



## Beyond the $Q$ -process: various ways of conditioning the multitype Galton-Watson process

Sophie Pénisson

Université Paris-Est, LAMA (UMR 8050), UPEMLV, UPEC, CNRS,  
94010 Créteil, France.

*E-mail address:* [sophie.penisson@u-pec.fr](mailto:sophie.penisson@u-pec.fr)

*URL:* <http://www.sophiepenisson.com/>

**Abstract.** Conditioning a multitype Galton-Watson process to stay alive into the indefinite future leads to what is known as its associated  $Q$ -process. We show that the same holds true if the process is conditioned to reach a positive threshold or a non-absorbing state. We also demonstrate that the stationary measure of the  $Q$ -process, obtained by construction as two successive limits (first by delaying the extinction in the original process and next by considering the long-time behavior of the obtained  $Q$ -process), is as a matter of fact a double limit. Finally, we prove that conditioning a multitype branching process on having an infinite total progeny leads to a process presenting the features of a  $Q$ -process. It does not however coincide with the original associated  $Q$ -process, except in the critical regime.

### 1. Introduction

The benchmark of our study is the  $Q$ -process associated with a multitype Galton-Watson (GW) process, obtained by conditioning the branching process  $\mathbf{X}_k$  on not being extinct in the distant future ( $\{\mathbf{X}_{k+n} \neq \mathbf{0}\}$ , with  $n \rightarrow +\infty$ ) and on the event that extinction takes place ( $\{\lim_l \mathbf{X}_l = \mathbf{0}\}$ ) (see [Nakagawa, 1978](#)). Our goal is to investigate some seemingly comparable conditioning results and to relate them to the  $Q$ -process.

After a description of the basic assumptions on the multitype GW process, we start in Subsection 1.4 by describing the “associated” branching process, which will be a key tool when conditioning on the event that extinction takes place, or when conditioning on an infinite total progeny.

We shall first prove in Section 3 that by replacing in what precedes the conditioning event  $\{\mathbf{X}_{k+n} \neq \mathbf{0}\}$  by  $\{\mathbf{X}_{k+n} \in S\}$ , where  $S$  is a subset which does not

---

*Received by the editors September 23, 2015; accepted March 4, 2016.*

*2010 Mathematics Subject Classification.* 60J80, 60F05.

*Key words and phrases.* Multitype branching process, Conditioned limit theorem, Quasi-stationary distribution,  $Q$ -process, Size-biased distribution, Total progeny.

contain  $\mathbf{0}$ , the obtained limit process remains the  $Q$ -process. This means in particular that conditioning in the distant future on reaching a non-zero state or a positive threshold, instead of conditioning on non-extinction, does not alter the result.

In a second instance, we focus in the noncritical case on the stationary measure of the positive recurrent  $Q$ -process. Formulated in a loose manner, this measure is obtained by considering  $\{\mathbf{X}_k \mid \mathbf{X}_{k+n} \neq \mathbf{0}\}$ , by delaying the extinction time ( $n \rightarrow \infty$ ), and by studying the long-time behavior of the limit process ( $k \rightarrow \infty$ ). It is already known (Nakagawa, 1978) that inverting the limits leads to the same result. We prove in Section 4 that the convergence to the stationary measure still holds even if  $n$  and  $k$  simultaneously grow to infinity. This requires an additional second-order moment assumption if the process is subcritical.

Finally, we investigate in Section 5 the distribution of the multitype GW process conditioned on having an infinite total progeny. This is motivated by Kennedy's result, who studies in Kennedy (1975) the behavior of a monotype GW process  $X_k$  conditioned on the event  $\{N = n\}$  as  $n \rightarrow +\infty$ , where  $N = \sum_{k=0}^{+\infty} X_k$  denotes the total progeny. Note that the latter conditioning seems comparable to the device of conditioning on the event that extinction occurs but has not done so by generation  $n$ . It is indeed proven in the aforementioned paper that in the critical case, conditioning on the total progeny or on non-extinction indifferently results in the  $Q$ -process. This result has since then been extended for instance to monotype GW trees and to other conditionings: in the critical case, conditioning a GW tree by its height, by its total progeny or by its number of leaves leads to the same limiting tree (see e.g. Abraham and Delmas, 2014; Janson, 2012). However, in the noncritical case, the two methods provide different limiting results: the limit process is always the  $Q$ -process of some critical process, no matter the class of criticality of the original process. Under a moment assumption (depending on the number of types of the process), we generalize this result to the multitype case. For this purpose we assume that the total progeny increases to infinity according to the "typical" limiting type proportions of the associated critical GW process, by conditioning on the event  $\{\mathbf{N} = \lfloor n\mathbf{w} \rfloor\}$  as  $n \rightarrow \infty$ , where  $\mathbf{w}$  is a left eigenvector related to the maximal eigenvalue 1 of the mean matrix of the critical process.

1.1. *Notation.* Let  $d \geq 1$ . In this paper, a generic point in  $\mathbb{R}^d$  is denoted by  $\mathbf{x} = (x_1, \dots, x_d)$ , and its transpose is written  $\mathbf{x}^T$ . By  $\mathbf{e}_i = (\delta_{i,j})_{1 \leq j \leq d}$  we denote the  $i$ -th unit vector in  $\mathbb{R}^d$ , where  $\delta_{i,j}$  stands for the Kronecker delta. We write  $\mathbf{0} = (0, \dots, 0)$  and  $\mathbf{1} = (1, \dots, 1)$ . The notation  $\mathbf{xy}$  (resp.  $\lfloor \mathbf{x} \rfloor$ ) stands for the vector with coordinates  $x_i y_i$  (resp.  $\lfloor x_i \rfloor$ , the integer part of  $x_i$ ). We denote by  $\mathbf{x}^{\mathbf{y}}$  the product  $\prod_{i=1}^d x_i^{y_i}$ . The obvious partial order on  $\mathbb{R}^d$  is  $\mathbf{x} \leq \mathbf{y}$ , when  $x_i \leq y_i$  for each  $i$ , and  $\mathbf{x} < \mathbf{y}$  when  $x_i < y_i$  for each  $i$ . Finally,  $\mathbf{x} \cdot \mathbf{y}$  denotes the scalar product in  $\mathbb{R}^d$ ,  $\|\mathbf{x}\|_1$  the  $L^1$ -norm and  $\|\mathbf{x}\|_2$  the  $L^2$ -norm.

1.2. *Multitype GW processes.* Let  $(\mathbf{X}_k)_{k \geq 0}$  denote a  $d$ -type GW process, with  $n$ -th transition probabilities  $P_n(\mathbf{x}, \mathbf{y}) = \mathbb{P}(\mathbf{X}_{k+n} = \mathbf{y} \mid \mathbf{X}_k = \mathbf{x})$ ,  $k, n \in \mathbb{N}$ ,  $\mathbf{x}, \mathbf{y} \in \mathbb{N}^d$ . Let  $\mathbf{f} = (f_1, \dots, f_d)$  be its offspring generating function, where for each  $i = 1 \dots d$  and  $\mathbf{r} \in [0, 1]^d$ ,  $f_i(\mathbf{r}) = \mathbb{E}_{\mathbf{e}_i}(\mathbf{r}^{\mathbf{X}_1}) = \sum_{\mathbf{k} \in \mathbb{N}^d} p_i(\mathbf{k}) \mathbf{r}^{\mathbf{k}}$ , the subscript  $\mathbf{e}_i$  denoting the initial condition, and  $p_i$  the offspring probability distribution of type  $i$ . For each  $i$ , we denote by  $\mathbf{m}^i = (m_{i1}, \dots, m_{id})$  (resp.  $\Sigma^i$ ) the mean vector (resp. covariance matrix) of the offspring probability distribution  $p_i$ . The mean matrix is then given

by  $\mathbf{M} = (m_{ij})_{1 \leq i, j \leq d}$ . If it exists, we denote by  $\rho$  its Perron's root, and by  $\mathbf{u}$  and  $\mathbf{v}$  the associated right and left eigenvectors (i.e. such that  $\mathbf{M}\mathbf{u}^T = \rho\mathbf{u}^T$ ,  $\mathbf{v}\mathbf{M} = \rho\mathbf{v}$ ), with the normalization convention  $\mathbf{u} \cdot \mathbf{1} = \mathbf{u} \cdot \mathbf{v} = 1$ . The process is then called critical (resp. subcritical, supercritical) if  $\rho = 1$  (resp.  $\rho < 1$ ,  $\rho > 1$ ). In what follows we shall denote by  $\mathbf{f}_n$  the  $n$ -th iterate of the function  $\mathbf{f}$ , and by  $\mathbf{M}^n = (m_{ij}^{(n)})_{1 \leq i, j \leq d}$  the  $n$ -th power of the matrix  $\mathbf{M}$ , which correspond respectively to the generating function and mean matrix of the process at time  $n$ . By the branching property, for each  $\mathbf{x} \in \mathbb{N}^d$ , the function  $\mathbf{f}_n^{\mathbf{x}}$  then corresponds to the generating function of the process at time  $n$  with initial state  $\mathbf{x}$ , namely  $\mathbb{E}_{\mathbf{x}}(\mathbf{r}^{\mathbf{X}^n}) = \mathbf{f}_n(\mathbf{r})^{\mathbf{x}}$ . Finally, we define the extinction time  $T = \inf\{k \in \mathbb{N}, \mathbf{X}_k = \mathbf{0}\}$ , and the extinction probability vector  $\mathbf{q} = (q_1, \dots, q_d)$ , given by  $q_i = \mathbb{P}_{\mathbf{e}_i}(T < +\infty)$ ,  $i = 1 \dots d$ .

