

# Binary Galton–Watson trees with mutations

Qiao Huang\*      Nicolas Privault†

December 31, 2025

## Abstract

We consider a multitype Galton–Watson process that allows for the mutation and reversion of individual types in discrete and continuous time. In this setting, we explicitly compute the time evolution of quantities such as the mean and distributions of different types. This allows us in particular to estimate the proportions of different types in the long run, as well as the distribution of the first time of occurrence of a given type as the tree size or time increases. Our approach relies on the recursive computation of the joint distribution of types conditionally to the value of the total progeny. In comparison with the literature on related multitype models, we do not rely on approximations.

*Keywords:* Evolutionary branching, multitype branching process, population dynamics, Galton–Watson, mutation, reversion.

*Mathematics Subject Classification (2020):* 34-04, 05C05, 60J80, 60J85.

## 1 Introduction

Multitype Galton–Watson processes have been used in population genetics and evolutionary biology to model the propagation and extinction of mutant types. In [INM06] and [DM10], the mutation of a type- $i$  cancer cell mutates into a type  $i + 1$  cell has been modeled using a continuous-time process that branches at an exponential rate depending on  $i \geq 0$ . In [SMJV13], evolutionary branching processes modeling subpopulations with different traits or genotypes have been analyzed under small mutational step sizes. More recently, the diffusion limit of Galton–Watson branching processes modeling allele types has been analyzed in [BW18]. This analysis relies mainly on the classical literature on birth-death and Galton–Watson processes, see e.g. [Ken48], [Ott49], [Har63], [AN72], and [BS84], which is used to

---

\*School of Mathematics, Southeast University, Nanjing 211189, P.R. China, and School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371. [qiao.huang@seu.edu.cn](mailto:qiao.huang@seu.edu.cn)

†School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371. [nprivault@ntu.edu.sg](mailto:nprivault@ntu.edu.sg)

derive the distribution and generating function properties of the progeny of random trees or branching processes.

In this paper, we consider multitype random binary trees in which every node bears an integer type  $j \geq 0$  and may generate either no offspring, or two offsprings, one with type 0 and one with type  $j + 1$ . Our analysis covers both the discrete-time case, in which two offsprings are generated with probability  $p \in (0, 1)$  and no offspring is generated with probability  $q := 1 - p$ , and the continuous-time setting, which yields a multitype Galton—Watson process.

Figure 1 presents a sample of the corresponding multitype branching process in discrete time.

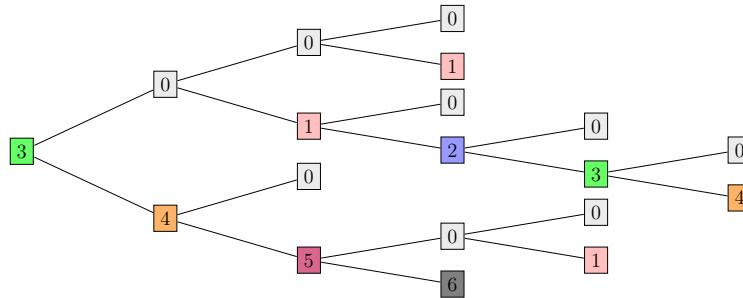


Figure 1: Marked random tree sample started from the initial type  $j = 3$ .

In terms of population genetics, such trees provide a way to model mutation and reversion, by considering “wild type” individuals with type 0, and “mutant” individuals with type  $j \geq 1$  which have been the object of  $j \geq 1$  mutations. In this setting, wild type 0 individuals can have offspring of both wild type 0 and mutant type 1, whereas mutants of type  $j \geq 1$  can have offsprings of wild type 0 (revertants), or mutant type  $j + 1$ . See for example [AO21] and references therein for the study of related population models in the framework of evolutionary rescue.

In this model, type 0 individuals follow the same Galton—Watson dynamics as in [INM06], [DM10], and the interbranching times are also exponentially distributed. However, higher type transitions are modeled differently using a marked tree, and the growth of population types over times can be subexponential, see Figure 6, whereas it is exponential in Theorem 4 of [DM10]. In addition, due to model complexity, the computation of waiting times distributions in [INM06], [DM10] relies on approximation formulas.

Our main results are presented in Sections 2 and 3 and are proved in Appendices A

and  $\mathbf{B}$ , respectively in discrete and continuous time. After recalling the computation of the distribution of tree progenies in Propositions 2.1 and 3.1, we derive recursive expressions for the distribution of any finite vector  $(X^{(1)}, \dots, X^{(n)})$  of type counts given the size of the random tree, see Theorems 2.2 and 3.2, respectively in discrete and continuous time.

In particular, we provide closed form and recursive expressions for the distribution of type counts in both discrete and continuous time, which allows us to determine the evolution of quantities such as:

- expected counts and proportions of types, see Figures 2, 3, 6 and 9, and
- the distributions of the first occurrence times of a given type count, see Figures 7 and 8.

We also derive identities for the expectation of product functionals on random trees, which in turn yield integrability conditions for generating functionals, see Sections 2.3 and 3.3.

In Figure 2 we display the computed values of the conditional mean proportions of types as the size of the discrete-time tree increases. Figure 9 displays the continuous-time evolution of those proportions. We note in particular that the (wild) type 0 remains predominant in discrete time, see Figure 2, whereas in continuous time it is the initial type  $j$  which remains predominant over time in Figure 9. Figures 7 and 8 present the tail distribution functions and probability density functions of the occurrence times of given types.

The closed-form expressions of Theorems 2.2 and 3.2 are then applied to the computation of the expectation of product functionals on random trees in Proposition 2.5 and Corollaries 2.6, 2.7 in discrete time, and in Propositions 3.3, 3.4 and Corollary 3.5 in continuous time. In particular, Corollaries 2.7 and 3.5 yield sufficient conditions for the integrability of random product functionals involving marks. Such results are applicable to problems where the generation of random trees is used in Monte Carlo integration, see for example [HP25], [HP26] for an application to Monte Carlo methods for differential equations.

The recurrence formulas proved in Theorems 2.2 and 3.2 are implemented in Mathematica and Python notebooks which are used to produce Figures 1-4 and 6-9, and are available at

<https://github.com/nprivaul/branching/>

in discrete and continuous time.

All analytical results are confirmed by Monte Carlo simulations that can also be run in the above notebooks.

## 2 Discrete-time setting

In what follows, we let  $\mathbb{N} = \{0, 1, 2, \dots\}$ .

### 2.1 Marked Galton–Watson process

We consider a branching chain in which every individual has either no offspring with probability  $q$ , or two offsprings with probability  $p$ . For this, let  $(\xi_{n,k})_{n,k \geq 1}$  denote a family of independent  $\{0, 2\}$ -valued Bernoulli random variables with the common distribution

$$q = \mathbb{P}(\xi_{n,k} = 0) \quad \text{and} \quad p = \mathbb{P}(\xi_{n,k} = 2), \quad n, k \geq 1,$$

with  $p + q = 1$  and  $0 < p, q < 1$ , where  $\xi_{n,k}$  represents the number of offsprings of the  $k$ -th individual of generation  $n - 1$ , see e.g. [Har63], [AN72].

In this framework, the branching chain  $(Z_n)_{n \geq 0}$  is recursively defined as

$$Z_0 = 1, \quad Z_n = \sum_{k=1}^{Z_{n-1}} \xi_{n,k}, \quad n \geq 1, \quad (2.1)$$

and represents the population size at generation  $n \geq 0$ . We let

$$S_\infty^{\neq 0} := \frac{1}{2} \sum_{k=1}^{\infty} Z_k$$

denote the total count of nodes with non-zero types, excluding the initial node, i.e.  $1 + 2S_\infty^{\neq 0}$  represents the total progeny of the chain  $(Z_n)_{n \geq 0}$ . Note that  $S_\infty^{\neq 0}$  is also equal to the number of nodes with type 0 excluding the initial node, since each node with non-zero type has a co-twin with type 0.

Using the sequence  $(\xi_{n,k})_{n,k \geq 1}$  we construct a marked random binary tree  $\mathcal{T}$  in which a node  $k \in \{1, \dots, Z_{n-1}\}$  at generation  $n - 1$  yields either two branches if  $\xi_{n,k} = 2$ , or zero branch if  $\xi_{n,k} = 0$ . In addition, the nodes of  $\mathcal{T}$  receive marks that represent individual types, as described below.

- i) The initial node has the type  $j \geq 0$ ;
- ii) if a node of type  $i \in \mathbb{N}$  splits, its two offsprings respectively receive the types 0 and  $i + 1$ ;

as shown in Figure 1. Proposition 2.1 recovers the distribution of the number of vertices of the random binary tree  $\mathcal{T}$  using classical results of [Ott49], and is proved in Appendix A for

completeness. In what follows, we let

$$C_n = \frac{1}{n+1} \binom{2n}{n}, \quad n \geq 0,$$

denote the  $n$ -th Catalan number (see [Aig07]), which represents the number of different rooted binary trees with  $n+1$  leaves.

**Proposition 2.1.** *The distribution of the count  $S_\infty^{\neq 0}$  of nodes with non-zero types is given by*

$$\mathbb{P}(S_\infty^{\neq 0} = n) = q(pq)^n C_n, \quad n \geq 0, \quad (2.2)$$

with the probability generating function

$$\mathbb{E}[\delta^{S_\infty^{\neq 0}}] = \frac{1 - \sqrt{1 - 4pq\delta}}{2p\delta}, \quad |\delta| \leq 1/(4pq), \quad (2.3)$$

and we have  $\mathbb{P}(S_\infty^{\neq 0} < \infty) = 1$  if  $p \leq 1/2$ .

In addition, it follows from (2.3) that if  $p < 1/2$ , we have

$$\mathbb{E}[S_\infty^{\neq 0}] = p/(q-p). \quad (2.4)$$

## 2.2 Conditional type distribution

We let  $X^{(k)}$  denote the count of types equal to  $k \geq 1$  in the random tree  $\mathcal{T}$ , excluding the initial node, with

$$X^{(k)} = 0 \text{ for } k > S_\infty^{\neq 0}.$$

For example, in Figure 1 with  $j = 3$  we have  $S_\infty^{\neq 0} = 9$ , and

$$X^{(1)} = 3, \quad X^{(2)} = 1, \quad X^{(3)} = 1, \quad X^{(4)} = 2, \quad X^{(5)} = X^{(6)} = 1.$$

We also let  $\mathbb{P}_j$ , resp.  $\mathbb{E}_j$ , denote conditional probabilities and expectations given that  $\mathcal{T}$  is started from the initial type  $j \in \mathbb{N}$ .

In Theorem 2.2, which is proved in Appendix A, we compute recursively the conditional type distribution of  $(X^{(1)}, \dots, X^{(n)})$  given that their summation equals  $S_\infty^{\neq 0}$  and  $X^{(k)} = 0$  for all  $k > n$ , and show that it does not depend on  $p, q$ . In what follows, we use the notation  $\mathbf{1}_A$  to denote the indicator function taking the value 1, resp. 0 when condition A is satisfied, resp. not satisfied.

**Theorem 2.2.** For  $j \geq 0$  and  $n \geq 1$ , the conditional type distribution

$$\mathbb{P}_j(X^{(1)} = m_1, \dots, X^{(n)} = m_n \mid S_\infty^{\neq 0} = m_1 + \dots + m_n) = \frac{b_j(m_1, \dots, m_n)}{C_{m_1 + \dots + m_n}}, \quad (2.5)$$

$m_1, \dots, m_n \geq 0$ , satisfies the recursive relation

$$b_j(m_1, \dots, m_n) = \sum_{l=1}^{n-j} \mathbf{1}_{\{m_{j+l} > m_{j+l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{j < i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0(m_1^k, \dots, m_n^k) \quad (2.6)$$

where  $m_{n+1} := 0$  in the last indicator  $\mathbf{1}_{\{m_n > m_{n+1}\}}$ ,

$$b_j(\emptyset) = 1, \quad b_j(m_1, \dots, m_{n-1}, 0) = b_j(m_1, \dots, m_{n-1}),$$

and

$$b_j(m_1, \dots, m_n) = 0 \text{ if } 1 \leq n < j \text{ and } m_1 + \dots + m_n \geq 1.$$

From Theorem 2.2, the conditional type distribution (2.5) can be computed first by applying the recursion (2.6) to  $j = 0$ , as

$$b_0(m_1, \dots, m_n) = \sum_{l=1}^n \mathbf{1}_{\{m_l > m_{l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0(m_1^k, \dots, m_n^k),$$

and then applying again (2.6) to  $j \geq 1$ . Note also that when  $m_1 + \dots + m_n < n - j$ , the summation range in (2.6) is empty, whence  $b_j(m_1, \dots, m_n) = 0$ . In addition, setting

$$\mathbb{K}_{j,n} := \begin{cases} \{\emptyset\} \cup \{(m_1, \dots, m_n) : m_1 \geq \dots \geq m_n \geq 1\}, & j = 0, n \geq 0, \\ \{(m_1, \dots, m_n) : m_1 \geq \dots \geq m_j \geq 0, m_j + 1 \geq m_{j+1} \geq \dots \geq m_n \geq 1\}, & 1 \leq j < n, \\ \{(m_1, \dots, m_j) : m_1 = \dots = m_j = 0\}, & j = n \geq 1, \\ \emptyset, & 1 \leq n < j, \end{cases}$$

for  $j \geq 0$ ,  $1 \leq n \leq m + j$ ,  $m \geq 1$ , and any weight function  $f_n : \mathbb{N}^n \rightarrow \mathbb{R}$ , we have

$$\mathbb{E}_j[f_n(X^{(1)}, \dots, X^{(n)}) \mathbf{1}_{\{X^{(1)} + \dots + X^{(n)} = m\}} \mid S_\infty^{\neq 0} = m] = \sum_{\substack{(m_1, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_n = m}} \frac{b_j(m_1, \dots, m_n)}{C_m} f_n(m_1, \dots, m_n).$$

In particular, the following corollary provides a way to solve the recursion (2.6) for the computation of mean type counts given the value of  $S_\infty^{\neq 0}$ .

