TAIWANESE JOURNAL OF MATHEMATICS
Vol. 22, No. 1, pp. 245–261, February 2018
DOI: 10.11650/tjm/8123

Coalescence on Supercritical Bellman-Harris Branching Processes

Krishna B. Athreya and Jyy-I Hong*

Abstract. 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. Let $k \geq 2$ be a positive integer. If $Z(t) \geq k$, we pick k individuals from those who are alive at time t by simple random sampling without replacement and trace their lines of descent backward in time until they meet for the first time. Let $D_k(t)$ be the coalescence time (the death time of the most recent common ancestor) and let $X_k(t)$ be the generation number of the most recent common ancestor of these k random chosen individuals. In this paper, we study the distributions of $D_k(t)$ and $X_k(t)$ and their limit distributions as $t \to \infty$.

1. Introduction

When a population evolves in time, its size can be viewed as a branching process. For such an evolution, there are two directions to investigate its changes. One is to look at its future and the other one is to study its history. In this paper, we consider a continuous time process and the past of the population is of our interest.

1.1. Bellman-Harris branching processes

We consider a population starting with an individual and in which each individual lives for a random amount of time, say L, with distribution function G and, upon death, produces a random number ξ of children according to the probability distribution $\{p_j\}_{j\geq 0}$. We assume that the life time and the reproduction of each individual are independent of its lifetime and of other individuals (see Athreya and Ney [4]).

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 $\{p_j\}_{j\geq 0}$.

For a Bellman-Harris process, if Y_n is the number of individuals in the nth generation, then $\{Y_n\}_{n\geq 0}$ is called its embedded Galton-Watson branching process.

Received November 8, 2016; Accepted May 23, 2017.
Communicated by Nikolaos Zygouras.
2010 *Mathematics Subject Classification*. 60J80.
Key words and phrases. branching processes, coalescence, supercritical, Bellman, Harris, age-dependent, line of descent.
*Corresponding author.

Let

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

and the Bellman-Harris branching process is called supercritical, critical, subcritical or explosive process according as $1 < m < \infty$, $m = 1$, $m < 1$, or $m = \infty$, respectively.

Moreover, a probability space $(\Omega, \mathcal{F}, P)$ is constructed by T. E. Harris [9] (see chapter VI) for the branching process $\{Z(t) : t \geq 0\}$ to live. Each point $\mathcal{T}$ of the sample space Ω is called a family tree of the process $\{Z(t) : t \geq 0\}$ and it can be thought of as a "tree" whose branch lengths represent individual life lengths; the number of branches at a given vertex representing the number of offspring of a given individual. Each family tree $\mathcal{T}$ provides the complete family information including the time of birth, life length, ancestors and descendants of each individual in the family.

1.2. The coalescence problem

A branching process is often used to describe the evolution of a population in time. One way to investigate the population is to look at its future. But, when a population grows so old, it is always interesting to know what happened to it in the past. The coalescence problem provides a way to understand the structure of the population and the ancestry of the individuals in it such as the "closeness" of any number of randomly chosen individuals. For a supercritical Bellman-Harris age-dependent processes, we now address the problem of coalescence.

Pick k individuals at random from the population alive at time t by simple random sampling without replacement. Trace their lines of descent backward in time till they meet. Let $X_k(t)$ be the generation number of the coalescence time of these k individuals alive. We call the common ancestor of these chosen individuals in the $X_k(t)$th generation their most recent common ancestor. In this paper, the limit behavior of the distributions of $X_k(t)$, for any integer $k \geq 2$, is studied for the supercritical case. Also, we investigate the coalescence time (the death time of the most recent common ancestor) $D_k(t)$ and its limit distribution as $t \to \infty$.

The coalescence problem has been studied for different branching processes. Athreya [1,2] has the results on the single-type Galton-Watson branching processes. Hong [10–12] extends them to multi-type Galton-Watson branching processes and also has studied the subcritical case for Bellman-Harris processes. In addition, many related works, such as the degree of relationship and other family structures, have been done for different cases in Markov branching processes, see Bühler [5,6], Durrett [7], Le [16] and O'Connell [17]. Fleischmann and Siegmund-Schultze [8] also study the coalescence in a reduced critical Galton-Watson tree. Lambert [14] obtains the limit distribution of the coalescence time in subcritical case with a more general settings (both discrete and continuous time and state space). Moreover, Lambert and Popovic [15] and Popovic [18] define a coalescent point process with the coalescence time of two successive individuals alive at the same time as the first point mass in it and study the limit of this coalescent point process.

In this paper, some classical limit theorems for Bellman-Harris processes are stated in Section 2 and will be used for the proofs. The main results on the generation number $X_k(t)$ of the most recent common ancestor are in Section 3.1 while the results on the death time $D_k(t)$ of the most recent common ancestor are provided in Section 3.2. In Section 4, Lemmas are listed. The proofs for the main results can be seen in Section 5 and, at the end of the paper, the lemmas are proved in Section 6.

2. Preliminaries

In this section, we introduce the notations which will be used in theorems, lemmas or proofs. Also, some well-known results regarding the population growth and age distribution for a Bellman-Harris process are provided here for later.