1.3. *Basic assumptions.*

- (A<sub>1</sub>) The mean matrix  $\mathbf{M}$  is finite. The process is nonsingular ( $\mathbf{f}(\mathbf{r}) \neq \mathbf{M}\mathbf{r}$ ), positive regular (there exists some  $n \in \mathbb{N}^*$  such that each entry of  $\mathbf{M}^n$  is positive), and is such that  $\mathbf{q} > \mathbf{0}$ .

The latter statement will always be assumed. It ensures in particular the existence of the Perron's root  $\rho$  and that (Karlin, 1966),

$$\lim_{n \rightarrow +\infty} \rho^{-n} m_{ij}^{(n)} = u_i v_j. \tag{1.1}$$

When necessary, the following additional assumptions will be made.

- (A<sub>2</sub>) For each  $i, j = 1 \dots d$ ,  $\mathbb{E}_{\mathbf{e}_i}(X_{1,j} \ln X_{1,j}) < +\infty$ .
- (A<sub>3</sub>) The covariance matrices  $\Sigma^i$ ,  $i = 1 \dots d$ , are finite.

1.4. *The associated process.* For any vector  $\mathbf{a} > \mathbf{0}$  such that for each  $i = 1 \dots d$ ,  $f_i(\mathbf{a}) < +\infty$ , we define the generating function  $\bar{\mathbf{f}} = (\bar{f}_1, \dots, \bar{f}_d)$  on  $[0, 1]^d$  as follows:

$$\bar{f}_i(\mathbf{r}) = \frac{f_i(\mathbf{a}\mathbf{r})}{f_i(\mathbf{a})}, \quad i = 1 \dots d.$$

We then denote by  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  the GW process with offspring generating function  $\bar{\mathbf{f}}$ , which will be referred to as the *associated process* with respect to  $\mathbf{a}$ . We shall denote by  $\bar{P}_n, \bar{p}_i$  etc. its transition probabilities, offspring probability distributions etc. We easily compute that for each  $n \geq 1$ ,  $i = 1 \dots d$ ,  $\mathbf{k} \in \mathbb{N}^d$  and  $\mathbf{r} \in [0, 1]^d$ , denoting by  $*$  the convolution product,

$$\bar{p}_i^{*n}(\mathbf{k}) = \frac{\mathbf{a}^{\mathbf{k}}}{f_i(\mathbf{a})^n} p_i^{*n}(\mathbf{k}), \quad \bar{f}_{n,i}(\mathbf{r}) = \frac{f_{n,i}(\mathbf{a}\mathbf{r})}{f_i(\mathbf{a})^n}. \tag{1.2}$$

*Remark 1.1.* It is known (Jagers and Lagerås, 2008) that a supercritical GW process conditioned on the event  $\{T < +\infty\}$  is subcritical. By construction, its offspring generating function is given by  $\mathbf{r} \mapsto f_i(\mathbf{q}\mathbf{r})/q_i$ . Since the extinction probability vector satisfies  $\mathbf{f}(\mathbf{q}) = \mathbf{q}$  (Harris, 1963), this means that the associated process  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  with respect to  $\mathbf{q}$  is subcritical.

2. Classical results: conditioning on non-extinction

2.1. *The Yaglom distribution (Joffe and Spitzer, 1967, Theorem 3).* Let  $(\mathbf{X}_k)_{k \geq 0}$  be a subcritical multitype GW process satisfying  $(A_1)$ . Then for all  $\mathbf{x}_0, \mathbf{z} \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,

$$\lim_{k \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_k = \mathbf{z} \mid \mathbf{X}_k \neq \mathbf{0}) = \nu(\mathbf{z}), \quad (2.1)$$

where  $\nu$  is a probability distribution on  $\mathbb{N}^d \setminus \{\mathbf{0}\}$  independent of the initial state  $\mathbf{x}_0$ . This quasi-stationary distribution is often referred to as the Yaglom distribution associated with  $(\mathbf{X}_k)_{k \geq 0}$ . We shall denote by  $g$  its generating function  $g(\mathbf{r}) = \sum_{\mathbf{z} \neq \mathbf{0}} \nu(\mathbf{z}) \mathbf{r}^{\mathbf{z}}$ . Under  $(A_2)$ ,  $\nu$  admits finite and positive first moments

$$\frac{\partial g(\mathbf{1})}{\partial r_i} = v_i \gamma^{-1}, \quad i = 1 \dots d, \quad (2.2)$$

where  $\gamma > 0$  is a limiting quantity satisfying for each  $\mathbf{x} \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,

$$\lim_{k \rightarrow +\infty} \rho^{-k} \mathbb{P}_{\mathbf{x}}(\mathbf{X}_k \neq \mathbf{0}) = \gamma \mathbf{x} \cdot \mathbf{u}. \quad (2.3)$$

2.2. *The Q-process (Nakagawa, 1978, Theorem 2).* Let  $(\mathbf{X}_k)_{k \geq 0}$  be a multitype GW process satisfying  $(A_1)$ . Then for all  $\mathbf{x}_0 \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,  $k_1 \leq \dots \leq k_j \in \mathbb{N}$ , and  $\mathbf{x}_1, \dots, \mathbf{x}_j \in \mathbb{N}^d$ ,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{X}_{k_j+n} \neq \mathbf{0}, T < +\infty) \\ = \frac{1}{\rho^{k_j}} \frac{\mathbf{x}_j \cdot \bar{\mathbf{u}}}{\mathbf{x}_0 \cdot \bar{\mathbf{u}}} \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j), \end{aligned} \quad (2.4)$$

where  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  is the associated process with respect to  $\mathbf{q}$ . As told in the introduction, this limiting process is the  $Q$ -process associated with  $(\mathbf{X}_k)_{k \geq 0}$ . It is Markovian with transition probabilities

$$Q_1(\mathbf{x}, \mathbf{y}) = \frac{1}{\rho} \frac{\mathbf{y} \cdot \bar{\mathbf{u}}}{\mathbf{x} \cdot \bar{\mathbf{u}}} \bar{P}_1(\mathbf{x}, \mathbf{y}) = \frac{1}{\rho} \mathbf{q}^{\mathbf{y}-\mathbf{x}} \frac{\mathbf{y} \cdot \bar{\mathbf{u}}}{\mathbf{x} \cdot \bar{\mathbf{u}}} P_1(\mathbf{x}, \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{N}^d \setminus \{\mathbf{0}\}.$$

If  $\rho > 1$ , the  $Q$ -process is positive recurrent. If  $\rho = 1$ , it is transient. If  $\rho < 1$ , the  $Q$ -process is positive recurrent if and only if  $(A_2)$  is satisfied. In the positive recurrent case, the stationary measure for the  $Q$ -process is given by the size-biased Yaglom distribution

$$\bar{\mu}(\mathbf{z}) = \frac{\mathbf{z} \cdot \mathbf{u} \bar{\nu}(\mathbf{z})}{\sum_{\mathbf{y} \in \mathbb{N}^d \setminus \{\mathbf{0}\}} \mathbf{y} \cdot \mathbf{u} \bar{\nu}(\mathbf{y})}, \quad \mathbf{z} \in \mathbb{N}^d \setminus \{\mathbf{0}\}, \quad (2.5)$$

where  $\bar{\nu}$  is the Yaglom distribution associated with the subcritical process  $(\bar{\mathbf{X}}_k)_{k \geq 0}$ .

2.3. *A Yaglom-type distribution (Nakagawa, 1978, Theorem 3).* Let  $(\mathbf{X}_k)_{k \geq 0}$  be a noncritical multitype GW process satisfying  $(A_1)$ . Then for all  $\mathbf{x}_0, \mathbf{z} \in \mathbb{N}^d \setminus \{\mathbf{0}\}$  and  $n \in \mathbb{N}$ ,  $\lim_k \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_k = \mathbf{z} \mid \mathbf{X}_{k+n} \neq \mathbf{0}, T < +\infty) = \bar{\nu}^{(n)}(\mathbf{z})$ , where  $\bar{\nu}^{(n)}$  is a probability distribution on  $\mathbb{N}^d \setminus \{\mathbf{0}\}$  independent of the initial state  $\mathbf{x}_0$ . In particular,  $\bar{\nu}^{(0)} = \bar{\nu}$  is the Yaglom distribution associated with  $(\bar{\mathbf{X}}_k)_{k \geq 0}$ , the associated subcritical process with respect to  $\mathbf{q}$ . Moreover, assuming in addition  $(A_2)$  if  $\rho < 1$ , then for each  $\mathbf{z} \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,  $\lim_n \bar{\nu}^{(n)}(\mathbf{z}) = \bar{\mu}(\mathbf{z})$ .

### 3. Conditioning on reaching a certain state or threshold

In this section we shall generalize (2.4) by proving that by replacing the conditioning event  $\{\mathbf{X}_{k_j+n} \neq \mathbf{0}\}$  by  $\{\mathbf{X}_{k_j+n} \in S\}$ , where  $S$  is a subset of  $\mathbb{N}^d \setminus \{\mathbf{0}\}$ , the obtained limit process remains the  $Q$ -process. In particular, conditioning the process on reaching a certain non-zero state or positive threshold in a distant future, i.e. with

$$S = \{\mathbf{y}\}, S = \{\mathbf{x} \in \mathbb{N}^d, \|\mathbf{x}\|_1 = m\} \text{ or } S = \{\mathbf{x} \in \mathbb{N}^d, \|\mathbf{x}\|_1 \geq m\},$$

( $\mathbf{y} \neq \mathbf{0}, m > 0$ ), leads to the same result as conditioning the process on non-extinction.

