ON MUTATIONS IN THE BRANCHING MODEL FOR MULTITYPE POPULATIONS

LOÏC CHAUMONT * ** AND THI NGOC ANH NGUYEN,* *** *Université d'Angers*

Abstract

The forest of mutations associated to a multitype branching forest is obtained by merging together all vertices in each of its clusters and by preserving connections between them. (Here, by cluster, we mean a maximal connected component of the forest in which all vertices have the same type.) 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 we describe some of its crucial properties in terms of the initial progeny distribution. We also obtain the limiting behaviour of the number of mutations both when the total number of individuals tends to ∞ and when the number of roots tends to ∞ . The continuous-time case is then investigated by considering multitype branching forests with edge lengths. When mutations are nonreversible, we give a representation of their emergence times which allows us to describe the asymptotic behaviour of the latter, under certain conditions on the mutation rates. These results have potential relevance for emergence of mutations in population cells, particularly for genetic evolution of cancer or development of infectious diseases.

Keywords: Multitype branching forest; mutation; forest of mutations; emergence time

2010 Mathematics Subject Classification: Primary 60J80

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, we refer the reader to, for example, [2], [10], [11], [13], [14], and [16].

In this paper we are 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 tends to ∞ .

Received 29 October 2015; revision received 1 February 2018.

^{*} Postal address: LAREMA – UMR CNRS 6093, Université d'Angers, 2 bd Lavoisier, 49045 Angers cedex 01, France.

^{**} Email address: loic.chaumont@univ-angers.fr

^{***} Email address: nguyen@math.univ-angers.fr

When time is continuous, we are mainly interested in emergence times of new mutations in the nonreversible case. In this paper we give explicit and exact expressions for successive emergence times of mutations. More specifically, we use a recent extension of the Lamperti representation in higher dimensions in order to express successive emergence times in terms of integral functionals of compound Poisson processes. These exact expressions are then exploited to give approximations of emergence times as sums of independent random variables whose law is easily evaluated. Moreover, under some conditions on the mutation rates, we are able to compute the Laplace transform of emergence times and derive from this expression some accurate estimates of their mean. Among previous works in this direction, Durrett [10], [11] used an approximation of the branching process itself for large times in order to provide explicit expressions for the tail distribution of successive emergence times. Exact results can also be found in [1] for the emergence time of some particular type. In this paper we prove that the distribution of this time satisfies some differential equation from which it can be numerically evaluated. Our work can also be compared to [19] and [20], where the waiting time to escape extinction was studied in discrete time. In these works, the reversible case was also considered but this more general assumption only allows us to obtain a functional equation satisfied by the generating function of the waiting time. Some approximations on the mutation rates yield nicer expressions leading to numerical evaluations.

In Section 2.1 we start with some preliminaries on the coding of multitype branching forests by multivariate random walks. Then we state and prove our results on the total mutations sizes of branching forests in Sections 2.2 and 3.2. Results bearing on emergence times are presented in Section 3.3. Then we conclude in Section 4 with a discussion on potential applications of our results to the genetic evolution of cancer and development of infectious diseases.

2. Mutations and their asymptotics in discrete multitype forests

2.1. Preliminaries on discrete multitype forests

Throughout we use the notation $\mathbb{Z}_+ = \{0, 1, 2, ...\}$, and for any positive integer *d* we set $[d] = \{1, ..., d\}$. We will denote by e_i the *i*th unit vector of \mathbb{Z}_+^d . We define the following partial order on \mathbb{R}^d by setting

$$x = (x_1, ..., x_d) \ge y = (y_1, ..., y_d)$$
 if $x_i \ge y_i$ for all $i \in [d]$.

The convention inf $\emptyset = +\infty$ will be valid throughout this paper. Then $(\Omega, \mathcal{F}, \mathbb{P})$ is a reference probability space on which all the stochastic processes involved in this paper are defined.

Let us first recall the coding of multitype forests, as defined in [9]. A (plane) tree t is a directed planar graph with no loops, on a possibly infinite and nonempty set of vertices v(t). All directed paths of a tree t have a common terminal vertex called the *root* of t. For two vertices $u, v \in v(t)$, if (u, v) is a directed edge of t, we say that u is a *child* of v and that v is the *parent* of u. Each vertex has a finite number of children and at most one parent. The only vertex with no parent is the root.

A forest f is a possibly infinite set of trees. We will denote by v(f) the set of vertices of f. The roots of the trees of a forest f are called the roots of f. We give an order to the trees of the forest f and denote them by $t_1(f), t_2(f), \ldots, t_k(f), \ldots$ (we will usually write $t_1, t_2, \ldots, t_k, \ldots$ if no confusion is possible). Then we rank (a part of) the vertices of faccording to the breadth-first search order, by ranking first the vertices of t_1 , then the vertices of t_2 , and so on; see the labelling of the two forests in Figure 2. Note that if t_k , for $k \ge 1$ is the first infinite tree, then the vertices of t_{k+1}, \ldots have no label according to this procedure. To each forest f, we associate the application $c_f : v(f) \to [d]$ such that if we place $u_i, u_{i+1}, \ldots, u_{i+j} \in v(f)$ from left to right and each having the same parent, then $c_f(u_i) \leq c_f(u_{i+1}) \leq \cdots \leq c_f(u_{i+j})$. For $v \in v(f)$, the integer $c_f(v)$ is called the *type* (or the *colour*) of v. The couple (f, c_f) is called a *d-type forest*. When no confusion is possible, we will simply write f. The set of *d*-type forests will be denoted by \mathfrak{F}_d .

A cluster or a subtree of type $i \in [d]$ of a d-type forest $(f, c_f) \in \mathcal{F}_d$ is a maximal connected subgraph of (f, c_f) whose vertices are all of type *i*. Formally, *t* is a cluster of type *i* of (f, c_f) if it is a connected subgraph whose vertices are all of type *i* and such that either the root of *t* has no parent or the type of its parent is different from *i*. Moreover, if the parent of a vertex $v \in v(t)^c$ belongs to v(t) then $c_f(v) \neq i$. Clusters of type *i* in t_1 are ranked according to the order of their roots in the breadth-first search order of t_1 ; see Figures 1 and 2. If the number of clusters of type *i* is finite in t_1 then we continue by ranking clusters of type *i* in t_2 , and so on. Note that with this procedure it is possible that clusters of t_k, t_{k+1}, \ldots , for some *k*, are not ranked. We denote by $t_1^{(i)}, t_2^{(i)}, \ldots, t_k^{(i)}, \ldots$ the sequence of clusters of type *i* in (f, c_f) . The forest $f^{(i)} := \{t_1^{(i)}, t_2^{(i)}, \ldots, t_k^{(i)}, \ldots\}$ is called *the subforest of type i* of (f, c_f) . We denote by $u_1^{(i)}, u_2^{(i)}, \ldots$ the elements of $v(f^{(i)})$, ranked in the breadth-first search order of $f^{(i)}$. The subforests of the 2-type forest from Figure 1 are represented in Figure 2.

To any forest $(f, c_f) \in \mathcal{F}_d$, we associate the *forest of mutations*, denoted by $(f, c_{\bar{f}}) \in \mathcal{F}_d$, which is the forest of \mathcal{F}_d obtained by aggregating all the vertices of each subtree of (f, c_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 presented in Figure 1.

