{k} \tilde{Z}_i(a_i + u) \tilde{Z}_j(a_j + u)}{\left(\sum_{i=1}^{k} \tilde{Z}_i(a_i + u)\right)\left(\sum_{i=1}^{k} \tilde{Z}_i(a_i + u) - 1\right)} I_{\left(\sum_{i=1}^{k} \tilde{Z}_i(a_i+u) \geq 2\right)} \right)$$

for any positive integer k and any positive real numbers $a_1, a_2, \cdots, a_k$.

Since, for any fixed u, $\phi(\cdot, u)$ is bounded and continuous and by Theorem 1.13 (b),

$$E\Big(\phi(A(t-u), u)\Big|Z(t-u) > 0\Big) \to E\Big(\phi(\tilde{A}, u)\Big) \quad \text{as } t \to \infty.$$

Moreover, by Theorem 1.13 (a) $P(Z(t) > 0) \sim ce^{-\alpha t}$ for some $c > 0$, we have that

$$\begin{aligned}
& \lim_{t\to\infty} P(t - D_2(t) > u\big|Z(t) \geq 2) \\
= & \lim_{t\to\infty} \frac{1}{P(Z(t) \geq 2\big|Z(t) > 0)} \frac{ce^{\alpha(t-u)}}{ce^{\alpha t}} E\Big(\phi_2(A(t-u), u)\Big|Z(t-u) > 0\Big) \\
= & \frac{1}{P(Y \geq 2)} e^{-\alpha u} E\Big(\phi(\tilde{A}, u)\Big) \\
\equiv & \ 1 - H_2(u).
\end{aligned}$$

It is remains to show that H_2 is a proper probability distribution, i.e., $H_2(u) \to 1$ as $u \to \infty$. It suffices to prove that

$$\lim_{u\to\infty} e^{-\alpha u} E\Big(\phi(\tilde{A}, u)\Big) = 0.$$

First, we have

$$\begin{aligned}
& E\Big(\phi(\tilde{A}, u)\Big) \\
= & E\left(\frac{\sum_{i\neq j=1}^{Y} \tilde{Z}_i(\tilde{a}_i + u)\tilde{Z}_j(\tilde{a}_j + u)}{\Big(\sum_{i=1}^{Y} \tilde{Z}_i(\tilde{a}_i + u)\Big)\Big(\sum_{i=1}^{Y} \tilde{Z}_i(\tilde{a}_i + u) - 1\Big)} I_{\left(\sum_{i=1}^{Y} \tilde{Z}_i(\tilde{a}_i+u)\geq 2\right)}\right) \\
= & E\left(E\left(\frac{\sum_{i\neq j=1}^{Y} \tilde{Z}_i(\tilde{a}_i + u)\tilde{Z}_j(\tilde{a}_j + u)}{\Big(\sum_{i=1}^{Y} \tilde{Z}_i(\tilde{a}_i + u)\Big)\Big(\sum_{i=1}^{Y} \tilde{Z}_i(\tilde{a}_i + u) - 1\Big)} I_{\left(\sum_{i=1}^{Y} \tilde{Z}_i(\tilde{a}_i+u)\geq 2\right)}\middle|\tilde{A}\right)\right) \\
= & E\Big(P\Big(\text{ there exist } 1 \leq i, j \leq Y \text{ s.t. } i \neq j, \tilde{Z}_i(\tilde{a}_i + u) > 0, \text{ and } \tilde{Z}_j(\tilde{a}_j + u) > 0\Big|\tilde{A}\Big)\Big) \\
\leq & E\Big(1 - P\Big(\tilde{Z}_i(\tilde{a}_i + u) = 0 \text{ for all } i = 1, 2, \cdots, Y\Big|\tilde{A}\Big) - P\Big(\tilde{Z}_i(\tilde{a}_i + u) > 0 \text{ for some } i, \tilde{Z}_j(\tilde{a}_j + u) \text{ for all } j \neq i\Big|\tilde{A}\Big)\Big)
\end{aligned}$$

For any $0 \leq s \leq 1$ and $t \geq 0$,Let $F(s,t) = \sum_{j=0}^{\infty} P(Z(t) = j)s^j$ and by Theorem 1.10, we have that

$$\lim_{t\to\infty} e^{-\alpha t}(1 - F(s,t)) \equiv \mathbf{Q}(s) \quad \text{exists for } 0 \leq s \leq 1.$$

So,

$$\begin{aligned} &= e^{-\alpha u} E\Big(\phi(\tilde{A}, u)\Big) \\ &\leq e^{-\alpha u} E\Bigg(1 - \prod_{i=1}^{Y} F(0, \tilde{a}_i + u) - \sum_{i=1}^{Y} (1 - F(0, \tilde{a}_i + u)) \prod_{j \neq i} F(0, \tilde{a}_j + u)\Bigg). \end{aligned}$$

Note that the assumption of $\sum_{j=1}^{\infty} (j \log j) p_j < \infty$ implies $0 < EY < \infty$ and hence $P(0 < Y < \infty) = 1$. Now, conditioned on the limit age chart $\tilde{A}$, we have that

$$\begin{aligned} \lim_{u\to\infty} e^{-\alpha u}\Bigg(1 - \prod_{i=1}^{Y} F(0, \tilde{a}_i + u)\Bigg) &= \lim_{u\to\infty} e^{-\alpha u}\Bigg(1 - \prod_{i=1}^{Y} \Big(1 - \mathbf{Q}(0) e^{\alpha(\tilde{a}_i + u)}\Big)\Bigg) \\ &= \lim_{u\to\infty} \frac{1 - \prod_{i=1}^{Y} \Big(1 - \mathbf{Q}(0) e^{\alpha(\tilde{a}_i + u)}\Big)}{e^{\alpha u}} \\ &= \lim_{u\to\infty} \frac{- \sum_{i=1}^{Y} \big(- \mathbf{Q}(0) e^{\alpha \tilde{a}_i} \alpha e^{\alpha u}\big) \prod_{j\neq i} \Big(1 - \mathbf{Q}(0) e^{\alpha(\tilde{a}_j + u)}\Big)}{\alpha e^{\alpha u}} \\ &= \lim_{u\to\infty} \sum_{i=1}^{Y} \mathbf{Q}(0) e^{\alpha \tilde{a}_i} \prod_{j\neq i} \Big(1 - \mathbf{Q}(0) e^{\alpha(\tilde{a}_j + u)}\Big) \\ &= \mathbf{Q}(0) \sum_{i=1}^{Y} e^{\alpha \tilde{a}_i} \end{aligned}$$

and

$$\begin{aligned} &\lim_{u\to\infty} e^{-\alpha u} \sum_{i=1}^{Y} (1 - F(0, \tilde{a}_i + u)) \prod_{j\neq i} F(0, \tilde{a}_j + u) \\ &\qquad = \lim_{u\to\infty} e^{-\alpha u} \sum_{i=1}^{Y} \mathbf{Q}(0) e^{\alpha(\tilde{a}_i + u)} \prod_{j\neq i} (1 - \mathbf{Q}(0) e^{\alpha(\tilde{a}_j + u)}) \\ &\qquad \geq \lim_{u\to\infty} e^{-\alpha u} \sum_{i=1}^{Y} \mathbf{Q}(0) e^{\alpha(\tilde{a}_i + u)} \prod_{j\neq i} (1 - \mathbf{Q}(0) e^{\alpha u}) \\ &\qquad = \lim_{u\to\infty} e^{-\alpha u} \sum_{i=1}^{Y} \mathbf{Q}(0) e^{\alpha(\tilde{a}_i + u)} (1 - \mathbf{Q}(0) e^{\alpha u})^{Y-1} \\ &\qquad = \lim_{u\to\infty} \sum_{i=1}^{Y} \mathbf{Q}(0) e^{\alpha \tilde{a}_i} (1 - \mathbf{Q}(0) e^{\alpha u})^{Y-1} \\ &\qquad = \mathbf{Q}(0) \sum_{i=1}^{Y} e^{\alpha \tilde{a}_i}. \end{aligned}$$

Hence, conditioned on $\tilde{A}$,

$$\begin{aligned}
0 &\leq \lim_{u\to\infty} e^{-\alpha u}\Big(1 - \prod_{i=1}^{Y} F(0, \tilde{a}_i + u) - \sum_{i=1}^{Y} (1 - F(0, \tilde{a}_i + u)) \prod_{j\neq i} F(0, \tilde{a}_j + u)\Big) \\
&= \lim_{u\to\infty} e^{-\alpha u}\Big(1 - \prod_{i=1}^{Y} F(0, \tilde{a}_i + u)\Big) - \lim_{u\to\infty} e^{-\alpha u} \sum_{i=1}^{Y} (1 - F(0, \tilde{a}_i + u)) \prod_{j\neq i} F(0, \tilde{a}_j + u) \\
&\leq \mathbf{Q}(0) \sum_{i=1}^{Y} e^{\alpha \tilde{a}_i} - \mathbf{Q}(0) \sum_{i=1}^{Y} e^{\alpha \tilde{a}_i} \\
&= 0 \quad \text{w.p.1.}
\end{aligned}$$

Therefore, by the bounded convergence theorem,

$$\begin{aligned}
&\lim_{u\to\infty} e^{-\alpha u} E\Big(\phi(\tilde{A}, u)\Big) \\
= \; &\lim_{u\to\infty} e^{-\alpha u} E\Big(1 - \prod_{i=1}^{Y} F(0, \tilde{a}_i + u) - \sum_{i=1}^{Y} (1 - F(0, \tilde{a}_i + u)) \prod_{j\neq i} F(0, \tilde{a}_j + u)\Big) \\
= \; &0
\end{aligned}$$

and the proof is complete.

CHAPTER 5. COALESCENCE IN CONTINUOUS-TIME MULTI-TYPE AGE-DEPENDENT BELLMAN-HARRIS BRANCHING PROCESSES

5.1 Introduction

In this chapter, we consider a continuous-time d-type ($2 \le d < \infty$) age-dependent Bellman-Harris branching process $\{\mathbf{Z}(t) : t \ge 0\}$, where

$$\mathbf{Z}(t) = (Z_1(t), Z_2(t), \cdots, Z_d(t))$$

is the population vector of the individuals alive at time t and $Z_i(t)$ is the number of individuals of type i alive at time t, $t \ge 0$.

Recall that in a continuous-time multi-type Bellman-Harris branching process, each type i individual, upon its death, produces $\xi_{i,j}$ children of type j, $j = 1, 2, \cdots, d$, according to the probability distribution $\{p^{(i)}(\mathbf{j}) \equiv p^{(i)}(j_1, j_2, \cdots, j_d)\}_{\mathbf{j} \in \mathbb{N}^d}$ and independently of other individual, where $p^{(i)}(j_1, j_2, \cdots, j_d)$ is the probability that a type i parent produces j_1 children of type 1, j_2 children of type 2, $\cdots$, j_d children of type d.

As in the single-type Bellman-Harris process, there is a embedded generation process $\{\mathbf{Y}_n\}_{n \ge 0}$ for the multi-type Bellman-Harris branching process, where $\mathbf{Y}_n = (Y_{n,1}, Y_{n,2}, \cdots, Y_{n,d})$ and $Y_{n,i}$ is the number of individuals of type i in the nth generation. It is clear that $\{\mathbf{Y}_n\}_{n \ge 0}$ is a discrete-time multi-type Galton-Watson branching process.

Throughout this section, we will adopt all the definitions and notations from Section 1.5 and have the following assumptions:

(1) this process is initiated by one individual of type i_0 of age 0, i.e., $\mathbf{Z}(0) = \mathbf{e}_{i_0}$ and $a_{0,1} = 0$.

(2) $\mathbf{M} = \{m_{ij} : i, j = 1, 2, \cdots, d\}$ is nonsingular and positively regular and write ρ for its Perron-Frobenius root (the maximal eigenvalue).

Also, we denote α the Malthusian parameter for the matrix $\widehat{\mathbf{M}}(\alpha) = ((m_{ij}\widehat{G}_i(\alpha)))_{i,j=1}^d$ where $\widehat{G}(\alpha) = \int_0^\infty e^{-\alpha t}G(dt)$.

5.2 Results in The Supercritical Case

For any $k \geq 2$, we pick two individuals at random from those alive at time t and trace their lines of descent backward in time until they meet.

Let $D_k(t)$ be the death time of the last common ancestor of these randomly chosen individuals at time t. We investigate the limit behavior of $D_k(t)$ when t gets large. The result for the supercritical caser is stated in Theorem 5.1.

First , we assume that $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ for any $i = 1, 2, \cdots, d$, and $E(Z_{1,j}\big|\mathbf{Z}_0 = \mathbf{e}_i) \equiv m_{ij} < \infty$ for all $1 \leq i, j \leq d$.

5.2.1 The statements of Results

Let $\xi_{s,i,p} = (\xi_{s,i,p,1}, \xi_{s,i,p,2}, \cdots, \xi_{s,i,p,d})$ be the offspring vector of the pth individual of type i alive at time s.

Let $L_{s,i,p}$ be the total lifetime of the pth individual of type i alive at time s. Then $\left\{L_{s,i,p}\right\}_{p\geq 1}$ are i.i.d. copies of the lifetime random variable with distribution G_i.

Let $a_{s,i,p}$ be the corresponding age and $R_{s,i,p}$ be the corresponding residual lifetime at time s. That is, $R_{s,i,p} = L_{s,i,p} - a_{s,i,p}$ for any $p \geq 1$, $i = 1, 2, \cdots, d$ and any $s \geq 0$.

Theorem 5.1. *Let $1 < \rho < \infty$ and the life time distribution G_i is non-lattice with $G_i(0+) = 0$ for $i = 1, 2, \cdots, d$. If $E(\|Z_1\| \log \|Z_1\| \big|\mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$, then, for any integer $k \geq 2$,*

(a) for almost all trees $\mathcal{T}$ and any $s \geq 0$,

$$P(\tilde{D}_k \leq s\big|\mathcal{T}) \equiv H_k(s, \mathcal{T}) = 1 - \frac{\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\left(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\right)^k}{\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\right)^k}$$

where $\{\tilde{W}_{r,i,p}\}_{p\geq 1}$ are the i.i.d copies of the sum $\displaystyle\sum_{j=1}^{d}\sum_{q=i}^{\xi_{s,i,p,j}} W_{s,i,p,j,q}$, and $\{W_{s,i,p,j,q} : q \geq 1, 1 \leq j \leq d\}$ are i.i.d. copies of W as defined in Theorem 1.17.

(b) *there exists a random variable* $\tilde{D}_k$ *on the set of non-negative real numbers such that* $D_k(t) \xrightarrow{d} \tilde{D}_k$ *as* $t \to \infty$ *and*

$$P(\tilde{D}_k \le s) = 1 - E\left(\frac{\sum_{i=1}^{d} \sum_{p=1}^{Z_i(s)} \left(e^{-\alpha R_{s,i,p}} \tilde{W}_{s,i,p} \right)^k}{\left(\sum_{i=1}^{d} \sum_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}} \tilde{W}_{s,i,p} \right)^k} \right) \equiv H_k(s).$$

for any $s \ge 0$.

Next, we investigate the generation number $X_k(t)$ of the last common ancestor of any k randomly chosen individuals alive at time t.

Consider the case in which every individual in the branching process has the same lifetime distribution G no matter what kind of type it is of.

Let $L_{r,i,p,q}$ be the lifetime of the qth-generation ancestor of the pth individual of type i in the rth generation, $q = 0, 1, \cdots, r-1$, $p \ge 1$, $i = 1, 2, \cdots, d$ and $r \ge 0$.

Let $S_{r,i,p} = \sum_{q=0}^{r-1} L_{r,i,p,q}$, then $S_{r,i,p}$ is the birth time of the pth individual of type i in the rth generation.

