DOCTORAL PROGRAMME IN MATHEMATICAL SCIENCES

Department of Mathematics "Federigo Enriques"

PH.D. in MATHEMATICS – XXVII cycle

DISSERTATION

Statistics on multitype Galton-Watson trees

MAT/06

SUBMITTED BY:
Antonio Daziario

SUPERVISOR
Prof. Giacomo Aletti

PH.D. COORDINATOR
Prof. Lambertus Nicolaas van Geemen

A.A. 2013/2014

Introduction

In this thesis is proposed a statistical study of the multitype Galton-Watson trees in order to obtain data on their offspring distribution $\boldsymbol{\mu}$. The investigation is motivated by some parametric simplified models, based on particular two-type Galton-Watson trees, that we propose for the biological process called angiogenesis, i.e. the growth of new blood vessels (Figure 1). This process may occur both in the healthy body, i.e. for curing wounds, and in the sick body. The normal, healthy body keeps a balance of angiogenesis and controls it by some angiogenesis-stimulating growth factors and some angiogenesis inhibitors; conversely, in serious diseases states, the body looses control over angiogenesis.
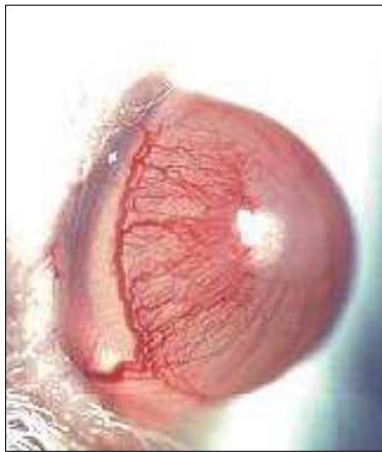

Figure 1: Angiogenesis on a rat cornea (from [9]).

In particular, when new blood vessels grow excessively or insufficiently,

angiogenesis-dependent diseases occur and moreover angiogenesis is very widespread studied in relation with the growth of tumours.

The basic idea of the models proposed is to simplify the structure of a blood vessel as union of its head and the body of the vessel itself. Moreover, the body of the vessel is conceived as union of essential units, all with the same size. Then, we apply the structure of certain two-type Galton-Watson trees to the growth of a blood vessel, where the two-type particles are the heads and the essential units of a blood vessel respectively.

Much has been done in literature concerning models of random trees and multitype Galton-Watson trees, while the statistical point of view has been less studied. The random trees are recently used for statistical mechanics and mathematical physics models, for instance in [18], [22] a ferromagnetic model is studied on locally tree-like random graphs. For more details on the random graphs see [21]. In [19] and [20] are studied broadcasting problems on random trees and there are found relations with some relevant mathematical physics topics. Regarding the multitype Galton-Watson trees, they have been used to achieve the Dawson-Watanabe superprocesses, that are called more simply superprocesses (see for more details [5]). For instance, from the one-type Galton-Watson trees, J.-F. Le Gall has defined the superprocesses and deduced also a tree structure for them (see [1] and [2]). Moreover, in reference to the subject of our models, the angiogenic system has been studied profusely from the mathematical point of view and so several models have been proposed for the growth of the blood vessels. For example, in the works [14], [15] of M.J. Plank and B.D. Sleeman a blood vessel keeps its biological structure and the growth is due to the movement of the endothelial cells (EC), which composed the blood vessel. In the above mentioned works, a reinforced random walk model and a circular random walk model are proposed for the movement of the EC. Moreover, in the work [9] of V. Capasso and D. Morale a blood vessel keeps its biological structure and, similarly to our models, the vessels are modelled as the trajectories developed by the heads, while the behavior of the heads of the vessels is modelled as a stochastic marked

counting process.

Since our statistical study arises from angiogenesis models based on two-type Galton-Watson trees, the main interest is oriented towards the *contour process*, $\mathscr{C}_\tau(s)$, related to a multitype Galton-Watson tree τ and introduced by J.-F. Le Gall (see for instance [2]). In particular, we are interested in the period of $\mathscr{C}_\tau(s)$ and it can be defined as $2\|\tau\|$, where $\|\tau\|$ is the number of edges of τ. Indeed, in our models such a period gives information on the finiteness a.s. of an angiogenic process starting with the head of a new blood vessel, i.e. a type 1 particle. In accordance with the literature of the Galton-Watson processes, one main tool that we have proved and used to investigate the statistical properties of the period of a contour process related to a multitype Galton-Watson tree τ_i, which starts with a type i particle, is the following characterization of the moment-generating function of the period through the offspring distribution $\mu^{(i)}$ of the tree τ_i, for every $i = 1, \ldots, r$,

$$F_i(s) = \mathbb{E}\left[e^{2\|\tau_i\|\cdot s}\right] = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \left(\mu^{(i)}(\boldsymbol{\alpha}) \cdot e^{2s|\boldsymbol{\alpha}|} \cdot \prod_{k=1}^{r} F_k(s)^{\alpha_k}\right), \qquad (1)$$

where r is the number of the types of the particles and $|\boldsymbol{\alpha}| = \alpha_1 + \ldots + \alpha_r$. In particular, the use of a certain scaling in the structure of the trees makes the periods of the related contour processes *triangular arrays*, for which is possible to apply a CLT that has a central rôle in the statistical study of the parameters of the offspring distribution of the trees. Thus, from the equation (1), here we achieve the following relevant result of weak convergence

$$n^{\frac{1}{2}}\left[\sum_{j=1}^{n}\left(\frac{Y_{n,j}^{(i)}}{n^{\alpha+1}}\right) - A_i + \frac{o(n^{\alpha+1})}{n^{\alpha+1}}\right] \xrightarrow{d} \mathscr{N}(0, B_i), \ n \to +\infty, \qquad (2)$$

where $Y_{n,1}^{(i)}, Y_{n,2}^{(i)}, \ldots, Y_{n,n}^{(i)}$ are n i.i.d. copies of $Y_n^{(i)}$, which is the period of the contour process of a multitype Galton-Watson tree starting from a type i particle at the *n-th* step of a scaling deduced by the Feller's one and with

a certain offspring distribution $\boldsymbol{\mu}_n$. The parameters $A_i \in \mathbb{R}$ and $B_i > 0$ concern the expected value and the variance of $Y_n^{(i)}$, for every $i = 1, \ldots, r$. Moreover, to improve and simplify our models, we also have treated two-type Galton-Watson trees with a particular offspring distribution $\boldsymbol{\mu}$ such that a type i particle may produce no particles, one particle of type i or two particles, the first of type 1 and the second of type 2, for every $i = 1, 2$. All these kind of trees are called *full binary trees with survivals* and, in the particular case when a particle may produce none or two particles, we talk about *full binary trees*. We have conducted a combinatorial investigation on the *full binary trees with survivals* and a relevant family of integers has appeared, the *Narayana numbers* $N(k, l)$ (see for more details [12]),

$$N(k, l) = \frac{1}{k} \binom{k}{l} \binom{k}{l-1}, k \geq 1, l = 1, \ldots, k.$$

The Narayana numbers are a sort of generalization of the *Catalan numbers* (see [10]), and from them we have obtained, in a non linear way and in the a.s. finite case, the likelihood of the number of left and right vertices having exactly two children,

$$\mathscr{L}(P, Q | n, m) = N(n + m, m + 1) \cdot P^{m+1} (1 - P)^n Q^n (1 - Q)^m, \qquad (3)$$

where $n \geq 1, m \geq 0$ are the number of the left and right vertices respectively, and the paramters $P, Q \in (0, 1)$ depend only on the offspring distribution $\boldsymbol{\mu}$ of the tree. Moreover, through the Narayana numbers, we have outlined an interesting rappresentation of the *full binary trees* as two-dimensional decompositions of the integers (see for more details [11]).

Now we present the structure of the thesis. In Chapter 1 we recall the fundamental theory of the multitype Galton-Watson processes and the multitype Galton-Watson trees. Then we give also a brief introduction to the superprocesses. We generalize and use the scaling limit which forms a superprocess in Chapter 5. Chapter 2 presents some relevant theoretical results on the period of the contour process related to the multitype Galton-Watson

trees, like the fundamental equation (1), the proof of the weak convergence illustred in (2), and the results on the *full binary trees* and the *full binary trees with survivals*, one of which is the likelihood (3). In Chapter 3 we introduce three stochastic models for the growth of the blood vessels, that is the reinforecd random walk and the circular random walk models used by M.J. Plank and B.D. Sleeman in [14] and [15], and the model proposed by V. Capasso and D. Morale in [9]. Our first model is presented in Chapter 4. It is based on particular two-type Galton-Watson processes, both in discrete and in continuous time. Then, in Chapter 5 we present other models based on the *full binary trees* and on the *full binary trees with survivals* introduced above. Eventually, we apply the scaling of the particles system related to the convergence (2) on the *full binary trees with survivals* model, that we will call *scaling* model.

Contents

List of Figures

List of Tables

Part I

Background and Theoretical Results

Chapter 1

Galton-Watson trees and Superprocesses

In this chapter we recall some basic notions about the multytipe Galton-Watson processes both in discrete and in continuous time, the multitype Galton-Watson trees and the construction of the superprocesses. In particular, in Section 1.1 we see the fundamental definitions and properties of the Galton-Watson processes as they are presented in the classical literature ([13]). In Section 1.2 we introduce the multitype Galton-Watson trees, i.e. the genealogical structure for the multitype Galton-Watson processes in the discrete time. At first we see the defintions and the construction of the multitype trees and then we consider the multitype Galton-Watson trees. For Section 1.2 we refer to [8]. Further details on random trees and their applications are also presented by J.-F. Le Gall ([3]), and J.-F. Le Gall and G. Miermont ([4]). Finally, in Section 1.3 we give a brief explanation of superprocesses, indeed their construction is due to a scaling from which we prove a relevant weak-convergence theorem in Chapter 2. For the superprocesses we will refer to the work of J.-F. Le Gall, [1], and L.G. Gorostiza - J.A. Lopez-Mambela, [6].

1.1 Multitype Galton-Watson processes

In this Section we report and use notations from the classical work about the branching processes by K.B. Athreya - P.E. Ney (see [13, Chapter V, Sec. 1, 2 and 3 for Section 1.1.1, and Chapter V, Sec. 7 for Section 1.1.2]).

1.1.1 Discrete time case

Let $r \geq 1$, and consider a r-type system of particles which have unit-time lifetimes and every unit-time interval is called generation. At each generation each particle makes children (i.e. offspring) according to its own type. A multitype Galton-Watson process in discrete time counts the number of particles, for each type, in each generation. As shown in [13, Chapter V], it is well known that to completely define the offspring of the particles we need a vector of r generating function. The j-th genereting function, $f^{(j)}$, determines the offspring distribution of particles produced by a type j particle. So, $\forall \, j = 1, \ldots, r$, we let

$$f^{(j)}(s_1, \ldots, s_r) = \sum_{i_1, \ldots, i_r \geq 0} p^{(j)}(i_1, \ldots, i_r) \, s_1^{i_1} \cdots s_r^{i_r}, \; 0 \leq s_j \leq 1, \qquad (1.1)$$

where $p^{(j)}(i_1, \ldots, i_r)$ is the probability that a type j particle produces i_1 particles of type 1,..., i_r particles of type r, that is $p^{(j)}(i_1, \ldots, i_r)$ is called the *offspring distribution* of the type j particles. To lighten the notations, we adopt the vectorial form

$$\boldsymbol{i} = (i_1, \ldots, i_r) \in \mathbb{N}^r,$$

$$\boldsymbol{p}(\boldsymbol{i}) = \left(p^{(1)}(\boldsymbol{i}), \ldots, p^{(r)}(\boldsymbol{i}) \right), \; \text{with} \; \sum_{\boldsymbol{i} \in \mathbb{N}^r} p^{(j)}(\boldsymbol{i}) = 1, \; \forall \, j = 1, \ldots, r,$$

$$\boldsymbol{f}(\boldsymbol{s}) = \left(f^{(1)}(\boldsymbol{s}), \ldots, f^{(r)}(\boldsymbol{s}) \right),$$

and so we may write (1.1) in the following way

$$\boldsymbol{f}\left(\boldsymbol{s}\right) = \sum_{\boldsymbol{i}\in\mathbb{N}^r} \boldsymbol{p}\left(\boldsymbol{i}\right)\boldsymbol{s}^{\boldsymbol{i}}, \ \ \boldsymbol{s}\in[0,1]^r. \tag{1.2}$$

Now we give the definition of a multitype Galton-Watson process in discrete time.

Definition 1.1.1. A process $\{\boldsymbol{Z}_n : \mathbb{N} \to \mathbb{N}^r, n \geq 0\}$ is a r-type Galton-Watson process if it is a Markov process , with transition function probabilities

$$\mathbb{P}\left(\boldsymbol{i},\boldsymbol{j}\right) = \mathbb{P}\left(\boldsymbol{Z}_{n+1} = \boldsymbol{j}|\boldsymbol{Z}_n = \boldsymbol{i}\right) = \text{coefficient of } \boldsymbol{s}^{\boldsymbol{j}} \text{ in } [\boldsymbol{f}\left(\boldsymbol{s}\right)]^{\boldsymbol{i}}, \ \ \boldsymbol{i},\boldsymbol{j}\in\mathbb{N}^r,$$

where $\boldsymbol{i} = (i_1,\ldots,i_r)$ and

$$[\boldsymbol{f}\left(\boldsymbol{s}\right)]^{\boldsymbol{i}} = \left(\prod_{k=1}^{i_1} f^{(1)}\left(\boldsymbol{s}\right)\right)\cdots\left(\prod_{k=1}^{i_r} f^{(r)}\left(\boldsymbol{s}\right)\right) = \left(f^{(1)}\left(\boldsymbol{s}\right)\right)^{i_1}\cdots\left(f^{(r)}\left(\boldsymbol{s}\right)\right)^{i_r}.$$

When the process is initiated in state $\boldsymbol{j}$, we will denote it by $\boldsymbol{Z}_n^{(j)}$ and in vectorial form we write

$$\boldsymbol{Z}_n = (Z_{n,1},\ldots,Z_{n,r}) \ \text{ and } \ \boldsymbol{Z}_n^{(j)} = \left(Z_{n,1}^{(j)},\ldots,Z_{n,r}^{(j)}\right),$$

where $Z_{n,k}^{(j)}$ is the number of type k particles in the n-th generation for a process with $\boldsymbol{Z}_0 = \boldsymbol{j}$. In particular when $\boldsymbol{j} = \boldsymbol{e}_i$ we write $\boldsymbol{Z}_n^{(e_i)} = \boldsymbol{Z}_n^{(i)}$.

Remark 1. Note that, given the offspring distribution $\boldsymbol{p}\left(\boldsymbol{i}\right)$, the whole process $\boldsymbol{Z}_n$, or $\boldsymbol{Z}_n^{(j)}$, is completely described.

Remark 2. The process $\{\boldsymbol{Z}_n^{(i)}; n \geq 0\}$ is equipped with the *additive property*, that is $\{\boldsymbol{Z}_n^{(i)}; n \geq 0\}$ is the sum of $i_1 + \ldots + i_r$ independet process, i_k of which are initiated in state $\boldsymbol{e}_k$, $k = 1,\ldots,r$.

Remark 3. Although most of the studies of the branching process are directly in terms of the generating function $\boldsymbol{f}\left(\boldsymbol{s}\right)$, it is important the probabilistic structure underlying this analytical setting. From classic results (see [13,

Chapter I, pag.3 and Chapter V, pag. 184]) we can make reference to the basic probability space $(\Omega, \mathbb{F}, \mathbb{P}_j)$ for the process $\boldsymbol{Z}_n^{(j)}(\omega) : \mathbb{N} \times \Omega \to \mathbb{N}^r$, we shall leave out ω throughout. Here Ω is the space of *trees* that begin from a type j particle, i.e. $\omega \in \Omega$ represents the generation number, the ancestors, and the offspring of each type particle of a process initiated from a type j particle. $\mathbb{F}$ is generated by the cylinder sets of Ω, and $\mathbb{P}_j$ is the probability measure on $(\Omega, \mathbb{F})$ when the process is initiated with $\boldsymbol{Z}_0 = \boldsymbol{e}_j$, $j = 1, \ldots, r$.

As shown in [13, Chapter V], exists a $r \times r$ matrix $\boldsymbol{M} = \{m_{jk}; j, k = 1, \ldots, r\}$ such that

$$\mathbb{E}[\boldsymbol{Z}_n | \boldsymbol{Z}_0] = \boldsymbol{Z}_0 \boldsymbol{M}^n \tag{1.3}$$

where m_{jk} is the expected number of type k particles from a type j particle in one generation. So, by the definitions of the probabilities $p^{(j)}(i_1, \ldots, i_r)$ and the generating functions $f^{(j)}(s_1, \ldots, s_r)$, we have

$$m_{jk} = \sum_{i_1, \ldots, i_r \geq 0} p^{(j)}(i_1, \ldots, i_r) i_k = \left. \frac{\partial f^{(j)}(\boldsymbol{s})}{\partial s_k} \right|_{\boldsymbol{s} = \boldsymbol{1}}.$$

If $\boldsymbol{M}$ is stricly positive, then the process $\{\boldsymbol{Z}_n; n = 0, 1, \ldots\}$ is called *positive regular*.

Moreover, if $\boldsymbol{f}(\boldsymbol{s}) = \boldsymbol{M} \boldsymbol{s}^T$ then each particle has exactly one-child offspring, and so the branching process is called *singular* process. We assume nonsingularity throughout.

Another important quantity to introduce is the extinction probability from a type j particle

$$q^{(j)} = \mathbb{P}\left(\boldsymbol{Z}_n^{(j)} = \boldsymbol{0} \text{ for some } n\right) \tag{1.4}$$

and we will denote $\boldsymbol{q} = (q^1, \ldots, q^r)$.

Now we report the most important theorem about the extinction probability (see [13, Chapter V, pag. 186, Theorem 2])

Theorem 1.1.1. *Let $\{\boldsymbol{Z}_n; n = 0, 1, \ldots\}$ be an r-type Galton-Watson positive regular and nonsingular process, and let ρ be the maximum eigenvalue of the matrix $\boldsymbol{M}$. Then*

(i) *if $\rho \leq 1$ then $\boldsymbol{q} = \boldsymbol{1}$. If $\rho > 1$ then $\boldsymbol{q} < \boldsymbol{1}$.*

(ii) *the only solution of $\boldsymbol{f}(\boldsymbol{s}) = \boldsymbol{s}$ in $[0, 1[^r$ is $\boldsymbol{q}$.*

Moreover, we shall call the process supercritical, critical, or subcritical according as $\rho >, = or < 1$.

Remark 4. Since $\boldsymbol{f}(\boldsymbol{1}) = \boldsymbol{1}$, if the equation $\boldsymbol{f}(\boldsymbol{s}) = \boldsymbol{s}$ has no solution in $[0, 1[^r$, then $\boldsymbol{q} = \boldsymbol{1}$.

1.1.2 Continuous time case

As in Section 1.1.1, here we consider a finite number, r, of particle types. In continuous time case the particles have random lifetimes and so the concept of generations does not make sense. According to the offspring distribution of the particles, a continuous time r-type Galton-Watson process counts the number of particles, for each type, at time $t \geq 0$. Moreover, in reference to [13, Chapter V, Sec. 7] the lifetimes of the particles are exponentially distribuited and the process is a continuous time Markov process on $\mathbb{N}^r$.

Let $Z_j(t)$ be the number of type j particles at time t, and set

$$\boldsymbol{Z}(t) = (Z_1(t), \ldots, Z_r(t)) : \mathbb{R}^+ \to \mathbb{N}^r.$$

Like in discrete time case, we can refer to the basic probability space $(\Omega, \mathbb{F}, \mathbb{P})$ for the process $\boldsymbol{Z}(t, \omega) : \mathbb{R}^+ \times \Omega \to \mathbb{N}^r$, we shall leave out ω throughout, and $\boldsymbol{Z}(t)$ is equipped with the *additive property* seen in the previous section. Now we give the definition of the continuous time Galton-Watson branching process

Definition 1.1.2. A stochastic process $\{\boldsymbol{Z}(t); t \geq 0\}$ is called *r-dimensional continuous Markov branching process* (i.e. r-type continuous time Galton-Watson branching process) if:

(i) its state space is $\mathbb{N}^r$;

(ii) it is a stationary strong Markov process with respect to the filtration

$$\mathbb{F}_t = \sigma\{\{\boldsymbol{Z}\left(s\right); s \le t\};$$

(iii) the transition function probabilities

$$\mathbb{P}\left(\boldsymbol{i},\boldsymbol{j};t\right) = \mathbb{P}\left(\boldsymbol{Z}\left(t+h\right) = \boldsymbol{j} \mid \boldsymbol{Z}\left(h\right) = \boldsymbol{i}\right), \forall\ h \ge 0,\ t > 0$$

satisfy

$$\sum_{\boldsymbol{j} \in \mathbb{N}^r} \mathbb{P}\left(\boldsymbol{i},\boldsymbol{j};t\right) \boldsymbol{s}^{\boldsymbol{j}} = \prod_{k=1}^{r} \left[\sum_{\boldsymbol{j} \in \mathbb{N}^r} \mathbb{P}\left(\boldsymbol{e}_k,\boldsymbol{j};t\right) \boldsymbol{s}^{\boldsymbol{j}} \right]^{i_k},$$

$\forall\ \boldsymbol{i} \in \mathbb{N}^r$ and $\boldsymbol{s} \in \left[0,1\right]^r$.

Remark 5. As in the discrete case, $\{\boldsymbol{Z}^{(j)}\left(t\right); t \ge 0\}$ represents a process initiated in the state $\boldsymbol{j}$; and when $\boldsymbol{j} = \boldsymbol{e}_i$ we write $\boldsymbol{Z}^{(e_i)}\left(t\right) \equiv \boldsymbol{Z}^{(i)}\left(t\right)$.

In accordance to [13, Chapter V, pag.200-201], we know that the transition probabilities in (iii) are solutions of the Kolmogorv forward and backward equations, whose parameters are

$$\boldsymbol{a} = \left(a_1,\ldots,a_r\right) \in \mathbb{R}_+^r, \tag{1.5}$$

where a_i is the strictly positive parameter of the exponential distribution of lifetime of the type i particles and

$$\boldsymbol{p}\left(\boldsymbol{j}\right) = \left(p^{(1)}\left(\boldsymbol{j}\right),\ldots,p^{(r)}\left(\boldsymbol{j}\right)\right), \tag{1.6}$$

with

$$\forall\ i = 1,\ldots,r, \sum_{\boldsymbol{j} \in \mathbb{N}^r} p^{(i)}\left(\boldsymbol{j}\right) = 1 \text{ and } \boldsymbol{j} = \left(j_1,\ldots,j_r\right) \in \mathbb{N}^r.$$

where $p^{(i)}\left(\boldsymbol{j}\right)$ is the probability that a type i particle produces j_1 particles of type 1,..., j_r particles of type r at the end of its lifetime, and $\boldsymbol{p}\left(\boldsymbol{j}\right)$ is called the offspring distribution.

In other words, the process $\boldsymbol{Z}\left(t\right)$ is completely determinated by the param-

eters (1.5) and (1.6).

Remark 6. From the theory (see [13], Chapter V, pag.201) we know that a sufficient condition to be sure that a.s. there cannot be an infinite number of particles produced in a finite time is the following

$$\left. \frac{\partial f^{(i)}(\boldsymbol{s})}{\partial s_j} \right|_{\boldsymbol{s=1}} < \infty, \ \forall \, i, j = 1, \dots, r,$$

where, as in the discrete case,

$$f^{(i)}(\boldsymbol{s}) = \sum_{\boldsymbol{j} \in \mathbb{N}^r} p^{(i)}(\boldsymbol{j}) \, \boldsymbol{s^j}, \ \boldsymbol{s} \in [0,1]^r.$$

Then, from Remark 6, we know that the expected value of j type particles produced by a single type i particle after an interval time t is finite a.s., and so is well defined the *mean* matrix $\boldsymbol{M}(t), t \geq 0$ (the analogous of the matrix $\boldsymbol{M}^n$ in (1.3))

$$\boldsymbol{M}(t) = \{m_{ij}(t) \, ; i, j = 1, \dots, r\}, \tag{1.7}$$

where

$$m_{ij}(t) := \mathbb{E}\left[Z_j^{(i)}(t) \right] < \infty.$$

It is well known that $\{\boldsymbol{M}(t) \, ; t \geq 0\}$ is a semigroup ([13], Chapter V, pag.202) and this implies the existence of a matrix $\boldsymbol{A} = \{a_{ij}; i, j = 1, \dots, r\}$, called the *infinitesimal generator* of $\{\boldsymbol{M}(t) \, ; t \geq 0\}$, such that

$$\boldsymbol{M}(t) \equiv e^{\boldsymbol{A}t} = \sum_{k=0}^{+\infty} \frac{t^k \boldsymbol{A}^k}{k!}, \tag{1.8}$$

where

$$a_{ij} = a_i \left(\left. \frac{\partial f^{(i)}(\boldsymbol{s})}{\partial s_j} \right|_{\boldsymbol{s=1}} - \delta_{ij} \right), \ \forall \, i, j = 1, \dots, r. \tag{1.9}$$

From classic results (see [13], chapter V, pag.202-203) we know that the matrix $\boldsymbol{A}$ has r eigenvalues, $\lambda_1, \dots, \lambda_r$, which can be arranged in the following

way

$$\lambda_1 > \operatorname{Re} \lambda_2 \geq \operatorname{Re} \lambda_3 \geq \ldots \geq \operatorname{Re} \lambda_r,$$

and, likewise the discrete case (Theorem 1.1.1), we shall call the process $\boldsymbol{Z}(t)$ *supercritical*, *critical*, or *subcritical* according as $\lambda_1 >, =$ or < 0.

Finally, as in Section 1.1.1, from [13, equation (25), Chapter 5, Section 7] we have that the extinction probability vector $\boldsymbol{q} = \left(q^{(1)}, q^{(2)}, \ldots, q^{(r)} \right)$, where

$$q^{(i)} = \mathbb{P}\left(\boldsymbol{Z}^{(i)}(t) = \boldsymbol{0}, \text{ for some } t \right),$$

is the unique solution of the equation

$$\boldsymbol{u}(\boldsymbol{s}) = \boldsymbol{0}, \ \boldsymbol{s} \in [0, 1[^r,$$

where

$$\boldsymbol{u}(\boldsymbol{s}) = \left(u^{(1)}(\boldsymbol{s}), \ldots, u^{(r)}(\boldsymbol{s}) \right) =$$

$$= \left(a_1 \left(f^{(1)}(\boldsymbol{s}) - s_1 \right), \ldots, a_r \left(f^{(r)}(\boldsymbol{s}) - s_r \right) \right).$$

Remark 7. Remark 4 holds also in this case.

1.2 Multitype Galton-Watson trees

In this Section we add a genealogical structure to the discrete-time multitype branching processes. We refer to G. Miermont [8, Sec. 1.3, 1.4] for the notions of the multitype trees and the multitype Galton-Watson trees. A complete dissertation and more details on the Galton-Watson trees can be found in [23].

1.2.1 Multitype trees

For $n \geq 0$, let U be the infinite-regular tree

$$U = \bigcup_{n \geq 0} \mathbb{N}^n,$$

where if $\boldsymbol{u} \in U$, then $\boldsymbol{u} = (u_1, \ldots, u_n) \in \mathbb{N}^n$, $u_i \in \mathbb{N}$, $i = 1, \ldots, n$. We use the convention $\mathbb{N}^0 = \{\varnothing\}$ throughout.

For $\boldsymbol{u} = (u_1, \ldots, u_n)$, $\boldsymbol{v} = (v_1, \ldots, v_m) \in U$, we let

$$\boldsymbol{u}\boldsymbol{v} = (u_1, \ldots, u_n, v_1, \ldots, v_m) \in U$$

be their concatenation and $|\boldsymbol{u}| = n, |\boldsymbol{v}| = m$ their length (with the convention $|\varnothing| = 0$).