For a forest $(f, c_f) \in \mathcal{F}_d$ and $u \in v(f)$, 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 $f^{(i)}$ of (f, c_f) . Then let us define the *d*-dimensional chain $x^{(i)} = (x^{i,1}, \ldots, x^{i,d})$, with length n_i and whose values belong to the set \mathbb{Z}^d , by $x_0^{(i)} = 0$ and



FIGURE 1: Upper: 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. Lower: the corresponding forest of mutations.



FIGURE 2: The subforests of the 2-type forest presented in Figure 1 with their breadth-first search labelling.

if $n_i \geq 1$,

$$\begin{aligned} x_{n+1}^{i,j} - x_n^{i,j} &= p_j(u_{n+1}^{(i)}) & \text{if } i \neq j, \\ x_{n+1}^{i,i} - x_n^{i,i} &= p_i(u_{n+1}^{(i)}) - 1, \qquad 0 \le n \le n_i - 1, \end{aligned}$$
(2.1)

where $(u_n^{(i)})_{n\geq 1}$ is the labelling of the subforest $f^{(i)}$ 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 skipfree chain, that is, $x_{n+1}^{i,i} - x_n^{i,i} \geq -1$ 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 [9], when trees of (f, c_f) are finite, the data of the chains $x^{(1)}, \ldots, x^{(d)}$ together with the sequence of ranked roots of (f, c_f) allow us to reconstruct this forest.

Let us now apply this coding to multitype branching forests. Let $v := (v_1, \ldots, v_d)$, where v_i is some distribution on \mathbb{Z}^d_+ . We consider a branching process with progeny distribution v, 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 $v_i(n_1, \ldots, n_d)$. 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*, that is,

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

We say that ν is nonsingular 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 $n \ge 1$ 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 $\rho \le 1$ then the population will become extinct almost surely (a.s.), 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 ν is irreducible, primitive, (sub)critical or supercritical, when this is the case for M. By multitype branching forest with progeny distribution v, we mean a sequence with a finite (deterministic) or infinite number of independent multitype branching trees with progeny distribution v. A multitype branching forest will be considered as a random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in \mathfrak{F}_d . To any multitype branching forest F, we associate the random sequences $X = \{X^{(i)}, i \in [d]\}$, where $X^{(i)} = \{(X_n^{i,1}, \ldots, X_n^{i,d}), 0 \le n \le n_i\}$, which are constructed as in (2.1). It was proved in [9, Theorem 3.1] that if F is a multitype branching forest with a finite number of trees, whose progeny distribution is primitive and (sub)critical, then $X^{(i)}$, $i \in [d]$, are independent random walks whose step distribution \tilde{v}_i is defined by

$$\tilde{\nu}_i(k_1, \dots, k_d) := \nu_i(k_1, \dots, k_{i-1}, k_i + 1, k_{i+1}, \dots, k_d), \qquad k_i \ge -1, \ k_j \ge 0, \ \text{if} \ i \ne j$$
(2.2)

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

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

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 $x = (x_1, ..., x_d)$. Note that the variables N_i are random, whereas the x_i are deterministic.

2.2. The total number of mutations and its asymptotics

A mutation producing 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 the same cluster, is itself a branching forest if and only if, for each $i \in [d]$, one of the following conditions is satisfied:

$$(A_i) \ m_{ii} \leq 1;$$

 (B_i) $m_{ii} > 1$ and, for all $j \neq i$, $m_{ij} = 0$.

Moreover, its progeny distribution can be expressed in terms of this for the initial forest. Note that the branching property of the forest of mutations is intuitively clear. In the neutral case, this was pointed out in [21].

In the next statement, we denote by v_i^{*n} the *n*-fold convolution product of the probability v_i by itself.

Theorem 2.1. Let F be any multitype branching forest with progeny distribution given by $v = (v_1, \ldots, v_d)$ and \overline{F} the associated forest of mutations. Assume that, for all $i \in [d]$, one of conditions (A_i) or (B_i) holds. Then \overline{F} is a multitype branching forest with progeny distribution $\mu = (\mu_1, \ldots, \mu_d)$, where μ_i is a probability on the set $\mathscr{S}_i := \{k \in \mathbb{Z}_+^d : k_i = 0\}$, which is defined by

$$\mu_i(k) = \sum_{n \ge 1} n^{-1} \nu_i^{*n} (k + (n-1)e_i), \qquad k \in \mathscr{S}_i,$$
(2.4)

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.

(i) Let $\overline{M} = (\overline{m}_{ij})$ be the mean matrix of μ and let $r \ge 1$. Then μ_i admits moments of order r if and only if either, for all $j \ne i$, $m_{ij} = 0$ or v_i admits moments of order r and $m_{ii} < 1$. In the latter case, for all i, j such that $i \ne j$, $\overline{m}_{ij} = m_{ij}/(1 - m_{ij})$.

- (ii) 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.
- (iii) Assume that \overline{M} is primitive then \overline{M} is subcritical (respectively, critical, supercritical) if and only if M is subcritical (respectively, critical, supercritical).

If, for some $i \in [d]$, neither of 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 relates to 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 $\{0, 1, ...\}$. Note that their definition slightly extends the definition of [9]. Indeed, without any more assumptions on v, 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, we can check exactly as in [9] that $X^{(i)}$, $i \in [d]$, are independent random walks and that the step distribution of $X^{(i)}$ is \tilde{v}_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.

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

$$\tau_k^{(i)} = \inf\{n \ge 0 \colon X_n^{i,i} = -k\}, \qquad k \ge 0.$$

Since $m_{ii} \leq 1$ then, from the law of large numbers, $\lim \inf_{n \to \infty} X_n^{i,i} = -\infty$ a.s., so that $\tau_k^{(i)}$ is a.s. finite for all $k \geq 0$ and $\lim_{k \to \infty} \tau_k^{(i)} = \infty$ a.s. Moreover, it is clear from definition (2.1) that, for all $i, j \in [d]$,

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

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 $\overline{X}^{(i)}$, $i \in [d]$, are independent random walks. Assume with no loss of generality that the root of the first tree in \overline{F} is type 1. Then a slight extension of Theorems 2.7 and 3.1 of [9] to any progeny distribution allows us to show that this first tree is coded by the processes $(\overline{X}_k^{(i)}, 0 \le k \le \overline{N}_i), i \in [d]$, where $(\overline{N}_1, \ldots, \overline{N}_d)$ is the smallest solution of the system

$$r_j + \sum_{i=1}^d \overline{X}^{i,j}(\overline{N}_i) = 0, \qquad j \in [d],$$

and $(r_1, \ldots, r_d) = (1, 0, \ldots, 0)$. Note that in our case, \overline{N}_i can be infinite. This extended notion of the smallest solution is defined in [8, Lemma 1]. This coding result implies that the first tree in \overline{F} can be reconstructed from the processes $(\overline{X}_k^{(i)}, 0 \le k \le \overline{N}_i), i \in [d]$. Moreover, applying part 3 of Theorem 3.1 of [9], we see that this tree is a branching tree whose progeny distribution $\mu = (\mu_i, i \in [d])$ is given by

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

Then in order to make this law explicit in terms of v, we apply the Ballot theorem for cyclically exchangeable sequences due to Takács [22]. Since, conditionally on $X^{i,j}$, $i \neq j$, $X^{i,i}$ is downward skip-free with cyclical exchangeable increments, we have, for all $(k_1, \ldots, k_d) \in \delta_i$,

$$\mathbb{P}(\overline{X}_{1}^{(i)} = (k_{1}, \dots, k_{i-1}, -1, k_{i+1}, \dots, k_{d}))$$

$$= \sum_{n \ge 1} \mathbb{P}(X_{n}^{(i)} = (k_{1}, \dots, k_{i-1}, -1, k_{i+1}, \dots, k_{d}), \tau_{1}^{(i)} = n)$$

$$= \sum_{n \ge 1} \frac{1}{n} \mathbb{P}(X_{n}^{(i)} = (k_{1}, \dots, k_{i-1}, -1, k_{i+1}, \dots, k_{d})),$$

from which we obtain (2.4) using (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 $n \ge 0$; see (2.1). In this case, the conclusion follows immediately.

