

Coalescence on Supercritical Multi-type Branching Processes

Jyy-I Hong

Iowa State University, Ames, USA

Abstract

Let $\mathbf{Z}_n = (Z_n^{(1)}, Z_n^{(2)}, \dots, Z_n^{(d)})$ be a d -type ($d < \infty$) Galton-Watson branching process. For a positive integer $k \geq 2$. Pick k individuals at random from the n th generation by simple random sampling without replacement. Trace their lines of descent backward in time till they meet. Let $X_{n,k}$ be the generation number of the coalescence time of these k individuals of the n th generation. We call the common ancestor of these chosen individuals in the $X_{n,k}$ th generation their last common ancestor. In this paper, the limit behaviors of the distributions of $X_{n,k}$, for any integer $k \geq 2$, is studied for the supercritical cases. Also, we investigate the limit distribution of joint distribution of the generation number and the type of the last common ancestor of these randomly chosen individuals and their types in the supercritical case.

AMS (2000) subject classification. Primary 60J80; Secondary 60G50.

Keywords and phrases. Branching processes, Coalescence, Supercritical, Multitype, Line of descent.

1 Introduction

1.1. The Coalescence Problem. We consider a branching process with finite number d of individual types (see Section 1.3 for a precise definition). Such processes arise in a variety of applications in biology and physics.

For these processes, we address the problem of *coalescence*.

Pick k individuals at random from the n th generation by *simple random sampling without replacement* (SRSWOR). Trace their lines of descent backward in time till they meet. Let $X_{n,k}$ be the generation number of that time. Call this *the coalescence time* of these k individuals of the n th generation. We call the common ancestor of these chosen individuals in the $X_{n,k}$ th generation their *last common ancestor*. In this paper, the limit behaviors of the distributions of $X_{n,k}$, for any integer $k \geq 2$, is studied for the supercritical case. Also, we investigate the limit distribution of the joint distribution of the generation number and the type of the last common ancestor of these randomly chosen individuals and their types. Finally, the Markov property

of the limit behavior of types along the line of descent backward in time of an individual randomly chosen by SRSWOR from the n th generation, as $n \rightarrow \infty$, is also discussed.

1.2. Notations. Throughout this paper, we adopt the following notations.

1. \mathbb{N}_0 is the set of all nonnegative integers.
2. $\mathbb{N}_0^d \equiv \{\mathbf{j} \equiv (j_1, j_2, \dots, j_d) : j_i \in \mathbb{N}_0, i = 1, 2, \dots, d\}$
3. $\mathbf{0} = (0, 0, \dots, 0)$ and $\mathbf{1} = (1, 1, \dots, 1)$ in \mathbb{N}_0^d
4. $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{N}_0^d$ with the 1 in the i th component.
5. Let $\mathbf{u} = (u_1, u_2, \dots, u_d)$ and $\mathbf{v} = (v_1, v_2, \dots, v_d)$ be d -vectors with $u_i, v_i \in \mathbb{R}$, $i = 1, 2, \dots, d$. Then $\mathbf{u} \leq \mathbf{v}$ means $u_i \leq v_i$ for $i = 1, 2, \dots, d$ 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 absolute value of the vector \mathbf{x} is

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

7. The sup norm of the vector \mathbf{x} is

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

8. For a vector \mathbf{x} and a \mathbf{y} in \mathbb{N}_0^d ,

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

9. For a matrix \mathbf{M} , the super norm is

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

1.3. Definition of Branching Processes. Let $\mathbf{Z}_n = (Z_n^{(1)}, Z_n^{(2)}, \dots, Z_n^{(d)})$ be the population vector in the n th generation, $n = 0, 1, 2, \dots$, where $Z_n^{(i)}$ is the number of individuals of type i in the n th generation. We assume that each individual of type i , $i = 1, 2, \dots, d$, lives a unit of time and, upon death, produces children of all types according to the offspring distribution

$\{p^{(i)}(\mathbf{j}) \equiv p^{(i)}(j_1, j_2, \dots, j_d)\}_{\mathbf{j} \in \mathbb{N}_0^d}$ and independently of other individuals, where $p^{(i)}(j_1, j_2, \dots, j_d)$ is the probability that a type i parent produces j_1 children of type 1, j_2 children of type 2, \dots , j_d children of type d .