Let $\boldsymbol{u} \in U$ and $A \subseteq U$, we let

$$\boldsymbol{u}A = \{\boldsymbol{u}\boldsymbol{v} | \boldsymbol{v} \in A\},$$

and say that $\boldsymbol{u}$ is a *prefix* of $\boldsymbol{w}$ if

$$\boldsymbol{w} \in \boldsymbol{u}U,$$

and we write $\boldsymbol{u} \vdash \boldsymbol{w}$.

Now we give the definition of a *planar tree*

Definition 1.2.1 (Planar tree)**.** A planar tree is a finite subset τ of U such that

   (*i*) $\varnothing \in \tau$, and it is called the *root* of τ,

   (*ii*) $\forall\, \boldsymbol{u} \in U$ and $i \in \mathbb{N}$, if $\boldsymbol{u}i \in \tau \Rightarrow \boldsymbol{u} \in \tau$, and $\boldsymbol{u}j \in \tau$ for every $1 \leq j \leq i$.

Moreover, an element $\boldsymbol{u} \in \tau$ is called a *vertex* of τ, and $\|\tau\|$ is the number of edges of the tree τ.

We let T be the set of all planar trees, which we refer to as *trees* in the sequel. Now we give some important definitions about trees.

Definition 1.2.2 (Number of children of a vertex)**.** Let $\tau \in T$ and $\boldsymbol{u} \in \tau$, the number

$$c_\tau(\boldsymbol{u}) = \max\{i \in \mathbb{N}^+ | \boldsymbol{u}i \in \tau\}, \text{ with } \boldsymbol{u}0 = \boldsymbol{u},$$

is the number of children of $\boldsymbol{u}$.

Definition 1.2.3 (Leaves)**.** Let $\tau \in T$ then

$$\{\boldsymbol{u} \in \tau | c_\tau(\boldsymbol{u}) = 0\}$$

is the set of the *leaves* of τ.

Definition 1.2.4 (Ancestors)**.** Let $\tau \in T$ and $\boldsymbol{u}, \boldsymbol{v} \in \tau$, then $\boldsymbol{u}$ is an *ancestor* of $\boldsymbol{v}$ if $\boldsymbol{u} \vdash \boldsymbol{v}$.

Any tree $\tau \in T$ is endowed with the *depth-first order*,

Definition 1.2.5 (Depth-first order $\prec$)**.** Let $\tau \in T$ and $\boldsymbol{u}, \boldsymbol{v} \in \tau$, then

$$\boldsymbol{u} \prec \boldsymbol{v} \text{ if } \boldsymbol{u} \vdash \boldsymbol{v} \text{ or } \boldsymbol{u} = \boldsymbol{w}\boldsymbol{u}', \ \boldsymbol{v} = \boldsymbol{w}\boldsymbol{v}', \text{where } u_1' < v_1'.$$

Now we show an example of the depth-first order for a planar tree, in which we use the *left-right* notation to label the children of each vertex from the holder one to the younger one.
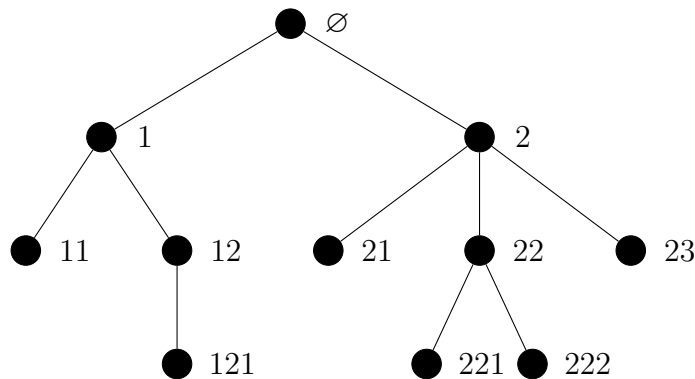
Figure 1.1: A planar tree.

Example 1.2.1. The tree $\tau \in T$ in Figure 1.1 can be written according to the depth-first order in the following way

$$\tau = \{\varnothing, 1, 11, 12, 121, 2, 21, 22, 221, 222, 23\}.$$

For instance, we write $2 \prec 221$ because if we denote $\boldsymbol{u} = 2$ and $\boldsymbol{v} = 221$, then we have that

$$\boldsymbol{v} = \boldsymbol{u}21 \Rightarrow \boldsymbol{v} \in \boldsymbol{u}U,$$

so $\boldsymbol{u}$ is a prefix of $\boldsymbol{v}$, i.e. $\boldsymbol{u} \vdash \boldsymbol{v}$ and, by the Defintion 1.2.5, $\boldsymbol{u} \prec \boldsymbol{v}$. Moreover, we write $12 \prec 221$ because if we denote $\boldsymbol{u} = 12$ and $\boldsymbol{v} = 221$, then we have that

$$\boldsymbol{u} = \boldsymbol{w}\boldsymbol{u}' \text{ and } \boldsymbol{v} = \boldsymbol{w}\boldsymbol{v}',$$

where

$$\boldsymbol{w} = \varnothing, \ \boldsymbol{u}' = \boldsymbol{u} = 12, \ \boldsymbol{v}' = \boldsymbol{v} = 221,$$

and so

$$u_1' = 1 < v_1' = 2,$$

and, by the Defintion 1.2.5, $\boldsymbol{u} \prec \boldsymbol{v}$.

Remark 8. From now on, we use the *left-right* order to represent the children of each vertex from the holder one to the younger one.

Now we are able to introduce the r-type planar trees, or simply the r-type trees.

Definition 1.2.6 (r-type planar trees)**.** Let $r \geq 1$, then a r-type planar tree is a pair (τ, e_τ), where

1. $\tau \in T$,

2. $e_\tau : \tau \longrightarrow \{1, \ldots, r\}$, i.e. $\forall \ \boldsymbol{u} \in \tau, \ e_\tau(\boldsymbol{u}) \in \{1, \ldots, r\}$ is called the type of the vertex $\boldsymbol{u}$.

Moreover, let $T^{(r)}$ be the set of r-type trees and we let

$$T_i^{(r)} = \{\tau \in T^{(r)} | e_\tau(\varnothing) = i\} \ \forall \ i \in \{1, \ldots, r\}.$$

Now the purpose is to count the children of a vertex, according to the type. We use the following notation $\{1, \ldots, r\} = [r]$. To do this, at first we define the *counter map*.

Definition 1.2.7 (Counter map)**.** Let $r \geq 1$ and

$$W_r = \bigcup_{n \geq 0} [r]^n ,$$

be the set of finite, possibly empty, $[r]$-valued sequences, then the counter map, $\boldsymbol{p} : W_r \longrightarrow \mathbb{N}^r$, is such that

$$\boldsymbol{p}(\boldsymbol{w}) = (p_1(\boldsymbol{w}), \dots, p_r(\boldsymbol{w})) \ \forall \ \boldsymbol{w} \in W_r,$$

where $p_i(\boldsymbol{w})$ is the number of i in $\boldsymbol{w}$, $\forall \ i = 1, \dots, r$.

So, $\forall \ (\tau, e_\tau) \in T^{(r)}$ and $\forall \ \boldsymbol{u} \in \tau$ we can define the following vector,

$$\boldsymbol{w}_\tau(\boldsymbol{u}) = (e_\tau(\boldsymbol{u}j), 1 \leq j \leq c_\tau(\boldsymbol{u})) \in W_r,$$

and then

$$\boldsymbol{p}(\boldsymbol{w}_\tau(\boldsymbol{u})) \in \mathbb{N}^r$$

is the vector of the number of children of $\boldsymbol{u}$ for each type.

Remark 9. Note that if u is a leave (i.e. $c_\tau(\boldsymbol{u}) = 0$), then $\boldsymbol{w}_\tau(\boldsymbol{u}) = \{\varnothing\}$ and $\boldsymbol{p}(\boldsymbol{w}_\tau(\boldsymbol{u})) = 0 \in \mathbb{N}^r$.

Remark 10. Note that the graphical representation of a multitype tree is the same of a one-type planar tree (Figure 1.1). Indeed, neglecting the type of the particles, it may be considered as one-type tree. Moreover, the set of the vertices of a multitype tree is ordered according to the first-depth order.

1.2.2 Galton-Watson trees

In this section we treat the multitype planar trees where each vertex has a number of children of certain type according to the offspring distribution.

Definition 1.2.8 (Offspring distribution)**.** Let $r \geq 1$ and $\boldsymbol{\xi} = \left(\xi^{(1)}, \dots, \xi^{(r)}\right)$ be a family of probabilities on the σ-algebra $\sigma(W_r)$.

We say that the family of probabilities

$$\boldsymbol{\mu} = \boldsymbol{p} * \boldsymbol{\xi} = \left(\boldsymbol{p} * \xi^{(1)}, \ldots, \boldsymbol{p} * \xi^{(r)} \right),$$

is an offspring distribution on the σ-algebra $\sigma \left(N^r \right)$

$$
\begin{array}{ccc}
W_r & \xrightarrow{\ \boldsymbol{p}\ } & \mathbb{N}^r \\
{\scriptstyle \xi} \downarrow & \swarrow {\scriptstyle \boldsymbol{\mu} = \boldsymbol{p} * \boldsymbol{\xi}} & \\
{[0,1]} & &
\end{array}
$$

where $\boldsymbol{p}$ is the counter map defined in Definition 1.2.7 and $\boldsymbol{p} * \xi^{(i)} \equiv \xi^{(i)} \circ \boldsymbol{p}^{-1}$ (i.e. $\boldsymbol{p} * \xi^{(i)}$ is the push-forward of $\xi^{(i)}$ by $\boldsymbol{p}$). Now, we build a distribution on $T_i^{(r)}$, $\forall\, i \in [r]$, such that

1. different vertices have independent offspring

2. type j vertices have a set of children with types given by a sequence $\boldsymbol{w} \in W_r$ with probability $\xi^{(j)} \left(\boldsymbol{w} \right)$.

To do this, $\forall\, j \in [r]$ and $\forall\, \boldsymbol{u} \in U$, let $\boldsymbol{C}_{\boldsymbol{u},j} = \left(C_{\boldsymbol{u},j} \left(l \right), 1 \leq l \leq \left| \boldsymbol{C}_{\boldsymbol{u},j} \right| \right)$ be a family of independent random variables , such that $\boldsymbol{C}_{\boldsymbol{u},j}$ has law $\xi^{(j)}$. Now, recursively, we construct a subset $\tau \subset U$ and a mark-map $e_\tau : \tau \to [r]$ in the following way

1. $\varnothing \in \tau$

2. $e_\tau \left(\varnothing \right) = i$

3. if $\boldsymbol{u} \in \tau$, $e \left(\boldsymbol{u} \right) = j$, then, with probability $\xi^{(j)} \left(\boldsymbol{C}_{\boldsymbol{u},j} \right)$, $\boldsymbol{u}l \in \tau$ if and only if $1 \leq l \leq \left| \boldsymbol{C}_{\boldsymbol{u},j} \right|$ and then $e \left(\boldsymbol{u}l \right) = C_{\boldsymbol{u},j} \left(l \right)$.

A pair $(\tau, e_\tau) \in T^{(r)}$ thus defined is called a *Galton-Watson multitype trees*.

Remark 11. It is easy to check that the subset $\tau \subset U$ has the properties of a planar tree (it might be infinite). Moreover, from the construction we have that

$$\boldsymbol{Z}_n \left(\tau \right) = \left(\# \{ \boldsymbol{u} \in \tau : \left| \boldsymbol{u} \right| = n, e \left(\boldsymbol{u} \right) = i \}, \ i \in [r] \right), \ n \geq 0, \qquad (1.10)$$

is a multitype Galton-Watson process (with discrete time) with offspring distribution $\boldsymbol{\mu} = \boldsymbol{p} * \boldsymbol{\xi}$.

From now on, we call $\boldsymbol{\mu}$-*GW tree* a Galton-Watson multitype tree that induces a multitype Galton-Watson process, in discrete time, with offspring distribution $\boldsymbol{\mu} = \boldsymbol{p} * \boldsymbol{\xi}$.

1.3   Superprocesses

A part of our statistical results in Chapter 2 is related to a scaling limit from which the superprocesses arise. In this Section we briefly introduce both the one-type and the multi-type superprocesses. In particular, in Section 1.3.1 we see how J.-F. Le Gall deduce the one-type superprocesses (we refer to [1, Sec. 1.2, 1.3 and 2.1]). In Section 1.3.2 we introduce the multitype superprocesses and we refer to L.G. Gorostiza - J.A. Lopez-Mimbela ([6, Sec. 3 and 4]). Because our interest is focused on the scaling limit, we refer to the above mentioned works for further details.

1.3.1   One type superprocesses

We start with the definition of spatial branching processes. A spatial branching process is a combination of a branching process with a spatial motion, which is usually a E-valued Markov process W, where E is a Polish space. In the discrete-time case, the branching process is a Galton-Watson process and so the particles alive at time (generation) n move from time n to time $n + 1$ according to the law of W. At time $n + 1$ the new born particles start from the final position of they parents, and so on. In this section we show the approximation of the continuous-state spatial branching process. We begin introducing the continuous-state branching processes (CSBP). A CSBP is the continuous-time analogue of the Galton-Watson process, i.e. a CSBP describes the behavior in continuous time of a particles system with values in $\mathbb{R}_+$. Now consider a sequence of Galton-Watson processes

$\left(Z_n^{(h)}, n \geq 0 \right)_{h \geq 1}$, with initial value z_h and offspring distribution p_h. We know from literature (see [1]) that if there exists a sequence $a_h \nearrow \infty, h \to \infty$, such that the scaled process

$$\left(\frac{1}{a_h} Z_{\lfloor ht \rfloor}^{(h)}, t \geq 0 \right) \tag{1.11}$$

converge to a process $(Y_t, t \geq 0)$, at least in the sense of weak convergence of the finite-dimensional marginals, then Y must be a continuous-state branching process (CSBP). The process in (1.11) is called the *scaled process* at the step h, and moreover, at the *h-th* step of the scaling each particles has "mass" equal to $1/a_h$.

A very interesting case is when $p_h = p, \ \forall \ h$, with $\sum n \cdot p(n) = 1$. Then the convergence above holds with $a_h = h$ and $h^{-1} z_h \longrightarrow x \geq 0$, and this is called *the Feller approximation* for banching processes. Now we show an approximation method to construct a continuous-state spatial branching process. Like in (1.11), consider a sequence of Galton-Watson process $\left(Z_n^{(h)}, n \geq 0 \right)_{h \geq 1}$ such that

$$\left(\frac{1}{a_h} Z_{\lfloor ht \rfloor}^{(h)}, t \geq 0 \right) \xrightarrow{f.d.} (Y_t, t \geq 0), \tag{1.12}$$

where $(Y_t, t \geq 0)$ is a CSBP.

Then, if $Z_0^{(h)} = z_h$, we consider z_h points $x_1^h, \ldots, x_{z_h}^h$ in the space E, and we assume that the z_h initial particles start respectively at $x_1^h, \ldots, x_{z_h}^h$ and then move according to the law of W between times $t = 0$ and $t = 1/h$. At time $t = 1/h$ each particles is replaced by its children, which also move from $t = 1/h$ to $t = 2/h$ according to the law of W, indipendently of each other, and so on. Then, for every $t \geq 0$ consider $\xi_t^{h,i}$ the position in the space E of the i-th particle alive at time t, $i = 1, \ldots, Z_{\lfloor ht \rfloor}^{(h)}$, and the random measure $N_t^{(h)}$

$$N_t^{(h)} = \frac{1}{a_h} \sum_{i=1}^{Z_{\lfloor ht \rfloor}^{(h)}} \delta_{\xi_t^{h,i}}. \tag{1.13}$$

If $\mathscr{M}_f(E)$ is the space of the finite measures on E, then $N_t^{(h)} \in \mathscr{M}_f(E)$, which is equipped with the topology of weak convergence. From the (1.13), we have that the total mass of $N_t^{(h)}$ is

$$N_t^{(h)}(E) = \frac{1}{a_h} Z_{\lfloor ht \rfloor}^{(h)},$$

which, from (1.12), converges to a CSBP. Now, if we suppose that

$$N_0^{(h)} = \frac{1}{a_h} \sum_{i=1}^{z_h} \delta_{x_i^h} \longrightarrow \theta \in \mathscr{M}_f(E), \ h \to \infty,$$

and that the spatial motion W has certain regularity properties (satisfied if W is the Brownian motion in $\mathbb{R}^d$), then exists an $\mathscr{M}_f(E)$-valued Markov process N_t such that

$$\left(N_t^{(h)}, t \geq 0 \right) \xrightarrow{f.d.} (N_t, t \geq 0), \ h \to \infty,$$

and the the process $(N_t, t \geq 0)$ is called (p_h, ξ)-superprocess. When ξ is the Browian motion in $\mathbb{R}^d$ and we use the Feller approximation, then N_t is called super-Browian motion.

Remark 12. Note that $Z_0^{(1)} = Z_0$. According to the Feller approximation, if $Z_0^{(1)} = 1$ then $N_0^{(1)}(E) = 1$. Moreover, since the initial mass is conserved in the scaling, at the *h-th* step we have that

$$N_0^{(1)}(E) = N_0^{(h)}(E) = \frac{1}{h} z_h. \tag{1.14}$$

Thus, if $Z_0^{(1)} = Z_0 = 1$ then

$$z_h = h,$$

that is, if the process begins with one particle in $x \in \mathbb{R}^d$, then at the *h-th* step of the scaling, in $t = 0$, we have h particles in x, each of which has "mass" $1/h$.

1.3.2   Multitype superprocesses

Let $r > 1$ and, with same settings of the system particles of the previous Section, consider a sequence of multitype Galton-Watson process

$$\left(\boldsymbol{Z}_n^{(h)}, n \geq 0 \right),$$

with initial value $\boldsymbol{z}_h$ and offspring distribution $\boldsymbol{p}_h$ such that

$$
\begin{cases}
m_i^{h,(1)}(k) = \delta_{ik} + h^{-1} \cdot d_{ik}^h, \ \text{with} \ \lim_{h \to \infty} d_{ik}^h = d_{ik}, \\
\lim_{h \to \infty} m_i^{h,(2)}(k,l) = m_i^{(2)}(k,l), \\
\sup_{h \geq 1} m_i^{h,(3)}(k,l,n) < \infty, \quad i,k,l,n = 1,\dots,r,
\end{cases}
\tag{1.15}
$$

where

$$
\begin{cases}
m_i^{h,(1)}(k) = \sum_{\boldsymbol{j} \in \mathbb{N}^r} p_h^{(i)}(\boldsymbol{j}) \, j_k, \\
m_i^{h,(2)}(k,l) = \sum_{\boldsymbol{j} \in \mathbb{N}^r} p_h^{(i)}(\boldsymbol{j}) \, j_k \left(j_l - \delta_{lk} \right), \\
m_i^{h,(3)}(k,l,n) = \sum_{\boldsymbol{j} \in \mathbb{N}^r} p_h^{(i)}(\boldsymbol{j}) \, j_k \left(j_l - \delta_{lk} \right) \left(j_n - \delta_{nk} - \delta_{nl} \right)
\end{cases}
$$

Consider the process

$$\left(\frac{1}{h} \boldsymbol{Z}_{\lfloor ht \rfloor}^{(h)}, \ t \geq 0 \right),$$

it is the *scaled process* at the step h, in which every particle has mass equal to $1/h$. Now we define the following random measures,

$$\boldsymbol{N}_t^{(h)} = \left(N_{t,1}^{(h)}, \dots, N_{t,r}^{(h)} \right),$$

with

$$N_{t,i}^{(h)} = \frac{1}{h} \sum_{l=1}^{Z_{\lfloor ht \rfloor, i}^h} \delta_{\xi_t^{h,l}}, \ \forall i = 1,\dots,r \ , \tag{1.16}$$

where $\xi_t^{h,l}$ is the position in $\mathbb{R}^d$ of the l-th particle of type i alive at time t, according to the law of the spatial motion of type i particles, W_i. Moreover, suppose that

$$\boldsymbol{N}_0^{(h)} \longrightarrow \boldsymbol{\theta} \in \mathscr{M}_f\left(\mathbb{R}^d\right)^r,$$

then, under certain regularity proporties of the spatial motions W_i, we know, from the main result proved in L.G. Gorostiza - J.A. Lopez-Mimbela [6] (Theorem 4.1, (a), Sec. 4), that

$$\boldsymbol{N}_t^{(h)} \longrightarrow \boldsymbol{N}_t,$$

where $\boldsymbol{N}_t$ is called *multitype superprocess* and it is also the unique continuous solution of a martingale problem (L.G. Gorostiza - J.A. Lopez-Mimbela, [6], Theorem 4.1, (b), Sec. 4) and it can be characterized by its Laplace functional (L.G. Gorostiza - J.A. Lopez-Mimbela, [6], Theorem 4.1, (c), Sec. 4)

Remark 13. Note that the (1.14) holds also for the multitype Galton-Watson processes, i.e. if $\boldsymbol{Z}_0^{(1)} = \boldsymbol{e}_i$ then $\boldsymbol{z}_h = h \cdot \boldsymbol{e}_i$, $\forall i = 1, \ldots, r$, and all the h initial particles have "mass" equal to $1/h$.

Chapter 2

Theoretical results on μ-GW trees and scaling limit weak-convergence

In this chapter we propose our statistical study of the μ-GW trees and we introduce our results concerning the period of the contour process of such trees, the scaling limit of suitable scaled periods and the full binary trees with survivals and the full binary trees. In Section 2.1 we recall the important notion of the *contour process* of the one-type Galton-Watson trees and, because we are interested in a scaling limit in the multitype case, we give the main example on the scaling limits of the contuor process, i.e. the Aldous' Theorem. We refer to J.-F. Le Gall ([2]) both for the defintion of the contour process related to the one-type Galton-Watson trees and for the Aldous' Theorem. Then, we apply the idea of the contour process also to the μ-GW trees and we deduce some important results about the period of the contour process. In particular, in accordance with the literature of the Galton-Watson processes, we give an iterative definition of the period of the contour process and there are shown results on its moment-generating function and relations between it and the extinction probability of a μ-GW tree. Then, from Feller's idea (see Section 1.3), in Section 2.2 we present a scaling for the μ-GW

trees. In particular, at the step n of the scaling we consider n $\boldsymbol{\mu}_n$-GW trees *accelerated* ($\boldsymbol{\mu}_n$ is the offspring distribution at the step n), and we prove the weak convergence of the total period of the contour processes of the n trees for $n \to \infty$. To prove that, we essentially apply the well known Lyapounovs Theorem to a particular triangular arrays family, that is a sequence of suitable scaled periods of contour process of the $\boldsymbol{\mu}_n$-GW trees at the step n of the scaling. For the Lyapounovs Theorem on triangular arrays we refer to [24, Section 27], and for the scaling adopted used in the proof we refer to [5, Chapter 3, Sec. 3.2]. Then, in Section 2.3 we consider two particular kind of multitype Galton-Watson trees, the *full binary trees* and the *full binary trees with survivals*, and we see interesting relations with the Narayana numbers (see [12]) and the two-dimensional decompositions of the integers (see [11] and [10]). In particular, we show relevant statistical results, as the maximum-likelihood estimator of parameters of a particular distribution related to the full binary trees and the estimation of the probability to have full binary trees still alive at the generation $M \geq 1$.

2.1   On the period of the contour process of the μ-GW trees

Let $(\tau, e_\tau) \in T_i^{(r)}$ and consider the following order for the offspring of a vertex, according to the type.

Definition 2.1.1 (Type-ordering offspring). Let $(\tau, e_\tau) \in T_i^{(r)}$. Then

$$
\begin{cases}
e\left(\boldsymbol{u}l\right) = 1, \ 1 \le l \le p_1\left(\boldsymbol{w}_\tau\left(\boldsymbol{u}\right)\right) \\
\ldots \\
e\left(\boldsymbol{u}l\right) = k, \ \sum_{i=1}^{k-1} p_i\left(\boldsymbol{w}_\tau\left(\boldsymbol{u}\right)\right) + 1 \le l \le \sum_{i=1}^{k} p_i\left(\boldsymbol{w}_\tau\left(\boldsymbol{u}\right)\right) \\
\ldots \\
e\left(\boldsymbol{u}l\right) = r, \ \sum_{i=1}^{r-1} p_i\left(\boldsymbol{w}_\tau\left(\boldsymbol{u}\right)\right) + 1 \le l \le c_\tau\left(\boldsymbol{u}\right),
\end{cases}
\tag{2.1}
$$

where $\boldsymbol{p} = (p_1, \ldots, p_r)$ is the counter map introduced in Chapter 1 and $\boldsymbol{u} \in \tau$. Note that if $p_j\left(\boldsymbol{w}_\tau\left(\boldsymbol{u}\right)\right) = 0$ for some $j \in [r]$, then $\{\boldsymbol{u}l \in \tau | e\left(\boldsymbol{u}l\right) = j\} = \{\varnothing\}$.

Remark 14. In other words, we order the offspring of a vertex $\boldsymbol{u} \in \tau$, with $\tau \in (\tau, e_\tau)$, from the older one to the younger one (*left-right* order) according to $\boldsymbol{p}\left(\boldsymbol{w}_\tau\left(\boldsymbol{u}\right)\right)$

From now on, a $\boldsymbol{\mu}$-GW tree is equipped with the Type-ordering offspring. Now, remembering that, in the graphical point of view, a multitype tree is like a one-type planar tree (Figure 1.1), we can derive the *contour process*

$$
\mathscr{C}_\tau\left(s\right) : [0, 2\|\tau\|] \longrightarrow \mathbb{N}
$$

of $(\tau, e_\tau) \in T_i^{(r)}$ (Figure 2.1). We recall, from J.-F. Le Gall ([2, Sec. 1.1]), that the contour process of a one-type tree is easy to visualize: imagine a particle on the root of the planar tree at time 0 and then it moves on the tree according to the following rules. The particle jumps from a vertex to its first not yet visited child, if any, and if none to the father of the vertex. Eventually, the particle comes back to the root of the tree after having visited all the vertices of the tree. The value $\mathscr{C}_\tau\left(n\right)$ at the time n is the generation of the vertex visited at the step n in this evolution. Because we are interested in the scaling limit of the period of the contuor process of a $\boldsymbol{\mu}$-GW tree,

(a)

(b)
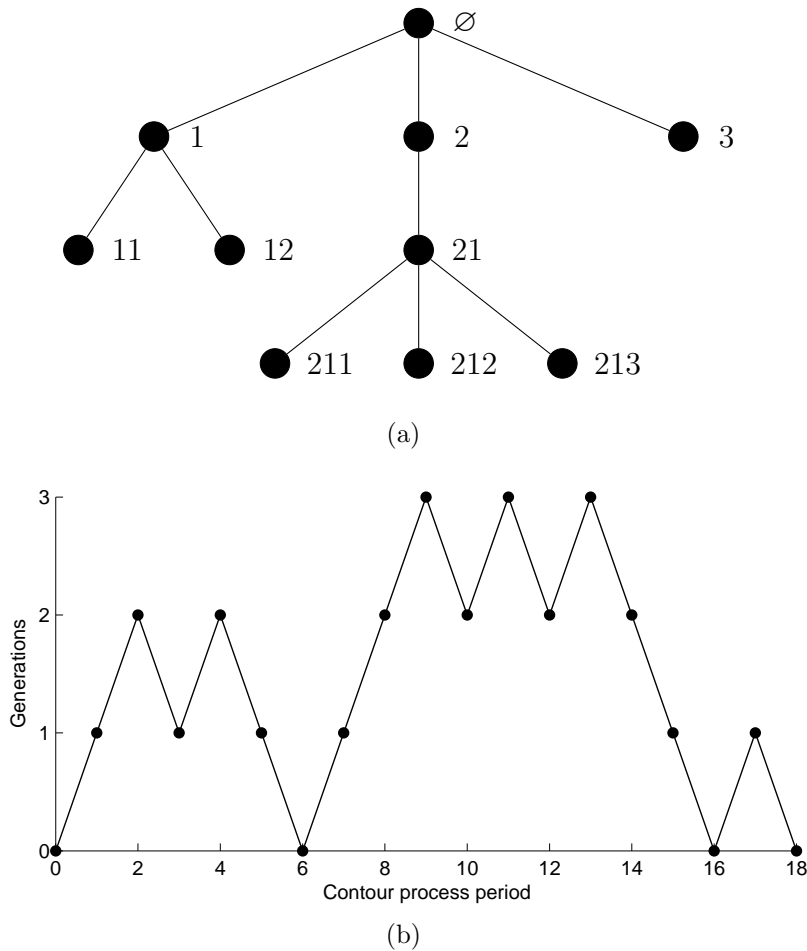
Figure 2.1: A planar tree (a) and its contour process (b)

consider, as an important example, the Aldous' Theorem about the scaling limit of the contour function related to a one-type Galton-Watson tree ([2])

Theorem 2.1.1 (Aldous' Theorem). *Let τ_p a one-type Galton-Watson tree, with offspring distribution μ, conditioned to have p edges. Then*

$$\left(\frac{1}{\sqrt{2p}} \mathscr{C}_{\tau_p} (2ps) \right)_{0 \leq s \leq 1} \xrightarrow{d} \left(\frac{\sqrt{2}}{\sigma} \beta_s \right)_{0 \leq s \leq 1} , \ p \to \infty,$$

where $\sigma^2 = Var[\mu]$ and $(\beta_s, 0 \leq s \leq 1)$ is a normalized Brownian excursion.