In what follows we call a subset  $S$  accessible if for any  $\mathbf{x} \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ , there exists some  $n \in \mathbb{N}$  such that  $\mathbb{P}_{\mathbf{x}}(\mathbf{X}_n \in S) > 0$ . For any subset  $S$  we shall denote  $S^c = \mathbb{N}^d \setminus (\{\mathbf{0}\} \cup S)$ .

**Theorem 3.1.** *Let  $(\mathbf{X}_k)_{k \geq 0}$  be a multitype GW process satisfying  $(A_1)$ , and let  $S$  be a subset of  $\mathbb{N}^d \setminus \{\mathbf{0}\}$ . If  $\rho \leq 1$  we assume in addition one of the following assumptions:*

- (a<sub>1</sub>)  $S$  is finite and accessible,
- (a<sub>2</sub>)  $S^c$  is finite,
- (a<sub>3</sub>)  $(\mathbf{X}_k)_{k \geq 0}$  is subcritical and satisfies  $(A_2)$ .

Then for all  $\mathbf{x}_0 \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,  $k_1 \leq \dots \leq k_j \in \mathbb{N}^*$  and  $\mathbf{x}_1, \dots, \mathbf{x}_j \in \mathbb{N}^d$ ,

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{X}_{k_j+n} \in S, T < +\infty) \\ &= \frac{1}{\rho^{k_j}} \frac{\mathbf{x}_j \cdot \bar{\mathbf{u}}}{\mathbf{x}_0 \cdot \bar{\mathbf{u}}} \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j), \end{aligned} \quad (3.1)$$

where  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  is the associated process with respect to  $\mathbf{q}$ .

*Proof:* Note that if  $\rho > 1$ , then  $\mathbf{q} < \mathbf{1}$  (Athreya and Ney, 1972) which implies that  $\mathbb{E}_{\mathbf{e}_i}(\bar{X}_{1,j} \ln \bar{X}_{1,j}) < +\infty$ , meaning that  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  automatically satisfies  $(A_2)$ . Thanks to Remark 1.1, we can thus assume without loss of generality that  $\rho \leq 1$  and simply consider the limit

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{X}_{k_j+n} \in S) \\ &= \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j) \frac{\mathbb{P}_{\mathbf{x}_j}(\mathbf{X}_n \in S)}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_j+n} \in S)}. \end{aligned} \quad (3.2)$$

Let us recall here some of the technical results established for  $\rho \leq 1$  in Nakagawa (1978), essential to our proof. First, for each  $\mathbf{0} \leq \mathbf{b} < \mathbf{c} \leq \mathbf{1}$  and  $\mathbf{x} \in \mathbb{N}^d$ ,

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{v} \cdot (\mathbf{f}_{n+2}(\mathbf{0}) - \mathbf{f}_{n+1}(\mathbf{0}))}{\mathbf{v} \cdot (\mathbf{f}_{n+1}(\mathbf{0}) - \mathbf{f}_n(\mathbf{0}))} = \rho, \quad (3.3)$$

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{f}_n(\mathbf{c})^{\mathbf{x}} - \mathbf{f}_n(\mathbf{b})^{\mathbf{x}}}{\mathbf{v} \cdot (\mathbf{f}_n(\mathbf{c}) - \mathbf{f}_n(\mathbf{b}))} = \mathbf{x} \cdot \mathbf{u}, \quad (3.4)$$

Moreover, for each  $\mathbf{x}, \mathbf{y} \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,

$$\lim_{n \rightarrow +\infty} \frac{1 - \mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}}}{1 - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}}} = \rho, \quad (3.5)$$

$$P_n(\mathbf{x}, \mathbf{y}) = (\pi(\mathbf{y}) + \varepsilon_n(\mathbf{x}, \mathbf{y})) (\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}}), \quad (3.6)$$

where  $\lim_n \varepsilon_n(\mathbf{x}, \mathbf{y}) = 0$  and  $\pi$  is the unique measure (up to multiplicative constants) on  $\mathbb{N}^d \setminus \{\mathbf{0}\}$  not identically zero satisfying  $\sum_{\mathbf{y} \neq \mathbf{0}} \pi(\mathbf{y}) P(\mathbf{y}, \mathbf{z}) = \rho \pi(\mathbf{z})$  for each  $\mathbf{z} \neq \mathbf{0}$ . In particular, if  $\rho < 1$ ,  $\pi = (1 - \rho)^{-1} \nu$ , where  $\nu$  is the probability distribution defined by (2.1).

Let us first assume (a<sub>1</sub>). By (3.6)

$$\begin{aligned} \frac{\mathbb{P}_{\mathbf{x}_j}(\mathbf{X}_n \in S)}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_j+n} \in S)} &= \frac{\sum_{\mathbf{z} \in S} P_n(\mathbf{x}_j, \mathbf{z})}{\sum_{\mathbf{z} \in S} P_{n+k_j}(\mathbf{x}_0, \mathbf{z})} \\ &= \frac{\pi(S) + \varepsilon_n(\mathbf{x}_j)}{\pi(S) + \varepsilon_{n+k_j}(\mathbf{x}_0)} \frac{\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}_j} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}_j}}{\mathbf{f}_{n+k_j+1}(\mathbf{0})^{\mathbf{x}_0} - \mathbf{f}_{n+k_j}(\mathbf{0})^{\mathbf{x}_0}}, \end{aligned} \quad (3.7)$$

where  $\lim_n \varepsilon_n(\mathbf{x}) = \lim_n \sum_{\mathbf{z} \in S} \varepsilon_n(\mathbf{x}, \mathbf{z}) = 0$  since  $S$  is finite. On the one hand, we can deduce from (3.3) and (3.4) that

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}_j} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}_j}}{\mathbf{f}_{n+k_j+1}(\mathbf{0})^{\mathbf{x}_0} - \mathbf{f}_{n+k_j}(\mathbf{0})^{\mathbf{x}_0}} = \frac{1}{\rho^{k_j}} \frac{\mathbf{x}_j \cdot \mathbf{u}}{\mathbf{x}_0 \cdot \mathbf{u}}.$$

On the other hand,  $\pi$  being not identically zero, there exists some  $\mathbf{y}_0 \in \mathbb{N}^d \setminus \{\mathbf{0}\}$  such that  $\pi(\mathbf{y}_0) > 0$ . Since  $S$  is accessible, there exists some  $\mathbf{z}_0 \in S$  and  $k \in \mathbb{N}^*$  such that  $P_k(\mathbf{y}_0, \mathbf{z}_0) > 0$ , and thus

$$+\infty > \pi(S) \geq \pi(\mathbf{z}_0) = \rho^{-k} \sum_{\mathbf{y} \in \mathbb{N}^d \setminus \{\mathbf{0}\}} \pi(\mathbf{y}) P_k(\mathbf{y}, \mathbf{z}_0) \geq \rho^{-k} \pi(\mathbf{y}_0) P_k(\mathbf{y}_0, \mathbf{z}_0) > 0.$$

From (3.7) we thus deduce that (3.2) leads to (3.1).

Let us now assume (a<sub>2</sub>). We can similarly deduce from (3.6) that

$$\begin{aligned} \frac{\mathbb{P}_{\mathbf{x}_j}(\mathbf{X}_n \in S)}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_j+n} \in S)} &= \frac{1 - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}_j} - (\pi(S^c) + \varepsilon_n(\mathbf{x}_j)) (\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}_j} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}_j})}{1 - \mathbf{f}_{n+k_j}(\mathbf{0})^{\mathbf{x}_0} - (\pi(S^c) + \varepsilon_{n+k_j}(\mathbf{x}_0)) (\mathbf{f}_{n+k_j+1}(\mathbf{0})^{\mathbf{x}_0} - \mathbf{f}_{n+k_j}(\mathbf{0})^{\mathbf{x}_0})}, \end{aligned} \quad (3.8)$$

with  $0 \leq \pi(S^c) < +\infty$  and  $\lim_n \varepsilon_n(\mathbf{x}) = \lim_n \sum_{\mathbf{z} \in S^c} \varepsilon_n(\mathbf{x}, \mathbf{z}) = 0$  since  $S^c$  is finite. Note that (2.4) implies that

$$\lim_{n \rightarrow +\infty} \frac{1 - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}_j}}{1 - \mathbf{f}_{n+k_j}(\mathbf{0})^{\mathbf{x}_0}} = \frac{1}{\rho^{k_j}} \frac{\mathbf{x}_j \cdot \mathbf{u}}{\mathbf{x}_0 \cdot \mathbf{u}},$$

which together with (3.5) enables to show that (3.8) tends to  $\rho^{-k_j} \frac{\mathbf{x}_j \cdot \mathbf{u}}{\mathbf{x}_0 \cdot \mathbf{u}}$  as  $n$  tends to infinity, leading again to (3.1).

