Reinforced Branching Processes

(based on lecture notes by Cécile Mailler, Peter Mörters, and Anna Senkevitch)

Cécile Mailler

Vienna, September 6–9, 2016

Contents

Introduction

The aim of this course is to define the notion of reinforced branching processes and give tools to study them, starting with a crash course on continuous-time martingales.

Although reinforced branching processes are interesting objects in themselves, we will to motivate their definition by looking at one particular case: the preferential attachment tree with fitnesses introduced by Bianconi and Barabási as a model for complex networks. This model is a discrete-time random process on the set of rooted trees, but if we embed this process in continuous time with the help of random exponential clocks, we obtain a continuous-time process called a reinforced branching process.

Reinforced branching processes can be seen as population processes with immortal particles. They are a particular case of the more general Crump-Mode-Jägers processes for which tools are available, especially when they admit a *Malthusian* parameter. We will show how to apply these methods, but will then focus on the case when there is no Malthusian parameter.

In this latter case, a phenomenon called *condensation* occurs: in terms of networks, it means that there exists a *small* set of nodes (typically sublinear in the size of the network) such that the sum of the degrees of these nodes is linear in the size of the network. If there exists one node having linear degree in terms of the size of the network, we say that *the winner takes it all* or that there is *extensive condensation.*

Our aim is to understand under what condition there is condensation, and whether this is extensive or non-extensive condensation. We will briefly introduce the main tools to study reinforced branching processes: martingales, Crump-Mode-Jägers processes, random point processes.

1 A crash course on continuous time martingales

As understood from its title, this section does not aim at being a complete course on martingales. We refer the reader to standard textbooks in which the whole theory is available: for example,

D. Williams: **Probability with Martingales**. *Cambridge University Press*, 1991.

1.1 Definitions and first properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space.

Definition 1.1: A continuous time process $(M_t)_{t\geq 0}$ is a martingale for the filtration $(\mathcal{F}_t)_{t\geq 0}$ if and only if, for all $t \geq 0$,

(i) M_t is $\mathcal{F}_t$-measurable;

(ii) M_t is integrable; and

(iii) for all $s < t$, $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$.

Definition 1.2: Replacing (iii) in the above definition by

- for all $s < t$, $\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s$ gives the definition of a **super-martingale**.
- for all $s < t$, $\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s$ gives the definition of a **sub-martingale**.

Example 1.1: The Yule tree (cf. Figure 1)
Let us consider the stochastic process $(Y_t)_{t\geq 0}$ defined as follows. At time zero, there is one particle in the system: $Y_0 = 1$. Each particle dies and gives birth to two new particles after an exponentially distributed random time, independently from the other particles. Let us denote by Y_t the number of particles alive at time t.

Can you find $(m_t)_{t\geq 0}$ a function such that $M_t := m_t^{-1} Y_t$ is a martingale?

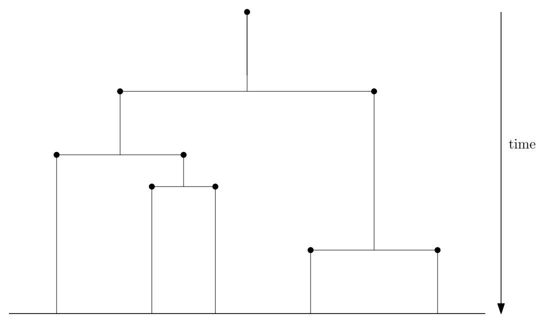

Figure 1 – *A realisation of the Yule tree*

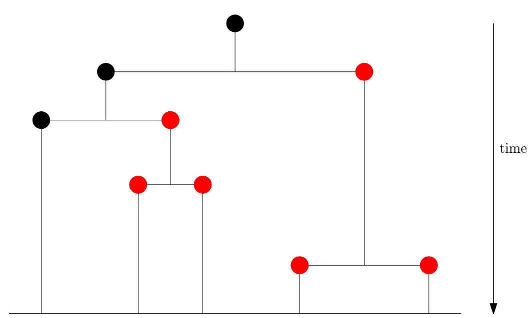

Figure 2 – *A realisation of the multi-type branching process defined by the initial composition* ${}^t(0,1)$ *and the replacement matrix* $R = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$.

Example 1.2: Multi-type branching process (cf. Figure 2)

A multi-type branching process is the embedding in continuous time of a Pólya urn. It is defined by an initial composition $U(0) = {}^t(\alpha, \beta)$ and a replacement matrix

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The vector composition of the urn at time t is given by $U(t) = {}^t(X_t, Y_t)$, where X_t is the number of red balls and Y_t the number of black balls at time t in the urn. Each ball in the urn will split after an exponentially distributed random time into

- $a+1$ red balls and b black balls if it is a red ball;
- or c red balls and $d+1$ black balls if it is a black ball,

independently for the other balls.

Assume that the replacement matrix is balanced: $a+b = c+d = S$. What can you say about the total number of balls in the urn at time t? Can you prove that $M_t := \mathsf{e}^{-tA}U(t)$ is a vector valued martingale, where $A = {}^tR$?

1.2 Doob's inequalities

Proposition 1.3: Let $(M_t)_{t\geq 0}$ a non-negative sub-martingale such that $\mathbb{E}M_0 < +\infty$. Then, for all $\alpha > 0$,

$$\mathbb{P}(\max_{s\leq t} M_s \geq \alpha) \leq \frac{\mathbb{E}M_t}{\alpha}$$

Corollary 1.4: Let $(M_t)_{t\geq 0}$ be a square integrable martingale. Then, for all $\alpha > 0$,

$$\mathbb{P}(\max_{s\leq t} M_s \geq \alpha) \leq \frac{\mathbb{E}M_t^2}{\alpha^2}.$$

1.3 Convergence of continuous time martingales

Definition 1.5: A sequence of random variables $(X_t)_{n\geq 0}$ is **bounded in** L^p if and only if

$$\sup_{t\geq 0} \mathbb{E}|X_t|^p < +\infty.$$

The sequence is uniformly integrable if and only if

$$\lim_{x\to+\infty} \sup_{t\geq 0} \mathbb{E}[X_t \mathbb{1}_{X_t > x}] \to 0,$$

when $x \to +\infty$.

