



HAL
open science

On mutations in the branching model for multitype populations

Loïc Chaumont, Thi Ngoc Anh Nguyen

► **To cite this version:**

Loïc Chaumont, Thi Ngoc Anh Nguyen. On mutations in the branching model for multitype populations. 2015. <hal-01211198>

HAL Id: hal-01211198

<https://hal.science/hal-01211198v1>

Preprint submitted on 3 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

ON MUTATIONS IN THE BRANCHING MODEL FOR MULTITYPE POPULATIONS

LOÏC CHAUMONT AND THI NGOC ANH NGUYEN

ABSTRACT. The forest of mutations associated to a multitype branching forest is obtained by merging together all vertices of its clusters and by preserving connections between them. We first show that the forest of mutations of any multitype branching forest is itself a branching forest. Then we give its progeny distribution and describe some of its crucial properties in terms the initial progeny distribution. We also obtain the limiting behaviour of the number of mutations both when the total number of individuals tends to infinity and when the number of roots tends to infinity. The continuous time case is then investigated by considering multitype branching forests with edge lengths. When mutations are non reversible, we give a representation of their emergence times which allows us to describe the asymptotic behaviour of the latters, when the ratios of successive mutation rates tend to 0.

1. INTRODUCTION

The homogeneous multitype branching hypothesis provides a relevant model of population growth in the absence of any competitive or environmental constraint. In particular, it is widely used in population genetics, when studying successive mutations whose accumulation leads to the development of cancer. Then determining the statistics of the emergence times of mutations, or evaluating the distribution of the population size of mutant cells at any time become important challenges. In the extensive literature on the subject, let us simply cite [12], [11], [9], and [8].

This work is concerned with the mathematical study of mutations in multitype branching frameworks. We first focus on the problem of the total number of mutations under very general assumptions. This number is not a functional of the associated branching process and its study requires the complete knowledge of the multitype branching structure, that is the underlying plane forest. Then we show that the forest of mutations associated to any multitype forest, is itself a multitype branching forest whose progeny distribution can be explicitly computed. This result allows us to investigate the asymptotic behaviour of the number of mutations, when either the total population or the initial number of individuals tend to infinity.

When time is continuous, we are mainly interested in emergence times of new mutations in the non reversible case. The characterisation of these times requires a good knowledge of the corresponding multitype branching process and the main tool in this study consists in a recent extension of the Lamperti representation in higher dimensions. Emergence times are then expressed in terms of the underlying multivariate compound Poisson process, which allows us to obtain some accurate approximations of their law.

We start with some preliminaries on the coding of multitype branching forests by multivariate random walks in Section 2.1. Then we state and prove our results on the

Date: October 3, 2015.

2010 Mathematics Subject Classification. 60J80.

Key words and phrases. Multitype branching forest, mutation, forest of mutations, emergence time.

total mutations sizes of branching forests in Sections 2.2 and 3.2. Results bearing on emergence times are presented in Section 3.3.

2. MUTATIONS AND THEIR ASYMPTOTICS IN DISCRETE MULTITYPE FORESTS

2.1. Preliminaries on discrete multitype forests. In all this work, we use the notation $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ and for any positive integer d , we set $[d] = \{1, \dots, d\}$. We will denote by e_i is the i -th unit vector of \mathbb{Z}_+^d . We define the following partial order on \mathbb{R}^d by setting $x = (x_1, \dots, x_d) \geq y = (y_1, \dots, y_d)$, if $x_i \geq y_i$, for all $i \in [d]$. The convention $\inf \emptyset = +\infty$ will be valid all along this paper. Then (Ω, \mathcal{F}, P) is a reference probability space on which all the stochastic processes involved in this work are defined.

Let us first recall the coding of multitype forests, as it has been defined in [7]. A (plane) forest \mathbf{f} is a directed planar graph with no loops on a possibly infinite and non empty set of vertices $\mathbf{v} = \mathbf{v}(\mathbf{f})$, such that each vertex has a finite inner degree and an outer degree equals to 0 or 1. The connected components of a forest are called the *trees*. In a tree \mathbf{t} , the only vertex with outer degree equal to 0 is called the *root* of \mathbf{t} . The roots of the connected components of a forest \mathbf{f} are called the roots of \mathbf{f} . For two vertices u and v of a forest \mathbf{f} , if (u, v) is a directed edge of \mathbf{f} , then we say that u is a *child* of v , or that v is the *parent* of u . We first give an order to the trees of the forest \mathbf{f} and denote them by $\mathbf{t}_1(\mathbf{f}), \mathbf{t}_2(\mathbf{f}), \dots, \mathbf{t}_k(\mathbf{f}), \dots$ (we will usually write $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_k, \dots$ if no confusion is possible). Then we rank (a part of) the vertices of \mathbf{f} according to the breadth first search order, by ranking first the vertices of \mathbf{t}_1 , then the vertices of \mathbf{t}_2 , and so on, see the labeling of the two forests in Figure 2. Note that if \mathbf{t}_k , for $k \geq 1$ is the first infinite tree, then the vertices of \mathbf{t}_{k+1}, \dots have no label according to this procedure.

To each forest \mathbf{f} , we associate the application $c_{\mathbf{f}} : \mathbf{v}(\mathbf{f}) \rightarrow [d]$ such that if $u_i, u_{i+1}, \dots, u_{i+j} \in \mathbf{v}(\mathbf{f})$ have the same parent and are placed from left to right, then $c_{\mathbf{f}}(u_i) \leq c_{\mathbf{f}}(u_{i+1}) \leq \dots \leq c_{\mathbf{f}}(u_{i+j})$. For $v \in \mathbf{v}(\mathbf{f})$, the integer $c_{\mathbf{f}}(v)$ is called the *type* (or the *color*) of v . The couple $(\mathbf{f}, c_{\mathbf{f}})$ is called a d -*type forest*. When no confusion is possible, we will simply write \mathbf{f} . The set of d -type forests will be denoted by \mathcal{F}_d .

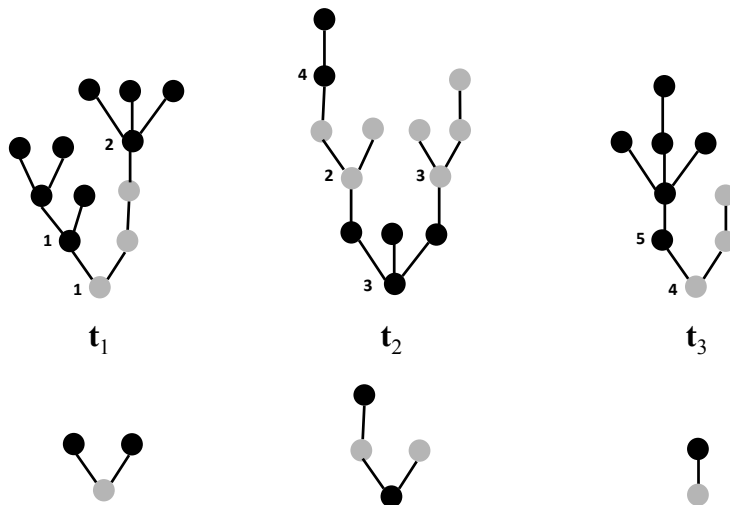


FIGURE 1. On the top, a discrete 2-type forest. Roots of clusters are ranked in the breadth first search order of the forest. The rank is written on the left of these roots. Below, the corresponding forest of mutations.

A *cluster* or a *subtree of type* $i \in [d]$ of a d -type forest $(\mathbf{f}, c_{\mathbf{f}}) \in \mathcal{F}_d$ is a maximal connected subgraph of $(\mathbf{f}, c_{\mathbf{f}})$ whose all vertices are of type i . Formally, \mathbf{t} is a cluster of type i of $(\mathbf{f}, c_{\mathbf{f}})$, if it is a connected subgraph whose all vertices are of type i and such that either the root of \mathbf{t} has no parent or the type of its parent is different from i . Moreover, if the parent of a vertex $v \in \mathbf{v}(\mathbf{t})^c$ belongs to $\mathbf{v}(\mathbf{t})$, then $c_{\mathbf{f}}(v) \neq i$. Clusters of type i in \mathbf{t}_1 are ranked according to the order of their roots in the breadth first search order of \mathbf{t}_1 , see Figures 1 and 2. Then if the number of clusters of type i is finite in \mathbf{t}_1 , we continue by ranking clusters of type i in \mathbf{t}_2 , and so on. Note that with this procedure, it is possible that clusters of $\mathbf{t}_k, \mathbf{t}_{k+1}, \dots$, for some k , are not ranked. We denote by $\mathbf{t}_1^{(i)}, \mathbf{t}_2^{(i)}, \dots, \mathbf{t}_k^{(i)}, \dots$ the sequence of clusters of type i in $(\mathbf{f}, c_{\mathbf{f}})$. The forest $\mathbf{f}^{(i)} := \{\mathbf{t}_1^{(i)}, \mathbf{t}_2^{(i)}, \dots, \mathbf{t}_k^{(i)}, \dots\}$ is called *the subforest of type* i of $(\mathbf{f}, c_{\mathbf{f}})$. We denote by $u_1^{(i)}, u_2^{(i)}, \dots$ the elements of $\mathbf{v}(\mathbf{f}^{(i)})$, ranked in the breadth first search order of $\mathbf{f}^{(i)}$. The subforests of the 2-type forest given in Figure 1 are represented in Figure 2.

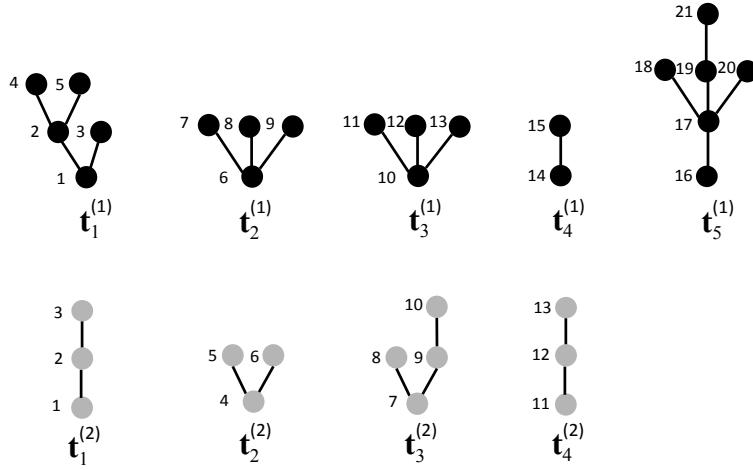


FIGURE 2. The subforests of the 2-type forest given in Figure 1 with their depth first search labeling.