Let

$$f^{(i)}(s_1, s_2, \dots, s_d) \equiv \sum_{j_1, j_2, \dots, j_d \geq 0} p^{(i)}(j_1, j_2, \dots, j_d) s_1^{j_1} s_2^{j_2} \dots s_d^{j_d} \quad (1.1)$$

where $0 \leq s_r \leq 1$, $r = 1, 2, \dots, d$, be the probability generating function of the numbers of various types produced by a type i individual.

Let

$$\mathbf{f} \equiv (f^{(1)}, f^{(2)}, \dots, f^{(d)}). \quad (1.2)$$

be the vector of generating functions.

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

$$P(\mathbf{i}, \mathbf{j}) = P(\mathbf{Z}_{n+1} = \mathbf{j} | \mathbf{Z}_n = \mathbf{i}) \quad \forall \mathbf{i}, \mathbf{j} \in \mathbb{N}_0^d \quad (1.3)$$

such that, for any \mathbf{i} , $\sum_{\mathbf{j} \in \mathbb{N}_0^d} P(\mathbf{i}, \mathbf{j}) \mathbf{s}^{\mathbf{j}} = \left(\mathbf{f}(\mathbf{s})\right)^{\mathbf{i}}$ (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

$$\mathbf{Z}_n^{(i)} = (Z_n^{(i)(1)}, Z_n^{(i)(2)}, \dots, Z_n^{(i)(d)})$$

where, for $j = 1, 2, \dots, d$, $Z_n^{(i)(j)}$ is the number of type j individuals in the n th generation for a process with $\mathbf{Z}_0 = \mathbf{e}_i$. The probability generating function of $\mathbf{Z}_n^{(i)}$ will be denoted by $\mathbf{f}_n^{(i)}(\mathbf{s})$.

Also, if we let $\xi_{n,r}^{(j)}$ be the vector of offsprings of the r th individual of type j in the n th generation then, for all r and n , $P(\xi_{n,r}^{(j)} = \cdot) = p^{(j)}(\cdot)$. Thus, the population in the $(n+1)$ th generation can be expressed as

$$\mathbf{Z}_{n+1} = \sum_{j=1}^d \sum_{r=1}^{Z_n^{(j)}} \xi_{n,r}^{(j)}. \quad (1.4)$$

This is a useful stochastic evolution relation. In particular, this generates the following result on the means $EZ_n^{(i)(j)}$.

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 $i, j = 1, 2, \dots, d$. Then,

$$\mathbf{M} \equiv \{m_{ij} : i, j = 1, 2, \dots, d\}. \quad (1.5)$$

is called the *mean matrix*.

From (1.4), it follows that $E(\mathbf{Z}_1 | \mathbf{Z}_0) = \mathbf{Z}_0 \mathbf{M}$ and hence by iteration $E(\mathbf{Z}_n | \mathbf{Z}_0) = \mathbf{Z}_0 \mathbf{M}^n$. Here, we denote the (i, j) th element of \mathbf{M}^n by $m_{ij}^{(n)}$.

We also impose the following assumptions on the process $\{\mathbf{Z}_n\}_{n \geq 0}$ for the rest of the paper:

1. The branching process $\{\mathbf{Z}_n\}_{n \geq 0}$ is *non-singular*, 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 *positive regular*. That is, there exists an n such that $m_{ij}^{(n)} > 0$ for all $1 \leq i, j \leq d$.
3. Each individual in this process produces at least one offspring with probability 1 upon death, that is, $P(\mathbf{Z}_1 = \mathbf{0} | \mathbf{Z}_0 = \mathbf{e}_i) = 0$ for all $i = 1, 2, \dots, d$. Thus, *the probability of extinction is zero*.