Then we have the following theorem.

Theorem 5.2. *Let* $1 < \rho < \infty$ *and the life time distribution* G *is non-lattice with* $G(0+) = 0$. *If* $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ *for all* $1 \le i \le d$, *then, for any integer* $k \ge 2$,

(a) *for almost all trees* $\mathcal{T}$ *and any* $s \ge 0$,

$$P(\tilde{X}_k \le s \big| \mathcal{T}) \equiv H_k(s, \mathcal{T}) = 1 - \frac{\sum_{i=1}^{d} \sum_{p=1}^{Z_i(s)} \left(e^{-\alpha S_{s,i,p}} \tilde{W}_{s,i,p} \right)^k}{W^k}$$

where $\{\tilde{W}_{r,i,p}\}_{p \ge 1}$ *are the i.i.d copies of* W *as defined in Theorem 1.17.*

(b) *there exists a random variable* $\tilde{X}_k$ *on the set of non-negative real numbers such that* $X_k(t) \xrightarrow{d} \tilde{X}_k$ *as* $t \to \infty$ *and*

$$P(\tilde{X}_k \le s) = 1 - E\left(\frac{\sum_{i=1}^{d} \sum_{p=1}^{Z_i(s)} \left(e^{-\alpha S_{s,i,p}} \tilde{W}_{s,i,p} \right)^k}{W^k} \right) \equiv \phi_k(s).$$

for any $s \ge 0$.

Moreover, let $\eta(t)$ be the type of the last common ancestor and $\zeta_1(t), \zeta_2(t)$ be the types of the two randomly chosen individuals at time t. We also have the limit joint distribution of $(X_2(t), \eta(t), \zeta_1(t), \zeta_2(t))$.

Theorem 5.3. *Let* $1 < \rho < \infty$ *and the life time distribution* G *is non-lattice with* $G(0+) = 0$*. If* $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ *for all* $1 \leq i \leq d$*, then*

$$\lim_{n\to\infty} P(X_2(t) = r, \eta(t) = j, \zeta_1(t) = i_1, \zeta_2(t) = i_2) \equiv \varphi_2(r, j, i_1, i_2) \quad \textit{exists}$$

and $\sum_{(r,j,i_1,i_2)} \varphi_2(r, j, i_1, i_2) = 1.$

5.2.2 The proof of Theorem 5.1

Let $\tilde{Z}_{t-s-R_{s,i,p},j,q}$ be the branching process initiated by the qth offspring of type j of the pth individual of type i alive at time s.

Pick k individuals randomly from those alive at time t and trace their lines of decent backward in time until they meet. Denote the coalescence time $D_k(t)$ which also means the death time of the last common ancestor of these randomly chosen individuals.

For almost all trees $\mathcal{T}$ and $s \geq 0$,

$$\begin{aligned}
& P\Big(D_k(t) \leq s \Big| \mathcal{T}\Big) \\
= \; & 1 - P\Big(D_k(t) > s \Big| \mathcal{T}\Big) \\
= \; & 1 - \frac{\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\left(\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right|\right)\left(\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right| - 1\right)\cdots\left(\sum_{j=1}^{\xi_{s,i}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right| - k + 1\right)}{\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right)\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right| - 1\right)\cdots\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right| - k + 1\right)} \\
= \; & 1 - \frac{\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\prod_{l=1}^{k}\left(e^{-\alpha R_{s,i,p}}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right|e^{-\alpha(t-s-R_{s,i,p})} - (l-1)e^{-\alpha(t-s-R_{s,i,p})}\right)}{\prod_{l=1}^{k}\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}}\left|\tilde{\mathbf{Z}}_{t-s-R_{s,i,p},j,q}\right|e^{-\alpha(t-s-R_{s,i,p})} - (l-1)e^{-\alpha(t-s-R_{s,i,p})}\right)}
\end{aligned}$$

and then, by Theroem 1.17,

$$\begin{aligned}
P\Big(D_k(t) \leq s \Big| \mathcal{T}\Big) \quad \to \quad & 1 - \frac{\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\left(e^{-\alpha R_{s,i,p}}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}} W_{s,i,,p,j,q}\right)^k}{\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\sum_{j=1}^{d}\sum_{q=1}^{\xi_{s,i,p,j}} W_{s,i,p,j,q}\right)^k} \quad \text{as } t \to \infty \\
= \quad & 1 - \frac{\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)}\left(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\right)^k}{\left(\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\right)^k} \equiv H_k(s, \mathcal{T})
\end{aligned}$$

where $\{W_{s,i,p,j,q} : q \geq 1, 1 \leq j \leq d\}$ are i.i.d copies of W in Theorem 1.17 and $\tilde{W}_{s,i,p} \equiv \sum_{j=1}^{d} \sum_{q=1}^{\xi_{s,i,p,j}} W_{s,i,p,j,q}$ for $p \geq 1$, $1 \leq i \leq d$ and $s \geq 0$.

So, by the bounded convergence theorem, as $t \to \infty$,

$$P\Big(D_k(t) \leq s\Big) = E\Big(P\Big(D_k(t) \leq s\Big|\mathcal{T}\Big)\Big) \to 1 - E\left(\frac{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)}\Big(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}{\Big(\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}\right) \equiv H_k(s).$$

Next, we need to show that H_k is a proper probability distribution, i.e. show that $H_k(s) \to 1$ as $s \to \infty$ and it is the same as showing that

$$E\left(\frac{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)}\Big(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}{\Big(\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}\right) \to 0 \quad \text{as } s \to \infty.$$

It suffices to prove that, as $s \to \infty$,

$$\frac{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)}\Big(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}{\Big(\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k} \to 0 \quad \text{in probability.}$$

Moreover, since

$$\left(\frac{\max\limits_{\substack{1\leq p\leq Z_i(s)\\ 1\leq i\leq d}} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}\right)^k \leq \frac{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)}\Big(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}{\Big(\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k} \leq \left(\frac{\max\limits_{\substack{1\leq p\leq Z_i(s)\\ 1\leq i\leq d}} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}\right)^{k-1},$$

it is enough to show that, as $s \to \infty$,

$$\frac{\max\limits_{\substack{1\leq p\leq Z_i(s)\\ 1\leq i\leq d}} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}} \to 0 \quad \text{in probability.}$$

For any fixed $k > 0$,

$$\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p} \geq e^{-\alpha k}\sum_{i=1}^{d}\sum_{p=1}^{Z_i(s)} \tilde{W}_{s,i,p} I_{(R_{s,i,p}\leq k)}$$

and then

$$\frac{\max\limits_{\substack{1\leq p\leq Z_i(s)\\ 1\leq i\leq d}} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}} \leq \frac{\max\limits_{\substack{1\leq p\leq Z_i(s)\\ 1\leq i\leq d}} \tilde{W}_{s,i,p}}{e^{-\alpha k}\sum\limits_{i=1}^{d}\sum_{p=1}^{Z_i(s)} \tilde{W}_{s,i,p} I_{(R_{s,i,p}\leq k)}} = \frac{e^{\alpha k}\frac{1}{|\mathbf{Z}(s)|}\max\limits_{\substack{1\leq p\leq Z_i(s)\\ 1\leq i\leq d}} \tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\frac{Z_i(s,k)}{|\mathbf{Z}(s)|}\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s)} \tilde{W}_{s,i,p} I_{(R_{s,i,p}\leq k)}} \tag{5.1}$$

where $Z_i(s,k)$ is the number of individuals of type i alive at time s with the residual lifetime less than or equal to k.

We need the following lemmas to prove Theorem 5.1.

Lemma 5.1. *Under the hypotheses of Theorem 5.1, as $s \to \infty$,*

$$\frac{1}{|\mathbf{Z}(s)|} \max_{\substack{1 \le p \le Z_i(s) \\ 1 \le i \le d}} \tilde{W}_{s,i,p} \to 0 \quad \text{in probability.}$$

Proof. In a continuous-time single-type age-dependent Bellman-Harris branching process, $\{\xi_{s,i,p} : p \ge 0\}$ are i.i.d. for any $i = 1, 2, \cdots, d$ and $s \ge 0$. Also, $\{W_{s,i,p,j,q} : p, q \ge 0, 1 \le i, j \le d\}$ are i.i.d. and independent of $\{\xi_{s,i,p} : p \ge 0, 1 \le i \le d\}$, so $\{\tilde{W}_{s,i,p} : p \ge 1, 1 \le i \le d\}$ are independent random variables and

$$E\tilde{W}_{s,i,p} = E\Big(\sum_{j=1}^{d} \sum_{q=1}^{\xi_{s,i,p,j}} W_{s,i,p,j,q}\Big) = E|\xi_{s,i,p}| \cdot EW_{s,i,p,1,1} \in (0, \infty). \tag{5.2}$$

Thus, since $E\tilde{W}_{s,i,p} < \infty$, for any $\epsilon > 0$,

$$nP(\tilde{W}_{s,i,p} > n\epsilon) \to 0 \quad \text{as } n \to \infty \tag{5.3}$$

and then,

$$\begin{aligned}
& P\Big(\frac{1}{|\mathbf{Z}(s)|} \max_{\substack{1 \le p \le Z_i(s) \\ 1 \le i \le d}} \tilde{W}_{s,i,p} > \epsilon\Big) \\
= \; & E\Big(P\Big(\frac{1}{|\mathbf{Z}(s)|} \max_{\substack{1 \le p \le Z_i(s) \\ 1 \le i \le d}} \tilde{W}_{s,i,p} > \epsilon \Big| \mathbf{Z}(s)\Big)\Big) \\
= \; & E\Big(P\Big(\max_{\substack{1 \le p \le Z_i(s) \\ 1 \le i \le d}} \tilde{W}_{s,i,p} > \epsilon|\mathbf{Z}(s)| \Big| \mathbf{Z}(s)\Big)\Big) \\
= \; & E\Big(1 - P\Big(\max_{\substack{1 \le p \le Z_i(s) \\ 1 \le i \le d}} \tilde{W}_{s,i,p} \le \epsilon|\mathbf{Z}(s)| \Big| \mathbf{Z}(s)\Big)\Big) \\
= \; & E\Big(1 - P\Big(\tilde{W}_{s,i,p} \le \epsilon|\mathbf{Z}(s)| \text{ for all } p = 1, 2, \cdots, Z_i(s), i = 1, 2, \cdots, d \Big| \mathbf{Z}(s)\Big)\Big) \\
= \; & E\Big(1 - \prod_{i=1}^{d} \prod_{p=1}^{Z_i(s)} P\Big(\tilde{W}_{s,i,p} \le \epsilon|\mathbf{Z}(s)| \Big| \mathbf{Z}(s)\Big)\Big) \\
= \; & E\Big(1 - \prod_{i=1}^{d} P\Big(\tilde{W}_{s,i,1} \le \epsilon|\mathbf{Z}(s)| \Big| \mathbf{Z}(s)\Big)^{Z_i(s)}\Big) \\
= \; & E\Big(1 - \prod_{i=1}^{d} \Big(1 - \frac{Z_i(s) P\Big(\tilde{W}_{s,i,1} > \epsilon|\mathbf{Z}(s)| \Big| \mathbf{Z}(s)\Big)}{Z_i(s)}\Big)^{Z_i(s)}\Big)
\end{aligned}$$

Since $Z_i(s) \to \infty$ w.p.1 as $n \to \infty$ for $i = 1, 2, \cdots, d$, by the bounded convergence theorem, as $s \to \infty$,

$$\begin{aligned} P\Big(\frac{1}{|\mathbf{Z}(s)|} \max_{\substack{1\le p\le Z_i(s)\\ 1\le i\le d}} \tilde{W}_{s,i,p} > \epsilon\Big) &= E\Big(P\Big(\frac{1}{|\mathbf{Z}(s)|} \max_{\substack{1\le p\le Z_i(s)\\ 1\le i\le d}} \tilde{W}_{s,i,p} > \epsilon \Big| \mathbf{Z}(s)\Big)\Big) \\ &\to E\Big(1 - \prod_{i=1}^{d} e^0\Big) = 0. \end{aligned}$$

Then, Lemma 5.1 is proved.

□

Lemma 5.2. *For any $i = 1, 2, \cdots, d$ and any $k > 0$, let $Z_i(s, k)$ be the number of individuals of type i alive at time s with the residual lifetime less than or equal to k. Then, under the hypotheses of Theorem 5.1, as $s \to \infty$,*

$$\frac{Z_i(s,k)}{Z_i(s)} \to B_i(k) \quad \textit{in probability}$$

where

$$B_i(k) = \frac{\int_{[0,\infty)} e^{-\alpha x}[G_i(x+k) - G_i(x)]dx}{\int_{[0,\infty)} e^{-\alpha x}[1 - G_i(x)]dx}.$$

Proof. Recall that $L_{s,i,p}$ is the total lifetime of the pth individual of type i alive at time s, $a_{s,i,p}$ is the corresponding age and $R_{s,i,p}$ is the corresponding residual lifetime at time s.

For any fixed $i = 1, 2, \cdots, d$ and any fixed $k > 0$, consider a function g such that

$$\begin{aligned} g(a) &\equiv P\Big(R_{s,i,p} \le k \Big| a_{s,i,p} = a\Big) = P\Big(L_{s,i,p} - a_{a,i,p} \le k \Big| L_{s,i,p} > a\Big) = P\Big(a < L_{s,i,p} \le a + k \Big| L_{s,i,p} > a\Big) \\ &= \frac{P\Big(a < L_{s,i,p} \le a + k\Big)}{P\Big(L_{s,i,p} > a\Big)} = \frac{G_i(a+k) - G_i(a)}{1 - G_i(a)}. \end{aligned}$$

Let $\mathfrak{F}_s$ be the $\sigma-$algebra generated all the history of this branching process up to time s.