To any forest $(\mathbf{f}, c_{\mathbf{f}}) \in \mathcal{F}_d$, we associate the *forest of mutations*, denoted by $(\bar{\mathbf{f}}, c_{\bar{\mathbf{f}}}) \in \mathcal{F}_d$, which is the forest of \mathcal{F}_d obtained by aggregating all the vertices of each subtree of $(\mathbf{f}, c_{\mathbf{f}})$ with a given type, in a single vertex with the same type, and preserving an edge between each pair of connected subtrees. An example is given in Figure 1.

For a forest $(\mathbf{f}, c_{\mathbf{f}}) \in \mathcal{F}_d$ and $u \in \mathbf{v}(\mathbf{f})$, when no confusion is possible, we denote by $p_i(u)$ the number of children of type i of u . For each $i \in [d]$, let $n_i \in \mathbb{Z}_+ \cup \{\infty\}$ be the number of vertices in the subforest $\mathbf{f}^{(i)}$ of $(\mathbf{f}, c_{\mathbf{f}})$. Then let us define the d -dimensional chain $x^{(i)} = (x^{i,1}, \dots, x^{i,d})$, with length n_i and whose values belong to the set \mathbb{Z}^d , by $x_0^{(i)} = 0$ and if $n_i \geq 1$,

$$(2.1) \quad x_{n+1}^{i,j} - x_n^{i,j} = p_j(u_{n+1}^{(i)}), \quad \text{if } i \neq j \quad \text{and} \quad x_{n+1}^{i,i} - x_n^{i,i} = p_i(u_{n+1}^{(i)}) - 1, \quad 0 \leq n \leq n_i - 1,$$

where $(u_n^{(i)})_{n \geq 1}$ is the labeling of the subforest $\mathbf{f}^{(i)}$ in its own breadth first search order. Note that the chains $(x_n^{i,j})$, for $i \neq j$ are nondecreasing whereas $(x_n^{i,i})$ is a downward

skip free chain, i.e. $x_{n+1}^{i,i} - x_n^{i,i} \geq -1$, for $0 \leq n \leq n_i - 1$. Besides, if n_i is finite, then $n_i = \min\{n : x_n^{i,i} = \min_{0 \leq k \leq n_i} x_k^{i,i}\}$. Let us also mention that from Theorem 2.7 of [7], when trees of $(\mathbf{f}, c_{\mathbf{f}})$ are finite, the data of the chains $x^{(1)}, \dots, x^{(d)}$ together with the sequence of ranked roots of $(\mathbf{f}, c_{\mathbf{f}})$, allow us to reconstruct this forest.

Let us now apply this coding to multitype branching forests. Let $\nu := (\nu_1, \dots, \nu_d)$, where ν_i is some distribution on \mathbb{Z}_+^d . We consider a branching process with progeny distribution ν , that is a population of individuals which reproduce independently of each other at each generation. Individuals of type i give birth to n_j children of type $j \in [d]$ with probability $\nu_i(n_1, \dots, n_d)$. For $i, j \in [d]$, we denote by m_{ij} the mean number of children of type j , given by an individual of type i , i.e.

$$m_{ij} = \sum_{(n_1, \dots, n_d) \in \mathbb{Z}_+^d} n_j \nu_i(n_1, \dots, n_d).$$

We say that ν is non singular if there is $i \in [d]$ such that $\nu_i(n : n_1 + \dots + n_d = 1) < 1$. The matrix $M = (m_{ij})$ is said to be irreducible if for all i, j , $m_{ij} < \infty$ and there exists $n \geq 1$ such that $m_{ij}^{(n)} > 0$, where $m_{ij}^{(n)}$ is the ij entry of the matrix M^n . If moreover the power n does not depend on (i, j) , then M is said to be primitive. In the latter case, according to Perron-Frobenius theory, the spectral radius ρ of M is the unique eigenvalue which is positive, simple and with maximal modulus. If $\rho \leq 1$, then the population will become extinct almost surely, whereas if $\rho > 1$, then with positive probability, the population will never become extinct. We say that ν is subcritical if $\rho < 1$, critical if $\rho = 1$ and supercritical if $\rho > 1$. We sometimes say that μ is irreducible, primitive, (sub)critical or supercritical, when this is the case for M .

By multitype branching forest with progeny distribution ν , we mean a sequence with a finite (deterministic) or infinite number of independent multitype branching trees with progeny distribution ν . A multitype branching forest will be considered as a random variable defined on the probability space (Ω, \mathcal{F}, P) and with values in \mathcal{F}_d . To any multitype branching forest F , we associate the random sequences $X = \{X^{(i)}, i \in [d]\}$, where $X^{(i)} = \{(X_n^{i,1}, \dots, X_n^{i,d}), 0 \leq n \leq n_i\}$, which are constructed as in (2.1). It has been proved in [7], Theorem 3.1 that if F is a primitive and (sub)critical branching forest with a finite number of trees, then $X^{(i)}, i \in [d]$ are independent random walks whose step distribution $\tilde{\nu}_i$ is defined by

$$(2.2) \quad \tilde{\nu}_i(k_1, \dots, k_d) := \nu_i(k_1, \dots, k_{i-1}, k_i + 1, k_{i+1}, \dots, k_d), \quad \text{for all } (k_1, \dots, k_d) \in \mathbb{Z}_+^d,$$

and stopped at the smallest solution (N_1, \dots, N_d) of the system

$$(2.3) \quad x_j + \sum_{i=1}^d X^{i,j}(N_i) = 0, \quad j \in [d].$$

In this equation, N_i is the total number of vertices of type i in F and x_i is the total number of trees in this forest whose root is of type i . We will say that F is issued from $x = (x_1, \dots, x_d)$. Note that the variables N_i are random, whereas the x_i 's are deterministic.

2.2. The total number of mutations and its asymptotics. A mutation of type i , is the birth event of an individual of type i from an individual of any type $j \neq i$. The aim of this section is to study the evolution of mutations in a multitype branching forest. Our main result asserts that the forest of mutations, that is the forest obtained

by merging together all the vertices of a same cluster, is itself a branching forest if and only if for each $i \in [d]$, one of the following conditions is satisfied,

$$\begin{aligned} (A_i) \quad & m_{ii} \leq 1, \\ (B_i) \quad & m_{ii} > 1 \text{ and for all } j \neq i, m_{ij} = 0. \end{aligned}$$

Moreover, its progeny distribution can be expressed in terms of this of the initial forest. Note that the branching property of the forest of mutations is intuitively clear. In the neutral case, it has been pointed out in [19].

Theorem 2.1. *Let F be any multitype branching forest with progeny distribution $\nu = (\nu_1, \dots, \nu_d)$ and denote by \overline{F} the associated forest of mutations. Assume that for all $i \in [d]$, one of the conditions (A_i) or (B_i) holds. Then \overline{F} is a multitype branching forest with progeny distribution $\mu = (\mu_1, \dots, \mu_d)$ on $\mathcal{S}_i := \{\mathbf{k} \in \mathbb{Z}_+^d : k_i = 0\}$, which is defined by*

$$(2.4) \quad \mu_i(\mathbf{k}) = \sum_{n \geq 1} n^{-1} \nu_i^{*n}(\mathbf{k} + (n-1)e_i), \quad \mathbf{k} \in \mathcal{S}_i,$$

if (A_i) is satisfied. If (B_i) is satisfied, then μ_i is the Dirac mass at 0. Moreover μ satisfies the following properties:

1. Let $\overline{M} = (\overline{m}_{ij})$ be the mean matrix of μ and let $r \geq 1$. Then μ_i admits moments of order r if and only if either for all $j \neq i$, $m_{ij} = 0$ or ν_i admits moments of order r and $m_{ii} < 1$. In the latter case, for all i, j such that $i \neq j$, $\overline{m}_{ij} = \frac{m_{ij}}{1-m_{ii}}$.
2. Assume that $\overline{m}_{ij} < \infty$, for all $i, j \in [d]$. Then \overline{M} is irreducible if and only if M is irreducible. If \overline{M} is primitive, then so is M . The converse is not true.
3. Assume that \overline{M} is primitive, then \overline{M} is subcritical (resp. critical, resp. supercritical) if and only if M is subcritical (resp. critical, resp. supercritical).

If for some $i \in [d]$, none of the conditions (A_i) and (B_i) holds, then there is $j \neq i$ such that individuals of type i in \overline{F} give birth to an infinite number of children of type j with positive probability. Therefore \overline{F} is not a branching forest in our sense.

Proof. Since the result only bears on the progeny law of forests, we do not lose any generality by assuming that F has an infinite number of trees. Then the stochastic processes $X = \{X^{(i)}, i \in [d]\}$ obtained from F , as in (2.1) are defined on the whole integer line $\{0, 1, \dots\}$. Note that their definition slightly extends the definition which is given in [7]. Indeed, without any more assumption on ν , trees of the forest can be infinite, so that the process X is not necessarily a coding of the forest, that is, if some trees are infinite then it is not possible to reconstruct the whole forest from X and the sequence of its roots. However, it is straightforward to check that $X^{(i)}, i \in [d]$ are independent random walks and that the step distribution of $X^{(i)}$ is $\tilde{\nu}_i$, which is defined in (2.2). In particular, the law of X characterizes this of F .

Now, let us consider the forest of mutations \overline{F} . By construction, this forest is composed of an infinite number of independent and identically distributed trees. Hence, in order to show that \overline{F} is a branching forest, it suffices to show that its trees are branching trees.