Let us finally assume (a<sub>3</sub>). Then we know from Nakagawa (1978, Remark 2) that  $\pi(\mathbf{z}) > 0$  for each  $\mathbf{z} \neq \mathbf{0}$ , hence automatically  $0 < \pi(S) = (1 - \rho)^{-1} \nu(S) < +\infty$ . Moreover,  $\nu$  admits finite first-order moments (see (2.2)). Hence for any  $a > 0$ , by Markov's inequality,

$$\begin{aligned} \left| \sum_{\mathbf{z} \in S} \varepsilon_n(\mathbf{x}, \mathbf{z}) \right| &\leq \sum_{\substack{\mathbf{z} \in S \\ \|\mathbf{z}\|_1 < a}} |\varepsilon_n(\mathbf{x}, \mathbf{z})| + \sum_{\substack{\mathbf{z} \in S \\ \|\mathbf{z}\|_1 \geq a}} \left| \frac{P_n(\mathbf{x}, \mathbf{z})}{\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}}} - \pi(\mathbf{z}) \right| \\ &\leq \sum_{\substack{\mathbf{z} \in S \\ \|\mathbf{z}\|_1 < a}} |\varepsilon_n(\mathbf{x}, \mathbf{z})| + \frac{1}{a} \frac{\mathbb{E}_{\mathbf{x}}(\|\mathbf{X}_n\|_1)}{\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}}} + \frac{1}{1 - \rho} \frac{1}{a} \sum_{i=1}^d \frac{\partial g(\mathbf{1})}{\partial r_i}. \end{aligned}$$

We recall that by (2.3),  $\lim_n \rho^{-n} (\mathbf{f}_{n+1}(\mathbf{0})^{\mathbf{x}} - \mathbf{f}_n(\mathbf{0})^{\mathbf{x}}) = (1 - \rho) \gamma \mathbf{x} \cdot \mathbf{u}$ , while by (1.1),  $\lim_n \rho^{-n} \mathbb{E}_{\mathbf{x}} (\|\mathbf{X}_n\|_1) = \sum_{i,j=1}^d x_i u_i v_j$ . Hence the previous inequality ensures that  $\lim_n \sum_{\mathbf{z} \in S} \varepsilon_n(\mathbf{x}, \mathbf{z}) = 0$ . We can thus write (3.7) even without the finiteness assumption of  $S$ , and prove (3.1) as previously.  $\square$

**4. The size-biased Yaglom distribution as a double limit**

From Subsection 2.2 and Subsection 2.3 we know that in the noncritical case, assuming (A<sub>2</sub>) if  $\rho < 1$ ,

$$\begin{aligned} \lim_{k \rightarrow +\infty} \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0} (\mathbf{X}_k = \mathbf{z} \mid \mathbf{X}_{k+n} \neq \mathbf{0}, T < +\infty) &= \lim_{k \rightarrow +\infty} Q_k(\mathbf{x}_0, \mathbf{z}) = \bar{\mu}(\mathbf{z}), \\ \lim_{n \rightarrow +\infty} \lim_{k \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0} (\mathbf{X}_k = \mathbf{z} \mid \mathbf{X}_{k+n} \neq \mathbf{0}, T < +\infty) &= \lim_{n \rightarrow +\infty} \bar{v}^{(n)}(\mathbf{z}) = \bar{\mu}(\mathbf{z}). \end{aligned}$$

We prove here that, under the stronger assumption (A<sub>3</sub>) if  $\rho < 1$ , this limiting result also holds when  $k$  and  $n$  simultaneously tend to infinity.

**Theorem 4.1.** *Let  $(\mathbf{X}_k)_{k \geq 0}$  be a noncritical multitype GW process satisfying (A<sub>1</sub>). If  $\rho < 1$ , we assume in addition (A<sub>3</sub>). Then for all  $\mathbf{x}_0 \in \mathbb{N}^d \setminus \{\mathbf{0}\}$  and  $\mathbf{z} \in \mathbb{N}^d$ ,*

$$\lim_{\substack{n \rightarrow +\infty \\ k \rightarrow +\infty}} \mathbb{P}_{\mathbf{x}_0} (\mathbf{X}_k = \mathbf{z} \mid \mathbf{X}_{k+n} \neq \mathbf{0}, T < +\infty) = \bar{\mu}(\mathbf{z}),$$

where  $\bar{\mu}$  is the size-biased Yaglom distribution of  $(\bar{\mathbf{X}}_k)_{k \geq 0}$ , the associated process with respect to  $\mathbf{q}$ .

*Remark 4.2.* This implies in particular that for any  $0 < t < 1$ ,

$$\lim_{k \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0} (\mathbf{X}_{\lfloor kt \rfloor} = \mathbf{z} \mid \mathbf{X}_k \neq \mathbf{0}, T < +\infty) = \bar{\mu}(\mathbf{z}).$$

*Remark 4.3.* In the critical case, the  $Q$ -process is transient and the obtained limit is degenerate. A suitable normalization in order to obtain a non-degenerate probability distribution is of the form  $\mathbf{X}_k/k$ . However, even with this normalization, the previous result does not hold in the critical case. Indeed, we know for instance that in the monotype case, a critical process with finite variance  $\sigma^2 > 0$  satisfies for each  $z \geq 0$  (Lamperti and Ney, 1968),

$$\begin{aligned} \lim_{k \rightarrow +\infty} \lim_{n \rightarrow +\infty} \mathbb{P}_1 \left( \frac{X_k}{k} \leq z \mid X_{k+n} \neq 0 \right) &= 1 - e^{-\frac{2z}{\sigma^2}}, \\ \lim_{n \rightarrow +\infty} \lim_{k \rightarrow +\infty} \mathbb{P}_1 \left( \frac{X_k}{k} \leq z \mid X_{k+n} \neq 0 \right) &= 1 - e^{-\frac{2z}{\sigma^2}} - \frac{2z}{\sigma^2} e^{-\frac{2z}{\sigma^2}}. \end{aligned}$$

*Proof:* Thanks to Remark 1.1 and to the fact that if  $\rho > 1$ ,  $\mathbb{E}_{\mathbf{e}_i}(\bar{X}_{1,j} \bar{X}_{1,l}) < +\infty$ , we can assume without loss of generality that  $\rho < 1$ . For each  $n, k \in \mathbb{N}$  and  $\mathbf{r} \in [0, 1]^d$ ,

$$\mathbb{E}_{\mathbf{x}_0} (\mathbf{r}^{\mathbf{X}_k} \mathbf{1}_{\mathbf{X}_{k+n}=\mathbf{0}}) = \sum_{\mathbf{y} \in \mathbb{N}^d} \mathbb{P}_{\mathbf{x}_0} (\mathbf{X}_k = \mathbf{y}) \mathbf{r}^{\mathbf{y}} \mathbb{P}_{\mathbf{y}} (\mathbf{X}_n = \mathbf{0}) = \mathbf{f}_k(\mathbf{r} \mathbf{f}_n(\mathbf{0}))^{\mathbf{x}_0},$$

which leads to

$$\begin{aligned} \mathbb{E}_{\mathbf{x}_0} [\mathbf{r}^{\mathbf{X}_k} \mid \mathbf{X}_{k+n} \neq \mathbf{0}] &= \frac{\mathbb{E}_{\mathbf{x}_0} (\mathbf{r}^{\mathbf{X}_k}) - \mathbb{E}_{\mathbf{x}_0} (\mathbf{r}^{\mathbf{X}_k} \mathbf{1}_{\mathbf{X}_{k+n}=\mathbf{0}})}{1 - \mathbb{P}_{\mathbf{x}_0} (\mathbf{X}_{k+n} = \mathbf{0})} \\ &= \frac{\mathbf{f}_k(\mathbf{r})^{\mathbf{x}_0} - \mathbf{f}_k(\mathbf{r} \mathbf{f}_n(\mathbf{0}))^{\mathbf{x}_0}}{1 - \mathbf{f}_{k+n}(\mathbf{0})^{\mathbf{x}_0}}. \end{aligned} \tag{4.1}$$

By Taylor's theorem,

$$\begin{aligned} \mathbf{f}_k(\mathbf{r})^{\mathbf{x}_0} - \mathbf{f}_k(\mathbf{r}_n(\mathbf{0}))^{\mathbf{x}_0} &= \sum_{i=1}^d \frac{\partial \mathbf{f}_k^{\mathbf{x}_0}(\mathbf{r})}{\partial r_i} r_i (1 - f_{n,i}(\mathbf{0})) \\ &- \sum_{i,j=1,\dots,d} \frac{r_i r_j (1 - f_{n,i}(\mathbf{0})) (1 - f_{n,j}(\mathbf{0}))}{2} \int_0^1 (1-t) \frac{\partial^2 \mathbf{f}_k^{\mathbf{x}_0}(\mathbf{r} - t\mathbf{r}(1 - \mathbf{f}_n(\mathbf{0})))}{\partial r_i \partial r_j} dt, \end{aligned} \quad (4.2)$$

with

$$\frac{\partial \mathbf{f}_k^{\mathbf{x}_0}(\mathbf{r})}{\partial r_i} = \sum_{j=1}^d x_{0,j} \frac{\partial f_{k,j}(\mathbf{r})}{\partial r_i} \mathbf{f}_k(\mathbf{r})^{\mathbf{x}_0 - \mathbf{e}_j}. \quad (4.3)$$

Let us first prove the existence of  $\lim_k \rho^{-k} \frac{\partial f_{k,j}(\mathbf{r})}{\partial r_i}$  for each  $i, j$  and  $\mathbf{r} \in [0, 1]^d$ . For each  $k, p \in \mathbb{N}$  and  $a > 0$ ,

$$\begin{aligned} &\left| \rho^{-k} \frac{\partial f_{k,j}(\mathbf{r})}{\partial r_i} - \rho^{-(k+p)} \frac{\partial f_{k+p,j}(\mathbf{r})}{\partial r_i} \right| \\ &\leq \sum_{\substack{\mathbf{z} \in \mathbb{N}^d \\ \|\mathbf{z}\|_2 < a}} z_i \mathbf{r}^{\mathbf{z} - \mathbf{e}_i} \left| \rho^{-k} P_k(\mathbf{e}_j, \mathbf{z}) - \rho^{-(k+p)} P_{k+p}(\mathbf{e}_j, \mathbf{z}) \right| \\ &+ \rho^{-k} \mathbb{E}_{\mathbf{e}_j} (X_{k,i} \mathbf{1}_{\|\mathbf{X}_k\|_2 \geq a}) + \rho^{-(k+p)} \mathbb{E}_{\mathbf{e}_j} (X_{k+p,i} \mathbf{1}_{\|\mathbf{X}_{k+p}\|_2 \geq a}). \end{aligned} \quad (4.4)$$