Then, for any $\epsilon > 0$,

$$
\begin{aligned}
& P\left(\left|\frac{Z_i(s,k)}{Z_i(s)} - B_i(k)\right| > \epsilon\right) \\
= \;& E\left(P\left(\left|\frac{Z_i(s,k)}{Z_i(s)} - B_i(k)\right| > \epsilon \middle| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\right)\right) \\
= \;& E\left(P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} I_{(R_{s,i,p}\leq k)} - B_i(k)\right| > \epsilon \middle| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\right)\right) \\
= \;& E\left(P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} \Big(I_{(R_{s,i,p}\leq k)} - g(a_{s,i,p})\Big) + \frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} \Big(g(a_{s,i,p}) - B_i(k)\Big)\right| > \epsilon \middle| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\right)\right) \\
\leq \;& E\left(P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} \Big(I_{(R_{s,i,p}\leq k)} - g(a_{s,i,p})\Big)\right| > \frac{\epsilon}{2} \middle| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\right)\right) \\
& \qquad + E\left(P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} \Big(g(a_{s,i,p}) - B_i(k)\Big)\right| > \frac{\epsilon}{2} \middle| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_(s)})\right)\right) \\
= \;& E\left(P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_(s)} \Big(I_{(R_{s,i,p}\leq k)} - g(a_{s,i,p})\Big)\right| > \frac{\epsilon}{2} \middle| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\right)\right) \\
& \qquad + P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} \Big(g(a_{s,i,p}) - B_i(k)\Big)\right| > \frac{\epsilon}{2}\right) \qquad (5.4)
\end{aligned}
$$

Note that, for any $i \geq 1$,

$$
E\Big(I_{(R_{s,i,p}\leq k)} - g(a_{s,i,p}) \Big| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\Big) = 0
$$

and

$$
\begin{aligned}
& Var\Big(I_{(R_{s,i,p}\leq k)} - g(a_{s,i,p}) \Big| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\Big) \\
= \;& E\Big(\Big|I_{(R_{s,i,p}\leq k)} \leq k) - g(a_{s,i,p})\Big|^2 \Big| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\Big) \\
= \;& E\Big(I_{(R_{s,i,p}\leq k)} \leq k) \Big| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})\Big) \\
& \qquad - 2g(a_{s,i,p})E\Big(I_{(R_{s,i,p}\leq k)} \leq k) \Big| \mathfrak{F}_s, (a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z(s)})\Big) + \Big(g(a_{s,i,p})\Big)^2 \\
= \;& g(a_{s,i,p}) - 2\Big(g(a_{s,i,p})\Big)^2 + \Big(g(a_{s,i,p})\Big)^2 = g(a_{s,i,p}) - \Big(g(a_{s,i,p})\Big)^2 \leq \frac{1}{4}.
\end{aligned}
$$

Conditioned on $\mathfrak{F}_s$ and $(a_{s,i,1}, a_{s,i,2}, \cdots, a_{s,i,Z_i(s)})$, we have that $\{I_{(R_{s,i,p}\leq k) - g(a_{s,i,p})} : p = 1, 2, \cdots, Z_i(s)\}$

are independent. So, by Chebycheve's inequality, for any $\epsilon > 0$,

$$
\begin{aligned}
& P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}\left(I_{(R_{s,i,p}\le k)} - g(a_{s,i,p})\right)\right| > \frac{\epsilon}{2}\middle|\mathfrak{F}_s,(a_{s,i,1},a_{s,i,2},\cdots,a_{s,i,Z_i(s)})\right) \\
\le\ & \frac{4}{\epsilon^2}Var\left(\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}\left(I_{(R_{s,i,p}\le k)} - g(a_{s,i,p})\right)\middle|\mathfrak{F}_s,(a_{s,i,1},a_{s,i,2},\cdots,a_{s,i,Z_i(s)})\right) \\
=\ & \frac{4}{\epsilon^2}\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}Var\left(I_{(R_{s,i,p}\le k)} - g(a_{s,i,p})\middle|\mathfrak{F}_s,(a_{s,i,1},a_{s,i,2},\cdots,a_{s,i,Z_i(s)})\right) \\
\le\ & \frac{4}{\epsilon^2}\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}\frac{1}{4} \\
=\ & \frac{1}{\epsilon^2}Z_i(s) \to 0 \quad \text{as } s \to \infty
\end{aligned}
$$

and hence

$$P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}\left(I_{(R_{s,i,p}\le k)} - g(a_{s,i,p})\right)\right| > \frac{\epsilon}{2}\middle|\mathfrak{F}_s,(a_{s,i,1},a_{s,i,2},\cdots,a_{s,i,Z_i(s)})\right) \to 0 \quad \text{w.p.1}$$

as $s \to \infty$. Then, by the bounded convergence theorem,

$$E\left(P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}\left(I_{(R_{s,i,p}\le k)} - g(a_{s,i,p})\right)\right| > \frac{\epsilon}{2}\middle|\mathfrak{F}_s,(a_{s,i,1},a_{s,i,2},\cdots,a_{s,i,Z_i(s)})\right)\right) \to 0 \quad \text{as } s \to \infty. \tag{5.5}$$

It remains to prove that

$$P\left(\left|\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)}\left(g(a_{s,i,p}) - B_i(k)\right)\right| > \frac{\epsilon}{2}\right) \to 0 \quad \text{as } s \to \infty.$$

Let $A_i(x,s) = \frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} I_{(a_{s,i,p}\le x)}$. Then, by Theorem 1.20, as $s \to \infty$

$$\sup_x\left|A_i(x,s) - A_i(x)\right| \to 0 \quad \text{w.p.1.}$$

where A_i is as defined in Section 1.5.

Since $g(x) = \frac{G_i(x+k) - G_i(x)}{1 - G_i(x)}$ is a bounded and continuous function, by Theorem 1.20 again, we have, as $s \to \infty$

$$\frac{1}{Z_i(s)}\sum_{p=1}^{Z_i(s)} g(a_{s,i,p}) \equiv \int_{[0,\infty)} g(x)dA_i(x,s) \to \int_{[0,\infty)} g(x)dA_i(x) \quad \text{w.p.1}$$

where

$$\int_{[0,\infty)} g(x)dA_i(x) = \frac{\int_0^\infty \frac{G_i(x+k)-G_i(x)}{1-G_i(x)}e^{-\alpha x}(1-G_i(x))dx}{\int_0^\infty e^{-\alpha x}(1-G_i(x))dx} = \frac{\int_0^\infty e^{-\alpha x}(G_i(x+k)-G_i(x))dx}{\int_0^\infty e^{-\alpha x}(1-G_i(x))dx} = B_i(k).$$

So, as $s \to \infty$, $\frac{1}{Z_i(s)} \sum_{p=1}^{Z_i(s)} g(a_{s,i,p}) \to B_i(k)$ w.p.1 and hence in probability. Therefore, for any $\epsilon > 0$,

$$P\Big(\Big|\frac{1}{Z_i(s)} \sum_{p=1}^{Z_i(s)} g(a_{s,i,p}) - B_i(k)\Big| > \frac{\epsilon}{2}\Big) \to 0 \quad \text{as } s \to \infty. \tag{5.6}$$

From (5.4), (5.5) and (5.6), we have that, for any $\epsilon > 0$,

$$P\Big(\Big|\frac{Z_i(s,k)}{Z_i(s)} - B_i(k)\Big| > \epsilon\Big) \to 0 \quad \text{as } s \to \infty$$

and the proof is complete. □

Lemma 5.3. *For any $i = 1, 2, \cdots, d$, let $\tilde{W}_{s,i,p}$ and $Z_i(s,k)$ be the random variables defined in Lemma 5.1 and Lemma 5.2, respectively. Then, under the hypotheses of Theorem 5.1, these exists a $\theta_i > 0$ such that, as $s \to \infty$,*

$$P\Big(\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \tilde{W}_{s,i,p} I_{(R_{s,i,p} \le k)} \ge \theta_i\Big) \to 1.$$

Proof. Let $n_{s,i,1} = \min\Big\{1 \le j \le Z_i(s) : R_{s,i,j} \le k\Big\}$ and $n_{s,i,p} = \min\Big\{n_{s,i,p-1} < j \le Z_i(s) : R_{s,i,j} \le k\Big\}$ for $i \ge 2$. Then

$$\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s)} \tilde{W}_{s,i,p} I_{(R_{s,i,p} \le k)} = \frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \tilde{W}_{s,i,n_{s,i,p}} I_{(R_{s,i,n_{s,i,p}} \le k)} = \frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \tilde{W}_{s,i,n_{s,i,p}}. \tag{5.7}$$

It is known from (5.2) that $E\tilde{W}_{s,i,1} > 0$ and hence there exists an $\eta > 0$ such that $P\Big(\tilde{W}_{s,i,1} \ge \eta\Big) > 0$.

Let $\mathfrak{F}_s$ be the $\sigma-$algebra generated by all the information of this Bellman-Harris branching process up to time s. Then, for any $p \ge 1$,

$$\begin{aligned}
P\Big(\tilde{W}_{s,i,n_{s,i,p}} \ge \eta\Big) &= E\Big(P\Big(\tilde{W}_{s,i,n_{s,i,p}} \ge \eta\Big|\mathfrak{F}_s\Big)\Big) = E\Big(\sum_{j=1}^{Z_i(s,k)} P\Big(\tilde{W}_{s,i,n_{s,i,p}} \ge \eta, n_{s,i,p} = j\Big|\mathfrak{F}_s\Big)\Big) \\
&= E\Big(\sum_{j=1}^{Z_i(s,k)} P\Big(\tilde{W}_{s,i,j} \ge \eta, n_{s,i,p} = j\Big|\mathfrak{F}_s\Big)\Big) \\
&= E\Big(\sum_{j=1}^{Z_i(s,k)} P\Big(\tilde{W}_{s,i,j} \ge \eta\Big|\mathfrak{F}_s\Big)P\Big(n_{s,i,p} = j\Big|\mathfrak{F}_s\Big)\Big) \\
&= E\Big(\sum_{j=1}^{Z_i(s,k)} P\Big(\tilde{W}_{s,i,j} \ge \eta\Big|\mathfrak{F}_s\Big)P\Big(n_{s,i,p} = j\Big|\mathfrak{F}_s\Big)\Big) \\
&= P\Big(\tilde{W}_{s,i,1} \ge \eta\Big).
\end{aligned}$$

Let

$$X_{s,i,p} = \begin{cases} 1 & , \text{if } \tilde{W}_{s,n_{s,i,p}} \geq \eta \\ 0 & . \text{if } \tilde{W}_{s,n_{s,i,p}} < \eta \end{cases}$$

then

$$\begin{aligned} \frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \tilde{W}_{s,i,n_{s,i,p}} &\geq \frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \eta X_{s,i,p} \\ &= \eta\Big(\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big) + \eta\Big(\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\Big) \\ &= \eta\Big(\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big) + \eta P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta). \end{aligned} \tag{5.8}$$

Conditioned on $\mathfrak{F}_s$, $\{X_{s,i,p}\}_{p\geq 1}$ are independent with $E\Big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,p}} \geq \eta)\Big|\mathfrak{F}_s\Big) = 0$ and

$$\begin{aligned} Var\Big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\Big|\mathfrak{F}_s\Big) &= P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta\big|\mathfrak{F}_s)\big(1 - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta\big|\mathfrak{F}_s)\big) \\ &= P(\tilde{W}_{s,i,1} \geq \eta\big|\mathfrak{F}_s)\big(1 - P(\tilde{W}_{s,i,1} \geq \eta\big|\mathfrak{F}_s)\big) \\ &\leq \frac{1}{4}, \end{aligned}$$

then, by Chebychev's inequality and Lemma 5.2, for any $\epsilon > 0$,

$$\begin{aligned} & P\Big(\Big|\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big| > \epsilon\Big|\mathfrak{F}_s\Big) \\ \leq\ & \frac{1}{\epsilon^2} Var\Big(\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big|\mathfrak{F}_s\Big) \\ =\ & \frac{1}{\epsilon^2} \frac{1}{Z_i(s,k)^2} \sum_{p=1}^{Z_i(s,k)} Var\Big(\big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big|\mathfrak{F}_s\Big) \\ \leq\ & \frac{1}{4\epsilon^2 Z_i(s,k)} \\ =\ & \frac{1}{4\epsilon^2 Z_i(s)} \frac{Z_i(s)}{Z_i(s,k)} \to 0 \quad \text{in probability} \quad \text{as } s \to \infty. \end{aligned}$$

Therefore, by the bounded convergence theorem,

$$\begin{aligned} & P\Big(\Big|\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big| > \epsilon\Big) \\ =\ & E\Big(P\Big(\Big|\frac{1}{Z_i(s,k)} \sum_{p=1}^{Z_i(s,k)} \big(X_{s,i,p} - P(\tilde{W}_{s,i,n_{s,i,p}} \geq \eta)\big)\Big| > \epsilon\Big|\mathfrak{F}_s\Big)\Big) \to 0 \quad \text{as } s \to \infty. \end{aligned}$$

Hence,

$$P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\Big(X_{s,i,p}-P(\tilde{W}_{s,i,n_{s,i,p}}\geq\eta)\Big)<-\epsilon\Big)\to 0 \quad \text{as } s\to\infty. \tag{5.9}$$

Let $\theta_i=\frac{1}{2}\eta P(\tilde{W}_{s,i,1}\geq\eta)$, then $\theta_i>0$. Also, (5.7), (5.8) and (5.9) together imply that

$$\begin{aligned}
& P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\leq k)}\geq\theta_i\Big)\\
\geq\ & P\Big(\eta\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\Big(X_{s,i,p}-P(\tilde{W}_{s,i,n_{s,i,p}}\geq\eta)\Big)\Big)+\eta P(\tilde{W}_{s,i,n_{s,i,p}}\geq\eta)\geq\frac{1}{2}\eta P(\tilde{W}_{s,i,1}\geq\eta)\Big)\\
=\ & P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\Big(X_{s,i,p}-P(\tilde{W}_{s,i,n_{s,i,p}}\geq\eta)\Big)\geq-\frac{1}{2}P(\tilde{W}_{s,i,1}\geq\eta)\Big)\\
=\ & 1-P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\Big(X_{s,i,p}-P(\tilde{W}_{s,i,n_{s,i,p}}\geq\eta)\Big)<-\frac{1}{2}P(\tilde{W}_{s,i,1}\geq\eta)\Big)\to 1 \quad \text{as } s\to\infty.
\end{aligned}$$

So, we have that

$$P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\leq k)}\geq\theta_i\Big)\to 1 \text{ as } s\to\infty$$

and hence Lemma 5.3 is proved.

□

Now, we can continue to prove Theorem 5.1.

From Lemma 5.2, we know that, as $s\to\infty$,

$$\frac{Z_i(s,k)}{Z_i(s)}\to B_i(k) \quad \text{in probability}$$

for any $i=1,2,\cdots,d$ and, from Theorem 1.17, we have that

$$\frac{Z_i(s)}{|\mathbf{Z}(s)|}\to\frac{v_i}{\mathbf{1}\cdot\mathbf{v}} \quad \text{w.p.1} \quad \text{as } s\to\infty.$$

Hence,

$$P\Big(\frac{Z_i(s,k)}{|\mathbf{Z}(s)|}<\frac{1}{2}B(k)\frac{v_i}{\mathbf{1}\cdot\mathbf{v}}\Big)\to 0 \quad \text{as } s\to\infty.$$

Also, from Lemma 5.3, we have that, for some $\theta_i>0$, as $s\to\infty$,

$$P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\leq k)}\geq\theta_i\Big)\to 1.$$

So, for any $\delta > 0$ and any $i = 1, 2, \cdots, d$, there exists an $M > 0$ such that for every $s > M$,

$$P\Big(\frac{Z_i(s,k)}{|\mathbf{Z}(s)|} < \frac{1}{2}B_i(k)\frac{v_i}{\mathbf{1}\cdot\mathbf{v}}\Big) < \frac{\delta}{2d}$$

and

$$P\Big(\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z(s,k)} \tilde{W}_{s,i,p} I_{(R_{s,i,p}\le k)} < \theta_i\Big) < \frac{\delta}{2d}.$$

Let

$$A = \Big\{\frac{Z_i(s,k)}{|\mathbf{Z}(s)|} \ge \frac{1}{2}B_i(k)\frac{v_i}{\mathbf{1}\cdot\mathbf{v}} \text{ for all } i = 1, 2, \cdots, d\Big\}$$

and

$$B = \Big\{\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s,k)} \tilde{W}_{s,i,p} I_{(R_{s,i,p}\le k)} \ge \theta_i \text{ for all } i = 1, 2, \cdots, d\Big\}.$$