Let us denote by $\{\overline{X}^{(i)}, i \in [d]\}$ the process which is defined from \overline{F} as in (2.1). Let $i \in [d]$ and assume first that (A_i) holds. Then we define the first passage time process of the random walks $X^{i,i}, i \in [d]$ by,

$$\tau_k^{(i)} = \inf\{n \geq 0 : X_n^{i,i} = -k\}, \quad k \geq 0.$$

Since $m_{ii} \leq 1$, then from the law of large numbers, $\liminf_{n \rightarrow \infty} X_n^{i,i} = -\infty$, a.s., so that $\tau_k^{(i)}$ is almost surely finite for all $k \geq 0$ and $\lim_{k \rightarrow \infty} \tau_k^{(i)} = \infty$, a.s. Moreover, for all $i, j \in [d]$,

$$\overline{X}_k^{i,j} = X^{i,j}(\tau_k^{(i)}), \quad k \geq 0.$$

Indeed, the effect of the time change by $\tau_k^{(i)}$ is to merge all vertices of a same cluster of type i into a single vertex. Note that $\overline{X}_k^{(i)}$, $i \in [d]$ are independent random walks. Assume with no loss of generality that the root of the first tree in \overline{F} has type 1, then a slight extension Theorems 2.7 and 3.1 in [7] to any progeny distribution, allows us to show that this first tree is coded by the processes $(\overline{X}_k^{(i)}, 0 \leq k \leq N_i)$, $i \in [d]$, where (N_1, \dots, N_d) is the smallest solution of the system

$$(2.5) \quad r_j + \sum_{i=1}^d \overline{X}^{i,j}(N_i) = 0, \quad j \in [d],$$

and $(r_1, \dots, r_d) = (1, 0, \dots, 0)$. Note that in our case, N_i can be infinite. This extended notion of smallest solution is defined in [6], see Lemma 1 therein. This coding result implies that the first tree in \overline{F} can be reconstructed from the processes $(\overline{X}_k^{(i)}, 0 \leq k \leq N_i)$, $i \in [d]$ and applying part 3. of Theorem 3.1 in [7], we obtain that this tree is a branching tree whose progeny distribution $\mu = (\mu_i, i \in [d])$ is given by

$$\mu_i(k_1, \dots, k_d) = P(\overline{X}_1^{(i)} = (k_1, \dots, k_{i-1}, -1, k_{i+1}, k_d)), \quad (k_1, \dots, k_d) \in \mathcal{S}_i.$$

Then in order to make this law explicit in terms of ν , we apply the Ballot theorem for cyclically exchangeable sequences due to Takács [18]. Since conditionally on $X^{i,j}$, $i \neq j$, $X^{i,i}$ is downward skip free with cyclical exchangeable increments, we have for all $(k_1, \dots, k_d) \in \mathcal{S}_i$,

$$\begin{aligned} & P(\overline{X}_1^{(i)} = (k_1, \dots, k_{i-1}, -1, k_{i+1}, \dots, k_d)) \\ &= \sum_{n \geq 1} P(X_n^{(i)} = (k_1, \dots, k_{i-1}, -1, k_{i+1}, \dots, k_d), \tau_1^{(i)} = n) \\ &= \sum_{n \geq 1} \frac{1}{n} P(X_n^{(i)} = (k_1, \dots, k_{i-1}, -1, k_{i+1}, \dots, k_d)), \end{aligned}$$

which gives (2.4) from (2.2). If (B_i) holds, then by definition, individuals of type i in \overline{F} are all leaves and hence, $\overline{X}^{i,j} \equiv 0$, for all $j \neq i$ and $\overline{X}_n^{i,i} = -n$, for all $n \geq 0$, see (2.1). In this case, the conclusion follows immediately.

Let us now prove properties 1–3 of μ . First note that for all $i \neq j$, $m_{ij} = 0$ if and only if $\bar{m}_{ij} = 0$. Then let $r \geq 1$, assume that μ_i admits moments of order r and that there is $j \neq i$ such that $m_{ij} = E(X_1^{i,j}) > 0$. The variable $\tau_1^{(i)}$ is a stopping time in the filtration generated by $X^{(i)}$ to which the increasing random walk $X^{i,j}$ is adapted. Then by applying Theorem 5.4 in [10], we obtain that $E((X_1^{i,j})^r) < \infty$ and $E((\tau_1^{(i)})^r) < \infty$. In particular $\tau_1^{(i)} < \infty$, a.s. Now by definition, the random walk $(X_n^{i,i})$ can be written as $X_n^{i,i} = Y_n^{i,i} - n$, where $(Y_n^{i,i})$ is an increasing random walk. Since $Y^{i,i}(\tau_1^{(i)}) = \tau_1^{(i)} - 1$ and $E((\tau_1^{(i)})^r) < \infty$, we have $E(|Y^{i,i}(\tau_1^{(i)})|^r) < \infty$ and by applying Theorem 5.4 in [10] again, we obtain that $E(|Y_1^{i,i}|^r) < \infty$, and hence $E(|X_1^{i,i}|^r) < \infty$. So we have proved that ν admits moments of order r . Then it follows from the definition of $\tau_1^{(i)}$

and from Lemma 3.1 in [13] that $E((\tau_1^{(i)})^r) < \infty$ implies $\lim_{n \rightarrow \infty} X_n^{i,i} = -\infty$, and hence $m_{ii} < 1$, from the law of large numbers.

Conversely, if $m_{ij} = 0$ for all $j \neq i$, then $\bar{m}_{ij} = 0$ for all $j \neq i$ and μ_i is the Dirac mass at 0, so it admits moments of order r . Now assume that ν_i admits moments of order r and $m_{ii} < 1$. Then it follows directly from Lemma 3.1 in [13] that $E((\tau_1^{(i)})^r) < \infty$. Moreover from Theorem 5.2 in [10], $E(X_1^{i,j}(\tau_1^{(i)})^r) < \infty$, for all $j \neq i$, which means that μ_i admits moments of order r . If ν_i admits moments of order 1 and $m_{ii} < 1$, then it follows from the optional stopping theorem applied to the martingale $(X_n^{i,j} - nE(X_1^{i,j}))$, that $E(X_1^{i,i}(\tau_1^{(i)})) = -1 = E(X_1^{i,i})E(\tau_1^{(i)}) = (m_{ii} - 1)E(\tau_1^{(i)})$, and when $i \neq j$, $E(X_1^{i,j}(\tau_1^{(i)})) = E(X_1^{i,j})E(\tau_1^{(i)}) = \frac{m_{ij}}{1 - m_{ii}}$ and part 1 is proved.

If \bar{M} is irreducible, then for all i , there is $j \neq i$ such that $\bar{m}_{ij} > 0$. From part 1., ν_i admits moments of order 1 and $m_{ii} < 1$, for all i . In this case,

$$\bar{M} + \Delta_2 = \Delta_1 M, \quad \text{where } \Delta_1 = \text{diag}\left(\frac{1}{1 - m_{ii}}\right) \text{ and } \Delta_2 = \text{diag}\left(\frac{m_{ii}}{1 - m_{ii}}\right),$$

and we derive from this identity that M is irreducible. Conversely if M is irreducible, then for all i , there is $j \neq i$ such that $m_{ij} > 0$ and hence $\bar{m}_{ij} > 0$. Since by assumption, $\bar{m}_{ij} < \infty$, for all i, j , then from part 1., $m_{ii} < 1$, and $\bar{M} + \Delta_2 = \Delta_1 M$ holds. We derive from this identity that \bar{M} is irreducible.

Now if \bar{M} is primitive, then it is irreducible and as before, $m_{ii} < 1$ for all $i \in [d]$. Moreover,

$$M = (I - \text{diag}(m_{ii}))\bar{M} + \text{diag}(m_{ii}).$$

Therefore M is primitive. The converse cannot be true since there are nonnegative, irreducible matrices whose main diagonal is zero and which are not primitive. We can find distributions ν such that it is the case for \bar{M} and hence for $(I - \text{diag}(m_{ii}))\bar{M}$. If $m_{ii} > 0$, for all i , then it follows from general theory of nonnegative matrices that $M = (I - \text{diag}(m_{ii}))\bar{M} + \text{diag}(m_{ii})$ becomes primitive, see [15].

Let us now prove 3. Recall that by definition, since \bar{M} is primitive, μ_i admits moments of order 1 for all $i \in [d]$. Then from the same arguments as in part 2., $M = (I - \text{diag}(m_{ii}))\bar{M} + \text{diag}(m_{ii})$ and $m_{ii} < 1$ for all $i \in [d]$. Assume that M is supercritical, then there is a positive vector x such that $Mx > x$. Therefore, $(I - \text{diag}(m_{ii}))\bar{M}x > (I - \text{diag}(m_{ii}))x$ and since $m_{ii} < 1$, we obtain $\bar{M}x > x$. Hence \bar{M} is supercritical. Conversely, assume that \bar{M} is supercritical. Then there is a positive vector x such that $\bar{M}x > x$, so that $Mx = (I - \text{diag}(m_{ii}))\bar{M}x + \text{diag}(m_{ii})x > (I - \text{diag}(m_{ii}))x + \text{diag}(m_{ii})x = x$ and thus M is supercritical. Then the identity $M = (I - \text{diag}(m_{ii}))\bar{M} + \text{diag}(m_{ii})$ allows us to derive that M is critical if and only if this is the case for \bar{M} .

Finally assume that $m_{ii} > 1$ for some $i \in [d]$. If $m_{ij} = 0$, for all $j \neq i$, then it is clear that individuals of type i in \bar{F} are leaves. If $m_{ij} > 0$, for some $j \in [d]$, then since clusters of type i are supercritical, some of them have infinitely many children with positive probability. Conditionally to this event, such a cluster produces almost surely infinitely many children of type j , which is equivalent to say that individuals of type i in \bar{F} give birth to an infinite number of children of type j with positive probability. \square

Let us now consider a multitype branching forest F with progeny distribution ν , with a finite number of trees and let $Z_n = (Z_n^{(1)}, \dots, Z_n^{(d)})$, $n \geq 0$ be the associated branching process, that is for each $i \in [d]$, $Z_n^{(i)}$ is the total number of individuals of

type i present in F at generation n . For $x = (x_1, \dots, x_d) \in \mathbb{Z}_+^d$, we denote by \mathbf{P}_x the law on (Ω, \mathcal{F}) under which F is issued from x . In particular, $\mathbf{P}_x(Z_0 = x) = 1$. Then the next result gives the law of the total number of mutations in the forest F , that is the number of mutations up to the last generation whose rank is the extinction time, $T := \inf\{n : Z_n = 0\}$. For $i, j \in [d]$, let us denote by M_i the total number of mutations of type i in F , up to time T and by M_{ij} the total number of mutations of type j produced by individuals of type i . In particular, $M_{ii} = 0$ and M_i and M_{ij} satisfy the relations

$$M_j = \sum_{i=1}^d M_{ij}, \quad j \in [d].$$

Note that if ν is primitive and supercritical, then $\mathbf{P}_x(T = \infty) > 0$ for all $x \in \mathbb{Z}_+^d$, so that under \mathbf{P}_x , M_i and M_{ij} are infinite with positive probability, for some $i, j \in [d]$. We also emphasize that M_i and M_{ij} are not functionals of the branching process (Z_n) .