Theorem 1.6

A martingale bounded in L^2 converges in L^2, meaning that there exists a random variable M_∞ such that

$$\lim_{t\to+\infty} \mathbb{E}[|M_t - M_\infty|^2] = 0.$$

Theorem 1.7 (*Doob's Theorem*)

Let $(M_t)_{t\geq 0}$ be a sub-martingale such that

$$\sup_{t\geq 0} \mathbb{E}X_t \mathbb{1}_{X_t\geq 0} < +\infty.$$

Then, M_t converges almost surely to an integrable random variable M_∞.

Corollary 1.8: All non negative super-martingale $(M_t)_{t\geq 0}$ converges almost surely to an integrable random variable M_∞ and

$$\mathbb{E}M_\infty \leq \liminf_{t\to+\infty} \mathbb{E}M_t.$$

Theorem 1.9

Let $(M_t)_{t\geq 0}$ be a martingale. The three following propositions are equivalent:

(i) M_t converges in L^1 to an integrable random variable M_∞;

(ii) $(M_t)_{t\geq 0}$ is bounded in L^1 and there exists a random variable M_∞ such that

$$\mathbb{E}[M_\infty|\mathcal{F}_t] = M_t \qquad (\textit{for all } t \geq 0);$$

(iii) $(M_t)_{t\geq 0}$ is uniformly integrable.

Such a martingale is called **regular**. *It implies in particular that, for all $t \geq 0$, $\mathbb{E}M_t = \mathbb{E}M_\infty$.*

Corollary 1.10: Any martingale bounded in L^p $(p > 1)$ converges in L^p.

2 The Bianconi and Barabási model for complex networks

2.1 Definition of the model

The Barabási and Albert model – The Barabási and Albert model [BA99] was originally introduced in order to try and explain the emergence of power-law degree distributions in complex networks such as the internet, the WWW or social networks. It is indeed observed that many of those real-life networks verify the following heuristic: the proportion of nodes of degree k is close to $k^{-\rho}$ where $\rho > 0$ (when k is large). Graphs verifying this property are called scale-free graphs, and among them, the Barabási and Albert model has the advantage to be dynamic.

The Barabási and Albert model is defined as a Markov chain on the space of rooted trees. We denote it by $(\mathtt{BA}_n)_{n\geq 1}$ and define it recursively as follows (see Figure 3):

- $\mathtt{BA}_1$ is composed of one root-node linked to a root half-edge;
- given $\mathtt{BA}_n$, pick a node at random among the nodes of $\mathtt{BA}_n$, with probability proportional to its degree, and attach a new node to this randomly chosen graph; the obtained graph is $\mathtt{BA}_{n+1}$.

The intuition behind this model is that old nodes will have larger and larger degrees while recently added nodes will have small degrees. This model is a reinforcement model since nodes that already have a large degree are more likely to attract new links.

Many results are known about the Barabási and Albert model (diameter, height, profile, insertion depth, etc); but since this is not the main object of the course, let us just mention its scale-free property:

Theorem 2.1 (*Scale-free property*)

For all $n \geq 0$, we denote by $D_k(n)$ the number of nodes of degree k in $\mathtt{BA}_n$. We have the following asymptotic result:

$$\lim_{n\to\infty} \frac{D_k(n)}{n} \sim k^{-3} \qquad \textit{when } k \to \infty.$$

Remark: Note that with the above definition, the Barabási and Albert graph is actually a tree, which is sometimes called the preferential attachment tree. Of course, a tree is obviously not a realistic model for complex networks since they typically contain cycles (in a social network, for example, two of your friends are quite likely to be friends

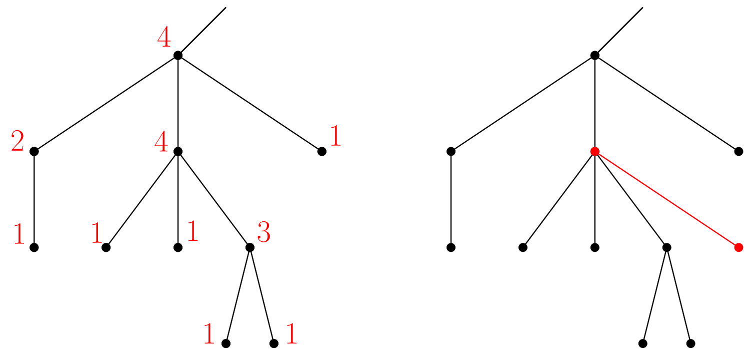

FIGURE 3 – *Barabási and Albert model: To build* $\mathtt{BA}_{n+1}$ *from* $\mathtt{BA}_n$*, each node is weighted by its degree, we pick a node with probability proportional to its weight, and add a child to this randomly chosen node.*

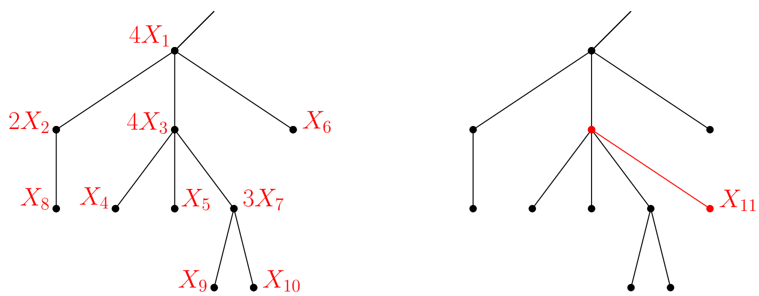

FIGURE 4 – *Bianconi and Barabási model: To build* $\mathtt{BB}_{n+1}$ *from* $\mathtt{BB}_n$*, each node is weighted by its degree times its fitness, we pick a node with probability proportional to its weight, and add a child to this randomly chosen node.*

too). It is thus usual to consider a generalisation of the model above in which you connect every new node with two (or even more) of the old nodes (chosen with probability proportional to their degree).

The Bianconi and Barabási model – The Bianconi and Barabási model [BB01] is a generalisation of the Barabási and Albert model. The idea behind it is to take into account the *intrinsic quality* of the nodes and not only their age into the network: a recent but popular node should be able to increase its degree quickly and compete with older but less popular nodes. This notion of intrinsic quality is modelled by i.i.d. random fitnesses $(X_n)_{n\geq 1}$.