By the Perron-Frobenius theorem (see Karlin and Taylor, 1975 or Athreya and Ney, 2004), the matrix \mathbf{M} has a maximal eigenvalue ρ which is positive, simple (it has 1 as its algebraic and geometric multiplicities) and has associated strictly positive right and left eigenvectors \mathbf{u} and \mathbf{v} . Moreover, these can be normalized so that the inner products

$$\mathbf{u} \cdot \mathbf{v} = 1 \quad \text{and} \quad \mathbf{u} \cdot \mathbf{1} = 1. \quad (1.6)$$

Further, one can write

$$\mathbf{M}^n = \rho^n \mathbf{P} + \mathbf{R}^n \quad (1.7)$$

where \mathbf{P} is the matrix whose (i, j) th entry is $u_i v_j$ and $\mathbf{R} \equiv \{r_{ij} : i, j = 1, 2, \dots, d\}$ is a matrix such that $\mathbf{P}\mathbf{R} = \mathbf{R}\mathbf{P} = \mathbf{0}$ and $r_{ij}^{(n)} \leq c\rho_0^n$, for all $n \geq 1, i, j = 1, 2, \dots, d$, for some $c < \infty$ and $0 < \rho_0 < \rho$.

In a discrete-time multi-type Galton-Watson branching process, the maximal eigenvalue ρ of the mean matrix \mathbf{M} plays a crucial role. The process is called a *subcritical*, *critical*, *supercritical* or *explosive* branching process according as $\rho < 1$, $\rho = 1$, $1 < \rho < \infty$ or $\rho = \infty$, respectively. It is known (Athreya and Ney, 2004) that if $\rho \leq 1$ then the process dies out with probability one and if $\rho > 1$ then this probability is less than one.

2 Main Results

For a supercritical branching process, we assume that $E(Z_{1,j} | \mathbf{Z}_0 = \mathbf{e}_i) \equiv m_{ij} < \infty$ for all $1 \leq i, j \leq d$.

Let ρ be the maximal eigenvalue of $\mathbf{M} = \{m_{ij} : i, j = 1, 2, \dots, d\}$. Then we have the following results.

THEOREM 2.1. *Let $1 < \rho < \infty$, $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ and $E(\|Z_1\| \log \|Z_1\| | \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$. Then, for $k = 2, 3, \dots$,*

- (a) (*Quenched version*) *for almost all trees \mathcal{T} and $r = 1, 2, \dots$, there exists positive real-valued random variables $W_{r,i}^{(l)}, i = 1, 2, \dots, Z_r^{(l)}, l = 1, 2, \dots, d$ such that*

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

as $n \rightarrow \infty$. The random variables $\{W_{r,i} : i = 1, 2, \dots, Z_r^{(l)}, l = 1, 2, \dots, d\}$ are all functions of the tree \mathcal{T} . Further, conditioned on \mathbf{Z}_r and averaged over all trees \mathcal{T} , they are independent random variables.

- (b) (*Annealed version*) *there exists random variable \tilde{X}_k such that $X_{n,k} \xrightarrow{d} \tilde{X}_k$ as $n \rightarrow \infty$, where*

$$P(\tilde{X}_k < r) \equiv \phi_k(r) = 1 - E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)}\right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)}\right)^k} \right)$$

for any $r = 1, 2, \dots$. Further, $\lim_{r \rightarrow \infty} \phi_k(r) = 1$ so that \tilde{X}_k is a proper random variable.

REMARK 2.1. Athreya (2012) proved the coalescence in the single-type case. Many ideas of the proof of the above theorem came from his paper and this theorem can be viewed as an extension to the multi-type case.

REMARK 2.2. Theorem 2.1 (a) and (b) should be valid just with $1 < \rho < \infty$. That is, the assumption $E\|Z_1\| \log \|Z_1\| < \infty$ should not be needed.

This will need Hoppe's result in Hoppe (1976) 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 Schuh (2003). The multi-type extension of this needs to be investigated.

Theorem 2.2 is a result on the limit distribution of \tilde{X}_k as $k \rightarrow \infty$. Not surprisingly, it says that the coalescence time goes away back to one generation before the first time when this process began to split.