By Cauchy-Schwarz and Markov's inequalities,  $\mathbb{E}_{\mathbf{e}_j} (X_{k,i} \mathbf{1}_{\|\mathbf{X}_k\|_2 \geq a}) \leq \frac{1}{a} \mathbb{E}_{\mathbf{e}_j} (\|\mathbf{X}_k\|_2^2)$ . For each  $\mathbf{x} \in \mathbb{N}^d$ , let  $\mathbf{C}_{\mathbf{x},k}$  be the matrix  $(\mathbb{E}_{\mathbf{x}}(X_{k,i} X_{k,j}))_{1 \leq i, j \leq d}$ . According to Harris (1963),

$$\mathbf{C}_{\mathbf{x},k} = (\mathbf{M}^T)^k \mathbf{C}_{\mathbf{x},0} \mathbf{M}^k + \sum_{n=1}^k (\mathbf{M}^T)^{k-n} \left( \sum_{i=1}^d \Sigma^i \mathbb{E}_{\mathbf{x}}(X_{n-1,i}) \right) \mathbf{M}^{k-n}. \quad (4.5)$$

Thanks to (1.1) this implies the existence of some  $C > 0$  such that for all  $k \in \mathbb{N}$ ,  $\rho^{-k} \mathbb{E}_{\mathbf{e}_j} (\|\mathbf{X}_k\|_2^2) = \rho^{-k} \sum_{i=1}^d [\mathbf{C}_{\mathbf{e}_j, k}]_{ii} \leq C$ , and the two last right terms in (4.4) can be bounded by  $2Ca^{-1}$ . As for the first right term in (4.4), it is thanks to (2.3) and (3.6) as small as desired for  $k$  large enough. This proves that  $(\rho^{-k} \frac{\partial f_{k,j}(\mathbf{r})}{\partial r_i})_k$  is a Cauchy sequence.

Its limit is then necessarily, for each  $\mathbf{r} \in [0, 1]^d$ ,

$$\lim_{k \rightarrow +\infty} \rho^{-k} \frac{\partial f_{k,j}(\mathbf{r})}{\partial r_i} = \gamma u_j \frac{\partial g(\mathbf{r})}{\partial r_i}, \quad (4.6)$$

where  $g$  is defined in Subsection 2. Indeed, since assumption  $(A_3)$  ensures that  $(A_2)$  is satisfied, we can deduce from (2.1) and (2.3) that  $\lim_k \rho^{-k} (f_{k,j}(\mathbf{r}) - f_{k,j}(\mathbf{0})) = \gamma u_j g(\mathbf{r})$ . Hence, using the fact that  $0 \leq \rho^{-k} \frac{\partial f_{k,j}(\mathbf{r})}{\partial r_i} \leq \rho^{-k} m_{ji}^{(k)}$ , which thanks to (1.1) is bounded, we obtain by Lebesgue's dominated convergence theorem that for each  $h \in \mathbb{R}$  such that  $\mathbf{r} + h\mathbf{e}_i \in [0, 1]^d$ ,

$$\begin{aligned} &\gamma u_j g(\mathbf{r} + h\mathbf{e}_i) - \gamma u_j g(\mathbf{r}) \\ &= \lim_{k \rightarrow +\infty} \int_0^h \rho^{-k} \frac{\partial f_{k,j}(\mathbf{r} + t\mathbf{e}_i)}{\partial r_i} dt = \int_0^h \lim_{k \rightarrow +\infty} \rho^{-k} \frac{\partial f_{k,j}(\mathbf{r} + t\mathbf{e}_i)}{\partial r_i} dt, \end{aligned}$$

proving (4.6).



In view of (4.2), let us note that for each  $\mathbf{r} \in [0, 1]^d$ , there exists thanks to (1.1) and (4.5) some  $C > 0$  such that for each  $k \in \mathbb{N}$ ,  $0 \leq \rho^{-k} \frac{\partial^2 \mathbf{f}_k^{\mathbf{x}}(\mathbf{r})}{\partial r_i \partial r_j} \leq \rho^{-k} \mathbb{E}_{\mathbf{x}}[X_{k,j}(X_{k,i} - \delta_{ij})] \leq C$ , hence for each  $k, n \in \mathbb{N}$ ,

$$\rho^{-k} \int_0^1 (1-t) \frac{\partial^2 \mathbf{f}_k^{\mathbf{x}}(\mathbf{r} - t\mathbf{r}(1 - \mathbf{f}_n(\mathbf{0})))}{\partial r_i \partial r_j} dt \leq \frac{C}{2}.$$

Together with (2.3) this entails that the last right term in (4.2) satisfies

$$\lim_{\substack{n \rightarrow +\infty \\ k \rightarrow +\infty}} \rho^{-(k+n)} \sum_{i,j=1 \dots d} \frac{r_i r_j (1 - f_{n,i}(\mathbf{0})) (1 - f_{n,j}(\mathbf{0}))}{2} \int_0^1 \dots dt = 0.$$

Moreover, we deduce from (4.3), (4.6) and  $\lim_n \mathbf{f}_n(\mathbf{r}) = \mathbf{1}$  that the first right term in (4.2) satisfies

$$\lim_{\substack{n \rightarrow +\infty \\ k \rightarrow +\infty}} \rho^{-(k+n)} \sum_{i=1}^d \frac{\partial \mathbf{f}_k^{\mathbf{x}_0}(\mathbf{r})}{\partial r_i} r_i (1 - f_{n,i}(\mathbf{0})) = \gamma^2 \mathbf{x}_0 \cdot \mathbf{u} \sum_{i=1}^d r_i u_i \frac{\partial g(\mathbf{r})}{\partial r_i}.$$

Recalling (4.1) and (2.3), we have thus proven that for each  $\mathbf{r} \in [0, 1]^d$ ,

$$\lim_{\substack{n \rightarrow +\infty \\ k \rightarrow +\infty}} \mathbb{E}_{\mathbf{x}_0} [\mathbf{r}^{\mathbf{X}_k} \mid \mathbf{X}_{k+n} \neq \mathbf{0}] = \gamma \sum_{i=1}^d r_i u_i \frac{\partial g(\mathbf{r})}{\partial r_i} = \gamma \sum_{\mathbf{z} \in \mathbb{N}^d \setminus \{\mathbf{0}\}} \mathbf{z} \cdot \mathbf{u} \nu(\mathbf{z}) \mathbf{r}^{\mathbf{z}}.$$

Finally, (2.2) leads to  $\gamma \sum_{i=1}^d u_i \frac{\partial g(\mathbf{1})}{\partial r_i} = 1$ , and thus

$$\lim_{\substack{n \rightarrow +\infty \\ k \rightarrow +\infty}} \mathbb{E}_{\mathbf{x}_0} [\mathbf{r}^{\mathbf{X}_k} \mid \mathbf{X}_{k+n} \neq \mathbf{0}] = \frac{\sum_{\mathbf{z} \in \mathbb{N}^d \setminus \{\mathbf{0}\}} \mathbf{z} \cdot \mathbf{u} \nu(\mathbf{z}) \mathbf{r}^{\mathbf{z}}}{\sum_{\mathbf{y} \in \mathbb{N}^d \setminus \{\mathbf{0}\}} \mathbf{y} \cdot \mathbf{u} \nu(\mathbf{y})},$$

which by (2.5) is a probability generating function.  $\square$

## 5. Conditioning on the total progeny

Let  $\mathbf{N} = (N_1, \dots, N_d)$  denote the total progeny of the process  $(\mathbf{X}_k)_{k \geq 0}$ , where for each  $i = 1 \dots d$ ,  $N_i = \sum_{k=0}^{+\infty} X_{k,i}$ , and  $N_i = +\infty$  if the sum diverges. Our aim is to study the behavior of  $(\mathbf{X}_k)_{k \geq 0}$  conditioned on the event  $\{\mathbf{N} = \lfloor n\mathbf{w} \rfloor\}$ , as  $n$  tends to infinity, for some specific positive vector  $\mathbf{w}$ . We recall that in the critical case, the GW process suitably normalized and conditioned on non-extinction in the same fashion as in (2.1), converges to a limit law supported by the ray  $\{\lambda \mathbf{v} : \lambda \geq 0\} \subset \mathbb{R}_+^d$ . In this sense, its left eigenvector  $\mathbf{v}$  describes “typical limiting type proportions”, as pointed out in [Vatutin and Flyaishmann \(2000\)](#). As we will see in [Lemma 5.3](#), conditioning a GW process on a given total progeny size comes down to conditioning an associated critical process on the same total progeny size. For this reason, the vector  $\mathbf{w}$  will be chosen to be the left eigenvector of the associated critical process. It then appears that, similarly as in the monotype case ([Kennedy, 1975](#)), the process conditioned on an infinite total progeny  $\{\mathbf{N} = \lfloor n\mathbf{w} \rfloor\}$ ,  $n \rightarrow \infty$ , has the structure of the  $Q$ -process of a critical process, and is consequently transient. This is the main result, stated in [Theorem 5.1](#).