The model is defined as follows: let $(X_n)_{n\geq 1}$ a sequence of i.i.d. random variables of distribution μ being supported on $[0,1]$. Given the random fitnesses, we build the random graph as follows (see Figure 4):

- $\mathtt{BB}_1$ is composed of a root-node of fitness X_1 and a root half-edge linked to this node;
- given $\mathtt{BB}_n$, pick a node in $\mathtt{BB}_n$ with probability proportional to its degree times its fitness, add a new node of fitness X_{n+1} to the graph and link it to the randomly chosen node. The obtained tree is $\mathtt{BB}_{n+1}$.

One can prove that for some fitness distributions μ, the Bianconi and Barabási tree is scale-free. When they introduced this model, Bianconi and Barabási also made the conjecture that *the winner takes it all*, i.e. they claimed that there exists a positive constant c such that

$$\liminf_{n\to\infty} \frac{\max_{i=1..n} \mathtt{deg}(i)}{n} \geq c > 0,$$

where $\mathtt{deg}(i)$ is the degree of the ith node (in order of addition into the graph).

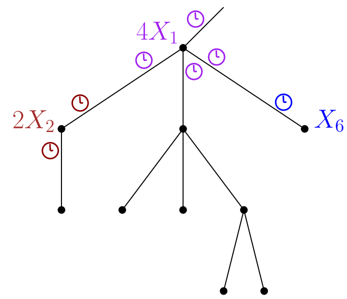

FIGURE 5 – *Bianconi and Barabási model embedded in continuous time: each half edge is equipped with an exponential clock of parameter the fitness of the node it is linked to. All the exponential clocks are independent from each other. When a clock rings, we add a new child with a new random fitness (here X_{11}) to the node linked to the corresponding half-edge (and thus add two new half-edges to the graph).*

The main objective of this mini-course is to show how one can try to prove or disprove this conjecture. The method chosen relies on the embedding in continuous time of the BB model. This continuous time analogue is what we call a *reinforced branching process.* We will define the general model and show how one can study it with the help of martingale theory.

2.2 Embedding into continuous time

It is sometimes convenient to embed the Bianconi and Barabási model into continuous time: we will actually focus on this continuous-time version for our study. Let us consider the following continuous time model (see Figure 5):

- at time $t = 0$, the graph $\mathcal{G}_0$ is composed of a unique node of fitness X_1 to which is linked a half edge;
- given the graph $\mathcal{G}_t$, we equip each of the half-edges present in $\mathcal{G}_t$ with an exponential clock of parameter equal to the fitness of the node it is linked to. Let us denote by τ the (random) time we have to wait starting at time t before one of the clock rings. For all $t \leq s < t + \tau$, we let $\mathcal{G}_s = \mathcal{G}_t$. The graph $\mathcal{G}_{t+\tau}$ is obtained by adding a child with a new random fitness (of distribution μ) to the node linked to the half-edge whose exponential clock rang at time $t + \tau$.

Thanks to the choice of the exponential distribution (see Exercise 5.1), one can check that if we denote by $(\sigma_n)_{n\geq 1}$ the sequence of random times at which a clock rings, then we can couple this continuous-time process with the Bianconi and Barabási model such that

$$(\mathcal{G}_{\sigma_n})_{n\geq 1} = (\mathtt{BB}_n)_{n\geq 1} \quad \text{almost surely.}$$

The main advantage of the embedding in continuous time is that the evolution of the degree of a given node (given its time of birth) is independent from the rest of the graph: as an example, the degree of the root of the tree is a Yule process of rate X_1, and thus behaves as $\mathsf{e}^{X_1 t}\xi_1$ (almost surely when t goes to infinity), where ξ_1 is an exponential of parameter 1 random variable (see Exercise 5.2)

The main disadvantage is that it is hard to get good estimates of the random times $(\sigma_n)_{n\geq 1}$, or equivalently of the number of nodes in the tree at time t.

One can see that the continuous-time model can be seen as a population process in which

- particles correspond to the half-edge of the graph:
  - each particle reproduces at rate equal to its fitness,
  - when it reproduces, it creates a new particle inheriting the same fitness and a new particle having a new fitness drawn according to the fitness distribution μ;

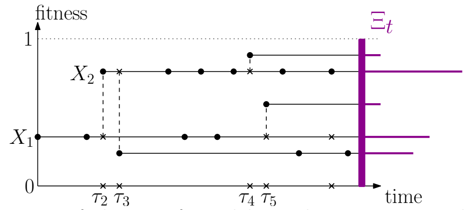

FIGURE 6 – *a realisation of the reinforced branching process: the dots correspond to selection events, the crosses to mutations.*

- particles are grouped into families: a family correspond to the set of half-edges linked to a given node (note that particles in the same family share the same fitness).

Note that in this population process framework, the degrees correspond to the sizes of the families, and thus, the largest degree is the the size of the largest family. This process is what we call a *reinforced branching process*.

2.3 A more general model: reinforced branching processes

A reinforced branching process is a population process $(\mathcal{X}(t))_{t\geq 0}$ that depends on three parameters: the fitness distribution μ, being a distribution on $[0,1]$, the mutation probability $\beta \in [0,1]$ and the selection probability $\gamma \in [0,1]$.

Each particle in the population has a fitness, and the population is partitioned into families such that all particle in one family share the same fitness. At time $t = 0$, the population consists of one family composed of one individual of fitness $X_1 \sim \mu$. We denote by $N(t)$ the number of particles in the population at time t and by $M(t)$ the number of families in the population at time t. We number the families from the eldest to the youngest and for all $1 \leq n \leq M(t)$, we denote by $Z_n(t)$ the size of the nth family at time t, by X_n its fitness and by τ_n the time of birth of the oldest particle of the family (by definition τ_n is increasing in n)

At time t, each family triggers a birth event at rate equal to its size times its fitness. When a birth event happens:

- either a new particle is added to the family that triggered the birth event (we call this particle a *selectant*) – this happens with probability γ;
- or a new family is founded by adding one particle of random fitness $X_{M(t)+1}$ to the population (we call this particle a *mutant*) – this happens with probability β.

Note that the probability that both a selectant and a mutant are born at a given birth event is $\beta + \gamma - 1 \geq 0$; the probability that only a selectant is born is $1 - \beta$ and the probability that only a mutant is born is $1 - \gamma$.