Corollary 2.2. *Assume that (A_i) or (B_i) holds for all $i \in [d]$. Then for all integers $x_i, n_i, k_{ij}, i, j \in [d]$, such that $x_i \geq 0$, $n_i = -k_{ii}$, for $i \neq j$, $k_{ij} \geq 0$, and for all $j \in [d]$, $n_j = x_j + \sum_{i \neq j} k_{ij}$,*

$$\begin{aligned} & \mathbf{P}_x(M_1 = n_1 - x_1, \dots, M_d = n_d - x_d, M_{ij} = k_{ij}, \forall i \neq j) \\ &= \frac{\det(K)}{\bar{n}_1 \dots \bar{n}_d} \prod_{i=1}^d \mu_i^{*n_i}(k_{i1}, \dots, k_{i(i-1)}, 0, k_{i(i+1)}, \dots, k_{id}), \end{aligned}$$

where μ_i is defined in Theorem 2.1 and $\mu_i^{*0} = \delta_0$, $\bar{n}_i = n_i \vee 1$, K is the matrix $(-k_{ij})_{i,j}$ to which we removed the line i and the column i for all i such that $n_i = 0$.

Proof. This result is a direct consequence of Theorem 1.2 in [7] and Theorem 2.1 applied to the forest of mutations \bar{F} . Indeed, it suffices to note that $x_i + M_i$ corresponds to the total number of individuals of type i in \bar{F} . Note however that Theorem 1.2 in [7] is proved only in the case where ν is primitive and (sub)critical. But using the coding which is presented in Section 2.1 and applying Lemma 1 in [6], we can check that it is still valid in the general case by following along the lines the proof which is given in [7]. \square

If for some $i \in [d]$, none of the conditions (A_i) and (B_i) holds, then the definition of the vector of mutation sizes (M_1, \dots, M_d) still makes sense. In this case, it is possible to obtain its law by extending Theorem 2.1 to branching forests whose progeny laws give mass to infinity. Note also that Corollary 2.2 can be considered as an extension of Theorem 1 in [4], where a similar formula is given in the neutral case.

We now turn our attention to the asymptotic behaviour of the number of mutations, when the total population is growing to infinity. Our first result is concerned with the critical case and is a direct consequence of Proposition 2 in [14] and Theorem 2.1. If M is primitive, then we denote by u and v the unique right and left positive eigenvectors of M which are associated to the eigenvalue 1 and normalized by $u \cdot 1 = u \cdot v = 1$. Recall that, for a multitype branching forest F , when no confusion is possible, N_i denotes the total population of type i in F and M_i denotes the total number of mutations of type i in F . Note also that when ν is primitive and critical, then (A_i) necessarily holds for all $i \in [d]$, so that from Theorem 2.1, the forest of mutations \bar{F} associated to F is a branching forest with progeny distribution μ defined by (2.4).

Corollary 2.3. *Let F be a branching forest with a non singular, primitive and critical progeny distribution ν . Assume that for all $i \in [d]$, μ_i admits moments of order $d + 1$. If moreover \overline{M} is primitive and the covariance matrices $\Sigma^i, \overline{\Sigma}^i$ of ν_i and μ_i , respectively are positive definite. Then $m_{ii} < 1$, for all $i \in [d]$ and there are constants $C_1, C_2 > 0$ such that for all $x_0 \in \mathbb{Z}_+^d$,*

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{d/2+1} \mathbf{P}_{x_0}(M_i = \lfloor n(1 - m_{ii})v_i \rfloor, i \in [d]) &= C_1 x_0 \cdot u, \\ \lim_{n \rightarrow \infty} n^{d+1} \mathbf{P}_{x_0}(M_i = \lfloor n(1 - m_{ii})v_i \rfloor, N_i = \lfloor nv_i \rfloor, i \in [d]) &= C_2 x_0 \cdot u. \end{aligned}$$

Proof. Since by assumption, \overline{M} is primitive, then for all i , there is $j \neq i$ such that $\overline{m}_{ij} > 0$, and hence $m_{ij} > 0$. Therefore, from part 1. of Theorem 2.1, $m_{ii} < 1$, for all i . Moreover, from our assumptions and part 3. of Theorem 2.1, μ is critical. Besides, it is plain that \overline{M} is non singular. Then conditions of Proposition 2 in [14] are satisfied for the multitype branching process associated to \overline{F} and the first assertion follows with \bar{u} and \bar{v} , the normalized, positive right and left eigenvectors of \overline{M} associated to the eigenvalue 1. Then recall from the proof of part 3. of Theorem 2.1 that $M = (I - \text{diag}(m_{ii}))\overline{M} + \text{diag}(m_{ii})$. We derive from this identity that $\bar{u} = u$ and $\bar{v} = cv(I - \text{diag}(m_{ii}))$, where $c = \|u \cdot v(I - \text{diag}(m_{ii}))\|^{-1}$ and the first assertion follows.

The proof of the second assertion follows the same lines as the proof of Proposition 2 in [14]. In this case, since the number of mutations is taken into account together with the total number of individuals, a $2d$ -dimensional random walk is involved in the proof, which explains that the rate of convergence is now $d + 1$. \square

Note that the constants C_1 and C_2 can be made explicit in terms of the distributions ν and μ by properly exploiting the proof of Proposition 2 in [14].

Through the next result we focus on the asymptotic behaviour of the number of mutations in a branching forest when the initial number of individuals $x = (x_1, \dots, x_d)$ tends to infinity along some given direction.

Theorem 2.4. *Let $F(x)$ be any family of multitype branching forests defined on the space (Ω, \mathcal{F}, P) , indexed by $x \in \mathbb{Z}_+^d$ and such that for each x , $F(x)$ has progeny distribution ν and is issued from x . For $i \in [d]$, let $N_i(x)$ (resp. $M_i(x)$) be the total number of individuals (resp. of mutations) of type i in $F(x)$. Assume that ν is primitive and let $w \in \mathbb{Z}_+^d \setminus \{0\}$.*

1. *If ν is critical, then*

$$\lim_{n \rightarrow \infty} \frac{N_i(nw)}{n} = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{M_i(nw)}{N_i(nw)} = 1 - m_{ii}, \quad \text{in probability.}$$

2. *If ν is subcritical, then*

$$\lim_{n \rightarrow \infty} \frac{N_i(nw)}{n} = c_i(w) \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{M_i(nw)}{n} = w_i + (1 - m_{ii})c_i(w), \quad \text{in probability,}$$

$$\text{where } c_i(w) := \sum_{k=1}^d w_k (I - M)_{ki}^{-1}.$$

In any case, $m_{ii} < 1$, for all $i \in [d]$.

Proof. In order to prove our result, it suffices to construct some particular family of forests $F(x)$, such that for each x , $F(x)$ has progeny distribution ν and is issued from $x \in \mathbb{Z}_+^d$, and to show that the limits in the statement hold.

Recall the coding of multitype branching forests which is presented at the end of Section 2.2 and let $X^{(i)} = \{X^{i,j}, j \in [d]\}$ be d independent random walks whose respective step distributions are $\tilde{\nu}_i, i \in [d]$ defined in (2.2). Then for each $x \in \mathbb{Z}_+^d$, we construct a forest $F(x)$ such that $F(x)$ is encoded by the random walks $X^{(i)}, i \in [d]$ and contains exactly x_i trees whose root is of type i . This construction is possible in the primitive, (sub)critical case, thanks to part 3. of Theorem 3.1 in [7].

Then $N_i(x)$ and $X^{(i)}, i \in [d]$, satisfy identity (2.3). Moreover, for $k \neq i$, the number of mutations of type i issued from an individual of type k is $X^{k,i}(N_k(x))$, so that the total number of mutations of type i is

$$M_i(x) = \sum_{k \neq i} X^{k,i}(N_k(x)) = -x_i - X^{i,i}(N_i(x)).$$

We derive from Lemma 2.2 in [7], that if $x_1, x_2 \in \mathbb{Z}_+^d$ are such that $x_1 \leq x_2$, then the couple of random variables $(N_i(x_2) - N_i(x_1), X^{i,i}(N_i(x_2) - N_i(x_1)))$ is independent of process $((N_i(x), X^{i,i}(N_i(x))), x \leq x_1)$ and has the same law as $(N_i(x_2 - x_1), X^{i,i}(N_i(x_2 - x_1)))$. Therefore, for any $w \in \mathbb{Z}_+^d, ((N_i(nw), X^{i,i}(N_i(nw))), n \geq 0)$ is a bivariate random walk whose step distribution is the law of $(N_i(w), X^{i,i}(N_i(w)))$.

Let $Z = (Z^{(1)}, \dots, Z^{(d)})$ be the branching process associated to $F(w)$. Then by definition of $N_i(w)$, we have $N_i(w) = \sum_{n=0}^{\infty} Z_n^{(i)}$. But $\mathbf{E}_w(Z_n) = wM^n$, so that $\mathbf{E}_w(Z_n^{(j)}) = \sum_{i=1}^d w_i m_{ij}^{(n)}$ and since ν is primitive, we have from Frobenius Theorem for primitive matrices, $m_{ij}^{(n)} \sim u_i v_j \rho^n$, see Theorem 1, Section V.2 in [2]. So we have proved that $E(N_i(w)) < \infty$ if and only if ν is subcritical. Moreover, if ν is subcritical, then $I - M$ is invertible and it follows from the above expressions that $E(N_i(w)) = \sum_{i=1}^d w_i (I - M)_{ij}^{-1}$. Then assertions 1. and 2. follow directly from the law of large numbers.

Finally, since ν and μ are primitive, by definition, they admit moments of order 1 and we derive from part 1. of Theorem 2.1 that $m_{ii} < 1$, for all $i \in [d]$. \square

3. WHEN CONTINUOUS TIME IS INVOLVED

3.1. The Lamperti representation. Let us now consider a d type population which is composed at time $t = 0$, of x_i individuals of type $i \in [d]$ and whose dynamics in continuous time behave according to a branching model. More specifically, at any time, all individuals in the population live, give birth and die independently of each other. Once it is born, any individual of type $i \in [d]$ gives birth after an exponential time with parameter $\lambda_i > 0$ to n_j individuals of type $j \in [d]$ with probability $\nu_i(n_1, \dots, n_d)$. Then this individual dies at the same time it gives birth. We emphasize that in this model, the probability for the population to become extinct does not depend on the rates λ_i .