**Corollary 2.3.** *The mean count of type  $l$  individuals given the value of  $S_\infty^{\neq 0} = m$  after starting from type  $j$  is given by*

$$\mathbb{E}_j[X^{(l)} \mid S_\infty^{\neq 0} = m] = \frac{1}{C_m} \mathbf{1}_{\{0 < l-j \leq m\}} \frac{l-j+1}{m+1} \binom{2m-l+j}{m} + \frac{1}{C_m} \mathbf{1}_{\{m \geq l\}} \binom{2m-l}{m+1}, \quad (2.7)$$

$l, m \geq 1, j \geq 0$ .

The proof of Corollary 2.3, which is given in Appendix A, also shows that

$$\mathbb{E}_j[X^{(l)}] = \mathbf{1}_{\{j < l\}} p^{l-j} + \frac{p^{l+1}}{q-p}, \quad j \geq 0, l \geq 1,$$

and

$$\sum_{k \geq 1} \mathbb{E}_j[X^{(k)}] = \frac{p}{q-p},$$

which recovers (2.4), with

$$\frac{\mathbb{E}_j[X^{(l)}]}{\sum_{k \geq 1} \mathbb{E}_j[X^{(k)}]} = \mathbf{1}_{\{j < l\}} (q-p) p^{l-j-1} + p^l, \quad j \geq 0, l \geq 1.$$

As a consequence of Corollary 2.3, the conditional mean proportions of non-zero types

$$\frac{1}{m} \mathbb{E}_j[X^{(l)} \mid S_\infty^{\neq 0} = m], \quad m \geq 1, \quad (2.8)$$

satisfy

$$\lim_{m \rightarrow \infty} \frac{1}{m} \mathbb{E}_j[X^{(l)} \mid S_\infty^{\neq 0} = m] = \frac{1}{2^l}, \quad j \geq 0, l \geq 1.$$

Figure 2 displays the evolution of computed values of the conditional mean proportions (2.8) of non-zero types for  $m = 1, \dots, 12$ , after starting from the initial types  $j = 0, 1, 2, 3$ .

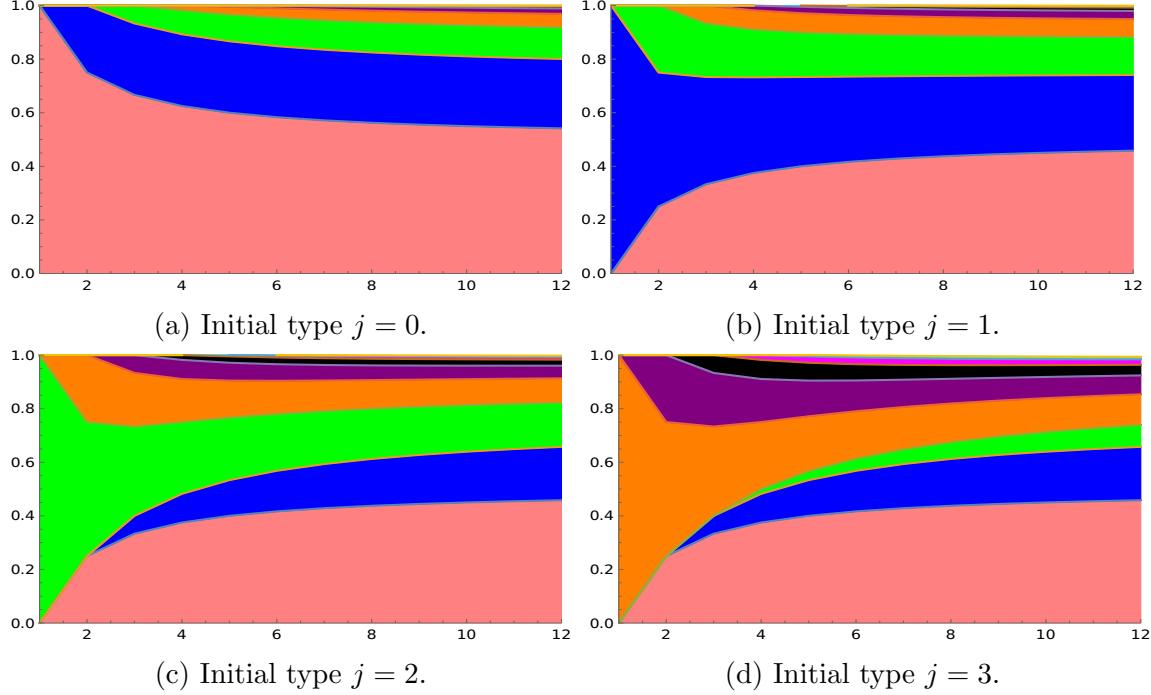


Figure 2: Conditional average type proportions (2.7) given the values of  $S_\infty^{\neq 0}$  in abscissa.

The color coding of types used in Figures 1-3 and 6-9 is shown below.

$$\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \quad (2.9)$$

The expected values of the conditional proportions (2.8) of non-zero types are computed as functions of  $p \in (0, 1/2)$  in Corollary 2.4. Here,

$$B(z; a, b) := \int_0^z u^{a-1} (1-u)^{b-1} du$$

denotes the incomplete beta function.

**Corollary 2.4.** *The conditional proportion of type  $k$  individuals after starting from the initial type  $j$  is given by*

$$\mathbb{E}_j \left[ \frac{X^{(k)}}{S_\infty^{\neq 0}} \mid S_\infty^{\neq 0} \geq 1 \right] = \frac{q}{p} B(p; k+1, -1) + \frac{q}{p} \mathbf{1}_{\{k>j\}} \left( (k+1-j) B(p; k-j, 0) - \frac{p^{k-j}}{q} \right), \quad (2.10)$$

$$k \geq 1, j \geq 0.$$

The proof of Corollary 2.4 is given in Appendix A, and the average proportions (2.10) are plotted in Figure 3 for  $p \in [0, 1/2)$ , after starting from the initial types  $j = 0, 1, 2, 3$ .

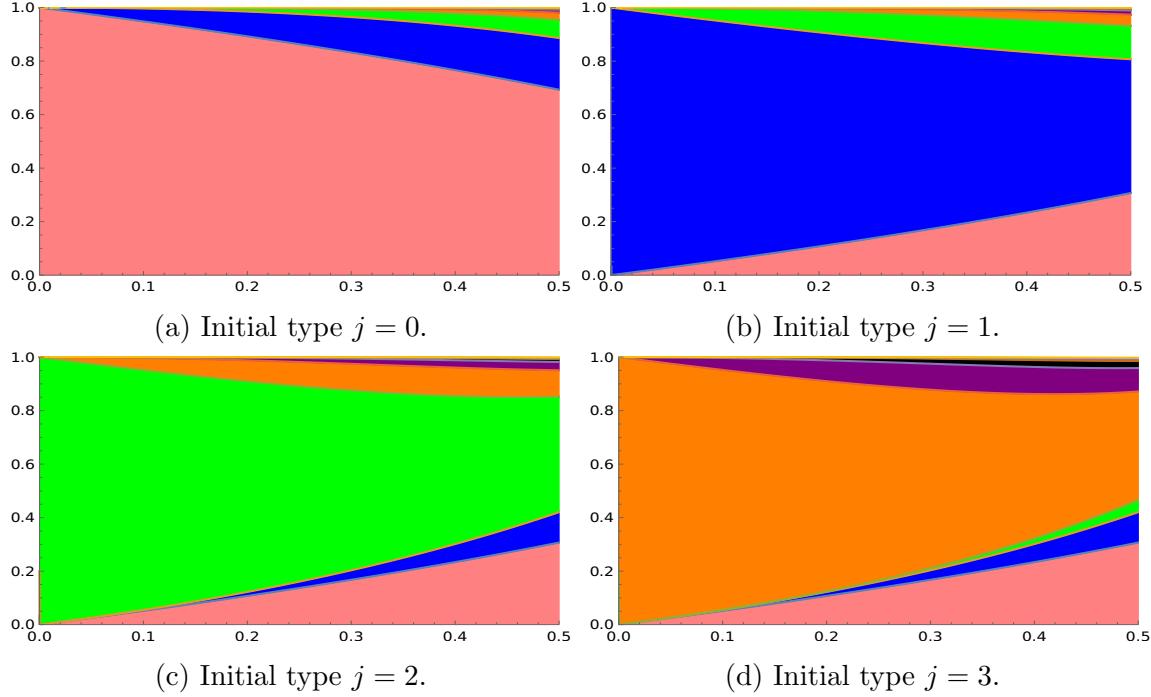


Figure 3: Average type proportions (2.10) as functions of  $p \in [0, 1/2]$ .

Corollary 2.4 also yields the limiting values of the mean proportions (2.10) as  $p$  tends to  $1/2$ , i.e.

$$\begin{aligned} \lim_{p \rightarrow 1/2} \mathbb{E}_j \left[ \frac{X^{(k)}}{S_\infty^{\neq 0}} \mid S_\infty^{\neq 0} \geq 1 \right] \\ = B \left( \frac{1}{2}, k+1, -1 \right) + \mathbf{1}_{\{k > j\}} \left( (k+1-j)B \left( \frac{1}{2}, k-j, 0 \right) - 2^{j-k+1} \right), \end{aligned} \quad (2.11)$$

as illustrated in Figure 4.

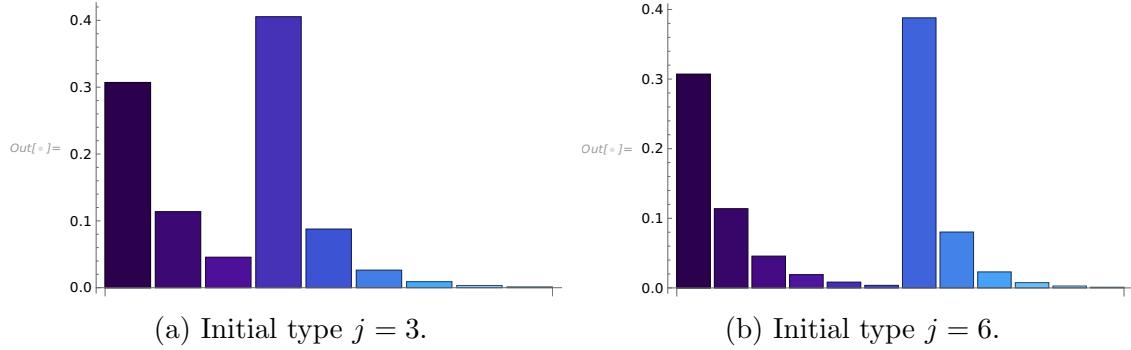


Figure 4: Limiting distributions (2.11) for  $p = 1/2$ .

## 2.3 Generating functions

Let  $F_0(p, r) = 1$ ,  $p, r \geq 0$ , and for  $n \geq 1$ ,

$$F_n(p, r) = \frac{r}{np+r} \binom{np+r}{n} = \frac{r}{n} \binom{np+r-1}{n-1} = \frac{r\Gamma(np+r)}{\Gamma(n+1)\Gamma((p-1)n+r+frac)}, \quad p, r \geq 0,$$

denote the generalized Catalan numbers, or two-parameter Fuss–Catalan numbers, see [Mlo10].

Then, the  $n$ -th Catalan number is given by

$$C_n = F_n(2, 1) = \frac{1}{n+1} \binom{2n}{n}, \quad n \geq 0.$$

In Proposition 2.5 we derive a closed-form conditional generating function expression using Fuss–Catalan numbers, which is proved in Appendix A.

**Proposition 2.5.** *For any  $\gamma \geq -1$  and  $m \geq 1$ , we have*

$$\mathbb{E}_0 \left[ \prod_{k=1}^{S_\infty^{\neq 0}} \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \mid S_\infty^{\neq 0} = m \right] = \frac{F_m(\gamma+2, \gamma+1)}{F_m(2, 1)}. \quad (2.12)$$

By differentiation of the generating function (2.12), we have

$$\mathbb{E}_0 \left[ \sum_{k=1}^{S_\infty^{\neq 0}} \frac{X^{(k)}}{k} \mid S_\infty^{\neq 0} = m \right] = \frac{\partial}{\partial \gamma} \frac{F_m(\gamma+2, \gamma+1)}{F_m(2, 1)} \Big|_{\gamma=0} = \sum_{j=1}^m \frac{m+1}{m+j},$$

hence

$$\lim_{m \rightarrow \infty} \frac{1}{m} \mathbb{E}_0 \left[ \sum_{k=1}^{S_\infty^{\neq 0}} \frac{X^{(k)}}{k} \mid S_\infty^{\neq 0} = m \right] = \lim_{m \rightarrow \infty} \sum_{j=1}^m \frac{1}{m+j} = \log 2.$$

The following corollary generalizes (2.3) from  $\gamma = 0$  to any  $\gamma \geq -1$ .