Remark 15. Note that $\|\tau\|$ is a non negative-integer-valued r.v. .

Now, in the next Theorem we prove an iterative definition of the period of the contour process $\mathscr{C}_\tau(s)$ of a $\boldsymbol{\mu}$-GW tree.

Definition 2.1.2 (Number of edges from a vertex of a $\boldsymbol{\mu}$-GW tree). Let $(\tau, e_\tau) \in T^{(r)}$ be a $\boldsymbol{\mu}$-GW tree. We define

$$L^{(j)}(\boldsymbol{u})$$

the r.v. that represents the number of edges from the vertex $\boldsymbol{u}$, $e_\tau(\boldsymbol{u}) = j$. Note that

$$L^{(i)}(\varnothing) = \|\tau\|,$$

where τ is a $\boldsymbol{\mu}$-GW tree rooted in a type i vertex.

Theorem 2.1.2. *Let $(\tau, e_\tau) \in T^{(r)}$ be a $\boldsymbol{\mu}$-GW tree, then, for every $\boldsymbol{u} \in \tau$,*

1. *$\forall\, l_1, \ldots, l_r,\ L^{(1)}(\boldsymbol{u}l_1), \ldots, L^{(r)}(\boldsymbol{u}l_r)$ are independent,*

2. *$\forall\, l = 1, \ldots, p_k(\boldsymbol{C_{u,j}}),\ L^{(k)}(\boldsymbol{u}l)$ are i.i.d. and $L^{(k)}(\boldsymbol{u}l) \stackrel{d}{=} L^{(k)}(\varnothing)$,*

3. *$\forall\, l, k,\ L^{(k)}(\boldsymbol{u}l)$ and $p_k(\boldsymbol{C_{u,j}})$ are independent,*

4.

$$L^{(j)}(\boldsymbol{u}) = \sum_{k=1}^{r} \left(p_k(\boldsymbol{C_{u,j}}) + \sum_{l=1}^{p_k(\boldsymbol{C_{u,j}})} L^{(k)}(\boldsymbol{u}l) \right). \tag{2.2}$$

Proof. Note that, in accordance with the literature of the Galton-Watson processes, 1, 2 and 3 are verified. In other words, they represent the property that different particles have independent offspring in the Galton-Watson prodesses. 4 can be proved for recursion on the vertices and on the vertices type. $\qquad\square$

In the next Theorem we show a characterization of the moment-generating function of the period of $\mathscr{C}_\tau(s)$, i.e. of $2L^{(i)}(\varnothing)$.

Theorem 2.1.3. $\forall i \in [r]$, let $F_i(s) = \mathbb{E}\left[e^{s \cdot 2L^{(i)}(\varnothing)}\right]$ be the moment-generating function of $2L^{(i)}(\varnothing)$, $s \in \mathbb{R}$. Then

$$F_i(s) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \left(\mu^{(i)}(\boldsymbol{\alpha}) \cdot e^{2s|\boldsymbol{\alpha}|} \cdot \prod_{k=1}^{r} F_k(s)^{\alpha_k} \right), \ \ with \ |\boldsymbol{\alpha}| = \sum_{i=1}^{r} \alpha_i. \quad (2.3)$$

Proof. Let $f_k(s) = \mathbb{E}\left[e^{s \cdot L^{(i)}}(\varnothing)\right]$. For the law of total probability we have,

$$f_i(s) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mathbb{E}\left[\exp\left(s \cdot L^{(i)}(\varnothing)\right) | \boldsymbol{p}\left(\boldsymbol{C}_{\varnothing,i}\right) = \boldsymbol{\alpha}\right] \cdot \mu^{(i)}(\boldsymbol{\alpha}) \stackrel{(2.2)}{=}$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mathbb{E}\left[\exp\left(s \cdot \sum_{k=1}^{r}\left(\alpha_k + \sum_{l=1}^{\alpha_k} L^{(k)}(l)\right)\right)\right] \cdot \mu^{(i)}(\boldsymbol{\alpha}) =$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \mathbb{E}\left[\prod_{k=1}^{r} \exp\left(s \cdot \left(\alpha_k + \sum_{l=1}^{\alpha_k} L^{(k)}(l)\right)\right)\right].$$

Using 1 of Theorem 2.1.2 we get

$$f_i(s) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} \mathbb{E}\left[\exp\left(s \cdot \left(\alpha_k + \sum_{l=1}^{\alpha_k} L^{(k)}(l)\right)\right)\right] =$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} \left(\exp\left(s \cdot \alpha_k\right) \mathbb{E}\left[\exp\left(s \cdot \sum_{l=1}^{\alpha_k} L^{(k)}(l)\right)\right]\right) =$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} \left(\exp\left(s \cdot \alpha_k\right) \mathbb{E}\left[\prod_{l=1}^{\alpha_k} \exp\left(s \cdot L^{(k)}(l)\right)\right]\right).$$

Now, from 2 of Theorem 2.1.2 we obtain

$$f_i(s) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} \left(\exp\left(s \cdot \alpha_k\right) \prod_{l=1}^{\alpha_k} \mathbb{E}\left[\exp\left(s \cdot L^{(k)}(l)\right)\right]\right) =$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} \left(\exp\left(s \cdot \alpha_k\right) \prod_{l=1}^{\alpha_k} f_k(s)\right) =$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} \left(\exp\left(s \cdot \alpha_k\right) f_k(s)^{\alpha_k} \right) =$$

$$= \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \left(\mu^{(i)}(\boldsymbol{\alpha}) \cdot e^{s|\boldsymbol{\alpha}|} \cdot \prod_{k=1}^{r} f_k(s)^{\alpha_k} \right)$$

And so

$$f_i(2s) = F_i(s) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \left(\mu^{(i)}(\boldsymbol{\alpha}) \cdot e^{2s|\boldsymbol{\alpha}|} \cdot \prod_{k=1}^{r} F_k(s)^{\alpha_k} \right).$$

$\square$

From Theorem 2.1.3 we can give a property of the moment-generating function $F_i(s), \forall i \in [r]$.

Lemma 2.1.4. *Define $F_i(0^-) := \lim\limits_{s \to 0^-} F_i(s), \forall i \in [r]$, then*

$$F_i\left(0^-\right) = \mathbb{P}\left(2L^{(i)}(\varnothing) < \infty\right).$$

Proof. Let $i \in [r]$ and $\{ \left(p_{2n}^{(i)} \right)_{n \geq 0} \cup p_\infty^{(i)} \}$ be the distribution of probability of $2L^{(i)}(\varnothing)$, that is $p_{2n}^{(i)} = \mathbb{P}\left(2L^{(i)}(\varnothing) = 2n\right), n \geq 0$, and $p_\infty^{(i)} = \mathbb{P}\left(2L^{(i)}(\varnothing) = \infty\right)$, with $p_\infty^{(i)} + \sum_{n \geq 0} p_{2n}^{(i)} = 1$. Then

$$F_i(s) = \mathbb{E}\left[e^{s \cdot 2L^{(i)}(\varnothing)} \right] = p_\infty^{(i)} \cdot e^{s \cdot \infty} + \sum_{n \geq 0} p_{2n}^{(i)} \cdot e^{2sn}, \ s \in \mathbb{R}.$$

So, $\forall s < 0$ we have

$$F_i(s) = \sum_{n \geq 0} p_{2n}^{(i)} \cdot e^{2sn}, \tag{2.4}$$

and $\forall n \geq 0$

$$\left| e^{2sn} \right| = e^{2sn} \leq 1.$$

Then, from the Dominated Convergence Theorem, we obtain

$$\lim_{s \to 0^-} F_i(s) = \sum_{n \geq 0} p_{2n}^{(i)} = \mathbb{P}\left(2L^{(i)}(\varnothing) < \infty\right).$$

$\square$

As corollary of the Lemma 2.1.4, we get a connection between $F_i(s), \forall i \in [r]$, and the extinction probability $\boldsymbol{q}$ (see (1.4), in Section 1.1.1).

Corollary 1. *Denote* $\boldsymbol{F}(0^-) = \left(\mathbb{P}\left(2L^{(1)}(\varnothing) < \infty\right), \ldots, \mathbb{P}\left(2L^{(r)}(\varnothing) < \infty\right)\right)$, *then*

$$\boldsymbol{F}(0^-) = \boldsymbol{f}\left(\boldsymbol{F}(0^-)\right),$$

where $\boldsymbol{f}(\boldsymbol{s})$ is the generating function (1.2) *related to the offspring distribution $\boldsymbol{\mu}$. Moreover,* $\boldsymbol{F}(0^-) = \boldsymbol{q}$.

Proof. Note that $\forall i \in [r], \forall s < 0$ and from (2.4) we have that

$$|F_i(s)| = F_i(s) = \sum_{n \geq 0} p_{2n}^{(i)} \cdot e^{2sn} \leq \sum_{n \geq 0} p_{2n}^{(i)} \leq 1,$$

and so

$$\left| \mu^{(i)}(\boldsymbol{\alpha}) \cdot e^{2s|\boldsymbol{\alpha}|} \cdot \prod_{k=1}^{r} F_k(s)^{\alpha_k} \right| = \mu^{(i)}(\boldsymbol{\alpha}) \cdot e^{2s|\boldsymbol{\alpha}|} \cdot \prod_{k=1}^{r} F_k(s)^{\alpha_k} \leq 1.$$

Thus, from the Theorem (2.1.3) and the Theorem of Dominated Convergnce, we have that

$$F_i(0^-) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \left(\mu^{(i)}(\boldsymbol{\alpha}) \cdot \prod_{k=1}^{r} F_k(0^-)^{\alpha_k} \right), \forall i \in [r],$$

and so

$$\boldsymbol{F}(0^-) = \boldsymbol{f}\left(\boldsymbol{F}(0^-)\right).$$

To proof the second part of the Corollary it's enough to note that $\forall i \in [r]$,

$$2L^{(i)}(\varnothing) < \infty \Longleftrightarrow \boldsymbol{Z}_n(\tau) = \boldsymbol{0}, \text{ for some } n \geq 0,$$

where $(\tau, e_\tau) \in T_i^{(r)}$ is a $\boldsymbol{\mu}$-GW tree, and $\boldsymbol{Z}_n(\tau)$ is the Galton-Watson process related to the tree (1.10). Then, by the Lemma 2.1.4 and the defintion of the extinction probability $\boldsymbol{q}$ (see (1.4), in Section 1.1.1) we get

$$F_i\left(0^-\right) = \mathbb{P}\left(2L^{(i)}(\varnothing) < \infty\right) = \mathbb{P}\left(\boldsymbol{Z}_n(\tau) = \boldsymbol{0}, \text{ for some } n \geq 0\right) = q^{(i)},$$

and so

$$\boldsymbol{F}\left(0^-\right) = \boldsymbol{q}.$$

$\square$

We consider two interesting examples concerning the results seen above.

Example 2.1.1 (One type geometric distribution). Let be $r = 1$, $p \in [0, 1)$ and use the notation $\mu^{(1)}(n) = \mu_n$. Consider the geometric offspring distribution

$$\begin{cases} \mu_n = b \cdot p^{n-1}, \ n \geq 1, \\ \mu_0 = \dfrac{1 - p - b}{1 - p}, \end{cases} \tag{2.5}$$

with $0 \leq b \leq 1 - p$. Two cases are examined, $p = 0$ and $p \in (0, 1)$.

a) Let $p = 0$ and so $0 \leq b \leq 1$. The distribution (2.5) becomes

$$\mu_0 = 1 - b, \mu_1 = b$$

and, from the Theorem 2.1.3, we have that

$$F(s) = \sum_{n \geq 0} \mu_n \left(e^{2s} F(s)\right)^n = (1 - b) + e^{2s} F(s) b, \tag{2.6}$$

and so

$$(1 - b) F\left(0^-\right) = 1 - b$$

Then, from the Lemma 2.1.4, we have

- if $0 \leq b < 1 \Rightarrow \mathbb{P}\left(2L\left(\varnothing\right) < \infty\right) = 1$,

- if $b = 1$, from (2.6), we have that $F\left(s\right) = e^{2s}F\left(s\right)$ and so, for every $s < 0$, $F\left(s\right) = 0$ and $\mathbb{P}\left(2L\left(\varnothing\right) < \infty\right) = 0$ (according to the offspring distribution $\mu_1 = 1$).

b) Let $p \in (0,1)$ and so $0 \leq b \leq 1 - p$.

$$F\left(s\right) = \sum_{n \geq 0} \mu_n \left(e^{2s}F\left(s\right)\right)^n =$$

$$= \frac{1 - p - b}{1 - p} + e^{2s}F\left(s\right)b\sum_{n \geq 1}\left(e^{2s}F\left(s\right)p\right)^{n-1}. \qquad (2.7)$$

For every $s < 0$, $\left|e^{2s}F\left(s\right)p\right| = e^{2s}F\left(s\right)p < 1$ and so (2.7) becomes

$$F\left(s\right) = \frac{1 - p - b}{1 - p} + \frac{e^{2s}F\left(s\right)b}{1 - e^{2s}F\left(s\right)p}.$$

Thus we obtain the following equation

$$p\left(1 - p\right)F\left(0^-\right)^2 + F\left(0^-\right)\left(p^2 + b - 1\right) + \left(1 - p - b\right) = 0$$

that has two solutions, 1 and $\dfrac{1 - p - b}{p\left(1 - p\right)}$, and so

- if $0 \leq b \leq \left(1 - p\right)^2$, then

$$\frac{1 - p - b}{p\left(1 - p\right)} \geq 1,$$

and from the Lemma 2.1.4,

$$\mathbb{P}\left(2L\left(\varnothing\right) < \infty\right) = 1,$$

- if $(1-p)^2 < b \le (1-p)$, then

$$0 \le \frac{1-p-b}{p(1-p)} < 1,$$

and from the Corollary 1 we can use the Theorem 1.1.1 ((ii)) to conclude that

$$\mathbb{P}\left(2L\left(\varnothing\right) < \infty\right) = \frac{1-p-b}{p(1-p)}.$$

Example 2.1.2 (A particular two type distribution). Let $r = 2, p, q \in (0,1)$. Consider the following offspring distribution $\boldsymbol{\mu} = \left(\mu^{(1)}, \mu^{(2)}\right)$,

$$\begin{cases} \mu^{(1)}(0,0) = p \\ \mu^{(1)}(1,1) = 1-p \end{cases}, \quad \begin{cases} \mu^{(2)}(0,0) = q \\ \mu^{(2)}(0,2) = 1-q \end{cases}.$$

Thus, from the Theorem 2.1.3 we have

$$\begin{cases} F_1(s) = (1-p)\,e^{4s}F_1(s)F_2(s) + p \\ F_2(s) = q + (1-q)\,e^{4s}F_2(s)^2 \end{cases},$$

and so we obtain the system

$$\begin{cases} F_1(0^-) = (1-p)\,F_1(0^-)F_2(0^-) + p \\ F_2(0^-) = q + (1-q)\,F_2(0^-)^2 \end{cases},$$

which has two solutions

$$\boldsymbol{a} = (1,1) \ , \quad \boldsymbol{b} = \left(\frac{p(1-q)}{1-q(2-p)}, \frac{q}{1-q}\right).$$

Now, from Lemma 2.1.4 we know that each solution must have non-negative

components, and so we need the following condition

$$1 - q\left(2 - p\right) > 0 \Leftrightarrow p > \frac{2q - 1}{q}, \tag{2.8}$$

which is verified for every $p \in (0, 1)$ if $0 < q < 1/2$.

With the condition (2.8) we have

$$\frac{p\left(1 - q\right)}{1 - q\left(2 - p\right)} \gtreqless 1 \Longleftrightarrow q \gtreqless \left(1 - q\right), \ \forall p \in (0, 1).$$

Thus, if $0 < q < 1/2$ then

$$\frac{2q - 1}{q} < 0 \text{ and } q < 1 - q,$$

and we obtain that the solution $\boldsymbol{b}$ is strictly less than $\boldsymbol{1}$, and if $1/2 \leq q < 1$ then

$$q \geq 1 - q,$$

and with the condition (2.8) we obtain that $\boldsymbol{b} \geq \boldsymbol{1}$.

Thus, we can conclude that

- $0 < q < 1/2 \Rightarrow \boldsymbol{b} < (1, 1) \Rightarrow \boldsymbol{F}\left(0^-\right) = \boldsymbol{b} = \left(\dfrac{p\left(1 - q\right)}{1 - q\left(2 - p\right)}, \dfrac{q}{1 - q}\right)$

- $\begin{cases} 1/2 \leq q < 1 \\ p > \dfrac{2q - 1}{q} \end{cases} \Rightarrow \boldsymbol{b} \geq (1, 1) \Rightarrow \boldsymbol{F}\left(0^-\right) = (1, 1)$

where $\boldsymbol{F}\left(0^-\right) = \left(\mathbb{P}\left(2L^{(1)}\left(\varnothing\right) < \infty\right), \mathbb{P}\left(2L^{(2)}\left(\varnothing\right) < \infty\right)\right)$.

2.2 Weak convergence of the period of a scaled contour process

From Theorem 2.1.3 we deduce that the moment-generating function $F_i\left(s\right)$ depends only on the offspring distribution $\boldsymbol{\mu}$, for every $i = 1, \ldots, r$.

Thus, we consider a scaling based on the one seen in Section 1.3.2, in particular, at the n-th step, the particles lifetimes in equal to $1/n$, each particle has "mass" $1/n$ and

$$\boldsymbol{\mu}_n = \left(\mu_n^{(1)}, \ldots, \mu_n^{(r)}\right)$$

is the offspring distribution.

Let $i \in [r]$ and $(\tau, e_\tau) \in T_i^{(r)}$ be a $\boldsymbol{\mu}$-GW tree. Thus, according to the Remark 13 in Section 1.3.2, at the step n of the scaling there will be n $\boldsymbol{\mu}_n$-GW trees, each of which starts with a type i particle. So, from the effects of the Feller scaling on to the contour process treated by A.M. Etheridge [5, Chapter 3, Sec. 3.2], we have to consider the concatenation of n contour processes and the total period of the concatenation is the sum of n i.i.d. periods. Thus, at the step n, we denote with $2L_n^{(i)}(\varnothing)$ the period of the contour process relative to a $\boldsymbol{\mu}_n$-GW tree starting with a type i particle and so the total period of the concatenation is equal to

$$\sum_{j=1}^{n} 2L_{n,j}^{(i)}(\varnothing), \tag{2.9}$$

where

$$2L_{n,j}^{(i)}(\varnothing), \ j = 1, \ldots, n,$$

are n i.i.d. copies of $2L_n^{(i)}(\varnothing)$.

Remark 16. Note that here, as in A.M. Etheridge [5, Sec. 3.2], the time of the contour process played no rôle because we are only interested in the period finiteness.

Now, we explain a result of weak convergence of (2.9) suitably scaled.

Theorem 2.2.1 (Weak convergence of a scaled contour process)**.** *Let $i \in [r]$ and $\boldsymbol{\mu}_n$ be an offspring distribution at the n-th step of the scaling seen above and let $\alpha \geq 1$. Use the notation $Y_n^{(i)} = 2L_n^{(i)}(\varnothing)$ and let*

$$\left(Y_{n,j}^{(i)}\right)_{j \geq 0}$$

be i.i.d. copies of $Y_n^{(i)}$. If

1. $\begin{cases} \mathbb{E}\left[Y_n^{(i)}\right] = \nu_n^{(i)} = A_i \cdot n^\alpha + o\left(n^\alpha\right), \\ Var\left[Y_n^{(i)}\right] = \left(\sigma_n^{(i)}\right)^2 = B_i \cdot n^\gamma + o\left(n^\gamma\right), \end{cases}$

   with $A_i \in \mathbb{R}$, $B_i > 0$, $\gamma \geq 0$,

2. *$\exists\, \delta > 0$ such that*

$$\mathbb{E}\left[\left|Y_n^{(i)} - \nu_n^{(i)}\right|^{2+\delta}\right] = o\left(n^{\frac{\delta + \gamma(2+\delta)}{2}}\right),$$

then

$$\frac{n^\alpha \cdot Z_n^{(i)}}{\sqrt{Var\left[Y_n^{(i)}\right]}} \xrightarrow{d} \mathcal{N}(0,1), \ n \to \infty,$$

where

$$Z_n^{(i)} = n^{\frac{1}{2}}\left[\sum_{j=1}^n \left(\frac{Y_{n,j}^{(i)}}{n^{\alpha+1}}\right) - A_i + \frac{o(n^{\alpha+1})}{n^{\alpha+1}}\right],$$

is called the α-scaling for the (2.9).

Moreover, if $\gamma = 2\alpha$, applying the Slutsky's theorem (see [25, Theorem 11.4]), we obtain

$$Z_n^{(i)} \xrightarrow{d} \mathcal{N}(0, B_i), n \to +\infty.$$

Proof. The idea is to apply the Lyapounov's Theorem ([24, Section 27]) to the following triangular array $T_{n,1}^{(i)}, \ldots, T_{n,n}^{(i)}$, for every $i = 1, \ldots, r$ and $n \geq 1$,

$$T_{n,j}^{(i)} := \frac{Y_{n,j}^{(i)} - \nu_n^{(i)}}{n^{\alpha+1}}, \ \forall j = 1, \ldots, n.$$

Now we normalize $T_{n,j}^{(i)}$ in such a way

1. $\mathbb{E}\left[T_{n,j}^{(i)}\right] = 0$, for all n, j,

2. $\sum_{j=1}^n \mathbb{E}\left[\left(T_{n,j}^{(i)}\right)^2\right] = 1$, for all n.

By the definition, $T_{n,j}^{(i)}$ satisfies the condition (1). Thus, the condition (2) becomes

$$\sum_{j=1}^{n} \mathbb{E}\left[\left(T_{n,j}^{(i)}\right)^2\right] = \sum_{j=1}^{n} Var\left[T_{n,j}^{(i)}\right] = 1.$$

From the defintion of the array $T_{n,j}^{(i)}$ we have that

$$\sum_{j=1}^{n} Var\left[T_{n,j}^{(i)}\right] = \sum_{j=1}^{n} \frac{1}{n^{2\alpha+2}} Var\left[Y_{n,j}^{(i)}\right] =$$

$$= \frac{1}{n^{2\alpha+1}} Var\left[Y_{n,j}^{(i)}\right] = \frac{1}{n^{2\alpha+1}} Var\left[Y_n^{(i)}\right],$$

and we obtain

$$\sum_{j=1}^{n} \mathbb{E}\left[\frac{n^{2\alpha+1}}{Var\left[Y_n^{(i)}\right]} \left(T_{n,j}^{(i)}\right)^2\right] = 1.$$

Thus, if we define

$$\tilde{T}_{n,j}^{(i)} := \frac{n^{\alpha+\frac{1}{2}} \cdot T_{n,j}^{(i)}}{\sqrt{Var\left[Y_n^{(i)}\right]}}$$

then the triangular array $\tilde{T}_{n,j}^{(i)}$ satisfies the conditions 1 and 2.

Now, for the triangular array $\tilde{T}_{n,j}^{(i)}$ the Lyapounov's condition in [24, Section 27] becomes

$$\exists\, \delta > 0 \text{ such that } \lim_{n \to +\infty} \left(\sum_{j=1}^{n} \mathbb{E}\left[\left|\tilde{T}_{n,j}^{(i)}\right|^{2+\delta}\right]\right) = 0. \qquad (2.10)$$

We have that

$$\left|\tilde{T}_{n,j}^{(i)}\right|^{2+\delta} = \frac{\left|Y_{n,j}^{(i)} - \nu_n^{(i)}\right|^{2+\delta}}{\left(B_i \cdot n^{\gamma+1} + o(n^{\gamma+1})\right)^{\frac{2+\delta}{2}}},$$

and the condition (2.10) becomes

$$\exists\, \delta > 0 \text{ such that } \lim_{n \to +\infty} \frac{n \cdot \mathbb{E}\left[\left|Y_n^{(i)} - \nu_n^{(i)}\right|^{2+\delta}\right]}{\left(B_i \cdot n^{\gamma+1} + o(n^{\gamma+1})\right)^{\frac{2+\delta}{2}}} = 0. \qquad (2.11)$$

By the hypothesis *2* we know that $\exists\, \delta > 0$ such that

$$\frac{n \cdot \mathbb{E}\left[\left|Y_n^{(i)} - \nu_n^{(i)}\right|^{2+\delta}\right]}{\left(B_i \cdot n^{\gamma+1} + o(n^{\gamma+1})\right)^{\frac{2+\delta}{2}}} = \frac{o\left(n^{\frac{\delta+\gamma(2+\delta)}{2}}\right)}{n^{\frac{\delta+\gamma(2+\delta)}{2}}\left(B^i + \frac{o(n^{\gamma+1})}{n^{\gamma+1}}\right)^{\frac{2+\delta}{2}}},$$

and so the condition (2.11) is satisfied and, by the Lyapounov's theorem ([24, Section 27]), we obtain that

$$\sum_{j=1}^{n} \tilde{T}_{n,j}^{(i)} \xrightarrow{d} \mathcal{N}(0,1), n \to +\infty. \tag{2.12}$$

Note that

$$\sum_{j=1}^{n} \tilde{T}_{n,j}^{(i)} = \frac{n^{\alpha}}{\sqrt{Var\left[Y_n^{(i)}\right]}}\left(n^{\frac{1}{2}}\sum_{j=1}^{n} T_{n,j}^{(i)}\right) =$$

$$= \frac{n^{\alpha}}{\sqrt{Var\left[Y_n^{(i)}\right]}}\left(n^{\frac{1}{2}}\sum_{j=1}^{n}\left(\frac{Y_{n,j}^{(i)} - \nu_n^{(i)}}{n^{\alpha+1}}\right)\right) = \frac{n^{\alpha} \cdot Z_n^{(i)}}{\sqrt{Var\left[Y_n^{(i)}\right]}},$$

and so the (2.12) can be written in the following way

$$\frac{n^{\alpha} \cdot Z_n^{(i)}}{\sqrt{Var\left[Y_n^{(i)}\right]}} \xrightarrow{d} \mathcal{N}(0,1), n \to +\infty.$$

Moreover, if $\gamma = 2\alpha$, then $\displaystyle\lim_{n\to\infty} \frac{\sqrt{Var\left[Y_n^{(i)}\right]}}{n^{\alpha}} = \sqrt{B_i}$, and so we can apply the Slutsky's theorem ([25], Theorem 11.4) and we obtain

$$Z_n^{(i)} = \frac{n^{\alpha} \cdot Z_n^{(i)}}{\sqrt{Var\left[Y_n^{(i)}\right]}} \cdot \frac{\sqrt{Var\left[Y_n^{(i)}\right]}}{n^{\alpha}} \xrightarrow{d} \mathcal{N}(0, B_i), n \to +\infty.$$

$\square$

2.3   Full binary trees and full binary trees with survivals

In this section we conduce a combinatorial investigation on two kind of $\boldsymbol{\mu}$-GW trees, the *full binary trees* and the *full binary trees with survivals*. Important relations arise between particular distributions of these trees and the Narayana numbers (see [12, Abstract and Section 1.1]) and the two-dimensional decomposition of integers (see [11, Vol.2, Section IX, Chapter II, Paragraph 429] and [10, Chapter 24, Example 24.3]). Now, we report the defintions of the Narayana numbers and the full binary trees and the full binary trees with survivals.