We are interested in the *empirical fitness distribution*, being the random measure given by

$$\Xi_t := \frac{1}{N(t)} \sum_{n=1}^{M(t)} Z_n(t)\, \delta_{X_n},$$

where δ_x is the Dirac mass in x, for all $x \in \mathbb{R}$ (see Figure 6). Does the empirical fitness distribution converges when t goes to infinity? If yes, what is its limit?

Note that if we take $\beta = \gamma = 1$, we get the Bianconi and Barabási model. Other particular cases are worth mentioning:

Example 2.1: Branching process with selection and mutation.

This model is a stochastic *house-of-cards model* in a similar vein as Kingman's model (which is deterministic and much easier to analyse, see [Kin78, DM13]). We start with a single individual with a genetic fitness chosen according to μ. Individuals never die and give birth to new individuals with a rate equal to their genetic fitness. When a new individual is born it is a *mutant* with probability β, in which case it gets a fitness drawn independently of everything else from μ. If the new individual is not a mutant, it inherits the fitness of its parent. Note that when a new individual is born its parent is chosen from the individuals in the population with a probability proportional to their fitness. In other words the different reproduction rates cause the *selection* effect. The number of families $M(t)$ corresponds to the number of mutants in the population at time t.

The model corresponds to the parameter choice $\gamma = 1-\beta$ in our framework. Observe that a mutation causes the complete loss of genetic information in the affected individual's ancestry, pictorially speaking 'the genetic house of cards collapses'. This is why the term house-of-cards model is used for this process, see [HDRT15] for a discussion of the relevance of these models in the theory of evolution.

Example 2.2: Generalised Pólya urns.

A class of generalised Pólya urns also falls into our framework, with general parameters $\beta, \gamma > 0$ and μ as above. It can be described as an urn containing balls of different colours. Every colour has a given *activity* chosen independently according to μ. At time zero, the urn contains one ball of colour 1. At every time step, a ball is drawn at random from the urn with probability proportional to its activity. Then the drawn ball is put back into the urn together with one or two new balls, at most one ball of the same and one of a new colour. A ball with the same colour is chosen with probability γ, and a ball of a new colour with probability β. New colours are chosen independently according to μ. To embed the urn model into our framework we again look at the times of birth events. Observe that Ξ_t is now the empirical distribution of activities in the urn at time t.

Such generalised Pólya urns have apparently not been studied so far in full generality. Janson [Jan04] is looking at the case where μ is finitely supported. A related model has been studied by Chung et al. [CHJ03] who draw balls depending in a non-linear way on the distribution of colours in the urn, and by Collevecchio et al. [CCL13] who allow for a time-dependent replacement rule. Their main focus is on the question whether there can be an unbounded number of balls of more than one colour, and if not which colour eventually dominates. In our setup all colours will have an unbounded number of balls.

3 Condensation

3.1 Our process as a Crump-Mode-Jägers process

We show here how the reinforced branching process falls into the Crump-Mode-Jägers framework and how Nerman's asymptotic results [Ner81] translate in our particular case.

Poisson point process – Given a positive measure ξ on $\mathbb{R}$, we define the Poisson point process of intensity ξ as the unique random measure $\sum_{i=1}^{\infty} \delta_{P_i}$ (where the P_i are real random variables) such that,

- for all borelian set B, the random variable $\mathtt{PPP}_\xi(B)$ is Poisson distributed of parameter $\xi(B)$;
- for all disjoint borelian sets $B_1, \ldots, B_r$ (for all integers r), the random variables $\mathtt{PPP}_\xi(B_1), \ldots, \mathtt{PPP}_\xi(B_r)$ are independent random variables

The reinforced branching process is defined by a sequence (X_n, Y_n, Π_n) of independent random variables such that

- X_n is a μ-distributed random variable in $[0,1]$,
- given X_n the process $Y_n = (Y_n(t))_{t\geq 0}$ is an independent Yule process with rate γX_n,
- given X_n and Y_n the process $\Pi_n = (\Pi_n(t))_{t\geq 0}$ is an inhomogeneous Poisson process with intensity measure given by
$$\frac{\beta+\gamma-1}{\gamma}\,\delta Y_n(t) + (1-\gamma)X_n Y_n(t)\,dt \quad (\forall t \geq 0).$$

Recall that Y_n determines the birth of family members of the nth family relative to the foundation time of the family, and Π_n the birth times of mutant offspring from this family. For greater generality we enrich this triple (X, Y, Π) by a fourth component $\Phi = (\Phi(t))_{t\geq 0}$, a càdlàg process taking values in $\mathcal{N}$ assigning some kind of score to the family t time units after its foundation. In all our examples (below) Φ is a function of (X, Y, Π) but this does not have to be the case. We use the convention that $\Phi(t) = 0$ if $t < 0$.

We let $\tau_1 = 0$ and

$$\tau_n = \inf\{t > \tau_{n-1} : \exists m \in \{1, \ldots, n-1\} \text{ with } \Delta\Pi_m(t - \tau_m) = 1\}.$$

Note that τ_n corresponds to the time of birth of the nth family. We then define

$$Z^{\Phi}(t) = \sum_{n\colon \tau_n < t} \Phi_n(t - \tau_n),$$

the *score of the population* at time t. Here are the main examples of interest to us.

(1) Let

$$\Phi_n^{(1)}(t) = \begin{cases} Y_n(t), & \text{if } t \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then $\Phi_n^{(1)}(t - \tau_n) = Z_n(t)$ is the size of the nth family at time t and hence $Z^{\Phi^{(1)}}(t) = N(t)$.

(2) Let $\Phi_n^{(2)}(t) = 1$ if $t \geq 0$ and zero otherwise. Then $Z^{\Phi^{(2)}}(t) = M(t)$ is the total number of families in the system at time t.

(3) Let $a > 0$ and

$$\Phi_n^{(a)}(t) = \begin{cases} Y_n(t), & \text{if } 0 \leq t < a, \\ 0, & \text{otherwise.} \end{cases}$$

Then $Z^{\Phi^{(a)}}(t)$ is the number of individuals in families younger than a at time t.