First, we introduce a parameter α which will describe the growth rate of the population. The Malthusian parameter for m and G is the root α in $\mathbb{R}$ (provided it exists) such that

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

Let f be the generating function of the offspring distribution, i.e.,

$$f(s) = \sum_{j=0}^{\infty} p_j s^j, \quad 0 \le s \le 1$$

and

$$F(s,t) \equiv \sum_{j=0}^{\infty} P(Z(t) = j \mid Z(0) = 1) s^j, \quad 0 \le s \le 1.$$

Then $F(s,t)$ is the unique bounded solution of the following integral equation

$$F(s,t) = s[1 - G(t)] + \int_0^t f(F(s, t-x))\, dG(x), \quad 0 \le s \le 1.$$

Thus, F is fully determined by the pair (f, G).

Also, let q be the probability of the extinction, i.e.,

$$q = P(Z(t) = 0 \text{ for some } t \mid Z(0) = 1).$$

The following theorem gives the growth rate of the population size $Z(t)$ and the properties related to its limit distribution. See Theorem 2 on page 172 in Athreya and Ney [4] for the details.

Theorem 2.1. *Let* $1 < m < \infty$. *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* $t)$.

Another important and useful aspect of Bellman-Harris branching processes is the limit behavior of the age distribution. For any family history (in what follows by a family history we mean the full information of the tree initiated by the initial ancestor including information on the life times and the number of offspring of all individuals in the family tree) and $0 \le x < \infty$, let $A(x, t)$ be the proportion of the individuals, whose ages are less than or equal to x, at time t. Then Athreya and Kaplan have the following theorem.

Theorem 2.2. (Athreya and Kaplan [3]) *Let* $1 < m = \sum_{j=1}^{\infty} jp_j < \infty$ *and* $p_0 = 0$. *Then*

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

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

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

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

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

3. Main results

Note that the assumption $p_0 = 0$ is not necessary for all four main theorems in this chapter. By conditioning on the event of non-extinction, the proofs can be changed appropriately.

3.1. Results on the generation number

Let $L_{n,i,k}$ be the lifetime of the ancestor in the kth generation of the ith individual in the nth generation. Then $\{L_{n,i,k} : n \ge 0, i \ge 1, k = 0, 1, 2, \ldots, n-1\}$ are i.i.d. copies with distribution G.

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

Let $W_{n,i}$ be the limit of $e^{-\alpha t} Z_{n,i}(t)$ as $t \to \infty$ (see Theorem 2.1) where $Z_{n,i}(t)$ is the branching process initiated by the ith individual of the nth generation.

Theorem 3.1. *Let $1 < m < \infty$, $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$, $p_0 = 0$ and the life time distribution G be non-lattice with $G(0+) = 0$. Then, for any integer $k \geq 2$,*

(a) *(Quenched version) For almost all family history trees $\mathcal{T}$ and $r = 0, 1, 2, \ldots$,*

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

as $t \to \infty$, where $\{Y_n\}_{n \geq 0}$ is the embedded Galton-Watson branching process. (For properties on $\{W_{r,i}\}$, see Theorem 3.4 below.)

(b) *(Annealed version) For each k, $X_k(t) \xrightarrow{d} \widetilde{X}_k$ as $t \to \infty$ with*

$$P(\widetilde{X}_k < r) = 1 - E\left(\frac{\sum_{i=1}^{Y_r}(e^{-\alpha S_{r,i}} W_{r,i})^k}{\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}\right)^k}\right) \equiv \phi_k(r), \quad r = 0, 1, 2, \ldots.$$

(c) *Further, $\lim_{r \uparrow \infty} \phi_k(r, \mathcal{T}) = 1$ for almost all trees $\mathcal{T}$ and $\lim_{r \uparrow \infty} \phi_k(r) = 1$.*

When $k \to \infty$, we show that the random variable $\widetilde{X}_k$ converges in distribution to a proper random variable U which is the last generation consisting of only one individual. With the assumption $p_0 = 0$, no sub-family (or sub-tree) will die out, so the limit (in distribution) as $k \to \infty$ of $\widetilde{X}_k$ can be thought as the generation number of the youngest individual who is an ancestor of the total future population.

Theorem 3.2. *Let $1 < m < \infty$ and $U = \max\{n \geq 1 : Y_n = 1\}$. Under the same hypotheses of Theorem 3.1, then $\widetilde{X}_k \xrightarrow{d} U$ as $k \to \infty$.*

Remark 3.3. The distribution of the random variable U can be found by simple calculation. Especially, if $p_0 = 0$ and $p_1 > 0$, then, by the independent reproduction law of the branching process, we have that

$$P(U = k) = P(Y_1 = 1, Y_2 = 1, \ldots, Y_k = 1, Y_k > 1) = p_1^k(1 - p_1)$$

for all $k = 1, 2, \ldots$. That is, U is geometrically distributed with parameter p_1.