Definition 2.3.1 (Narayana numbers)**.** The Narayana numbers $N(n,k)$, $n \geq 1$ and $k = 1, \ldots, n$ are defined in the followinf way

$$N(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}, \tag{2.13}$$

and they are expecially usefull for the counting problems. For example from [12, Section 1.1], it is known that the Narayana number $N(n,k)$ is the number of expressions containing n pairs of parentheses which are correctly matched and which contain k distinct nestings. For instance, $N(4,3) = 6$ counts all the following expressions with 4 pairs of parentheses, which each contains three times the sub-pattern ( ),

$$(()()()), \quad (()())(), \quad (())()(), \quad ()(()()), \quad ()()(()), \quad ()(())().$$

We now define the full binary trees with survivals and the full binary trees.

Definition 2.3.2 (Full binary trees and full binary trees with survivals)**.** A full binary tree with survivals is a $\boldsymbol{\mu}$-GW tree $(\tau, e_\tau) \in T_i^{(2)}, i = 1, 2$, equipped with the Type-ordering offspring (Definition 2.1.1), and such that

the offspring distribution $\boldsymbol{\mu}$ is the following

$$
\begin{cases}
\mu^{(1)}(0,0) = p_0 \\
\mu^{(1)}(1,0) = p_1 \\
\mu^{(1)}(1,1) = p_2
\end{cases}
,
\begin{cases}
\mu^{(2)}(0,0) = q_0 \\
\mu^{(2)}(0,1) = q_1 \\
\mu^{(2)}(1,1) = q_2
\end{cases}
$$

where $p_i, q_i \in (0,1)$, $i = 0, 1, 2$ and $\sum p_i = \sum q_i = 1$.

The vertices which produce only one vertex is called the *survivals* and the vertices which produced two vertices are called *fathers*.

Moreover, a full binary tree is a full binary tree with survivals such that there aren't *survivals*, i.e.

$$
\mu^{(1)}(1,0) = \mu^{(2)}(0,1) = 0.
$$

Remark 17. Note that, according to the Type-ordering offspring defined in (2.1) in Section 2.1, when a vertex of type i produces two vertices then the first one (the left one) is a type 1 vertex and the second one (the right one) is a type 2 vertex. This holds both for the full binary trees and for the full binary trees with survivals.

Remark 18. From now on, if a full binary tree with survivals (or without survivals) starts with a type j vertex, then the root is considered a left vertex or a right vertex respectively if $j = 1$ or $j = 2$. Morevoer, in the following sections we consider trees starting with a type 1 vertex (the same arguments are verified with a type 2 root).

2.3.1 Number of fathers of full binary trees with survivals

Now, we consider the full binary trees with survivals. Our purpose is to find the likelihood of the number of type 1 and type 2 fathers, conditioning to the a.s. finiteness of the tree. In the next Corollary we give the condition for the a.s. finiteness of full binary trees with survivals.

Corollary 2. *Let $(\tau, e_\tau) \in T^{(2)}$ be a full binary tree with survivals, then*

$$\text{if } p_0 q_0 - p_2 q_2 \geq 0, \text{ then } \tau \text{ is finite a.s. },$$

else

$$\text{if } p_0 q_0 - p_2 q_2 < 0, \text{ then } \tau \text{ is finite with a probability less than } 1.$$

Proof. From Theorem 2.1.3 we have that

$$
\begin{cases}
F_1(s) = \mu^{(1)}(0,0) + \mu^{(1)}(1,0) e^{2s} F_1(s) + \mu^{(1)}(1,1) e^{4s} F_1(s) F_2(s) \\
F_2(s) = \mu^{(2)}(0,0) + \mu^{(2)}(0,1) e^{2s} F_2(s) + \mu^{(2)}(1,1) e^{4s} F_1(s) F_2(s)
\end{cases},
$$

and so, passing to the limit $s \to 0^-$, we obtain

$$
\begin{cases}
F_1(0^-) = p_0 + p_1 F_1(0^-) + p_2 F_1(0^-) F_2(0^-) \\
F_2(0^-) = q_0 + q_1 F_2(0^-) + q_2 F_1(0^-) F_2(0^-)
\end{cases}.
$$

The system above has two solutions

$$\boldsymbol{F}_1(0^-) = (1,1) \text{ and } \boldsymbol{F}_2(0^-) = \left(\frac{p_0(q_0 + q_2)}{q_2(p_0 + p_2)}, \frac{q_0(p_0 + p_2)}{p_2(q_0 + q_2)} \right)$$

Note that if $p_0 q_0 > p_2 q_2$ then

$$
\begin{cases}
\dfrac{p_0(q_0 + q_2)}{q_2(p_0 + p_2)} > \dfrac{p_2 q_2 + p_0 q_2}{q_2(p_0 + p_2)} = 1 \\[2ex]
\dfrac{q_0(p_0 + p_2)}{p_2(q_0 + q_2)} > \dfrac{p_2 q_2 + q_0 p_2}{p_2(q_0 + q_2)} = 1
\end{cases},
$$

and, by the Lemma 2.1.4, we have that $\boldsymbol{F}_2(0^-)$ is not acceptable, then the only solution is $(1,1)$ and so

$$\boldsymbol{F}(0^-) = \left(\mathbb{P}\left(2L^{(1)}(\varnothing) < \infty \right), \mathbb{P}\left(2L^{(2)}(\varnothing) < \infty \right) \right) = (1,1).$$

If $p_0 q_0 = p_2 q_2$ then

$$
\begin{cases}
\dfrac{p_0 (q_0 + q_2)}{q_2 (p_0 + p_2)} = \dfrac{p_2 q_2 + p_0 q_2}{q_2 (p_0 + p_2)} = 1 \\[4mm]
\dfrac{q_0 (p_0 + p_2)}{p_2 (q_0 + q_2)} = \dfrac{p_2 q_2 + q_0 p_2}{p_2 (q_0 + q_2)} = 1
\end{cases},
$$

so $\boldsymbol{F}_1 \left(0^- \right) = \boldsymbol{F}_2 \left(0^- \right)$ and then we have that

$$
\boldsymbol{F} \left(0^- \right) = \left(\mathbb{P} \left(2 L^{(1)} (\varnothing) < \infty \right), \mathbb{P} \left(2 L^{(2)} (\varnothing) < \infty \right) \right) = (1, 1).
$$

Finally, if $p_0 q_0 < p_2 q_2$ then

$$
\begin{cases}
\dfrac{p_0 (q_0 + q_2)}{q_2 (p_0 + p_2)} < \dfrac{p_2 q_2 + p_0 q_2}{q_2 (p_0 + p_2)} = 1 \\[4mm]
\dfrac{q_0 (p_0 + p_2)}{p_2 (q_0 + q_2)} < \dfrac{p_2 q_2 + q_0 p_2}{p_2 (q_0 + q_2)} = 1
\end{cases},
$$

and, from the Corollary 1, we can use the Theorem 1.1.1 (ii) obtaining that

$$
\boldsymbol{F} \left(0^- \right) = \boldsymbol{F}_2 \left(0^- \right) = \left(\mathbb{P} \left(2 L^{(1)} (\varnothing) < \infty \right), \mathbb{P} \left(2 L^{(2)} (\varnothing) < \infty \right) \right) =
$$

$$
= \left(\dfrac{p_0 (q_0 + q_2)}{q_2 (p_0 + p_2)}, \dfrac{q_0 (p_0 + p_2)}{p_2 (q_0 + q_2)} \right) < (1, 1).
$$

$\square$

Suppose that $(\tau, e_\tau) \in T_1^{(2)}$ is an a.s. finite full binary trees with survivals starting with a type 1 vertex and use the following notations

$$
D_1 = \text{number of type 1 fathers in } \tau,
$$

$$
D_2 = \text{number of type 2 fathers in } \tau,
$$

$$
S_1 = \text{number of type 1 survivals in } \tau,
$$

$$S_2 = \text{number of type 2 survivals in } \tau.$$

At first, we want to compute the joint distribution of the number of fathers and survivals for each type,

$$\mathbb{P}\left(D_1 = n, D_2 = m, S_1 = s_1, S_2 = s_2\right), \qquad (2.14)$$

where $n \geq 1$ and $m, \ s_1, \ s_2 \geq 0$.

The case $n = 0$ is not included for the Remark 18 (there couldn't be type 2 vertices in the tree). To compute the probability (2.14) we need the probability that an a.s. finite full binary tree with survivals has exactly $D_1 = n, D_2 = m, S_1 = s_1, S_2 = s_2$ and the number of such trees. It's easy to check that in this case the number of the type 1 leaves is equal to $m + 1$ and the number of the type 2 leaves is equal to n, so the probability that a finite full binary tree with survivals has exactly $D_1 = n, \ D_2 = m, \ S_1 = s_1, \ S_2 = s_2$ is

$$p_0^{m+1} \cdot p_1^{s_1} \cdot p_2^{n} \cdot q_0^{n} \cdot q_1^{s_2} \cdot q_2^{m}. \qquad (2.15)$$

Remark 19. Note that the total number of type 1 and type 2 vertices are $m + 1 + s_1 + n$ and $n + s_2 + m$ respectively.

Now we have to count such trees. To make this, we at first consider the case in which $s_1 = s_2 = 0$ and we add the survivals later. Now, starting from the root, we walk the trees in this way, as presented by R.P. Grimaldi in [10, Chapter 24, Example 24.3]: if we are in a vertex with children we at first visit the left one, if we are in a left leaf we visit its right brother and if we are in a right leaf we visit the older not visited right vertex having the youngest last ancestor in common. Now we write L or R for each left or right vertex visited, and note that the number of L and R is the same and it's equal to the number of the father vertices. This is what we call the "LR" encoding of the tree. Now, we substitute each L with an open parenthesis (, and each R with close one), and note that each couple of consecutive parentheses () (i.e. "LR") represents a left leaf. This is the "parentheses" encoding of the tree. Thus, the total number of (or), is the number of the fathers and the

number of couple "()" is equal to the exponent of p_0 in (2.15), i.e. $m + 1$.

Remark 20. Note that each full binary tree has an unique "LR", and so "parentheses", encoding.

Consider now an example of these encodings.

Example 2.3.1. Consider the full binary tree, with a type 1 root, in Figure 2.2. It has 5 left fathers (including the root) and 4 right fathers. The "LR"
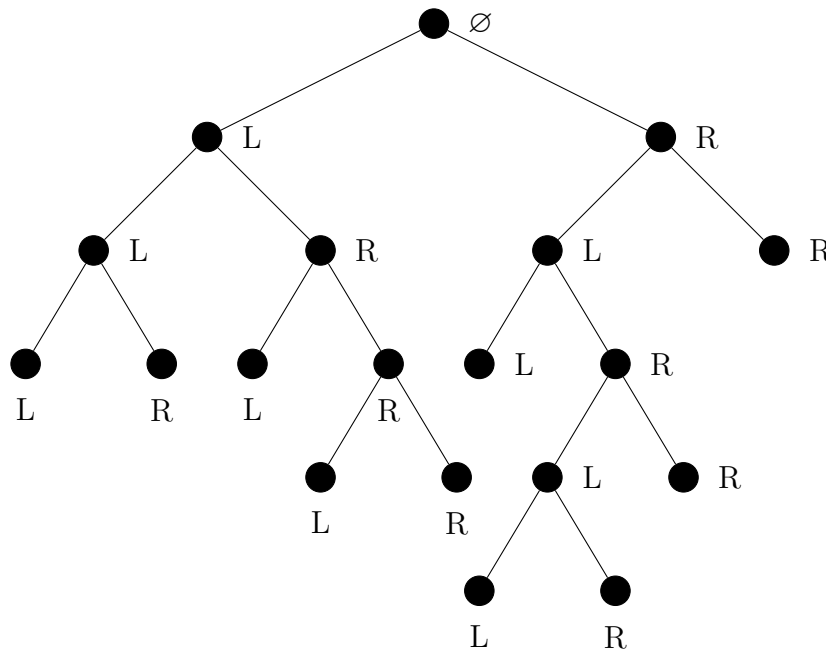
Figure 2.2: A full-binary tree with 5 left fathers (including the root) and 4 right fathers.

encoding is L, L, L, R, R, L, R, L, R, R, L, L, R, L, L, R, R, R, and its related "parentheses" encoding is ((()) () ())(() (())).

Now, using the "parentheses" encoding seen above, we know that the number of full binary trees having n left fathers and m right fathers is the number of expressions containing $n + m$ pairs of parentheses which are correctly matched and which contain $m + 1$ distinct nestings "()". Thus, from the Defintion 2.3.1, this counting problem can be solved by the Narayana

numbers seen above and the solution is the following number

$$N\left(n+m,m+1\right)=\frac{1}{n+m}\binom{n+m}{m+1}\binom{n+m}{m},\ n\geq 1, m\geq 0.$$

Now we insert the survivals, i.e. $s_1, s_2 \geq 0$. From Remark 19, we can choose the s_1 survivals type 1 vertices in $m + n + s_1$ ways (i.e. the total number of type 1 vertices of the tree, in the depth-first order, excluded the last one that is certainly a leaf), and, for the same reason, s_2 can be choosen in $m + n + s_2 - 1$ ways. Thus, the probability (2.14) becomes

$$\mathbb{P}\left(D_1 = n, D_2 = m, S_1 = s_1, S_2 = s_2\right) =$$

$$= N\left(n+m, m+1\right)\binom{m+n+s_1}{s_1}\binom{m+n+s_2-1}{s_2}p_0^{m+1}\cdot p_1^{s_1}\cdot p_2^n\cdot q_0^n\cdot q_1^{s_2}\cdot q_2^m$$

$$\tag{2.16}$$

where $n \geq 1$ and $m, s_1, s_2 \geq 0$.

Now, for the law of total probability and from the (2.16), we have that

$$\mathbb{P}\left(D_1 = n, D_2 = m\right) = \sum_{\substack{s_1 \geq 0 \\ s_2 \geq 0}} \mathbb{P}\left(D_1 = n, D_2 = m, S_1 = s_1, S_2 = s_2\right) =$$

$$= N\left(n+m, m+1\right) p_0^{m+1} p_2^n q_0^n q_2^m \cdot \sum_{\substack{s_1 \geq 0 \\ s_2 \geq 0}} \binom{m+n+s_1}{s_1}\binom{m+n+s_2-1}{s_2}p_1^{s_1} q_1^{s_2}$$

$$\tag{2.17}$$

Note that the number of the *not survivals* type 1 vertices is $n + m + 1$, and so we can say that

$$\sum_{s_1 \geq 0}\binom{m+n+s_1}{s_1}p_1^{s_1}\left(1-p_1\right)^{n+m+1}=1,$$

from which we obtain that

$$\sum_{s_1 \geq 0}\binom{m+n+s_1}{s_1}p_1^{s_1}=\frac{1}{\left(1-p_1\right)^{n+m+1}}=\frac{1}{\left(p_0+p_2\right)^{n+m+1}}.$$

Thus, the (2.17) becomes

$$\mathbb{P}\left(D_1 = n, D_2 = m\right) =$$

$$= N\left(n+m, m+1\right) \frac{p_0^{m+1} p_2^n}{\left(p_0 + p_2\right)^{n+m+1}} q_0^n q_2^m \cdot \sum_{s_2 \geq 0} \binom{m+n+s_2-1}{s_2} q_1^{s_2}.$$

Note again that the number of the *not survivals* type 2 vertices is $n+m$, and so we have that

$$\sum_{s_2 \geq 0} \binom{m+n+s_2-1}{s_2} q_1^{s_2} \left(1-q_1\right)^{n+m} = 1,$$

from which we obtain that

$$\sum_{s_2 \geq 0} \binom{m+n+s_2-1}{s_2} q_1^{s_2} = \frac{1}{\left(1-q_1\right)^{n+m}} = \frac{1}{\left(q_0 + q_2\right)^{n+m}}.$$

Finally, from the (2.17) we have that

$$\mathbb{P}\left(D_1 = n, D_2 = m\right) =$$

$$= N\left(n+m, m+1\right) \cdot \frac{p_0^{m+1} p_2^n}{\left(p_0 + p_2\right)^{n+m+1}} \cdot \frac{q_0^n q_2^m}{\left(q_0 + q_2\right)^{n+m}}. \qquad (2.18)$$

If we denote

$$P = \frac{p_0}{p_0 + p_2} \quad \text{and} \quad Q = \frac{q_0}{q_0 + q_2},$$

from the (2.18), we finally get the likelihood of the fathers of tha a.s. full binary trees with survivals.

Theorem 2.3.1. *Let* $(\tau, e_\tau) \in T_1^{(2)}$ *be an a.s. finite binary tree with survivals with n left fathers and m right fathers, with $n \geq 1$ and $m \geq 0$. Then*

$$\mathscr{L}\left(P, Q | n, m\right) = N\left(n+m, m+1\right) \cdot P^{m+1} \left(1-P\right)^n Q^n \left(1-Q\right)^m, \qquad (2.19)$$

where parameters $P, Q \in (0,1)$ *and depending only from the offspring distri-*

bution $\boldsymbol{\mu}$ *of the tree.*

From the theory of the MLE (Maximum-Likelihood Estimation) (see for instance [7]), now we are able to find the maximum-likelihood estimators of the parameters P, Q.

Consider the related log-likelihood

$$\ln \left(\mathscr{L} \left(P, Q | n, m \right) \right) = \ln \left(N \left(n + m, m + 1 \right) \right) + (m + 1) \cdot \ln \left(P \right) +$$

$$+ n \cdot \ln \left(1 - P \right) + n \cdot \ln \left(Q \right) + m \cdot \ln \left(1 - Q \right).$$

To find the maximum-likelihood estimators $\tilde{P}, \tilde{Q}$ of the parameters P, Q, we have to solve the following system

$$\begin{cases} \dfrac{\partial \ln \left(\mathscr{L} \left(\tilde{P}, \tilde{Q} | n, m \right) \right)}{\partial \tilde{P}} = 0 \\[4mm] \dfrac{\partial \ln \left(\mathscr{L} \left(\tilde{P}, \tilde{Q} | n, m \right) \right)}{\partial \tilde{Q}} = 0 \end{cases},$$

and thus obtain

$$\begin{cases} \dfrac{m + 1}{\tilde{P}} - \dfrac{n}{1 - \tilde{P}} = 0 \\[4mm] \dfrac{n}{\tilde{Q}} - \dfrac{m}{1 - \tilde{Q}} = 0 \end{cases} \Rightarrow \begin{cases} \dfrac{\left(1 - \tilde{P} \right) (m + 1) - n \cdot \tilde{P}}{\tilde{P} \left(1 - \tilde{P} \right)} = 0 \\[4mm] \dfrac{\left(1 - \tilde{Q} \right) n - m \cdot \tilde{Q}}{\tilde{Q} \left(1 - \tilde{Q} \right)} = 0 \end{cases} \Rightarrow$$

$$\begin{cases} \tilde{P} = \dfrac{m + 1}{m + n + 1} \\[4mm] \tilde{Q} = \dfrac{n}{m + n} \end{cases}, \text{ where } \tilde{P}, \tilde{Q} \in (0, 1) \text{ and } n \geq 1, \ m \geq 0. \qquad (2.20)$$

Note that the maximum-likelihood estimators of the parameters P and Q are

the number of the tipe i leaves over the sum of the type i leaves and fathers, respectively for $i = 1, 2$.

2.3.2 Relation between full binary trees and two dimensional decompositions of integers

In this section we outline, through the Narayana numbers, a particular relation between the full binary trees and the two-dimensional decomposition of the integers, conditioning to the a.s. finiteness of the trees, i.e. $p_0 q_0 \geq p_2 q_2$. For the decompositions we will refer to P.A. Macmahon [11, Vol.2, Section IX, Chapter II, Paragraph 429]. Note that in the case of the full binary trees the a.s. finite condition $p_0 q_0 \geq p_2 q_2$ is equivalent to say $p_2 + q_2 \leq 1$, indeed

$$p_0 q_0 \geq p_2 q_2 \Rightarrow (1 - p_2)(1 - q_2) \geq p_2 q_2 \Rightarrow 1 - q_2 - p_2 + p_2 q_2 \geq p_2 q_2 \Rightarrow$$

$$\Rightarrow p_2 + q_2 \leq 1.$$

From the previous section we know that the number of the full binary trees with exactly n left fathers vertices (included the root) and m right fathers vertices is the Narayana number $N(n + m, m + 1)$, with $n \geq 1$, $m \geq 0$. It is obvious that for every $l \geq 1$ and $k = 1, \ldots, l$, the Narayana number $N(l, k)$ counts the number of full binary trees with exactly $l - k + 1$ left fathers vertices and $k - 1$ right fathers vertices. Now, we introduce the notion of the two-dimensional decomposition of intgers as it is presented in P.A. Macmahon [11, Vol.2, Section IX, Chapter II, Paragraph 429].

Definition 2.3.3 (Two-dimensional decompostion of integers)**.** Let $d \geq 1$, $b \geq 1$, $c \geq 0$ and $w \geq 0$ be integers. Consider a matrix $b \times d$ with elements limited in magnitude to c (zero being included) and in descending order in each row and column, and such that the sum of all the elements is exactly w (see Figure 2.3). Then, each of these matrices is called a *two-dimensional decomposition* of w with parameters d, b, c.

In this section our interest is for the case $b = 2$. It is known from [11]

$$
\begin{bmatrix}
a_{1,1} & \geq & \cdots & \geq & a_{1,d} \\
\vee| & & \vee| & & \vee| \\
\cdots & & \cdots & & \cdots \\
\vee| & & \vee| & & \vee| \\
a_{b,1} & \geq & \cdots & \geq & a_{b,d}
\end{bmatrix},
$$

with $\forall\ i = 1, \ldots, b,\ j = 1, \ldots, d,\ a_{i,j} \in \{0, \ldots, c\}$ and $\sum a_{i,j} = w$

Figure 2.3: A two-dimensional decomposition with parameters a, b, c.

that, given the following function

$$
GF_{d,c}(x) = \frac{\left(1 - x^{c+2}\right) \cdots \left(1 - x^{c+d+1}\right) \cdot \left(1 - x^{c+1}\right) \cdots \left(1 - x^{c+d}\right)}{\left(1 - x^2\right) \cdots \left(1 - x^{d+1}\right) \cdot \left(1 - x\right) \cdots \left(1 - x^d\right)}, \quad (2.21)
$$

the number of two-dimensional decomposition of w is the coefficient C_w of x^w in $GF_{d,c}(x)$, written in the power series form. Note that we can only represent integers in the set $\{0, 1, \ldots, 2dc\}$, and so we have that

$$
GF_{d,c}(1) = \sum_{w=0}^{2dc} C_w, \quad (2.22)
$$

and moreover, using the equality

$$
(1 - x)(1 + x + \cdots + x^n) = 1 - x^{n+1}, \ \forall n \geq 0,
$$

the function (2.21) becomes

$$
GF_{d,c}(x) = \frac{\displaystyle\prod_{i=1}^{d} \left(1 + \cdots + x^{c+i}\right) \cdot \prod_{i=0}^{d-1} \left(1 + \cdots + x^{c+i}\right)}{\displaystyle\prod_{i=1}^{d} \left(1 + \cdots + x^{i}\right) \cdot \prod_{i=0}^{d-1} \left(1 + \cdots + x^{i}\right)}. \quad (2.23)
$$

Now, we compute (2.23) for $x = 1$ and get

$$GF_{d,c}(1) = \frac{(c+2)\cdots(c+d+1)\cdot(c+1)\cdots(c+d)}{(2)\cdots(d+1)\cdot(1)\cdots(d)} =$$

$$= \frac{(c+d+1)!\,(c+d)!}{(c+1)!\,(c)!\,(d+1)!\,(d)!} = \binom{c+d+1}{d+1}\frac{(c+d)!}{(c+1)!\,(d)!}.$$

Then, multipling the term on the right by $\dfrac{c+d+1}{c+d+1}$, we obtain

$$GF_{d,c}(1) = \binom{c+d+1}{d+1}\frac{(c+d)!}{(c+1)!\,(d)!}\cdot\frac{c+d+1}{c+d+1} =$$

$$= \binom{c+d+1}{d+1}\frac{(c+d+1)!}{(c+1)!\,(d)!}\cdot\frac{1}{c+d+1} =$$

$$= \binom{c+d+1}{d+1}\binom{c+d+1}{d}\frac{1}{c+d+1} =$$

$$= N(c+d+1, d+1). \tag{2.24}$$

In this way we can include the special case of decomposition with $d = 0$, indeed it's sufficient to fix $GF_{0,c}(1) = N(c+1, 1) = 1$. So, we have found an important relation between the a.s. finite full binary trees and the two-dimensional decompositions, i.e. fixed $d \geq 0$ and $c \geq 0$, let be $GF_{d,c}(x)$ the function in (2.21) and C_w the number of the two-dimensional decompositions of the integer $w \in \{0, \ldots, 2dc\}$ then, from (2.22) and (2.24), we have

$$\sum_{w=0}^{2dc} C_w = \tag{2.25}$$

$=$ number of a.s. finite full binary trees with $(c+1)$ left fathers (included the root) and d right fathers.

Moreover, extending the result expressed by (2.25), we are also able to pass from a full binary tree to a two-dimensional decompostion and vice versa.

Theorem 2.3.2. *Let $d \geq 1$, $c \geq 0$. Denote with R_{dc} and $FBT_{d,c+1}$ the*

set of the two-dimensional decompositions with parameters d, c of integers
$\{0, \ldots, 2dc\}$ *and the set of the full binary trees with $c+1$ left fathers (including*
the root) and d right fathers, respectively.

Then, exists an unique representation for each element of R_{dc} in $FBT_{d,c+1}$
and vice versa.

Proof. To pass from a full binary tree with $c+1$ left fathers and d right fathers
to its related two-dimensional decomposition we have to compute the "LR"
and the "parentheses"" encodings of the tree (see Section 2.3.1), i.e. a string
of $d+c+1$ couples of parentheses (,) with exactly $d+1$ distinct nestings ().
Therefore there are c separated couples (,) in the string. Then we consider
the following defintions: $a_{1,i}$ is the number of) of the separated couples (,)
that stay after the $(i + 1)$-*th* nesting (), for every $i = 1, \ldots, d$. If $a_{1,1} < c$
then the remaining $c - a_{1,1}$) are all between the first and second nesting ().
Similarly, the element $a_{2,i}$ is the number of (between the separated couples
(,) that stay after the i-*th* nesting (), for every $i = 1, \ldots, d$ and if $a_{2,1} < c$
then the remaining $c - a_{2,1}$ (are all before the first nesting (). So, according
to the defintions of $a_{i,j}$, the two-dimensional decomposition related to the
tree is

$$
\begin{bmatrix}
a_{1,1} & \geq & \cdots & \geq & a_{1,d} \\
\text{VI} & & \text{VI} & & \text{VI} \\
a_{2,1} & \geq & \cdots & \geq & a_{2,d}
\end{bmatrix}.
$$

Note that the descending order of the rows is verified from the definitions of
elements $a_{i,j}$ and, because the number of the open parentheses (can not be
greater than the closed one, even the descending order of the columns is also
verified.