Then, for any $\epsilon > 0$,

$$\begin{aligned}
& P\Bigg(\frac{e^{\alpha k}\frac{1}{|\mathbf{Z}(s)|}\max\limits_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\frac{Z_i(s,k)}{|\mathbf{Z}(s)|}\frac{1}{Z_i(s,k)}\sum\limits_{p=1}^{Z_i(s)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\le k)}} > \epsilon\Bigg) \\
= \;& P\Big(\frac{1}{|\mathbf{Z}(s)|}\max_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p} > \epsilon e^{-\alpha k}\sum_{i=1}^{d}\frac{Z_i(s,k)}{|\mathbf{Z}(s)|}\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\le k)}\Big) \\
\le \;& P\Big(\frac{1}{|\mathbf{Z}(s)|}\max_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p} > \epsilon e^{-\alpha k}\sum_{i=1}^{d}\frac{1}{2}B_i(k)\frac{v_i}{\mathbf{1}\cdot\mathbf{v}}\theta_i : A\cap B\Big) + P(A^C) + P(B^C) \\
\le \;& P\Big(\frac{1}{|Z(s)|}\max_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p} > \frac{1}{2}\epsilon e^{-\alpha k}\sum_{i=1}^{d}\theta_i B_i(k)\frac{v_i}{\mathbf{1}\cdot\mathbf{v}}\Big) + \delta
\end{aligned}$$

for every $s > M$. Thus, for any $\delta > 0$, by Lemma 5.1,

$$\begin{aligned}
& \limsup_{s\to\infty} P\Bigg(\frac{e^{\alpha k}\frac{1}{|\mathbf{Z}|(s)}\max\limits_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\frac{Z_i(s,k)}{|\mathbf{Z}(s)|}\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\le k)}} > \epsilon\Bigg) \\
\le \;& \limsup_{s\to\infty} P\Big(\frac{1}{|\mathbf{Z}(s)|}\max_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p} > \frac{1}{2}\epsilon\sum_{i=1}^{d}\theta_i e^{-\alpha k}B_i(k)\frac{v_i}{\mathbf{1}\cdot\mathbf{v}}\Big) + \delta = \delta
\end{aligned}$$

i.e.,

$$\lim_{s\to\infty} P\Bigg(\frac{e^{\alpha k}\frac{1}{|\mathbf{Z}(s)|}\max\limits_{\substack{1\le p\le Z_i(s)\\1\le i\le d}} \tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\frac{Z_i(s,k)}{|\mathbf{Z}(s)|}\frac{1}{Z_i(s,k)}\sum_{p=1}^{Z_i(s)}\tilde{W}_{s,i,p}I_{(R_{s,i,p}\le k)}} > \epsilon\Bigg) = 0 \quad \text{for any } \epsilon > 0$$

and hence, from (5.1), we have that

$$\lim_{s\to\infty} P\Bigg(\frac{\max\limits_{\substack{1\le p\le Z_i(s)\\ 1\le i\le d}} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}} > \epsilon\Bigg) = 0 \quad \text{for any } \epsilon > 0,$$

i.e., as $s \to \infty$,

$$\frac{\max\limits_{\substack{1\le p\le Z_i(s)\\ 1\le i\le d}} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}}{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}} \to 0 \quad \text{in probability.}$$

By the bounded convergence theorem,

$$E\Bigg(\frac{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)}\Big(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}{\Big(\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}\Bigg) \to 0 \quad \text{as } s \to \infty.$$

and thus H_k is a proper probability distribution on $\mathbb{R}_+ \equiv \{x \in \mathbb{R} : x \ge 0\}$. So, there exists a random variable $\tilde{D}_k$ on $\mathbb{R}_+$ such that $D_k(t) \xrightarrow{d} \tilde{D}_k$ as $t \to \infty$ and

$$P(\tilde{D}_k \le s) = 1 - E\Bigg(\frac{\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)}\Big(e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}{\Big(\sum\limits_{i=1}^{d}\sum\limits_{p=1}^{Z_i(s)} e^{-\alpha R_{s,i,p}}\tilde{W}_{s,i,p}\Big)^k}\Bigg) \equiv H_k(s).$$

for any $s \ge 0$. The proof of Theorem 5.1 is complete.

5.2.3 The proof of Theorem 5.2

Let $\{\mathbf{Y}_n\}_{n\ge 0}$ be the embedded generation process of the continuous-time multi-type Bellman-Harris process $\{\mathbf{Z}(t) : t \ge 0\}$.

Let $\{\mathbf{Z}_{r,i,p}(t) : t > 0\}$ be the the continuous-time multi-type age-dependent Bellman-Harris branching process initiated with the pth individual of type i in the rth generation when it is of age 0.

Let $L_{r,i,p,q}$ be the lifetime of the qth-generation ancestor of the pth individual of type i in the rth generation, then $\{L_{r,i,p,q} : r \ge 0, i = 1, 2, \cdots, d, p \ge 1, q = 1, 2, \cdots, r-1\}$ are i.i.d copies with the lifetime distribution G.

Let $S_{r,i,p} = \sum\limits_{q=0}^{r-1} L_{r,i,p,q}$, then $S_{r,i,p}$ is the birth time of the pth individual of type i in the rth generation.

(a) For almost all trees $\mathcal{T}$ and any $r = 0, 1, 2, \cdots$,

$$
\begin{aligned}
& P\Big(X_k(t) \geq r \Big| \mathcal{T}\Big) \\
= \quad & \frac{\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} \binom{Z_{r,i,p}(t-S_{r,i,p})}{k}}{\binom{Z(t)}{k}} \\
= \quad & \frac{\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} Z_{r,i,p}(t-S_{r,i,p})[Z_{r,i,p}(t-S_{r,i,p})-1]\cdots[Z_{r,i}(t-S_{r,i,p})-k+1]}{Z(t)[Z(t)-1]\cdots[Z(t)-k+1]} \qquad (5.10) \\
= \quad & \frac{\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha(t-S_{r,i,p})} Z_{r,i,p}(t-S_{r,i,p}) \cdot e^{-\alpha(t-S_{r,i,p})}[Z_{r,i,p}(t-S_{r,i,p})-1]\cdots}{e^{-\alpha t}Z(t)\cdot e^{-\alpha t}[Z(t)-1]\cdots} \\
& \cdot \frac{e^{-\alpha(t-S_{r,i,p})}[Z_{r,i,p}(t-S_{r,i,p})-k+1]\cdot e^{-k\alpha S_{r,i,p}}}{e^{-\alpha t}[Z(t)-k+1]} \qquad (5.11)
\end{aligned}
$$

where α is the Malthusian parameter for the offspring mean m and the lifetime distribution G.

It known from Theorem 1.17 that if $|\mathbf{Z}_0| = 1$, $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ for any $i = 1, 2, \cdots, d$ and $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$, then

$$e^{-\alpha t}\mathbf{Z}(t) \to \mathbf{v}W \quad \text{w.p.1} \quad \text{as } t \to \infty$$

where W is a random variable such that $P(W > 0) = 1$. So, as $t \to \infty$,

$$P(X_k \geq r \big| \mathcal{T}) \to \frac{\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} \left(e^{-\alpha S_{r,i,p}} W_{r,i,p}\right)^k}{W^k} \equiv 1 - \phi_k(r, \mathcal{T})$$

as $t \to \infty$, where $\left\{W_{r,i,p}\right\}_{i\geq 1, p\geq 1}$ are the i.i.d copies of W.

(b) Since $P(X_k(t) \geq r) = E(P(X_k(t) \geq r) \big| \mathcal{T})$ and hence, by the bounded convergence theorem,

$$P(X_k(t) \geq r) \to E\left(\frac{\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} \left(e^{-\alpha S_{r,i,p}} W_{r,i,p}\right)^k}{W^k}\right) \equiv 1 - \phi_k(r) \quad \text{as } t \to \infty$$

for $r = 1, 2, \cdots$.

To finish the proof, we need to show that ϕ_k is a proper probability distribution, i.e., $\phi_k(r) \to 1$ as $r \to \infty$, and it is sufficient to prove that

$$\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} \left(e^{-\alpha S_{r,i,p}} W_{r,i,p}\right)^k \to 0 \quad \text{in probability} \quad \text{as } r \to \infty.$$

We will follow the lines similar to the proof for the single-type Bellman-Harris process.

First, we have that

$$\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big)^k \le \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}}\Big(e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big)^k \tag{5.12}$$

$$\le \Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p} W_{r,i,p}}\Big)^{k-1} \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} \tag{5.13}$$

and

$$\begin{aligned} & E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big) \\ = \;& E\Big(E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big| L_{r,i,p,q}, 0\le q\le r-1, 1\le p\le Y_{r,i}, 1\le i\le d, \mathbf{Y}_0, \mathbf{Y}_1, \cdots, \mathbf{Y}_r\Big)\Big) \\ = \;& E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} E\Big(W_{r,i,p}\Big| L_{r,i,p,q}, 0\le q\le r-1, 1\le p\le Y_{r,i}, 1\le i\le d, \mathbf{Y}_0, \mathbf{Y}_1, \cdots, \mathbf{Y}_r\Big)\Big) \end{aligned}$$

Note that $\{W_{r,i,p}\}_{p\ge1,1\le i\le d}$ are i.i.d. copies of W and are independent of $\{L_{r,i,p,q}, 0\le q\le r-1, 1\le p\le Y_{r,i}, 1\le i\le d, \mathbf{Y}_0, \mathbf{Y}_1, \cdots, \mathbf{Y}_r\}$, so

$$\begin{aligned} & E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big) \\ = \;& E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} E(W)\Big) = EW\cdot E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}}\Big) \\ = \;& EW\cdot E\Big(E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}}\Big|\mathbf{Y}_r\Big)\Big) = EW\cdot E\Big(|\mathbf{Y}_r| E\Big(e^{-\alpha S_{r,1,1}}\Big|\mathbf{Y}_r\Big)\Big) \\ = \;& EW\cdot E\Big(|\mathbf{Y}_r| E\Big(e^{-\alpha S_{r,1,1}}\Big)\Big) = EW\cdot E|\mathbf{Y}_r|\cdot E\Big(e^{-\alpha S_{r,1,1}}\Big) = EW\cdot E|\mathbf{Y}_r|\cdot\Big(Ee^{-\alpha L}\Big)^r \end{aligned}$$

where $\Big\{S_{r,i,p}\equiv\sum_{q=0}^{r-1} L_{r,i,p,q}\Big\}_{p\ge1,1\le i\le d}$ are identically distributed and $\{L_{r,i,p,q} : 0\le q\le r-1\}$ are i.i.d copies of the lifetime random variable L for each $p\ge1$ and $1\le i\le d$.

Since $E(\|Z_1\|\log\|Z_1\|\big|\mathbf{Z}_0=\mathbf{e}_i)<\infty$, it is known that $0<EW<\infty$. Then,

$$\lim_{r\to\infty} E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big) = \lim_{r\to\infty}\Big(EW\cdot\rho^{-r}E|\mathbf{Y}_r|\cdot(\rho\cdot\varphi_L(\alpha))^r\Big) \tag{5.14}$$

$$= EW\cdot\lim_{r\to\infty}\Big(\rho^{-r}E|\mathbf{Y}_r|\Big) \tag{5.15}$$

$$= cEW \tag{5.16}$$

for some $0 < c < \infty$, where $\varphi_L(\alpha) \equiv \int_0^\infty e^{-\alpha u} dG(u)$ and hence $m\varphi_L(\alpha) = 1$ since α is the Mathusian parameter for m and G.

For any $\eta > 0$, by Chebyshev's inequality,

$$\lim_{r\to\infty} P\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \eta\Big) \le \lim_{r\to\infty} \frac{1}{\eta} E\Big(\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big) = \frac{cEW}{\eta} < \infty.$$

For any $\epsilon > 0$,

$$\begin{aligned}
&\lim_{r\to\infty} P\Big(\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p} W_{r,i,p}}\Big)^{k-1} \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon\Big) \\
= \;& \lim_{r\to\infty} P\Big(\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p} W_{r,i,p}}\Big)^{k-1} \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon, \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \eta\Big) \\
&\quad + \lim_{r\to\infty} P\Big(\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p} W_{r,i,p}}\Big)^{k-1} \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon, \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} W_{r,i,p} \le \eta\Big) \\
\le \;& \frac{cEW}{\eta} + \lim_{r\to\infty} P\Big(\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p} W_{r,i,p}}\Big)^{k-1} > \frac{\epsilon}{\eta}\Big) \qquad (5.17)
\end{aligned}$$

So, to prove that

$$\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} \Big(e^{-\alpha S_{r,i,p}} W_{r,i,p}\Big)^k \to 0 \quad \text{in probability} \quad \text{as } r\to\infty,$$

it suffices, from (5.12) and (5.17), to prove that

$$\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} \to 0 \quad \text{in probability} \quad \text{as } r\to\infty.$$

Let $\mathfrak{F}_r$ be the $\sigma-$algebra generated by all the information up to the rth generation in the embedded tree. Then, for any $\epsilon > 0$,

$$\begin{aligned}
P\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon \Big| \mathfrak{F}_r\Big) &= P\Big(\exists i = 1,2,\cdots,d, \exists p = 1,2,\cdots,Y_{r,i} \text{ s.t. } e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon \Big| \mathfrak{F}_r\Big) \\
&\le \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} P\Big(e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon \Big| \mathfrak{F}_r\Big) \\
&= \sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} P\Big(W_{r,i,p} > \epsilon e^{\alpha S_{r,i,p}} \Big| \mathfrak{F}_r\Big)
\end{aligned}$$

Let $\eta(y) = \sup_{x \geq y} xP(W > x)$. Since $EW < \infty$, $xP(W > x) \to 0$ as $x \to \infty$. So, for any $l > 0$, there exists $a > 0$ s.t. $yP(W > y) < \frac{l\epsilon}{cEW}$ for all $y \geq a$ and hence $\eta(a) \leq \frac{l\epsilon}{cEW}$.
Let $n > \frac{1}{\alpha} \ln \frac{a}{\epsilon}$, then $\epsilon e^{\alpha n} > a$. Hence,

$$
\begin{aligned}
& P\Big(\max_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon \Big) \\
= \; & P\Big(\max_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon, \min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) + P\Big(\max_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon, \min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} > n \Big) \\
\leq \; & P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) + E\Big(P\Big(\max_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon, \min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} > n \Big| \mathfrak{F}_r \Big) \Big) \\
\leq \; & P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) + E\Big(\sum_{i=1}^{d} \sum_{p=1}^{Y_{r,i}} P\Big(W_{r,i,p} > \epsilon e^{\alpha S_{r,i,p}}, \min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} > n \Big| \mathfrak{F}_r \Big) \Big) \\
= \; & P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) + \frac{1}{\epsilon} E\Big(\sum_{i=1}^{d} \sum_{p=1}^{Y_{r,i}} \epsilon e^{\alpha S_{r,i,p}} P\Big(W_{r,i,p} > \epsilon e^{\alpha S_{r,i,p}}, \min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} > n \Big| \mathfrak{F}_r \Big) \cdot e^{-\alpha S_{r,i,p}} \Big) \\
\leq \; & P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) + \frac{1}{\epsilon} E\Big(\sum_{i=1}^{d} \sum_{p=1}^{Y_{r,i}} \eta(a) e^{-\alpha S_{r,i,p}} \Big) \\
= \; & P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) + \frac{1}{\epsilon} \eta(a) E\Big(\sum_{i=1}^{d} \sum_{i=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}} \Big). \qquad (5.18)
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) &= \sum_{\mathbf{x} \in \mathbb{N}_+^d} P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big| \mathbf{Y}_r = \mathbf{x} \Big) P(\mathbf{Y}_r = \mathbf{x}) \\
&\leq \sum_{\mathbf{x} \in \mathbb{N}_+^d} |\mathbf{x}| P(S_{r,1,1} \leq n) P(\mathbf{Y}_r = \mathbf{x}) \\
&= P(S_{r,1,1} \leq n) E|\mathbf{Y}_r| \\
&= P(e^{-\theta S_{r,1,1}} \leq e^{-\theta n}) E|\mathbf{Y}_r|
\end{aligned}
$$

where $\theta > \alpha$ such that $\rho\varphi_L(\theta) < 1$. Then, by Markov inequality,

$$
\begin{aligned}
P\Big(\min_{\substack{1 \leq p < Y_{r,i} \\ 1 \leq i \leq d}} S_{r,i,p} \leq n \Big) &\leq \frac{E\Big(e^{-\theta S_{r,1,1}} \Big)}{e^{-\theta n}} E|\mathbf{Y}_r| = e^{\theta n} \Big(\rho^{-r} E|\mathbf{Y}_r| \Big) \Big(\rho E e^{-\theta L} \Big)^r \\
&= c e^{\theta n} \Big(\rho \varphi_L(\theta) \Big)^r \to 0 \qquad (5.19)
\end{aligned}
$$

as $r \to \infty$.