(4) Let $0 < x < 1$ and

$$\Phi_n^{(x)}(t) = \begin{cases} Y_n(t), & \text{if } X_n \geq 1 - x \text{ and } t \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then $Z^{\Phi^{(x)}}(t) = N(t)\, \Xi_t[1 - x, 1]$ and together with (1) this can be used to identify the limit of the empirical fitness distribution of particles.

(5) Let $k \in \mathbb{N}$ and

$$\Phi_n(t) = \begin{cases} 1, & \text{if } t \geq 0 \text{ and } Y_n(t) = k, \\ 0, & \text{otherwise.} \end{cases}$$

Then $Z^{\Phi}(t)$ is the number of families of size k at time t. In the Barabási and Bianconi tree this refers to the number of vertices of degree k and will allow the calculation of the empirical degree distribution.

The main result of this section is a convergence theorem under the following main assumption.

Assumption 1: Existence of a Malthusian parameter There exists an $\lambda^\star > \gamma$, called the *Malthusian parameter*, such that

$$1 = \int_0^\infty e^{-\lambda^\star s}\, \mathbb{E}\Pi(ds).$$

We shall see below what this condition means in terms of the model parameters β, γ and μ. We also formulate an assumption on the process Φ.

Assumption 2: Regularity of Φ The function $t \mapsto \mathbb{E}[\Phi(t)]$ is almost everywhere continuous and there exists $h\colon [0,\infty) \to (0,\infty)$ integrable, bounded and non-increasing such that

$$\mathbb{E}\Big[\sup_{t\geq 0} \frac{e^{-\lambda^\star t}\Phi(t)}{h(t)}\Big] < \infty.$$

Theorem 3.1

Under Assumptions 1 and 2, there exists a positive random variable W, not depending on Φ, such that

$$\lim_{t\to\infty} e^{-\lambda^\star t} Z_t^\Phi = W\, m_\infty^\Phi \qquad \textit{almost surely,}$$

where

$$m_\infty^\Phi = \frac{\int_0^\infty e^{-\lambda^\star t}\mathbb{E}\Phi(t)\,dt}{\int_0^\infty te^{-\lambda^\star t}\,\mathbb{E}\Pi(dt)}.$$

We now look at the consequences of Theorem 3.1. We first express Assumption 1 explicitly in terms of the model parameters β, γ and μ. We have, for any $\lambda^\star \geq \gamma$,

$$\begin{aligned}\int_0^\infty e^{-\lambda^\star s}\,\mathbb{E}\Pi(ds) &= \int d\mu(x)\Big\{\frac{\beta+\gamma-1}{\gamma}\int_0^\infty \mathsf{e}^{-\lambda^\star s}\,de^{\gamma xs} + (1-\gamma)x\int_0^\infty e^{-\lambda^\star s}e^{\gamma xs}\,ds\Big\}\\ &= \beta\int x\int_0^\infty e^{-\lambda^\star s+\gamma xs}\,ds\,d\mu(x)\\ &= \beta\int\frac{x}{\lambda^\star-\gamma x}\,d\mu(x).\end{aligned}$$

This is decreasing in $\lambda^\star$ and going to zero as $\lambda^\star \to \infty$. As $\lambda^\star \downarrow \gamma$ it converges to

$$\frac{\beta}{\gamma}\int\frac{x}{1-x}\,d\mu(x),$$

which has to be at least one for a Malthusian parameter to exists. Hence Assumption 1 translates to

$$\frac{\beta}{\gamma}\int\frac{x}{1-x}\,d\mu(x) > 1,$$

or, equivalently,

$$\frac{\beta}{\beta+\gamma}\int\frac{d\mu(x)}{1-x} > 1. \tag{1}$$

When (1) holds, the Malthusian parameter $\lambda^\star$ is defined by the equation

$$\beta\int\frac{x}{\lambda^\star-\gamma x}\,d\mu(x) = 1. \tag{2}$$

3.2 The Maulthusian case

Let us now look at the examples of scores Φ listed earlier and harvest the results. Until the end of this section, we assume that (1) holds and that there exists a Malthusian parameter.

(1) Almost surely,

$$\lim_{t\to\infty}\mathsf{e}^{-\lambda^\star t}N(t) = Wm_\infty^{\Phi^{(1)}}.$$

To confirm this result we check that Assumption 2 holds for $\Phi^{(1)}$: using Doob's maximal inequality and Exercise 5.2, we have

$$\mathbb{E}\Big[\sup_{t\geq 0}\mathsf{e}^{-\gamma Xt}\Phi^{(1)}(t)\Big] = \mathbb{E}\Big[\sup_{s\geq 0}\mathsf{e}^{-s}Y_s\Big] \leq 2\sup_{s\geq 0}\sqrt{\mathbb{E}[\mathsf{e}^{-2s}Y_s^2]} < \infty.$$

We see that the right hand side is strictly positive showing that the number of individuals has purely exponential growth. For later comparison we calculate the numerator of $m_\infty^{\Phi^{(1)}}$, i.e. the score dependent quantity. We get

$$\int d\mu(x) \int_0^\infty \mathsf{e}^{-\lambda^\star t} \mathbb{E}\Phi^{(1)}(t)\, dt = \int \frac{d\mu(x)}{\lambda^\star - \gamma x} = \frac{\beta+\gamma}{\lambda^\star \beta}.$$

(2) Almost surely,

$$\lim_{t\to\infty} \mathsf{e}^{-\lambda^\star t} M(t) = W m_\infty^{\Phi^{(2)}}.$$

To compare with (1) we calculate the score dependent numerator of $m_\infty^{\Phi^{(2)}}$. We get $\int_0^\infty \mathsf{e}^{-\lambda^\star t}\, dt = {}^1/_{\lambda^\star}$. Therefore, we get that $\lim {}^{M(t)}/_{N(t)} = {}^{\beta}/_{\beta+\gamma}$, almost surely when t goes to infinity, as expected by the law of large numbers.

(3) We see that the proportion of individuals in families born less than a time units ago is asymptotically equal to

$$\frac{\lambda^\star \beta}{\beta+\gamma} \int d\mu(x) \int_0^a \mathsf{e}^{-\lambda^\star t + \gamma x t}\, dt.$$

This limit goes to one as a goes to ∞, which shows that most individuals come from recently established families.