3.2. Results on the death time

Let $L_{s,i}$ be the total lifetime of the ith individual alive at time s. Then $\{L_{s,i}\}_{i \geq 1}$ 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 3.4. *Let $1 < m < \infty$, $\sum_{j=1}^{\infty}(j \log j)p_j < \infty$, $p_0 = 0$ and the life time distribution G be non-lattice with $G(0+) = 0$. Then, for any integer $k \geq 2$,*

(a) *(Quenched version) For almost all family trees $\mathcal{T}$ and any $s \geq 0$, there exists positive real-valued random variables $\widetilde{W}_{s,i}$, $i = 1, 2, \ldots, Z(s)$, such that*

$$P(D_k(t) \leq s \mid \mathcal{T}) \to H_k(s, \mathcal{T}) = 1 - \frac{\sum_{i=1}^{Z(s)} (e^{-\alpha R_{s,i}} \widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}\right)^k}$$

as $t \to \infty$. The random variables $\widetilde{W}_{r,i}$, $i = 1, 2, \ldots, Z(s)$, are all functions of the tree $\mathcal{T}$. Further, conditioned on $Z(s)$, they are the i.i.d. copies of the sum $\sum_{j=1}^{\xi} W_j$ where ξ is a random variable with the offspring distribution $\{p_j\}_{j\geq 0}$ and $\{W_j\}_{j\geq 0}$ are i.i.d. copies of W as defined in Theorem 2.1(b) and ξ and $\{W_j\}_{j\geq 1}$ are independent.

(b) *(Annealed version) For each $k \geq 2$, there exists a nonnegative real-valued random variable $\widetilde{D}_k$ such that $D_k(t) \xrightarrow{d} \widetilde{D}_k$ as $t \to \infty$ and*

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

for any $s \geq 0$.

The next theorem shows that the limit law of $\widetilde{D}_k$ converges to the first time when the process splits into more than one as $k \to \infty$.

Theorem 3.5. *Let $1 < m < \infty$ and $U = \max\{n \geq 1 : Y_n = 1\}$ where $\{Y_n\}_{n\geq 0}$ is the embedded Galton-Watson branching process. Under the hypotheses of Theorem 3.4, there exist a random variable $\widetilde{D}$ such that $\widetilde{D}_k \xrightarrow{d} \widetilde{D}$ as $k \to \infty$ and, for any $s \geq 0$,*

$$P(\widetilde{D} \leq s) = P(L_0 + L_1 + \cdots + L_U \leq s)$$

where $\{L_i\}_{i\geq 0}$ are i.i.d. random variables with distribution G and independent of U.

Remark 3.6. The limit (in distribution) as $k \to \infty$ of $\widetilde{D}_k$ describes the death time of the youngest ancestor of the total future population and it is the sum of U i.i.d. copies of lifetime law.

4. Lemmas

The following lemmas will be needed for proving the main results in this paper. The first lemma follows from a well-known fact that, as $n \to \infty$,

$$\frac{1}{n} \max_{1\leq i\leq n} X_i \xrightarrow{P} 0$$

where $\{X_i\}_{i\geq 1}$ are i.i.d. copies with finite mean.

Lemma 4.1. *For any $s \geq 0$, let $\{W_{s,i,j} : j \geq 1, i \geq 1\}$ be i.i.d. copies of W defined in Theorem* 2.1(b) *and be independent of $\{\xi_{s,i}\}_{i\geq 1}$ where $\xi_{s,i}$ is the number of offspring of the ith individual alive at time s. Let $\widetilde{W}_{s,i} = \sum_{j=1}^{\xi_{s,i}} W_{s,i,j}$. Then, under the hypotheses of Theorem* 3.4, *as $s \to \infty$,*

$$\frac{1}{Z(s)} \max_{1\leq i\leq Z(s)} \widetilde{W}_{s,i} \to 0 \quad \textit{in probability.}$$

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* 3.4, *as $s \to \infty$,*

$$\frac{Z(s,k)}{Z(s)} \to B(k) \quad \textit{in probability}$$

where

$$B(k) = \frac{\int_{[0,\infty)} e^{-\alpha x}[G(x+k) - G(x)]\, dx}{\int_{[0,\infty)} e^{-\alpha x}[1 - G(x)]\, dx}.$$

Lemma 4.3. *Fix $k > 0$, let $\widetilde{W}_{s,i}$ and $Z(s,k)$ be the random variables defined in Lemmas* 4.1 *and* 4.2, *respectively. Then, under the hypotheses of Theorem* 3.4, *there exists $\theta > 0$ such that, as $s \to \infty$,*

$$P\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,k} I_{(R_{s,i}\leq k)} \geq \theta\right) \to 1.$$

5. Proofs of main results

5.1. Proof of Theorem 3.1

Let $\{Z_{r,i}(t) : t > 0\}$ be the branching process initiated with the ith individual in the rth generation when it is of age 0, then $\{Z_{r,i}(t - S_{r,i}) : t \geq S_{r,i}\}$ denotes the size of the process initiated with the ith individual of the rth generation with birth time $S_{r,i}$.