**Corollary 2.6.** *The generating function*

$$G_0^\gamma(\delta) := \mathbb{E}_0 \left[ \delta^{S_\infty^{\neq 0}} \prod_{k=1}^{S_\infty^{\neq 0}} \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \right]$$

solves the equation

$$(1 - \delta p G_0^\gamma(\delta))^{1+\gamma} G_0^\gamma(\delta) = q. \quad (2.13)$$

*Proof.* From Propositions 2.1 and 2.5, we have

$$G_0^\gamma(\delta) := \mathbb{E}_0 \left[ \delta^{S_\infty^{\neq 0}} \prod_{k=1}^{S_\infty^{\neq 0}} \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \right]$$

$$\begin{aligned}
&= \sum_{m=0}^{\infty} \delta^m \mathbb{P}(S_{\infty}^{\neq 0} = m) \mathbb{E}_0 \left[ \prod_{k=1}^{S_{\infty}^{\neq 0}} \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \mid S_{\infty}^{\neq 0} = m \right] \\
&= q \sum_{m=0}^{\infty} (pq\delta)^m F_m(\gamma + 2, \gamma + 1) \\
&= \frac{1}{p\delta} \Phi_{\gamma}^{-1}(pq\delta)
\end{aligned}$$

by Lemma A.2 below, where  $\Phi_{\gamma}^{-1}$  the inverse function of

$$\Phi_{\gamma}(w) := w(1-w)^{1+\gamma}, \quad w \in \mathbb{C},$$

which yields (2.13).  $\square$

For example, taking  $\gamma = 1$ , (2.13) becomes a cubic equation that can be solved in closed form as

$$\begin{aligned}
\mathbb{E}_0 \left[ \delta^{S_{\infty}^{\neq 0}} \prod_{k=1}^{S_{\infty}^{\neq 0}} \left(1 + \frac{1}{k}\right)^{X^{(k)}} \right] &= \frac{2}{3p\delta} - \frac{1}{3 \times 2^{2/3} (27\delta^4 p^4 q - 2\delta^3 p^3 + 3\sqrt{3\delta^7 p^7 q (27\delta p q - 4)})^{1/3}} \\
&\quad - \frac{((27\delta^4 p^4 q - 2\delta^3 p^3 + 3\sqrt{3\delta^7 p^7 q (27\delta p q - 4)})^{1/3}}{6 \times 2^{1/3} \delta^2 p^2}.
\end{aligned}$$

As a consequence of Proposition 2.5, we also obtain the following integrability criterion for product functionals.

**Corollary 2.7.** *Let  $\delta > 0$  and  $\gamma \geq -1$ , and let  $(\sigma(k))_{k \geq 0}$  be a real sequence such that*

$$0 \leq \sigma(0) < \frac{(1+\gamma)^{1+\gamma}}{(2+\gamma)^{2+\gamma} pq\delta}, \quad \text{and} \quad 0 \leq \sigma(k) \leq \left(1 + \frac{\gamma}{k}\right) \delta, \quad k \geq 1. \quad (2.14)$$

*Then, we have*

$$\mathbb{E}_0 \left[ \sigma(0)^{S_{\infty}^{\neq 0}} \prod_{k=1}^{S_{\infty}^{\neq 0}} \sigma(k)^{X^{(k)}} \right] < \infty.$$

*Proof.* By Proposition 2.5, we have

$$\begin{aligned}
\mathbb{E}_0 \left[ (\sigma(0)\delta)^{S_{\infty}^{\neq 0}} \prod_{k=1}^{S_{\infty}^{\neq 0}} \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \right] &= \sum_{m=0}^{\infty} (\sigma(0)\delta)^m \mathbb{E}_0 \left[ \prod_{k=1}^m \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \mid S_{\infty}^{\neq 0} = m \right] \mathbb{P}_0(S_{\infty}^{\neq 0} = m) \\
&= q \sum_{m=0}^{\infty} (pq)^m (\sigma(0)\delta)^m C_m \mathbb{E}_0 \left[ \prod_{k=1}^m \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \mid S_{\infty}^{\neq 0} = m \right] \\
&= q \sum_{m=0}^{\infty} (pq\sigma(0)\delta)^m F_m(\gamma + 2, \gamma + 1).
\end{aligned} \quad (2.15)$$

From the relation  $\Gamma(x + \alpha)/\Gamma(x) = O(x^\alpha)$ , we obtain

$$\begin{aligned} \limsup_{m \rightarrow \infty} \frac{F_{m+1}(\gamma + 2, \gamma + 1)}{F_m(\gamma + 2, \gamma + 1)} &= \limsup_{m \rightarrow \infty} \frac{\Gamma((2 + \gamma)(m + 1) + \gamma + 1)\Gamma((1 + \gamma)(m + 1))}{(m + 2)\Gamma((1 + \gamma)(m + 2))\Gamma((2 + \gamma)m + \gamma + 1)} \\ &= \limsup_{m \rightarrow \infty} \frac{((2 + \gamma)m + \gamma + 1)^{2+\gamma}}{(m + 2)((1 + \gamma)(m + 1))^{1+\gamma}} \\ &= \frac{(2 + \gamma)^{2+\gamma}}{(1 + \gamma)^{1+\gamma}}, \end{aligned}$$

hence under (2.14) we have

$$\limsup_{m \rightarrow \infty} \frac{F_{m+1}(\gamma + 2, \gamma + 1)}{F_m(\gamma + 2, \gamma + 1)} < \frac{1}{pq\sigma(0)\delta},$$

and the series (2.15) converges absolutely. We conclude from the inequality

$$\mathbb{E}_0 \left[ \sigma(0)^{S_\infty^{\neq 0}} \prod_{k=1}^{S_\infty^{\neq 0}} \sigma(k)^{X^{(k)}} \right] \leq \mathbb{E}_0 \left[ (\sigma(0)\delta)^{S_\infty^{\neq 0}} \prod_{k=1}^{S_\infty^{\neq 0}} \left(1 + \frac{\gamma}{k}\right)^{X^{(k)}} \right] = G_0^\gamma(\sigma(0)\delta),$$

that follows from (2.14).  $\square$

### 3 Continuous-time setting

#### 3.1 Marked binary branching process

In this section, we consider an age-dependent continuous-time random tree  $\mathcal{T}_t$ ,  $t > 0$ , in which the lifetimes of branches are independent and identically distributed via a common exponential density function  $\rho(t) = \lambda e^{-\lambda t}$ ,  $t \geq 0$ , with parameter  $\lambda > 0$ . In addition to a type  $j \in \mathbb{N}$ , a label  $\mathbf{k}$  in

$$\mathbb{K} := \{\emptyset\} \cup \bigcup_{n \geq 1} \{1, 2\}^n,$$

is attached to every branch, as follows.

- At time 0 we start from a single branch with label  $\mathbf{k} = \emptyset$  and initial type  $j \in \mathbb{N}$ . At the end of its lifetime  $T_\emptyset$ , this branch yields either:
  - no offspring if  $T_\emptyset \geq t$ ;
  - two independent offsprings with respective labels (1), (2) and respective types 0,  $j + 1$  if  $T_\emptyset < t$ .

- At generation  $n \geq 1$ , a branch having a parent label  $\mathbf{k}_- := (k_1, \dots, k_{n-1})$  and type  $i \in \mathbb{N}$  starts at time  $T_{\mathbf{k}_-}$  and has the lifetime  $\tau_{\mathbf{k}}$ . At the end of its lifetime  $T_{\mathbf{k}} := T_{\mathbf{k}_-} + \tau_{\mathbf{k}}$ , this branch yields either:
  - no offspring if  $T_{\mathbf{k}} \geq t$ ;
  - two independent offsprings with respective labels  $(\mathbf{k}, 1) = (k_1, \dots, k_n, 1)$  and  $(\mathbf{k}, 2) = (k_1, \dots, k_n, 2)$ , and respective types 0,  $i + 1$  if  $T_{\mathbf{k}} < t$ ;

see Figure 5. In particular, when a branch  $\mathbf{k}$  with type  $i \geq 0$  splits, its two offsprings are respectively marked by 0 and  $i + 1$ .

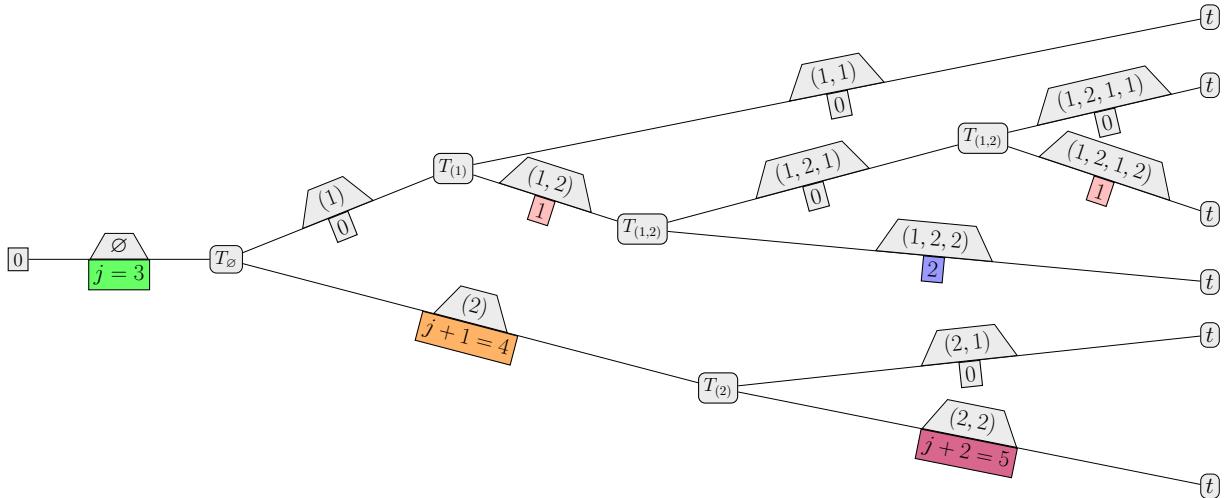


Figure 5: Sample of the marked random tree  $\mathcal{T}_t$ ,  $t > 0$ , started from the initial type  $j = 3$ .

We refer to e.g. [Ken48, Eq. (8) page 3], [Har63, Example 13.2 page 112], and [AN72, Example 5 page 109] or [INM06, Equation (5)] for the following result, whose proof is given in Appendix B for completeness.

**Proposition 3.1.** *The distribution of the count  $N_t$  of nodes with non-zero types in  $\mathcal{T}_t$ ,  $t \geq 0$ , excluding the initial node, is given by*

$$\mathbb{P}(N_t = m) = e^{-\lambda t} (1 - e^{-\lambda t})^m, \quad m \geq 0, \quad (3.1)$$

and probability generating function

$$G_t(\delta) = \mathbb{E}[\delta^{N_t}] = \frac{e^{-\lambda t}}{1 - (1 - e^{-\lambda t})\delta}, \quad t \geq 0. \quad (3.2)$$

In addition, it follows from (3.2) that  $\mathbb{E}[N_t] = e^{\lambda t} - 1$ ,  $t \geq 0$ . Note also that  $N_t$  equals the number of nodes with type 0 excluding the initial node, since each node with non-zero type has a co-twin with type 0. Hence, the total progeny of the random tree  $\mathcal{T}_t$ ,  $t \geq 0$ , is  $2N_t + 1$ .

### 3.2 Conditional type distribution

In what follows, we let  $X_t^{(i)}$  denote the count of types equal to  $i \geq 1$  until time  $t$ , excluding the initial node.

In Theorem 3.2, which is proved in Appendix B, we compute recursively the conditional type distribution of  $(X_t^{(1)}, \dots, X_t^{(n)})$  given that their summation equals  $N_t$  and  $X_t^{(k)} = 0$  for all  $k > n$ , and show that it does not depend on time  $t > 0$  and on the parameter  $\lambda > 0$ .

**Theorem 3.2.** *For  $j \geq 0$  and  $n \geq 1$ , the conditional type distribution*

$$a_j(m_1, \dots, m_n) := \mathbb{P}_j(X_t^{(1)} = m_1, \dots, X_t^{(n)} = m_n \mid N_t = m_1 + \dots + m_n) \quad (3.3)$$

is given by the recursion

$$a_j(m_1, \dots, m_n) = \sum_{l=1}^{n-j} \frac{1}{l!} \mathbf{1}_{\{m_{j+l} > m_{j+l+1}\}} \sum_{\substack{m_i^1 + \dots + m_i^l = m_i - 1 \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0(m_1^k, \dots, m_n^k)}{1 + m_1^k + \dots + m_n^k}, \quad (3.4)$$

$m_1, \dots, m_n \geq 0$ , with  $m_{n+1} := 0$  in the last indicator  $\mathbf{1}_{\{m_n > m_{n+1}\}}$ ,  $a_j(\emptyset) := 1$ ,  $a_j(m_1, \dots, m_n) = a_j(m_1, \dots, m_{n-1})$  if  $m_n = 0$ , and  $a_j(m_1, \dots, m_n) = 0$  if  $1 \leq n < j$ ,  $m_1 + \dots + m_n \geq 1$ .

From Theorem 3.2, the conditional type distribution (3.3) can be computed by first applying the recursion (3.4) to  $j = 0$ , as

$$a_0(m_1, \dots, m_n) = \sum_{l=1}^n \frac{1}{l!} \mathbf{1}_{\{m_l > m_{l+1}\}} \sum_{\substack{m_i^1 + \dots + m_i^l = m_i - 1 \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0(m_1^k, \dots, m_n^k)}{1 + m_1^k + \dots + m_n^k},$$

and then applying again (3.4) to  $j \geq 1$ . Note also that when  $m_1 + \dots + m_n < n - j$ , the summation range in (3.4) is empty, whence  $a_j(m_1, \dots, m_n) = 0$ . In addition, for  $j \geq 0$ ,  $m \geq 1$ ,  $1 \leq n \leq m + j$  and any weight function  $f_n : \mathbb{N}^n \rightarrow \mathbb{R}$ , we have

$$\mathbb{E}_j [f_n(X_t^{(1)}, \dots, X_t^{(n)}) \mathbf{1}_{\{X_t^{(1)} + \dots + X_t^{(n)} = m\}} \mid N_t = m] = \sum_{\substack{(m_1, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_n = m}} a_j(m_1, \dots, m_n) f_n(m_1, \dots, m_n). \quad (3.5)$$

Figure 6 displays the time evolution of the expected values

$$\mathbb{E}_j [X_t^{(l)}] = \sum_{m=1}^{\infty} \mathbb{E}_j [X_t^{(l)} \mid N_t = m] \mathbb{P}(N_t = m) \quad (3.6)$$

of the count of non-zero types computed as functions of  $t \in [0, 1]$  from

$$\mathbb{E}_j[X_t^{(l)} \mid N_t = m] = \sum_{n=\max(l,j)}^{m+j} \sum_{\substack{(m_1, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_n = m}} m_l a_j(m_1, \dots, m_n), \quad l = 1, \dots, m+j, \quad (3.7)$$

by truncation of the series (3.6) up to  $m = 12$ , after starting from the initial types  $j = 0, 1, 2, 3$ , together with Monte Carlo simulations over 10,000 samples. Color codings are consistent with (2.9) and those of Figures 1-3 and 5.