We now prove properties (i)–(iii) 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 $m_{ij} = \mathbb{E}(X_1^{i,j}) > 0$. 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 of [12], we obtain $\mathbb{E}((X_1^{i,j})^r) < \infty$ and $\mathbb{E}((\tau_1^{(i)})^r) < \infty$. 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 $\mathbb{E}((\tau_1^{(i)})^r) < \infty$, we have $\mathbb{E}(|Y^{i,i}(\tau_1^{(i)})|^r) < \infty$, and by applying Theorem 5.4 of [12] again, we obtain $\mathbb{E}(|Y_1^{i,i}|^r) < \infty$ and, hence, $\mathbb{E}(|X_1^{i,i}|^r) < \infty$. 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 of [15] that $\mathbb{E}((\tau_1^{(i)})^r) < \infty$ implies that $\lim_{n\to\infty} X_n^{i,i} = -\infty$ 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 of [15] that $\mathbb{E}((\tau_1^{(i)})^r) < \infty$. Moreover, from Theorem 5.2 of [12], $\mathbb{E}(X^{i,j}(\tau_1^{(i)})^r) < \infty$ 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 $(X_n^{i,j} - n\mathbb{E}(X_1^{i,j}))$ that $\mathbb{E}(X_1^{i,i}(\tau_1^{(i)})) = -1 = \mathbb{E}(X_1^{i,i})\mathbb{E}(\tau_1^{(i)}) = (m_{ii} - 1)\mathbb{E}(\tau_1^{(i)})$, and when $i \neq j$, $\mathbb{E}(X^{i,j}(\tau_1^{(i)})) = \mathbb{E}(X_1^{i,j})\mathbb{E}(\tau_1^{(i)}) = m_{ij}/(1 - m_{ii})$ and part (i) is proved.

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

$$M + \Delta_2 = \Delta_1 M,$$

where $\Delta_1 = \text{diag}(1/(1 - m_{ii}))$ and $\Delta_2 = \text{diag}(m_{ii}/(1 - m_{ii}))$, and from this identity we derive 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 (i), $m_{ii} < 1$, and $\overline{M} + \Delta_2 = \Delta_1 M$ holds. From this identity we derive that \overline{M} is irreducible.

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

$$M = (I - \operatorname{diag}(m_{ii}))\overline{M} + \operatorname{diag}(m_{ii}).$$

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 v

such that \overline{M} (and, hence, $(I - \text{diag}(m_{ii}))\overline{M}$) satisfies these properties. If $m_{ii} > 0$ for all *i* then it follows from general theory of nonnegative matrices that $M = (I - \text{diag}(m_{ii}))\overline{M} + \text{diag}(m_{ii})$ becomes primitive; see [18].

We now prove (iii). Recall that, by definition, since \overline{M} is primitive, μ_i admits moments of order 1 for all $i \in [d]$. Then, from the same arguments as in part (ii), $M = (I - \text{diag}(m_{ii}))\overline{M} + \text{diag}(m_{ii})$ 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, $(I - \text{diag}(m_{ii}))\overline{M}x > (I - \text{diag}(m_{ii}))x$ and since $m_{ii} < 1$, we obtain $\overline{M}x > x$. Hence, \overline{M} is supercritical. Conversely, assume that \overline{M} is supercritical. Then there is a positive vector x such that Mx > x. Hence, \overline{M} is supercritical. Conversely, assume that \overline{M} is supercritical. Then there is a positive vector x such that $\overline{M}x > x$, so that $Mx = (I - \text{diag}(m_{ii}))\overline{M}x + \text{diag}(m_{ii})x > (I - \text{diag}(m_{ii}))x + \text{diag}(m_{ii})x = x$ and, thus, M is supercritical. Then the identity $M = (I - \text{diag}(m_{ii}))\overline{M} + \text{diag}(m_{ii})$ allows us to derive that M is critical if and only if this is the case for \overline{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 \overline{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 on this event, such a cluster produces a.s. infinitely many children of type *j*, which is equivalent to saying that individuals of type *i* in \overline{F} give birth to an infinite number of children of type *j* with positive probability.

We now consider a multitype branching forest F with progeny distribution ν with a finite number of trees and let $Z_n = (Z_n^{(1)}, \ldots, Z_n^{(d)}), n \ge 0$, 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 $x = (x_1, \ldots, x_d) \in \mathbb{Z}_+^d$, we denote by \mathbb{P}_x the law on (Ω, \mathcal{F}) under which F is issued from x. In particular, $\mathbb{P}_x(Z_0 = x) = 1$. In the next result we state 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 $i, j \in [d]$, denote by M_i the total number of mutations producing type i in F up to time T and by M_{ij} the total number of mutations producing type j produced by individuals of type i. In particular, $M_{ii} = 0$ and M_i and M_{ij} satisfy the relation

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

Note that if ν is primitive and supercritical then $\mathbb{P}_x(T = \infty) > 0$ for all $x \in \mathbb{Z}^d_+ \setminus \{0\}$, so that under \mathbb{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 2.1. Assume that (A_i) or (B_i) holds for all $i \in [d]$. Then, for all $i, j \in [d]$ and all integers x_i, n_i, k_{ij} , such that $x_i \ge 0, n_i = -k_{ii}, k_{ij} \ge 0$ when $i \ne j$ and $n_j = x_j + \sum_{i \ne j} k_{ij}$,

$$\mathbb{P}_{x}(M_{1} = n_{1} - x_{1}, \dots, M_{d} = n_{d} - x_{d}, M_{ij} = k_{ij} \text{ for all } i \neq j)$$
$$= \frac{\det(K)}{\bar{n}_{1} \cdots \bar{n}_{d}} \prod_{i=1}^{d} \mu_{i}^{*n_{i}}(k_{i1}, \dots, k_{i(i-1)}, 0, k_{i(i+1),\dots,k_{id}}),$$

where μ_i is defined in Theorem 2.1, μ_i^{*0} is the Dirac mass at 0, $\bar{n}_i = n_i \vee 1$, and K is the matrix $(-k_{ij})_{i,j}$ from 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 of [9] and the first part of the statement of Theorem 2.1 applied to the forest of mutations \overline{F} . Indeed, it suffices to note

that $x_i + M_i$ corresponds to the total number of individuals of type *i* in \overline{F} . Note, however, that Theorem 1.2 of [9] is proved only in the case where ν is primitive and (sub)critical. But using the coding presented in Section 2.1 and applying Lemma 1 of [8], we can check that it is still valid in the general case by following the lines of the proof of [9].