**Theorem 5.1.** *Let  $(\mathbf{X}_k)_{k \geq 0}$  be a multitype GW process satisfying  $(A_1)$ . We assume in addition that*

- (A<sub>4</sub>) *there exists  $\mathbf{a} > \mathbf{0}$  such that the associated process with respect to  $\mathbf{a}$  is critical,*

(A<sub>5</sub>) for each  $j = 1 \dots d$ , there exist  $i = 1 \dots d$  and  $\mathbf{k} \in \mathbb{N}^d$  such that  $p_i(\mathbf{k}) > 0$  and  $p_i(\mathbf{k} + \mathbf{e}_j) > 0$ ,

(A<sub>6</sub>) the associated process with respect to  $\mathbf{a}$  admits moments of order  $d + 1$ , and its covariance matrices are positive-definite.

Then for all  $\mathbf{x}_0 \in \mathbb{N}^d \setminus \{\mathbf{0}\}$ ,  $k_1 \leq \dots \leq k_j \in \mathbb{N}$ , and  $\mathbf{x}_1, \dots, \mathbf{x}_j \in \mathbb{N}^d$ ,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{N} = \lfloor n\bar{\mathbf{v}} \rfloor) \\ = \frac{\mathbf{x}_j \cdot \bar{\mathbf{u}}}{\mathbf{x}_0 \cdot \bar{\mathbf{u}}} \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j), \end{aligned} \quad (5.1)$$

where  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  is the associated process with respect to  $\mathbf{a}$ .

The limiting process defined by (5.1) is thus Markovian with transition probabilities

$$\bar{Q}_1(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{y} \cdot \bar{\mathbf{u}}}{\mathbf{x} \cdot \bar{\mathbf{u}}} \bar{P}_1(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{a}^{\mathbf{y}}}{\mathbf{f}(\mathbf{a})^{\mathbf{x}}} \frac{\mathbf{y} \cdot \bar{\mathbf{u}}}{\mathbf{x} \cdot \bar{\mathbf{u}}} P_1(\mathbf{x}, \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{N}^d \setminus \{\mathbf{0}\},$$

and corresponds to the  $Q$ -process associated with the critical process  $(\bar{\mathbf{X}}_k)_{k \geq 0}$ .

*Remark 5.2.*

- If  $d = 1$ , the conditional event  $\{\mathbf{N} = \lfloor n\bar{\mathbf{v}} \rfloor\}$  reduces to  $\{N = n\}$ , as studied in Kennedy (1975), in which assumptions (A<sub>4</sub>)–(A<sub>6</sub>) are also required<sup>1</sup>.
- If  $(\mathbf{X}_k)_{k \geq 0}$  is critical, assumption (A<sub>4</sub>) is satisfied with  $\mathbf{a} = \mathbf{1}$ . This assumption is also automatically satisfied if  $(\mathbf{X}_k)_{k \geq 0}$  is supercritical. Indeed, as mentioned in Remark 1.1, the associated process with respect to  $\mathbf{0} < \mathbf{q} < \mathbf{1}$  is subcritical and thus satisfies  $\bar{\rho} < 1$ . The fact that  $\rho > 1$  and the continuity of the Perron’s root as a function of the mean matrix coefficients then ensures the existence of some  $\mathbf{q} \leq \mathbf{a} \leq \mathbf{1}$  satisfying (A<sub>4</sub>). Note however that such an  $\mathbf{a}$  is not unique.
- For any  $\mathbf{a} > \mathbf{0}$ ,  $p_i$  and  $\bar{p}_i$  share by construction the same support. As a consequence,  $(\mathbf{X}_k)_{k \geq 0}$  satisfies (A<sub>5</sub>) if and only if  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  does. Moreover, a finite covariance matrix  $\Sigma^i$  is positive-definite if and only if there does not exist any  $c \in \mathbb{R}$  and  $\mathbf{x} \neq \mathbf{0}$  such that  $\mathbf{x} \cdot \mathbf{X} = c$   $\mathbb{P}_{\mathbf{e}_i}$ -almost-surely, hence if and only if  $\mathbf{x} \cdot \bar{\mathbf{X}} = c$   $\mathbb{P}_{\mathbf{e}_i}$ -almost-surely. Consequently, provided it exists,  $\Sigma^i$  is positive-definite if and only if  $\bar{\Sigma}^i$  is positive-definite as well.

We shall first show in Lemma 5.3 that for any  $\mathbf{a}$ , the associated process  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  with respect to  $\mathbf{a}$ , conditioned on  $\{\bar{\mathbf{N}} = \mathbf{n}\}$ , has the same probability distribution as the original process conditioned on  $\{\mathbf{N} = \mathbf{n}\}$ , for any  $\mathbf{n} \in \mathbb{N}^d$ . It is thus enough to prove Theorem 5.1 in the critical case, which is done at the end of the article.

It follows from Proposition 1 in Good (1975) or directly from Theorem 1.2 in Chaumont (2016) that the probability distribution of the total progeny in the multitype case is given for each  $\mathbf{x}_0, \mathbf{n} \in \mathbb{N}^d$  with  $\mathbf{n} > \mathbf{0}, \mathbf{n} \geq \mathbf{x}_0$  by

$$\mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \mathbf{n}) = \frac{1}{n_1 \dots n_d} \sum_{\substack{\mathbf{k}^1, \dots, \mathbf{k}^d \in \mathbb{N}^d \\ \mathbf{k}^1 + \dots + \mathbf{k}^d = \mathbf{n} - \mathbf{x}_0}} \det \begin{pmatrix} n_1 \mathbf{e}_1 - \mathbf{k}^1 & & \\ & \dots & \\ & & n_d \mathbf{e}_d - \mathbf{k}^d \end{pmatrix} \prod_{i=1}^d p_i^{*n_i}(\mathbf{k}^i). \quad (5.2)$$

<sup>1</sup>Since the author’s work, it has been proved in Abraham et al. (2015) that in the critical case and under (A<sub>1</sub>), Theorem 5.1 holds true under the minimal assumptions of aperiodicity of the offspring distribution (implied by (A<sub>5</sub>)) and the finiteness of its first order moment.

**Lemma 5.3.** *Let  $(\mathbf{X}_k)_{k \geq 0}$  be a multitype GW process. Then, for any  $\mathbf{a} > \mathbf{0}$ , the associated process  $(\bar{\mathbf{X}}_k)_{k \geq 0}$  with respect to  $\mathbf{a}$  satisfies for any  $\mathbf{x}_0 \in \mathbb{N}^d$ ,  $k_1 \leq \dots \leq k_j \in \mathbb{N}$ ,  $\mathbf{x}_1, \dots, \mathbf{x}_j \in \mathbb{N}^d$  and  $\mathbf{n} \in \mathbb{N}^d$ ,*

$$\mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{N} = \mathbf{n}) = \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j \mid \bar{\mathbf{N}} = \mathbf{n}).$$

*Proof:* From (1.2) and (5.2),  $\mathbb{P}_{\mathbf{k}}(\bar{\mathbf{N}} = \mathbf{n}) = \frac{\mathbf{a}^{\mathbf{n}-\mathbf{k}}}{\mathbf{f}(\mathbf{a})^{\mathbf{n}}} \mathbb{P}_{\mathbf{k}}(\mathbf{N} = \mathbf{n})$ . For all  $n \in \mathbb{N}$ , we denote by  $\mathbf{N}_n = \sum_{k=0}^n \mathbf{X}_k$  (resp.  $\bar{\mathbf{N}}_n = \sum_{k=0}^n \bar{\mathbf{X}}_k$ ) the total progeny up to generation  $n$  of  $(\mathbf{X}_k)_{k \geq 0}$  (resp.  $(\bar{\mathbf{X}}_k)_{k \geq 0}$ ). Then

$$\begin{aligned} \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_j} = \mathbf{x}_j, \bar{\mathbf{N}}_{k_j} = \mathbf{l}) &= \sum_{\substack{\mathbf{i}_1, \dots, \mathbf{i}_{k_j-1} \in \mathbb{N}^d \\ \mathbf{i}_1 + \dots + \mathbf{i}_{k_j-1} = \mathbf{l} - \mathbf{x}_0 - \mathbf{x}_j}} \bar{P}_1(\mathbf{x}_0, \mathbf{i}_1) \dots \bar{P}_1(\mathbf{i}_{k_j-1}, \mathbf{x}_j) \\ &= \sum_{\substack{\mathbf{i}_1, \dots, \mathbf{i}_{k_j-1} \in \mathbb{N}^d \\ \mathbf{i}_1 + \dots + \mathbf{i}_{k_j-1} = \mathbf{l} - \mathbf{x}_0 - \mathbf{x}_j}} \frac{\mathbf{a}^{\mathbf{i}_1} P_1(\mathbf{x}_0, \mathbf{i}_1)}{\mathbf{f}(\mathbf{a})^{\mathbf{x}_0}} \dots \frac{\mathbf{a}^{\mathbf{x}_j} P_1(\mathbf{i}_{k_j-1}, \mathbf{x}_j)}{\mathbf{f}(\mathbf{a})^{\mathbf{i}_{k_j-1}}} \\ &= \frac{\mathbf{a}^{\mathbf{l} - \mathbf{x}_0} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_j} = \mathbf{x}_j, \mathbf{N}_{k_j} = \mathbf{l})}{\mathbf{f}(\mathbf{a})^{\mathbf{l} - \mathbf{x}_j}}, \end{aligned}$$

and similarly

$$\begin{aligned} \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j, \bar{\mathbf{N}}_{k_j} = \mathbf{l}) \\ = \frac{\mathbf{a}^{\mathbf{l} - \mathbf{x}_0} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j, \mathbf{N}_{k_j} = \mathbf{l})}{\mathbf{f}(\mathbf{a})^{\mathbf{l} - \mathbf{x}_j}}. \end{aligned}$$

Consequently, thanks to the Markov property,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j \mid \bar{\mathbf{N}} = \mathbf{n}) \\ = \sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ \mathbf{l} \leq \mathbf{n}}} \frac{\mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{X}}_{k_1} = \mathbf{x}_1, \dots, \bar{\mathbf{X}}_{k_j} = \mathbf{x}_j, \bar{\mathbf{N}}_{k_j} = \mathbf{l}) \mathbb{P}_{\mathbf{x}_j}(\bar{\mathbf{N}} = \mathbf{n} - \mathbf{l} + \mathbf{x}_j)}{\mathbb{P}_{\mathbf{x}_0}(\bar{\mathbf{N}} = \mathbf{n})} \\ = \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{N} = \mathbf{n}). \end{aligned}$$

□

Thanks to Lemma 5.3, it suffices to prove Theorem 5.1 in the critical case. For this purpose, we prove the following convergence result for the total progeny of a critical GW process.