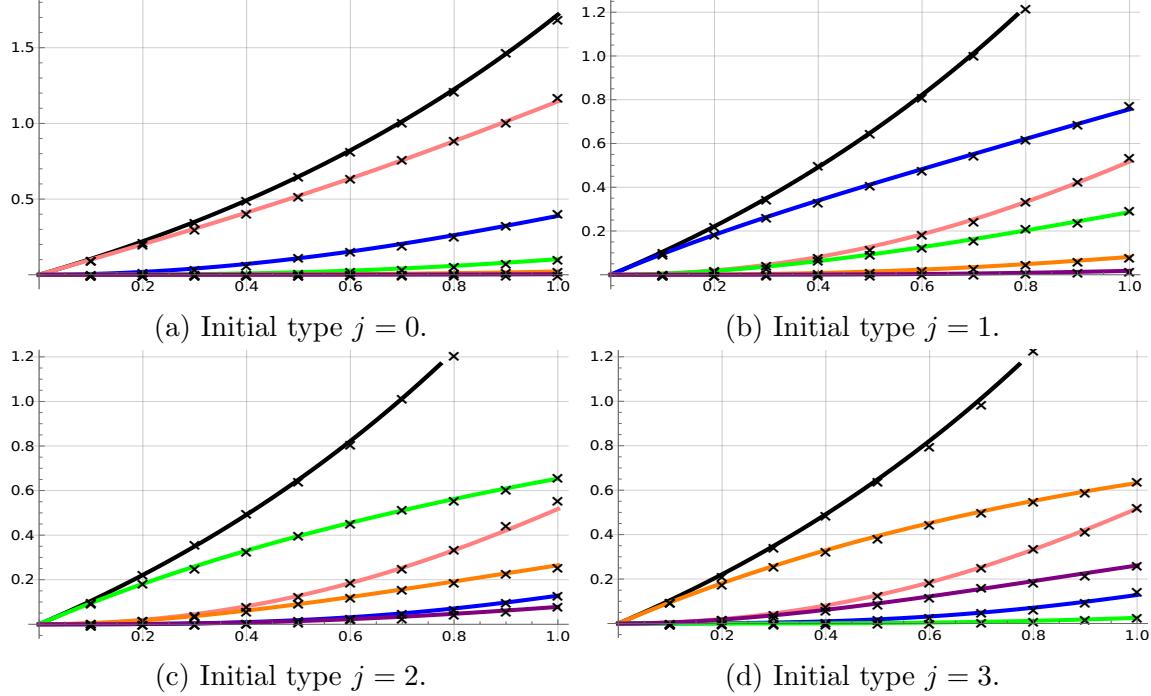


Figure 6: Expected counts (3.6) of types as functions of  $t \in [0, 1]$  with  $\lambda = 1$ .

Figure 7 displays the tail cumulative distribution functions

$$\mathbb{P}_j(\tau^{(l)} > t) = \mathbb{P}_j(X_t^{(l)} = 0) = \mathbb{P}(N_t = 0) + \sum_{m=1}^{\infty} \mathbb{P}_j(X_t^{(l)} = 0 \mid N_t = m) \mathbb{P}(N_t = m) \quad (3.8)$$

of the first time  $\tau^{(l)}$  of occurrence of type  $l$  which, according to (3.3), is computed as

$$\begin{aligned} \mathbb{P}_j(X_t^{(l)} = 0 \mid N_t = m) &= \sum_{n=j}^{l-1} \sum_{\substack{(m_1, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_n = m}} a_j(m_1, \dots, m_n) \\ &+ \sum_{n=\max(l,j)}^{m+j} \sum_{\substack{(m_1, \dots, m_{l-1}, 0, m_{l+1}, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_{l-1} + m_{l+1} + \dots + m_n = m}} a_j(m_1, \dots, m_{l-1}, 0, m_{l+1}, \dots, m_n), \quad l = 1, \dots, m+j. \end{aligned} \quad (3.9)$$

For this, we truncate of the series (3.8) up to  $m = 12$ , after starting from the initial types  $j = 0, 1, 2, 3$ . The closed-form expressions are confirmed by Monte Carlo simulations over 10,000 samples, and remain stable for  $t \in [0, 2]$ .

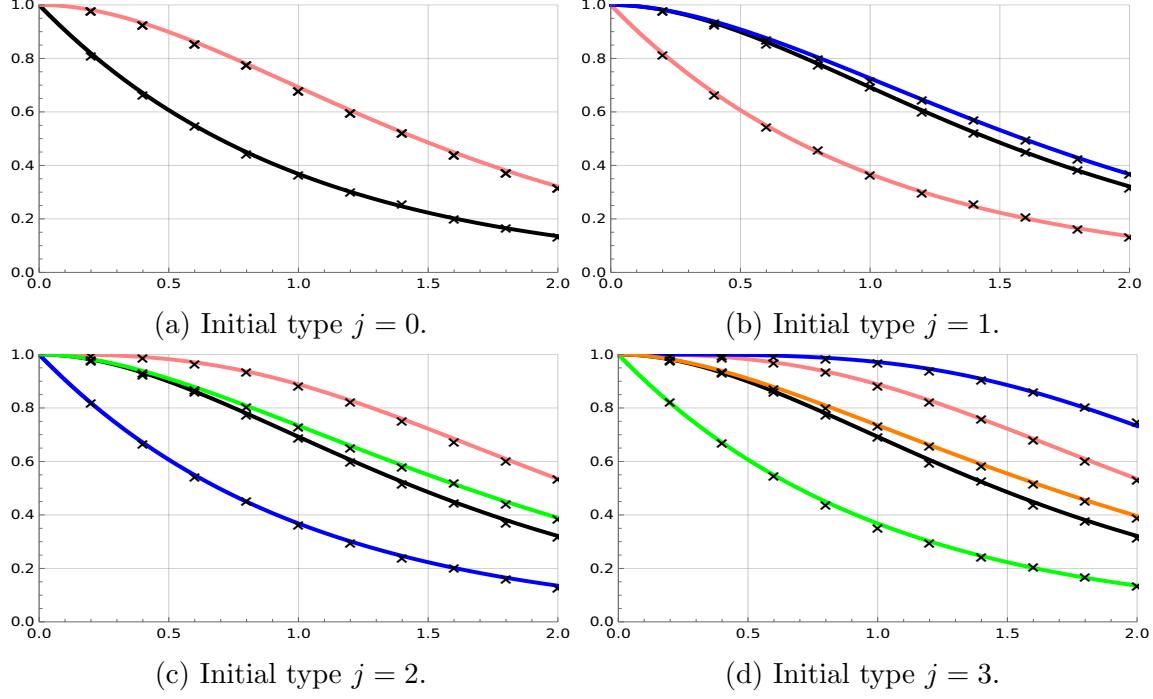


Figure 7: Tail CDFs (3.8) of the occurrence times of given types with  $\lambda = 1$ .

Figure 8 displays the evolution of the probability density functions of the first time  $\tau^{(l)}$  of occurrence of type  $l$ , as obtained from (3.8) and (3.9) for  $t \in [0, 2]$ , after starting from the initial types  $j = 0, 1, 2, 3$ .

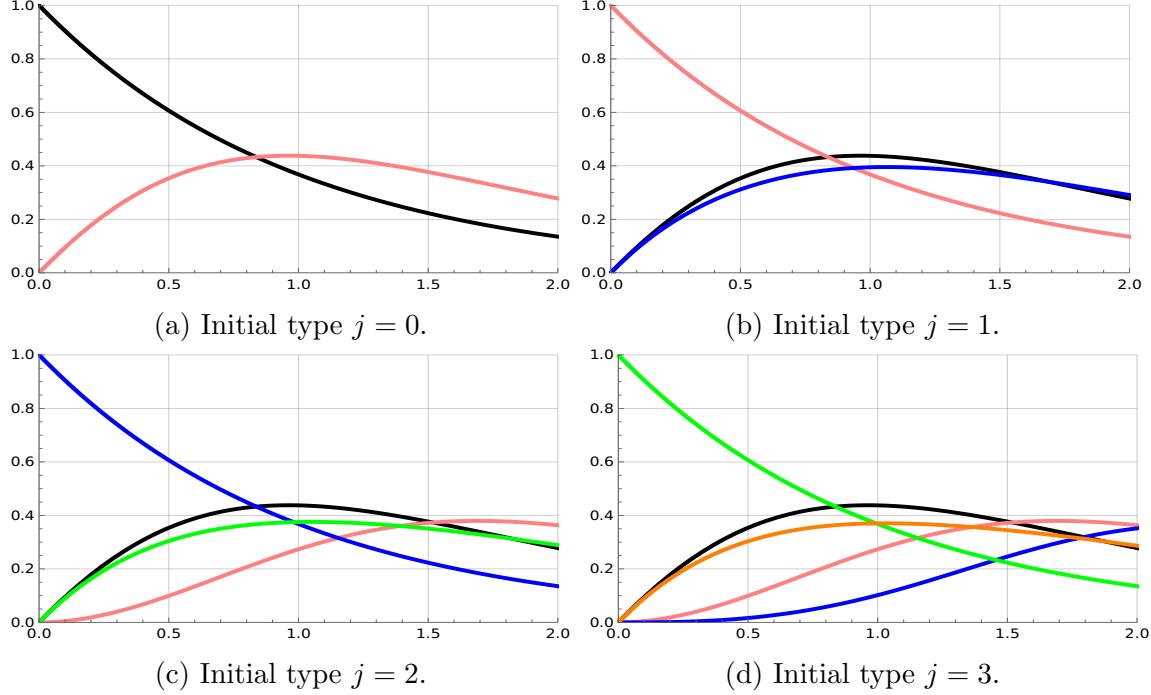


Figure 8: PDFs (3.8) of the occurrence times of given types with  $\lambda = 1$ .

Figure 9 displays the mean proportions

$$\mathbb{E}_j \left[ \frac{X_t^{(l)}}{N_t} \mid N_t \geq 1 \right] = \frac{1}{1 - e^{-\lambda t}} \sum_{m=1}^{\infty} \frac{1}{m} \mathbb{E}_j [X_t^{(l)} \mid N_t = m] \mathbb{P}(N_t = m) \quad (3.10)$$

of non-zero types computed as functions of  $t \in (0, 1)$  from (3.7) and truncation of the series (3.10) up to  $m = 12$ , after starting from the initial types  $j = 0, 1, 2, 3$ . Due to truncation, the computed proportions are accurate and add up to 100% only up to  $t = 1$ .

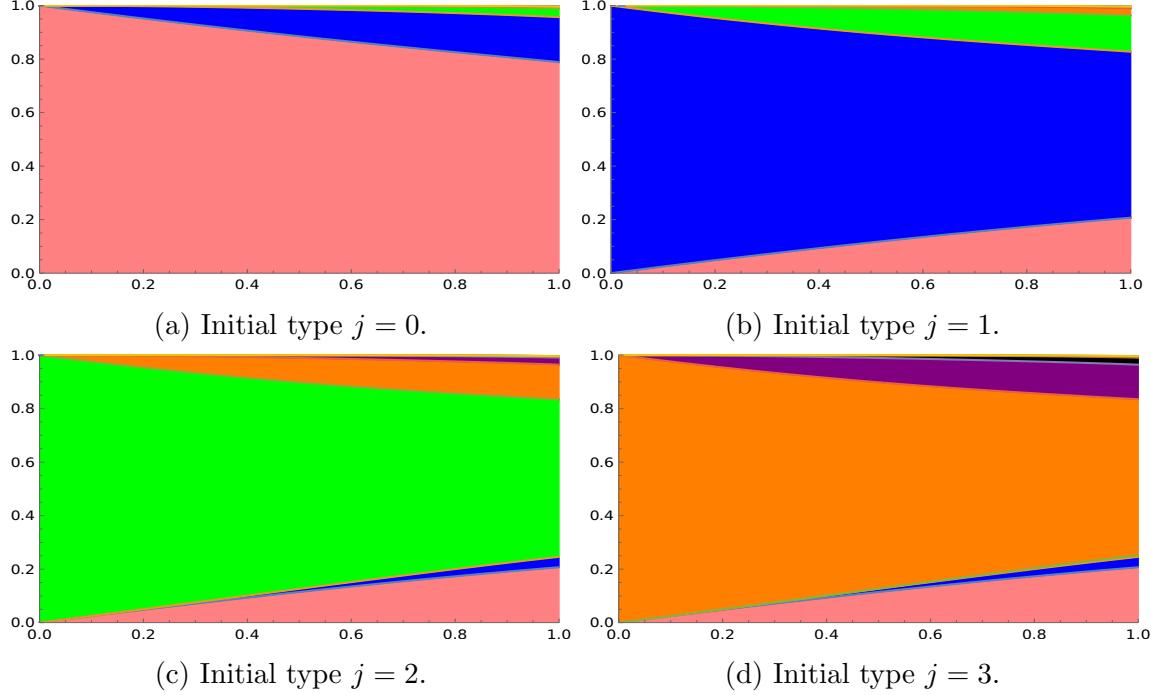


Figure 9: Mean proportions of types (3.10) as functions of  $t \in [0, 1]$  with  $\lambda = 1$ .