(4) Almost surely,

$$\lim_{t\to\infty} \Xi_t[1-x, 1] = \frac{\lambda^\star \beta}{\beta+\gamma} \int_{1-x}^1 \frac{1}{\lambda^\star - \gamma x}\, d\mu(x).$$

In other words, in the bulk driven phase, the empirical fitness distribution converges to a deterministic probability distribution which is absolutely continuous with respect to μ and has density

$$\frac{\beta}{\beta+\gamma} \frac{\lambda^\star}{\lambda^\star - \gamma x}.$$

Example (5) in treated in Exercise 5.3. A somewhat similar application of general branching processes to the study of preferential attachment networks (without fitness but with a nonlinear attachment rule) is carried out in Rudas et al. [RTV07].

3.3 Condensation

We have seen in Section 3.2 that when a Malthusian parameter exists, then one can obtain limit theorems for different measurable quantities of the system such as the number of families or of particles in the system, or the empirical fitness distribution, or the distribution of family sizes. This chapter is devoted to the study of reinforced branching processes which do not admit a Malthusian parameter. We will see that reinforced branching processes with no Malthusian parameter exhibit *condensation*, meaning that the empirical fitness distribution converges to the sum of an absolute continuous part, called the *bulk*, and a Dirac mass in the essential supremum of the support of the fitness distribution, called the *condensate*.

Recall the definition of the empirical fitness distribution:

$$\Xi_t = \frac{1}{N(t)} \sum_{n=1}^{M(t)} Z_n(t)\, \delta_{X_n}.$$

Theorem 3.2

Assume that

$$\frac{\beta}{\beta+\gamma} \int_0^1 \frac{d\mu(x)}{1-x} < 1, \tag{cond}$$

and let $\lambda^\star := \gamma$. Then

1. $\int x\, d\Xi_t(x) \to \lambda^\star/\beta+\gamma$ *almost surely when t goes to infinity;*

2. $\Xi_t \to \pi$ *almost surely weakly when t goes to infinity, where*

$$\mathrm{d}\pi(x) = \frac{\beta}{\beta+\gamma}\,\frac{1}{1-x}\,d\mu(x) + \varpi(\beta,\gamma)\delta_1,$$

with

$$\varpi(\beta,\gamma) = 1 - \frac{\beta}{\beta+\gamma}\int_0^1 \frac{d\mu(x)}{1-x} > 0.$$

Remark: Theorem 3.2(i) implies that

$$\lim_{t\to\infty}\frac{1}{t}\log N(t) = \gamma.$$

Moreover in the empirical fitness distribution we see the phenomenon of condensation, as a positive fraction of individuals are pushed toward the extreme fitness value.

The proof we develop here uses Theorem 3.1: the idea is to couple the branching process with a branching process admitting a Malthusian parameter and apply Theorem 3.1 to the latter. The two coupled branching processes are continuous-time branching processes, but the coupling only relates their discrete-time versions.

The coupling of the processes (lower bound).

We look at the reinforced brancing process with fitness distribution μ at the time (σ_n) of the birth events and abbreviate $\widehat{\Xi}_n := \Xi_{\sigma_n}$.

Fix $\varepsilon > 0$. We define a discrete-time branching process whose empirical fitness distribution $\widehat{\Xi}_n^{(\varepsilon)}$ has the property that for all $n \geq 0$, $(\widehat{\Xi}_n, \widehat{\Xi}_n^{(\varepsilon)}) \in \mathcal{S}$, where $\mathcal{S}$ is the subset of the set of pairs of counting measures on $[0,1]$ defined by

$$\mathcal{S} := \big\{(\nu,\mu) \colon \nu\big([0,1]\big) = \mu\big([0,1]\big) \text{ and } \nu\big([a,b]\big) \geq \mu\big([a,b]\big) \text{ for all } a,b \in [0,1-\varepsilon)\,\big\}.$$

Let $(U_n)_{n\geq 1}$ be a sequence of i.i.d. random variables uniformly distributed on $[0,1]$. At time zero, the new process contains one particle of fitness $X_1\mathbf{1}_{X_1<1-\varepsilon} + \mathbf{1}_{X_1\geq 1-\varepsilon}$ and thus $(\widehat{\Xi}_0, \widehat{\Xi}_0^{(\varepsilon)}) \in \mathcal{S}$.

Assume now that, $(\widehat{\Xi}_n, \widehat{\Xi}_n^{(\varepsilon)}) \in \mathcal{S}$. We construct the new process at time $n+1$ as follows:

- if a mutant of fitness f is born at time $n+1$ (in the original process), then we add in the (new) process a new particle of fitness $x\mathbf{1}\{x < 1-\varepsilon\} + \mathbf{1}\{x \geq 1-\varepsilon\}$ born at time $n+1$;
- if a selectant of fitness larger than $1-\varepsilon$ is born at time $n+1$ in the original process, then we add a new particle of fitness 1 born at time $n+1$;
- if a selectant of fitness $x < 1-\varepsilon$ is born at time $n+1$ in the original process, then if

$$U_{n+1} \leq \left(\frac{\widehat{\Xi}_n^{(\varepsilon)}(\{x\})}{\int_0^1 u\,\mathrm{d}\widehat{\Xi}_n^{(\varepsilon)}(u)}\right)\left(\frac{\widehat{\Xi}_n(\{x\})}{\int_0^1 u\,\mathrm{d}\widehat{\Xi}_n(u)}\right)^{-1},$$

  we add a particle of fitness f born at time $n+1$, otherwise, add a particle of fitness 1.