This model is represented as a plane forest with edge lengths, see Figure 3. (In each sibling, we rank individuals of type 1 to the left, then individuals of type 2, and so on.) Such a forest will be called a multitype branching forest with edge lengths issued from $x = (x_1, \dots, x_d)$, with progeny distribution $\nu := (\nu_1, \dots, \nu_d)$ and reproduction rates $(\lambda_1, \dots, \lambda_d)$. By construction, its discrete time skeleton is a multitype branching (plane) forest, as defined in the previous section, with progeny distribution ν , which is independent from the edge lengths. Edge lengths are independent between themselves

and the length of an edge issued from a vertex of type i follows an exponential distribution with parameter λ_i . We emphasize that the total number of individuals and the total number of mutations in a multitype branching forest with edge lengths are the same as in its discrete skeleton. Hence, the results of the previous section can be applied in the present setting.

Given a branching forest with edge lengths, as defined above, we denote by $Z = (Z^{(1)}, \dots, Z^{(d)})$ the corresponding multitype branching process, that is for $t \geq 0$ and $i \in [d]$, $Z_t^{(i)}$ is the number of individuals of type i at time t in the population. (Since no confusion is possible, for the branching process we have kept the same notation as in discrete time.) The process Z is a \mathbb{Z}_+^d -valued continuous time Markov process which satisfies the branching property, i.e., for $\lambda \in \mathbb{R}_+^d$, $t \geq 0$ and $x, y \in \mathbb{Z}_+^d$,

$$\mathbb{E}_{x+y}(e^{-\lambda Z_t}) = \mathbb{E}_x(e^{-\lambda Z_t})\mathbb{E}_y(e^{-\lambda Z_t}),$$

where \mathbb{P}_x is the law under which the forest is issued from x . In particular, $Z_0 = x$, \mathbb{P}_x -a.s. The process Z actually contains much less information than the original branching forest. In order to preserve the essential part of this information, we need to decompose Z as in the following definition.

Definition 3.1. *For $i \neq j$, we denote by $Z_t^{i,j}$ the total number of individuals of type j whose parent has type i and who were born before time t . For $i = j$, the definition of $Z_t^{i,i}$ is the same, except that to this number we add the number of individuals of type i at time 0 and we subtract the number of individuals of type i who died before time t .*

The processes $Z^{i,j}$ whose definition should be clear from the example given in Figure 3 will play a crucial role in our continuous time model. A more formal definition can be found in Section 4.2 of [6]. The interest of these processes is the following straightforward decomposition of the branching process $Z = (Z^{(1)}, \dots, Z^{(d)})$:

$$(3.6) \quad Z_t^{(j)} = \sum_{i=1}^d Z_t^{i,j}, \quad j \in [d].$$

Our model bears on a Lamperti type representation of these processes. According to Lamperti representation, any one dimensional branching process can be expressed as a Lévy process time changed by some integral functional. In this subsection, we will recall from [6] the extension of this transformation to multitype, continuous time, discrete valued branching processes. The latter involves time changed multidimensional compound Poisson processes which we now introduce.

Since our models of evolution are only concerned with mutations, individuals of type i having exactly one child of type i do not present any interest. Hence we can assume without loss of generality that

$$\nu_i(e_i) = 0, \text{ for all } i \in [d].$$

Then let $X = (X^{(1)}, \dots, X^{(d)})$, where $X^{(i)}$, $i \in [d]$ are d independent \mathbb{Z}^d -valued compound Poisson processes. We assume that $X_0^{(i)} = 0$ and that $X^{(i)}$ has rate λ_i and jump distribution $\tilde{\nu}_i$ which has been defined in (2.2). In particular, with the notation $X^{(i)} = (X^{i,1}, \dots, X^{i,d})$, the process $X^{i,i}$ is a \mathbb{Z} -valued, downward skip free, compound Poisson process, i.e. $\Delta X_t^{i,i} = X_t^{i,i} - X_{t-}^{i,i} \geq -1$, $t \geq 0$, with $X_{0-} = 0$ and for all $i \neq j$, the process $X^{i,j}$ is an increasing compound Poisson process. We emphasize that in this

definition, some of the processes $X^{i,j}$, $i, j \in [d]$ can be identically equal to 0.

The following extension of the Lamperti representation to multitype branching processes can be found in [6], see also [5] for the case of continuous state multitype branching processes.

Theorem 3.2. *Let us consider a multitype branching forest with edge lengths issued from $x = (x_1, \dots, x_d) \in \mathbb{Z}_+^d$, with progeny distribution $\nu := (\nu_1, \dots, \nu_d)$ and reproduction rates $(\lambda_1, \dots, \lambda_d)$. Then the processes $Z^{i,j}$, $i, j \in [d]$ introduced in Definition 3.1 admit the following representation:*

$$(3.7) \quad Z_t^{i,j} = \begin{cases} X_{\int_0^t Z_s^{(i)} ds}^{i,j}, & t \geq 0, \quad \text{if } i \neq j, \\ x_i + X_{\int_0^t Z_s^{(i)} ds}^{i,i}, & t \geq 0, \quad \text{if } i = j, \end{cases}$$

where the processes,

$$X^{(i)} = (X^{i,1}, X^{i,2}, \dots, X^{i,d}), \quad i = 1, \dots, d,$$

are independent \mathbb{Z}_+^d valued compound Poisson processes, with jump distribution $(\tilde{\nu}_1, \dots, \tilde{\nu}_d)$ and rates $(\lambda_1, \dots, \lambda_d)$. In particular from (3.6) and (3.7), the multitype branching process Z admits the following representation,

$$(3.8) \quad (Z_t^{(1)}, \dots, Z_t^{(d)}) = x + \left(\sum_{i=1}^d X_{\int_0^t Z_s^{(i)} ds}^{i,1}, \dots, \sum_{i=1}^d X_{\int_0^t Z_s^{(i)} ds}^{i,d} \right), \quad t \geq 0.$$

3.2. Further results on asymptotics of mutations. For $i \in [d]$ and $t \geq 0$, we will denote by $M_{i,t}$ the total number of mutations of type i which occurred up to time t . The definition of this quantity is illustrated on Figure 3. Let us also define a cluster of type i as the subtree corresponding to the descendance of type i of an individual of type i which is either a root or an individual whose parent is a type different from i . Then $x_i + M_{i,t}$ corresponds to the number of clusters of type i in the forest truncated at time t .

In Proposition 3.4, we describe the asymptotic behaviour of $M_{i,t}$, as t tends to ∞ in the case where the progeny distribution ν is primitive and supercritical. To this aim, we will need the joint representation of $M_{i,t}$ together with the number $Z_t^{(i)}$ of individuals of type i at time t which is presented in Proposition 3.3.

Proposition 3.3. *Recall from Section 3.1 the definition of the compound Poisson processes $X^{i,j}$, $i, j \in [d]$. Then for any $x = (x_1, \dots, x_d) \in \mathbb{Z}_+^d$, under \mathbb{P}_x , the stochastic process $(Z_t^{(i)}, M_{i,t})$ fulfills the following representation,*

$$\left(Z_t^{(i)}, M_{i,t} \right) = \left(x_i + \sum_{k=1}^d X_{\int_0^t Z_u^{(k)} du}^{k,i}, \sum_{k=1, k \neq i}^d X_{\int_0^t Z_u^{(k)} du}^{k,i} \right), \quad t \geq 0.$$

Proof. This result is a direct consequence of the representation which is recalled in Theorem 3.2. Indeed, recall from Section 3.1 the definition of $Z^{i,j}$, then the number of mutations of type i up to time t is

$$M_{i,t} = \sum_{k \neq i} Z_t^{k,i}.$$

The result follows from identity (3.7) in Theorem 3.2. \square

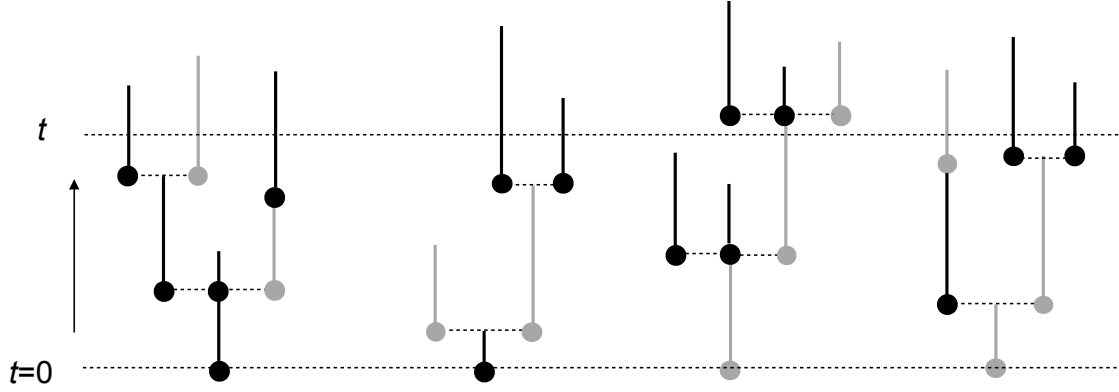


FIGURE 3. A two type forest with edge lengths issued from $x = (2, 2)$. Vertices of type 1 (resp. 2) are represented in black (resp. grey). At time t ,

$$Z_t^{(1)} = 6, \quad Z_t^{(2)} = 3, \quad Z_t^{1,1} = -2, \quad Z_t^{1,2} = 5, \quad Z_t^{2,1} = 8, \quad Z_t^{2,2} = -2,$$

$$\text{and } M_{1,t} = 8, \quad M_{2,t} = 5.$$

Let's us now turn to the limiting behavior of $M_{i,t}$, as t tends to infinity. The next result is concerned with the case where ν is primitive and supercritical. It allows us to evaluate the number of mutations which occurred up to time t (or equivalently the number of clusters in the forest truncated at time t), when t is large.

Let us define the matrix $A = \Lambda(M - I)$, where $\Lambda = \text{diag}(\lambda_i)$. If M is primitive, then so is A and it follows from Perron-Frobenius theory that the eigenvalues ρ_i , $i \in [d]$ of A can be arranged so that $\rho_1 > \text{Re}(\rho_2) \geq \dots \geq \text{Re}(\rho_d)$. Moreover, ν is subcritical, critical or supercritical according as $\rho_1 < 0$, $\rho_1 = 0$ or $\rho_1 > 0$. Then a well known result due to [3], see also Theorem 2, p. 206 in [2] asserts that when ν is non singular and primitive, there exists a nonnegative random variable W such that for all $i \in [d]$,

$$(3.9) \quad \lim_{t \rightarrow \infty} e^{-\rho_1 t} Z_t^{(i)} = v_i W, \quad \text{a.s.},$$

where v_i is the i -th coordinate of the normalized left eigenvector associated with ρ_1 .