### 3.3 Generating functions

In Proposition 3.3, which is proved in Appendix B, we derive a closed-form conditional generating function expression.

**Proposition 3.3.** *For any  $\gamma, t > 0$  and  $m, j \geq 0$  we have*

$$\mathbb{E}_j \left[ \prod_{k=1}^{N_t} (\gamma + k - 2)^{X_t^{(k)}} \mid N_t = m \right] = (-\gamma)^m \binom{-1 - (j-1)/\gamma}{m}. \quad (3.11)$$

In particular, for  $j = 0$ ,  $\gamma = 2$  and  $t > 0$ , we have

$$\mathbb{E}_0 \left[ \prod_{k=1}^{N_t} k^{X_t^{(k)}} \mid N_t = m \right] = \frac{(2m)!}{2^m (m!)^2}, \quad m \geq 0.$$

**Proposition 3.4.** *For any  $\delta, \gamma, t > 0$  such that  $(1 - e^{-\lambda t})\gamma\delta < 1$ , we have*

$$\mathbb{E}_j \left[ \delta^{N_t} \prod_{k=1}^{N_t} (\gamma + k - 2)^{X_t^{(k)}} \right] = \frac{e^{-\lambda t}}{(1 - (1 - e^{-\lambda t})\gamma\delta)^{1+(j-1)/\gamma}}, \quad j \geq 0. \quad (3.12)$$

*Proof.* By Propositions 3.1 and 3.3, we have

$$\mathbb{E}_j \left[ \delta^{N_t} \prod_{k=1}^{N_t} (\gamma + k - 2)^{X_t^{(k)}} \right] = \sum_{m=0}^{\infty} \mathbb{P}(N_t = m) \delta^m \mathbb{E}_j \left[ \prod_{k=1}^m (\gamma + k - 2)^{X_t^{(k)}} \mid N_t = m \right]$$

$$\begin{aligned}
&= e^{-\lambda t} \sum_{m=0}^{\infty} (1 - e^{-\lambda t})^m (-\gamma\delta)^m \binom{-1 - (j-1)/\gamma}{m} \\
&= \frac{e^{-\lambda t}}{(1 - (1 - e^{-\lambda t})\gamma\delta)^{1+(j-1)/\gamma}}, \quad j \geq 0.
\end{aligned}$$

□

In particular, for  $\delta = 1$ ,  $\gamma = 2$  and  $t > 0$  we have

$$\mathbb{E}_j \left[ \prod_{k=1}^{N_t} k^{X_t^{(k)}} \right] = \frac{e^{-\lambda t}}{(2e^{-\lambda t} - 1)^{(j+1)/2}}, \quad j \geq 0.$$

As a consequence of Proposition 3.4, we obtain the following integrability criterion for product functionals.

**Corollary 3.5.** *Let  $t > 0$ ,  $j \geq 0$ ,  $\delta > 0$ ,  $\gamma > 1$ , and let  $(\sigma(k))_{k \geq 0}$  be a real sequence such that*

$$0 \leq \sigma(0) < \frac{1}{(1 - e^{-\lambda t})\gamma\delta} \quad \text{and} \quad 0 \leq \sigma(k) \leq (\gamma + k - 2)\delta, \quad k \geq 1. \quad (3.13)$$

*Then, we have the bound*

$$\mathbb{E}_j \left[ \sigma(0)^{N_t} \prod_{k=1}^{N_t} \sigma(k)^{X_t^{(k)}} \right] \leq \frac{e^{-\lambda t}}{(1 - (1 - e^{-\lambda t})\gamma\delta\sigma(0))^{1+(j-1)/\gamma}} < \infty.$$

*Proof.* By (3.13) we have

$$\mathbb{E}_j \left[ \sigma(0)^{N_t} \prod_{k=1}^{N_t} \sigma(k)^{X_t^{(k)}} \right] \leq \mathbb{E}_j \left[ (\sigma(0)\delta)^{N_t} \prod_{k=1}^{N_t} (\gamma + k - 2)^{X_t^{(k)}} \right], \quad j \geq 0,$$

and we conclude from (3.12). □

## Conclusion

We have presented a multitype Galton–Watson process that can model mutation and reversion in discrete and continuous time. Through a recursive computation of the joint distribution of types conditionally to the value of the total progeny, we have determined the evolution of various expected quantities, such as the mean proportions of different types as the tree size or time increases, and the distribution of the first time of occurrence of a given type. In comparison with the literature on related multitype models, our approach does not rely on approximations.

## A Proofs - discrete-time setting

*Proof of Proposition 2.1.* By [Ott49, Theorem 2], the probability generating function  $G$  of the total progeny  $1 + 2S_\infty^{\neq 0}$  of  $\mathcal{T}$  satisfies the quadratic equation

$$G(\delta) = \delta q + \delta p G(\delta)^2$$

in a neighborhood of 0, and admits the solution (2.3), in which the choice of minus sign follows from the initial condition  $p_0 = \lim_{\delta \rightarrow 0} G(\delta) = 0$ . Letting  $g(w) := q + pw^2$ , by [Ott49, Corollary 3] we have  $\mathbb{P}(S_\infty^{\neq 0} < \infty) = 1$  if and only if  $g'(1) \leq 1$ , i.e.  $p \leq 1/2$ , and

$$\begin{aligned} \mathbb{P}(S_\infty^{\neq 0} < \infty) &= G(1) \\ &= \frac{1 - \sqrt{1 - 4pq}}{2p} \\ &= \frac{1 - \sqrt{1 - 4q + 4q^2}}{2p} \\ &= \frac{1 - |1 - 2q|}{2p} \\ &= \begin{cases} \frac{q}{p} & p \geq 1/2, \\ 1 & p \leq 1/2. \end{cases} \end{aligned}$$

Finally, by Lagrange inversion, see e.g. Theorem 2.10 in [Drm09], and the binomial theorem, we have

$$\begin{aligned} \mathbb{P}(S_\infty^{\neq 0} = n) &= \frac{1}{n!} G^{(n)}(0) \\ &= \frac{1}{n!} \frac{\partial^{n-1}}{\partial w^{n-1}} (g(w))^n \Big|_{w=0} \\ &= \frac{1}{n!} \frac{\partial^{n-1}}{\partial w^{n-1}} \sum_{k=0}^n \binom{n}{k} q^{n-k} p^k w^{2k} \Big|_{w=0} \\ &= \frac{1}{n!} \sum_{k=\lceil(n-1)/2\rceil}^n \binom{n}{k} \frac{q^{n-k} p^k (2k)!}{(2k-n+1)!} w^{2k-n+1} \Big|_{w=0}, \end{aligned}$$

from which (2.2) follows.  $\square$

*Proof of Theorem 2.2.* Recall that  $X^{(k)}$  denotes the count of types equal to  $k \geq 1$  in  $\mathcal{T}$ , excluding the initial node, with  $X^{(k)} = 0$  for  $k > S_\infty^{\neq 0}$ . In what follows, we let

$$p_j(m_1, \dots, m_n) := \mathbb{P}_j(X^{(1)} = m_1, \dots, X^{(n)} = m_n, S_\infty^{\neq 0} = m_1 + \dots + m_n) \quad (\text{A.1})$$

$$= \mathbb{P}_j(X^{(1)} = m_1, \dots, X^{(n)} = m_n, X^{(i)} = 0 \text{ for all } i \geq n+1), \quad j \geq 0.$$

Our proof proceeds by induction on the value of  $m_1 + \dots + m_n$ , noting that when  $m_1 = \dots = m_n = 0$ , we have  $p_j(0, \dots, 0) = \mathbb{P}(S_\infty^{\neq 0} = 0) = q$ .

(i) From the branching mechanism defining the random tree  $\mathcal{T}$ , we have

$$\begin{aligned} p_0(m_1, \dots, m_n) &= p \mathbf{1}_{\{m_1 > m_2\}} p_0(m_1 - 1, m_2, \dots, m_n) p_1(0, \dots, 0) \\ &\quad + p \sum_{\substack{m'_i + m''_i = m_i - 1 \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n, i \neq 3 \\ 0 \leq m'_3 \leq m'_2 + 1}} p_0(m'_1, \dots, m'_n) p_1(m''_1, m''_2 + 1, m''_3, \dots, m''_n), \end{aligned} \quad (\text{A.2})$$

and, for  $1 \leq j < n-1$ ,

$$\begin{aligned} p_j(m_1, \dots, m_j, m_{j+1} + 1, m_{j+2}, \dots, m_n) &= p \mathbf{1}_{\{m_{j+1} \geq m_{j+2}\}} p_0(m_1, \dots, m_n) p_{j+1}(0, \dots, 0) \\ &\quad + p \sum_{\substack{m'_i + m''_i = m_i - 1 \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n, i \neq j+3 \\ 0 \leq m''_{j+3} \leq m''_{j+2} + 1}} p_0(m'_1, \dots, m'_n) \\ &\quad \times p_{j+1}(m''_1, \dots, m''_{j+1}, m''_{j+2} + 1, m''_{j+3}, \dots, m''_n), \end{aligned} \quad (\text{A.3})$$

while for  $j = n-1$  we have

$$p_{n-1}(m_1, \dots, m_{n-1}, m_n + 1) = p p_0(m_1, \dots, m_n) p_n(0, \dots, 0). \quad (\text{A.4})$$

We apply (A.3) with  $j = 1$  to (A.2) to get, since  $p_j(0, \dots, 0) = q$ ,

$$\begin{aligned} p_0(m_1, m_2, \dots, m_n) &= p q \mathbf{1}_{\{m_1 > m_2\}} p_0(m_1 - 1, m_2, \dots, m_n) \\ &\quad + p^2 q \sum_{\substack{m'_i + m''_i = m_i - 1 \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n, i \neq 3 \\ 0 \leq m''_3 \leq m''_2 + 1}} p_0(m'_1, \dots, m'_n) \mathbf{1}_{\{m''_2 \geq m''_3\}} p_0(m''_1, \dots, m''_n) \\ &\quad + p^2 \sum_{\substack{m'_i + m''_i + m'''_i = m_i - 1 \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n \\ 0 \leq m'''_i \leq m'''_{i-1}, 2 \leq i \leq n, i \neq 4 \\ 0 \leq m'''_4 \leq m'''_3 + 1}} p_0(m'_1, \dots, m'_n) p_0(m''_1, \dots, m''_n) p_2(m'''_1, m'''_2, m'''_3 + 1, m'''_4, \dots, m'''_n). \end{aligned}$$

By repeated application of (A.3) with  $j = 2, \dots, n-2$  as well as (A.4) and using the fact that  $m_1^k + \dots + m_n^k \leq m - l$  for all  $k = 1, \dots, n$ , we obtain

$$p_0(m_1, \dots, m_n) = q \sum_{l=1}^n \mathbf{1}_{\{m_l > m_{l+1}\}} p^l \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l p_0(m_1^k, \dots, m_n^k).$$

Next, by the recurrence assumption (2.5) and Proposition 2.1, we have

$$\begin{aligned} p_0(m_1^k, \dots, m_n^k) &= \frac{1}{C_m} b_0(m_1^k, \dots, m_n^k) \mathbb{P}(S_\infty^{\neq 0} = m_1^k + \dots + m_n^k) \\ &= b_0(m_1^k, \dots, m_n^k) q^{1+m_1^k+\dots+m_n^k} p^{m_1^k+\dots+m_n^k}, \end{aligned}$$

hence

$$\begin{aligned} p_0(m_1, \dots, m_n) &= q \sum_{l=1}^n \mathbf{1}_{\{m_l > m_{l+1}\}} p^l \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0(m_1^k, \dots, m_n^k) q^{1+m_1^k+\dots+m_n^k} p^{m_1^k+\dots+m_n^k} \\ &= q(pq)^{m_1+\dots+m_n} \sum_{l=1}^n \mathbf{1}_{\{m_l > m_{l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0(m_1^k, \dots, m_n^k), \end{aligned}$$

which shows (2.5) for  $j = 0$  from (2.2) and the recursive definition (2.6) of  $b_0$ .