Since (5.18) and (5.19) together imply that

$$\begin{aligned}
&\lim_{r\to\infty} P\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon\Big) \\
\le\ & \lim_{r\to\infty} P\Big(\min_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} S_{r,i,p} \le n\Big) + \lim_{r\to\infty} \frac{1}{\epsilon}\eta(a)E\Big(\sum_{i=1}^{d}\sum_{i=1}^{Y_{r,i}} e^{-\alpha S_{r,i,p}}\Big) \\
=\ & \lim_{r\to\infty} P\Big(\min_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} S_{r,i,p} \le n\Big) + \frac{1}{\epsilon}\eta(a)cEW \\
\le\ & \lim_{r\to\infty} P\Big(\min_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} S_{r,i,p} \le n\Big) + l \\
=\ & l
\end{aligned}$$

for any $l > 0$. Hence, for any $\epsilon > 0$,

$$P\Big(\max_{\substack{1\le p<Y_{r,i}\\1\le i\le d}} e^{-\alpha S_{r,i,p}} W_{r,i,p} > \epsilon\Big) \to 0 \quad \text{as } r\to\infty,$$

and so ϕ_k is a proper probability distribution and hence there exists a random variable $\tilde{X}_k$ on $\{0, 1, 2, \cdots\}$ such that $X_k(t) \xrightarrow{d} \tilde{X}_k$ as $t \to \infty$ and

$$P(\tilde{X}_k < r) = 1 - E\left(\frac{\sum_{i=1}^{d}\sum_{p=1}^{Y_{r,i}} \left(e^{-\alpha S_{r,i,p}} W_{r,i,p}\right)^k}{W^k}\right) \equiv \phi_k(r).$$

for any $r = 0, 1, 2, \cdots$.

The proof of Theorem 5.2 is complete.

5.2.4 The proof of Theorem 5.3

Let $\mathbf{Y}_n = (Y_{n,1}, Y_{n,2}, \cdots, Y_{n,d})$ be the embedded Galton-Watson branching process.

Let $Y_{n,i}(t)$ be the number of individuals of type i in the nth generation alive at time t. Then

$$Z_i(t) = \sum_{n=0}^{\infty} Y_{n,i}(t).$$

Let $\xi_{n,i,p} = (\xi_{n,i,p,1}, \xi_{n,i,p,2}, \cdots, \xi_{n,i,p,d})$ be the offspring vector of the pth individual of type i in the nth generation.

Let $\xi_{n,i,p}(t)$ be the vector of alive offspring of the pth individual of type i in the nth generation.

Let $\mathbf{Z}_{n,i,p,j,q}(t)$ be the continuous-time multi-type Bellman-Harris branching process initiated by the qth child of type j of the pth individual of type i in the nth generation. Then $\{\mathbf{Z}_{n,i,p,j,q}(t) : t \geq 0\}$ is distributed as $\{\mathbf{Z}(t)|\mathbf{Z}(0) = e_j : t \geq 0\}$.

Let $D_2(t)$ be the death time of the last common ancestor of these two randomly chosen individuals alive at time t. By Theorem 5.1, we have that

$$D_2(t) \xrightarrow{d} \tilde{D}_2 \quad \text{as } t \to \infty.$$

Let $X_2(t)$ be the generation number of this last common ancestor. Recall that every individual has the same lifetime distribution no matter what type it is of according to the assumption. So,

$$X_2(t) \xrightarrow{d} \tilde{X}_2 \quad \text{as } t \to \infty.$$

Let $A_i(t)$ be the type of the ancestor in the next generation after the last common ancestor of the ith chosen individual, $i = 1, 2$. Then, for almost all trees $\mathcal{T}$,

$$\begin{aligned}
& P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = \zeta_2(t) = i_1, A_1(t) = A_2(t)\Big|\mathcal{T}\Big) \\
= \quad & \frac{\sum_{p=1}^{Y_{r,j}-Y_{r,j}(t)} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} Z_{r,j,p,i,m,i_1}(t - D_2(t)) \cdot Z_{r,j,p,i,n,i_1}(t - D_2(t))}{|\mathbf{Z}(t)| \cdot (|\mathbf{Z}(t)| - 1)} \\
\to \quad & E\left(\frac{\sum_{p=1}^{Y_{r,j}} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} v_{i_1} W_{r,p,m} v_{i_1} W_{r,p,n}}{W^2}\right) \quad \text{as } t \to \infty.
\end{aligned}$$

So,

$$\begin{aligned}
& P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = \zeta_2(t) = i_1, A_1(t) = A_2(t)\Big) \\
= \quad & E\Big(P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = \zeta_2(t) = i_1, A_1(t) = A_2(t)\Big|\mathcal{T}\Big)\Big) \\
= \quad & E\left(\frac{\sum_{p=1}^{Y_{r,j}-Y_{r,j}(t)} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} Z_{r,j,p,i,m,i_1}(t - D_2(t)) \cdot Z_{r,j,p,i,n,i_1}(t - D_2(t))}{|\mathbf{Z}(t)| \cdot (|\mathbf{Z}(t)| - 1)}\right) \\
= \quad & E\left(\frac{\sum_{p=1}^{Y_{r,j}-Y_{r,j}(t)} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} e^{-\alpha(t-D_2(t))} Z_{r,j,p,i,m,i_1}(t - D_2(t)) \cdot e^{-\alpha(t-D_2(t))} Z_{r,j,p,i,n,i_1}(t - D_2(t)) e^{-2\alpha D_2(t)}}{e^{-\alpha t}|\mathbf{Z}(t)| \cdot e^{-\alpha t}(|\mathbf{Z}(t)| - 1)}\right) \\
= \quad & E\left(e^{-2\alpha D_2(t)} E\left(\frac{\sum_{p=1}^{Y_{r,j}-Y_{r,j}(t)} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i,i_1}} e^{-\alpha(t-D_2(t))} Z_{r,j,p,i,m,i_1}(t - D_2(t)) \cdot e^{-\alpha(t-D_2(t))} Z_{r,j,p,i,n,i_1}(t - D_2(t))}{e^{-\alpha t}|\mathbf{Z}(t)| \cdot e^{-\alpha t}(|\mathbf{Z}(t)| - 1)}\middle| D_2(t)\right)\right).
\end{aligned}$$

Let $h(D_2(t)) = e^{-2\alpha D_2(t)}$. Note that $0 \le h(D_2(t)) \le 1$.

Let $g(t, D_2(t)) = E\left(\frac{\sum_{p=1}^{Y_{r,j}-Y_{r,j}(t)} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} e^{-\alpha(t-D_2(t))} Z_{r,j,p,i,m,i_1}(t-D_2(t)) \cdot e^{-\alpha(t-D_2(t))} Z_{r,j,p,i,n,i_1}(t-D_2(t))}{e^{-\alpha t}|\mathbf{Z}(t)| \cdot e^{-\alpha t}(|\mathbf{Z}(t)|-1)} \middle| D_2(t) \right)$.

Let $g = E\left(\frac{\sum_{p=1}^{Y_{r,j}} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} v_{i_1} W_{r,p,m} v_{i_1} W_{r,p,n}}{W^2} \right)$. Note that g is a constant.

Then, since $D_2(t) \xrightarrow{d} \tilde{D}(t)$ as $t \to \infty$, by the bounded convergence theorem, we have that

$$g(t, D_2(t)) \to g \quad \text{w.p.1} \quad \text{as } t \to \infty$$

and, also, h is a bounded continuous function, hence

$$Eh(D_2(t)) \to Eh(\tilde{D}_2) \quad \text{as } t \to \infty.$$

Therefore,

$$\begin{aligned}
& \left| E(h(D_2(t) g(t, D_2(t)))) - gEh(\tilde{D}_2) \right| \\
= & \left| E\Big(h(D_2(t))[g(t, D_2(t)) - g] + g[h(D_2(t)) - h(\tilde{D}_2)] \Big) \right| \\
\le & \left| E\Big(h(D_2(t))[g(t, D_2(t)) - g] \Big) \right| + \left| g[Eh(D_2(t)) - Eh(\tilde{D}_2)] \right| \\
\le & Eh(D_2(t)) E\left| g(t, D_2(t)) - g \right| + |g| \left| Eh(D_2(t)) - Eh(\tilde{D}_2) \right| \\
\le & E\left| g(t, D_2(t)) - g \right| + |g| \left| Eh(D_2(t)) - Eh(\tilde{D}_2) \right| \\
\to & 0 \quad \text{as } t \to \infty,
\end{aligned}$$

by the bounded convergence theorem.

That is,

$$\begin{aligned}
& P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = \zeta_2(t) = i_1, A_1(t) = A_2(t) \Big) \\
\to \quad & E\Big(e^{-2\alpha \tilde{D}_2}\Big) E\left(\frac{\sum_{p=1}^{Y_{r,j}} \sum_{i=1}^{d} \sum_{m\neq n=1}^{\xi_{r,j,p,i}} v_{i_1} W_{r,p,m} v_{i_1} W_{r,p,n}}{W^2} \right) \quad \text{as } t \to \infty.
\end{aligned}$$

Similarly, as $t \to \infty$,

$$\begin{aligned}
& P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = \zeta_2(t) = i_1, A_1(t) \neq A_2(t) \Big) \\
\to \quad & E\Big(e^{-2\alpha \tilde{D}_2}\Big) E\left(\frac{\sum_{p=1}^{Y_{r,j}} \sum_{k\neq l=1}^{d} \sum_{m=1}^{\xi_{r,j,p,k}} \sum_{n=1}^{\xi_{r,j,p,l}} v_{i_1} W_{r,p,m} v_{i_1} W_{r,p,n}}{W^2} \right),
\end{aligned}$$

$$P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = i_1, \zeta_2(t) = i_2, i_1 \neq i_2, A_1(t) = A_2(t)\Big)$$
$$\rightarrow \quad E\Big(e^{-2\alpha\tilde{D}_2}\Big)E\Bigg(\frac{\sum_{p=1}^{Y_{r,j}}\sum_{i=1}^{d}\sum_{m\neq n=1}^{\xi_{r,j,p,i}} v_{i_1}W_{r,p,m}v_{i_2}W_{r,p,n}}{W^2}\Bigg)$$

and

$$P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = i_1, \zeta_2(t) = i_2, i_1 \neq i_2 A_1(t) \neq A_2(t)\Big)$$
$$\rightarrow \quad E\Big(e^{-2\alpha\tilde{D}_2}\Big)E\Bigg(\frac{\sum_{p=1}^{Y_{r,j}}\sum_{k\neq l=1}^{d}\sum_{m=1}^{\xi_{r,j,p,k}}\sum_{n=1}^{\xi_{r,j,p,l}} v_{i_1}W_{r,p,m}v_{i_2}W_{r,p,n}}{W^2}\Bigg).$$

So, for any $r \geq 0$ and any $j, i_1, i_2 = 1, 2, \cdots, d$, we have that

$$P\Big(X_2(t) = r, \eta(t) = j, \zeta_1(t) = i_1, \zeta_2(t) = i_2\Big)$$
$$\rightarrow \quad v_{i_1}v_{i_2}E\Big(e^{-2\alpha\tilde{D}_2}\Big)E\Bigg(\frac{\sum_{p=1}^{Y_{r,j}}\sum_{m\neq n=1}^{|\xi_{r,j,p}|} W_{r,p,m}W_{r,p,n}}{W^2}\Bigg) \equiv \varphi_2(r, j, i_1, i_2) \quad \text{as } t \rightarrow \infty.$$

Since $X_2(t) \xrightarrow{d} \tilde{X}_2$ as $t \rightarrow \infty$ and $\tilde{X}_2$ is a proper probability distribution, $\{X_2(t) : t \geq 0\}$ is tight.

Also, $\eta(t)$, $\zeta_1(t)$ and $\zeta_2(t)$ only take values on a finite set $\{1, 2, \cdots, d\}$, so we know that $\{(X_2(t), \eta(t), \zeta_1(t), \zeta_2(t)) : t \geq 0\}$ is tight. Thus, the limit φ_2 is a probability distribution. Hence,

$$\sum_{(r,j,i_1,i_2)} \varphi_2(r, j, i_1, i_2) = 1.$$

The proof is complete.

5.3 The Generation Problem in Supercritical Case

We know that if $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ for any $i = 1, 2, \cdots, d$, then $|\mathbf{Z}(t)| \rightarrow \infty$ w.p.1.

For a continuous-time Bellman-Harris branching process, since the lifetime is a random quantity, individuals alive at time t may belong to different generations. It is clear that the population will grow old as time t gets large but the question is how fast the generation number grows.

Now, we pick an individual at random from those alive at time t, let $M(t)$ be the generation number of this individual. Our interest is that to determine the growth rate of M_t with t.

In this section, we assume that all individuals of various types have the same lifetime distribution G although their offspring distributions may be different.

5.3.1 The statement of Result

Theorem 5.4. *Let* $1 < \rho < \infty$ *and the lifetime distribution* G *is non-lattice with* $G(0+) = 0$ *for* $i = 1, 2, \cdots, d$*. If* $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ *for all* $1 \le i \le d$*. Then,*

$$\frac{M(t)}{t} \to \frac{1}{\mu_\alpha} \quad \text{in probability} \quad \text{as } t \to \infty$$

where $\mu_\alpha = \rho \int_{[0,\infty)} x e^{-\alpha x} dG(x)$*.*

5.3.2 The proof of Theorem 5.4

We need the following lemma to prove the theorem.

Lemma 5.4. *(Athreya, Athreya and Iyer [11]) Let* $\{L_i\}_{i\ge1}$ *be i.i.d. positive random variable with distribution* G *and* $G(0) = 0$*. Let* $\rho > 1$ *and* $0 < \alpha < \infty$ *be the Malthusian parameter given by* $\rho \int_0^\infty e^{-\alpha t} dG(t) = 1$*. Let* $\{\tilde{L}_i\}_{i\ge1}$ *be i.i.d. positive random variables with distribution function* $G_\alpha(x) = \rho \int_0^x e^{-\alpha t} dG(t)$*,* $x \ge 0$*. Let* $S_0 = 0$ *and* $S_n = \sum_{i=1}^{n} L_i$*,* $n \ge 1$*. For* $t \ge 0$*, let* $N(t) = k$ *if* $S_k \le t < S_{k+1}$*. Further, let* $R_t = S_{N(t)+1} - t$ *be the residual lifetime at time* t *for* $\{L_i\}_{i\ge1}$*. Let* $\tilde{N(t)}$ *and* $\tilde{R}_t$ *be the corresponding objects for* $\{\tilde{L}_i\}_{i\ge1}$*. Then*

(a) for any $k \ge 1$*, and any bounded Borel measurable function* $\phi : \mathbb{R}^k \to \mathbb{R}$*,*

$$E\Big(\phi(\tilde{L}_1, \tilde{L}_2, \cdots, \tilde{L}_k)\Big) = E\Big(e^{-\alpha S_k} \rho^k \phi(L_1, L_2, \cdots, L_k)\Big) \quad \text{and}$$

(b) $\lim_{l\to\infty} \lim_{t\to\infty} E\Big(e^{\alpha \tilde{R}_t} : \tilde{R}_t > l\Big) = 0.$

Now, we can begin the proof.

Let $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ for some $i_0 = 1, 2, \cdots, d$.

Let $\mathbf{Y}_n = (Y_{n,1}, Y_{n,2}, \cdots, Y_{n,d})$ be the embedded Galton-Watson branching process.