Proposition 3.4. *Assume that ν is non singular, primitive and supercritical. Then for all $i \in [d]$,*

$$\lim_{t \rightarrow \infty} e^{-\rho_1 t} M_{i,t} = K_i W, \quad \text{a.s.},$$

where $K_i = v_i(1 + (1 - m_{ii})(\lambda_i \rho_1)^{-1})$.

Proof. We derive from Proposition 3.3 that,

$$Z_t^{(i)} - M_{i,t} = X_{\int_0^t Z_u^{(i)} du}^{i,i}, \quad \text{a.s.}$$

On the other hand, in the supercritical case, ρ_1 is strictly positive. Hence it follows from (3.9) that

$$\int_0^t Z_u^{(i)} du \sim \rho_1^{-1} W v_i e^{\rho_1 t}, \quad \text{a.s.}, \quad \text{as } t \rightarrow \infty.$$

Then the desired result is a consequence of the latter equivalence and the law of large numbers applied to the compound Poisson process $X^{i,i}$. \square

Under conditions of Proposition 3.4, assume moreover that for some $i \in [d]$, K_i is positive, that is

$$m_{ii} < 1 + \lambda_i \rho_1,$$

and that for some j , $\mathbb{P}_{e_j}(W > 0) = 1$. Then using Proposition 3.4, we can compare the asymptotic behaviour of the number of mutations prior to t with this of $Z_t^{(i)}$, under \mathbb{P}_{e_j} , that is

$$(3.10) \quad M_{i,t} \sim K_i Z_t^{(i)}, \quad \mathbb{P}_{e_j}\text{-a.s., as } t \rightarrow \infty.$$

Regarding the condition $\mathbb{P}_{e_j}(W > 0) = 1$, note that Theorem 2, p. 206 in [2] also asserts that $\mathbb{P}_{e_k}(W > 0) > 0$, for some (hence for all) $k \in [d]$, if and only if

$$(3.11) \quad E(\xi_{ij} \log \xi_{ij}) < \infty, \text{ for all } i, j \in [d],$$

where $(\xi_{i1}, \dots, \xi_{id})$ is a random vector with law ν_i . Moreover, $1 - \mathbb{P}_{e_k}(W > 0)$ corresponds to the probability of extinction, when the forest is issued from e_k .

3.3. Emergence times of mutations. In this section, we shall assume that mutations are not reversible, that is for all $i = 1, \dots, d-1$, individuals of type i can only have children of type i or $i+1$. In particular ν is not irreducible. Moreover when giving birth, individuals of type $i = 1, \dots, d-1$ have at least one child of type i with probability one, and have children of type $i+1$ with positive probability. These conditions can be explicitated in terms of the progeny distribution ν_i as follows

$$(3.12) \quad \begin{cases} \nu_i(\mathbf{k}) > 0 \Rightarrow k_j = 0, \text{ for } j \notin \{i, i+1\}, \\ \sum_{\mathbf{k} \in \mathbb{Z}_+^d : k_i = 0} \nu_i(\mathbf{k}) = 0 \text{ and } \sum_{\mathbf{k} \in \mathbb{Z}_+^d : k_{i+1} = 0} \nu_i(\mathbf{k}) < 1. \end{cases}$$

We are interested in the waiting time until an individual of type i first emerges in the population, that is

$$\tau_i := \inf\{t \geq 0 : Z_t^{(i)} \geq 1\}.$$

The problem of determining a general expression for the law of τ_i is quite challenging. As far as we know, there is no explicit expression for this law in terms of the progeny distribution and the reproduction rates. Various results in this direction can be found in [16], [17], [9] and [1] for instance. Most of them provide approximations of this law, using martingale convergence theorems [9] or through numerical methods for the inversion of the generating function [1]. In Proposition 3.5 we first give a relationship between the successive emergence times τ_2, τ_3, \dots in terms of the underlying compound Poisson process in the Lamperti representation of Z . We also characterize the joint law under $\mathbb{P}_{e_{i-1}}$ of the time τ_i and the number of individuals of type $i-1$ at this time. In Theorem 3.7 we derive an approximation of the time τ_i , under \mathbb{P}_{e_1} , as the mutation rate of type k increases faster than that of type $k-1$, for all $k = 3, \dots, i$. Then in Corollary 3.8 we focus on a case where these law can be explicitated.

In the following developments, we use the notation of Section 3.1 from which we recall the Lamperti representation of the multitype branching process $Z = (Z^{(1)}, \dots, Z^{(d)})$ in terms of the compound Poisson processes $X^{(i)}$. Let us also introduce a few more

notation. For $i, j \in [d]$, we denote by $\lambda_{i,j}$ the parameter of the compound Poisson process $X^{i,j}$, that is

$$\lambda_{i,j} := \lambda_i \left(1 - \sum_{\mathbf{k} \in \mathbb{Z}_+^d : k_j = 0} \tilde{\nu}_i(\mathbf{k}) \right).$$

Note that from our assumptions (3.12), for all $i = 1, \dots, d-1$, $\lambda_{i,i+1} > 0$ and for $j \notin \{i, i+1\}$, $\lambda_{i,j} = 0$, that is $X^{i,j}$ is identically equal to 0. In particular, $\lambda_i = \lambda_{i,i} + \lambda_{i,i+1}$, for $i \leq d-1$ and $\lambda_d = \lambda_{d,d}$. The parameter $\lambda_{i,i+1}$ will be call the mutation rate of type $i+1$. For $i \geq 2$, let

$$\gamma_i := \inf\{t : X_t^{i-1,i} \geq 1\}$$

be the time of the first jump by the process $X^{i-1,i}$ and note that this time is exponentially distributed with parameter $\lambda_{i-1,i}$.

Proposition 3.5. *Assume that (3.12) holds and define $Z^{0,1}$ as the process identically equal to 1 and set $\tau_1 = 0$.*

1. *For $i = 2, \dots, d$, the emergence time τ_i of type i admits the following representation under \mathbb{P}_{e_1} ,*

$$(3.13) \quad \tau_i = \tau_{i-1} + \int_0^{\gamma_i} \frac{1}{X_s^{i-1,i-1} + Z_{\kappa_{i-1}(s)}^{i-2,i-1}} ds, \quad \mathbb{P}_{e_1}\text{-a.s.},$$

where κ_{i-1} is the right continuous inverse of the functional $t \mapsto \int_0^t Z_s^{(i-1)} ds$, i.e. $\kappa_{i-1}(t) = \inf\{s > 0 : \int_0^s Z_u^{(i-1)} du > t\}$.

2. *Under $\mathbb{P}_{e_{i-1}}$, the joint law of the emergence time τ_i of type i together with the number of individuals of type $i-1$ in the population at time τ_i admits the following representation,*

$$(3.14) \quad (\tau_i, Z_{\tau_i}^{(i-1)}) \stackrel{(d)}{=} \left(\int_0^{\gamma_i} \frac{ds}{1 + X_s^{i-1,i-1}}, 1 + X_{\gamma_i}^{i-1,i-1} \right).$$

3. *Let us define $\theta_k = \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + 1} ds$, for $k \geq 2$. Then the random variables θ_k , $k \geq 2$ are independent and for $i = 2, \dots, d$,*

$$(3.15) \quad \mathbb{P}_{e_1}(\tau_i > t) \leq P \left(\sum_{k=2}^i \theta_k > t \right), \quad \text{for all } t > 0.$$

Proof. Since $X^{i,j}$ is identically equal to 0 whenever $j \notin \{i, i+1\}$, then under \mathbb{P}_{e_1} , the representation (3.8) admits the simpler form

$$(3.16) \quad (Z_t^{(1)}, \dots, Z_t^{(d)}) = e_1 + \left(X_{\int_0^t Z_s^{(1)} ds}^{1,1}, X_{\int_0^t Z_s^{(2)} ds}^{2,2} + X_{\int_0^t Z_s^{(1)} ds}^{1,2}, \dots, X_{\int_0^t Z_s^{(d)} ds}^{d,d} + X_{\int_0^t Z_s^{(d-1)} ds}^{d-1,d} \right).$$

In particular, for $i = 2, \dots, d$,

$$Z_t^{(i)} = X_{\int_0^t Z_s^{(i)} ds}^{i,i} + X_{\int_0^t Z_s^{(i-1)} ds}^{i-1,i}, \quad t \geq 0.$$

Since $X_0^{i,i} = 0$, for $i \geq 2$, we see that the time τ_i corresponds to the first hitting time of level 1 by the process $t \mapsto X_{\int_0^t Z_s^{(i-1)} ds}^{i-1,i}$, that is

$$(3.17) \quad \tau_i = \kappa_{i-1}(\gamma_i),$$

where γ_i has been defined as the time of the first jump of the process $X^{i-1,i}$. For t such that $\kappa_{i-1}(t) < \infty$, we have $t = \int_0^{\kappa_{i-1}(t)} Z_s^{(i-1)} ds$, so that $dt = Z_{\kappa_{i-1}(t)}^{(i-1)} d\kappa_{i-1}(t)$, and

since $\kappa_{i-1}(0) = \tau_{i-1}$, we obtain

$$\begin{aligned}
\kappa_{i-1}(t) &= \tau_{i-1} + \int_0^t \frac{ds}{Z_{\kappa_{i-1}(s)}^{(i-1)}} \\
(3.18) \quad &= \tau_{i-1} + \int_0^t \frac{ds}{X_s^{i-1,i-1} + X_{\int_0^{\kappa_{i-1}(s)} Z_u^{(i-2)} du}^{i-2,i-1}}.
\end{aligned}$$

The latter identity together with (3.17) prove identity (3.13).

The second part of the proposition is easily derived from the same arguments. More specifically, it follows from (3.17) and the following identities

$$Z_t^{(i-1)} = 1 + X_{\int_0^t Z_s^{(i-1)} ds}^{i-1,i-1} \quad \text{and} \quad \kappa_{i-1}(t) = \int_0^t \frac{ds}{1 + X_s^{i-1,i-1}}, \quad t \geq 0,$$

which hold $\mathbb{P}_{e_{i-1}}$ -a.s.