THEOREM 2.2. *Let $1 < \rho < \infty$ 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 $k \rightarrow \infty$.*

Next, we pick two individuals (i.e. consider $k = 2$) at random by SR-SWOR from the n th 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 types 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 $n \rightarrow \infty$ to a proper distribution. It is necessarily the annealed version.

THEOREM 2.3. *Let $1 < \rho < \infty$, $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ and $E(\|Z_1\| \log \|Z_1\| | \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$. Then*

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

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

The next result is an extension of the above theorem for $k \geq 2$.

THEOREM 2.4. *Let $1 < \rho < \infty$, $\mathbf{Z}_0 = \mathbf{e}_{i_0}$ and $E(\|Z_1\| \log \|Z_1\| | \mathbf{Z}_0 = \mathbf{e}_i) < \infty$ for all $1 \leq i \leq d$. Then, for any $2 \leq k < \infty$,*

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_{n,k} = r, \eta_n = j, \zeta_{n,1} = i_1, \zeta_{n,2} = i_2, \dots, \zeta_{n,k} = i_k) \\ \equiv \varphi_k(r, j, i_1, i_2, \dots, i_k) \end{aligned}$$

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

3 Proofs of Main Results

In order to prove our main theorems, we need the following limit theorem that shows that the population of a supercritical multi-type branching

process grows geometrically under the condition $E\|Z_1\| \log \|Z_1\| < \infty$. (Recall that we have assumed that there is no extinction.)

THEOREM 3.1. (Kesten and Stigum, 1966) *Let $1 < \rho < \infty$.*

(a) *Let $W_n = \frac{\mathbf{u} \cdot \mathbf{Z}_n}{\rho^n}$ 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 $W \equiv \lim_{n \rightarrow \infty} W_n$ exists with probability 1.*

(b) *Furthermore,*

(i) $P(0 < W < \infty) = 1$ *if and only if* $E\|Z_1\| \log \|Z_1\| < \infty$.

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

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

(c) *For any initial $|Z_0| \neq 0$,*

$$\lim_{n \rightarrow \infty} \left(\frac{\mathbf{Z}_n}{\mathbf{u} \cdot \mathbf{Z}_n} \right) = \mathbf{v} \quad \text{with probability 1.}$$

3.1. *Proof of Theorem 2.1.* We need the following lemma to prove this theorem.

LEMMA 3.1. (O'Brien, 1980) *Assume W_1, W_2, \dots are pairwise independent, and identically distributed and positive random variables. Then,*

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

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

Now, we begin to prove Theorem 2.1.

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

For any $k \geq 2$, we pick k individuals by SRSWOR from the population in the n th generation and let $X_{n,k}$ be the generation number of their last common ancestor.

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

$$\begin{aligned}
 & P(X_{n,k} \geq r | \mathcal{T}) \\
 &= \frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} |\mathbf{Z}_{r,i,n-r}^{(l)}| \left(|\mathbf{Z}_{r,i,n-r}^{(l)}| - 1 \right) \cdots \left(|\mathbf{Z}_{r,i,n-r}^{(l)}| - k + 1 \right)}{|\mathbf{Z}_n| (|\mathbf{Z}_n| - 1) \cdots (|\mathbf{Z}_n| - k + 1)} \\
 &= \frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}_{r,i,n-r}^{(l)}|}{\rho^{n-r}} \frac{|\mathbf{Z}_{r,i,n-r}^{(l)}| - 1}{\rho^{n-r}} \cdots \frac{|\mathbf{Z}_{r,i,n-r}^{(l)}| - k + 1}{\rho^{n-r}}}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}_{r,i,n-r}^{(l)}|}{\rho^{n-r}} \right) \left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}_{r,i,n-r}^{(l)}|}{\rho^{n-r}} - \frac{1}{\rho^{n-r}} \right) \cdots \left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \frac{|\mathbf{Z}_{r,i,n-r}^{(l)}|}{\rho^{n-r}} - \frac{k-1}{\rho^{n-r}} \right)} \quad (3.1)
 \end{aligned}$$