If, for some $i \in [d]$, neither of conditions (A_i) and (B_i) hold then the definition of the vector of mutation sizes (M_1, \ldots, 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 ∞ . However, we do not consider this case since it makes no biological sense. Note also that Corollary 2.1 can be considered as an extension of Theorem 1 of [5], where a similar formula can be found for the neutral case.

We now turn our attention to the asymptotic behaviour of the number of mutations, when the total population is growing to ∞ . Our first result is concerned with the critical case and is a direct consequence of Proposition 2 of [17] 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 producing type i in F. Note also that when v is primitive and critical then (A_i) necessarily holds for all $i \in [d]$, so that from Theorem 2.1, the forest of mutations \overline{F} associated to F is a branching forest with progeny distribution μ defined by (2.4).

Corollary 2.2. Let *F* be a branching forest with a nonsingular, primitive, and critical progeny distribution *v*. 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 Σ^i and $\overline{\Sigma}^i$ of v_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 $x_0 \in \mathbb{Z}_+^d$,

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

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 Theorem 2.1(i), $m_{ii} < 1$ for all *i*. Moreover, from our assumptions and Theorem 2.1(iii), μ is critical. Besides, it is plain that \overline{M} is nonsingular. Then the conditions of Proposition 2 of [17] are satisfied for the multitype branching process associated to \overline{F} and the first assertion follows with \overline{u} and \overline{v} , the normalized, positive right and left eigenvectors of \overline{M} associated to the eigenvalue 1. Then recall from the proof of Theorem 2.1(iii) that $M = (I - \text{diag}(m_{ii}))\overline{M} + \text{diag}(m_{ii})$. From this identity we see that $\overline{u} = u$ and $\overline{v} = cv(I - \text{diag}(m_{ii}))$, 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 of [17]. In this case, since the number of mutations is taken into account together with the total number of individuals, a two-dimensional random walk is involved in the proof, which explains that the rate of convergence is now d + 1.

Note that 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 of [17].

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 $x = (x_1, ..., x_d)$ tends to ∞ along some given direction.

Theorem 2.2. Let F(x) be any family of multitype branching forests defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$, indexed by $x \in \mathbb{Z}_+^d$, and such that, for each x, F(x) has progeny distribution v and is issued from x. For $i \in [d]$, let $N_i(x)$ (respectively, $M_i(x)$) be the total number of individuals (respectively, of mutations) of type i in F(x). Assume that v is primitive and let $w \in \mathbb{Z}_+^d \setminus \{0\}$.

(i) If v is critical then

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

(ii) If v is subcritical then

$$\lim_{n \to \infty} \frac{N_i(nw)}{n} = c_i(w)$$

and

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

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{v}_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 of [9].

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

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

From Lemma 2.2 of [9], we see that if $x_1, x_2 \in \mathbb{Z}_+^d$ are such that $x_1 \leq x_2$, then the couple of random variables $(N_i(x_2) - N_i(x_1), X^{i,i}(N_i(x_2)) - X^{i,i}(N_i(x_1)))$ is independent of the 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 $Z = (Z^{(1)}, \ldots, Z^{(d)})$ be the branching process associated to F(w). Then, by the definition of $N_i(w)$, we have $N_i(w) = \sum_{n=0}^{\infty} Z_n^{(i)}$. But $\mathbb{E}_w(Z_n) = wM^n$, so that $\mathbb{E}_w(Z_n^{(j)}) = \sum_{i=1}^{d} w_i m_{ij}^{(n)}$ and since v is primitive, we have, from the Frobenius theorem for primitive matrices, $m_{ij}^{(n)} \sim u_i v_j \rho^n$; see Theorem 1 of [4, Section V.2]. So we have proved that $\mathbb{E}(N_i(w)) < \infty$ if and only if v is subcritical. Moreover, if v is subcritical then I - M is invertible and it follows from the above expressions that $\mathbb{E}(N_i(w)) = \sum_{i=1}^{d} w_i(I - M)_{ij}^{-1}$. Then assertions (i) and (ii) 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 Theorem 2.1(i), that $m_{ii} < 1$ for all $i \in [d]$.

3. When continuous time is involved

3.1. The Lamperti representation

We 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 $v_i(n_1, \ldots, n_d)$. Then this individual dies at the same time as 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 $x = (x_1, \ldots, x_d)$ with progeny distribution $v := (v_1, \ldots, v_d)$ and with reproduction rates $(\lambda_1, \ldots, \lambda_d)$. By construction, its discrete-time skeleton is a multitype branching (plane) forest, as defined in the previous section, with progeny distribution v, 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 (defined above), by $Z = (Z^{(1)}, \ldots, Z^{(d)})$ we denote the corresponding multitype branching process, that is, for $t \ge 0$ 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, that is, for $\lambda \in \mathbb{R}_+^d$, $t \ge 0$, and $x, y \in \mathbb{Z}_+^d$,

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

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.



FIGURE 3: A 2-type forest with edge lengths issued from x = (2, 2). Vertices of type 1 (respectively, 2) are represented as solid (respectively, shaded). At time t, $Z_t^{(1)} = 6$, $Z_t^{(2)} = 3$, $Z_t^{1,1} = -2$, $Z_t^{1,2} = 5$, $Z_t^{2,1} = 8$, $Z_t^{2,2} = -2$, and $M_{1,t} = 8$, $M_{2,t} = 5$.

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 parents have 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 presented in Figure 3, will play a crucial role in our continuous-time model. A more formal definition can be found in [8, Section 4.2]. The interest of these processes is the following straightforward decomposition of the branching process $Z = (Z^{(1)}, \ldots, Z^{(d)})$:

$$Z_t^{(j)} = \sum_{i=1}^d Z_t^{i,j}, \qquad j \in [d].$$
(3.1)

Our model relies on a Lamperti-type representation of these processes. According to a Lamperti representation, any one-dimensional branching process can be expressed as a Lévy process timechanged by some integral functional. In this subsection we will recall from [8] 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

$$v_i(e_i) = 0$$
 for all $i \in [d]$.

Then, let $X = (X^{(1)}, \ldots, X^{(d)})$, where $X^{(i)} = (X_t^{(i)})_{t \ge 0}$, $i \in [d]$, 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 was defined in (2.2). In particular, with the notation $X^{(i)} = (X^{i,1}, \ldots, X^{i,d})$, the process $X^{i,i}$ is a \mathbb{Z} -valued, downward skip-free, compound Poisson process, that is, $\Delta X_t^{i,i} = X_t^{i,i} - X_{t-}^{i,i} \ge -1$, $t \ge 0$, with $X_{0-} = 0$ and, for all $i \ne j$, 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 [8]; see also [7] for the case of continuous-state multitype branching processes.