Independence between the variables θ_k , $k \geq 2$ is a direct consequence of the independence between the processes $X^{(i)}$, $i \in [d]$. We derive from the representation of τ_i in part 1. of this proposition that

$$(3.19) \quad \tau_i = \sum_{k=2}^i \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2,k-1}} ds, \quad \text{a.s.}$$

Note that since $\kappa_{k-1}(0) = \tau_{k-1}$, then from (3.17), for all $k \geq 2$, $\int_0^{\kappa_{k-1}(0)} Z_u^{(k-2)} du = \gamma_{k-1}$, so that by definition of γ_{k-1} ,

$$(3.20) \quad X_{\int_0^{\kappa_{k-1}(0)} Z_u^{(k-2)} du}^{k-2,k-1} = X_{\gamma_{k-1}}^{k-2,k-1} \geq 1, \quad \text{a.s.}$$

Besides, since $s \mapsto X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2,k-1}$ are increasing processes, then inequality (3.15) is a direct consequence of identities (3.19) and (3.20). \square

Note that the law of θ_k or equivalently, the law of τ_k under $\mathbb{P}_{e_{k-1}}$ can be made explicit in some instances through its Laplace transform, see Corollary 3.8 below.

For the remainder of this section we will assume moreover that at each mutation, individuals of type i do not give birth to more than one child of type $i+1$ in a same litter. More specifically, assumptions (3.12) are replaced by,

$$(3.21) \quad \begin{cases} \nu_i(\mathbf{k}) > 0 \Rightarrow k_{i+1} = 0 \text{ or } 1 \text{ and } k_j = 0, \text{ for } j \notin \{i, i+1\}, \\ \sum_{\mathbf{k} \in \mathbb{Z}_+^d : k_i=0} \nu_i(\mathbf{k}) = 0 \text{ and } \sum_{\mathbf{k} \in \mathbb{Z}_+^d : k_{i+1}=0} \nu_i(\mathbf{k}) < 1. \end{cases}$$

In particular, under these assumptions, the process $X^{i,i+1}$ is a standard Poisson process. Then we will need the next lemma in order to derive our main result on the estimation of the time τ_i , as the mutation rates $\lambda_{k-1,k}$, $k = 2, \dots, d$ grow faster.

Lemma 3.6. *Assume that (3.21) holds, let $k \geq 3$ and fix $\lambda_{1,2} > 0$, then*

$$(3.22) \quad \mathbb{P}_{e_1} \left(X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} du}^{k-2,k-1} = 1 \right) \longrightarrow 1, \quad \text{as } \lambda_{n-2,n-1} / \lambda_{n-1,n} \rightarrow 0, \text{ for } n = 3, \dots, k.$$

Proof. First set $\gamma_{k-1}^{(1)} = \inf\{t > \gamma_{k-1} : X_t^{k-2,k-1} = 2\}$ and note that

$$\begin{aligned} \{X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} du}^{k-2,k-1} = 1\} &= \left\{ \int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} du < \gamma_{k-1}^{(1)} \right\} \\ &= \{\kappa_{k-1}(\gamma_k) < \kappa_{k-2}(\gamma_{k-1}^{(1)})\}. \end{aligned}$$

It is easy to check that $\kappa_{k-2}(\gamma_{k-1}^{(1)}) = \tau_{k-1}^{(1)}$, where

$$\tau_{k-1}^{(1)} := \inf\{t > \tau_{k-1} : Z_t^{k-2,k-1} - Z_{\tau_{k-1}}^{k-2,k-1} = 1\}.$$

(Note that from our assumptions $Z_{\tau_{k-1}}^{k-2,k-1} = 1$ and $Z_{\tau_{k-1}^{(1)}}^{k-2,k-1} = 2$, \mathbb{P}_{e_1} -a.s.) So from (3.17), we have showed that

$$(3.23) \quad \left\{ X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} du}^{k-2,k-1} = 1 \right\} = \{\tau_k < \tau_{k-1}^{(1)}\}.$$

The event $\{\tau_k < \tau_{k-1}^{(1)}\}$ means that before the first time when an individual of type k appears in the population, there has been only one birth of type $k-1$. From the Markov property applied at time τ_{k-1} , we have

$$(3.24) \quad \mathbb{P}_{e_1}(\tau_k \leq \tau_{k-1}^{(1)}) = \int \mathbb{P}_z(\tau_k \leq \tau_{k-1}^{(1)}) \mathbb{P}_{e_1}(Z_{\tau_{k-1}} \in dz).$$

The support in the integral of (3.24) is included in the set $\{z : z_{k-1} = 1\}$, so from (3.23), (3.24) and the Lebesgue theorem of dominated convergence, all we need to prove is

$$(3.25) \quad \mathbb{P}_z(\tau_k \leq \tau_{k-1}^{(1)}) \rightarrow 1, \quad \text{as } \lambda_{n-2,n-1}/\lambda_{n-1,n} \rightarrow 0, \text{ for } n = 3, \dots, k,$$

for all z such that $z_{k-1} = 1$. (Note that if z is such that $z_1 = \dots = z_{k-2} = 0$, or such that $z_k \geq 1$, then it is clear that $\mathbb{P}_z(\tau_k \leq \tau_{k-1}^{(1)}) = 1$, since in the first case $Z^{k-2,k-1}$ is identically equal to 0, so that $\tau_{k-1}^{(1)} = \infty$, \mathbb{P}_z -a.s. and in the second case, $\tau_k = 0$, \mathbb{P}_z -a.s.)

Let z be such that $z_{k-1} = 1$. Without loss of generality we can assume that $z_i \geq 1$, for $i = 1, \dots, k-2$. For $i = 1, \dots, k-1$, let us denote by U_i the first time that the lineage of one of the z_{k-i} initial individuals of type $k-i$ gives birth to an individual of type $k-i+1$. Then from the branching property, under \mathbb{P}_z , the r.v.'s U_i are independent and from part 2. of Proposition 3.5, U_i has the same law as $\int_0^{\gamma_{k-i+1}} \frac{ds}{X_s^{k-i,k-i} + z_{k-i}}$. Then set $Y_s^{(i)} := X_s^{k-i,k-i} + z_{k-i}$ and note the inclusions,

$$\left\{ \gamma_k \leq \min\left(\gamma_{k-1}/Y_{\gamma_{k-1}}^{(2)}, \dots, \gamma_2/Y_{\gamma_2}^{(k-1)}\right) \right\} \subset \{U_1 \leq \min(U_2, \dots, U_{k-1})\} \subset \{\tau_k \leq \tau_{k-1}^{(1)}\},$$

which imply the inequality,

$$P(\gamma_k/\gamma_{k-1} \leq \min(1/Y_{\gamma_{k-1}}^{(2)}, \gamma_{k-2}/(\gamma_{k-1}Y_{\gamma_{k-2}}^{(3)}), \dots, \gamma_2/(\gamma_{k-1}Y_{\gamma_2}^{(k-1)})) \leq \mathbb{P}_z(\tau_k \leq \tau_{k-1}^{(1)}).$$

But when $\lambda_{n-2,n-1}/\lambda_{n-1,n} \rightarrow 0$, for $n = 3, \dots, k$, the parameter $\lambda_{1,2} > 0$ being fixed, we necessarily have $\lim \lambda_{n-1,n} = \infty$, for $n = 3, \dots, k$. Hence γ_k/γ_{k-1} converges in probability toward 0, $1/Y_{\gamma_{k-1}}^{(2)}$ converges in probability toward $1/z_{k-2}$ and $\gamma_{n-1}/(\gamma_n Y_{\gamma_{n-1}}^{(k-n+2)})$, for $n = 3, \dots, k-1$ converge in probability toward $+\infty$. Therefore, the left hand side of the above inequality tends to 1, which proves (3.25) and the lemma is proved. \square

In the following theorem, the assumption $\frac{\lambda_{k-1,k}}{\lambda_{k,k+1}} \rightarrow 0$ is quite adapted to several biological models such as cancer growth, for instance. Indeed, cancer is often the result of a series of successive mutations, [12], [9], [8]. Each new mutation is itself more unstable than the previous ones, and in particular, the successive mutation rates can

increase very fast. It would be interesting to study the asymptotic behavior of τ_i , when $\frac{\lambda_{k,k}}{\lambda_{k+1,k+1}} \rightarrow 0$, that is when the intrinsic reproduction rates increase very fast. This assumption also fits to the model of cancer, since mutations are always more sensitive to proliferate.

Theorem 3.7. *Assume that (3.21) holds. Recall the definition of θ_k in Proposition 3.5 and let us fix $\lambda_{1,2} > 0$, then under \mathbb{P}_{e_1} ,*

$$\frac{\tau_i}{\sum_{k=2}^i \theta_k} \xrightarrow{P} 1, \quad \text{as } \frac{\lambda_{k-2,k-1}}{\lambda_{k-1,k}} \rightarrow 0, \text{ for } k = 3, \dots, i.$$

Besides, the expectation of τ_i fulfills the following approximation:

$$\mathbb{E}_{e_1}(\tau_i) \sim \sum_{k=2}^i E(\theta_k), \quad \text{as } \frac{\lambda_{k-2,k-1}}{\lambda_{k-1,k}} \rightarrow 0, \text{ for } k = 3, \dots, i.$$

Proof. Since $s \mapsto X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2,k-1}$ are increasing processes, then from (3.20), \mathbb{P}_{e_1} -almost surely on the set $\{X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} du}^{k-2,k-1} = 1\}$, we have

$$\int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2,k-1}} ds = \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + 1} ds.$$

Hence it follows from Lemma 3.6 that for fixed $\lambda_{1,2} > 0$, as $\lambda_{n-2,n-1}/\lambda_{n-1,n} \rightarrow 0$, for all $n = 3, \dots, k$,

$$\left(\int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + 1} ds \right)^{-1} \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2,k-1}} ds \xrightarrow{P} 1,$$

and the first part of the theorem is easily derived from this convergence and (3.13) (or equivalently (3.19)).

In order to prove the second part, let us first set

$$H_k := \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2,k-1}} ds \quad \text{and} \quad A_k := \{X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} du}^{k-2,k-1} = 1\}.$$

Then from (3.13), $\mathbb{E}_{e_1}(\tau_i) = \sum_{k=2}^i \mathbb{E}_{e_1}(H_k)$, so it suffices to prove that for all $k = 2, \dots, i$,

$$(3.26) \quad \mathbb{E}_{e_1}(H_k) \sim E(\theta_k), \quad \text{as } \lambda_{n-2,n-1}/\lambda_{n-1,n} \rightarrow 0, \text{ for } n = 3, \dots, k.$$

Observe that $\mathbb{E}_{e_1}(H_k) = E(\theta_k \mathbf{1}_{A_k}) + \mathbb{E}_{e_1}(H_k \mathbf{1}_{A_k^c})$. Moreover, $\mathbb{E}_{e_1}(H_k \mathbf{1}_{A_k^c}) \leq E(\theta_k \mathbf{1}_{A_k^c})$. Then to obtain (3.26), it is enough to prove that

$$(3.27) \quad \frac{E(\theta_k \mathbf{1}_{A_k^c})}{E(\theta_k)} \rightarrow 0, \quad \text{as } \lambda_{n-2,n-1}/\lambda_{n-1,n} \rightarrow 0, \text{ for } n = 3, \dots, k.$$

But for any $p, q \geq 1$, such that $p^{-1} + q^{-1} = 1$, we have from Holder inequality $E(\theta_k \mathbf{1}_{A_k^c}) \leq E(\theta_k^p)^{1/p} P(A_k^c)^{1/q}$. Moreover, we clearly have $E(\theta_k^p)^{1/p} \sim 1/\lambda_{k-1,k}$, as $\lambda_{k-1,k} \rightarrow \infty$. Hence, (3.27) is satisfied thanks to Lemma 3.6. \square

We end this section with an example where the distribution of τ_i can be estimated a bit more specifically. We consider the case of binary fission with mutations, where each individual of type i can give birth to either two individuals of type i or one individual

of type i and one individual of type $i + 1$. In particular, all jumps of $Z^{i,i}$ have size 1 and $X^{i,i}$ is a Poisson process with parameter $\lambda_{i,i}$.