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

$$\begin{aligned} P(X_k(t) \geq r \mid \mathcal{T}) &= \frac{\sum_{i=1}^{Y_r} \prod_{j=0}^{k-1} (Z_{r,i}(t - S_{r,i}) - j)}{\prod_{j=0}^{k-1} (Z(t) - j)} \\ &= \frac{\sum_{i=1}^{Y_r} \prod_{j=0}^{k-1} (Z_{r,i}(t - S_{r,i}) - j)}{\prod_{j=0}^{k-1} \left(\sum_{i=1}^{Y_r} Z_{r,i}(t - S_{r,i}) - j\right)} \\ &= \frac{\sum_{i=1}^{Y_r} \left[e^{-k\alpha S_{r,i}} \prod_{j=0}^{k-1} e^{-\alpha(t-S_{r,i})}(Z_{r,i}(t - S_{r,i}) - j)\right]}{\prod_{j=0}^{k-1} \left[\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} e^{-\alpha(t-S_{r,i})}(Z_{r,i}(t - S_{r,i}) - j)\right]} \end{aligned}$$

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

It is known from Theorem 2.1 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 $t \to \infty$,

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

as $t \to \infty$, where, conditioned on Y_r and averaged over all trees $\mathcal{T}$, $\{W_{r,i}\}_{i\geq 1}$ are the i.i.d. copies of W.

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

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

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

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

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

and then by the bounded convergence theorem, this will yield

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

First, note that, conditioned on Y_r and averaged over all trees $\mathcal{T}$, $\{W_{r,i}\}_{i\geq 1}$ are i.i.d. copies of W and are independent of $L_0, L_1, \ldots, L_{r-1}$ and $Y_0, Y_1, \ldots, Y_r$ and since $\left\{S_{r,i} \equiv \sum_{k=0}^{r-1} L_{r,i,k}\right\}_{i\geq 1}$ 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 $i \geq 1$, we have that

$$\begin{aligned}
&E\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}W_{r,i}\right) \\
&= E\left(\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}}E(W_{r,i} \mid L_0, L_1, \ldots, L_{r-1}, Y_0, Y_1, \ldots, Y_r)\right) \\
&= EY_r \cdot (Ee^{-\alpha L})^r = m^r \cdot (E(e^{-\alpha L}))^r = 1 < \infty
\end{aligned}$$

where $EW = 1$ and $mE(e^{-\alpha L}) = 1$ by the definition of the Mathusian parameter for m and G. Therefore, any $\eta > 0$, by Markov's inequality,

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

So, for any $\epsilon > 0$,

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

and hence it suffices to prove that

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

Let $\mathfrak{G}_r$ be the σ-algebra generated by all the information up to the rth generation in the embedded Galton-Watson tree. Moreover, let $\eta(y) = \sup_{x\geq y} xP(W > x)$. The fact that $EW < \infty$ implies $xP(W > x) \to 0$ as $x \to \infty$ and hence, for any $\epsilon > 0$ and $l > 0$, there exists $a > 0$ such that $yP(W > y) < l\epsilon$ for all $y \geq a$ and therefore $\eta(a) \leq l\epsilon$. Now, let $n > \frac{1}{\alpha} \ln \frac{a}{\epsilon}$, then we have that $\epsilon e^{\alpha n} > a$ and

$$\begin{aligned}
&P\left(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon\right)\\
&\leq P\left(\min_{1\leq i\leq Y_r} S_{r,i} \leq n\right) + E\left(P\left(\max_{1\leq i<Y_r} e^{-\alpha S_{r,i}} W_{r,i} > \epsilon, \min_{1\leq i\leq Y_r} S_{r,i} > n \,\middle|\, \mathfrak{G}_r\right)\right)\\
&\leq P\left(\min_{1\leq i\leq Y_r} S_{r,i} \leq n\right) + E\left(\sum_{i=1}^{Y_r} P\left(W_{r,i} > \epsilon e^{\alpha S_{r,i}}, \min_{1\leq i\leq Y_r} S_{r,i} > n \,\middle|\, \mathfrak{G}_r\right)\right)\\
&\leq P\left(\min_{1\leq i\leq Y_r} S_{r,i} \leq n\right) + \frac{1}{\epsilon} E\left(\sum_{i=1}^{Y_r} \eta(a) e^{-\alpha S_{r,i}}\right)\\
&< P\left(\min_{1\leq i\leq Y_r} S_{r,i} \leq n\right) + l.
\end{aligned}$$

Moreover,

$$\begin{aligned}P\left(\min_{1\le i\le Y_r} S_{r,i}\le n\right) &= \sum_{x=0}^{\infty} P\left(\min_{1\le i\le Y_r} S_{r,i}\le n \;\middle|\; Y_r=x\right)P(Y_r=x)\\ &\le \sum_{x=0}^{\infty} xP(S_{r,1}\le n)P(Y_r=x) = P(S_{r,1}\le n)EY_r\\ &= P\left(e^{-\theta S_{r,1}}\ge e^{-\theta n}\right)EY_r\\ &\le \frac{E(e^{-\theta S_{r,1}})}{e^{-\theta n}}m^r = e^{\theta n}(Ee^{-\theta L})^r m^r = e^{\theta n}(m\varphi_L(\theta))^r\\ &\to 0 \qquad \text{as } r\to\infty\end{aligned}$$

where $\varphi_L(\theta)\equiv E(e^{-\theta L})$ is the Laplace transform of the lifetime L and $\theta>\alpha$ such that $m\varphi_L(\theta)<1$.

Therefore, for any $\epsilon>0$,

$$\overline{\lim_{r\to\infty}} P\left(\max_{1\le i<Y_r} e^{-\alpha S_{r,i}}W_{r,i}>\epsilon\right)<l$$

for any $l>0$. Hence, for any $\epsilon>0$,

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

The proof of Theorem 3.1 is complete.