Theorem 3.1. We consider a multitype branching forest with edge lengths issued from $x = (x_1, ..., x_d) \in \mathbb{Z}^d_+$ with progeny distribution $v := (v_1, ..., v_d)$ and with reproduction rates $(\lambda_1, ..., \lambda_d)$. Then the processes $Z^{i,j}$, $i, j \in [d]$, introduced in Definition 3.1 admit the following representation:

$$Z_{t}^{i,j} = \begin{cases} X_{\int_{0}^{i} Z_{s}^{(i)} \, \mathrm{d}s}^{i,j} & \text{if } i \neq j, \\ x_{i} + X_{\int_{0}^{i} Z_{s}^{(i)} \, \mathrm{d}s}^{i,j} & \text{if } i = j, \end{cases} \qquad t \ge 0,$$
(3.2)

where the processes

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

are independent \mathbb{Z}_{+}^{d} -valued compound Poisson processes with jump distribution $(\tilde{v}_{1}, \ldots, \tilde{v}_{d})$ and rates $(\lambda_{1}, \ldots, \lambda_{d})$. In particular, from (3.1) and (3.2), the multitype branching process Z admits the following representation:

$$(Z_t^{(1)}, \dots, Z_t^{(d)}) = x + \left(\sum_{i=1}^d X_{\int_0^t Z_s^{(i)} \, \mathrm{d}s}^{i,1}, \dots, \sum_{i=1}^d X_{\int_0^t Z_s^{(i)} \, \mathrm{d}s}^{i,d}\right), \qquad t \ge 0.$$
(3.3)

3.2. Further results on asymptotics of mutations

For $i \in [d]$ and $t \ge 0$, we will denote by $M_{i,t}$ the total number of mutations producing type *i* which occurred up to time *t*. The definition of this quantity is illustrated in Figure 3. Let us specify that if a parent of type $j \ne i$ gives birth to more than one child of type *i* simultaneously, then each child counts as a mutation. Let us also define a cluster of type *i* as the subtree corresponding to the lineage 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 3.2 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 3.1.

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

$$(Z_t^{(i)}, M_{i,t}) = \left(x_i + \sum_{k=1}^d X_{\int_0^t Z_u^{(k)} \, \mathrm{d}u}^{k,i}, \sum_{k=1, \, k \neq i}^d X_{\int_0^t Z_u^{(k)} \, \mathrm{d}u}^{k,i}\right), \qquad t \ge 0$$

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

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

The result follows from (3.2) in Theorem 3.1.

We now turn to the limiting behaviour of $M_{i,t}$ as t tends to ∞ . 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.

We 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 the Perron–Frobenius theory that the eigenvalues ρ_i , $i \in [d]$, of Acan be arranged so that $\rho_1 > \text{Re}(\rho_2) \ge \cdots \ge \text{Re}(\rho_d)$. Moreover, ν is subcritical, critical, or supercritical accordingly as $\rho_1 < 0$, $\rho_1 = 0$, or $\rho_1 > 0$. Then a well-known result due to [3] (see also Theorem 2 of [4, p. 206]) asserts that when ν is nonsingular and primitive, there exists a nonnegative random variable W such that, for all $i \in [d]$,

$$\lim_{t \to \infty} e^{-\rho_1 t} Z_t^{(i)} = v_i W \quad \text{a.s.}, \tag{3.4}$$

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

 \square

Proposition 3.2. Assume that v is nonsingular, primitive, and supercritical. Then, for all $i \in [d]$,

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

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

Proof. From Proposition 3.1, we derive

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

On the other hand, in the supercritical case, ρ_1 is strictly positive. Hence, from (3.4), it follows that

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

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

Under the conditions of Proposition 3.2, 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 3.2, we can compare the asymptotic behaviour of the number of mutations prior to t with that of $Z_t^{(i)}$, under \mathbb{P}_{e_i} , that is,

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

Regarding the condition $\mathbb{P}_{e_j}(W > 0) = 1$, note that Theorem 2 of [4, p. 206] also asserts that $\mathbb{P}_{e_k}(W > 0) > 0$ for some (hence, for all) $k \in [d]$ if and only if

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

where $(\xi_{i1}, \ldots, \xi_{id})$ is a random vector with law v_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 will assume that mutations are not reversible, that is, for all i = 1, ..., d-1, individuals of type i can only have children of type i or i + 1. In particular, v is not irreducible. Moreover, when giving birth, individuals of type i = 1, ..., d-1 have at least one child of type i with probability 1, and have children of type i + 1 with positive probability. These conditions can be made explicit in terms of the progeny distribution v_i as follows:

$$\begin{aligned}
\nu_i(\mathbf{k}) &> 0 \implies k_j = 0 \quad \text{for } j \notin \{i, i+1\}, \\
\sum_{\mathbf{k} \in \mathbb{Z}_+^d: \ k_i = 0} \nu_i(\mathbf{k}) &= 0 \quad \text{and} \quad \sum_{\mathbf{k} \in \mathbb{Z}_+^d: \ k_{i+1} = 0} \nu_i(\mathbf{k}) < 1.
\end{aligned} (3.5)$$

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

$$z_i := \inf\{t \ge 0 \colon Z_t^{(i)} \ge 1\}.$$

The problem of determining a general expression for the law of τ_i is quite challenging. To the best of the authors' knowledge, 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, for example, in [1], [11], [19], and [20]. Most of them provide approximations of this law using martingale convergence theorems [11] or through numerical methods involving the generating function [1]. In Proposition 3.3 we first give a relationship between the successive emergence times τ_2, τ_3, \ldots 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 3.2 we derive an approximation of the time τ_i , under \mathbb{P}_{e_1} , as the mutation rate to type k increases faster than the mutation rate to type k - 1 for all $k = 3, \ldots, i$. Then, in Corollary 3.1, we focus on a case where these laws can be made explicit.

In the following developments, we use the notation of Section 3.1 from which we recall the Lamperti representation of the multitype branching process $Z = (Z^{(1)}, \ldots, Z^{(d)})$ in terms of the compound Poisson processes $X^{(i)}$. We also introduce additional notation. For $i, j \in [d]$, we denote by $\lambda_{i,j}$ the parameter of the compound Poisson process $X^{i,j}$, that is,

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

Note that from our assumptions in (3.5), for all i = 1, ..., 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 called the mutation rate to type i + 1. For $i \geq 2$, let

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

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

(i) For i = 2,..., d, the emergence time τ_i of type i admits the following representation under P_{e1}:

$$\tau_i = \tau_{i-1} + \int_0^{\gamma_i} \frac{1}{X_s^{i-1, i-1} + Z_{\kappa_{i-1}(s)}^{i-2, i-1}} \,\mathrm{d}s, \qquad \mathbb{P}_{e_1}\text{-}a.s., \tag{3.6}$$

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

(ii) 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:

$$(\tau_i, Z_{\tau_i}^{(i-1)}) \stackrel{\mathrm{D}}{=} \left(\int_0^{\gamma_i} \frac{1}{X_s^{i-1, i-1}} \,\mathrm{d}s, 1 + X_{\gamma_i}^{i-1, i-1} \right)$$

where $\stackrel{\text{o}}{=}$ 'denotes equality in distribution.