To pass from a two dimensional decomposition with parameters d and c to
its related full binary tree with $c + 1$ left fathers (including the root) and
d right fathers it is enough to consider the definitions of the elements $a_{i,j}$
given above and write the related string of parentheses which encode a full
binary tree with $c + 1$ left fathers and d right fathers. The uniqueness of the
representations is given respectively by the definitions of the elements of the

decompositions and the uniqueness of the "LR" encoding of the trees (see Remark 20). $\qquad\square$

Example 2.3.2. Let $d = 2, c = 1$, so by (2.25) we have that

$$\sum_{w=0}^{4} C_w = N(4,3) = 6.$$

$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, w = 0$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, w = 1$	$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, w = 2$
LLRRLRLR	LLRLRRLR	LLRLRLRR
(()) () ()	(() ()) ()	(() () ())

Table 2.1: (a)-Representations of finite full binary trees with 2 left fathers (including the root) and 2 right fathers as two-dimensional decompositions of integers
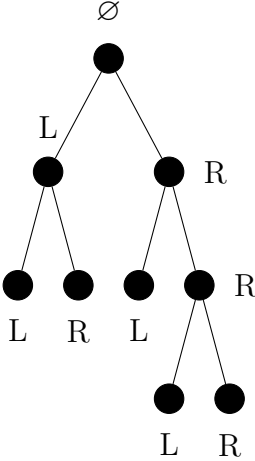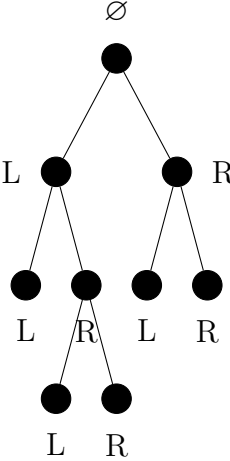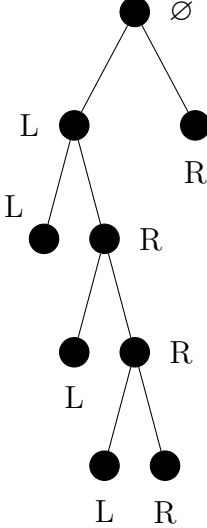
$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, w = 4$	$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, w = 2$	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, w = 3$
LRLRLLRR	LRLLRRLR	LRLLRLRR
() () (())	() (()) ()	() (() ())

Table 2.2: (*b*)-Representations of finite full binary trees with 2 left fathers (including the root) and 2 right fathers as two-dimensional decompositions of integers

In Tables 2.1 and 2.2 you can see the rappresentations of all the 6 full binary trees with 2 left fathers and 2 right fathers, in accordance with the algorithm seen above. For each tree is shown the "LR" encoding, the related "parentheses" encoding (see Section 2.1) and the two-dimensional decomposition related and moreover the integer w rappresented in such decomposition.

2.3.3 Upper and lower bounds for the full binary trees still alive in the *M-th* generation

In this section we make a statistical investigation on full binary trees that have some vertices at the generation M, i.e. we want to find the probability that a full binary tree has vertices in the generation M, $M \geq 1$. We define such trees in the following way

Definition 2.3.4 (Full binary trees still alive at the M-*th* generation)**.** A full binary tree that has at least one vertex in the generation M is called a full binary tree still alive in the M-*th* generation.

It is elaborate to count exactly all the full binary trees that are still alive at the generation M, so what we are going to do is an estimation of that. At first, we compute the minimum and maximum generations reached by a full binary tree with n left fathers and m right fathers, where $n \geq 1$, $m \geq 0$. It's easy to check that the maximum generation is reached when there's one and only one vertex in each generation and so it's equal to $n + m$. To find the minimum generation, we define the *Triangular Configuration* of a full binary tree

Definition 2.3.5 (Triangular Configuration)**.** A full binary tree verifies the triangular configuration at the generation T if every vertex of the generation K is a father, for $K = 0, \ldots, T - 1$. Moreover, in the triangular configuration at the generation T, the full binary tree has

$$\sum_{k=0}^{T-1} 2^k = 2^T - 1 \text{ fathers.}$$

Thus, if a full binary tree verifies the Triangular Configuration at the generation T then the minimum generation is equal to T. If there are remaining fathers then they are located in the vertices of the generation T.

Remark 21. Note that the remaining fathers are located not in all the vertices of the generation T, otherwise the tree verifies the Triangular Configuration

at the generation $T + 1$).

So, we consider a full binary tree with $n \geq 1$ left fathers and $m \geq 0$ right fathers such that it verifies the Traingular Configuration at the generation T then, denoting with x the number of the remaining fathers, we have that

$$n + m = \sum_{k=0}^{T-1} 2^k + x = 2^T - 1 + x, \tag{2.26}$$

where $x = 0, \ldots, 2^T - 1$.

Remark 22. Note, from the Definition of the Triangular Configuration, that if $x = 0$ then the minimum generation is T, and if $x = 1, \ldots, 2^T - 1$ the minimum generation is $T + 1$.

From the (2.26) we get

$$\log_2 (n + m + 1) = \log_2 \left(2^T + x\right), \tag{2.27}$$

where $T \leq \log_2 \left(2^T + x\right) < T + 1$, for $x = 0, \ldots, 2^T - 1$.

Finally from the (2.27) and the Remark 22, we have that the minimum generation is

$$\lceil \log_2 (n + m + 1) \rceil.$$

We can sumarize the results on the maximum and minimum generation of the full binary trees. Let $(\tau, e_\tau) \in T_1^{(2)}$ be a full binary tree with n left fathers (including the root) and m right fathers, $n \geq 1$, $m \geq 0$, then the minimum generation reached by the tree is equal to

$$\lceil \log_2 (n + m + 1) \rceil,$$

and the maximum generation is equal to

$$n + m.$$

Because the minimum and the maximum generations of a full binary tree depend on the total number of fathers, we consider this changing of parameters

$$\begin{cases} k = n + m \geq 1, \text{ the total number of the fathers of the tree} \\ m \geq 0, \text{ the number of the right fathers} \end{cases}.$$

Now, we want to compute a lower bound for the probability that a full binary tree is still alive at the generation $M \geq 1$. Consider all the binary trees with k total fathers (for every $m = 0, \ldots, k-1$) that reach at least the generation M, i.e. with the minimum generation

$$\lceil \log_2 (k + 1) \rceil \geq M. \tag{2.28}$$

From the theory of the number of digits in a certain base, we have that the condition (2.28) becomes

$$\lceil \log_2 (k + 1) \rceil = \lfloor \log_2 (k) \rfloor + 1 \geq M,$$

and so

$$\lfloor \log_2 (k) \rfloor \geq M - 1 \Rightarrow k \geq 2^{M-1}.$$

Note that this is a lower bound because we don't consider, for example, the full binary trees such that

$$M \leq k \leq 2^{M-1} - 1.$$

For instance, if we consider $M = 4$, we count all the full binary trees with a total number of fathers $k \geq 8$, i.e. all the Narayana numbers $N(k, m + 1)$ with $k \geq 8$ and $m = 0, \ldots, k-1$, but we don't consider the full binary tree in the first column of the Table 2.2 that has 4 total fathers and is still alive at the generation 4.

So, if we denote with

$$\mathbb{P}_\tau (M)$$

the probability that a full binary tree τ is still alive in the generation M, we get the following lower bound

$$\sum_{k \geq 2^{M-1}} \sum_{m=0}^{k-1} N\left(k, m+1\right) p_0^{m+1} p_2^{k-m} q_0^{k-m} q_2^m < \mathbb{P}_\tau\left(M\right).$$

Now, for the upper bound of the probability that a full binary tree is still alive in the generation $M \geq 1$, we consider the full binary trees with k total fathers (for every $m = 0, \ldots, k-1$) that reach the generation M with the maximum generation, i.e.

$$k \geq M.$$

Note that this is an upper bound becasue we consider, for example, also the full binary trees with a total number of fathers $k \geq M$ but such that they do not reach the generation M. For instance, if we consider $M = 4$, we count all the full binary trees with a total number of fathers $k \geq 4$, i.e. all the Narayana numbers $N\left(k, m+1\right)$ with $k \geq 4$ and $m = 0, \ldots, k-1$, but we consider also the full binary trees in the first two columns of the Table 2.1. So, we get the following upper bound of $\mathbb{P}_\tau\left(M\right)$

$$\sum_{k \geq M} \sum_{m=0}^{k-1} N\left(k, m+1\right) p_0^{m+1} p_2^{k-m} q_0^{k-m} q_2^m > \mathbb{P}_\tau\left(M\right).$$

Finally, using the computational software program *Wolfram Mathematica*, we can simplify the bounds of $\mathbb{P}_\tau\left(M\right)$ as follows

$$\sum_{k \geq 2^{M-1}} p_0 p_2^k q_0^k \left({}_2F_1\left(1-k, -k; 2; \frac{p_0 q_2}{p_2 q_0}\right)\right) <$$

$$< \mathbb{P}_\tau\left(M\right) < \tag{2.29}$$

$$< \sum_{k \geq M} p_0 p_2^k q_0^k \left({}_2F_1\left(1-k, -k; 2; \frac{p_0 q_2}{p_2 q_0}\right)\right), \ \forall M \geq 1,$$

where

$$_2F_1\left(a,b;c;d\right)=\sum_{n\geq0}\frac{\left(a\right)_n\left(b\right)_n}{\left(c\right)_n}\cdot\frac{z^n}{n!},$$

with

$$\left(q\right)_n=\begin{cases}1, & n=0\\q\left(q+1\right)\cdots\left(q+n-1\right), & n>0\end{cases},$$

is the hypergeometric function, defined for $|z|<1$.

Note that the condition

$$\frac{p_0q_2}{p_2q_0}<1$$

is equal to

$$p_2>q_2, \tag{2.30}$$

indeed

$$\frac{p_0q_2}{p_2q_0}<1\Leftrightarrow\frac{\left(1-p_2\right)q_2-p_2\left(1-q_2\right)}{p_2\left(1-q_2\right)}<0\Leftrightarrow p_2>q_2.$$

So, combining the condition (2.30) with the a.s. finite condition seen above $p_2+q_2\leq1$, we get the following condition

$$\begin{cases}p_2>q_2, & 0<p_2<1/2\\p_2\leq1-q_2, & 1/2\leq p_2<1\end{cases}.$$

Part II

Stochastic Models for Angiogenesis

Chapter 3

Selected topics about angiogenesis from literature

In literature we may find many mathematical models dealing with some of the features of the angiogenic process, and in particular there are relevant model used to describe the growth of the blood vessels. The purpose of most of these models is to describe the network of the blood vessels, by the coupling of the stochastic movement of the blood vessels and the biological structure of the tissue in which the vessel move. Otherwise, in our models we are focused on the statistical behavior of the blood vessels. Thus, in Chapters 4 and 5 we propose some parametric simplified models based on two-type Galton-Watson processes and trees which allow us to obtain statistical information on the offspring distribution μ, i.e. the branching mechanism of the blood vessels. In this Chapter, neglecting the biological point of view, we present three relevant stochastic models from literature and relate our approach to them.

3.1 A Geometrical Model

We start with the model presented in [9] by V. Capasso and D. Morale. It is a geometrical model whose discrete approximation is a branching process

and so, when the bifurcation/death rates are homogeneous (time independent), it can be considered as a superprocess, i.e. the weak limit of processes that we have treated and that we use in our models. Thus, our study can be considered as statistics on the trees generated by the processes of V. Capasso and D. Morale. Relative to the model presented in [9], one of the strengths of our approach for the angiogenesis modelling is to be a parametric statistical model, while one of its weakness is the homogeneity request (at least in this thesis). Now we have a look to the settings of the model proposed. The blood vessels are modelled as the trajectories developed by the heads of the vessels, while the behavior of the heads is modeled as a stochastic marked counting process. Specifically, as regards the vessels network, let $N(t)$ be the jump process that counts the number of heads at the time $t \geq 0$, and $X^i(t) \in \mathbb{R}^d$ the location at time t of the i-th head. Then, denoting with T_i the birth time of the i-th head, i.e. the time when an existing vessel branches and i-th blood vessel springs up, the trajectory-set

$$\{X^i(s), T_i \leq s \leq t\}, \ \forall i = 1, \ldots, N(t),$$

represents the blood vessel starting at time T_i from the i-th head, at time t, and so

$$X(t) = \bigcup_{i=1}^{N(t)} \{X^i(s), T_i \leq s \leq t\}$$

is the blood vessels network, i.e. the union of the trajectories of the heads of the vessels. The counting jump process $N(t)$ is defined starting from the following marked point process

$$\Phi = \sum_n \delta_{(T^n, Y^n)},$$

where T^n and Y^n are the birth time and location, respectively, of the n-th head. Indeed, from the process Φ is determinated the random measure $\Phi(A)$

on $\mathscr{B}_{\mathbb{R}^+ \times \mathbb{R}^d}$

$$\Phi\left(A\right) := \sum_n \delta_{(T^n, Y^n)}\left(A\right), \ A \subseteq \mathscr{B}_{\mathbb{R}^+ \times \mathbb{R}^d},$$

and so the process $N\left(t\right)$ is defined by

$$N\left(t\right) = \Phi\left([0, t] \times \mathbb{R}^d\right).$$

Then, starting from this stochastic settings, the vessels extension is treated by SDEs, depending on the biological structure and the nature of the angiogenic process.

3.2 Random Walk Models

In this section we consider two models presented by M.J. Plank and B.D. Sleeman in [14] and in [15]. They are both geometrical models whose discrete approximations are given by particular random walks, i.e. the movements of the blood vessels are considered random walks in suitable spaces. Thus, also our work can be seen as a statistical investigation on the multytipe Galton-Watson trees based on particular random walks (i.e. the contour processes) and their limit, with the difference that M.J. Plank and B.D. Sleeman use random walks in a geometrical way, whereas we do it from a statistical point of view. Relative to the models presented in [14] and [15], one of the strengths of our approach for the angiogenesis modelling is parametric statistical use of particular random walks, while one of its weakness is the lack of geometrical point of view.

3.2.1 A Reinforced Random Walk Model

Now, we explain a lattice model based on the reinforced random walk (RRW) for the movement of the heads of the blood vessel. This model has been presented by M.J. Plank and B.D. Sleeman in [14]. The purpose of this short presentation is to show the main importance of the RRW to determine

the probability density function for the movement of the heads. At first we give a theoretical introduction of the RRW and then we expose the model. For the theory of the RRW we refer to [16]. Essentially a RRW is a classic random walk in which the transition probability from time n to time $n+1$ depends on some random quantities called *weigths* regarding the history of the process until the time n.

Definition 3.2.1 (Reinforced Random Walk). A reinforced random walk is a couple $(\boldsymbol{X}, \boldsymbol{W})$, where $\boldsymbol{X}$ is a sequence $\boldsymbol{X} = \{X_i, i \geq 0\}$ of integer random variables and $\boldsymbol{W} = \{w(n,j), n \geq 0, j \in \mathbb{Z}\}$ is a matrix of positive random variables, all defined on the same probability space, such that if F_n is the σ-field defined as follows

$$F_n = \sigma\left(\{X_i, 0 \leq i \leq n; w(i,j), j \in \mathbb{Z}, 0 \leq i \leq n\}\right),$$

then

i) $w(n+1, j) - w(n, j) \geq 0$,

ii) $\mathbb{P}(X_{n+1} = j+1 | X_n = j, F_n) = 1 - \mathbb{P}(X_{n+1} = j-1 | X_n = j, F_n) =$

$$= \frac{w(n,j)}{w(n,j) + w(n,j-1)}. \tag{3.1}$$

The element $w(n, j)$ of the matrix $\boldsymbol{W}$ is the *weigth* at time n of the interval $(j, j+1)$, for every $j \in \mathbb{Z}$, and so i) means that the map $n \mapsto w(n, j)$ is increasing, in particular it is strictly increasing if the walk crosses the interval $(j, j+1)$ between times n and $n+1$, i.e.

$$X_n = j+1, \; X_{n+1} = j \text{ or } X_n = j, \; X_{n+1} = j+1,$$

otherwise $w(n, j) = w(n+1, j)$.

Moreover, we say that $(\boldsymbol{X}, \boldsymbol{W})$ is *initially fair* if the *initial weights* $w(0, j)$, $\forall j \in \mathbb{Z}$, are all equal to 1.

Condition ii) means that if at the time n we have $X_n = j$ then the probability

that $X_{n+1} = j + 1$ (or $j - 1$) is equal to the weight of the interval $(j, j + 1)$ (or $(j - 1, j)$) at the time n over the total weight of the interval $(j - 1, j + 1)$ at the time n.

In [14], M.J. Plank and B.D. Sleeman introduce an one dimensional lattice model in which the heads of the blood vessels move on a regular mesh and their density at mesh point n at time t, $p_n(t)$, is such that

$$\frac{\partial p_n(t)}{\partial t} = \widehat{\tau}_{n-1}^+ \cdot p_{n-1}(t) + \widehat{\tau}_{n+1}^- \cdot p_{n+1}(t) - \left(\widehat{\tau}_n^+ + \widehat{\tau}_n^-\right) \cdot p_n(t), \qquad (3.2)$$

where $\widehat{\tau}_n^+$ and $\widehat{\tau}_n^-$ are the transition rates of the heads moving from mesh point n to mesh point $n+1$ and $n-1$ respectively (i.e. a sort of *infinitesimal* transistion probabilities from mesh point n to mesh point $n + 1$ or $n - 1$). At this point, in accordance with the (3.1), the transition rates are defined in the following way

$$\widehat{\tau}_n^{\pm}(\boldsymbol{w}) = 2\lambda \frac{\tau\left(w_{n\pm 1/2}\right)}{\tau\left(w_{n-1/2}\right) + \tau\left(w_{n+1/2}\right)}, \lambda > 0,$$

where $\boldsymbol{w}$ is the vector of control substances related to the angiogenic process and the transition probabilities $\tau(\cdot)$ depend only on $\boldsymbol{w}$.

3.2.2 A Circular Random Walk Model

Here we explain another model proposed by M.J. Plank and B.D. Sleeman for the movement of the heads of the blood vessels. It is a non-lattice model introduced in [15] and based on the circular random walk. In [15], M.J. Plank and B.D. Sleeman refer to one of the first work on the circular random walk [17] by N. A. Hill and D.-P. Häder. As RRW model, our aim is to show the main importance of the circular random walk in determining the probability density function for the movement of the heads. In this model each head of a blood vessel performs a random walk on the unit circle, described by the random variable $\Theta(t)$, whose value $\theta(t)$ is the direction of the head at time $t \geq 0$. Denoting with k the lenght of the time step, at each step a head of

a blood vessel has a probability, $a(\theta)$, of turning clockwise through an angle δ, a probability, $b(\theta)$, of turning anti-clockwise through an angle δ, and a probability, $1 - a(\theta) - b(\theta)$, of continuing in the same direction,

$$\mathbb{P}(\Theta(t + k) - \Theta(t) = \delta) = a(\theta),$$

$$\mathbb{P}(\Theta(t + k) - \Theta(t) = -\delta) = b(\theta),$$

$$\mathbb{P}(\Theta(t + k) - \Theta(t) = 0) = 1 - a(\theta) - b(\theta),$$

where $a(\theta)$ and $b(\theta)$ are functions.

The probability density function, $f(\theta, t)$, for $\Theta(t)$ is defined in the usual way

$$f(\theta, t)\, d\theta = \mathbb{P}(\theta \leq \Theta(t) < \theta + d\theta),$$

indeed, from the Mean Value Theorem of Integration, we have that exists γ in $[\alpha, \alpha + \beta]$ such that, $\forall \alpha, \beta \in [0, 2\pi[$ with $\alpha + \beta \in [0, 2\pi[$,

$$f(\gamma, t) = \frac{1}{\beta} \int_{\alpha}^{\alpha+\beta} f(\xi, t)\, d\xi,$$

and thus

$$f(\theta, t) = \frac{1}{d\theta} \int_{\theta}^{\theta+d\theta} f(\xi, t)\, d\xi = \frac{1}{d\theta} \cdot \mathbb{P}(\theta \leq \Theta(t) < \theta + d\theta).$$

It is shown in [17] that, in the limit $k \to 0$, $\delta \to 0$ (such that $\delta^2/k = A$, for some fixed constant A), $f(\theta, t)$ satisfies the following equation

$$\frac{\partial}{\partial t} f(\theta, t) = -\frac{\partial}{\partial \theta}(\mu(\theta) f(\theta, t)) + \frac{1}{2}\frac{\partial^2}{\partial \theta^2}\left(\sigma^2(\theta) f(\theta, t)\right), \qquad (3.3)$$

where

$$\mu(\theta) = \lim_{k \to 0, \delta \to 0}\left(\frac{1}{k} \cdot \mathbb{E}\left[\Theta(t + k) - \Theta(t)\right]\right),$$

$$\sigma^2(\theta) = \lim_{k \to 0, \delta \to 0}\left(\frac{1}{k} \cdot Var\left[\Theta(t + k) - \Theta(t)\right]\right).$$

Moreover, M.J. Plank and B.D. Sleeman in [15] prove that the equation (3.3) is the continuous limit, $k \to 0$, $\delta \to 0$, of the reinforced random walk equation (3.2).

Chapter 4

Models based on Two-Type Galton-Watson Processes

Our approach to model the growth of the blood vessels is the following. Consider a blood vessel, at the time $t \geq 0$, as the composition of two-type objects, the head of the blood vessel, i.e. the material point that could grow up while the time is running and the body of the blood vessel itself (Figure 4.1). In this way, the body of a blood vessel can be seen as the result of the
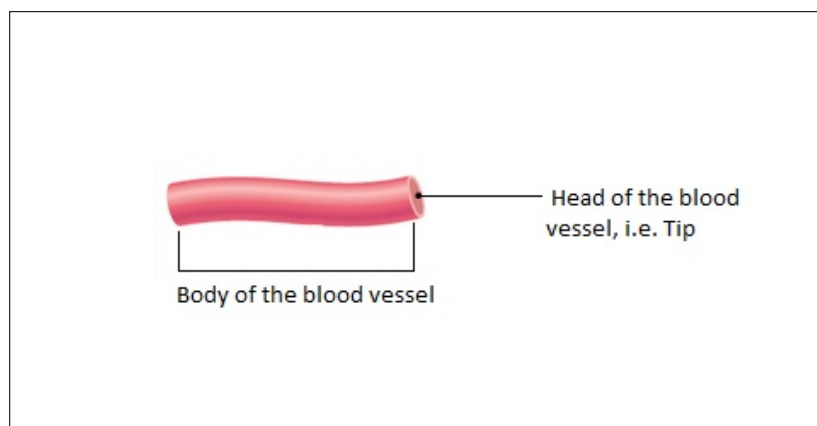

Figure 4.1: Tip and body of a blood vessel

growth of a head, if it happens, i.e. the body of a blood vessel represents the trajectory of the head, conditionally to the growth. In the following

model we consider the body of a vessel blood as union of parts, each one with the same dimensions and produced by the growth of a tip in a unit time interval. Now, considering the heads and the parts of the blood vessels as two-type particles, each of which has an appropriate offspring distribution, the model of the growth of the blood vessels is given by the two-type branching processes, where type 1 particles are the heads of the blood vessels (*tips*) and the type 2 particles are the essential units of the body of the blood vessels (*branches*) (see Figure 4.2). In this Chapter we start with a model based
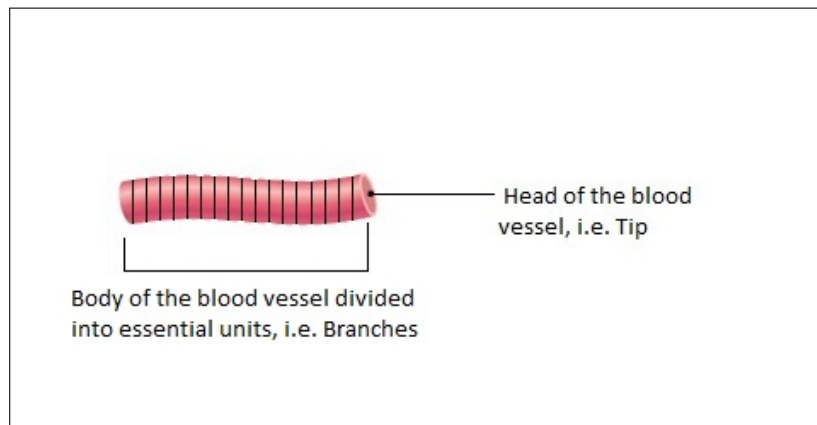
Figure 4.2: Tip and branches of a blood vessel

on a discrete time two-type Galton-Watson process, where the two-types particles are defined as above and with a particular offspring distribution. Thus, we are able to compute all the fundamental quantities that describe the process. Then, we present a model based on a continuous time two-type Galton-Watson process, where particular definitions of the lifetimes of the particles are used and the fundamental quantities are deduced using a technique of inverse problem.

4.1   The Discrete Time Model

We consider a discrete time two-type Galton-Watson process

$$(\boldsymbol{Z}_n; n \geq 0) = (Z_{1,n}, Z_{2,n}; n \geq 0),$$

where $Z_{1,n}$ and $Z_{2,n}$ are the number of *tips* and *branches* alive at the n-*th* generation. From the Section 1.1.1, we know that with the offspring distribution $p^{(i)}(j_1, j_2)$ we are able to describe the whole process. Now, we introduce the marginal distributions of the tips and the branches respectively and then we deduce the joint distributions $p^{(i)}(j_1, j_2)$.

4.1.1 Tips distribution

For an alive tip at the generation n we consider the following marginal distribution

1) the tip dies and produces no particles (i.e. the blood vessel stops growing, for example by the chemotherapy, and the last branch created dies too), with probability p_0, i.e. at the generation $n + 1$ there are no tips or branches alive,

2) the tip grows up, creating a branch behind itself (i.e. the blood vessel grows and creates a branch), with probability p_1, i.e. at the generation $n + 1$ there are the survived tip and the branch created by its growth,

3) the tip splits itself in two new tips (i.e. the blood vessel splits itself in new two others), with probability p_2, i.e. at the generation $n + 1$ there are the two new tips created by the ancestor tip,

4) the tip survives (i.e. the blood vessel does not grow but it is still alive), with probability p_3, i.e. at the generation $n+1$ there is only the survived tip.

where p_0, p_1, p_2, $p_3 \geq 0$ and $p_0 + p_1 + p_2 + p_3 = 1$.

4.1.2 Branches distribution

For an alive branch at the generation n we introduce the following marginal distribution

1) the branch dies and produces no particles (i.e. the last branch of an existing blood vessel dies, for instance by fibrinoid necrosis of the blood vessel induced by chemioterapy), with probability q_0, i.e. at the generation $n+1$ there are no tips or branches alive,

2) the branch survives (i.e. the last branch of an existing blood vessel doesn't die), with probability q_1, i.e. at the generation $n+1$ there is only the survived branch,

3) the branch survives and produces a new tip from itself (i.e. the last branch of an existing blood vessel doesn't die and a new one could born from it), with probability q_2, i.e. at the generation $n+1$ there are the survived branch and the new tip born from it,

where q_0, q_1, $q_2 \geq 0$ and $q_0 + q_1 + q_2 = 1$.

Remark 23. When a particle produces only a particle of its same type, we say that the particle *survives*.

4.1.3   Joint distribution

Now, we deduce the joint distribution of tips and branches. In general, in accordance with Section 1.1.1, the distributions in Sections 4.1.1 and 4.1.2 can be considered the really offspring distributions $p^{(1)}(j_1, j_2)$, $p^{(2)}(j_1, j_2)$ of the process $\boldsymbol{Z}_n$, but in this model it is partially true. Indeed, as regards the branches, we assume that the new tip born from a branch at the generation n, with probability q_2, is visible in the generation $n+1$ and so we have that the marginal distribution of the branches in Section 4.1.2 is exactly the offspring distribution $p^{(2)}(j_1, j_2)$. As regards the tips, on the contrary, we assume that if a tip grows up (with probability p_1) at the generation n, then at the generation $n+1$ we consider, in addition to the survived tip, the offspring of the new branch created. Thus, if a tip grows up at the generation n, then at the generation $n+1$ we count the sum of the tip with the offspring of the new branch created from the growth of the tip.