5.2. Proof of Theorem 3.2

Let $U=\max\{n\ge 1: Y_n=1\}$. Then

$$\begin{aligned}P(\widetilde{X}_k\ge r) &= E\left(\frac{\sum_{i=1}^{Y_r}(e^{-\alpha S_{r,i}}W_{r,i})^k}{\left(\sum_{i=1}^{Y_r}e^{-\alpha S_{r,i}}W_{r,i}\right)^k}I_{(r\le U)}\right)+E\left(\frac{\sum_{i=1}^{Y_r}(e^{-\alpha S_{r,i}}W_{r,i})^k}{\left(\sum_{i=1}^{Y_r}e^{-\alpha S_{r,i}}W_{r,i}\right)^k}I_{(r>U+1)}\right)\\ &= P(r\le U)+E\left(E\left(\frac{\sum_{i=1}^{Y_r}(e^{-\alpha S_{r,i}}W_{r,i})^k}{\left(\sum_{i=1}^{Y_r}e^{-\alpha S_{r,i}}W_{r,i}\right)^k}I_{(r>U+1)}\;\middle|\; Y_r\right)\right)\\ &= P(r\le U)+E\left(Y_rE\left(\left(\frac{e^{-\alpha S_{r,1}}W_{r,1}}{\sum_{i=1}^{Y_r}e^{-\alpha S_{r,i}}W_{r,i}}\right)^k I_{(r>U+1)}\;\middle|\; Y_r\right)\right).\end{aligned}$$

Since $p_0=0$ and $\sum_{j=1}^{\infty}(j\log j)p_j<\infty$, $P(0<W_{r,j}<\infty)=1$ for all $j\ge 0$ and all $r=0,1,2,\ldots$. So, on the set $\{r\ge U+1\}$,

$$0<\frac{e^{-\alpha S_{r,1}}W_{r,1}}{\sum_{i=1}^{Y_r}e^{-\alpha S_{r,i}}W_{r,i}}<1 \quad \text{w.p.1}$$

and hence, for any $r = 0, 1, 2, \ldots$, as $k \to \infty$,

$$\left(\frac{e^{-\alpha S_{r,1}} W_{r,1}}{\sum_{i=1}^{Y_r} e^{-\alpha S_{r,i}} W_{r,i}} \right)^k \to 0 \quad \text{w.p.1.}$$

Therefore, by the bounded convergence theorem, we have that

$$P(\widetilde{X}_k \geq r) \to P(U \geq r) \quad \text{as } k \to \infty$$

and the proof is complete.

5.3. Proof of Theorem 3.4

Let $\{\widetilde{Z}_{t-s-R_{s,i},j} : t \geq s + R_{s,i}\}$ be the branching process initiated by the jth offspring of the ith individual alive at time s.

For almost all trees $\mathcal{T}$ and $s \geq 0$,

$$\begin{aligned}
&P(D_k(t) \leq s \mid \mathcal{T}) \\
&= 1 - P(D_k(t) > s \mid \mathcal{T}) \\
&= 1 - \frac{\sum_{i=1}^{Z(s)} \left(\sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} \right) \left(\sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} - 1 \right) \cdots \left(\sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} - k + 1 \right)}{\left(\sum_{i=1}^{Z(s)} \sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} \right) \left(\sum_{i=1}^{Z(s)} \sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} - 1 \right) \cdots \left(\sum_{i=1}^{Z(s)} \sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} - k + 1 \right)} \\
&= 1 - \frac{\sum_{i=1}^{Z(s)} \prod_{l=1}^{k} \left(e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} e^{-\alpha(t-s-R_{s,i})} - (l-1) e^{-\alpha(t-s-R_{s,i})} \right)}{\prod_{l=1}^{k} \left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} \widetilde{Z}_{t-s-R_{s,i},j} e^{-\alpha(t-s-R_{s,i})} - (l-1) e^{-\alpha(t-s-R_{s,i})} \right)}
\end{aligned}$$

and then, by Theorem 2.1,

$$\begin{aligned}
P(D_k(t) \leq s \mid \mathcal{T}) &\to 1 - \frac{\sum_{i=1}^{Z(s)} \left(e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} W_{s,i,j} \right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \sum_{j=1}^{\xi_{s,i}} W_{s,i,j} \right)^k} \quad \text{as } t \to \infty \\
&= 1 - \frac{\sum_{i=1}^{Z(s)} \left(e^{-\alpha R_{s,i}} \widetilde{W}_{s,i} \right)^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i} \right)^k} \equiv H_k(s, \mathcal{T})
\end{aligned}$$

where $\widetilde{W}_{s,i} \equiv \sum_{i=1}^{\xi_{s,i}} W_{s,i,j}$ for $i \geq 1$ and $s \geq 0$ and, conditioned on $Z(s)$ and averaged over all trees $\mathcal{T}$, $\{W_{s,i,j}\}_{j \geq 1}$ are i.i.d. copies of W in Theorem 2.1.