By construction, $(\widehat{\Xi}_{n+1}, \widehat{\Xi}_{n+1}^{(\varepsilon)}) \in \mathcal{S}$. It is now easy to check that the new process is the discrete-time version of the reinforced branching process with fitness distribution $\mu_\varepsilon := \mathbf{1}_{[0,1-\varepsilon)}\mu + \mu(1-\varepsilon,1)\delta_1$, and falls into the framework of Section 3.2. Since

$$\frac{\beta}{\beta+\gamma}\int_0^1 \frac{d\mu_\varepsilon(x)}{1-x} = \infty,$$

the new process admits a Malthusian parameter λ_ε and $\lambda_\varepsilon \downarrow \gamma$ as $\varepsilon \to 0$. We thus deduce that, for all $0 \le a, b < 1-\varepsilon$, we have

$$\lim_{n\to\infty} \widehat{\Xi}_n^{(\varepsilon)}\big([a,b]\big) = \lim_{t\to\infty} \widehat{\Xi}_t^{(\varepsilon)}\big([a,b]\big) = \frac{\beta}{\beta+\gamma}\int_a^b \frac{\lambda_\varepsilon}{\lambda_\varepsilon - \gamma x}\, d\mu(x)$$

almost surely. For all $0 \le a, b < 1$ and $0 < \varepsilon < 1-b$, we thus have

$$\liminf_{t\to\infty} \Xi_t\big([a,b]\big) = \liminf_{n\to\infty} \widehat{\Xi}_n\big([a,b]\big) \ge \lim_{n\to\infty} \widehat{\Xi}_n^{(\varepsilon)}\big([a,b]\big) = \frac{\beta}{\beta+\gamma}\int_a^b \frac{\lambda_\varepsilon}{\lambda_\varepsilon - \gamma x}\, d\mu(x).$$

Letting ε tend to 0 concludes the proof of the lower bound.

The coupling of the processes (upper bound).

Fix $\varepsilon > 0$, and let $(\Xi_t^{[\varepsilon]})_{t\ge 0}$ be the reinforced branching process of fitness distribution

$$\mu^{[\varepsilon]} = \mathbf{1}_{[0,1-\varepsilon)}\mu + \mu(1-\varepsilon, 1)\delta_{1-\varepsilon},$$

and $\widehat{\Xi}_n^{[\varepsilon]} = \Xi_{\sigma_n}^{[\varepsilon]}$ its discrete-time version. Denote by $X_n^{[\varepsilon]}$ the i.i.d. sequence of fitnesses in this reinforced branching process and by $\lambda^{[\varepsilon]}$ the Malthusian parameter.

We construct a coupling of $\widehat{\Xi}_n$ and $\widehat{\Xi}_n^{[\varepsilon]}$ such that $(\widehat{\Xi}_n^{[\varepsilon]}, \widehat{\Xi}_n,) \in \mathcal{S}$. Let $(V_n)_{n\ge 1}$ be a sequence of i.i.d. random variables uniformly distributed on $[0,1]$ and $(W_n, W_n')_{n\ge 1}$ be independent sequences of i.i.d. random variables of distribution $\mathbf{1}_{(1-\varepsilon,1]}\mu/\mu(1-\varepsilon,1)$.

We construct $\widehat{\Xi}_n$ from $\widehat{\Xi}_n^{[\varepsilon]}$. At time zero, $\widehat{\Xi}_0 = \delta_{X_1}$, where $X_1 = X_1^{[\varepsilon]}\mathbf{1}\{X_1^{[\varepsilon]} < 1-\varepsilon\} + W_1\mathbf{1}\{X_1^{[\varepsilon]} = 1-\varepsilon\}$ and hence $(\widehat{\Xi}_0^{[\varepsilon]}, \widehat{\Xi}_0) \in \mathcal{S}$.

Assume now that $(\widehat{\Xi}_n^{[\varepsilon]}, \widehat{\Xi}_n) \in \mathcal{S}$. We define $\widehat{\Xi}_{n+1}$ as follows:

- if a mutant of fitness x is born at time $n+1$ in the ε-truncated process, then

$$\widehat{\Xi}_{n+1} = \widehat{\Xi}_n + \delta_x\mathbf{1}\{x < 1-\varepsilon\} + W_{n+1}\mathbf{1}\{x = 1-\varepsilon\};$$

- if a selectant of fitness $1-\varepsilon$ is born at time $n+1$ in the ε-truncated process, let

$$\widehat{\Xi}_{n+1} = \widehat{\Xi}_n + \delta_{W_{n+1}'}$$

- if a selectant of fitness $x < 1-\varepsilon$ is born at time $n+1$ in the ε-truncated process, then if

$$V_{n+1} \le \frac{\Xi_n\big(\{x\}\big)}{\int_0^1 u\,\mathrm{d}\widehat{\Xi}_n(u)}\left(\frac{\widehat{\Xi}_n^{[\varepsilon]}\big(\{x\}\big)}{\int_0^1 u\,\mathrm{d}\widehat{\Xi}_n^{[\varepsilon]}(u)}\right)^{-1},$$

  then $\Xi_{n+1} = \Xi_n + \delta_x$, otherwise, $\Xi_{n+1} = \Xi_n + \delta_{W_{n+1}'}$.

By construction, $(\widehat{\Xi}_{n+1}, \widehat{\Xi}_{n+1}^{(\varepsilon)}) \in \mathcal{S}$, and it is easy to check that $(\widehat{\Xi}_n)_{n\ge 0}$ has indeed the same law as the empirical fitness distribution of the original reinforced branching process.

We get that, for all $0 < a < b < 1-\varepsilon$,

$$\limsup_{n\to\infty} \widehat{\Xi}_n(a,b) \le \lim_{n\to\infty} \widehat{\Xi}_n^{[\varepsilon]}(a,b) = \frac{\beta}{\beta+\gamma}\int_a^b \frac{\lambda^{[\varepsilon]}}{\lambda^{[\varepsilon]} - \gamma x}\, d\mu^{[\varepsilon]}(x).$$

Observing that $\mu^{[\varepsilon]} \to \mu$ weakly and $\lambda^{[\varepsilon]} \to \gamma$, as $\varepsilon \downarrow 0$, is enough to conclude the proof of the lower bound and hence of Theorem 3.2.

4 The size of the largest family

In this section we follow [DMM] and study asymptotics for the features (like size or fitness) of the largest family in the system at time t. Such results require regularity assumptions on μ at the upper end. We assume here that $w(\mu) < \infty$ and then, without loss of generality, $w(\mu) = 1$. We further assume that μ has a regularly varying tail in one, meaning that

$$\frac{\mu(1 - x\varepsilon, 1)}{\mu(1 - \varepsilon, 1)} \to x^{\alpha}, \qquad \text{for all } x > 0 \text{ as } \varepsilon \to 0,$$

or equivalently

$$\mu(1 - \varepsilon, 1) = \varepsilon^{\alpha} \ell(\varepsilon), \tag{RV}$$

for a slowly varying function ℓ and some $\alpha > 1$ (see [BGT89]). This corresponds to the most common type of behaviour of μ at its tip that allows a condensation phase.