Let $Y_{n,i}(t)$ be the number of individuals of type i in the nth generation alive at time t. Then

$$Z_i(t) = \sum_{n=0}^{\infty} Y_{n,i}(t).$$

Let $L_{n,i,j}$ be the lifetime of the jth individual of type i in the nth generation.

Let $L_{n,i,j,k}$ be the lifetime of the kth-generation ancestor of the jth individual of type i in the nth generation, then $\{L_{n,i,j,k} : n \geq 0, i = 1, 2, \cdots, d, j \geq 1, k = 1, 2, \cdots, r-1\}$ are i.i.d copies with the lifetime distribution G.

Let $S_{n,i,j} = \sum_{k=0}^{r-1} L_{n,i,j,k}$, then $S_{n,i,j}$ is the birth time of the jth individual of type i in the nth generation.

Recall that ρ is the maximal eigenvalue of the offspring mean matrix $\mathbf{M}$ with right eigenvector $\mathbf{u}$ and left eigenvector $\mathbf{v}$.

Let α be the Mathusian parameter for $\mathbf{M}$ and the lifetime distribution G. Then, since all the individuals have the same lifetime distribution,

$$\rho \int_{[0,\infty)} e^{-\alpha t} dG(t) = 1.$$

Let $dG_\alpha(t) = \rho e^{-\alpha t} dG(t)$, then

$$\mu_\alpha \equiv \int_{[0,\infty) t dG_\alpha(t)} = \rho \int_{[0,\infty)} t e^{-\alpha t} dG(t).$$

For any $c > \frac{1}{\mu_\alpha}$, we have

$$\begin{aligned}
P(M(t) > ct) &= E\Big(\frac{1}{|\mathbf{Z}(t)|} \sum_{i=1}^{d} \sum_{n>ct} Y_{n,i}(t)\Big) \\
&\leq \sum_{i=1}^{d} E\Big(\frac{1}{Z_i(t)} \sum_{n>ct} Y_{n,i}(t)\Big) \\
&= \sum_{i=1}^{d} \Big(E\Big(\frac{1}{Z_i(t)} \sum_{n>ct} Y_{n,i}(t) : Z_i(t) \leq e^{\alpha t} v_i \epsilon\Big) + E\Big(\frac{1}{Z_i(t)} \sum_{n>ct} Y_{n,i}(t)\Big) : Z_i(t) > e^{\alpha t} v_i \epsilon\Big) \\
&\equiv \sum_{i=1}^{d} (a_i(t) + b_i(t)) \qquad (5.20)
\end{aligned}$$

for any $0 < \epsilon < \infty$.

We first claim that $\overline{\lim}_{\epsilon \downarrow 0} \overline{\lim}_{t \to \infty} b_i(t) = 0$ for $1 \leq i \leq d$.

Let

$$\delta_{n,i,r}(t) = \begin{cases} 1, & \text{if the } r\text{th individual of type } i \text{ in the } n\text{th generation is alive at time } t \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned}
E\Big(Y_{n,i}(t)\Big) &= E\Big(\sum_{r=1}^{Y_{n,i}} \delta_{n,i,r}(t)\Big) \\
&= E\Big(E\Big(\sum_{r=1}^{Y_{n,i}} \delta_{n,i,r}(t)\Big| Y_{n,i}, n \geq 0\Big)\Big) \\
&= E\Big(Y_{n,i} E\Big(\delta_{n,i,1}(t)\Big| Y_{n,i}, n \geq 0\Big)\Big) \\
&= E(Y_{n,i}) E(\delta_{n,i,1}(t)) \\
&= m_{i_0 i}^{(n)} P(S_{n,i,1} \leq t < S_{n,i,1} + L_{n,i,1}) \\
&= m_{i_0 i}^{(n)} P(S_n \leq t < S_{n+1}) \\
&= m_{i_0 i}^{(n)} P(N(t) = n)
\end{aligned}$$

where $S_k = \sum_{i=1}^{k} L_i$, $k \geq 1$, $\{L_i\}_{i \geq 1}$ are i.i.d. random variables with the lifetime distribution G and $N(t) = n$ if $S_n \leq t < S_{n+1}$.

So, by Lemma 5.4 (a), we have

$$\begin{aligned}
b_i(t) &= E\Big(\frac{1}{Z_i(t)} \sum_{n>ct} Y_{n,i}(t)\Big) : Z_i(t) > e^{\alpha t} v_i \epsilon\Big) \\
&\leq \frac{1}{e^{\alpha t} v_i \epsilon} \sum_{n>ct} E(Y_{n,i}(t)) \\
&= \frac{e^{-\alpha t}}{v_i \epsilon} \sum_{n>ct} m_{i_0 i}^{(n)} P(S_n \leq t < S_{n+1}) \\
&= \frac{1}{v_i \epsilon \rho} \sum_{n>ct} \frac{m_{i_0 i}^{(n)}}{\rho^n} E\Big(e^{\alpha(S_{n+1}-t)} e^{-\alpha S_{n+1}} \rho^{n+1} I_{(S_n \leq t < S_{n+1})}\Big) \\
&= \frac{1}{v_i \epsilon \rho} \sum_{n>ct} \frac{m_{i_0 i}^{(n)}}{\rho^n} E\Big(e^{\alpha(\tilde{S}_{n+1}-t)} I_{(\tilde{S}_n \leq t < \tilde{S}_{n+1})}\Big) \\
&= \frac{1}{v_i \epsilon \rho} \sum_{n>ct} \frac{m_{i_0 i}^{(n)}}{\rho^n} E\Big(e^{\alpha \tilde{R}_t} I_{(\tilde{N}(t)=n)}\Big).
\end{aligned}$$

Let $\beta_{i_0,i} = \sup_{n \geq 1} \frac{m_{i_0 i}^{(n)}}{\rho^n}$, then

$$\frac{m_{i_0 i}^{(n)}}{\rho^n} \to u_{i_0 i} v_i \quad \text{as } n \to \infty$$

and hence $0 < \beta_{i_0,i} < \infty$. Therefore,

$$\begin{aligned} b_i(t) &\leq \frac{1}{v_i\epsilon\rho}\sum_{n>ct}\frac{m_{i_0 i}^{(n)}}{\rho^n}E\Big(e^{\alpha\tilde{R}_t}I_{(\tilde{N}(t)=n)}\Big) \leq \frac{1}{v_i\epsilon\rho}\sum_{n>ct}\beta_{i_0,i}E\Big(e^{\alpha\tilde{R}_t}I_{(\tilde{N}(t)=n)}\Big) \\ &= \frac{\beta_{i_0,i}}{v_i\epsilon\rho}E\Big(e^{\alpha\tilde{R}_t}I_{(\tilde{N}(t)>ct)}\Big) \end{aligned}$$

From Lemma 5.4 (b), we have that, for any $\epsilon > 0$, there exists an $l > 0$ such that

$$\lim_{t\to\infty} E\Big(e^{\alpha\tilde{R}_t} : \tilde{R}_t > l\Big) < \epsilon^2.$$

So,

$$\begin{aligned} b_i(t) &\leq \frac{\beta_{i_0,i}}{v_i\epsilon\rho}\Big(E\Big(e^{\alpha\tilde{R}_t}I_{(\tilde{N}(t)>ct)} : \tilde{R}_t > l\Big) + E\Big(e^{\alpha\tilde{R}_t}I_{(\tilde{N}(t)>ct)} : \tilde{R}_t \leq l\Big)\Big) \\ &\leq \frac{\beta_{i_0,i}}{v_i\epsilon\rho}\Big(E\Big(e^{\alpha\tilde{R}_t} : \tilde{R}_t > l\Big) + e^{\alpha l}P(\tilde{N}(t) > ct)\Big). \end{aligned}$$

Moreover, by the strong law of large numbers,

$$\frac{\tilde{N}(t)}{t} \to \frac{1}{\mu_\alpha} \quad \text{w.p.1} \quad \text{as } t \to \infty$$

and hence

$$\overline{\lim_{t\to\infty}} P\Big(\frac{\tilde{N}(t)}{t} > c\Big) = 0 \quad \text{for any } c > \frac{1}{\mu_\alpha}.$$

Therefore, we have that

$$0 \leq \overline{\lim_{t\to\infty}} b_i(t) \leq \frac{\beta_{i_0,i}}{v_i\epsilon\rho}\Big(\overline{\lim_{t\to\infty}} E\Big(e^{\alpha\tilde{R}_t} : \tilde{R}_t > l\Big) + e^{\alpha l}\overline{\lim_{t\to\infty}} P(\tilde{N}(t) > ct)\Big) < \frac{\beta_{i_0,i}\epsilon}{v_i\rho}$$

for any $1 \leq i \leq d$. Hence

$$\overline{\lim_{\epsilon\downarrow 0}}\ \overline{\lim_{t\to\infty}} b_i(t) = 0 \quad \text{for all } i = 1, 2, \cdots, d.$$

Next, we claim that $\overline{\lim_{\epsilon\downarrow 0}}\ \overline{\lim_{t\to\infty}} a_i(t) = 0$ for all $i = 1, 2, \cdots, d$.

Since

$$a_i(t) \equiv E\Big(\frac{1}{Z_i(t)}\sum_{n>ct} Y_{n,i}(t) : Z_i(t) \leq e^{\alpha t}v_i\epsilon\Big) \leq P\Big(Z_i(t) < e^{\alpha t}v_i\epsilon\Big)$$

and

$$e^{-\alpha t}Z_i(t) \to v_i W \quad \text{w.p.1} \quad \text{as } t \to \infty$$

and, under the assumptions that $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ and $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$, $P(0 < W < \infty) = 1$. Then $P(W \leq \epsilon) \to 0$ as $\epsilon \downarrow 0$. Hence,

$$\overline{\lim_{\epsilon \downarrow 0}}\ \overline{\lim_{t \to \infty}} a_i(t) \leq \overline{\lim_{\epsilon \downarrow 0}}\ \overline{\lim_{t \to \infty}} P\Big(Z_i(t) < e^{\alpha t} v_i \epsilon\Big) = \overline{\lim_{\epsilon \downarrow 0}}\ \overline{\lim_{t \to \infty}} P(W \leq \epsilon) = 0$$

for all $1 \leq i \leq d$.

Then, from (5.20), we have that

$$0 \leq \overline{\lim_{t \to \infty}} P(M(t) > ct) \leq \sum_{i=1}^{d} \left(\overline{\lim_{t \to \infty}} a_i(t) + \overline{\lim_{t \to \infty}} b_i(t) \right)$$

for any $\epsilon > 0$ and hence,

$$0 \leq \overline{\lim_{t \to \infty}} P(M(t) > ct) \leq \sum_{i=1}^{d} \left(\overline{\lim_{\epsilon \downarrow 0}}\ \overline{\lim_{t \to \infty}} a_i(t) + \overline{\lim_{\epsilon \downarrow 0}}\ \overline{\lim_{t \to \infty}} b_i(t) \right) = 0,$$

i.e., $\lim_{t \to \infty} P(M(t) > ct) = 0$ for any $c > \dfrac{1}{\mu_\alpha}$.

By the similar argument, we can prove that,

$$\lim_{t \to \infty} P(M(t) < ct) = 0$$

for any $c < \dfrac{1}{\mu_\alpha}$.

Since, for any $\epsilon > 0$,

$$P\left(\left| \frac{M(t)}{t} - \frac{1}{\mu_\alpha} \right| > \epsilon \right) = P\left(\frac{M(t)}{t} > \frac{1}{\mu_\alpha} + \epsilon \right) + P\left(\frac{M(t)}{t} < \frac{1}{\mu_\alpha} - \epsilon \right),$$

we have that

$$\lim_{t \to \infty} P\left(\left| \frac{M(t)}{t} - \frac{1}{\mu_\alpha} \right| > \epsilon \right) = \lim_{t \to \infty} P\left(\frac{M(t)}{t} > \frac{1}{\mu_\alpha} + \epsilon \right) + \lim_{t \to \infty} P\left(\frac{M(t)}{t} < \frac{1}{\mu_\alpha} - \epsilon \right) = 0$$

for any $\epsilon > 0$. So,

$$\frac{M(t)}{t} \to \frac{1}{\mu_\alpha} \quad \text{in probability} \quad \text{as } t \to \infty$$

and hence the proof is complete.

CHAPTER 6. APPLICATION TO BRANCHING RANDOM WALKS

6.1 Introduction

A branching random walk is a branching tree such that with each line of descent a random walk is associated.

Let $\{Z_n\}_{n\geq 0}$ be a discrete-time single-type Galton-Watson branching process with offspring distribution $\{p_j\}_{j\geq 0}$. Let $Z_0 = 1$, then there is a unique probability measure on the family tree initiated by this ancestor.

On this family tree, we impose the following movement structure.

If an individual is located at x in the real line $\mathbb{R}$, and, upon death, produces k children, then these k children move to $x + X_{kj}$, for $1 \leq j \leq k$, where $(X_{k1}, X_{k2}, \cdots, X_{kk})$ is a random vector with a joint distribution π_k on $\mathbb{R}^k$, for each k. The random vector $X_k \equiv (X_{k1}, X_{k2}, \cdots, X_{kk})$ is stochastically independent of the history up to that generation as well as the movement of the offspring of other individuals.

Let $\zeta_n \equiv \{x_{ni} : 1 \leq i \leq Z_n\}$ be the positions of the Z_n individuals of the n-th generation. For each $n \geq 0$, ζ_n is a collection of random numbers of random points on $\mathbb{R}$ and hence is a point process. The sequence of pairs of $\{Z_n, \zeta_n\}_{n\geq 0}$ is called *branching random walk*. The probability distribution of this process is completely specified by

1. the offspring distribution $\{p_j\}_{j\geq 0}$;

2. the family of probability measures $\{\pi_k\}_{k\geq 1}$;

3. the initial population size Z_0; and

4. the locations $\zeta_0 \equiv \{x_{0i}, 1 \leq i \leq Z_0\}$ of the initial ancestors.

It is clear that $\{\zeta_n\}_{n\geq 0}$ is also a Markov chain whose state space is the set of all finite subsets of $\mathbb{R}$.

The problem of our interest is what happens to the point process ζ_n as $n \to \infty$. In particular,

(1) If $Z_n(x)$ is the number of points in ζ_n that are less than or equal to x, then how does $Z_n(x)$ behave as $n \to \infty$?

(2) Does there exist $\{x_n\}_{n\geq 0}$ such that the proportion $\frac{Z_n(x_n)}{Z_n}$ has a nontrivial limit as $n \to \infty$?

It is clear that the movement along any one line of descent is that of a classical random walk. Thus, if X_{ki} are identically distributed with mean μ and finite variance σ^2 then the location of an individual of the n-th generation should be approximately Gaussian with mean $n\mu$ and variance $n\sigma^2$ by the central limit theorem.

This suggests that if $Z_n \to \infty$ as $n \to \infty$ and if $x_n = \sigma\sqrt{n}x + n\mu$, then $\frac{Z_n(x_n)}{Z_n}$ could have $\Phi(x)$, the standard $N(0,1)$ cumulative distribution function (c.d.f.), as its limit. Or, if $X_{k,1} \in D(\alpha)$ with $0 < \alpha \leq 2$, i.e., $X_{k,1}$ is in the domain of attraction of a stable law of order α, then there exist a_n and b_n such that $\frac{Z_n(a_n + b_n y)}{Z_n}$ converges to a standard stable law c.d.f. as $n \to \infty$. It turns out to be true in the supercritical case ($1 < m = \sum_{j=1}^{\infty} jp_j < \infty$), but the same result doesn't hold for the explosive case ($m = \infty$).

Recall that the limit of the coalescence time of two randomly chosen individuals in the nth generation in the supercritical Galton-Watson branching process is very close to the beginning of the tree but its rate of growth is n in the explosive case when n gets large. Surprisingly, this causes the difference on the limit behavior of the proportion $\frac{Z_n(x_n)}{Z_n}$ between the supercritical and explosive cases.