So, by the bounded convergence theorem, as $t \to \infty$,

$$P(D_k(t) \leq s) = E\left(P(D_k(t) \leq s \mid \mathcal{T}) \right) \to 1 - E\left(\frac{\sum_{i=1}^{Z(s)} (e^{-\alpha R_{s,i}} \widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i} \right)^k} \right) \equiv H_k(s).$$

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

$$E\left(\frac{\sum_{i=1}^{Z(s)} (e^{-\alpha R_{s,i}} \widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i} \right)^k} \right) \to 0 \quad \text{as } s \to \infty.$$

It suffices to prove that, as $s \to \infty$,

$$\frac{\sum_{i=1}^{Z(s)} (e^{-\alpha R_{s,i}} \widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}\right)^k} \to 0 \quad \text{in probability.}$$

Moreover, since

$$\left(\frac{\max_{1\le i\le Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}\right)^k \le \frac{\sum_{i=1}^{Z(s)} (e^{-\alpha R_{s,i}} \widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}\right)^k} \le \left(\frac{\max_{1\le i\le Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}\right)^{k-1},$$

it is enough to show that, as $s \to \infty$,

$$\frac{\max_{1\le i\le Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}} \to 0 \quad \text{in probability.}$$

First of all, by Lemmas 4.2 and 4.3, we know that, there exists $\theta > 0$ such that, for any $\delta > 0$, there exists an $M > 0$ such that for any $s > M$,

$$P\left(\frac{Z(s,k)}{Z(s)} < \frac{1}{2}B(k)\right) < \frac{\delta}{2}$$

and

$$P\left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,k} I_{(R_{s,i}\le k)} < \theta\right) < \frac{\delta}{2}.$$

Then, we let $A = \left\{\frac{Z(s,k)}{Z(s)} \ge \frac{1}{2}B(k)\right\}$ and $B = \left\{\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,k} I_{(R_{s,i}\le k)} \ge \theta\right\}$ be two events. So, for any $\epsilon > 0$, $s > M$ and $k > 0$,

$$\begin{aligned}
&P\left(\frac{\max_{1\le i\le Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}} > \epsilon\right) \\
&\le P\left(\frac{e^{\alpha k} \frac{1}{Z(s)} \max_{1\le i\le Z(s)} \widetilde{W}_{s,i}}{\frac{Z(s,k)}{Z(s)} \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s)} \widetilde{W}_{s,i} I_{(R_{s,i}\le k)}} > \epsilon\right) \\
&= P\left(\frac{1}{Z(s)} \max_{1\le i\le Z(s)} \widetilde{W}_{s,i} > \epsilon e^{-\alpha k} \frac{Z(s,k)}{Z(s)} \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s)} \widetilde{W}_{s,i} I_{(R_{s,i}\le k)}\right) \\
&\le P\left(\frac{1}{Z(s)} \max_{1\le i\le Z(s)} \widetilde{W}_{s,i} > \epsilon e^{-\alpha k} \frac{1}{2} B(k)\theta : A\cap B\right) + P(A^C) + P(B^C) \\
&\le P\left(\frac{1}{Z(s)} \max_{1\le i\le Z(s)} \widetilde{W}_{s,i} > \frac{1}{2}\epsilon\theta e^{-\alpha k} B(k)\right) + \delta.
\end{aligned}$$

Thus, by Lemma 4.1 and since δ is arbitrary, we have that, for any $\epsilon > 0$,

$$\lim_{s\to\infty} P\left(\frac{\max_{1\le i\le Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}} \widetilde{W}_{s,i}} > \epsilon\right) = 0$$

and the proof of Theorem 3.4 is complete.

5.4. Proof of Theorem 3.5

From Theorem 3.4, for any $s > 0$, we have that

$$\begin{aligned}
P(\widetilde{D}_k(t) > s) &= E\left(\frac{\sum_{i=1}^{Z(s)}(e^{-\alpha R_{s,i}}\widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}\right)^k} I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U}>s)}\right) \\
&\quad + E\left(\frac{\sum_{i=1}^{Z(s)}(e^{-\alpha R_{s,i}}\widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}\right)^k} I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U}\le s)}\right) \\
&= P(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U} > s) \\
&\quad + E\left(E\left(\frac{\sum_{i=1}^{Y_r}(e^{-\alpha R_{s,i}}\widetilde{W}_{s,i})^k}{\left(\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}\right)^k} I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U}\le s)} \,\middle|\, Z(s)\right)\right) \\
&= P(L_0+L_1+\cdots+L_U > s) \\
&\quad + E\left(\sum_{i=1}^{Z(s)} E\left(\left(\frac{e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}}\right)^k I_{(L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U}\le s)} \,\middle|\, Z(s)\right)\right)
\end{aligned}$$

where $\{L_i\}_{i\ge 0}$ 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 $\widetilde{W}_{s,j} = \sum_{k=1}^{\xi_{s,j}} W_{s,j,k}$, we have that $P(0 < \widetilde{W}_{s,j} < \infty) = 1$ for all $j \ge 0$ and $s > 0$. So, on the set $\{L_{s,i,0}+L_{s,i,1}+\cdots+L_{s,i,U} \le s\}$,

$$0 < \frac{e^{-\alpha R_{s,j}}\widetilde{W}_{s,j}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}} < 1 \quad \text{w.p.1}$$

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

$$\left(\frac{e^{-\alpha R_{s,j}}\widetilde{W}_{s,j}}{\sum_{i=1}^{Z(s)} e^{-\alpha R_{s,i}}\widetilde{W}_{s,i}}\right)^k \to 0 \quad \text{w.p.1.}$$

Therefore, by the bounded convergence theorem again, the proof is complete.