Since $1 < \rho < \infty$ and $E\|Z_1\| \log \|Z_1\| < \infty$, by Theorem 3.1, we know that $\frac{|\mathbf{Z}_{r,i,n-r}^{(l)}|}{\rho^{n-r}} \rightarrow (\mathbf{1} \cdot \mathbf{v}) W_{r,i}^{(l)}$ with probability 1 as $n \rightarrow \infty$, for all r, i, l . So, as $n \rightarrow \infty$,

$$\begin{aligned}
 P(X_{n,k} \geq r | \mathcal{T}) &\rightarrow \frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left((\mathbf{1} \cdot \mathbf{v}) W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} (\mathbf{1} \cdot \mathbf{v}) W_{r,i}^{(l)} \right)^k} \quad \text{with probability 1} \\
 &= \frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} \equiv 1 - \phi_k(r, \mathcal{T})
 \end{aligned}$$

and hence (a) is proved.

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

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

for $r = 1, 2, \dots$. Now, averaged over all trees \mathcal{T} , conditioned on \mathbf{Z}_r , the random variables $W_{r,i}^{(l)}$, $i = 1, 2, \dots, Z_r^{(l)}$ are i.i.d for each $l = 1, 2, \dots, d$.

Moreover, since $E\|Z_1\| \log \|Z_1\| < \infty$, by Theorem 3.1, $EW^{(l)} < \infty$ for $l = 1, 2, \dots, d$. Hence, dropping l , if we let

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

then $L(x)$ is slowly varying at ∞ since EW is finite. That is, for $l = 1, 2, \dots, d$, the function $E(W_{r,1}^{(l)} : W_{r,1}^{(l)} \leq x)$ in x is slowly varying at ∞ .

Therefore, by Lemma 3.1, for each $l = 1, 2, \dots, d$

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

as $n \rightarrow \infty$. So, since $|Z_r| \rightarrow \infty$ with probability 1 as $r \rightarrow \infty$, by the bounded convergence theorem, we have

$$E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} \right) \rightarrow 0 \quad \text{as } r \rightarrow \infty.$$

Thus, ϕ_k is a proper probability distribution. So, there exists a random variable \tilde{X}_k with $P(\tilde{X}_k < r) = \phi_k(r)$ for any $r \geq 1$ such that $X_{n,k} \xrightarrow{d} \tilde{X}_k$ as $n \rightarrow \infty$ and we have completed the proof of Theorem 2.1.

3.2. Proof of Theorem 2.2. 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 $r = 1, 2, \dots$, we have that

$$\phi_k(r, \mathcal{T}) = 1 - \frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} = \begin{cases} 0 & \text{if } r \leq U - 1 \\ 1 - \frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} & \text{if } r \geq U \end{cases}$$

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

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

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

$$\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} \leq \left(\frac{\max_{\substack{1 \leq i \leq Z_r^{(l)} \\ 1 \leq l \leq d}} W_{r,i}^{(l)}}{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)}} \right)^{k-1} \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

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

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

Step 2:

We have that

$$\begin{aligned} & E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} \right) \\ &= E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} I(r \leq U - 1) \right) + E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} I(r \geq U) \right) \\ &= P(r \leq U - 1) + E \left(E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} I(r \geq U) \middle| \mathbf{Z}_r \right) \right) \end{aligned}$$

Since, averaged over all trees \mathcal{T} and conditioned on \mathbf{Z}_r , $\{W_{r,i}^{(l)} : i = 1, 2, \dots, Z_r^{(l)}\}$ are i.i.d., for each $l = 1, 2, \dots, d$,

$$\begin{aligned} & E \left(E \left(\frac{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} \left(W_{r,i}^{(l)} \right)^k}{\left(\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)} \right)^k} I(r \geq U) \middle| \mathbf{Z}_r \right) \right) \\ &= E \left(\sum_{l=1}^d Z_r^{(l)} E \left(\left(\frac{W_{r,1}^{(l)}}{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)}} I(r \geq U) \right)^k \middle| \mathbf{Z}_r \right) \right) \end{aligned}$$