**Proposition 5.4.** *Let  $(\mathbf{X}_k)_{k \geq 0}$  be a critical multitype GW process satisfying  $(A_1)$ ,  $(A_5)$  and  $(A_6)$ . Then there exists  $C > 0$  such that for all  $\mathbf{x}_0 \in \mathbb{N}^d$ ,*

$$\lim_{n \rightarrow +\infty} n^{\frac{d}{2}+1} \mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor) = C \mathbf{x}_0 \cdot \mathbf{u}. \quad (5.3)$$

*Proof:* From (5.2), for each  $n \geq \max_i v_i^{-1}$ ,  $n \geq \max_i x_{0,i} v_i^{-1}$ ,

$$\begin{aligned} \mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor) &= \frac{1}{\prod_{i=1}^d \lfloor nv_i \rfloor} \mathbb{E} \left[ \det \begin{pmatrix} \lfloor nv_1 \rfloor \mathbf{e}_1 - \mathbf{S}_{\lfloor nv_1 \rfloor}^1 \\ \cdots \\ \lfloor nv_d \rfloor \mathbf{e}_d - \mathbf{S}_{\lfloor nv_d \rfloor}^d \end{pmatrix} \mathbf{1}_{\sum_{i=1}^d \mathbf{S}_{\lfloor nv_i \rfloor}^i = \lfloor n\mathbf{v} \rfloor - \mathbf{x}_0} \right] \\ &= \frac{1}{\lfloor nv_d \rfloor} \mathbb{E} \left[ \det \begin{pmatrix} \mathbf{e}_1 - \mathbf{S}_{\lfloor nv_1 \rfloor}^1 / \lfloor nv_1 \rfloor \\ \cdots \\ \mathbf{e}_{d-1} - \mathbf{S}_{\lfloor nv_{d-1} \rfloor}^{d-1} / \lfloor nv_{d-1} \rfloor \\ \mathbf{x}_0 \end{pmatrix} \mathbf{1}_{\sum_{i=1}^d \mathbf{S}_{\lfloor nv_i \rfloor}^i = \lfloor n\mathbf{v} \rfloor - \mathbf{x}_0} \right], \end{aligned}$$

where the family  $(\mathbf{S}_{\lfloor nv_i \rfloor}^i)_{i=1 \dots d}$  is independent and is such that for each  $i$ ,  $\mathbf{S}_{\lfloor nv_i \rfloor}^i$  denotes the sum of  $\lfloor nv_i \rfloor$  independent and identically distributed random variables with probability distribution  $p_i$ .

Let us consider the event  $A_n = \{ \sum_{i=1}^d \mathbf{S}_{\lfloor nv_i \rfloor}^i = \lfloor n\mathbf{v} \rfloor - \mathbf{x}_0 \}$ . We define the covariance matrix  $\Sigma = \sum_{i=1}^d v_i \Sigma^i$ , which since  $\mathbf{v} > \mathbf{0}$  is positive-definite under  $(A_6)$ . Theorem 1.1 in Bentkus (2004) for nonidentically distributed independent variables ensures that  $\sum_{i=1}^d (\mathbf{S}_{\lfloor nv_i \rfloor}^i - \lfloor nv_i \rfloor \mathbf{m}^i) n^{-\frac{1}{2}}$  converges in distribution as  $n \rightarrow +\infty$  to the multivariate normal distribution  $\mathcal{N}_d(\mathbf{0}, \Sigma)$  with density  $\phi$ . Under  $(A_5)$  we have

$$\limsup_n \frac{n}{\min_{j=1 \dots d} \sum_{i=1}^d \frac{n_i}{d} \sum_{\mathbf{k} \in \mathbb{N}^d} \min(p_i(\mathbf{k}), p_i(\mathbf{k} + \mathbf{e}_j))} < +\infty,$$

which by Theorem 2.1 in Davis and McDonald (1995) ensures the following local limit theorem for nonidentically distributed independent variables:

$$\lim_{n \rightarrow \infty} \sup_{\mathbf{k} \in \mathbb{N}^d} \left| n^{\frac{d}{2}} \mathbb{P} \left( \sum_{i=1}^d \mathbf{S}_{\lfloor nv_i \rfloor}^i = \mathbf{k} \right) - \phi \left( \frac{\mathbf{k} - \sum_{i=1}^d \lfloor nv_i \rfloor \mathbf{m}^i}{\sqrt{n}} \right) \right| = 0. \quad (5.4)$$

In the critical case, the left eigenvector  $\mathbf{v}$  satisfies for each  $j$ ,  $v_j = \sum_{i=1}^d v_i m_{ij}$ , hence  $0 \leq |\lfloor nv_j \rfloor - \sum_{i=1}^d \lfloor nv_i \rfloor m_{ij}| < \max(1, \sum_{i=1}^d m_{ij})$  and (5.4) implies in particular that

$$\lim_{n \rightarrow +\infty} n^{\frac{d}{2}} \mathbb{P}(A_n) = \phi(\mathbf{0}) = \frac{1}{(2\pi)^{\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}}}. \quad (5.5)$$

Now, denoting by  $\mathfrak{S}_d$  the symmetric group of order  $d$  and by  $\epsilon(\sigma)$  the signature of a permutation  $\sigma \in \mathfrak{S}_d$ , we obtain by Leibniz formula that

$$\begin{aligned} \lfloor nv_d \rfloor \mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor) &= \sum_{\sigma \in \mathfrak{S}_d} \epsilon(\sigma) x_{0,\sigma(d)} \mathbb{E} \left[ \prod_{i=1}^{d-1} \left( \delta_{i,\sigma(i)} - \frac{S_{\lfloor nv_i \rfloor, \sigma(i)}^i}{\lfloor nv_i \rfloor} \right) \right] \\ &= \sum_{I \subset \{1, \dots, d-1\}} \sum_{\sigma \in \mathfrak{S}_d} \epsilon(\sigma) x_{0,\sigma(d)} \mathbb{E} \left[ \prod_{i \in I} \left( -\frac{S_{\lfloor nv_i \rfloor, \sigma(i)}^i}{\lfloor nv_i \rfloor} + m_{i,\sigma(i)} \right) \mathbf{1}_{A_n} \right] \prod_{i \notin I} (\delta_{i,\sigma(i)} - m_{i,\sigma(i)}). \end{aligned} \quad (5.6)$$

Let  $\varepsilon > 0$ . Since on the event  $A_n$  each  $S_{\lfloor nv_i \rfloor, j}^i / \lfloor nv_i \rfloor$  is bounded, there exists some constant  $A > 0$  such that for each  $i, j = 1 \dots d$ ,

$$\begin{aligned} \mathbb{E} \left( \left| \frac{S_{\lfloor nv_i \rfloor, j}^i}{\lfloor nv_i \rfloor} - m_{i,j} \right| \mathbf{1}_{A_n} \right) &\leq \varepsilon \mathbb{P}(A_n) + \frac{A}{\varepsilon^{d+1}} \mathbb{E} \left( \left| \frac{S_{\lfloor nv_i \rfloor, j}^i}{\lfloor nv_i \rfloor} - m_{i,j} \right|^{d+1} \right) \\ &\leq \varepsilon \mathbb{P}(A_n) + \frac{AB}{\varepsilon^{d+1} \lfloor nv_i \rfloor^{\frac{d+1}{2}}} \mathbb{E} \left( |S_{1,j}^i - m_{i,j}|^{d+1} \right), \end{aligned}$$

for some constant  $B > 0$ . The second inequality on the  $d + 1$ -th central moment can be found for instance in [Dharmadhikari and Jogdeo \(1969\)](#), Theorem 2. From (5.5) it thus appears that for each non-empty subset  $I \subset \{1, \dots, d - 1\}$ ,

$$\lim_{n \rightarrow +\infty} n^{\frac{d}{2}} \sum_{\sigma \in \mathfrak{S}_d} \varepsilon(\sigma) x_{0, \sigma(d)} \mathbb{E} \left[ \prod_{i \in I} \left( -\frac{S_{\lfloor nv_i \rfloor, \sigma(i)}^i}{\lfloor nv_i \rfloor} + m_{i, \sigma(i)} \right) \mathbf{1}_{A_n} \right] \prod_{i \notin I} (\delta_{i, \sigma(i)} - m_{i, \sigma(i)}) = 0.$$