Corollary 3.8. *With the above assumptions, the law of τ_i can be specified as follows.*

1. Under $\mathbb{P}_{e_{i-1}}$, the Laplace transform of τ_i is expressed as,

$$\mathbb{E}_{e_{i-1}}(e^{-\alpha\tau_i}) = \lambda_{i-1,i} \sum_{n \geq 0} \frac{\lambda_{i-1,i-1}^n}{\prod_{k=0}^n (\alpha_k + \dots + \alpha_n + \bar{\alpha}_{n+1})}, \quad \alpha \geq 0,$$

where $\alpha_0 = 0$, $\alpha_k = \frac{\alpha}{k(k+1)}$ and $\bar{\alpha}_k = \lambda_{i-1} + \frac{\alpha}{k}$, for $k \geq 1$.

2. The expectation of τ_i is given by $\mathbb{E}_{e_{i-1}}(\tau_i) = \frac{1}{\lambda_{i-1,i} \lambda_{i-1,i-1}} \ln \frac{\lambda_{i-1}}{\lambda_{i-1,i}}$. In particular, for fixed $\lambda_{1,2} > 0$, under \mathbb{P}_{e_1} , the expectation of τ_i fulfills the following approximation:

$$\mathbb{E}_{e_1}(\tau_i) \sim \sum_{k=2}^i \lambda_{k-1,k}^{-2}, \quad \text{as } \frac{\lambda_{k-2,k-1}}{\lambda_{k-1,k}} \rightarrow 0, \text{ for } k = 3, \dots, i.$$

Proof. From part 2. of Proposition 3.5, for all $\beta \geq 0$,

$$\begin{aligned} \mathbb{E}_{e_{i-1}}(e^{-\alpha\tau_i}) &= E \left(e^{-\alpha \int_0^{\gamma_i} \frac{1}{1+X_s^{i-1,i-1}} ds} \right) \\ (3.28) \quad &= \lambda_{i-1,i} \int_0^{+\infty} E \left(e^{-\alpha \int_0^x \frac{1}{1+X_s^{i-1,i-1}} ds} \right) e^{-\lambda_{i-1,i}x} dx. \end{aligned}$$

Under $\mathbb{P}_{e_{i-1}}$, $X^{i-1,i-1}$ is a standard Poisson process with parameter $\lambda_{i-1,i-1}$ starting at 0. So if we denote by $(J_n)_{n \geq 1}$ the sequence of jump times of $X^{i-1,i-1}$ and set $J_0 = 0$, then developing the expression $E \left(e^{-\alpha \int_0^x \frac{1}{1+X_s^{i-1,i-1}} ds} \right)$, we obtain with the convention that $\sum_{k=0}^{-1} = 0$,

$$\begin{aligned} E \left(e^{-\alpha \int_0^x \frac{1}{1+X_s^{i-1,i-1}} ds} \right) &= \sum_{n \geq 0} E \left(X_x^{i-1,i-1} = n, e^{-\alpha \left(\frac{x-J_n}{n+1} + \sum_{k=0}^{n-1} \frac{J_{k+1}-J_k}{k+1} \right)} \right) \\ &= e^{-(\alpha + \lambda_{i-1,i-1})x} + \sum_{n \geq 1} e^{-\lambda_{i-1,i-1}x} \frac{(\lambda_{i-1,i-1}x)^n}{n!} \\ &\quad \times \int_{0 \leq x_1 \leq \dots \leq x_n \leq x} e^{-\alpha \left(\frac{x}{n+1} + \sum_{k=1}^n \frac{x_k}{k(k+1)} \right)} \frac{n!}{x^n} dx_1 \dots dx_n \\ &= e^{-(\alpha + \lambda_{i-1,i-1})x} + \sum_{n \geq 1} \lambda_{i-1,i-1}^n e^{-(\lambda_{i-1,i-1} + \frac{\alpha}{n+1})x} \\ &\quad \times \int_{0 \leq x_1 \leq \dots \leq x_n \leq x} e^{-\alpha \sum_{k=1}^n \frac{x_k}{k(k+1)}} dx_1 \dots dx_n. \end{aligned}$$

Then coming back to expression (3.28), we obtain with the convention that $\sum_{k=1}^0 = 0$,

$$\mathbb{E}_{e_{i-1}}(e^{-\alpha\tau_i}) = \lambda_{i-1,i} \sum_{n \geq 0} \lambda_{i-1,i-1}^n \int_{0 \leq x_1 \leq \dots \leq x_{n+1}} e^{-(\bar{\alpha}_{n+1}x_{n+1} + \sum_{k=1}^n \alpha_k x_k)} dx_1 \dots dx_{n+1},$$

where $\alpha_1, \dots, \alpha_n, \bar{\alpha}_{n+1}$ are defined in the statement. (Here we used the fact that $\lambda_{i-1} = \lambda_{i-1,i} + \lambda_{i-1,i-1}$.) The computation of the integral is easily done.

Then using again part 2. of Proposition 3.5, we obtain the expectation of τ_i under $\mathbb{P}_{e_{i-1}}$, after easy computations,

$$\begin{aligned} \mathbb{E}_{e_{i-1}}(\tau_i) &= \int_0^{+\infty} dx \lambda_{i-1,i} e^{-\lambda_{i-1,i}x} \int_0^x e^{-\lambda_{i-1,i-1}s} \sum_{k \geq 0} \frac{(\lambda_{i-1,i-1}s)^k}{(k+1)!} ds \\ &= \frac{1}{\lambda_{i-1,i} \lambda_{i-1,i-1}} \ln \frac{\lambda_{i-1}}{\lambda_{i-1,i}}. \end{aligned}$$

We conclude from Theorem 3.7. \square

REFERENCES

- [1] ALEXANDER, H.K. (2013). *Conditional distributions and waiting times in Multitype branching processes*. Adv. Appl. Prob. 45, 692–718.
- [2] ATHREYA, K.B. AND NEY, P.E. (1972). *Branching Processes*. Springer, Berlin.
- [3] ATHREYA, K.B. (1967). *Some results on multitype continuous time Markov branching processes*. AMS, 39, 347–357.
- [4] BERTOIN, J. (2009) The structure of the allelic partition of the total population for Galton-Watson processes with neutral mutations. *Ann. Probab.* 37, no. 4, 1502–1523.
- [5] M.E. CABALLERO, J.L. PÉREZ GARMENDIA, AND G.URIBE BRAVO. (2015) Affine processes on $\mathbb{R}_+^n \times \mathbb{R}^n$ and multiparameter time changes. *Preprint*, arXiv:1501.03122.
- [6] CHAUMONT, L. *Breadth first search coding of multitype forests with application to Lamperti representation*. To appear in Séminaire de Probabilités XLVII, volume In Memoriam Marc Yor, (2015).
- [7] CHAUMONT, L. AND LIU, R. *Coding multitype forests: application to the law of the total population of branching forests*. To appear in Transactions of the American Mathematical Society, (2015).
- [8] DURRETT, R. (2013) *Population genetics of neutral mutations in exponentially growing cancer cell populations*. *The Annals of Applied Probability*, Vol. 23, No. 1, 230–250.
- [9] DURRETT, R. AND MOSELEY, S. (2010). *Evolution of resistance and progression to disease during clonal expansion of cancer*. Theor. Pop. Biol. 77 42–48.
- [10] GUT, A. *Stopped random walks. Limit theorems and applications*. SECOND EDITION. SPRINGER SERIES IN OPERATIONS RESEARCH AND FINANCIAL ENGINEERING. SPRINGER, NEW YORK, 2009.
- [11] HAENO, H., IWASA, Y., MICHOR, F. (2007). *The evolution of two mutations during clonal expansion*. GENETICS 177, 2209–2221.
- [12] IWASA, Y., MICHOR, F., KOMAROVA, N.L., NOWAK, M.A. (2005). *Population genetics of tumor suppressor genes*. J. THEORET. BIOL. 233, 15–23.
- [13] KESTEN, H. AND MALLER, R. A. (1996) Two renewal theorems for general random walks tending to infinity. *Probab. Theory Related Fields* 106, no. 1, 1–38.
- [14] PENISSON, S. (2015) Various ways of conditioning multitype Galton-Watson processes, *Preprint*, arXiv:1412.3322v2
- [15] SENETA, E. (2006) Non-negative matrices and Markov chains. Springer Series in Statistics. Springer, New York, 2006.
- [16] SERRA, M.C. (2006). *On the Waiting Time to Escape*. Journal of applied Probability, vol. 43, No. 1, 296–302.
- [17] SERRA, M.C. AND HACCOU, P.(2007). *Dynamics of escape mutants*. Theoret. Pop.Biol., 72, 167–178.

- [18] TAKÁCS, L. (1961) The probability law of the busy period for two types of queuing processes. *Operations Res.*, 9, 402–407.
- [19] TAĪB, Z. (1992) Branching processes and neutral evolution. Lecture Notes in Biomathematics, 93. *Springer-Verlag*, Berlin.

LOÏC CHAUMONT – LAREMA – UMR CNRS 6093, UNIVERSITÉ D'ANGERS, 2 BD LAVOISIER,
49045 ANGERS CEDEX 01

E-mail address: loic.chaumont@univ-angers.fr

THI NGOC ANH NGUYEN – LAREMA – UMR CNRS 6093, UNIVERSITÉ D'ANGERS, 2 BD
LAVOISIER, 49045 ANGERS CEDEX 01

E-mail address: nguyen@math.univ-angers.fr