6. Proofs of lemmas

6.1. Proof of Lemma 4.2

For any fixed $k > 0$, consider the function g defined by

$$g(a) \equiv P(R_{s,i} \le k \mid a_{s,i} = a) = \frac{G(a+k) - G(a)}{1 - G(a)}.$$

Let $\mathfrak{F}_s$ be the σ-algebra generated by all the history of this branching process up to time s.

Then, for any $\epsilon > 0$,

$$
\begin{aligned}
& P\left(\left|\frac{Z(s,k)}{Z(s)} - B(k)\right| > \epsilon\right) \\
&= E\left(P\left(\left|\frac{Z(s,k)}{Z(s)} - B(k)\right| > \epsilon \,\Big|\, \mathfrak{F}_s\right)\right) \\
&= E\left(P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} I_{(R_{s,i}\le k)} - B(k)\right| > \epsilon \,\Big|\, \mathfrak{F}_s\right)\right) \\
&= E\left(P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} (I_{(R_{s,i}\le k)} - g(a_{s,i})) + \frac{1}{Z(s)}\sum_{i=1}^{Z(s)} (g(a_{s,i}) - B(k))\right| > \epsilon \,\Big|\, \mathfrak{F}_s\right)\right) \qquad (6.1)\\
&\le E\left(P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} (I_{(R_{s,i}\le k)} - g(a_{s,i}))\right| > \frac{\epsilon}{2} \,\Big|\, \mathfrak{F}_s\right)\right) \\
&\quad + P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} (g(a_{s,i}) - B(k))\right| > \frac{\epsilon}{2}\right).
\end{aligned}
$$

First, by the weak law of large numbers for Bernoulli random variables and the bounded convergence theorem, as $s \to \infty$,

$$
E\left(P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} (I_{(R_{s,i}\le k)} - g(a_{s,i}))\right| > \frac{\epsilon}{2} \,\Big|\, \mathfrak{F}_s\right)\right) \to 0. \tag{6.2}
$$

So, it remains to prove that

$$
P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} (g(a_{s,i}) - B(k))\right| > \frac{\epsilon}{2}\right) \to 0 \quad \text{as } s \to \infty.
$$

Let $A(x,s) = \frac{1}{Z(s)}\sum_{i=1}^{Z(s)} I_{(a_{s,i}\le x)}$ as defined in Section 2 and since $g(x) = \frac{G(x+k)-G(x)}{1-G(x)}$ is a bounded function that is continuous except on a countable set and hence bounded a.e. with respect to $A(\cdot)$ defined in Theorem 2.2, by Theorem 2.2(c), we have, as $s \to \infty$

$$
\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} g(a_{s,i}) \equiv \int_{[0,\infty)} g(x)\, dA(x,s) \to \int_{[0,\infty)} g(x)\, dA(x) = B(k) \quad \text{w.p.1.}
$$

Therefore, for any $\epsilon > 0$,

$$
P\left(\left|\frac{1}{Z(s)}\sum_{i=1}^{Z(s)} g(a_{s,i}) - B(k)\right| > \frac{\epsilon}{2}\right) \to 0 \quad \text{as } s \to \infty. \tag{6.3}
$$

From (6.1), (6.2) and (6.3), 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.

6.2. Proof of Lemma 4.3

Let $n_{s,1} = \min\{1 \le j \le Z(s) : R_{s,j} \le k\}$ and $n_{s,i} = \min\{n_{s,i-1} < j \le Z(s) : R_{s,j} \le k\}$ for $i \ge 2$. Then

$$\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s)} \widetilde{W}_{s,i} I_{(R_{s,i} \le k)} = \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,n_{s,i}} I_{(R_{s,n_{s,i}} \le k)}$$
$$= \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,n_{s,i}}.$$

Note that $E\widetilde{W}_{s,1} > 0$ and hence there exists an $\eta > 0$ such that $P(\widetilde{W}_{s,1} \ge \eta) > 0$.

Let $\mathfrak{F}_s$ be the σ-algebra generated by all the information of this Bellman-Harris branching process up to time s. Then

$$P(\widetilde{W}_{s,n_{s,i}} \ge \eta)$$
$$= E\left(P(\widetilde{W}_{s,n_{s,i}} \ge \eta \mid \mathfrak{F}_s)\right) = E\left(\sum_{j=1}^{Z(s,k)} P(\widetilde{W}_{s,n_{s,i}} \ge \eta, n_{s,i} = j \mid \mathfrak{F}_s)\right)$$
$$= E\left(\sum_{j=1}^{Z(s,k)} P(\widetilde{W}_{s,j} \ge \eta, n_{s,i} = j \mid \mathfrak{F}_s)\right) = E\left(\sum_{j=1}^{Z(s,k)} P(\widetilde{W}_{s,j} \ge \eta \mid \mathfrak{F}_s) P(n_{s,i} = j \mid \mathfrak{F}_s)\right)$$
$$= E\left(\sum_{j=1}^{Z(s,k)} P(\widetilde{W}_{s,1} \ge \eta \mid \mathfrak{F}_s) P(n_{s,i} = j \mid \mathfrak{F}_s)\right) = P(\widetilde{W}_{s,1} \ge \eta).$$