Also, $P \left(0 < \frac{W_{r,i}^{(l)}}{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)}} < 1 \middle| r \geq U \right) = 1$ implies that $\left(\frac{W_{r,i}^{(l)}}{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)}} I(r \geq U) \right)^k \rightarrow 0$ with probability 1 as $k \rightarrow \infty$, and hence

$$E \left(\sum_{l=1}^d Z_r^{(l)} E \left(\left(\frac{W_{r,1}^{(l)}}{\sum_{l=1}^d \sum_{i=1}^{Z_r^{(l)}} W_{r,i}^{(l)}} I(r \geq U) \right)^k \middle| \mathbf{Z}_r \right) \right) \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

by the bounded convergence theorem. Therefore, as $k \rightarrow \infty$,

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

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

3.3. Proof of Theorem 2.3. The following proof also can be extended to prove Theorem 2.4.

Let $\xi_{n,j}^{(i)} = (\xi_{n,j}^{(i)1}, \xi_{n,j}^{(i)2}, \dots, \xi_{n,j}^{(i)d})$ be the vector of the offsprings of the j th individual of the type i in the n th generation. Let $\{\mathbf{Z}_{p,r,s,n}^{j(l)}\}_{n \geq 0}$ be the multi-type Galton-Watson branching process initiated by the s th child of type l of the p th individual of type j in the r th generation. So, $\{\mathbf{Z}_{p,r,s,n}^{j(l)} = (Z_{p,r,s,n}^{j(l)1}, Z_{p,r,s,n}^{j(l)2}, \dots, Z_{p,r,s,n}^{j(l)d})\}_{n \geq 0}$ 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 i th chosen individual, $i = 1, 2$. Then

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2}) \\
&= E\left(P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} = A_{n,2} | \mathcal{T})\right) \\
&= E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} Z_{p,r,s,n-r-1}^{j(l)i} Z_{p,r,t,n-r-1}^{j(l)i}}{|\mathbf{Z}_n|(|\mathbf{Z}_n| - 1)}\right) \\
&= E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} \frac{Z_{p,r,s,n-r-1}^{j(l)i}}{\rho^{n-r-1}} \frac{Z_{p,r,t,n-r-1}^{j(l)i}}{\rho^{n-r-1}}}{\frac{|\mathbf{Z}_n|}{\rho^{n-r-1}} \frac{|\mathbf{Z}_n| - 1}{\rho^{n-r-1}}}\right) \\
&\rightarrow E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} (v_i W_{p,r,s})(v_i W_{p,r,t})}{\left(\sum_{l=1}^d \sum_{s=1}^{Z_{r+1}^{(l)}} W_{r+1,s}^{(l)}\right)^2}\right) \quad \text{as } n \rightarrow \infty \\
&= E\left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} v_i^2 W_{p,r,s} W_{p,r,t}}{\left(\sum_{l=1}^d \sum_{s=1}^{Z_{r+1}^{(l)}} W_{r+1,s}^{(l)}\right)^2}\right) \tag{3.2}
\end{aligned}$$

Conditioned on \mathbf{Z}_r and \mathbf{Z}_{r+1} and averaged over all trees \mathcal{T} , the random variables $\{W_{p,r,s} : s = 1, 2, \dots, \xi_{r,p}^{(j)l}, l = 1, 2, \dots, d, p = 1, 2, \dots, Z_r^{(j)}\}$ are independent. Further, conditioned on \mathbf{Z}_{r+1} and averaged over all trees \mathcal{T} , the random variables $\{W_{r+1,s}^{(l)} : s = 1, 2, \dots, Z_{r+1}^{(l)}, l = 1, 2, \dots, d\}$ are independent as well.