6.2 Review of Results in The Supercritical Case

Consider a supercritical Galton-Watson branching process $\{Z_n\}_{n\geq 0}$, the following theorems are results on its corresponding branching random walk.

Theorem 6.1. *(Athreya [9]) Let* $p_0 = 0$, $1 < m \equiv \sum_{j=1}^{\infty} jp_j < \infty$ *and* π_k *be such that* $\{X_{k,i} : i = 1, 2, \cdots, k\}_{k\geq 1}$ *are identically distributed. Let* $EX_{k,1} = 0$ *and* $EX_{k,1}^2 = \sigma^2 < \infty$. *Then,*

(a) for any $y \in \mathbb{R}$,

$$\frac{Z_n(\sqrt{n}\sigma y)}{Z_n} \to \Phi(y) \quad \textit{in mean square,}$$

where $\Phi(y)$ *is the c.d.f. of the standard normal* $N(0,1)$.

(b) if Y_n is the position of a randomly chosen individual from the nth generation, then, for any $y \in \mathbb{R}$,

$$P(Y_n \leq \sqrt{n}\sigma y) \to \Phi(y).$$

Theorem 6.2. *(Athreya [9]) Let $p_0 = 0$, $1 < m \equiv \sum_{j=1}^{\infty} jp_j < \infty$ and π_k be such that $\{X_{k,i} : i = 1, 2, \cdots, k\}_{k\geq 1}$ are identically distributed. Let $X_{k,1} \in D(\alpha)$, $0 < \alpha \leq 2$, then*

(a) there exist a_n, b_n such that

$$\frac{Z_n(a_n + b_n y)}{Z_n} \to G_\alpha(y) \quad \text{in mean square,}$$

where $G_\alpha(\cdot)$ is a standard stable law c.d.f. (of order α).

(b) if Y_n is the position of a randomly chosen individual from the nth generation, then, for any $y \in \mathbb{R}$,

$$P(Y_n \leq \sqrt{n}\sigma y) \to \Phi(y).$$

The results depend on the fact when $p_0 = 0$ and $1 < m \equiv \sum_{j=1}^{\infty} jp_j < \infty$, the coalescence time $X_{n,2}$ is way back in time and so the positions of two randomly chosen individuals in the nth generation are essentially independent and has the marginal distribution of a random walk at step n.

Remark 6.1. Theorem 6.1 and Theorem 6.2 holds under the following weaker assumption about π_k, the distribution of $(X_{k,1}, X_{x,2}, \cdots, X_{k,k})$, that does not require $\{X_{k,1}\}_{k\geq 1}$ to be identically distributed. It suffices to assume:

(i) $\forall k \geq 1$, $(X_{k,1}, X_{x,2}, \cdots, X_{k,k})$ has a distribution that is invariant under permutation.

(ii) If $\{p_k\}_{k\geq 1}$ is the offspring distribution with

$$\sum_{k=1}^{\infty} p_k E X_{k,1}^2 < \infty, \quad 1 < m = \sum_{k=1}^{\infty} kp_k < \infty, \quad p_0 = 0.$$

6.3 Results in The Explosive Case

In this section, we consider the explosive Galton-Watson branching process such that the offspring distribution $\{p_j\}_{j\geq 0}$ is in the domain of a stable law of order α with $0 < \alpha < 1$.

6.3.1 The Statements of Theorems in The Explosive Case

First, we pick an individual at random from the nth generation.

Recall the following notations:

(1) Y_n is the position of this randomly chosen individual;

(2) $Z_n(x)$ is the number of points in ζ_n that are less than or equal to x for any $x \in \mathbb{R}$;

(3) $X_k \equiv (X_{k,1}, X_{k,2}, \cdots, X_{k,k})$ are the movements of all the offspring of an individual with k offspring and have the joint distribution π_k;

(4) $\zeta_n \equiv \{x_{ni} : 1 \le i \le Z_n\}$ are the positions of the Z_n individuals of the n-th generation.

Theorem 6.3. *Let $m = \infty$, $p_0 = 1$, $\{p_j\}_{j\ge0} \in D(\alpha)$, $0 < \alpha < 1$. Let $\{X_{k,i} : 1 \le i \le k\}_{k\ge1}$ be identically distributed. Let $EX_{k,1} = 0$ and $EX_{k,1}^2 = \sigma^2 < \infty$. Then, for any fixed $y \in \mathbb{R}$,*

(a) $P(Y_n \le \sqrt{n}\sigma y) \to \Phi(y)$ *as* $n \to \infty$;

(b) $\dfrac{Z_n(\sqrt{n}\sigma y)}{Z_n} \xrightarrow{d} \delta_y$ *as* $n \to \infty$, *where* δ_y *is Bernoulli*$(\Phi(y))$, *i.e.* $P(\delta_y = 1) = \Phi(y) = 1 - P(\delta_y = 0)$.

The result in Theorem 6.3 (b) can be strengthened to the joint convergence of

$$\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}, \quad i = 1, 2, \cdots, k,$$

and hence we have the following theorem.

Theorem 6.4. *Under the hypothesis of Theorem 6.3,*

(a) for any $-\infty < y_1 < y_2 < \infty$,

$$\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n}, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n}\right) \xrightarrow{d} (\delta_1(\Phi(y_1)), \delta_2(\Phi(y_2)))$$

which takes values $(0,0)$, $(0,1)$ *and* $(1,1)$ *with probabilities* $1 - \Phi(y_2)$, $\Phi(y_2) - \Phi(y_1)$ *and* $\Phi(y_1)$, *respectively.*

(b) for any $-\infty < y_1 < y_2 < \cdots < y_k < \infty$,

$$\left(\frac{Z_n(\sqrt{n}\sigma y_i)}{Z_n} : 1 \le i \le k\right) \xrightarrow{d} (\delta_1, \cdots, \delta_k))$$

where each δ_i *is* 0 *or* 1 *and further* $\delta_i = 1 \Rightarrow \delta_j = 1$ *for* $j \geq i$ *and*

$$\begin{aligned} & P(\delta_1 = 0, \delta_2 = 0, \cdots, \delta_{j-1} = 0, \delta_j = 1, \cdots, \delta_k = 1) \\ = \quad & P(\delta_{j-1} = 0, \delta_j = 1) = \Phi(y_j) - \Phi(y_{j-1}). \end{aligned}$$

Remark 6.2. Theorem 6.4 suggests that

$$\left\{ Z_n(y) = \frac{Z_n(\sqrt{n}\sigma y)}{Z_n}, -\infty < y < \infty \right\}$$

converges in the Skorohod Space $D(-\infty, \infty)$ weakly to

$$\{X(y) \equiv I_{(N \leq y)}, -\infty < y < \infty\}$$

where N is a $N(0, 1)$ r.v.

So, we have the following result and only tightness needs to be established:

If Y_n is the position of a randomly chosen individual in the nth generation, then in all cases (as long as $p_0 = 0$), given the tree (random walk) $\mathcal{T}$, $\forall y \in \mathbb{R}$,

$$P(Y_n \leq \sqrt{n}\sigma y | \mathcal{T}) \xrightarrow{d} \delta_y \sim Bernoulli(\Phi(y)).$$

6.3.2 The Proof of Theorem 6.3

To prove Theorem 6.3, we need the following lemma.

Lemma 6.1. *Let* $\{\mu_n\}_{n \geq 1}$ *be probability distributions on* $[0, 1]$ *and such that, as* $n \to \infty$*,*

$$\int_{[0,1]} x d\mu_n \to \lambda \quad and \quad \int_{[0,1]} x^2 d\mu_n \to \lambda$$

for some $0 < \lambda < 1$*. Then,*

$$\mu_n \xrightarrow{w} \mu \quad as\ n \to \infty,$$

where μ *is a probability distribution on* $[0, 1]$ *with* $\mu(\{1\}) = \lambda$ *and* $\mu(\{0\}) = 1 - \lambda$*.*

Proof. First, we have that, for any $n \in \mathbb{N}$,

$$\int_{[0,1]} x d\mu_n = \mu_n(\{1\}) + \int_{(0,1)} x dx$$

and

$$\int_{[0,1]} x^2 d\mu_n = \mu_n(\{1\}) + \int_{(0,1)} x^2 dx.$$

So,

$$\int_{(0,1)} x d\mu_n - \int_{(0,1)} x^2 d\mu_n = \int_{[0,1]} x dx - \int_{[0,1]} x^2 d\mu_n$$

and hence

$$\begin{aligned} & \lim_{n\to\infty} \int_{(0,1)} x - x^2 d\mu_n \\ = & \lim_{n\to\infty} \int_{(0,1)} x d\mu_n - \lim_{n\to\infty} \int_{(0,1)} x^2 d\mu_n = \lim_{n\to\infty} \int_{[0,1]} x dx - \lim_{n\to\infty} \int_{[0,1]} x d\mu_n \\ = & \lambda - \lambda = 0 \end{aligned}$$

Now, for any $0 < a < b < 1$, we have that

$$\int_{(0,1)} x - x^2 d\mu_n \geq \int_{(a,b]} x - x^2 d\mu_n \geq \int_{(a,b]} r - r^2 d\mu_n = (r - r^2)\mu_n((a,b])$$

where $r = \min\left\{a, 1-b\right\}$ and thus

$$\lim_{n\to\infty} \mu_n((a,b]) \leq \frac{1}{r - r^2} \int_{(0,1)} x - x^2 d\mu_n = 0 = \mu((a,b]).$$

Since $\lim_{n\to\infty} \mu_n((a,b]) = \mu((a,b])$ for any $a, b \in [0,1]$ with $\mu(\{a\}) = \mu(\{b\}) = 0$, we have that

$$\mu_n \xrightarrow{v} \mu \quad \text{as } n \to \infty.$$

Also, μ is a probability measure on $[0,1]$, so

$$\mu_n \xrightarrow{w} \mu \quad \text{as } n \to \infty.$$

□

Now, we begin the proof of Theorem 6.3.

(a) Recall that $\zeta_n \equiv \{x_{ni} : 1 \leq i \leq Z_n\}$ are the positions of the Z_n individuals of the n-th generation.

For any fixed $y \in \mathbb{R}$, let

$$\delta_{n,i} = \begin{cases} 1 & \text{, if } x_{n,i} \leq \sqrt{n}\sigma y \\ 0 & \text{, otherwise.} \end{cases}$$

Then we have that

$$Z_n(\sqrt{n}\sigma y) = \sum_{i=1}^{Z_n} \delta_{n,i}.$$

So,

$$\begin{aligned} E\Big(\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}\Big) &= E\Big(\frac{1}{Z_n}\sum_{i=1}^{Z_n}\delta_{n,i}\Big) = E\Big(\frac{1}{Z_n}\sum_{i=1}^{Z_n}E\Big(\delta_{n,i}\Big|Z_n\Big)\Big) = E\Big(\frac{1}{Z_n}\sum_{i=1}^{Z_n}E\Big(\delta_{n,1}\Big)\Big) \\ &= E(\delta_{n,1}) = P(x_{n,1} \le \sqrt{n}\sigma y) = P(x_{0,1} + S_n \le \sqrt{n}\sigma y) \\ &= P(S_n \le \sqrt{n}\sigma y - x_{0,1}) \end{aligned}$$

where $S_n = \sum_{i=1}^{n} \eta_i$, $\{\eta_i\}_{i\ge 1}$ are i.i.d copies with distribution π_1 and $x_{0,1}$ is the location of the initial ancestor of the nth generation individual located at the position $x_{n,1}$. Since $EX_{k,1} = 0$ and $EX_{k,1}^2 = \sigma^2 < \infty$, by the central limit theorem, we have

$$P\Big(\frac{S_n}{\sqrt{n}\sigma} \le y - \frac{x_{0,1}}{\sqrt{n}\sigma}\Big) \to \Phi(y) \quad \text{as } n \to \infty.$$

Hence,

$$\begin{aligned} P(Y_n \le \sqrt{n}\sigma y) &= P(Y_n \le \sqrt{n}\sigma y \big| Z_n) = E\Big(\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}\Big) \\ &\to \Phi(y) \quad \text{as } n \to \infty. \end{aligned}$$

(b) We will prove that, for any $y \in \mathbb{R}$,

$$\frac{Z_n(\sqrt{n}\sigma y)}{Z_n} \xrightarrow{d} Bernoulli(\Phi(y)) \quad \text{as } n \to \infty.$$

From (a), we already know that, for any fixed $y \in \mathbb{R}$,

$$E\Big(\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}\Big) \to \Phi(y) \quad \text{as } n \to \infty.$$

It suffices to show that, for any fixed $y \in \mathbb{R}$, we also have

$$E\Big(\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}\Big)^2 \to \Phi(y) \quad \text{as } n \to \infty.$$

Recall that, for any fixed $y \in \mathbb{R}$,

$$\delta_{n,i} = \begin{cases} 1 & \text{, if } x_{n,i} \le \sqrt{n}\sigma y \\ 0 & \text{, otherwise.} \end{cases}$$

and then

$$E\Big(\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}\Big)^2 = E\Big(\frac{1}{Z_n}\sum_{i=1}^{Z_n}\delta_{n,i}\Big) = E\Big(\frac{1}{Z_n^2}\sum_{i=1}^{Z_n}\delta_{n,i}^2\Big) + E\Big(\frac{1}{Z_n^2}\sum_{i\neq j=1}^{Z_n}\delta_{n,i}\delta_{n,j}\Big).$$

First, it is known that, in the explosive case under the assumption that $p_0 = 0$, $P(Z_n \to \infty) = 1$. Also, we have that

$$P\Big(0 < \frac{1}{Z_n^2}\sum_{i=1}^{Z_n}\delta_{n,i}^2 < \frac{1}{Z_n}\Big) = 1.$$

Hence,

$$P\Big(\frac{1}{Z_n^2}\sum_{i=1}^{Z_n}\delta_{n,i}^2 \to 0\Big) = 1$$

and then, by the bounded convergence theorem,

$$E\Big(\frac{1}{Z_n^2}\sum_{i=1}^{Z_n}\delta_{n,i}^2\Big) \to 0 \quad \text{as } n \to \infty. \tag{6.1}$$

Secondly, by the symmetry consideration conditioned on the branching tree (but not the random walk), we have that

$$\begin{aligned} & E\Big(\frac{1}{Z_n^2}\sum_{i\neq j=1}^{Z_n}\delta_{n,i}\delta_{n,j}\Big) \\ = \; & E\Big(\frac{1}{Z_n^2}\sum_{i\neq j=1}^{Z_n}E\Big(\delta_{n,i}\delta_{n,j}\Big|Z_n\Big)\Big) = E\Big(\frac{1}{Z_n^2}\sum_{i\neq j=1}^{Z_n}E\Big(\delta_{n,1}\delta_{n,2}\Big|Z_n\Big)\Big) = E\Big(\frac{Z_n(Z_n-1)}{Z_n^2}E\Big(\delta_{n,1}\delta_{n,2}\Big)\Big) \\ = \; & E\Big(\frac{Z_n(Z_n-1)}{Z_n^2}\Big)E\Big(\delta_{n,1}\delta_{n,2}\Big) \end{aligned}$$

Note that, by the bounded convergence theorem,

$$E\Big(\frac{Z_n(Z_n-1)}{Z_n^2}\Big) \to 1 \quad \text{as } n \to \infty. \tag{6.2}$$

Now, let $\tau_{n,2}$ be the generation number of the last common ancestor of any two randomly chosen individuals in the nth generation. Then, by Theorem 2.4, we have

$$n - \tau_{n,2} \xrightarrow{d} \tilde{\tau}_2 \quad \text{as } n \to \infty$$

for some random variable $\tilde{\tau}_2$.