Let

$$X_{s,i} = \begin{cases} 1 & \text{if } \widetilde{W}_{s,n_{s,i}} \ge \eta, \\ 0 & \text{if } \widetilde{W}_{s,n_{s,i}} < \eta \end{cases}$$

then

$$\begin{aligned}
\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,n_{s,i}} &\ge \frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \eta X_{s,i} \\
&= \eta \left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\widetilde{W}_{s,n_{s,i}} \ge \eta) \right) \right) \\
&\quad + \eta \left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} P(\widetilde{W}_{s,n_{s,i}} \ge \eta) \right) \\
&= \eta \left(\frac{1}{Z(s,k)} \sum_{i=1}^{Z(s,k)} \left(X_{s,i} - P(\widetilde{W}_{s,n_{s,i}} \ge \eta) \right) \right) \\
&\quad + \eta P(\widetilde{W}_{s,n_{s,i}} \ge \eta).
\end{aligned} \tag{6.4}$$

Then we know from the weak law of large numbers again that the first term converges to 0 in probability. Also, the second term in (6.4) does not depend on s, so if let $\theta = \frac{1}{2}\eta P(\widetilde{W}_{s,1} \geq \eta)$, then $\theta > 0$ and, as $s \to \infty$,

$$P\left(\frac{1}{Z(s,k)}\sum_{i=1}^{Z(s,k)} \widetilde{W}_{s,i} I_{(R_{s,i} \leq k)} \geq \theta\right) \to 1.$$

Hence, Lemma 4.3 is proved.

Acknowledgments

The authors would like to give many of their thanks to the referee for the suggestions which make this paper more clear and informative.

References

[1] K. B. Athreya, *Coalescence in critical and subcritical Galton-Watson branching processes*, J. Appl. Probab. **49** (2012), no. 3, 627–638.

[2] ———, *Coalescence in the recent past in rapidly growing populations*, Stochastic Process. Appl. **122** (2012), no. 11, 3757–3766.

[3] K. B. Athreya and N. Kaplan, *Convergence of the age distribution in the one-dimensional supercritical age-dependent branching process*, Ann. Probability **4** (1976), no. 1, 38–50.

[4] K. B. Athreya and P. E. Ney, *Branching Processes*, Dover, Mineola, NY, 2004.

[5] W. J. Bühler, *Generations and degree of relationship in supercritical Markov branching processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **18** (1971), 141–152.

[6] ———, *The distribution of generations and other aspects of the family structure of branching processes*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, 463–480. Univ. California Press, Berkeley, Calif., 1972.

[7] R. Durrett, *The genealogy of critical branching processes*, Stochastic Process. Appl. **8** (1978), no. 1, 101–116.

[8] K. Fleischmann and R. Siegmund-Schultze, *The structure of reduced critical Galton-Watson processes*, Math. Nachr. **79** (1977), 233–241.

[9] T. E. Harris, *The Theory of Branching Processes*, Die Grundlehren der Mathematischen Wissenschaften, Bd. 119 Springer-Verlag, Berlin, 1963.

[10] J.-I. Hong, *Coalescence in subcritical Bellman-Harris age-dependent branching processes*, J. Appl. Probab. **50** (2013), no. 2, 576–591.

[11] ______, *Coalescence on supercritical multi-type branching processes*, Sankhya A **77** (2015), no. 1, 65–78.

[12] ______, *Coalescence on critical and subcritical multitype branching processes*, J. Appl. Probab. **53** (2016), no. 3, 802–817.

[13] P. Jagers and O. Nerman, *The growth and composition of branching populations*, Adv. in Appl. Probab. **16** (1984), no. 2, 221–259.

[14] A. Lambert, *Coalescence times for the branching process*, Adv. in Appl. Probab. **35** (2003), no. 4, 1071–1089.

[15] A. Lambert and L. Popovic, *The coalescent point process of branching trees*, Ann. Appl. Probab. **23** (2013), no. 1, 99–144.

[16] V. Le, *Coalescence times for the Bienaymé-Galton-Watson process*, J. Appl. Probab. **51** (2014), no. 1, 209–218.

[17] N. O'Connell, *The genealogy of branching processes and the age of our most recent common ancestor*, Adv. in Appl. Probab. **27** (1995), no. 2, 418–442.

[18] L. Popovic, *Asymptotic genealogy of a critical branching process*, Ann. Appl. Probab. **14** (2004), no. 4, 2120–2148.

Krishna B. Athreya
Department of Mathematics and Department of Statistics, Iowa State University, 396 Carver Hall, 411 Morrill Road, Ames, IA 50011, U.S.A.
E-mail address: `kbathreya@gmail.com`

Jyy-I Hong
Department of Applied Mathematics, National Sun Yat-sen University, No. 70, Lienhai Rd., Kaohsiung 80424, Taiwan
E-mail address: `hongjyyi@gmail.com`