(iii) Define $\theta_k = \int_0^{\gamma_k} 1/(X_s^{k-1,k-1}+1) \, ds$ for $k \ge 2$. Then the random variables θ_k , $k \ge 2$, are independent and, for i = 2, ..., d,

$$\mathbb{P}_{e_1}(\tau_i > t) \le \mathbb{P}\left(\sum_{k=2}^i \theta_k > t\right) \quad \text{for all } t > 0.$$
(3.7)

Proof. Since $X^{i,j}$ is identically equal to 0 whenever $j \notin \{i, i + 1\}$, then under \mathbb{P}_{e_1} , representation (3.3) admits the simpler form

$$(Z_t^{(1)},\ldots,Z_t^{(d)}) = e_1 + \left(X_{\int_0^t Z_s^{(1)} \,\mathrm{d}s}^{1,1}, X_{\int_0^t Z_s^{(2)} \,\mathrm{d}s}^{2,2} + X_{\int_0^t Z_s^{(1)} \,\mathrm{d}s}^{1,2}, \ldots, X_{\int_0^t Z_s^{(d)} \,\mathrm{d}s}^{d,d} + X_{\int_0^t Z_s^{(d-1)} \,\mathrm{d}s}^{d-1,d}\right).$$

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

$$Z_t^{(i)} = X_{\int_0^t Z_s^{(i)} \, \mathrm{d}s}^{i,i} + X_{\int_0^t Z_s^{(i-1)} \, \mathrm{d}s}^{i-1,i}, \qquad t \ge 0.$$

Since $X_0^{i,i} = 0$ for $i \ge 2$, 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,

$$\tau_i = \kappa_{i-1}(\gamma_i), \tag{3.8}$$

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

$$\kappa_{i-1}(t) = \tau_{i-1} + \int_0^t \frac{\mathrm{d}s}{Z_{\kappa_{i-1}(s)}^{(i-1)}} = \tau_{i-1} + \int_0^t \frac{\mathrm{d}s}{X_s^{i-1,\,i-1} + X_{\int_0^{\kappa_{i-1}(s)} Z_u^{(i-2)} \,\mathrm{d}u}^{(i-2)}}$$

The latter identity together with (3.8) prove (3.6).

The second part of the proposition is easily derived from the same arguments. More specifically, it follows from (3.8) and the identities

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

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

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

$$\tau_i = \sum_{k=2}^{l} \int_0^{\gamma_k} \frac{1}{X_s^{k-1, \, k-1} + X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} \, \mathrm{d}u}^{k-1, \, k-1} \, \mathrm{d}s \quad \text{a.s.}$$
(3.9)

Note that since $\kappa_{k-1}(0) = \tau_{k-1}$, then from (3.8), for all $k \ge 2$, $\int_0^{\kappa_{k-1}(0)} Z_u^{(k-2)} du = \gamma_{k-1}$, so that by the definition of γ_{k-1} ,

$$X_{\int_{0}^{k_{k-1}(0)} Z_{u}^{(k-2)} du}^{k-2, k-1} = X_{\gamma_{k-1}}^{k-2, k-1} \ge 1 \quad \text{a.s.}$$
(3.10)

Besides, since $s \mapsto X_{\int_0^{\kappa_{k-1}(s)} Z_u^{(k-2)} du}^{k-2, k-1}$ are increasing processes then (3.7) is a direct consequence of (3.9) and (3.10).

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 3.1 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 the same litter. More specifically,

assumptions (3.5) are replaced by

$$\nu_{i}(\boldsymbol{k}) > 0 \implies k_{i+1} = 0 \text{ or } 1 \text{ and } k_{j} = 0 \text{ for } j \notin \{i, i+1\},$$

$$\sum_{\boldsymbol{k} \in \mathbb{Z}_{+}^{d}: k_{i} = 0} \nu_{i}(\boldsymbol{k}) = 0 \text{ and } \sum_{\boldsymbol{k} \in \mathbb{Z}_{+}^{d}: k_{i+1} = 0} \nu_{i}(\boldsymbol{k}) < 1.$$
(3.11)

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, ..., d, grow faster.

Lemma 3.1. Assume that (3.11) holds, let $k \ge 3$, and fix $\lambda_{1,2} > 0$. Then

$$\mathbb{P}_{e_1}\left(X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} \, \mathrm{d}u}^{k-2} = 1\right) \to 1 \quad as \; \frac{\lambda_{n-2,n-1}}{\lambda_{n-1,n}} \to 0 \, for \, n = 3, \dots, k.$$

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

$$\{X_{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} \, \mathrm{d}u}^{k-2, \, k-1} = 1\} = \left\{\int_0^{\kappa_{k-1}(\gamma_k)} Z_u^{(k-2)} \, \mathrm{d}u < \gamma_{k-1}^{(1)}\right\} = \{\kappa_{k-1}(\gamma_k) < \kappa_{k-2}(\gamma_{k-1}^{(1)})\}.$$

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

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

(Note that, from our assumptions, $Z_{\tau_{k-1}}^{k-2, k-1} = 1$ and $Z_{\tau_{k-1}}^{k-2, k-1} = 2$, \mathbb{P}_{e_1} -a.s.) So from (3.8), we have showed that

$$\{X_{\int_{0}^{\kappa_{k-1}(\gamma_{k})}Z_{u}^{(k-2)} du}^{k-2, k-1} = 1\} = \{\tau_{k} < \tau_{k-1}^{(1)}\}.$$
(3.12)

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

$$\mathbb{P}_{e_1}(\tau_k \le \tau_{k-1}^{(1)}) = \int \mathbb{P}_z(\tau_k \le \tau_{k-1}^{(1)}) \mathbb{P}_{e_1}(Z_{\tau_{k-1}} \in \mathrm{d}z).$$
(3.13)

The support in the integral of (3.13) is included in the set $\{z : z_{k-1} = 1\}$, so from (3.12), (3.13), and the Lebesgue theorem of dominated convergence, all we need to prove is

$$\mathbb{P}_{z}(\tau_{k} \leq \tau_{k-1}^{(1)}) \to 1 \quad \text{as} \ \frac{\lambda_{n-2,n-1}}{\lambda_{n-1,n}} \to 0 \text{ for } n = 3, \dots, k$$
 (3.14)

for all z such that $z_{k-1} = 1$. (Note that if z is such that $z_1 = \cdots = z_{k-2} = 0$, or such that $z_k \ge 1$, then it is clear that $\mathbb{P}_z(\tau_k \le \tau_{k-1}^{(1)}) = 1$, 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 $z_i \ge 1$ for i = 1, ..., k-2. For i = 1, ..., k-1, we 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 random variables U_i are independent, and from an obvious extension of Proposition 3.3(ii), U_i has the same law as $\int_0^{\gamma_{k-i+1}} (ds/(X_s^{k-i, k-i} + z_{k-i})))$. Then set $Y_s^{(i)} := X_s^{k-i, k-i} + z_{k-i}$ and note the inclusions

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

which imply the inequality

$$\mathbb{P}\left(\frac{\gamma_{k}}{\gamma_{k-1}} \le \min\left(\frac{1}{Y_{\gamma_{k-1}}^{(2)}}, \frac{\gamma_{k-2}}{(\gamma_{k-1}Y_{\gamma_{k-2}}^{(3)})}, \dots, \frac{\gamma_{2}}{(\gamma_{k-1}Y_{\gamma_{2}}^{(k-1)})}\right)\right) \le \mathbb{P}_{z}(\tau_{k} \le \tau_{k-1}^{(1)}).$$

But when $\lambda_{n-2,n-1}/\lambda_{n-1,n} \to 0$ for n = 3, ..., k, the parameter $\lambda_{1,2} > 0$ being fixed, we necessarily have $\lim \lambda_{n-1,n} = \infty$ for n = 3, ..., k. Hence, γ_k/γ_{k-1} converges in probability to $0, 1/Y_{\gamma_{k-1}}^{(2)}$ converges in probability to $1/z_{k-2}$, and $\gamma_{n-1}/(\gamma_n Y_{\gamma_{n-1}}^{(k-n+2)})$ for n = 3, ..., k-1 converge in probability to $+\infty$. Therefore, the left-hand side of the above inequality tends to 1, which proves (3.14) and the lemma is proved.