(ii) We iterate (A.3) over  $n - j - 1$  steps and then use (A.4) to obtain

$$\begin{aligned} p_j(m_1, \dots, m_j, m_{j+1} + 1, m_{j+2}, \dots, m_n) &= q \sum_{l=1}^{n-j} \mathbf{1}_{\{m_{j+l} - \mathbf{1}_{\{l \geq 2\}} \geq m_{j+l+1}\}} p^l \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{j+2 \leq i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l p_0(m_1^k, \dots, m_n^k) \\ &= q(pq)^{1+m_1+\dots+m_n} \sum_{l=1}^{n-j} \mathbf{1}_{\{m_{j+l} - \mathbf{1}_{\{l \geq 2\}} \geq m_{j+l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{j+2 \leq i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0(m_1^k, \dots, m_n^k), \end{aligned}$$

which shows (2.5) for  $j \geq 1$  from (2.2) and (2.6).  $\square$

*Proof of Corollary 2.3.* Let

$$B_j^\sigma(m) := C_m \mathbb{E}_j \left[ \prod_{k=1}^{m+j} \sigma(k)^{X^{(k)}} \mid S_\infty^{\neq 0} = m \right], \quad j \geq 0, \quad (\text{A.5})$$

with  $B_j^\sigma(0) = 1$ . By Theorem 2.2, we have

$$B_j^\sigma(m) = \sum_{n=1}^{m+j} \sum_{\substack{(m_1, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_n = m}} b_j^\sigma(m_1, \dots, m_n),$$

where

$$b_j^\sigma(m_1, \dots, m_n) := b_j(m_1, \dots, m_n) \prod_{k=1}^n \sigma(k)^{m_k}.$$

By the induction relation (2.6), i.e.

$$b_j^\sigma(m_1, \dots, m_n) = \sum_{l=1}^{n-j} \mathbf{1}_{\{m_{j+l} > m_{j+l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{j < i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0^\sigma(m_1^k, \dots, m_n^k)$$

we have

$$\begin{aligned} B_j^\sigma(m+1) &= \sum_{\substack{m_1 + \dots + m_n = m+1, n \geq 1, \\ 1 \leq m_i \leq m_{i-1}, 2 \leq i \leq n}} b_j^\sigma(m_1, \dots, m_n) \\ &= \sum_{n=j+1}^{m+j+1} \sum_{\substack{m_1 + \dots + m_n = m+1 \\ 1 \leq m_i \leq m_{i-1}, 2 \leq i \leq n}} \sum_{l=1}^{n-j} \mathbf{1}_{\{m_{j+l} > m_{j+l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{j < i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l b_0^\sigma(m_1^k, \dots, m_n^k) \\ &= \sum_{l=1}^{m+1} \sum_{n'=1}^{m+1-l} \sum_{\substack{m'_1 + \dots + m'_{n'} = m+1-l \\ 1 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n' \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n', 1 \leq k \leq l}} \sum_{\substack{\sum_{k=1}^l m_i^k = m'_i, 1 \leq i \leq n' \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n', 1 \leq k \leq l}} \prod_{k=1}^l b_0^\sigma(m_1^k, \dots, m_{n'}^k) \\ &= \sum_{l=1}^{m+1} \sum_{\substack{m_1 + \dots + m_l = m+1-l \\ m_1, \dots, m_l \geq 0}} \sum_{n' \geq 1} \sum_{\substack{m_1^k + \dots + m_{n'}^k = m_k, 1 \leq k \leq l \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n', 1 \leq k \leq l \\ \text{at least one of } m_{n'}^k, 1 \leq k \leq l \text{ is nonzero}}} \prod_{k=1}^l b_0^\sigma(m_1^k, \dots, m_{n'}^k) \\ &= \sum_{l=1}^{m+1} \sum_{\substack{m_1 + \dots + m_l = m+1-l \\ m_1, \dots, m_l \geq 0}} \prod_{k=1}^l \sum_{n_k \geq 0} \sum_{\substack{m_1^k + \dots + m_{n_k}^k = m_k \\ 1 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n_k}} b_0^\sigma(m_1^k, \dots, m_{n_k}^k) \\ &= \sum_{l=1}^{m+1} \left( \prod_{k=j+1}^{j+l} \sigma(k) \right) \sum_{\substack{m_1 + \dots + m_l = m+1 \\ m_1 \geq 1, \dots, m_l \geq 1}} \prod_{k=1}^l B_0^\sigma(m_k - 1), \quad m \geq 0, \end{aligned} \tag{A.6}$$

where in the third equality we made the change of variables  $m'_i = m_i - \mathbf{1}_{\{j < i \leq j+l\}}$ . Let now

$$D_j^{(k)}(m) := C_m \mathbb{E}_j[X^{(k)} \mid S_\infty^{\neq 0} = m]$$

$$\begin{aligned}
&= \sum_{n=\max(k,j)}^{m+j} \sum_{\substack{(m_1, \dots, m_n) \in \mathbb{K}_{j,n} \\ m_1 + \dots + m_n = m}} m_k b_j(m_1, \dots, m_n) \\
&= \frac{\partial}{\partial \sigma(k)} \Big|_{\sigma=1} B_j^\sigma(m), \quad l = 1, \dots, m+j, \quad j, m \geq 0,
\end{aligned}$$

with initial values  $D_j^{(k)}(0) = 0$ . By (A.6), for  $m \geq 0$  we have

$$\begin{aligned}
D_j^{(k)}(m+1) &= \frac{\partial}{\partial \sigma(k)} \Big|_{\sigma=1} [x^{m+1}] \sum_{l=1}^{\infty} \left( \prod_{k'=j+1}^{j+l} \sigma(k') \right) \left( \sum_{n=1}^{\infty} B_0^\sigma(n-1)x^n \right)^l \\
&= [x^{m+1}] \sum_{l=1}^{\infty} \mathbf{1}_{\{j < k \leq j+l\}} \left( \prod_{k'=j+1, k' \neq k}^{j+l} \sigma(k') \right) \left( \sum_{n=1}^{\infty} B_0^\sigma(n-1)x^n \right)^l \Big|_{\sigma=1} \\
&\quad + [x^{m+1}] \sum_{l=1}^{\infty} \left( \prod_{k'=j+1}^{j+l} \sigma(k') \right) l \left( \sum_{n=1}^{\infty} B_0^\sigma(n-1)x^n \right)^{l-1} \left( \sum_{n=1}^{\infty} \frac{\partial}{\partial \sigma(k)} B_0^\sigma(n-1)x^n \right) \Big|_{\sigma=1} \\
&= \mathbf{1}_{\{j < k\}} [x^{m+1}] \sum_{l=k-j}^{\infty} \left( \sum_{n=1}^{\infty} B_0^1(n-1)x^n \right)^l \\
&\quad + [x^{m+1}] \sum_{l=1}^{\infty} l \left( \sum_{n=1}^{\infty} B_0^1(n-1)x^n \right)^{l-1} \left( \sum_{n=1}^{\infty} D_k^{(0)}(n)x^{n+1} \right),
\end{aligned}$$

where  $[x^{m+1}]$  is the operator extracting the coefficient of the term  $x^{m+1}$  from the series following it. Thus,

$$\begin{aligned}
\sum_{m=0}^{\infty} D_j^{(k)}(m+1)x^{m+1} &= \mathbf{1}_{\{j < k\}} \sum_{l=k-j}^{\infty} \left( \sum_{n=1}^{\infty} B_0^1(n-1)x^n \right)^l \\
&\quad + \sum_{l=1}^{\infty} l \left( \sum_{n=1}^{\infty} B_0^1(n-1)x^n \right)^{l-1} \left( \sum_{n=1}^{\infty} D_k^{(0)}(n)x^n \right).
\end{aligned}$$

By (A.5) and Proposition 2.1, we have

$$\sum_{n=1}^{\infty} B_0^1(n-1)x^n = \sum_{n=1}^{\infty} C_{n-1}x^n = \frac{1 - \sqrt{1 - 4x}}{2},$$

which implies

$$\sum_{l=k}^{\infty} \left( \sum_{n=1}^{\infty} B_0^1(n-1)x^n \right)^l = x^k \left( \frac{1 - \sqrt{1 - 4x}}{2x} \right)^{k+1},$$

and

$$\sum_{l=1}^{\infty} l \left( \sum_{n=1}^{\infty} B_0^1(n-1)x^n \right)^{l-1} = \left( \frac{1 - \sqrt{1 - 4x}}{2x} \right)^2.$$

Hence, the unconditional expected value of  $X^{(k)}$  is given by

$$\begin{aligned}
\mathbb{E}_j[X^{(k)}] &= \sum_{m=0}^{\infty} \mathbb{E}_j[X^{(k)} \mid S_{\infty}^{\neq 0} = m] \mathbb{P}(S_{\infty}^{\neq 0} = m) \\
&= q \sum_{m=0}^{\infty} D_j^{(k)} (m+1) (pq)^{m+1} \\
&= \mathbf{1}_{\{j < k\}} \frac{1}{p} \left( \frac{1 - \sqrt{1 - 4pq}}{2} \right)^{k+1-j} + \frac{1}{pq} \left( \frac{1 - \sqrt{1 - 4pq}}{2} \right)^2 \mathbb{E}_0[X^{(k)}] \\
&= \mathbf{1}_{\{j < k\}} p^{k-j} + \frac{p}{q} \mathbb{E}_0[X^{(k)}].
\end{aligned}$$

When  $j = 0$ , this yields

$$\mathbb{E}_0[X^{(k)}] = \frac{q}{\sqrt{1 - 4pq}} \left( \frac{1 - \sqrt{1 - 4pq}}{2} \right)^k = \frac{qp^k}{q - p},$$

and in general we obtain

$$\begin{aligned}
\mathbb{E}_j[X^{(k)}] &= \frac{1}{p} \mathbf{1}_{\{j < k\}} \left( \frac{1 - \sqrt{1 - 4pq}}{2} \right)^{k+1-j} + \frac{1}{p\sqrt{1 - 4pq}} \left( \frac{1 - \sqrt{1 - 4pq}}{2} \right)^{k+2} \\
&= \mathbf{1}_{\{j < k\}} p^{k-j} + \frac{p^{k+1}}{q - p}.
\end{aligned}$$

Hence, when  $j = 0$  we have

$$\mathbb{E}_0[X^{(k)}] = q \sum_{n=k}^{\infty} \binom{2n - k}{n} (pq)^n,$$

and in general we obtain

$$\mathbb{E}_j[X^{(k)}] = q \mathbf{1}_{\{j < k\}} \sum_{n=k-j}^{\infty} \frac{k+1-j}{n+1} \binom{2n - k + j}{n} (pq)^n + q \sum_{n=k}^{\infty} \binom{2n - k}{n+1} (pq)^n,$$

which yields (2.7).  $\square$

*Proof of Corollary 2.4.* Using (2.7), we have

$$\begin{aligned}
\mathbb{E}_j \left[ \frac{X^{(k)}}{S_{\infty}^{\neq 0}} \mid S_{\infty}^{\neq 0} \geq 1 \right] &= \frac{1}{p} \sum_{m=1}^{\infty} \frac{1}{m} \mathbb{E}_j[X^{(k)} \mid S_{\infty}^{\neq 0} = m] \mathbb{P}(S_{\infty}^{\neq 0} = m) \\
&= \frac{q}{p} \mathbf{1}_{\{j < k\}} \sum_{m=k-j}^{\infty} \frac{k+1-j}{m+1} \binom{2m - k + j}{m} \frac{(pq)^m}{m} + \frac{q}{p} \sum_{m=k}^{\infty} \binom{2m - k}{m+1} \frac{(pq)^m}{m} \\
&= \frac{q}{p} \mathbf{1}_{\{j < k\}} \int_0^{pq} \sum_{m=k-j}^{\infty} \frac{k+1-j}{m+1} \binom{2m - k + j}{m} x^{m-1} dx + \frac{q}{p} \int_0^{pq} \sum_{m=k}^{\infty} \binom{2m - k}{m+1} x^{m-1} dx
\end{aligned}$$

$$\begin{aligned}
&= \frac{q}{p} \mathbf{1}_{\{j < k\}} \int_0^{pq} \frac{1}{x^2} \left( \frac{1 - \sqrt{1 - 4x}}{2} \right)^{k+1-j} dx + \frac{q}{p} \int_0^{pq} \frac{1}{x^2 \sqrt{1 - 4x}} \left( \frac{1 - \sqrt{1 - 4x}}{2} \right)^{k+2} dx \\
&= \frac{q}{p} \mathbf{1}_{\{j < k\}} \left( (k+1-j)B\left(\frac{1 - \sqrt{1 - 4pq}}{2}; k-j; 0\right) - \frac{1}{pq} \left( \frac{1 - \sqrt{1 - 4pq}}{2} \right)^{k+1-j} \right) \\
&\quad + \frac{q}{p} B\left(\frac{1 - \sqrt{1 - 4pq}}{2}; 1+k, -1\right).
\end{aligned}$$

□

*Proof of Proposition 2.5.* Taking  $j = 0$  and

$$\sigma(k) := 1 + \frac{\gamma}{k}, \quad k \geq 1,$$

in (A.6) and denoting  $B_j^\sigma$  by  $B_j^\gamma$ , we have

$$B_0^\gamma(n+1) = \sum_{l=1}^{n+1} \binom{l+\gamma}{l} \sum_{\substack{m_1 + \dots + m_l = n+1 \\ m_1, \dots, m_l \geq 1}} \prod_{k=1}^l B_0^\gamma(m_k - 1),$$

and by the Faà di Bruno formula in Lemma A.1 below we find that  $B_0^\gamma(n)$  is the coefficient of  $x^n$  in the series

$$\sum_{l=1}^{\infty} \binom{l+\gamma}{l} \left( \sum_{n=1}^{\infty} \binom{(2+\gamma)n-2}{n-1} \frac{x^n}{n} \right)^l.$$

By Lemma A.2 below, denoting by  $\Phi_\gamma^{-1}$  the inverse function of

$$\Phi_\gamma(w) := w(1-w)^{1+\gamma}, \quad w \in \mathbb{C},$$

we have

$$\begin{aligned}
\sum_{l=1}^{\infty} \binom{l+\gamma}{l} \left( \sum_{n=1}^{\infty} \binom{(2+\gamma)n-2}{n-1} \frac{x^n}{n} \right)^l &= \sum_{l=1}^{\infty} \binom{l+\gamma}{l} \left( \sum_{n=1}^{\infty} F_{n-1}(\gamma+2, \gamma+1)x^n \right)^l \\
&= \sum_{l=1}^{\infty} \binom{l+\gamma}{l} (\Phi_\gamma^{-1}(x))^l \\
&= (1 - \Phi_\gamma^{-1}(x))^{-\gamma-1} - 1 \\
&= \frac{1}{x} \Phi_\gamma^{-1}(x) - 1 \\
&= \sum_{n=0}^{\infty} F_n(\gamma+2, \gamma+1)x^n,
\end{aligned}$$

which yields (2.12). □

We also recall the following version of the Faà di Bruno formula which is used in the proofs of Propositions 2.5 and 3.3, see for example Theorem 5.1.4 in [Sta99].