So, from the marginal distributions seen above, we obtain the following joint distribution from an alive tip at the generation n, for every $n \geq 0$ (see Table 4.1). Thus, we can write the offspring distribution of the process $\boldsymbol{Z}_n$,

$\boldsymbol{Z}_n$	$\boldsymbol{Z}_{n+1}$	*Probability*
$(1,0)$	$(0,0)$	p_0
$(1,0)$	$(1,0)$	$p_1 q_0 + p_3$
$(1,0)$	$(1,1)$	$p_1 q_1$
$(1,0)$	$(2,1)$	$p_1 q_2$
$(1,0)$	$(2,0)$	p_2

Table 4.1: Joint distribution from an alive tip at the generation n, for every $n \geq 0$.

$$
\begin{cases}
p^{(1)}(0,0) = p_0 \\
p^{(1)}(1,0) = p_1 q_0 + p_3 \\
p^{(1)}(1,1) = p_1 q_1 \\
p^{(1)}(2,1) = p_1 q_2 \\
p^{(1)}(2,0) = p_2
\end{cases}
, \quad
\begin{cases}
p^{(2)}(0,0) = q_0 \\
p^{(2)}(0,1) = q_1 \\
p^{(2)}(1,1) = q_2
\end{cases}
. \tag{4.1}
$$

Before introducing the generating functions $f^{(1)}(s_1, s_2)$, $f^{(2)}(s_1, s_2)$ and the mean matrix $\boldsymbol{M}$, we summarize in Table 4.2 the meaning of the offspring distribution shown in (4.1) concerning the growth of the blood vessels, i.e. we explain how we model angiogenesis by the offspring distribution (4.1).

Remark 24. Remark that the probability $p^{(i)}(j_1, j_2)$ represents the probability that are j_1 type 1 particles and j_2 type 2 particles at the generation $n+1$ from a single alive type i particle at the generation $n \geq 0$, for every $i = 1, 2$.

Probability	*Angiogenesis*
$p^{(1)}(0,0)$	The head of a blood vessel dies. So the vessel stops growing and the last branch created dies, i.e. by a chemioterapy action
$p^{(1)}(1,0)$	The head of a blood vessel survives and no branches are created
$p^{(1)}(1,1)$	The head of a blood vessel grows creating a new branch of the blood vessel
$p^{(1)}(2,1)$	The head of a blood vessel grows creating a new branch of the blood vessel from which is already born a new head of a blood vessel
$p^{(1)}(2,0)$	The head of a blood vessel splits itself in other two and so two new blood vessels are born from the croosroad
$p^{(2)}(0,0)$	A branch of a blood vessel dies and so no one else is born from it
$p^{(2)}(0,1)$	A branch of a blood vessel survives but no one else is born from it
$p^{(2)}(1,1)$	A branch of a blood vessel survives and a new one is born from it

Table 4.2: Meaning of the offspring distribution of tips and branches.

Remark 25. Remark that when we say that a new blood vessel is born from a branch of a blood vessel we mean that the head of a new blood vessel is born from the branch.

The probabilities $p^{(1)}(0,0)$, $p^{(1)}(1,0)$, $p^{(2)}(0,0)$ and $p^{(2)}(0,1)$ can be called *chemotherapic* probabilities, i.e. they can be considered respectively

- the probability that a chemotherapic action (CA) kills the last part created of a blood vessel (branch) and stops the growth of the blood vessel,

- the probability that a CA allows the growth of a blood vessel killing the last branch created or stops the growth of the vessel without killing the head,

- the probability a CA kills a branch of a blood vessel,

- the probability a CA may keep alive a branch of a blood vessel but it does not allow the creation of new blood vessels from it.

Similarly, the probabilities $p^{(1)}(1,1)$, $p^{(1)}(2,1)$, $p^{(1)}(2,0)$ and $p^{(2)}(1,1)$ can be called *proliferation* probabiliies, i.e. they can be considered respectively

- the probability that a head of a blood vessel grows, producing a new branch,

- the probability that a head of a blood vessel grows and another blood vessel is born from the last branch created,

- the probability that a head of a blood vessel splits in two,

- the probability that a branch of a blood vessel produces a new blood vessel from it.

4.1.4 Generating functions and mean matrix $\boldsymbol{M}$

Now, we are able to write the generating functions $f^{(1)}(s_1, s_2)$, $f^{(2)}(s_1, s_2)$ with $s_1, s_2 \in [0, 1]$, and the mean matrix $\boldsymbol{M}$. From the offspring distributions (4.1) and the definitions of $f^{(1)}(s_1, s_2)$, $f^{(2)}(s_1, s_2)$ and $\boldsymbol{M}$ (see Section 1.1.1), we get

$$f^{(1)}(s_1, s_2) = p_0 + (p_1 q_0 + p_3) s_1 + (p_1 q_1) s_1 s_2 + (p_1 q_2) s_1^2 s_2 + p_2 s_1^2,$$

$$f^{(2)}(s_1, s_2) = q_0 + q_1 s_2 + q_2 s_1 s_2,$$

and

$$\boldsymbol{M} = \begin{pmatrix} p_1 q_2 + 2p_2 + p_1 + p_3 & p_1 (q_1 + q_2) \\ q_2 & q_1 + q_2 \end{pmatrix}.$$

Thus, the process $\boldsymbol{Z}_n$ is *positive regular* and *no singular*. Moreover, the

maximal eigenvalue of the mean matrix $\boldsymbol{M}$ is

$$\rho = \frac{p_1 q_2 + 2p_2 + p_1 + p_3 + q_1 + q_2}{2} +$$

$$+ \frac{\sqrt{4q_2 p_1 (q_1 + q_2) + (p_1 q_2 + 2p_2 + p_1 + p_3 - q_1 - q_2)^2}}{2},$$

and so, by the Theorem 1.1.1, we obtain the following theoretical classification of process $\boldsymbol{Z}_n$,

$$p_0 + q_0 \leq p_1 q_2 + p_2 \text{ or } \begin{cases} p_0 + q_0 > p_1 q_2 + p_2 \\ q_0 (p_0 - p_2) < p_1 q_2 \end{cases} \implies \text{SuperCritical}, \quad (4.2)$$

$$\begin{cases} p_0 + q_0 > p_1 q_2 + p_2 \\ q_0 (p_0 - p_2) = p_1 q_2 \end{cases} \implies \text{Critical},$$

and

$$\begin{cases} p_0 + q_0 > p_1 q_2 + p_2 \\ q_0 (p_0 - p_2) > p_1 q_2 \end{cases} \implies \text{SubCritical}.$$

Thus, we can make some considerations. By the offspring distribution (4.1) shown in the Section 4.1.3, $p_0 + q_0$ and $p_1 q_2 + p_2$ may be considered respectively the *death probability* (i.e. the probability of minimum production) and the *probability of maximum production*. Moreover, $p_1 q_2$ is the probability of maximum production from a tip and the sign of $q_0 (p_0 - p_2)$ is related to the death/procution rate of a tip. Moreover, note that with probability p_2 a tip splits itself in other two and so the effect of the growth of the tips is amplified. Finally in the Exapmle 4.1.1 we explain that this model is *unbalanced* to the SuperCrital case, i.e. if tips and branches have the same marginal distributions then the process $\boldsymbol{Z}_n$ is SuperCritical.

Remark 26. Note that the extinction probability $\boldsymbol{q} = \left(q^{(1)}, q^{(2)}\right)$, defined in Section 1.1.1, represents the probability that an angiogenic process starting from a tip $(q^{(1)})$ or a branch $(q^{(1)})$ ends.

Example 4.1.1 (Equiprobable marginal distributions). Consider the following equiprobable marginal distributions,

$$p_i = \frac{1}{4}, \ \forall i = 0, 1, 2, 3,$$

and

$$q_i = \frac{1}{3}, \ \forall i = 0, 1, 2.$$

The offspring distributions (4.1) become

$$
\begin{cases}
p^{(1)}(0,0) = 1/4 \\
p^{(1)}(1,0) = 1/3 \\
p^{(1)}(1,1) = 1/12 \\
p^{(1)}(2,1) = 1/12 \\
p^{(1)}(2,0) = 1/4
\end{cases}
,
\begin{cases}
p^{(2)}(0,0) = 1/3 \\
p^{(2)}(0,1) = 1/3 \\
p^{(2)}(1,1) = 1/3
\end{cases}
.
$$

So, the generating functions $f^{(1)}(s_1, s_2)$, $f^{(2)}(s_1, s_2)$ and the mean matrix $\boldsymbol{M}$ are

$$f^{(1)}(s_1, s_2) = \frac{1}{4} + \frac{1}{3}s_1 + \frac{1}{12}s_1 s_2 + \frac{1}{12}s_1^2 s_2 + \frac{1}{4}s_1^2,$$

$$f^{(2)}(s_1, s_2) = \frac{1}{3} + \frac{1}{3}s_2 + \frac{1}{3}s_1 s_2,$$

$$M = \begin{pmatrix} \dfrac{13}{12} & \dfrac{1}{6} \\[2ex] \dfrac{1}{3} & \dfrac{2}{3} \end{pmatrix}.$$

The maximum eigenvalue of the mean matrix $\boldsymbol{M}$ is equal to

$$\frac{21 + \sqrt{57}}{24} > \frac{21 + 7}{24} > 1,$$

and so the process is SuperCritical. The same result is obtained considering

the following quantities

$$p_0 + q_0 = \frac{7}{12}, \quad p_1 q_2 + p_2 = \frac{1}{3}, \quad q_0 \left(p_0 - p_2 \right) = 0, \quad p_1 q_2 = \frac{1}{12},$$

indeed they satisfy the (4.2) and so the process is SuperCritical.

Moreover, from the Theorem 1.1.1 in the Section 1.1.1, we know that the extinction probability $\boldsymbol{q}$ of the process is the only solution of $\boldsymbol{f}\left(\boldsymbol{s}\right) = \boldsymbol{s}$ in $[0,1[^2$. This is equivalent to solve in $[0,1[^2$ the following system

$$\begin{cases} \dfrac{1}{4} + \dfrac{1}{3}s_1 + \dfrac{1}{12}s_1 s_2 + \dfrac{1}{12}s_1^2 s_2 + \dfrac{1}{4}s_1^2 = s_1 \\[4mm] \dfrac{1}{3} + \dfrac{1}{3}s_2 + \dfrac{1}{3}s_1 s_2 = s_2 \end{cases} . \tag{4.3}$$

The system (4.3) has three solutions

$$\left(1,1\right), \quad \left(2 - \sqrt{2}, \frac{\sqrt{2}}{2}\right), \quad \left(2 + \sqrt{2}, -\frac{\sqrt{2}}{2}\right).$$

Note that

$$\left(2 + \sqrt{2}, -\frac{\sqrt{2}}{2}\right)$$

is not acceptable and that

$$\left(0,0\right) < \left(2 - \sqrt{2}, \frac{\sqrt{2}}{2}\right) < \left(1,1\right),$$

and so $\boldsymbol{q} = \left(2 - \sqrt{2}, \dfrac{\sqrt{2}}{2}\right)$.

Moreover, since

$$2 - \sqrt{2} < \frac{\sqrt{2}}{2},$$

then the probability that an angiogenic process starting from a branch ends is greater than the probability that an angiogenic process starting from a tip

ends.

4.2 The Continuous Time Model

In this Section we introduce a discrete approximation of a two-type continuous time Galton-Watson process, $\boldsymbol{Z}(t) = (Z_1(t), Z_2(t), t \geq 0)$, where the two-type particles are the same of Section 4.1, i.e. tips and branches respectively. In particular, we propose a discrete model with time interval $\Delta t > 0$ and then we pass to the limit $\Delta t \to 0$.

Moreover, using a technique of inverse problem, we compute the mean matrix $\boldsymbol{M}(t)$, for every $t \geq 0$, and from that we show the theoretical classification of the process (see Section 4.1.4).

Remembering, from Section 1.1.2, that a two-type continuous time Galton-Watson process is totally described by the stricly positive parameters of the exponential distributions of the particles lifetimes

$$\boldsymbol{a} = (a_1, a_2) \in \mathbb{R}_+^2,$$

and the offspring distribution

$$\boldsymbol{p}(\boldsymbol{j}) = \left(p^{(1)}(j_1, j_2), p^{(2)}(j_1, j_2) \right),$$

we at first define the lifetimes and the infinitesimal marginal and joint distributions of the particles, then we are able to compute the parameters a_1, a_2 and the offspring distribution $\boldsymbol{p}(\boldsymbol{j})$.

4.2.1 Lifetimes

In this model we consider a different defintion of the usual lifetimes of the particles, i.e. the lifetime Γ_i of the type i particles is the first time in which a process $\boldsymbol{Z}(t)$ starting from a single type i particle changes state. Thus, we

define the following lifetimes for the two-type particles, for every $i = 1, 2$,

$$\Gamma_i := \inf\{\tau > t | \boldsymbol{Z}(\tau) \neq \boldsymbol{e}_i\}, \text{ conditioning to } \boldsymbol{Z}(t) = \boldsymbol{e}_i. \qquad (4.4)$$

When Γ_i occurs, we say that the type i particle *jumps*.

Remark 27. The definitions of lifetimes outline that our interest is focused when the particles *do something* and not when they survive.

Like the usual defintions of the lifetimes, we suppose the *memoryless property* for Γ_1, Γ_2 and so we know, from Section 1.1.2, that they are exponentially distributed, with parameters $a_1, a_2 > 0$ respectively.

Finally, the offspring distribution of the particles

$$\boldsymbol{p}\left(\boldsymbol{j}\right) = \left(p^{(1)}\left(j_1, j_2\right), p^{(2)}\left(j_1, j_2\right)\right),$$

are the production probabilities of each particles when the lifetimes Γ_1, Γ_2 occur, i.e. the probability that a type i particles jumps and produces j_1 type 1 particles and j_2 type 2 particles.

Remark 28. Note that from definitions of lifetimes Γ_1, Γ_2, we get $p^{(i)}\left(\boldsymbol{e}_i\right) = 0$, for every $i = 1, 2$.

4.2.2 The Infinitesimal Marginal and Joint Distributions

The idea is to consider the marginal distributions seen in the Sections 4.1.1 and 4.1.2 not in a unit time interval but in the time interval Δt.

Thus, we replace the probabilities p_i, $i = 0, \ldots, 3$ and q_i, $i = 0, \ldots, 2$, of Section 4.1.1 and 4.1.2, like in Table 4.3. Moreover, in this model we assume that a CA operates in the same way on the tips and on the branches, i.e. we get

$$q_0 = p_0.$$

Probability in $[n, n+1]$	Probability in $[t, t + \Delta t]$
p_0	$p_0 \cdot \Delta t + o\left(\Delta t\right)$
p_1	$p_1 \cdot \Delta t + o\left(\Delta t\right)$
p_2	$p_2 \cdot \Delta t + o\left(\Delta t\right)$
p_3	$1 - \left(p_0 + p_1 + p_2\right)\Delta t + o\left(\Delta t\right)$
q_0	$p_0 \cdot \Delta t + o\left(\Delta t\right)$
q_1	$1 - \left(p_0 + q_2\right)\cdot \Delta t + o\left(\Delta t\right)$
q_2	$q_2 \cdot \Delta t + o\left(\Delta t\right)$

Table 4.3: Marginal distribution in $[t, t + \Delta t]$

Now, according to Section 4.1.1, we introduce the infinitesimal marginal distribution of a tip alive at the time $t \geq 0$:

1) the tip jumps, producing no particles (the blood vessel stops growing, for example by a CA, and the last part branch of the blood vessel created dies too), with probability $p_0 \cdot \Delta t + o\left(\Delta t\right)$, i.e. at time $t + \Delta t$ there are no tips or branches alive.

2) the tip jumps, growing up and creating a branch behind itself (the blood vessel grows and creates a branch), with probability $p_1 \cdot \Delta t + o\left(\Delta t\right)$, i.e. at time $t + \Delta t$ there are the survived tip and the branch created by its growth.

3) the tip jumps, splitting itself in two new tips (the blood vessel splits itself in new two others), with probability $p_2 \cdot \Delta t$, i.e. at time $t + \Delta t$ there are the two new tips created by the ancestor tip.

4) the tip does not jump and survives (the blood vessel does not grow but it is still alive), with probability $1 - \left(p_0 + p_1 + p_2\right)\Delta t + o\left(\Delta t\right)$, i.e. at the time $t + \Delta t$ there is only the survived tip.

Then, according to Section 4.1.2, we consider a branch alive at the time $t \geq 0$:

1) the branch jumps, producing no particles (an existing branch of a blood vessel dies, for instance by fibrinoid necrosis of the blood vessel induced

by chemioterapy), with probability $p_0 \cdot \Delta t$, i.e. at time $t + \Delta t$ there are no tips or branches alive,

2) the branch does not jumps (an existing branch of a blood vessel doesn't die), with probability $1 - (p_0 + q_2) \cdot \Delta t$, i.e. at time $t + \Delta t$ there is only the survived branch,

3) the branch jumps, surviving and producing a new tip from itself (an existing branch of a blood vessel doesn't die and a new one could born from it), with probability $q_2 \cdot \Delta t$, i.e. at time $t + \Delta t$ there are the survived branch and the new tip born from it.

Event	*Probability*
A tip jumps, producing no particles	$p_0 \cdot \Delta t + o(\Delta t)$
A tip jumps, growing up and producing a branch	$p_1 \cdot \Delta t + o(\Delta t)$
A tip jumps, splitting in two other tips	$p_2 \cdot \Delta t + o(\Delta t)$
A tip does not jump and survives	$1 - (p_0 + p_1 + p_2)\Delta t$ $+o(\Delta t)$

Table 4.4: Infinitesimal marginal distribution of the tips

In Tables 4.4 and 4.5 we summarize the infinitesimal marginal distributions of the two-type particles, where $p_0,\ p_2,\ p_1,\ q_2 \geq 0,\ p_0 + p_1 + p_2 < 1$ and $p_0 + q_2 < 1$.

Event	*Probability*
A branch jumps, producing no particles	$p_0 \cdot \Delta t + o(\Delta t)$
A branch does not jump, surviving	$1 - (p_0 + q_2) \cdot \Delta t + o(\Delta t)$
A branch jumps, surviving and creating a new tip from it	$q_2 \cdot \Delta t + o(\Delta t)$

Table 4.5: Infinitesimal marginal distribution of the branches

Remark 29. Note that, at limit $\Delta t \to 0$, both the probabilities of tips and branches have sum equal to 1.

Now, in accordance with the joint distribution in the discrete model shown in Section 4.1.3 (see Table 4.1), we can write the following infinitesimal joint distributions for tips and branches respectively

$$
\begin{cases}
\mathbb{P}\left((1,0),(0,0),\Delta t\right) = p_0 \cdot \Delta t + o\left(\Delta t\right) \\
\mathbb{P}\left((1,0),(1,0),\Delta t\right) = 1 - \left(p_0 + p_1 + p_2\right) \cdot \Delta t + o\left(\Delta t\right) \\
\mathbb{P}\left((1,0),(1,1),\Delta t\right) = p_1 \cdot \Delta t + o\left(\Delta t\right) \\
\mathbb{P}\left((1,0),(2,1),\Delta t\right) = o\left(\Delta t\right) \\
\mathbb{P}\left((1,0),(2,0),\Delta t\right) = p_2 \cdot \Delta t + o\left(\Delta t\right)
\end{cases}
, \qquad (4.5)
$$

and

$$
\begin{cases}
\mathbb{P}\left((0,1),(0,0),\Delta t\right) = p_0 \cdot \Delta t + o\left(\Delta t\right) \\
\mathbb{P}\left((0,1),(0,1),\Delta t\right) = 1 - \left(p_0 + q_2\right) \cdot \Delta t + o\left(\Delta t\right) \\
\mathbb{P}\left((0,1),(1,1),\Delta t\right) = q_2 \cdot \Delta t + o\left(\Delta t\right)
\end{cases}
. \qquad (4.6)
$$

Remark 30. Note that the infinitesimal joint distributions (4.5) and (4.6) represent the transition probabilities, introduced in Section 1.1.2, calculated in the time interval Δt.

4.2.3 Exponential parameters

In this Section, from the probabilities (4.5), (4.6), we are able to find out the parameters a_1 and a_2 of the exponential distribuited lifetimes $\Gamma_1\,\Gamma_2$. We start with the parameter a_2. From the definition (4.4), we know that

$$
\mathbb{P}\left(\Gamma_2 \leq T\right) = \int_0^T a_2 e^{-a_2 \cdot t} dt = 1 - e^{-a_2 \cdot T}, \ \forall\, T \geq 0.
$$

Thus, when $T = \Delta t$, the probability $\mathbb{P}\left(\Gamma_2 \leq \Delta t\right)$ is equal to the right term in the equation above written in Taylor series, and we get

$$
\mathbb{P}\left(\Gamma_2 \leq \Delta t\right) = a_2 \cdot \Delta t + o\left(\Delta t\right). \qquad (4.7)
$$

From the probabilities (4.6), we have another way to exprime $\mathbb{P}\left(\Gamma_2 \leq \Delta t\right)$, indeed in the interval $t + \Delta t$ a branch can change the state of the process in two ways, when it produces no paticles and when it produces two. So we get

$$\mathbb{P}\left(\Gamma_2 \leq \Delta t\right) = 1 - \mathbb{P}\left(\left(0,1\right), \left(0,1\right), \Delta t\right) =$$

$$= \left(p_0 + q_2\right) \Delta t + o\left(\Delta t\right). \tag{4.8}$$

Thus, from (4.7) and (4.8), we obtain the value of a_2,

$$a_2 \Delta t + o\left(\Delta t\right) = \left(p_0 + q_2\right) \Delta t + o\left(\Delta t\right) \Rightarrow$$

$$\Rightarrow a_2 + \frac{o\left(\Delta t\right)}{\Delta t} = p_0 + q_2 + \frac{o\left(\Delta t\right)}{\Delta t} \Rightarrow$$

$$\Rightarrow a_2 = p_0 + q_2, \text{ for } \Delta t \to 0.$$

Now consider the parameter a_1. From the definition of Γ_1, we have that

$$\mathbb{P}\left(\Gamma_1 \leq \Delta t\right) = a_1 \Delta t + o\left(\Delta t\right).$$

According to (4.5), we have that and we get

$$\mathbb{P}\left(\Gamma_1 \leq dt\right) = 1 - \mathbb{P}\left(\left(1,0\right), \left(1,0\right), \Delta t\right) =$$

$$= \left(p_0 + p_1 + p_2\right) \Delta t + o\left(\Delta t\right),$$

and so in accordance to a_2, we obtain that

$$a_1 = p_0 + p_1 + p_2.$$

4.2.4 Inverse Technique for the Mean Matrix $M\left(t\right)$

From (1.8) and (1.9) in Section 1.1.2, we know that

$$\boldsymbol{M}\left(t\right) := e^{t \cdot \boldsymbol{A}} = \sum_{i \geq 0} \frac{\left(t \cdot \boldsymbol{A}\right)^i}{i!} = \boldsymbol{I} + t \cdot \boldsymbol{A} + \boldsymbol{o}\left(t\right), \quad \forall t \geq 0,$$

where the elements a_{ij} of the matrix $\boldsymbol{A}$ are defined in the following way

$$a_{ij} := a_i \left(\left. \frac{\partial f^{(i)}(s_1, s_2)}{\partial s_j} \right|_{s_1=1, s_2=1} - \delta_{ij} \right), i, j = 1, 2,$$

and

$$\boldsymbol{o}(t) = \begin{pmatrix} o_{11}(t) & o_{12}(t) \\ o_{21}(t) & o_{22}(t) \end{pmatrix}.$$

In this section, using a technique of inverse problem and from (4.5), (4.6) and parameters a_1, a_2, we show that is possible to compute the mean matrix $\boldsymbol{M}(t)$, for every $t \geq 0$, without the generating functions $f^{(1)}(s_1, s_2)$, $f^{(2)}(s_1, s_2)$.

From the infinitesimal transition probabilities (4.6) and (4.5) we can write the following expected values

$$\begin{cases} \mathbb{E}[Z_1^{(1)}(\Delta t)] = p_1 \Delta t + 1 - (p_0 + p_1 + p_2) \Delta t + 2p_2 \Delta t + o(\Delta t) = \\ \qquad = 1 - (p_0 - p_2) \Delta t + o(\Delta t) \\ \mathbb{E}[Z_2^{(1)}(\Delta t)] = p_1 \Delta t + o(\Delta t) \\ \mathbb{E}[Z_1^{(2)}(\Delta t)] = q_2 \Delta t + o(\Delta t) \\ \mathbb{E}[Z_2^{(2)}(\Delta t)] = 1 - (p_0 + q_2) \Delta t + q_2 \Delta t + o(\Delta t) = \\ \qquad = 1 - p_0 \Delta t + o(\Delta t) \end{cases}, \quad (4.9)$$

and so, from (4.9) and (1.7) (see Section 1.1.2), we can write

$$\boldsymbol{M}(\Delta t) = \begin{pmatrix} 1 - (p_0 - p_2) \Delta t + o(\Delta t) & p_1 \Delta t + o(\Delta t) \\ q_2 \Delta t + o(\Delta t) & 1 - p_0 \Delta t + o(\Delta t) \end{pmatrix}. \quad (4.10)$$

From (1.8) in Section 1.1.2, we know that exists a matrix $\boldsymbol{A}$ such that

$$\boldsymbol{A} = \lim_{\Delta t \to 0} \frac{\boldsymbol{M}(\Delta t) - \boldsymbol{I}}{\Delta t}, \quad (4.11)$$

and so, from (4.10) and (4.11), we get the matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} -(p_0 - p_2) & p_1 \\ q_2 & -p_0 \end{pmatrix}. \tag{4.12}$$

Finally, from (1.8) and (4.12), we can write the mean matrix $\boldsymbol{M}(t)$

$$\boldsymbol{M}(t) = e^{\frac{1}{2}(-2p_0 + p_2)t} \left(\boldsymbol{H}(t) + \boldsymbol{E}(t) \right),$$

where

$$\boldsymbol{H}(t) = \cosh \left(\frac{t}{2} \sqrt{p_2^2 + 4p_1 q_2} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} +$$

$$+ \frac{p_2 \sinh \left(\frac{t}{2} \sqrt{p_2^2 + 4p_1 q_2} \right)}{\sqrt{p_2^2 + 4p_1 q_2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and

$$\boldsymbol{E}(t) = \frac{e^{-\frac{t}{2}\sqrt{p_2^2 + 4p_1 q_2}}}{\sqrt{p_2^2 + 4p_1 q_2}} \begin{pmatrix} 0 & p_1 \left(-1 + e^{t\sqrt{p_2^2 + 4p_1 q_2}} \right) \\ q_2 \left(-1 + e^{t\sqrt{p_2^2 + 4p_1 q_2}} \right) & 0 \end{pmatrix}.$$

Remark 31. Recall that the mean matrix $\boldsymbol{M}(t)$, for every $t \geq 0$, as the matrix $\boldsymbol{M}^n$ in the discrete time case (see the equation (1.3) in Section 1.1.1), is such that

$$\mathbb{E}\left[\boldsymbol{Z}(t) \,|\, \boldsymbol{Z}(0) \right] = \boldsymbol{Z}(0) \cdot \boldsymbol{M}(t).$$

4.2.5   Theoretical Classification and Growth-Speed of Tips

Now, from Section 1.1.2, we know that the matrix $\boldsymbol{A}$ gives basic information about the process $\boldsymbol{Z}(t)$. Indeed, denoting with λ the maximum eigenvalue of $\boldsymbol{A}$, we have the following theoretical classification

- $\lambda < 0 \Rightarrow$ the process is SubCritical, i.e. exists a finite and strictly positive t such that $\boldsymbol{Z}(t) = \boldsymbol{0}$ a.s. ,

- $\lambda = 0 \Rightarrow$ the process is Critical, i.e. exists a finite and strictly positive t such that $\boldsymbol{Z}(t) = \boldsymbol{0}$ a.s. ,

- $\lambda > 0 \Rightarrow$ the process is SuperCritical, i.e. with a positive probability exists a finite and strictly positive t such that $\boldsymbol{Z}(t) = \boldsymbol{0}$ a.s. .