The following theorem intuitively means that when $\lambda_{k-2, k-1}/\lambda_{k-1,k} \to 0$ for k = 3, ..., i, the emergence time τ_i , when starting from an individual of type 1, can be approximated by the sum of independent random variables $\tau_{1,2} + \cdots + \tau_{i-1,i}$, where $\tau_{k-1,k}$ is the emergence time of type k when starting from an individual of type k - 1. The assumption $\lambda_{k-2, k-1}/\lambda_{k-1,k} \to 0$ is quite adapted to several biological models such as cancer growth. Indeed, cancer is often the result of a series of successive mutations; see [10], [11], [14], and [16]. The successive mutation rates can increase very quickly. It would be interesting to study the asymptotic behaviour of τ_i when $\lambda_{k,k}/\lambda_{k+1, k+1} \to 0$, that is, when the intrinsic reproduction rates increase very quickly. This assumption also fits to the model of cancer since mutations are often more sensitive to proliferation.

Theorem 3.2. Assume that (3.11) holds. Recall the definition of θ_k in Proposition 3.3 and fix $\lambda_{1,2} > 0$. Then, under \mathbb{P}_{e_1} , for $i \ge 3$,

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

where $\stackrel{{}_{\circ}\mathbb{P}}{\rightarrow}$ denotes convergence in probability. Besides, the expectation of τ_i fulfills the following approximation:

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

Proof. Since $s \mapsto X_{\int_{0}^{k_{k-1}(s)} Z_{u}^{(k-2)} du}^{k-2, k-1}$ are increasing processes then, from (3.10), $\mathbb{P}_{e_{1}}$ -a.s. on the set $\{X_{\int_{0}^{k_{k-1}(y_{k})} Z_{u}^{(k-2)} du}^{k-2, k-1} = 1\}$, we have

$$\int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + X_{\int_0^{k_{k-1}(s)} Z_u^{(k-2)} \, \mathrm{d}u}} \, \mathrm{d}s = \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1} + 1} \, \mathrm{d}s.$$

Hence, from Lemma 3.1, it follows that, for fixed $\lambda_{1,2} > 0$ as $\lambda_{n-2,n-1}/\lambda_{n-1,n} \to 0$ for all n = 3, ..., k,

$$\left(\int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1}+1} \,\mathrm{d}s\right)^{-1} \int_0^{\gamma_k} \frac{1}{X_s^{k-1,k-1}+X_s^{k-2,k-1}} \,\mathrm{d}s \xrightarrow{\mathbb{P}} 1$$

and the first part of the theorem is easily derived from this convergence and (3.6) (or, equivalently, (3.9)).

In order to prove the second part, set

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

Then, from (3.6), $\mathbb{E}_{e_1}(\tau_i) = \sum_{k=2}^{i} \mathbb{E}_{e_1}(H_k)$, so it suffices to prove that, for all k = 2, ..., i,

$$\mathbb{E}_{e_1}(H_k) \sim \mathbb{E}(\theta_k) \quad \text{as } \frac{\lambda_{n-2,n-1}}{\lambda_{n-1,n}} \to 0 \text{ for } n = 3, \dots, k.$$
(3.15)

Observe that $\mathbb{E}_{e_1}(H_k) = \mathbb{E}(\theta_k \mathbf{1}_{A_k}) + \mathbb{E}_{e_1}(H_k \mathbf{1}_{A_k^c})$, where **1** is the indicator function. Moreover, $\mathbb{E}_{e_1}(H_k \mathbf{1}_{A_k^c}) \leq \mathbb{E}(\theta_k \mathbf{1}_{A_k^c})$. Then in order to obtain (3.15), it is enough to prove that

$$\frac{\mathbb{E}(\theta_k \mathbf{1}_{A_k^c})}{\mathbb{E}(\theta_k)} \to 0 \quad \text{as } \frac{\lambda_{n-2,n-1}}{\lambda_{n-1,n}} \to 0 \text{ for } n = 3, \dots, k.$$
(3.16)

But, for any $p, q \ge 1$ such that $p^{-1} + q^{-1} = 1$, we have, from Holder's inequality, $\mathbb{E}(\theta_k \mathbf{1}_{A_k^c}) \le \mathbb{E}(\theta_k^p)^{1/p} \mathbb{P}(A_k^c)^{1/q}$. Moreover, we clearly have $\mathbb{E}(\theta_k^p)^{1/p} \sim 1/\lambda_{k-1,k}$ as $\lambda_{k-1,k} \to \infty$. Hence, (3.16) is satisfied thanks to Lemma 3.1.

We end this section with an example of when the distribution of τ_i can be estimated 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 standard Poisson process with parameter $\lambda_{i,i}$.

Corollary 3.1. With the above assumptions, the law of τ_i can be specified as follows.

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

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

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

(ii) The expectation of τ_i is given by $\mathbb{E}_{e_{i-1}}(\tau_i) = (1/\lambda_{i-1,i}\lambda_{i-1,i-1})\ln(\lambda_{i-1,i})$. In particular, for fixed $\lambda_{1,2} > 0$, under \mathbb{P}_{e_1} , the expectation of τ_i , for $i \ge 3$, fulfills the following approximation:

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

Proof. From Proposition 3.3(ii) and the fact that $X^{i-1,i}$ is a standard Poisson process, for all $\alpha \ge 0$,

$$\mathbb{E}_{e_{i-1}}(\mathrm{e}^{-\alpha\tau_i}) = \mathbb{E}\left(\exp\left(-\alpha\int_0^{\gamma_i}\frac{1}{1+X_s^{i-1,i-1}}\,\mathrm{d}s\right)\right)$$
$$= \lambda_{i-1,i}\int_0^{+\infty}\mathbb{E}\left(\exp\left(-\alpha\int_0^x\frac{1}{1+X_s^{i-1,i-1}}\,\mathrm{d}s\right)\right)\mathrm{e}^{-\lambda_{i-1,i}x}\,\mathrm{d}x.$$
(3.17)

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 $\mathbb{E}(\exp(-\alpha \int_0^x (1/(1 + X_s^{i-1, i-1})) \, ds))$, we obtain, with the convention that $\sum_{k=0}^{-1} = 0$,

$$\mathbb{E}\left(\exp\left(-\alpha \int_{0}^{x} \frac{1}{1+X_{s}^{i-1,i-1}} \, \mathrm{d}s\right)\right)$$

$$= \sum_{n \ge 0} \mathbb{E}\left(X_{x}^{i-1,i-1} = n, \exp\left(-\alpha \left(\frac{x-J_{n}}{n+1} + \sum_{k=0}^{n-1} \frac{J_{k+1} - J_{k}}{k+1}\right)\right)\right)$$

$$= e^{-(\alpha+\lambda_{i-1,i-1}x}$$

$$+ \sum_{n \ge 1} e^{-\lambda_{i-1,i-1}x} \frac{(\lambda_{i-1,i-1}x)^{n}}{n!}$$

$$\times \int_{0 \le x_{1} \le \dots \le x_{n} \le x} \exp\left(-\alpha \left(\frac{x}{n+1} + \sum_{k=1}^{n} \frac{x_{k}}{k(k+1)}\right)\right) \frac{n!}{x^{n}} \, \mathrm{d}x_{1} \cdots \, \mathrm{d}x_{n}$$

$$= e^{-(\alpha+\lambda_{i-1,i-1})x}$$

$$+ \sum_{n \ge 1} \lambda_{i-1,i-1}^{n} \exp\left(-\left(\lambda_{i-1,i-1} + \frac{\alpha}{n+1}\right)x\right)$$

$$\times \int_{0 \le x_{1} \le \dots \le x_{n} \le x} \exp\left(-\alpha \sum_{k=1}^{n} \frac{x_{k}}{k(k+1)}\right) \, \mathrm{d}x_{1} \cdots \, \mathrm{d}x_{n}.$$

Then returning to (3.17), we obtain, with the convention that $\sum_{k=1}^{0} = 0$,

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

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

Then using again Proposition 3.3(ii), we obtain the expectation of τ_i under $\mathbb{P}_{e_{i-1}}$, after easy computations:

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

We conclude from Theorem 3.2.