**Lemma A.1.** *For any two sequences  $(\alpha_n)_{n \geq 1}$ ,  $(\beta_n)_{n \geq 1}$ , the coefficient of  $x^m$ ,  $m \geq 1$ , in the series*

$$\sum_{l=1}^{\infty} \alpha_l \left( \sum_{n=1}^{\infty} \beta_n x^n \right)^l$$

*is given by*

$$\sum_{l=1}^m \alpha_l \sum_{\substack{m_1 + \dots + m_l = m \\ m_1, \dots, m_l \geq 1}} \beta_{m_1} \dots \beta_{m_l}.$$

The following lemma was used in the proof of Proposition 2.5.

**Lemma A.2.** *The inverse function  $\Phi_{\gamma}^{-1}$  of*

$$\Phi_{\gamma}(w) := w(1-w)^{1+\gamma}, \quad w \in \mathbb{C}, \quad (\text{A.7})$$

*admits the expansion*

$$\Phi_{\gamma}^{-1}(x) = \sum_{n=1}^{\infty} F_{n-1}(\gamma+2, \gamma+1) x^n.$$

*Proof.* Since  $\Phi_{\gamma}$  is analytic near  $w = 0$  and  $\Phi_{\gamma}(0) = 0$ ,  $\Phi'_{\gamma}(0) = 1 \neq 0$ , by the Lagrange inversion theorem, the inverse function of  $\Phi_{\gamma}$  is given by the power series

$$\Phi_{\gamma}^{-1}(z) = \sum_{n=1}^{\infty} \frac{\alpha_n}{n!} z^n,$$

where

$$\begin{aligned} \alpha_n &= \lim_{w \rightarrow 0} \frac{\partial^{n-1}}{\partial w^{n-1}} \left( \frac{w}{\Phi_{\gamma}(w)} \right)^n \\ &= \lim_{w \rightarrow 0} \frac{\partial^{n-1}}{\partial w^{n-1}} (1-w)^{-(1+\gamma)n} \\ &= \lim_{w \rightarrow 0} \frac{\partial^{n-1}}{\partial w^{n-1}} \sum_{k=0}^{\infty} \binom{k + (1+\gamma)n - 1}{k} w^k \\ &= (n-1)! \binom{(2+\gamma)n - 2}{n-1}. \end{aligned}$$

□

## B Proofs - continuous-time setting

*Proof of Proposition 3.1.* We denote by

$$\bar{F}_\rho(t) := \mathbb{P}(T_\emptyset > t) = \int_t^\infty \rho(r) dr, \quad t \geq 0,$$

the tail cumulative distribution function of  $\rho$ , and let  $p_t(n) := \mathbb{P}(N_t = n)$ ,  $n \geq 0$ , with

$$p_t(0) = \mathbb{P}(N_t = 0) = \mathbb{P}(T_\emptyset > t) = \bar{F}_\rho(t), \quad t \in \mathbb{R}_+.$$

For  $n \geq 1$ , by the relation  $\{N_t \geq 1\} \subset \{T_\emptyset \leq t\}$  and independence of branches, denoting by  $(N_t^1)_{t \in \mathbb{R}_+}$  and  $(N_t^2)_{t \in \mathbb{R}_+}$  two independent copies of  $(N_t)_{t \in \mathbb{R}_+}$ , we have

$$\begin{aligned} p_t(n) &= \mathbb{P}(N_t = n) \\ &= \mathbb{E}[\mathbb{P}(N_t = n, T_\emptyset \leq t \mid T_\emptyset)] \\ &= \mathbb{E}[\mathbb{P}(N_s^1 + N_s^2 = n - 1)_{|s=t-T_\emptyset} \mathbf{1}_{\{T_\emptyset \leq t\}}] \\ &= \mathbb{E}[p_s^{*2}(n - 1)_{|s=t-T_\emptyset} \mathbf{1}_{\{T_\emptyset \leq t\}}] \\ &= \int_0^t \rho(t - s) p_s^{*2}(n - 1) ds, \end{aligned}$$

where  $*$  is the discrete convolution product. As the distribution  $\rho$  is exponential with parameter  $\lambda$ , we have

$$p_t(n) = \begin{cases} e^{-\lambda t}, & n = 0, \\ \lambda \int_0^t e^{-(t-s)\lambda} p_s^{*2}(n - 1) ds = \lambda \int_0^t e^{-(t-s)\lambda} \sum_{\substack{n_1 + n_2 = n-1 \\ n_1, n_2 \geq 0}} p_s(n_1) p_s(n_2) ds, & n \geq 1. \end{cases} \quad (\text{B.1})$$

Multiplying both sides of the third equality in (B.1) by  $z^n$  and summing over  $n \geq 1$  gives

$$G_t(z) - z e^{-\lambda t} = z \lambda \int_0^t e^{-(t-s)\lambda} G_s(z)^2 ds,$$

which in turns yields the Bernoulli ODE

$$\frac{d}{dt} G_t(z) + \lambda G_t(z) = \lambda z G_t(z)^2, \quad t > 0, \quad (\text{B.2})$$

with initial condition  $G_0(z) = z$  since  $p_0(n) = \mathbf{1}_{\{n=0\}}$ . The solution of (B.2) is then obtained by a standard argument, which allows us to conclude to (3.2).  $\square$

*Proof of Theorem 3.2.* Recall that  $X_t^{(i)}$  denotes the count of types equal to  $i \geq 1$  until time  $t$ , excluding the initial node. Similarly to (A.1), we let

$$\begin{aligned} p_{t,j}(m_1, \dots, m_n) &:= \mathbb{P}_j(X_t^{(1)} = m_1, \dots, X_t^{(n)} = m_n, N_t = m_1 + \dots + m_n) \\ &= \mathbb{P}_j(X_t^{(1)} = m_1, \dots, X_t^{(n)} = m_n, X_t^{(i)} = m_n \text{ for all } i \geq n+1), \end{aligned}$$

$j \geq 0$ . Our proof proceeds by induction on the value of  $m_1 + \dots + m_n$ , with

$$p_{t,j}(0, \dots, 0) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$$

when  $m_1 = \dots = m_n = 0$ .

We note that the branching chain  $(X_t)_{t \geq 0}$  with initial type 0 has  $m_i$  branches with type  $i$  for each  $i \geq 1$ , then it must have  $(1 + m_1 + \dots + m_n)$  branches with type 0, since each branch with type 0, except the initial one, has one and only one brother with a positive type.

(i) For  $j = 0$ , we have

$$\begin{aligned} p_{t,0}(m_1, \dots, m_n) &= \mathbf{1}_{\{m_1 > m_2\}} \lambda \int_0^t e^{-(t-s)\lambda} p_{s,0}(m_1 - 1, m_2, \dots, m_n) p_{s,1}(1) ds \\ &+ \lambda \int_0^t e^{-(t-s)\lambda} \sum_{\substack{m'_i + m''_i = m_i - \mathbf{1}_{\{1 \leq i \leq 2\}}, 1 \leq i \leq n \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n, i \neq 3 \\ 0 \leq m''_3 \leq m''_2 + 1}} p_{s,0}(m'_1, \dots, m'_n) p_{s,1}(m''_1, m''_2 + 1, m''_3, \dots, m''_n) ds, \end{aligned} \quad (\text{B.3})$$

and, for  $1 \leq j < n-1$ ,

$$\begin{aligned} p_{t,j}(m_1, \dots, m_j, m_{j+1} + 1, m_{j+2}, \dots, m_n) &= \mathbf{1}_{\{m_{j+1} \geq m_{j+2}\}} \lambda \int_0^t e^{-(t-s)\lambda} p_{s,0}(m_1, \dots, m_n) p_{s,j+1}(0, \dots, 0) ds + \lambda \int_0^t e^{-(t-s)\lambda} \\ &\sum_{\substack{m'_i + m''_i = m_i - \mathbf{1}_{\{1 \leq i = j+2\}}, 1 \leq i \leq n \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n, i \neq j+3 \\ 0 \leq m''_{j+3} \leq m''_{j+2} + 1}} p_{s,0}(m'_1, \dots, m'_n) p_{s,j+1}(m''_1, \dots, m''_{j+1}, m''_{j+2} + 1, m''_{j+3}, \dots, m''_n) ds, \end{aligned} \quad (\text{B.4})$$

while for  $j = n-1$  we have

$$p_{t,n-1}(m_1, \dots, m_{n-1}, m_n + 1) = \lambda \int_0^t e^{-(t-s)\lambda} p_{s,0}(m_1, \dots, m_n) p_{s,n}(0, \dots, 0) ds. \quad (\text{B.5})$$

Since  $p_{t,j}(0, \dots, 0) = e^{-\lambda t}$ , we apply (B.4) with  $j = 1$  to (B.3) to get

$$p_{t,0}(m_1, \dots, m_n) = \mathbf{1}_{\{m_1 > m_2\}} \lambda e^{-\lambda t} \int_0^t p_{s,0}(m_1 - 1, m_2, \dots, m_n) ds$$

$$\begin{aligned}
& + \mathbf{1}_{\{m_2 > m_3\}} \lambda^2 e^{-\lambda t} \int_0^t \int_0^s \sum_{\substack{m_i^1 + m_i^2 = m_i - \mathbf{1}_{\{1 \leq i \leq 2\}}, 1 \leq i \leq n \\ 0 \leq m_i^1 \leq m_{i-1}^1, 2 \leq i \leq n \\ 0 \leq m_i^2 \leq m_{i-1}^2, 2 \leq i \leq n}} p_{s,0}(m_1^1, \dots, m_n^1) p_{r,0}(m_1^2, \dots, m_n^2) dr ds \\
& + \int_0^t \int_0^s \lambda^2 e^{(r-t)\lambda} \sum_{\substack{m_i^1 + m_i^2 + m_i^3 = m_i - \mathbf{1}_{\{1 \leq i \leq 3\}}, 1 \leq i \leq n \\ 0 \leq m_i^1 \leq m_{i-1}^1, 2 \leq i \leq n \\ 0 \leq m_i^2 \leq m_{i-1}^2, 2 \leq i \leq n \\ 0 \leq m_i^3 \leq m_{i-1}^3, 2 \leq i \leq n, i \neq 4 \\ 0 \leq m_4^3 \leq m_3^3 + 1}} p_{s,0}(m_1^1, \dots, m_n^1) p_{r,0}(m_1^2, \dots, m_n^2) p_{r,2}(m_1^3, m_2^3, m_3^3 + 1, m_4^3, \dots, m_n^3) dr ds.
\end{aligned}$$

By repeated application of (B.4) with  $j = 2, \dots, n-2$  as well as (B.5), we obtain

$$\begin{aligned}
& p_{t,0}(m_1, \dots, m_n) \\
& = e^{-\lambda t} \sum_{l=1}^n \lambda^l \mathbf{1}_{\{m_l > m_{l+1}\}} \int_{0 \leq s_l \leq \dots \leq s_1 \leq t} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l p_{s_k,0}(m_1^k, \dots, m_n^k) ds_l \dots ds_1 \\
& = e^{-\lambda t} \sum_{l=1}^n \frac{\lambda^l}{l!} \mathbf{1}_{\{m_l > m_{l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \int_0^t p_{s,0}(m_1^k, \dots, m_n^k) ds.
\end{aligned} \tag{B.6}$$

Observe that in multi-index notation, the constraint in the above summation reads

$$\sum_{k=1}^l (m_1^k, \dots, m_n^k) = (m_1, \dots, m_n) - (\underbrace{1, \dots, 1, 0, \dots, 0}_n).$$

Thus, the proof can be conducted by induction over the set of multi-indices

$$\{(m_1, \dots, m_n) : m_1 \geq \dots \geq m_n \geq 0\}$$

in the back-diagonal order. The induction starts from the initial multi-index  $\emptyset$ , in which case the result follows from  $a_0(\emptyset) = 1$  and  $p_{t,0}(0, \dots, 0) = e^{-\lambda t}$ . Writing the induction hypothesis as

$$p_{s,0}(m_1^k, \dots, m_n^k) = a_0(m_1^k, \dots, m_n^k) e^{-\lambda s} (1 - e^{-\lambda s})^{m_1^k + \dots + m_n^k}$$

and using (B.6), we obtain

$$\begin{aligned}
& p_{t,0}(m_1, \dots, m_n) \\
& = e^{-\lambda t} \sum_{l=1}^n \frac{\mathbf{1}_{\{m_l > m_{l+1}\}}}{l!} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \int_0^t p_{s,0}(m_1^k, \dots, m_n^k) ds
\end{aligned}$$

$$\begin{aligned}
&= e^{-\lambda t} \sum_{l=1}^n \frac{\mathbf{1}_{\{m_l > m_{l+1}\}}}{l!} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \int_0^t a_0(m_1^k, \dots, m_n^k) e^{-\lambda s} (1 - e^{-\lambda s})^{m_1^k + \dots + m_n^k} ds \\
&= e^{-\lambda t} (1 - e^{-\lambda t})^{m_1 + \dots + m_n} \sum_{l=1}^n \frac{\mathbf{1}_{\{m_l > m_{l+1}\}}}{l!} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0(m_1^k, \dots, m_n^k)}{1 + m_1^k + \dots + m_n^k} \\
&= \mathbb{P}(N_t = m_1 + \dots + m_n) \sum_{l=1}^n \frac{\mathbf{1}_{\{m_l > m_{l+1}\}}}{l!} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0(m_1^k, \dots, m_n^k)}{1 + m_1^k + \dots + m_n^k}
\end{aligned}$$

from (3.1), which yields (3.4) when  $j = 0$  and  $1 \leq m_i \leq m_{i-1}$ ,  $2 \leq i \leq n$ .