Similarly, we have, as $n \rightarrow \infty$,

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, \zeta_{n,1} = \zeta_{n,2} = i, A_{n,1} \neq A_{n,2}) \\
&= E \left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l \neq q=1}^d \sum_{s=1}^{\xi_{r,p}^{(j)l}} \sum_{t=1}^{\xi_{r,p}^{(j)q}} \frac{Z_{p,r,s,n-r-1}^{j(l)i}}{\rho^{n-r-1}} \frac{Z_{p,r,t,n-r-1}^{j(l)i}}{\rho^{n-r-1}}}{\frac{|\mathbf{Z}_n|}{\rho^{n-r-1}} \frac{|\mathbf{Z}_n|-1}{\rho^{n-r-1}}} \right) \\
&\rightarrow E \left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l \neq q=1}^d \sum_{s=1}^{\xi_{r,p}^{(j)l}} \sum_{t=1}^{\xi_{r,p}^{(j)q}} v_i^2 W_{p,r,s} W_{p,r,t}}{\left(\sum_{l=1}^d \sum_{s=1}^{Z_{r+1}^{(l)}} W_{r+1,s}^{(l)} \right)^2} \right) \quad \text{as } n \rightarrow \infty, \quad (3.3)
\end{aligned}$$

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, i_1 = \zeta_{n,1} \neq \zeta_{n,2} = i_2, A_{n,1} = A_{n,2}) \\
&= E \left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} \frac{Z_{p,r,s,n-r-1}^{j(l)i_1}}{\rho^{n-r-1}} \frac{Z_{p,r,t,n-r-1}^{j(l)i_2}}{\rho^{n-r-1}}}{\frac{|\mathbf{Z}_n|}{\rho^{n-r-1}} \frac{|\mathbf{Z}_n|-1}{\rho^{n-r-1}}} \right) \\
&\rightarrow E \left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} (v_{i_1} W_{p,r,s})(v_{i_2} W_{p,r,t})}{\left(\sum_{l=1}^d \sum_{s=1}^{Z_{r+1}^{(l)}} W_{r+1,s}^{(l)} \right)^2} \right) \quad \text{as } n \rightarrow \infty \\
&= E \left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l=1}^d \sum_{s \neq t=1}^{\xi_{r,p}^{(j)l}} v_{i_1} v_{i_2} W_{p,r,s} W_{p,r,t}}{\left(\sum_{l=1}^d \sum_{s=1}^{Z_{r+1}^{(l)}} W_{r+1,s}^{(l)} \right)^2} \right), \quad (3.4)
\end{aligned}$$

and

$$\begin{aligned}
& P(X_{n,2} = r, \eta_n = j, i_1 = \zeta_{n,1} \neq \zeta_{n,2} = i_2, A_{n,1} \neq A_{n,2}) \\
&\rightarrow E \left(\frac{\sum_{p=1}^{Z_r^{(j)}} \sum_{l \neq q=1}^d \sum_{s=1}^{\xi_{r,p}^{(j)l}} \sum_{t=1}^{\xi_{r,p}^{(j)q}} v_{i_1} v_{i_2} W_{p,r,s} W_{p,r,t}}{\left(\sum_{l=1}^d \sum_{s=1}^{Z_{r+1}^{(l)}} W_{r+1,s}^{(l)} \right)^2} \right) \quad \text{as } n \rightarrow \infty. \quad (3.5)
\end{aligned}$$

Therefore, (3.2), (3.3), (3.4) and (3.5) together yield, as $n \rightarrow \infty$,

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

By Theorem 2.1, 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 $\{1, 2, \dots, \alpha\}$. 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.

References

- ATHREYA, K. B. and SCHUH, H.-J. (2003). On the supercritical Bellman-Harris process with finite mean. *Indian J. Statist.*, **65**(2), 229–248.
- ATHREYA, K.B. and NEY, P. (2004). *Branching Processes*. Dover, New York.
- ATHREYA, K. B. (2012). Coalescence in recent past in rapidly growing populations. *Stoch. Process. Appl.*, **122**(2012), 3757–3766.
- HOPPE, F. M. (1976). Supercritical multitype branching processes. *Ann. Probab.*, **4**(3), 393–401.
- KARLIN, S. and TAYLOR, H. (1975). *A First Course in Stochastic Processes*, 2nd. Academic Press.
- KESTEN, H. and STIGUM, B.P. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.*, **37**, 1211–1223.
- O'BRIEN, G. L. (1980). A limit theorem for sample maxima and heavy branches in Galton-Watson trees. *J. Appl. Probab.*, **17**, 539–545.

JYY-I HONG
IOWA STATE UNIVERSITY
AMES, IA, USA
E-mail: hongjyyi@gmail.com