In our model we have that

$$\lambda = \frac{1}{2}\left(-2p_0 + p_2 + \sqrt{p_2^2 + 4q_2p_1}\right),$$

and we obtain the following theoretical classification of the process $\boldsymbol{Z}(t)$,

$$p_0 \le \frac{p_2}{2} \text{ or } \begin{cases} p_0 > \dfrac{p_2}{2} \\ p_0^2 < q_2p_1 + p_0p_2 \end{cases} \Rightarrow \text{SuperCritical}$$

$$\begin{cases} p_0 > \dfrac{p_2}{2} \\ p_0^2 = q_2p_1 + p_0p_2 \end{cases} \Rightarrow \text{Critical}$$

$$\begin{cases} p_0 > \dfrac{p_2}{2} \\ p_0^2 > q_2p_1 + p_0p_2 \end{cases} \Rightarrow \text{SubCritical}$$

Finally, let us make a consideration about the growth-speed of the tips. If we define Δx the lenght of the branches created by the growth of the tips in the interval $[t, t + \Delta t]$, $\forall t \ge 0$, then we have that the quantity

$$\mathbb{E}[Z_2^{(1)}(\Delta t)] \cdot \Delta x$$

is the average lenght of branches created by a tip in $[t, t + \Delta t]$, and so it can be considered the average growth-speed of a tip in $[t, t + \Delta t]$. Thus, we can define the following average growth-speed of tips

$$\overline{v} = \frac{\mathbb{E}[Z_2^{(1)}(\Delta t)] \cdot \Delta x}{\Delta t} = \frac{(p_1\Delta t + o(\Delta t)) \cdot \Delta x}{\Delta t} =$$

$$= p_1 \Delta x + \Delta x \cdot \frac{o\left(\Delta t\right)}{\Delta t}.$$

So, the related instantaneous growth-speed of tips v is

$$v = \lim_{\Delta t \to 0} \overline{v} = p_1 \Delta x.$$

4.2.6 Offspring Distribution, Generating Functions and Extinction Probability

Let be fixed the matrix $\boldsymbol{A}$ and the parameters a_1, a_2, so exist different offspring distributions $p^{(1)}\left(j_1, j_2\right)$, $p^{(2)}\left(j_1, j_2\right)$ that verify the condition (1.9) (see Section 1.1.1) and the Remark 28, for instance the distributions

$$\begin{cases} p^{(1)}\left(0, 0\right) = p \\\\ p^{(1)}\left(2, 0\right) = \dfrac{p_1 + 2p_2}{2\left(p_0 + p_1 + p_2\right)} \\\\ p^{(1)}\left(0, 1\right) = \dfrac{p_1}{p_0 + p_1 + p_2} \end{cases} , \quad \begin{cases} p^{(2)}\left(0, 0\right) = q \\\\ p^{(2)}\left(1, 0\right) = \dfrac{q_2}{p_0 + q_2} \\\\ p^{(2)}\left(0, 2\right) = \dfrac{q_2}{2\left(p_0 + q_2\right)} \end{cases} ,$$

for every p, $q \in [0, 1]$, produces the following generating functions

$$\begin{cases} f^{(1)}\left(s_1, s_2\right) = p + \dfrac{p_1 + 2p_2}{2\left(p_0 + p_1 + p_2\right)} \cdot s_1^2 + \dfrac{p_1}{p_0 + p_1 + p_2} \cdot s_2 \\\\ f^{(2)}\left(s_1, s_2\right) = q + \dfrac{q_2}{p_0 + q_2} \cdot s_1 + \dfrac{q_2}{2\left(p_0 + q_2\right)} \cdot s_2^2 \end{cases} ,$$

that verify (1.9).

For this model, we propose the most natural offspring distribution, i.e. for $p^{(1)}\left(j_1, j_2\right)$ we choose the probabilities (4.5) where $o\left(\Delta t\right)$ is omitted and Δt is replaced by

$$\mathbb{E}\left[\Gamma_1\right] = \frac{1}{a_1} = \frac{1}{p_0 + p_1 + p_2},$$

and for $p^{(2)}(j_1, j_2)$ we choose the probabilities (4.5) where $o(\Delta t)$ is omitted and Δt is replaced by

$$\mathbb{E}[\Gamma_2] = \frac{1}{a_2} = \frac{1}{p_0 + q_2}.$$

Thus, we obtain

$$\begin{cases} p^{(1)}(0,0) = \dfrac{p_0}{p_0 + p_1 + p_2} \\[2mm] p^{(1)}(1,1) = \dfrac{p_1}{p_0 + p_1 + p_2} \\[2mm] p^{(1)}(2,0) = \dfrac{p_2}{p_0 + p_1 + p_2} \end{cases}, \quad \begin{cases} p^{(2)}(0,0) = \dfrac{p_0}{p_0 + q_2} \\[4mm] p^{(2)}(1,1) = \dfrac{q_2}{p_0 + q_2} \end{cases}. \qquad (4.13)$$

Remark 32. Note that, according to the Remark 28, we have that

$$p^{(2)}(1,0) = p^{(1)}(1,0) = 0.$$

Remark 33. Note that the offspring distributions (4.13) and probabilites (4.5), (4.6) have the same meaning of the probabilities seen in Table 4.2.

Now, we prove that the offspring distributions (4.13) verify the equation (1.9). Let $f^{(1)}(s_1, s_2)$ and $f^{(2)}(s_1, s_2)$ be the related generating finctions and we get

$$f^{(1)}(s_1, s_2) = \sum_{j_1, j_2 \geq 0} p^{(1)}(j_1, j_2) s_1^{j_1} s_2^{j_2} =$$

$$= \frac{p_0}{p_0 + p_1 + p_2} + \frac{p_1}{p_0 + p_1 + p_2} \cdot s_1 s_2 + \frac{p_2}{p_0 + p_1 + p_2} \cdot s_1^2, \qquad (4.14)$$

and

$$f^{(2)}(s_1, s_2) = \sum_{j_1, j_2 \geq 0} p^{(2)}(j_1, j_2) s_1^{j_1} s_2^{j_2} =$$

$$= \frac{p_0}{p_0 + q_2} + \frac{q_2}{p_0 + q_2} \cdot s_1 s_2. \qquad (4.15)$$

Now, we see that applying (1.9) to the generating functions (4.14), (4.15),

we get the elements of the matrix $\boldsymbol{A}$ in (4.12). Thus,

- $a_1 \left(\left. \dfrac{\partial f^{(1)}(s_1, s_2)}{\partial s_1} \right|_{s_1=1, s_2=1} - 1 \right) =$

$$= (p_0 + p_1 + p_2) \left(\frac{p_1}{p_0 + p_1 + p_2} + \frac{2p_2}{p_0 + p_1 + p_2} - 1 \right) =$$

$$= p_1 + 2p_2 - p_0 - p_1 - p_2 = -p_0 + p_2 = a_{11},$$

- $a_1 \left(\left. \dfrac{\partial f^{(1)}(s_1, s_2)}{\partial s_2} \right|_{s_1=1, s_2=1} \right) = (p_0 + p_1 + p_2) \left(\frac{p_1}{p_0 + p_1 + p_2} \right) =$

$$= p_1 = a_{12},$$

- $a_2 \left(\left. \dfrac{\partial f^{(2)}(s_1, s_2)}{\partial s_1} \right|_{s_1=1, s_2=1} \right) = (p_0 + q_2) \left(\frac{q_2}{p_0 + q_2} \right) = q_2 = a_{21},$

- $a_2 \left(\left. \dfrac{\partial f^{(2)}(s_1, s_2)}{\partial s_2} \right|_{s_1=1, s_2=1} - 1 \right) = (p_0 + q_2) \left(\frac{q_2}{p_0 + q_2} - 1 \right) =$

$$= (p_0 + q_2) \left(\frac{q_2 - p_0 - q_2}{p_0 + q_2} \right) = -p_0 = a_{22}.$$

Chapter 5

μ-GW trees and scaling model

In this Chapter we consider, as in Section 4.1, the phenomenon of the angiogenesis as a system of two type particles that evolve according to a multitype Galton-Watson process with discrete time. Moreover, we add a genealogical structure to the branching process and so we consider a μ-GW tree, in particular we treat the special cases of the full binary trees and the full binary trees with survivals (see for more details Section 2.3).

In Section 5.1 we apply the theoretical results about the μ-GW tress seen in Section 2.1, then we consider also the results about the full binary trees, exposed in Section 2.3. From these results, we show important considerations about the a.s. *finished* and the *still alive* angiogenesis processes, rapresented by the full binary trees with survivals and and the full binary trees respectively.

In Section 5.2 we propose a scaling model for angiogenesis, such that we can apply the result of weak convergence seen in Section 2.2. In particular, under the a.s. finite condition, the total period of contour processes related to a system of n *accelerated* angiogenesis processes (described by μ_n-GW trees) has a normal distribution at the limit $n \to \infty$, i.e. we have informations on the *finiteness* of the processes for $n \to \infty$.

5.1 Full binary trees and full binary trees with survivals

In this Section we model the biological process of the angiogenesis as a $\boldsymbol{\mu}$-GW tree, defined in Section 1.2.2, with the same two-type particles introduced for the discrete time model in Section 4.1 of the Chapter 4, i.e. the tips and the branches of the blood vessels. The offspring distribution $\boldsymbol{\mu}$ is the *binarization* of (4.1) in Section 4.1.3, that is, using the notation $\boldsymbol{\mu}$ for the offspring distribution $p^{(i)}(j_1, j_2)$ for every $i = 1, 2$, we simplify the production of particles (4.1) by not considering $p^{(1)}(2, 1)$ and $p^{(1)}(2, 0)$. Moreover, instead of the discrete time model, here we give directly the offspring distribution $\boldsymbol{\mu}$ of the two-types particles without passing by the marginal distributions, and so we have the following offspring distribution

$$
\begin{cases}
\mu^{(1)}(0,0) = p_0 \\
\mu^{(1)}(1,0) = p_1 \quad \text{and} \\
\mu^{(1)}(1,1) = p_2
\end{cases}
\begin{cases}
\mu^{(2)}(0,0) = q_0 \\
\mu^{(2)}(0,1) = q_1 \quad , \\
\mu^{(2)}(1,1) = q_2
\end{cases}
\tag{5.1}
$$

with p_i and $q_i \in [0, 1]$, $i = 0, 1, 2$ and $\sum_i p_i = \sum_i q_i = 1$. In other words in this model we will consider the angiogenesis process as a full binary tree with survivals (see Section 2.3).

Remark 34. Before introducing the model, we recall that the offspring distribution $\boldsymbol{\mu}$, in the model of the angiogenesis, has the same meaning of the probabilities in Table 4.2 of Section 4.1.

In this Section we recall all the definitions and results about the $\boldsymbol{\mu}$-GW trees and the full binary trees with survivals seen in Section 1.2. In particular, remember that the $\boldsymbol{\mu}$-GW trees are equipped with the type-ordering offspring (see Defintion 2.1.1) and $2L^{(i)}(\varnothing)$ is the non-negative integer-valued r.v. representing the period of the contour process $\mathscr{C}_\tau(s)$ associated to a full binary tree τ starting with a type i particle, for every $i = 1, 2$. Recall also that in Corollary 2 of Section 2.3.1 we have proved the following theoretical

classification of a binary tree with survivals τ,

$$p_0 q_0 - p_2 q_2 \geq 0 \qquad \Longleftrightarrow \qquad \tau \text{ is finite } a.s. \qquad (5.2)$$

5.1.1 Estimations of Biological Mass and Crossroads of an Angiogenesis Process

In (5.2) we have recalled that if the product of the probabilites $\mu^{(1)}(0,0)$ and $\mu^{(2)}(0,0)$ is more or equal to the product of $\mu^{(1)}(1,1)$ and $\mu^{(2)}(1,1)$ then the process of angiogenesis starting with a new blood vessel (tip) or with a branch by an existing blood vessel (branch) will stop in a finite time interval a.s. . Otherwise, the process of angiogenesis will stop in a finite time interval only with a probability striclky less than 1.

Remark 35. Note that the growth time of an angiogenesis process is equal to the generations of the full binary tree with survivals corresponding to the biological process.

Now, in the case of the a.s. finiteness of the angiogenesis process, suppose that η is the time of extinction of the process (i.e. the last generation of the tree), then we want to know which are the most representative quantity of the process. From a picture of a finished process of growth of blood vessels, the first two natural features that you can note are the total biological mass produced (BM_η) and the number of crossroads of bloos vessels (VC_η). The total biological mass produced during the process is clearly composed of all the blood vessels produced and it may represent the dimension of the hypothetical tumor and so also the damege caused by it. The number of the crossroads of the blood vessels can be used to study the diffusion way of the process, for example if this number is sufficiently small we know that only a few blood vessels have grown and so a chemotherapic action may be more powerful than if there are a large number of crossroads of blood vessels. In fact, a joint knowledge of these two quantities can be used to distinguish angiogenesis processes like those in the Figure 5.1.

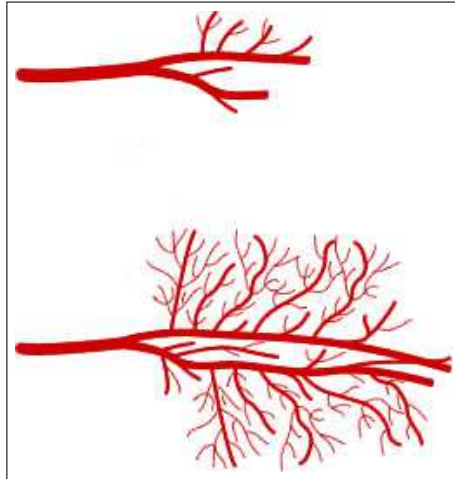In this model, a new branch of a blood vessel is produced only with the

Figure 5.1: Example of two angiogenesis processes that can be distinguished by the total biological mass produced and the number of the crossroads of the blood vessels.

profileration probability $\mu^{(1)}(1,1)$, corresponding to the tips that have exactly two children and so the total biological mass produced is equal to the a.s. finite number of type 1 vertices with two children. On the other hand, each crossroad of blood vessels corresponds to the *profileration* probability $\mu^{(2)}(1,1)$ and so the number of the crossroads of blood vessels is the a.s. finite number of type 2 vertices with exactly two children.

Here we suppose that the process of the angiogenesis starts from a tip and that at least one tip has two children (otherwise no blood vessel will be created) and so, if we denote with D_1^η and D_2^η the r.v. that represent the number of type 1 and type 2 vertices with two children respectivley at the time η, the likelihood (2.19) seen in Section 2.3.1 is exactly the joint distribution of the total biological mass produced and the number of crossroads of the blood vessels, and so we get

$$\mathbb{P}\left(D_1^\eta = n, D_2^\eta = m\right) = N\left(n+m, m+1\right) \cdot P^{m+1}\left(1-P\right)^n Q^n \left(1-Q\right)^m.$$

where $n \geq 1$, $m \geq 0$ and

$$P = \frac{p_0}{p_0 + p_2} \quad \text{and} \quad Q = \frac{q_0}{q_0 + q_2},$$

and $N(\cdot, \cdot)$ is a Narayana number (see (2.13) in Section 2.3).

Each tip generates a single tip and we may say that it remains in its own state with probability p_1. Eventually, it will change "its state" $(1, 0)$ by dying $(0, 0)$ or by creating a branch $(1, 1)$. Conditioned to this event, P is the probability that a tip dies. The same holds for the branches and so Q is the probability that a branch dies, conditioned to the change of its own state.

We set $K > 0$ the length of the branch created by the growth of a tip, then we have that

$$BM_\eta = K \cdot D_1^\eta \text{ and } VC_\eta = D_2^\eta.$$

From the theory of the Maximum-Likelihood Estimation (MLE) and the equations (2.20) in Section 2.3.1, if we know the values of BM_η and VC_η of an a.s. finite process of angiogenesis modeled with a full binary tree with survivals, then the maximum-likelihood estimators of the parameter P and Q are the following quantitites

$$\widehat{P} = \frac{VC_\eta + 1}{VC_\eta + 1 + \dfrac{BM_\eta}{K}} = \frac{K \cdot (VC_\eta + 1)}{K \cdot (VC_\eta + 1) + BM_\eta},$$

and

$$\widehat{Q} = \frac{\dfrac{BM_\eta}{K}}{VC_\eta + \dfrac{BM_\eta}{K}} = \frac{BM_\eta}{K \cdot VC_\eta + BM_\eta}.$$

5.1.2 Estimations on still alive Full Binary Trees

Now we want to consider the case in which we investigate the growth of the angiogenesis process in a generic time M. To further simplify this study, we consider only the process of angiogenesis modeled on the full binary trees,

i.e. $\mu^{(1)}(1,0) = \mu^{(2)}(0,1) = 0$. From the result (2.29) in Section 2.3.3, we can compute a lower and an upper estimation of the probability to have a process of angiogenesis, starting from a new blood vessel, in a generic time M, denoted with $\mathbb{P}_M$.

$$\sum_{k \geq 2^{M-1}} p_0 p_2^k q_0^k \left({}_2F_1 \left(1-k, -k; 2; \frac{p_0 q_2}{p_2 q_0} \right) \right)$$

$$\leq \mathbb{P}_M \leq \sum_{k \geq M} p_0 p_2^k q_0^k \left({}_2F_1 \left(1-k, -k; 2; \frac{p_0 q_2}{p_2 q_0} \right) \right).$$

where

$$_2F_1(a, b; c; d) = \sum_{n \geq 0} \frac{(a)_n (b)_n}{(c)_n} \cdot \frac{z^n}{n!},$$

with

$$(q)_n = \begin{cases} 1, & n = 0 \\ q(q+1)\cdots(q+n-1), & n > 0 \end{cases},$$

is the hypergeometric function, defined for $|z| < 1$.

In the case fo the full binary trees the condition $|z| < 1$ is equal to $p_2 > q_2$, indeed we have

$$\frac{p_0 q_2}{p_2 q_0} < 1 \Rightarrow p_0 q_2 < p_2 q_0 \Rightarrow (1 - p_2) q_2 < (1 - q_2) p_2 \Rightarrow q_2 < p_2.$$

Rememebr that the a.s. finiteness condition for the full binary trees with survivals $p_0 q_0 - p_2 q_2 \geq 0$ is more simple in the case without survivals, indeed it becomes

$$0 \leq p_0 q_0 - p_2 q_2 = (1 - p_2)(1 - q_2) - p_2 q_2 = 1 - p_2 - q_2 + p_2 q_2 - p_2 q_2$$

$$\implies \quad p_2 + q_2 \leq 1.$$

So, by combining the inequalities $q_2 < p_2$ and $p_2 + q_2 \leq 1$, we obtain the final parameters condition to get the upper and lower estimation of the probability

that a process of angiogenesis is growing at the time M,

$$\begin{cases} p_2 > q_2, & 0 < p_2 < 1/2 \\ p_2 \leq q_0, & 1/2 \leq p_2 \leq 1 \end{cases},$$

where $p_0 + p_2 = 1$, $q_0 + q_2 = 1$.

5.2   Scaling Model

Here, starting with the model of the full binary trees with survivals seen in the previous Section, we consider the scaling exposed in Section 1.3.2. In particular at the n-th step of the scaling each particles has "mass" equal to $1/n$ and they produces other particles, according to an offspring distribution $\boldsymbol{\mu}_n$, at time intervals equal to $1/n$. Because for us an agiogenesis process starts with a new blood vessel (i.e. from a tip), from the Remark 13 in Section 1.3.2, we know that a the n-th step of the scaling, for $t = 0$, we have n tips (with "mass" equal to $1/n$). The scaled offspring distribution $\boldsymbol{\mu}_n$ we use is the following

$$\begin{cases} \mu_n^{(1)}(0,0) = p_0^n = \frac{(1-p)W_1}{n^\alpha} \\ \mu_n^{(1)}(1,0) = p_1^n = 1 - \frac{W_1}{n^\alpha} \\ \mu_n^{(1)}(1,1) = p_2^n = \frac{pW_1}{n^\alpha} \end{cases} \text{and} \begin{cases} \mu_n^{(2)}(0,0) = q_0^n = \frac{(1-q)W_2}{n^\alpha} \\ \mu_n^{(2)}(0,1) = q_1^n = 1 - \frac{W_2}{n^\alpha} \\ \mu_n^{(2)}(1,1) = q_2^n = \frac{qW_2}{n^\alpha} \end{cases}, \quad (5.3)$$

where $p, q \in (0,1)$, $W_1, W_2 \in (0,1)$ and $\alpha \geq 1$, $n \geq 1$.

Before giving the results of this model, we explain the reason of such a choice of $\boldsymbol{\mu}_n$ and the meaning of the parameters p, q, W_1, W_2 and α. At first we have searched a scaled offspring ditribution such that the conditions (1.15) of Section 1.3.2 hold, in such a way we know that the random measures defined in (1.16) converge to a superprocess.

The offspring distribution we have considered is the following

$$\begin{cases} \mu_n^{(1)}(0,0) = \frac{(1-p)W_1}{n} \\ \mu_n^{(1)}(1,0) = 1 - \frac{W_1}{n} \\ \mu_n^{(1)}(1,1) = \frac{pW_1}{n} \end{cases} \quad \text{and} \quad \begin{cases} \mu_n^{(2)}(0,0) = \frac{(1-q)W_2}{n} \\ \mu_n^{(2)}(0,1) = 1 - \frac{W_2}{n} \\ \mu_n^{(2)}(1,1) = \frac{qW_2}{n} \end{cases}, \qquad (5.4)$$

where $p, q \in (0,1)$, $W_1, W_2 \in (0,1)$ and $n \geq 1$.

Remark 36. The offspring distributions (5.3) and (5.4), in the angiogenesis process, have the same meaning of the probabilities seen in Tables 4.2 of the Section 4.1 in Chapter 4.

We have defined the probabilities $\mu_n^{(1)}(1,0)$ and $\mu_n^{(2)}(0,1)$ to approximate at the limit $n \to \infty$ a time continuous process, indeed for every $i = 1, 2$, we have that

$$\left(1 - \frac{W_i}{n}\right)^n$$

is the probability that a type i particle survives in the unit interval time according to the offspring distribution (5.4), and so at the limit $n \to \infty$ we obtain

$$\lim_{n \to \infty} \left(1 - \frac{W_i}{n}\right)^n = e^{-W_i}.$$

Now, remembering the defintions (4.4) of the *lifetimes* Γ_1, Γ_2, if we consider W_1 and W_2 the exponential parameters of the distributions of Γ_1, Γ_2, we have that the probability that a type i particle survives in the unit interval time according to the defintions (4.4) is equal to

$$\mathbb{P}(\Gamma_i \geq 1) = \int_1^\infty W_i \cdot e^{-W_i t} dt = e^{-W_i}.$$

Thus, the paramters W_1, W_2 can be defined as the *survival parameters*, and moreover we have the following estimations of the probability that a type i particles, at the limit $n \to \infty$, may survive in a unit time interval, for every $i = 1, 2$,

$$e^{-W_i} \xrightarrow[W_i \to 0^+]{} 1,$$

and

$$e^{-W_i} \xrightarrow[W_i \to 1^-]{} \frac{1}{e} \sim 0.37.$$

From the biological point of view, the parameters W_1 and W_2 indicate the strength of the angiogenesis process to the chemoterapic actions. Moreover, p, q can be seen has a sort of parameters that estimate the probability of *death-proliferation* of the process.

Now, using the notation $h = n$ and $p_n^{(i)}(\boldsymbol{j}) = \mu_n^{(i)}(\boldsymbol{j})$, we check that the offspring distribution (5.4) satisfies the conditions (1.15),

$$m_1^{n,(1)}(1) = \mu_n^{(1)}(1,0) + \mu_n^{(1)}(1,1) = 1 + \frac{pW_1 - W_1}{n},$$

$$m_1^{n,(1)}(2) = \mu_n^{(1)}(1,1) = \frac{pW_1}{n},$$

$$m_2^{n,(1)}(1) = \mu_n^{(2)}(1,1) = \frac{qW_2}{n},$$

$$m_2^{n,(1)}(2) = \mu_n^{(2)}(0,1) + \mu_n^{(2)}(1,1) = 1 + \frac{qW_2 - W_2}{n},$$

and

$$\lim_{n\to\infty} m_1^{n,(2)}(1,1) = 0,$$

$$\lim_{n\to\infty} m_1^{n,(2)}(1,2) = \lim_{n\to\infty} \mu^{(1)}(1,1) = \lim_{n\to\infty} \frac{pW_1}{n} = 0,$$

$$\lim_{n\to\infty} m_1^{n,(2)}(2,1) = \lim_{n\to\infty} \mu^{(1)}(1,1) = \lim_{n\to\infty} \frac{pW_1}{n} = 0,$$

$$\lim_{n\to\infty} m_1^{n,(2)}(2,2) = 0,$$

$$\lim_{n\to\infty} m_2^{n,(2)}(1,1) = 0,$$

$$\lim_{n\to\infty} m_2^{n,(2)}(1,2) = \lim_{n\to\infty} \mu^{(2)}(1,1) = \lim_{n\to\infty} \frac{qW_2}{n} = 0,$$

$$\lim_{n\to\infty} m_2^{n,(2)}(2,1) = \lim_{n\to\infty} \mu^{(2)}(1,1) = \lim_{n\to\infty} \frac{qW_2}{n} = 0,$$

$$\lim_{n\to\infty} m_2^{n,(2)}(2,2) = 0,$$

and finally it's easy to see that

$$m_1^{n,(3)}(k,l,h) = m_2^{n,(3)}(k,l,h) = 0, \qquad \forall\, k,l,n = 1,2.$$

Then, the offspring distribution $\boldsymbol{\mu}_n$ in (5.3), that we shall call α-*scaled off-spring distribution*, is a generalization of the (5.4). A priori, we don't know anything about a convergence result to the superprocesses (maybe, even the conditions (1.15) have to be generalized), but we can prove here that $\boldsymbol{\mu}_n$ satisfies the assumptions of the Theorem 2.2.1 of Section 2.2. Indeed, in the next lemma we see that the hypothesis 1 of the Theorem (2.2.1) holds.