Consequently, considering the remaining term in (5.6) corresponding to  $I = \emptyset$ , we obtain that

$$\begin{aligned} &\lim_{n \rightarrow +\infty} n^{\frac{d}{2}+1} \mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor) \\ &= \lim_{n \rightarrow +\infty} n^{\frac{d}{2}} \mathbb{P}(A_n) \frac{1}{v_d} \sum_{\sigma \in \mathfrak{S}_d} \varepsilon(\sigma) x_{0, \sigma(d)} \prod_{i=1}^{d-1} (\delta_{i, \sigma(i)} - m_{i, \sigma(i)}) \\ &= \frac{1}{v_d (2\pi)^{\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}}} \det \begin{pmatrix} \mathbf{e}_1 - \mathbf{m}^1 \\ \dots \\ \mathbf{e}_{d-1} - \mathbf{m}^{d-1} \\ \mathbf{x}_0 \end{pmatrix} = \frac{\mathbf{x}_0 \cdot \mathbf{D}}{v_d (2\pi)^{\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}}}, \end{aligned}$$

where  $\mathbf{D} = (D_1, \dots, D_d)$  is such that  $D_i$  is the  $(d, i)$ -th cofactor of the matrix  $\mathbf{I} - \mathbf{M}$ . The criticality of  $(\mathbf{X}_k)_{k \geq 0}$  implies that  $\det(\mathbf{I} - \mathbf{M}) = (\mathbf{e}_d - \mathbf{m}^d) \cdot \mathbf{D} = 0$ . Moreover, for each  $j = 1 \dots d - 1$ ,  $(\mathbf{e}_j - \mathbf{m}^j) \cdot \mathbf{D}$  corresponds to the determinant of  $\mathbf{I} - \mathbf{M}$  in which the  $d$ -th row has been replaced by the  $j$ -th row, and is consequently null. We have thus proven that for each  $j = 1 \dots d$ ,  $(\mathbf{e}_j - \mathbf{m}^j) \cdot \mathbf{D} = 0$ , or equivalently that  $\sum_{i=1}^d m_{ji} D_i = D_j$ . Hence  $\mathbf{D}$  is a right eigenvector of  $\mathbf{M}$  for the Perron's root 1, which implies the existence of some nonnull constant  $c$  such that  $\mathbf{D} = c\mathbf{u}$ , leading to the desired result.  $\square$

*Proof of Theorem 5.1* Let us assume that  $(\mathbf{X}_k)_{k \geq 0}$  is critical and satisfies  $(A_1)$ ,  $(A_5)$  and  $(A_6)$ . Let  $\mathbf{x}_0 \in \mathbb{N}^d$ ,  $k_1 \leq \dots \leq k_j \in \mathbb{N}$ , and  $\mathbf{x}_1, \dots, \mathbf{x}_j \in \mathbb{N}^d$  and let us show that

$$\begin{aligned} &\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{N} = \lfloor n\mathbf{v} \rfloor) \\ &= \frac{\mathbf{x}_j \cdot \mathbf{u}}{\mathbf{x}_0 \cdot \mathbf{u}} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j). \quad (5.7) \end{aligned}$$

Let  $\frac{3}{4} < \varepsilon < 1$ . The Markov property entails that

$$\begin{aligned} & \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j \mid \mathbf{N} = \lfloor n\mathbf{v} \rfloor) \\ &= \sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ \mathbf{l} < \lfloor n^\varepsilon \mathbf{v} \rfloor}} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j, \mathbf{N}_{k_j} = \mathbf{l}) \frac{\mathbb{P}_{\mathbf{x}_j}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor - \mathbf{l} + \mathbf{x}_j)}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor)} \\ &+ \sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ \lfloor n^\varepsilon \mathbf{v} \rfloor \leq \mathbf{l} \leq \lfloor n\mathbf{v} \rfloor}} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j, \mathbf{N}_{k_j} = \mathbf{l}) \frac{\mathbb{P}_{\mathbf{x}_j}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor - \mathbf{l} + \mathbf{x}_j)}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor)}. \end{aligned} \quad (5.8)$$

Note that (5.4) ensures that

$$\lim_n n^{\frac{d}{2}} \mathbb{P} \left( \sum_{i=1}^d \mathbf{S}_{\lfloor n v_i \rfloor - l_i + x_{j,i}}^i = \lfloor n\mathbf{v} \rfloor - \mathbf{l} \right) = \frac{1}{(2\pi)^{\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}}},$$

uniformly in  $\mathbf{l} < \lfloor n^\varepsilon \mathbf{v} \rfloor$ , and that the proof of Proposition 5.4 can be used to show that

$$\lim_{n \rightarrow +\infty} n^{\frac{d}{2}+1} \mathbb{P}_{\mathbf{x}_j}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor - \mathbf{l} + \mathbf{x}_j) = \frac{C \mathbf{x}_j \cdot \mathbf{u}}{v_d (2\pi)^{\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}}},$$

uniformly in  $\mathbf{l} < \lfloor n^\varepsilon \mathbf{v} \rfloor$ . Together with Proposition 5.4, this shows that the first sum in (5.8) converges to

$$\begin{aligned} & \frac{\mathbf{x}_j \cdot \mathbf{u}}{\mathbf{x}_0 \cdot \mathbf{u}} \sum_{\mathbf{l} \in \mathbb{N}^d} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j, \mathbf{N}_{k_j} = \mathbf{l}) \\ &= \frac{\mathbf{x}_j \cdot \mathbf{u}}{\mathbf{x}_0 \cdot \mathbf{u}} \mathbb{P}_{\mathbf{x}_0}(\mathbf{X}_{k_1} = \mathbf{x}_1, \dots, \mathbf{X}_{k_j} = \mathbf{x}_j) \end{aligned} \quad (5.9)$$

as  $n \rightarrow +\infty$ . The second sum in (5.8) can be bounded by

$$\begin{aligned} \frac{\mathbb{P}_{\mathbf{x}_j}(\mathbf{N}_{k_j} \geq \lfloor n^\varepsilon \mathbf{v} \rfloor)}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor)} &\leq \frac{\mathbb{P}_{\mathbf{x}_j}(\|\mathbf{N}_{k_j}\|_1^{d+1} \geq n^{(d+1)\varepsilon} \|\mathbf{v}\|_1^{d+1})}{\mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor)} \\ &\leq \frac{\mathbb{E}_{\mathbf{x}_j}(\|\mathbf{N}_{k_j}\|_1^{d+1})}{\|\mathbf{v}\|_1^{d+1} n^{(d+1)\varepsilon} \mathbb{P}_{\mathbf{x}_0}(\mathbf{N} = \lfloor n\mathbf{v} \rfloor)}. \end{aligned}$$

Thanks to (A<sub>5</sub>), the moments of order  $d+1$  of the finite sum  $\mathbf{N}_{k_j}$  are finite, and since  $(d+1)\varepsilon > \frac{d}{2} + 1$ , the right term of the last inequality converges to 0 as  $n \rightarrow +\infty$  thanks to Proposition 5.4. This together with (5.9) in (5.8) finally proves (5.7).

## References

- R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electron. J. Probab.* **19**, no. 2, 19 (2014). [MR3164755](#).
- R. Abraham, J.-F. Delmas and H. Guo. Critical multi-type Galton-Watson processes conditioned to be large. *ArXiv Mathematics e-prints* (2015). [arXiv:1511.01721](#).
- K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg (1972). Die Grundlehren der mathematischen Wissenschaften, Band 196. [MR0373040](#).

- V. Bentkus. A Lyapunov type bound in  $\mathbf{R}^d$ . *Teor. Veroyatn. Primen.* **49** (2), 400–410 (2004). [MR2144310](#).
- L. Chaumont. Coding multitype forests: Application to the law of the total population of branching forests. *Trans. Amer. Math. Soc.* **368** (4), 2723–2747 (2016). [MR3449255](#).
- B. Davis and D. McDonald. An elementary proof of the local central limit theorem. *J. Theoret. Probab.* **8** (3), 693–701 (1995). [MR1340834](#).
- S. W. Dharmadhikari and K. Jogdeo. Bounds on moments of certain random variables. *Ann. Math. Statist.* **40**, 1506–1509 (1969). [MR0243587](#).
- I. J. Good. The Lagrange distributions and branching processes. *SIAM J. Appl. Math.* **28**, 270–275 (1975). [MR0359039](#).
- T. E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J. (1963). [MR0163361](#).
- P. Jagers and A. N. Lagerås. General branching processes conditioned on extinction are still branching processes. *Electron. Commun. Probab.* **13**, 540–547 (2008). [MR2453547](#).
- S. Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.* **9**, 103–252 (2012). [MR2908619](#).
- A. Joffe and F. Spitzer. On multitype branching processes with  $\rho \leq 1$ . *J. Math. Anal. Appl.* **19**, 409–430 (1967). [MR0212895](#).
- S. Karlin. *A first course in stochastic processes*. Academic Press, New York-London (1966). [MR0208657](#).
- D. P. Kennedy. The Galton-Watson process conditioned on the total progeny. *J. Appl. Probability* **12** (4), 800–806 (1975). [MR0386042](#).
- J. Lamperti and P. Ney. Conditioned branching processes and their limiting diffusions. *Teor. Veroyatnost. i Primenen.* **13**, 126–137 (1968). [MR0228073](#).
- T. Nakagawa. The  $Q$ -process associated with a multitype Galton-Watson process and the additional results. *Bull. Gen. Ed. Dokkyo Univ. School Medicine* **1**, 21–32 (1978). [MR653216](#).
- V. A. Vatutin and K. Flyaishmann. Deviations from typical proportions in multitype critical Galton-Watson branching processes. *Teor. Veroyatnost. i Primenen.* **45** (1), 30–51 (2000). [MR1810973](#).