(ii) By iterating (B.4) over  $n - j - 1$  steps and then using (B.5), we obtain

$$\begin{aligned}
&p_{t,j}(m_1, \dots, m_j, m_{j+1} + 1, m_{j+2}, \dots, m_n) \\
&= \mathbf{1}_{\{m_{j+1} \geq m_{j+2}\}} \lambda e^{-\lambda t} \int_0^t p_{s,0}(m_1, \dots, m_n) ds \\
&\quad + \lambda \int_0^t e^{-(t-s)\lambda} \\
&\quad \sum_{\substack{m'_i + m''_i = m_i - \mathbf{1}_{\{1 \leq i = j+2\}}, 1 \leq i \leq n \\ 0 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n \\ 0 \leq m''_i \leq m''_{i-1}, 2 \leq i \leq n \\ i \neq j+3 \\ 0 \leq m''_{j+3} \leq m''_{j+2} + 1}} p_{s,0}(m'_1, \dots, m'_n) p_{s,j+1}(m''_1, \dots, m''_{j+1}, m''_{j+2} + 1, m''_{j+3}, \dots, m''_n) ds \\
&= \mathbf{1}_{\{m_{j+1} \geq m_{j+2}\}} \lambda e^{-\lambda t} \int_0^t p_{s,0}(m_1, \dots, m_n) ds \\
&\quad + \mathbf{1}_{\{m_{j+2} > m_{j+3}\}} \lambda^2 e^{-\lambda t} \int_0^t \int_0^s \sum_{\substack{m_i^1 + m_i^2 = m_i - \mathbf{1}_{\{1 \leq i = j+2\}}, 1 \leq i \leq n \\ 0 \leq m_i^1 \leq m_{i-1}^1, 2 \leq i \leq n \\ 0 \leq m_i^2 \leq m_{i-1}^2, 2 \leq i \leq n}} p_{s,0}(m_1^1, \dots, m_n^1) p_{r,0}(m_1^2, \dots, m_n^2) dr ds \\
&\quad + \lambda^2 \int_0^t \int_0^s e^{(r-t)\lambda} \\
&\quad \sum_{\substack{m_i^1 + m_i^2 + m_i^3 = m_i - \mathbf{1}_{\{j+2 \leq i \leq j+3\}}, 1 \leq i \leq n \\ 0 \leq m_i^1 \leq m_{i-1}^1, 2 \leq i \leq n \\ 0 \leq m_i^2 \leq m_{i-1}^2, 2 \leq i \leq n \\ 0 \leq m_i^3 \leq m_{i-1}^3, 2 \leq i \leq n, i \neq j+4 \\ 0 \leq m_{j+4}^3 \leq m_{j+3}^3 + 1}} p_{s,0}(m_1^1, \dots, m_n^1) \\
&\quad p_{r,0}(m_1^2, \dots, m_n^2) p_{r,j+2}(m_1^3, \dots, m_{j+2}^3, m_{j+3}^3 + 1, m_{j+4}^3, \dots, m_n^3) dr ds \\
&= \dots \\
&= e^{-\lambda t} \sum_{l=1}^{n-j} \mathbf{1}_{\{m_{j+l} - \mathbf{1}_{\{l \geq 2\}} \geq m_{j+l+1}\}} \lambda^l
\end{aligned}$$

$$\begin{aligned}
& \int_{0 \leq s_l \leq \dots \leq s_1 \leq t} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - 1_{\{j+2 \leq i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l p_{s_k,0}(m_1^k, \dots, m_n^k) ds_l \dots ds_1 \\
&= e^{-\lambda t} (1 - e^{-\lambda t})^{1+m_1+\dots+m_n} \sum_{l=1}^{n-j} \frac{\mathbf{1}_{\{m_{j+l} - 1_{\{l \geq 2\}} \geq m_{j+l+1}\}}}{l!} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - 1_{\{j+2 \leq i \leq j+l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0(m_1^k, \dots, m_n^k)}{1 + m_1^k + \dots + m_n^k},
\end{aligned}$$

from which (3.4) follows.  $\square$

*Proof of Proposition 3.3.* We proceed by induction on  $m \geq 0$ . We let

$$A_j^\sigma(m) := \mathbb{E}_j \left[ \prod_{k=1}^{m+j} \sigma(k)^{X_t^{(k)}} \mid N_t = m \right], \quad j \geq 0,$$

with  $A_j^\sigma(0) = 1$ . By (3.5), we have

$$A_j^\sigma(m) = \sum_{\substack{m_1+\dots+m_n=m, n \geq 0, \\ 1 \leq m_i \leq m_{i-1}, 2 \leq i \leq n}} a_j^\sigma(m_1, \dots, m_n),$$

where

$$a_j^\sigma(m_1, \dots, m_n) := a_j(m_1, \dots, m_n) \prod_{k=1}^n \sigma(k)^{m_k},$$

and  $\sigma(k) := \gamma + k - 2$ ,  $k \geq 1$ . By the induction relation (3.4), similarly to (A.6), we have

$$\begin{aligned}
A_j^\sigma(m+1) &= \sum_{\substack{m_1+\dots+m_n=m+1, n \geq 1, \\ 1 \leq m_i \leq m_{i-1}, 2 \leq i \leq n}} a_j^\sigma(m_1, \dots, m_n) \\
&= \sum_{n=j+1}^{m+j+1} \sum_{\substack{m_1+\dots+m_n=m+1 \\ 1 \leq m_i \leq m_{i-1}, 2 \leq i \leq n}} \sum_{l=1}^{n-j} \frac{1}{l!} \mathbf{1}_{\{m_l > m_{l+1}\}} \sum_{\substack{\sum_{k=1}^l m_i^k = m_i - 1_{\{1 \leq i \leq l\}}, 1 \leq i \leq n \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n, 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0^\sigma(m_1^k, \dots, m_n^k)}{1 + m_1^k + \dots + m_n^k} \\
&= \sum_{l=1}^{m+1} \frac{1}{l!} \sum_{n'=1}^{m+1-l} \sum_{\substack{m'_1+\dots+m'_{n'}=m+1-l \\ 1 \leq m'_i \leq m'_{i-1}, 2 \leq i \leq n' \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n', 1 \leq k \leq l}} \sum_{\substack{\sum_{k=1}^l m_i^k = m'_i, 1 \leq i \leq n' \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n', 1 \leq k \leq l}} \prod_{k=1}^l \frac{a_0^\sigma(m_1^k, \dots, m_{n'}^k)}{1 + m_1^k + \dots + m_{n'}^k} \\
&= \sum_{l=1}^{m+1} \frac{1}{l!} \sum_{\substack{\sum_{k=1}^l m_k = m+1-l \\ m_1, \dots, m_l \geq 0}} \sum_{n' \geq 1} \sum_{\substack{m_1^k+\dots+m_{n'}^k=m_k, 1 \leq k \leq l \\ 0 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n', 1 \leq k \leq l \\ \text{at least one of } m_{n'}^k, 1 \leq k \leq l \text{ is nonzero}}} \prod_{k=1}^l \frac{a_0^\sigma(m_1^k, \dots, m_{n'}^k)}{m_k + 1} \\
&= \sum_{l=1}^{m+1} \frac{1}{l!} \sum_{\substack{m_1+\dots+m_l=m+1-l \\ m_1, \dots, m_l \geq 0}} \prod_{k=1}^l \left( \frac{1}{m_k + 1} \sum_{n_k \geq 0} \sum_{\substack{m_1^k+\dots+m_{n_k}^k=m_k \\ 1 \leq m_i^k \leq m_{i-1}^k, 2 \leq i \leq n_k}} a_0^\sigma(m_1^k, \dots, m_{n_k}^k) \right)
\end{aligned}$$

$$= \sum_{l=1}^{m+1} \frac{1}{l!} \left( \prod_{k=j+1}^{j+l} \sigma(k) \right) \sum_{\substack{m_1+\dots+m_l=m+1 \\ m_1 \geq 1, \dots, m_l \geq 1}} \prod_{k=1}^l \frac{A_0^\sigma(m_k - 1)}{m_k}, \quad m \geq 0, \quad (\text{B.7})$$

where in the third equality we made the change of variables  $m'_i = m_i - \mathbf{1}_{\{1 \leq i \leq l\}}$ . Plugging the relation

$$\sigma(k) = \gamma + k - 2, \quad k \geq 1,$$

in (B.7), we have

$$A_0^\sigma(m+1) = \sum_{l=1}^{m+1} \binom{l+\gamma-2}{l} \sum_{\substack{m_1+\dots+m_l=m+1 \\ m_1 \geq 1, \dots, m_l \geq 1}} \prod_{k=1}^l \left( \frac{(-\gamma)^{m_k-1}}{m_k} \binom{-1+1/\gamma}{m_k-1} \right), \quad m \geq 0,$$

and Lemma A.1 then shows that  $A_0^\gamma(m+1)$  is the coefficient of  $x^{m+1}$  in the series

$$\begin{aligned} \sum_{l=1}^{\infty} \binom{l+\gamma-2}{l} \left( \sum_{n=1}^{\infty} \frac{(-\gamma)^{n-1}}{n} \binom{-1+1/\gamma}{n-1} x^n \right)^l &= \sum_{l=1}^{\infty} (-l)^l \binom{1-\gamma}{l} (1 - (1 - \gamma x)^{1/\gamma})^l \\ &= (1 - (1 - (1 - \gamma x)^{1/\gamma}))^{1-\gamma} - 1 \\ &= (1 - \gamma x)^{-1+1/\gamma} - 1 \\ &= \sum_{m=1}^{\infty} (-\gamma)^m \binom{-1+1/\gamma}{m} x^m, \end{aligned}$$

which allows us to conclude when  $j = 0$ . When  $j \geq 1$ , we have

$$A_j^\sigma(m+1) = \sum_{l=1}^{m+1} \binom{j+l+\gamma-2}{l} \sum_{\substack{m_1+\dots+m_l=m+1 \\ m_1, \dots, m_l \geq 1}} \prod_{k=1}^l \left( -(-\gamma)^{m_k} \binom{1/\gamma}{m_k} \right), \quad m \geq 0,$$

hence Lemma A.1 shows that, letting

$$Z_\gamma(x) := - \sum_{n=1}^{\infty} (-\gamma)^n \binom{1/\gamma}{n} x^n = 1 - (1 - \gamma x)^{1/\gamma},$$

the quantity  $A_j^\sigma(m+1)$  is the coefficient of  $x^{m+1}$  in the series

$$\begin{aligned} \sum_{l=1}^{\infty} \binom{j+l+\gamma-2}{l} (Z_\gamma(x))^l &= \sum_{l=1}^{\infty} (-l)^l \binom{-(j-1+\gamma)}{l} (Z_\gamma(x))^l \\ &= \frac{1}{(1 - Z_\gamma(x))^{j-1+\gamma}} - 1 \\ &= \sum_{m=1}^{\infty} (-\gamma x)^m \binom{-1 - (j-1)/\gamma}{m}, \end{aligned}$$

which yields (3.11).  $\square$

**Acknowledgement.** This research is supported by the National Research Foundation, Singapore. The work of Q. Huang is supported by the National Natural Science Foundation of China under Grant No. 12501241, the Basic Research Program of Jiangsu under Grant No. BK20251280, the Zhishan Young Scholar Program of Southeast University, the Start-Up Research Fund of Southeast University under Grant No. RF1028624194, and the Jiangsu Provincial Scientific Research Center of Applied Mathematics under Grant No. BK20233002.

## References

- [Aig07] M. Aigner. *A course in enumeration*, volume 238 of *Graduate Texts in Mathematics*. Springer, Berlin, 2007.
- [AN72] K.B. Athreya and P.E. Ney. *Branching processes*, volume Band 196 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York-Heidelberg, 1972.
- [AO21] R.B.R. Azevedo and P. Olofsson. A branching process model of evolutionary rescue. *Math. Biosci.*, 341:108708, 2021.
- [BS84] G.G. Brown and B.O. Shubert. On random binary trees. *Math. Oper. Res.*, 9(1):43–65, 1984.
- [BW18] C.J. Burden and Y. Wei. Mutation in populations governed by a Galton–Watson branching process. *J. Theoret. Probab.*, 120:52–61, 2018.
- [DM10] R. Durrett and S. Moseley. Evolution of resistance and progression to disease during clonal expansion of cancer. *J. Theoret. Probab.*, 77:42–48, 2010.
- [Drm09] M. Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability.
- [Har63] T.E. Harris. *The theory of branching processes*, volume Band 119 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin; Prentice Hall, Inc., Englewood Cliffs, NJ, 1963.
- [HP25] Q. Huang and N. Privault. On the random generation of Butcher trees, 2025. Preprint arXiv:2404.05969, 20 pages, to appear in *Statistics of Random Processes and Optimal Control*. Springer, Heidelberg.
- [HP26] Q. Huang and N. Privault. Probabilistic representation of ODE solutions with quantitative estimates. *J. Math. Anal. Appl.*, 555:130195, 2026.
- [INM06] Y. Iwasa, M.A. Nowak, and F. Michor. Evolution of resistance during clonal expansion. *Genetics*, 172(4):2557–66, 2006.
- [Ken48] D.G. Kendall. On the generalized “birth-and-death” process. *Ann. Math. Statistics*, 19:1–15, 1948.
- [Mło10] W. Młotkowski. Fuss-Catalan numbers in noncommutative probability. *Doc. Math.*, 15:939–955, 2010.
- [Ott49] R. Otter. The multiplicative process. *The Annals of Mathematical Statistics*, 20(2):206–224, 1949.
- [SMJV13] S. Sagitov, B. Mehlig, P. Jagers, and V. Vatutin. Evolutionary branching in a stochastic population model with discrete mutational steps. *Theor. Popul. Biol.*, 83:145–154, 2013.
- [Sta99] R.P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.