Lemma 5.2.1. *Let* $X_n = 2L_n^{(1)}(\varnothing)$ *and* $Y_n = 2L_n^{(2)}(\varnothing)$ *be respectively the period of the contour process of a full binary tree at the n-th step of the scaling and let* $\boldsymbol{\mu}_n$ *defined in (5.3) be the offspring distribution. If* $p_0^n q_0^n > p_2^n q_2^n$ *then*

$$\mathbb{E}[X_n] = 2\left(\frac{pW_1 + (1-q)W_2}{(1-p-q)W_1W_2}n^\alpha - 1\right)$$
$$\mathbb{E}[Y_n] = 2\left(\frac{qW_2 + (1-p)W_1}{(1-p-q)W_1W_2}n^\alpha - 1\right) \tag{5.5}$$

and

$$Var[X_n] = K_1 \cdot n^{2\alpha} + K_2 \cdot n^\alpha$$
$$Var[Y_n] = H_1 \cdot n^{2\alpha} + H_2 \cdot n^\alpha$$

where

$$K_1 = \frac{-4p\left(2 - 2q - 3p + 3pq + p^2\right)W_1^2 + 8\left(q-1\right)\left(1+q-p\right)pW_1W_2}{(p+q-1)^3 W_1^2 W_2^2}$$
$$+ \frac{4\left(q-1\right)\left((q-1)^2 + (3q-1)p\right)W_2^2}{(p+q-1)^3 W_1^2 W_2^2} > 0, \tag{5.6}$$
$$K_2 = \frac{4\left(pW_1 + W_2 - qW_2\right)}{(-1+p+q)W_1W_2}, \tag{5.7}$$
$$H_1 = \frac{-4q\left(2 - 2p - 3q + 3pq + q^2\right)W_2^2 + 8\left(p-1\right)\left(1+p-q\right)qW_1W_2}{(p+q-1)^3 W_1^2 W_2^2} +$$
$$+ \frac{4\left(p-1\right)\left((p-1)^2 + (3p-1)q\right)W_1^2}{(p+q-1)^3 W_1^2 W_2^2} > 0,$$

$$H_2 = \frac{4\left(qW_2 + W_1 - pW_1\right)}{\left(-1 + p + q\right)W_1 W_2}.$$

Proof. At first, note that we have

$$p_0^n q_0^n > p_2^n q_2^n \implies \frac{(1-p)\,W_1}{n^\alpha} \cdot \frac{(1-q)\,W_2}{n^\alpha} > \frac{pW_1}{n^\alpha} \cdot \frac{qW_2}{n^\alpha}$$
$$\implies (1-p)\,(1-q) > pq$$
$$\implies 1 - p - q > 0.$$

Define $F_{1,n}\left(s\right) = \mathbb{E}\left[e^{s \cdot X_n}\right]$ and $F_{2,n}\left(s\right) = \mathbb{E}\left[e^{s \cdot Y_n}\right]$ respectively the moment-generating functions of $2L_n^{(1)}\left(\varnothing\right)$ and $2L_n^{(2)}\left(\varnothing\right)$. Thus, from the Theorem 2.1.3 we have that

$$\begin{cases} F_{1,n}\left(s\right) = p_0^n + p_1^n \cdot e^{2s} F_{1,n}\left(s\right) + p_2^n \cdot e^{4s} \cdot F_{1,n}\left(s\right) \cdot F_{2,n}\left(s\right) \\ F_{2,n}\left(s\right) = q_0^n + q_1^n \cdot e^{2s} F_{2,n}\left(s\right) + q_2^n \cdot e^{4s} \cdot F_{1,n}\left(s\right) \cdot F_{2,n}\left(s\right) \end{cases} \qquad (5.8)$$

Denoting
$$\frac{dF_{i,n}\left(s\right)}{ds} = F_{i,n}'\left(s\right), \ \forall\, i = 1, 2,$$

we get the following system

$$\begin{cases} F_{1,n}'\left(0\right) = p_1^n\left(2 + F_{1,n}'\left(0\right)\right) + p_2^n\left(4 + F_{1,n}'\left(0\right) + F_{2,n}'\left(0\right)\right) \\ F_{2,n}'\left(0\right) = q_1^n\left(2 + F_{2,n}'\left(0\right)\right) + q_2^n\left(4 + F_{1,n}'\left(0\right) + F_{2,n}'\left(0\right)\right) \end{cases}.$$

Then, remembering that

$$\left.\frac{dF_{i,n}\left(s\right)}{ds}\right|_{s=0} = F_{i,n}'\left(0\right) = \mathbb{E}\left[X_n\right], \ \forall\, i = 1, 2,$$

we obtain

$$\begin{cases} \mathbb{E}\left[X_n\right] = p_1^n\left(2 + \mathbb{E}\left[X_n\right]\right) + p_2^n\left(4 + \mathbb{E}\left[X_n\right] + \mathbb{E}\left[Y_n\right]\right) \\ \mathbb{E}\left[Y_n\right] = q_1^n\left(2 + \mathbb{E}\left[Y_n\right]\right) + q_2^n\left(4 + \mathbb{E}\left[X_n\right] + \mathbb{E}\left[Y_n\right]\right) \end{cases}$$

$$\implies \begin{cases} \mathbb{E}\left[X_n\right] = 2\left(\dfrac{pW_1 + (1-q)\,W_2}{(1-p-q)\,W_1W_2}n^\alpha - 1\right) \\ \mathbb{E}\left[Y_n\right] = 2\left(\dfrac{qW_2 + (1-p)\,W_1}{(1-p-q)\,W_1W_2}n^\alpha - 1\right) \end{cases} \tag{5.9}$$

Moreover, from the system (5.8) we have that

$$\begin{cases} F_{1,n}''\left(0\right) = p_1^n\left(4 + 4F_{1,n}'\left(0\right) + F_{1,n}''\left(0\right)\right) \\ \qquad + p_2^n\left(16 + 8F_{1,n}'\left(0\right) + 8F_{2,n}'\left(0\right) + 2F_{1,n}'\left(0\right)F_{2,n}'\left(0\right)\right) \\ \qquad + p_2^n\left(F_{1,n}''\left(0\right) + F_{2,n}''\left(0\right)\right) \\ F_{2,n}''\left(0\right) = q_1^n\left(4 + 4F_{2,n}'\left(0\right) + F_{2,n}''\left(0\right)\right) \\ \qquad + q_2^n\left(16 + 8F_{1,n}'\left(0\right) + 8F_{2,n}'\left(0\right) + 2F_{1,n}'\left(0\right)F_{2,n}'\left(0\right)\right) \\ \qquad + q_2^n\left(F_{1,n}''\left(0\right) + F_{2,n}''\left(0\right)\right) \end{cases}$$

and from the values (5.9) we have that

$$\begin{cases} F_{1,n}''\left(0\right) = 4 + n^\alpha \cdot C_1 + n^{2\alpha} \cdot D_1 \\ F_{2,n}''\left(0\right) = 4 + n^\alpha \cdot C_2 + n^{2\alpha} \cdot D_2 \end{cases},$$

where

$$C_1 = \frac{12\left(pW_1 + W_2 - qW_2\right)}{\left(-1 + p + q\right)W_1W_2},$$

$$D_1 = \frac{8\left(-1+q\right)\left(\left(1-p\right)pW_1^2 + 2\left(1-p\right)pW_1W_2\right)}{\left(-1+p+q\right)^3 W_1^2 W_2^2}$$

$$\qquad + \frac{8\left(-1+q\right)\left(\left(-1+q\right)^2 + p\left(-1+2q\right)\right)W_2^2}{\left(-1+p+q\right)^3 W_1^2 W_2^2},$$

$$C_2 = \frac{12\left(W_1 + qW_2 - pW_1\right)}{\left(-1+p+q\right)W_1W_2},$$

$$D_2 = \frac{8\left(-1+p\right)\left(\left(-1+p\right)^2 + \left(-1+2p\right)q\right)W_1^2}{\left(-1+p+q\right)^3 W_1^2 W_2^2}$$

$$\qquad + \frac{8\left(-1+p\right)\left(2\left(1-q\right)qW_1W_2 + \left(1-q\right)qW_2^2\right)}{\left(-1+p+q\right)^3 W_1^2 W_2^2}.$$

Then, remembering that

$$\left. \frac{d^2 F_{i,n}(s)}{ds^2} \right|_{s=0} = F_{i,n}''(0) = \mathbb{E}\left[X_n^2\right], \ \forall \ i = 1, 2,$$

and that

$$\begin{cases} Var\left[X_n\right] = F_{1,n}''(0) - \mathbb{E}\left[X_n\right]^2 \\ Var\left[Y_n\right] = F_{2,n}''(0) - \mathbb{E}\left[Y_n\right]^2 \end{cases},$$

from the values (5.9), we obtain the (5.5).

Moreover, the constants K_1 and H_1 are strictly positive with the condition $p + q < 1$. We prove that $K_1 > 0$, for $H_1 > 0$ the proof is the same. The idea is to consider K_1 as the sum of the following three quantites K_{11}, K_{12} and K_{13}

$$K_{11} := \frac{-4p\left(2 - 2q - 3p + 3pq + p^2\right)W_1^2}{\left(p + q - 1\right)^3 W_1^2 W_2^2},$$

$$K_{12} := \frac{8\left(q - 1\right)\left(1 + q - p\right)pW_1W_2}{\left(p + q - 1\right)^3 W_1^2 W_2^2},$$

and

$$K_{13} := \frac{4\left(q - 1\right)\left(\left(q - 1\right)^2 + \left(3q - 1\right)p\right)W_2^2}{\left(p + q - 1\right)^3 W_1^2 W_2^2},$$

and then to verify that are all strictly positive. We begin proving that

$$K_{11} > 0.$$

From the defintions of the parameters p, q, W_1, $W_2 \in (0, 1)$ and the condition $p + q < 1$, we have that

$$-4p < 0,$$

and

$$\left(p + q - 1\right)^3 < 0,$$

and so it is equal to prove that

$$2 - 2q - 3p + 3pq + p^2 > 0, \qquad \forall\, p, q \in (0, 1).$$

We have that

$$
\begin{aligned}
2 - 2q - 3p + 3pq + p^2 &= 2\,(1 - q) + p\,(p - 3 + 3q) = \\
&= 2\,(1 - q) + p\,(p + q - 1) + 2p\,(q - 1),
\end{aligned}
$$

and so, noting that $p + q - 1 < 0$ and $q - 1 < 0$, we get

$$
\begin{aligned}
2 - 2q - 3p + 3pq + p^2 &> 0 \\
2\,(1 - q) &> p\,(1 - p - q) + 2p\,(1 - q) \\
2\,(1 - q)\,(1 - p) &> p\,(1 - p - q) \\
2\,(1 - q)\,(1 - p) &> p\,(1 - p - q + pq - pq) \\
2\,(1 - q)\,(1 - p) &> p\,(1 - p - q + pq) - p^2 q \\
2\,(1 - q)\,(1 - p) &> p\,(1 - q)\,(1 - p) - p^2 q \\
(1 - q)\,(1 - p)\,(2 - p) &> -p^2 q,
\end{aligned}
$$

and it holds for every $p, q \in (0, 1)$.
Now we prove that

$$K_{12} > 0.$$

This is much easier that K_{11}, because, from the defintions of the parameters $p, q, W_1, W_2 \in (0, 1)$ and the condition $p + q < 1$, we have that

$$
\begin{aligned}
q - 1 &< 0, \\
(p + q - 1)^3 &< 0,
\end{aligned}
$$

and

$$pW_1 W_2 > 0,$$

and so it is equal to prove that

$$1 + q - p > 0,$$

but it is always true for every $p, q \in (0, 1)$.
Finally we have to prove that

$$K_{13} > 0.$$

From the defintions of the parameters p, q, W_1, $W_2 \in (0, 1)$ and the condition $p + q < 1$, we have that

$$q - 1 < 0,$$

and

$$(p + q - 1)^3 < 0,$$

and so it is equal to prove that

$$(q - 1)^2 + (3q - 1) p > 0.$$

We have that

$$(q - 1)^2 + (3q - 1) p = (q - 1)^2 + (2q + q - 1) p$$
$$= (q - 1)^2 + p (q - 1) + 2pq = (q - 1) (q - 1 + p) + 2pq.$$

From the defintion of q and the condition $p + q < 1$, we have that $q - 1 < 0$, and $q - 1 + p < 0$, thus, we have that

$$(q - 1)^2 + (3q - 1) p = (q - 1) (q - 1 + p) + 2pq > 0.$$

So, we obtain that

$$K_1 = K_{11} + K_{12} + K_{13} > 0.$$

$\square$

Now, from Lemma 5.2.1, we can apply the weak-convergence Theorem 2.2.1 in the Section 2.2 to the full binary trees with offspring distribution $\boldsymbol{\mu}_n$ at the step n of the scaling. Because in our model the angiogenesis process starts with the growth of a new blood vessel, in the next Theorem we show the weak convergence to a normal distribution of the period of the contour process of a full binary tree with survivals, with offspring distribution $\boldsymbol{\mu}_n$, starting with a type 1 particle and the period is appropriately scaled.

Theorem 5.2.2. *Consider a full binary tree with survivals at the n-th step of the scaling, with offspring distribution $\boldsymbol{\mu}_n$, defined in (5.3). As in the Lemma 5.2.1, we use the following notation*

$$X_n = 2L_n^{(1)}(\varnothing), \qquad Y_n = 2L_n^{(2)}(\varnothing).$$

Moreover, consider $(X_{n,j})_{j \geq 0}$ i.i.d. copies of X_n.
If $p_0^n q_0^n > p_2^n q_2^n$ then

$$Z_n^{(1)} \xrightarrow{d} \mathscr{N}(0, K_1), \ n \to +\infty,$$

where $K_1 > 0$ is the parameter defined in (5.6) in Lemma 5.2.1, and

$$Z_n^{(1)} = n^{\frac{1}{2}} \left[\sum_{j=1}^n \left(\frac{X_{n,j}}{n^{\alpha+1}} \right) - 2 \left(\frac{pW_1 + (1-q)W_2}{(1-p-q)W_1W_2} \right) - \frac{2}{n^{\alpha+1}} \right].$$

Proof. The idea is to apply the Theorem 2.2.1 when $r = 2$, $i = 1$ and the offspring distribution is $\boldsymbol{\mu}_n$ (5.3). From the Lemma 5.2.1, at the *n-th* step of the scaling, we have that if $p_0^n q_0^n > p_2^n q_2^n$ (i.e. $p + q < 1$) then

$$\mathbb{E}[X_n] = 2 \left(\frac{pW_1 + (1-q)W_2}{(1-p-q)W_1W_2} - 1 \right),$$

and

$$Var\left[X_n\right] = K_1 \cdot n^{2\alpha} + K_2 \cdot n^{\alpha},$$

where $p, q \in (0, 1)$, $W_1, W_2 \in (0, 1)$, $\alpha \geq 1$ and the parameters K_1, K_2 are defined from (5.6) and (5.7) respectively in the Lemma 5.2.1. Thus, the condition 1 of the Theorem 2.2.1 is satisfied for X_n, where

$$A_1 = 2 \cdot \frac{pW_1 + (1 - q)W_2}{(1 - p - q)W_1 W_2}, \qquad B_1 = K_1, \qquad \gamma = 2 \cdot \alpha.$$

For the condition 2 of the same theorem we have to prove that

$$\exists\, \delta > 0 \ \text{ such that } \mathbb{E}\left[|X_n - \mathbb{E}\left[X_n\right]|^{2+\delta}\right] = o\left(n^{\frac{\delta + 2\alpha(2+\delta)}{2}}\right). \tag{5.10}$$

In general the system (5.8) has two solutions, but from the Corollary 2 we know that if $p_0^n q_0^n > p_2^n q_2^n$ then there's only one solution (the one such that $F_{1,n}\left(0^-\right) = F_{2,n}\left(0^-\right) = 1$):

$$
\begin{aligned}
F_{1,n}\left(s\right) = {} & \frac{1}{2e^{4s}\left(-1 + e^{2s}p_1^n\right)q_2^n} \cdot \left(\left(p_1^n e^{2s} - 1\right)\left(1 - q_1^n e^{2s}\right) + e^{4s}\left(p_2^n q_0^n - p_0^n q_2^n\right)\right. \\
& + \left[\left(\left(1 - p_1^n e^{2s}\right)\left(1 - q_1^n e^{2s}\right) - e^{4s}\left(p_2^n q_0^n - p_0^n q_2^n\right)\right)^2\right. \\
& \left.\left. -4q_2^n p_0^n e^{4s}\left(1 - p_1^n e^{2s}\right)\left(1 - q_1^n e^{2s}\right)\right]^{1/2}\right), \\
F_{2,n}\left(s\right) = {} & \frac{1}{2e^{4s}\left(-1 + e^{2s}q_1^n\right)p_2^n} \cdot \left(\left(q_1^n e^{2s} - 1\right)\left(1 - p_1^n e^{2s}\right) + e^{4s}\left(q_2^n p_0^n - q_0^n p_2^n\right)\right. \\
& + \left[\left(\left(1 - q_1^n e^{2s}\right)\left(1 - p_1^n e^{2s}\right) - e^{4s}\left(q_2^n p_0^n - q_0^n p_2^n\right)\right)^2\right. \\
& \left.\left. -4p_2^n q_0^n e^{4s}\left(1 - q_1^n e^{2s}\right)\left(1 - p_1^n e^{2s}\right)\right]^{1/2}\right).
\end{aligned}
$$

Note that
$$\mathbb{E}\left[e^{s\cdot(X_n - \mathbb{E}[X_n])}\right] = F_{1,n}\left(s\right) \cdot e^{-s\cdot\mathbb{E}[X_n]},$$

and so we can write

$$\mathbb{E}\left[\left(X_n - \mathbb{E}\left(X_n\right)\right)^4\right] = \left.\frac{d^4}{ds^4}\left(\mathbb{E}\left[e^{s\cdot(X_n - \mathbb{E}[X_n])}\right]\right)\right|_{s=0} = \left.\frac{d^4}{ds^4}\left(F_{1,n}\left(s\right) \cdot e^{-s\cdot\mathbb{E}[X_n]}\right)\right|_{s=0}.$$

The computation of the quantity

$$\frac{d^4}{ds^4}\left(F_{1,n}\left(s\right)\cdot e^{-s\cdot\mathbb{E}[X_n]}\right)\bigg|_{s=0} \tag{5.11}$$

is very laborious and so, using the computational software program *Wolfram Mathematica* (see Appendix A), we have that the maximum power with which n appears in the expanded form of the expression (5.11) is 4α.

So we can conclude that the maximum degree of n of the following expression

$$\mathbb{E}\left[\left|X_n-\mathbb{E}\left(X_n\right)\right|^4\right]=\mathbb{E}\left[\left(X_n-\mathbb{E}\left(X_n\right)\right)^4\right]$$

is 4α, and so for $\delta=2$ the condition (5.10) is verified, indeed

$$n^{4\alpha}=o\left(n^{\frac{2+2\alpha(2+2)}{2}}\right)=o\left(n^{1+4\alpha}\right).$$

Thus we can apply the Theorem 2.2.1 and, remebering that $\gamma=2\cdot\alpha$, we have that

$$Z_n^{(1)}\xrightarrow{d}\mathscr{N}\left(0,K_1\right),\ n\to+\infty,$$

where $K_1>0$ is the parameter defined in (5.6) in the Lemma 5.2.1, and

$$Z_n^{(1)}=n^{\frac{1}{2}}\left[\sum_{j=1}^n\left(\frac{X_{n,j}}{n^{\alpha+1}}\right)-2\left(\frac{pW_1+(1-q)W_2}{(1-p-q)W_1W_2}\right)-\frac{2}{n^{\alpha+1}}\right]$$

$\square$

Conclusions and further perspectives

Essentially this work is a step through the statistical study of multi-type Galton-Watson trees, in particular the full binary trees with two-type vertices. Our main purpose has been to get informations of the offspring distribution $\boldsymbol{\mu}$ of the trees and important results have been proved, like the characterization of the moment-generating function of the contour process related to the full binary trees and the likelihood of the number of left and right fathers of such trees. It is interesting to recall that we have found some particular and interesting relations between arguments of Combinatorics (the Narayana numbers and the two-dimensional decompositions of intgers) and the full binary trees. Moreover, in this work has been shown also a scaling limit result, i.e. we have proved the weak convergence of the period of the contour process related to a full binary tree to a normal distribution.

As applications of this investigation, we have proposed also parametrical models for the angiogenesis process. The main idea has been modelling the process with a two-type Galton-Watson process, both in discrete and continuous time case. Then, for more completeness, we have also considered models based on the full binary trees with two-type vertices and so we could use previous theoretical results.

A basic request that we have done through this work is the homogeneity of the offspring distribution $\boldsymbol{\mu}$ of the full binary trees. In other words, in the models proposed it is equal to say that the branching of the blood vessels

has the same distribution during the angiogenic process.

A further extension may be the use of a non-homogeneous offspring distribution $\boldsymbol{\mu}$

$$\boldsymbol{\mu} = \boldsymbol{\mu}(t), \ \forall t \geq 0,$$

i.e. the offspring distribution (5.1) may be replaced with the following one

$$\begin{cases} \mu_t^{(1)}(0,0) = p_0(t) \\ \mu_t^{(1)}(1,0) = p_1(t) \\ \mu_t^{(1)}(1,1) = p_2(t) \end{cases} \text{and} \begin{cases} \mu_t^{(2)}(0,0) = q_0(t) \\ \mu_t^{(2)}(0,1) = q_1(t) \\ \mu_t^{(2)}(1,1) = q_2(t) \end{cases},$$

where

$$\sum_i p_i(t) = \sum_i q_i(t) = 1, \ \forall t \geq 0,$$

and $p_i(t), q_i(t) \in [0,1], \ \forall t \geq 0$.

In this way $\mu_t^{(i)}(k,l)$ is the probability that a type i particle produces respectively k and l type 1 and type 2 particles at the time $t \geq 0$, for every $i = 1, 2$ and k, l.

This choice may be very useful, indeed we can assume that during an angiogenic process the branching of the vessels may change its own behavior, for instance by a chemioterapic action or the own nature of the biological process. Note that in this case it is crucial the limit $t \to \infty$ and perhaps some quantities may need to be revised, like the moment-generating function of the period of the contour process (2.3).

Another further extension may be the use of k-ary trees, with $k \geq 3$. It could be interesting verify and possibly revise the results that we have outlined for full binary trees, the relations with the Narayana numbers and the decompositions of integers. The k-ary trees could be used also for modelling the angiogenic process, indeed a blood vessel could be seen as union of k-types particles each of which may produce other particles in acccordance to an

offspring distribution

$$\boldsymbol{\mu} = \left(\mu^{(1)}\left(j_1, \ldots, j_k\right), \ldots, \mu^{(k)}\left(j_1, \ldots, j_k\right) \right),$$

where $\mu^{(i)}\left(j_1, \ldots, j_k\right)$ is the probability that a type i particles produces j_1 type 1 particles,$\ldots$, j_k type k particles, for every $i = 1, \ldots, k$ and $k \geq 3$.

Appendix A

Wolfram Mathematica Code

In accordance to the proof of Theorem 5.2.2, we report the commands of the computational software program *Wolfram Mathematica* (see Figure A.1) that we have used to prove that the maximum degree of n of the expression (5.11)

$$\left. \frac{d^4}{ds^4} \left(F_{1,n}\left(s\right) \cdot e^{-s \cdot \mathbb{E}[X_n]} \right) \right|_{s=0}$$

is 4α.

In In [1] and Out [1] with F1n [s] we have denoted $F_{1,n}\left(s\right)$ when $p_0^n q_0^n > p_2^n q_2^n$, where p0, p1, p2, q0, q1 and q2 play the rôle of p_0^n, p_1^n, p_2^n, q_0^n, q_1^n and q_2^n. In In [2] and Out [2] we have verified that F1n [s] is the right solution of the system (5.8) when $p_0^n q_0^n > p_2^n q_2^n$. In In [3], with F [s] we have denoted

$$F_{1,n}\left(s\right) \cdot e^{-s \cdot \mathbb{E}[X_n]},$$

where $\mu = \mathbb{E}\left[X_n\right]$.

Then, in In [4], In [5] and In [6] we have computed in order the quantity (5.11), we have imposed the condition $\sum_i p_i^n = \sum_i q_i^n = 1$ and then we have replaced p0, p1, p2, q0, q1 and q2 with the values of p_0^n, p_1^n, p_2^n, q_0^n, q_1^n and q_2^n defined in the (5.3). Finally, in In [7] and Out [7], we have computed the maximum power of n in the simplified quatity (5.11).

```
In[1]:= F1n[s] = ─────────────────── e^(-4 s)
                  2 (-1 + e^(2 s) p1) q2

          (-1 + e^(2 s) (p1 + q1) - e^(4 s) (-p2 q0 + p1 q1 + p0 q2) +

           Sqrt[-4 e^(4 s) p0 (-1 + e^(2 s) p1) (-1 + e^(2 s) q1) q2 +

           (1 - e^(2 s) (p1 + q1) + e^(4 s) (-p2 q0 + p1 q1 + p0 q2))^2])
```

```
Out[1]= ─────────────────────
        2 (-1 + e^(2 s) p1) q2

        e^(-4 s) (-1 + e^(2 s) (p1 + q1) - e^(4 s) (-p2 q0 + p1 q1 + p0 q2) +

           √(-4 e^(4 s) p0 (-1 + e^(2 s) p1) (-1 + e^(2 s) q1) q2 +

           (1 - e^(2 s) (p1 + q1) + e^(4 s) (-p2 q0 + p1 q1 + p0 q2))^2))
```

```
In[2]:= Assuming[p0 q0 > p2 q2, FullSimplify[
           Limit[F1n[s], s → 0, Assumptions → s < 0] /.
           {p1 → 1 - p0 - p2, q1 → 1 - q0 - q2}]]
```

```
Out[2]= 1
```

```
In[3]:= F[s] = e^(-s μ) F1n[s]
```

```
In[4]:= G = D[F[s], {s, 4}] /. {s → 0}
```

```
In[5]:= H = Assuming[p0 q0 > p2 q2,
           Simplify[G /. {p1 → 1 - p0 - p2, q1 → 1 - q0 - q2}]]
```

```
In[6]:= L = Assuming[α ≥ 1 && W1 > 0 && W1 < 1 && W2 > 0 &&

           W2 < 1 && p > 0 && p < 1 && q > 0 && q < 1, Simplify[

           H /. {p0 → ── (1 - p), p2 → ── p, q0 → ── (1 - q),
                 W1        W1            W2
                 n^α       n^α           n^α

                 W2              2 n^α (p W1 + W2 - q W2)
           q2 → ── q, μ → (-2 - ─────────────────────────)}]]]
                 n^α             (-1 + p + q) W1 W2
```

```
In[7]:= Exponent[L, n]
```

```
Out[7]= Max[α, 2 α, 3 α, 4 α]
```

Figure A.1: Wolfram Mathematica commands showing that the maximum degree of n in the expression (5.11) is exactly equal to 4α.

Bibliography

[1] *Spatial branching processes, random snakes and partial differential equations*, Jean-Francois Le Gall, published by Birkhäuser in the series Lectures in Mathematics ETH Zürich in 1999.

[2] *Random trees and spatial branching processes*, Jean-Francois Le Gall, from a DEA course at Paris 6 in 2000 and a course given at Maphysto in 2000 (PDF, 80 pages).

[3] *Random trees and applications*, Jean-Francois Le Gall, Probability Surveys Vol.2 (2005), 245-311.

[4] *Scaling limits of random trees and planar maps*, Jean-Francois Le Gall and Gregory Miermont, Lecture notes for a course given at the 2010 Clay Mathematics Institute Summer school in Buzios, published in Probability and Statistical Physics in Two and More Dimensions, Clay Mathematics Proceedings, vol.15, pp.155-211. AMS-CMI, 2012.

[5] *An introduction to superprocesses*, Alison M. Etheridge, University Lecture Series, Vol.20, American Mathematical Society.

[6] *The multitype measure branching process*, Luis G. Gorostiza and José A. Lopez-Mimbela, Advances in Applied Probability, Vol.22, No.1(Mar.1990), pp. 49-67.

[7] *Likelihood (Expanded Edition)*, A. W. F. Edwards, Cambridge University Press, 1972.

[8] *Invariance principles for spatial multitype Galton-Watson trees*, Gregory Miermont, Annales de l'Institut Henri Poincaré -Probabilities et Statistiques, 2008, Vol.44, No.6, 1128-1161.

[9] *Stochastic modelling of tumour-induced angiogenesis*, Vincenzo Capasso and Daniela Morale, J.Math.Biol.(2009)58:219-233.

[10] *Fibonacci and Catalan numbers*, Ralph P. Grimaldi, Wiley, 2012.

[11] *Combinatory analysis: Volumes 1,2 bound in one volume*, Percy A. Macmahon, 3 ed., New York:Chelsea publishing company, 1984.

[12] *Narayana numbers and Schur-Szego composition*, V.P. Kostov, A.M.-Finkelshtein, B.Z. Shapiro, Journal of Approximation Theory 161 (2009) 464-476.

[13] *Branching processes*, K.B.Athreya and P.E.Ney, Springer-Verlag, 1972.

[14] *A reinforced random walk model of tumour angiogenesis and anti-angiogenic strategies*, M.J. Plank and B.D. Sleeman, Mathematical Medicine and Biology (2003) 20, 135-181.

[15] *Lattice and Non-Lattice Models of Tumour Angiogenesis*, M.J. Plank and B.D. Sleeman, Bulletin of Mathematical Biology (2004) 66, 1785-1819.

[16] *Reinforced random walk*, B. Davis, Probability Theory and Related Fields 84, 203-229 (1990).

[17] *A biased Random Walk Model for the Trajectories of Swimming Micro-organisms*, N. A. Hill and D.-P. Häder, Journal of Theoretical Biology (1997) 186, 503-526.

[18] *Ising critical exponents on random trees and graphs*, S. Dommers, C. Giardiná and R. van der Hofstad, Commun. Math. Phys. 328, 355-395 (2014).

[19] *Reconstruction on Trees and Spin Glass Transition*, M. Mézard and A. Montanari, Journal of Statistical Physics, Vol. 124, No. 6, September 2006.

[20] *Broadcasting on trees and the Ising model*, W. Evans, C. Kenyon, Y. Peres and L.J. Schulman, The Annals of Applied Probability , Vol. 10, No. 2, 410-433, 2000.

[21] *Random graphs and complex networks, Vol. I*, Lecture notes in preparation, R. van der Hofstad, July 11 2014.

[22] *The weak limit of Ising models on locally tree-like graphs*, A. Montanari, E. Mossel and A. Sly, Probab. Theory Relat. Fields, 152:31-51, 2012.

[23] *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, S. Janson, Probability Surveys, Vol. 9, 103-252, 2012.

[24] *Probability and Measure*, P. Billingsley, Wiley series in Probability and Mathematical Statistics (Third Edition), 1995.

[25] *Probability: A graduate course*, A. Gut, Springer, 2005.