4. Discussion

In this paper we addressed mutations in multitype branching processes and focused on two main points: the statistics of the total number of mutations in a multitype branching forest, and the distribution of emergence times of new mutations.

 \square

The first point regards finite multitype branching forests issued from r_i individuals of type i = 1, ..., d. In this paper, the lifetime of individuals is not involved and can be considered to be equal to 1. A mutation to type j occurs whenever some individual of type $i \neq j$ gives birth to an individual of type j. For a given progeny law, we stated the joint distribution of the total number of mutations to type j = 1, ..., d in Corollary 2.1. Our assumptions on the progeny law were very general: for each i, either families of individuals of type i become extinct almost surely (this is our assumption (A_i)) or lineages of type i may survive but do not have progeny of any other type (assumption (B_i)). Then in Theorem 2.2, we compared the total number of mutations with the total population size as the numbers of ancestors r_i tend to $+\infty$. These results have potential applications in epidemiology and more particularly in evolution of infectious diseases where the distribution of the final epidemic size is of great importance; see [6]. It may also provide a good complement to the study of emergence of infectious diseases as such; see [2].

The second point regarding emergence times of mutations concerns continuous-time multitype branching processes. In this case, each individual in the population has an exponentially distributed lifetime which is independent of all other variables and whose parameter only depends on its type. Unlike for the total number of mutations, in general, the distribution of the first time at which a mutation occurs cannot be made explicit in terms of the characteristics of the branching process. One has to make strong assumptions. In this paper, we considered multitype populations where only mutations from type i - 1 to type i can occur. We call mutation rate to type *i* the parameter $\lambda_{i-1,i}$ of the exponentially distributed time at which a generic individual of type i - 1 will give birth to an individual of type i. In Theorem 3.2, the law of the emergence time of type i was then estimated when the ratios $\lambda_{k-2,k-1}/\lambda_{k-1,k}$ tend to 0 for $k = 3, \ldots, i$. This model fits to the widely accepted idea that progression of cancer involves successive accumulation of mutations; see [10], [11], and [16]. In [11], the process $e^{-\lambda_i^*}Z_t^{(i)}$, for some rate λ_i^* , is approximated for large times t by its limit at ∞ . This provides an exact formulation of the Laplace transform of the emergence time τ_i and leads to an estimation of its distribution function. In this paper, considering $Z^{(i)}$ as a time-changed compound Poisson process allowed us to not make any approximation on the branching process itself. We derived an integral representation for τ_i . When mutation rates $\lambda_{k-1,k}$ are growing very quickly with k, that is, when the ratio $\lambda_{k-2, k-1}/\lambda_{k-1, k}$ is very small for each $k = 3, \ldots, i$, we obtained an even more exploitable representation of τ_i . An interesting counterpart to this paper would consist of the same study but with the proliferation rates $\lambda_{k,k}$ increasing.

References

- ALEXANDER, H. K. (2013). Conditional distributions and waiting times in multitype branching processes. Adv. Appl. Prob. 45, 692–718.
- [2] ANTIA, R., REGOES, R. R., KOELLA, J. C. AND BERGSTROM, C. T. (2003). The role of evolution in the emergence of infectious diseases. *Nature* 426, 658–661.
- [3] ATHREYA, K. B. (1968). Some results on multitype continuous time Markov branching processes. Ann. Math. Statist. 39, 347–357.
- [4] ATHREYA, K. B. AND NEY, P. E. (1972). Branching Processes. Springer, New York.
- [5] BERTOIN, J. (2009). The structure of the allelic partition of the total population for Galton-Watson processes with neutral mutations. *Ann. Prob.* **37**, 1502–1523.
- [6] BLACK, A. J. AND ROSS, J. V. (2015). Computation of epidemic final size distributions. J. Theoret. Biol. 367, 159–165.
- [7] CABALLERO, M. E., PÉREZ GARMENDIA, J. L. AND URIBE BRAVO, G. (2015). Affine processes on $\mathbb{R}^n_+ \times \mathbb{R}^n$ and multiparameter time changes. *Preprint*. Available at https://arxiv.org/abs/1501.03122.
- [8] CHAUMONT, L. (2015). Breadth first search coding of multitype forests with application to Lamperti representation. In *In Memoriam Marc Yor—Séminaire de Probabilités XLVII* (Lecture Notes Math. 2137), Springer, Cham, pp. 561–584.

- [9] CHAUMONT, L. AND LIU, R. (2016). Coding multitype forests: application to the law of the total population of branching forests. *Trans. Amer. Math. Soc.* 368, 2723–2747.
- [10] DURRETT, R. (2013). Population genetics of neutral mutations in exponentially growing cancer cell populations. Ann. Appl. Prob. 23, 230–250.
- [11] DURRETT, R. AND MOSELEY, S. (2010). Evolution of resistance and progression to disease during clonal expansion of cancer. *Theoret. Pop. Biol.* 77, 42–48.
- [12] GUT, A. (2009). Stopped Random Walks: Limit Theorems and Applications, 2nd edn. Springer, New York.
- [13] HAENO, H., IWASA, Y. AND MICHOR, F. (2007). The evolution of two mutations during clonal expansion. *Genetics* 177, 2209–2221.
- [14] IWASA, Y., MICHOR, F., KOMAROVA, N. L. AND NOWAK, M. A. (2005). Population genetics of tumor suppressor genes. J. Theoret. Biol. 233, 15–23.
- [15] KESTEN, H. AND MALLER, R. A. (1996). Two renewal theorems for general random walks tending to infinity. *Prob. Theory Relat. Fields* 106, 1–38.
- [16] LOEB, L. A. (2001). A mutator phenotype in cancer. Perspectives Cancer Res. 61, 3230–3239.
- [17] PÉNISSON, S. (2016). Beyond the Q-process: various ways of conditioning multitype Galton-Watson processes. ALEA Latin. Amer. J. Prob. Math. Statist. 13, 223–237.
- [18] SENETA, E. (2006). Non-negative matrices and Markov chains. Springer, New York.
- [19] SERRA, M. C. (2006). On the waiting time to escape. J. Appl. Prob. 43, 296-302.
- [20] SERRA, M. C. AND HACCOU, P. (2007). Dynamics of escape mutants. Theoret. Pop. Biol. 72, 167-178.
- [21] TAïB, Z. (1992). Branching Processes and Neutral Evolution (Lecture Notes Biomath. 93). Springer, Berlin.
- [22] TAKÁCS, L. (1961). The probability law of the busy period for two types of queuing processes. Operat. Res. 9, 402–407.