Let x_{τ_n} be the position of the last common ancestor of these two individuals corresponding to the positions $x_{n,1}$ and $x_{n,2}$. Then we can write

$$x_{n,i} = x_{\tau_n} + Y_{n,i} \quad i = 1, 2$$

where $Y_{n,i}$ is the net displacement of the individual with position $x_{n,i}$ from generation τ_n to n. Clearly, $Y_{n,1}$ and $Y_{n,2}$ are independent. Moreover, x_{τ_n}, $Y_{n,1}$ and $Y_{n,2}$ can be written as

$$x_{\tau_n} = x_{0,1} + \sum_{j=1}^{\tau_{n,2}} \eta_j \quad \text{and} \quad Y_{n,i} = \sum_{j=1}^{n-\tau_{n,2}} \eta_{i,j} \quad \text{for } i = 1, 2$$

respectively, where $\{\eta_j\}_{j\geq 1}$, $\{\eta_{1,i}\}_{j\geq 1}$ and $\{\eta_{2,i}\}_{j\geq 1}$ are i.i.d copies with distribution π_1 and are independent with each other.

Therefore,

$$\begin{aligned}
& E(\delta_{n,1}\delta_{n,2}) \\
= \; & E\Big(E(\delta_{n,1}\delta_{n,2}\Big|n-\tau_{n,2})\Big) \\
= \; & E\Big(E\Big(I_{(x_{n,1}\leq\sqrt{n}\sigma y)} I_{(x_{n,2}\leq\sqrt{n}\sigma y)}\Big|n-\tau_{n,2}\Big)\Big) \\
= \; & E\Big(E\Big(I_{(x_{0,1}+\sum_{j=1}^{\tau_{n,2}}\eta_j+\sum_{j=1}^{n-\tau_{n,2}}\eta_{1,j}\leq\sqrt{n}\sigma y)} I_{(x_{0,1}+\sum_{j=1}^{\tau_{n,2}}\eta_j+\sum_{j=1}^{n-\tau_{n,2}}\eta_{2,j}\leq\sqrt{n}\sigma y)}\Big|n-\tau_{n,2}\Big)\Big) \\
= \; & E\Big(E\Big(I_{(\sum_{j=1}^{\tau_{n,2}}\eta_j\leq\sqrt{n}\sigma y-x_{0,1}-\sum_{j=1}^{n-\tau_{n,2}}\eta_{1,j})} I_{(\sum_{j=1}^{\tau_{n,2}}\eta_j\leq\sqrt{n}\sigma y-x_{0,1}-\sum_{j=1}^{n-\tau_{n,2}}\eta_{2,j})}\Big|n-\tau_{n,2}\Big)\Big) \\
= \; & E\Big(P\Big(\sum_{j=1}^{\tau_{n,2}}\eta_j \leq \sqrt{n}\sigma y - x_{0,1} - \max\Big\{\sum_{j=1}^{n-\tau_{n,2}}\eta_{1,j}, \sum_{j=1}^{n-\tau_{n,2}}\eta_{2,j}\Big\}\Big|n-\tau_{n,2}\Big)\Big)
\end{aligned}$$

Since $n - \tau_{n,2} \xrightarrow{d} \tilde{\tau}_2$ as $n \to \infty$ and $P(\tilde{\tau}_2 < \infty) = 1$, we have that, for $i = 1, 2$,

$$\sum_{j=1}^{n-\tau_{n,2}} \eta_{i,j} \xrightarrow{d} \sum_{j=1}^{\tilde{\tau}_2} \eta_{i,j} \quad \text{as } n \to \infty.$$

Also, $\tau_{n,2} \xrightarrow{d} \infty$ and $\dfrac{\tau_{n,2}}{n} \xrightarrow{d} 1$ as $n \to \infty$. Hence, as $n \to \infty$,

$$P\Big(\sum_{j=1}^{\tau_{n,2}}\eta_j \leq \sqrt{n}\sigma y - x_{0,1} - \max\Big\{\sum_{j=1}^{n-\tau_{n,2}}\eta_{1,j}, \sum_{j=1}^{n-\tau_{n,2}}\eta_{2,j}\Big\}\Big|n-\tau_{n,2}\Big) \to \Phi(y) \quad \text{w.p.1.}$$

Then, by the bounded convergence theorem,

$$E(\delta_{n,1}\delta_{n,2}) \to \Phi(y) \quad \text{as } n \to \infty. \tag{6.3}$$

So, (6.1), (6.2) and (6.3) together imply that

$$E\left(\frac{Z_n(\sqrt{n}\sigma y)}{Z_n}\right)^2 \to \Phi(y) \quad \text{as } n \to \infty.$$

By Lemma 6.1, we have that, for any $y \in \mathbb{R}$,

$$\frac{Z_n(\sqrt{n}\sigma y)}{Z_n} \xrightarrow{d} Bernoulli(\Phi(y)) \quad \text{as } n \to \infty$$

and hence the proof is complete.

6.3.3 The Proof of Theorem 6.4

(a) Let $-\infty < y_1 < y_2 < \infty$ be any two fixed real numbers. Then,

$$P\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} \leq \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n}\right) = 1.$$

So,

$$P\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 0\right) = 0$$

for any $n = 1, 2, \cdots$, and hence

$$P\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 0\right) \to 0$$

as $n \to \infty$.

Also, by Theorem 6.3, we have that

$$\frac{Z_n(\sqrt{n}\sigma y_i)}{Z_n} \xrightarrow{d} \delta_i(\Phi(y_i)) \quad \text{as } n \to \infty$$

where $\delta_i(\Phi(y_i))$ is a Bernoulli random variable with $P(\delta_i(\Phi(y_i)) = 1) = \Phi(y_i) = 1 - P(\delta_i(\Phi(y_i)) = 0)$ for $i = 1, 2$.

Therefore, as $n \to \infty$,

$$P\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 0, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 0\right) = P\left(\frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 0\right) \to 1 - \Phi(y_2)$$

and

$$P\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 1\right) = P\left(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1\right) \to \Phi(y_1).$$

Moreover, since $\Big(\delta_1(\Phi(y_1)), \delta_2(\Phi(y_2))\Big)$ only take values on the set $\Big\{(0,0),(0,1),(1,0),(1,1)\Big\}$,

$$P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 0, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 1\bigg) \to \Phi(y_2) - \Phi(y_1).$$

Hence, (a) is proved.

(b) Let $k \in \mathbb{N}$ be any positive integer and $-\infty < y_1 < y_2 < \cdots < y_k < \infty$ be any fixed real numbers.

Let $i_1, i_2, \cdots, i_k \in \{0, 1\}$. If there exist l, m with $l < m$ such that $i_l = 1$ and $i_m = 0$, then

$$P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = i_1, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = i_2, \cdots, \frac{Z_n(\sqrt{n}\sigma y_k)}{Z_n} = i_k\bigg) = 0$$

for any $n = 1, 2, \cdots$ and hance

$$P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = i_1, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = i_2, \cdots, \frac{Z_n(\sqrt{n}\sigma y_k)}{Z_n} = i_k\bigg) \to 0$$

as $n \to \infty$.

Secondly, if $i_1 = i_2 = \cdots = i_k = 1$, then

$$P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 1, \cdots, \frac{Z_n(\sqrt{n}\sigma y_k)}{Z_n} = 1\bigg) = P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1\bigg) \to \Phi(y_1)$$

as $n \to \infty$, by Theorem 6.3.

Also, if $i_1 = i_2 = \cdots = i_k = 0$, then, by Theorem 6.3 again,

$$P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 0, \frac{Z_n(\sqrt{n}\sigma y_2)}{Z_n} = 0, \cdots, \frac{Z_n(\sqrt{n}\sigma y_k)}{Z_n} = 0\bigg) = P\bigg(\frac{Z_n(\sqrt{n}\sigma y_k)}{Z_n} = 0\bigg) \to 1 - \Phi(y_k)$$

as $n \to \infty$.

Moreover, if $i_1 = \cdots = i_{j-1} = 0 < 1 = i_j = \cdots = i_k$ for some $1 < j < k$, then

$$\begin{aligned} & P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 0, \cdots, \frac{Z_n(\sqrt{n}\sigma y_{j-1})}{Z_n} = 0, \frac{Z_n(\sqrt{n}\sigma y_j)}{Z_n} = 1, \cdots, \frac{Z_n(\sqrt{n}\sigma y_k)}{Z_n} = 1\bigg) \\ = \quad & P\bigg(\frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 0, \frac{Z_n(\sqrt{n}\sigma y_1)}{Z_n} = 1\bigg) \to \Phi(y_j) - \Phi(y_{j-1}) \end{aligned}$$

as $n \to \infty$, by (a). Therefore, the proof of part (b) is complete.

CHAPTER 7. OPEN PROBLEMS

7.1 Problems in Discrete-time Multi-type Galton-Watson Branching Processes

1. In the critical case, we are able to prove the results on the coalescence times using the convergence of the point process, but the problem regarding limit behavior of the joint distribution of the generation number and the type of the last common ancestor and the types of the randomly chosen individuals remains open.

2. The coalescence problem in the explosive case is still open.

7.2 Problems in Continuous-time Single-type Bellman-Harris Branching Processes

1. For the proofs in Chapter 4, we impose the condition $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$. Can we still prove the results without this hypotheses?

2. The coalescence problems in the critical case are open including the results on the limit distributions of the generation number and the death time of the last common ancestor of the randomly chosen individuals.

3. Prove the direct Riemann integrability of the $e^{-\alpha t}\xi_2(t)$ in Lemma 4.9 in the proof of Theorem 4.5 under some other sufficient conditions on the offspring and lifetime distributions.

4. The age chart for a continuous-time single-type Markov Bellman-Harris branching process, i.e., when the lifetime distribution G is exponentially distributed with the parameter λ.

5. The limit behavior of the generation number of the last common ancestor of the randomly chosen individuals are still open.

6. All the analogs on the coalescence problems in the explosive case are unknown.

7.3 Problems in Continuous-time Multi-type Bellman-Harris Branching Processes

1. To find the limits distribution of the generation number of the last common ancestor of the randomly chosen individuals, we assume that the lifetime distributions for individuals of different types are the same. Can we achieve without imposing this hypotheses?

2. What happens to the limits distribution of the generation number of the last common ancestor in a continuous-time multi-type Markov Bellman-Harris branching process?

3. The coalescence problems in the critical, subcritical and explosive cases for a continuous-time multi-type Bellman-Harris branching process remain open.

4. The results of the limit behavior of the generation number of any random chosen individual from those alive at time t in the critical, subcritical and explosive cases are still unknown.

5. Can we drop the hypotheses of $E(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$ for the proofs for a continuous-time multi-type Bellman-Harris branching process?

BIBLIOGRAPHY

[1] Athreya, K. B. (1971). A note on a functional equation arising in Galton-Watson branching processes. *Journal of Applied Probability*, *8*, 589–598.

[2] Athreya, K. B. and Kaplan, N. (1976). Convergence of the age distribution in the one-dimensional supercritical age-dependent branching process. *The Annals of Probability*, *40*(1), 38–50.

[3] Athreya, K. B. and Kaplan, N. (1978). The additive property and its applications in branching processes. *Advances in Probability*, Vol. 5, ed. A. Joffe and P. E. Ney, Dekker, New York.

[4] Athreya, K. B. and Schun, H.-J. (2003). On the supercritical Bellman-Harris process with finite mean. *The Indian Journal of Statistics*, *65*(2), 229–248.

[5] Athreya, K. B. and Ney, P. (2004). *Branching Processes*. New York: Dover.

[6] Athreya, K. B. and Lahiri, S. N. (2006). *Measure Theory and Probability Theory*. New York: Springer.

[7] Athreya, K. B., Athreya, S. R. and Iyer, S. K. (2010). Critical age dependent branching Markov processes and their scaling limits. *Proceedings of the Indian Academy of Sciences, Mathematical Sciences*, *120*(3), 1–25.

[8] Athreya, K. B. (2010). Ancestor problem for branching trees. *Ramanujan Mathematics Newsletter*, *19*(1), 1–10.

[9] Athreya, K. B. (2010). Branching random walks. *The Legacy of Alladi Ramakrishnan in the Mathematical Sciences*, (3), 337–349.

[10] Athreya, K. B. (2010). Coalescence in recent past in rapidly growing populations. *(Submitted)*.

[11] Athreya, K. B., Athreya, S. R. and Iyer, S. K. (2011). Supercritical age dependent branching Markov processes and their scaling limits. *(To appear in Bernoulli)*.

[12] Athreya, K. B. (2011). Coalescence in critical and subcritical Galton-Watson branching processes. *(To appear in Journal of Applied Probability)*.

[13] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press.

[14] Cohn, H. (1982). Norming constants for the finite mean supercritical Bellman-Harris process. *Z. Wahrscheinlichkeitsth*.

[16] Davies, P. L. (1978). The simple branching process: A note on convergence when the mean is infinite. *Journal of Applied Probability*, *15*, 466–480.

[16] Davies, P. L. (1978). The simple branching process: A note on convergence when the mean is infinite. *Journal of Applied Probability*, *15*, 466–480.

[17] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications, Vol. II*. New York: Wiley.

[18] Grey, D. R. (1978). Almost sure convergence in Markov branching processes with infinite mean. *Journal of Applied Probability*, *14*, 702–716.

[19] Grey, D. R. (1979). On regular branching processes with infinite mean. *Stochastic Processes and their Applications*, *8*, 257–267.

[20] Harris, T. E. (1963). *The Theory of Branching Processes*. Berlin: Springer-Verlag.

[21] Hoppe, F. M. (1976). Supercritical multitype branching processes. *The Annals of Probability*, *4*(3), 393–401.

[22] Jagers, P. (1975). *Branching Processes with Biological Applications*. London: Wiley.

[23] Jagers, P. and Nerman, O. (1984). The growth and composition of branching populations. *Advances in Applied Probability*, *16*(2), 221–259.

[24] Jagers, P. (1992). Stabilities and instabilities in population dynamics. *Journal of Applied Probability*, *29*(4), 770–780.

[25] Kallenberg, O. (1986). *Random Measures, 4th Ed.*. Berlin: Academic Press.

[26] Karlin, S. and McGregor, J. (1966). Spectral theory of branching processes. I. The case of discrete spectrum. *ZW*, *5*, 6–33.

[27] Lepage, R., Woodroffe, M. and Zinn, J. (1981). ***. *The Annals of Probability*, *9*(4), 624–632.

[28] Mode, C. J. (1971). *Multitypr Branching Processes*. New York: Elsevier.

[29] O'Brien, G. L. (1980). A limit theorem for sample maxima and heavy branches in Galton-Watson trees. *Journal of Applied Probability*, *17*, 539–545.

[30] Rama Murthy, K. (1980). Convergence of age and type distributions in multitype critical Bellman-Harris processes. *Journal of Applied Probability*, *17*, 948–955.

[31] Samuels, M. L. (1971). Distribution of the branching-process population among generations. *Journal of Applied Probability*, *8*, 655–667.

[32] Schuh, H.-J. and Barbour, A. D. (1977). On the asymptotic behavior of branching processes with infinite mean. *Advances in Applied Probability*, *9*, 681–723.

[33] Schuh, H.-J. (1982). Seneta constants for the supercritical Bellman-Harris process. *Advances in Applied Probability*, *14*, 732–751.

[34] Seneta, E. and Heyde, C. C. (1977). *I. J. Bienaymé: Statistical theory anticipated*. New York: Springer-Verlag.

[35] Sevastyanov, B. A. (1971). *Branching processes*. Moscow: Nauka. (Russ.)

[36] Zubkov, A. M. (1975). Limit distribution of the distance to the nearest common ancestor. *Theory of Probability and Its Applications*, *20*, 602–612.