We introduce the random times $T(t)$, $t > 0$, as

$$T(t) := \inf\Big\{ s \geq 0 : \; M(s) \geq n(t) \Big\} \quad \text{where} \quad n(t) := \left\lceil \frac{1}{\mu(1 - t^{-1}, 1)} \right\rceil.$$

Our intuition is that

- the largest families of the population at time t are born around time $T(t)$;
- $T(t)$ grows like $1/\lambda^\star \log n(t) \sim \frac{\alpha}{\lambda^\star} \log t$;
- the largest families at time t have fitness X_n with $1 - X_n$ of order $1/t$ and size of order $\mathsf{e}^{\gamma(t - T(t))}$.

To confirm our intuition we consider the point process

$$\Gamma_t = \sum_{n=1}^{M(t)} \delta\big(\tau_n - T(t), t(1 - X_n), \mathsf{e}^{-\gamma(t - T(t))} Z_n(t)\big),$$

where $\delta(x)$ is the Dirac mass in x.

Theorem 4.1 (*Poisson limit*)

Under assumption (RV), *the point process* $(\Gamma_t)_{t \geq 0}$ *converges vaguely on the space* $[-\infty, \infty] \times [0, \infty] \times (0, \infty]$ *to the Poisson point process* Π_ζ *with intensity measure*

$$d\zeta(s, x, z) = \alpha x^{\alpha - 1} \lambda^\star \mathsf{e}^{\lambda^\star s} \mathsf{e}^{-z \mathsf{e}^{\gamma(s+x)}} \mathsf{e}^{\gamma(s+x)} \, ds \, dx \, dz.$$

We give here some quick definitions that help understanding Theorem 4.1:

- A point process is a random measure of the form $\sum_{i=1}^{\infty} \delta_{P_i}$, where the P_i's are some random elements of $\mathbb{R}^3$.
- The Poisson point process of intensity ζ on $\mathbb{R}^3$ is characterised by the following properties:
  - for all Borel set B, $\mathtt{PPP}_\zeta(B) = \int_B d\mathtt{PPP}_\zeta$ is Poisson distributed of parameter $\zeta(B)$;
  - for all disjoints Borel sets $B_1, \ldots, B_r$, the random variables $\mathtt{PPP}_\zeta(B_i)$ are independent.
- Vague convergence means that for all compact set K of $[-\infty, \infty] \times [0, \infty] \times (0, \infty]$,

$$\int_K d\Gamma_t \to \int_K d\mathtt{PPP}_\zeta.$$

Remark: Note the compactifications at $\pm\infty$ in Theorem 4.1. As the limiting point process has a continuous density, Theorem 4.1 implies that all mass of Γ_t that asymptotically accumulates at infinity in one of the first two components, must escape at zero in the last component, meaning that the only way points can disappear in the limit is because the corresponding family size is small relative to the normalisation.

Remark: As there is no scaling in the first component of Γ_t, the limit theorem focuses on a time window of constant width around $T(t)$. The theorem shows that this is wide enough to capture the largest family at time t. However, it turns out that in the condensation phase this is *not* wide enough to capture *all* families that contribute to the condensate. This is why important questions on the emergence of the condensate remain open.

Corollary 4.2 (Limits of family characteristics)**:** Let $V(t)$ be the fitness and $S(t)$ the birth time of the family of maximal size at time t. There exist random variables U, V, Z such that, in distribution as $t \to \infty$,

(i) $e^{-\gamma(t-T(t))} \max_{n\in\mathbb{N}} Z_n(t) \to Z$,

(ii) $t(1 - V(t)) \to V$,

(iii) $S(t) - T(t) \to U$.

Remark: The birth time of the family of maximal size at time t is of asymptotic order $T(t) + \mathcal{O}(1)$ and hence of leading order $\alpha/\chi \log t$. This answers a question of Borgs et al. [BCDR] about the rate at which new nodes with higher fitness become the leading influence in the population. We prove Corollary 4.2 and give further details of the limit laws in Exercise 5.4.

A further problem that can be solved using Theorem 4.1 is about that *emergence* of the condensate, i.e. how the condensate manifests itself at large finite times. Following the discussion of Bose-Einstein condensation in van den Berg et al. [vdBLP86] two alternative scenarios are possible:

- For the largest family, the proportion of individuals belonging to this family in the overall population at time t is asymptotically positive. This phenomenon of *macroscopic occupancy* arises in condensation of the free Bose gas below a critical temperature (see [vdBLP86]).
- No individual family makes an asymptotically positive contribution. Instead, it is a collective effort of a growing number of families to form the condensate. This phenomenon is called *non-extensive condensation*. van den Berg et al. [vdBLP86] have shown that this occurs in the free Bose gas for an intermediate temperature range.

The following theorem ensures that the second scenario prevails:

Theorem 4.3 (*The winner does not take it all*)

Under assumption (RV) *the size of the largest family is negligible relative to the overall population size, i.e.*

$$\lim_{t\to\infty} \frac{\max_{n\in\{1,\dots,M(t)\}} Z_n(t)}{N(t)} = 0, \text{ in probability.}$$

Remark: Theorem 4.3 means that asymptotically no single family contributes a positive proportion of the total mass, hence if there is condensation it is always *non-extensive*. This means in the context of Example 2 that no vertex attracts a positive fraction of the edges in the network. This is at odds with the informal description of condensation in the preferential attachment networks by Bianconi and Barabasi [BB01], who are stating that *'the fittest node [is] acquiring a finite fraction of the links, independent of the size of the network.'* It is also at odds with more recent work of Godrèche and Luck [GL10] who use a non rigorous analysis on assumptions based on simulations to conclude that asymptotically there is even an unbounded number of macroscopic families. However the phenomenon we investigate here is too subtle to be reliably captured by non-rigorous techniques. In the context of Example 3 our theorem states that the proportion of balls of any colour goes to zero, uniformly over all colours.

Sketch of Theorem 4.3. Subject to a cut-off argument we have in view of Theorem 4.1,

$$\mathsf{e}^{-\gamma(t-T(t))} \sum_{n=1}^{M(t)} Z_n(t) = \int z\, d\Gamma_t(s,x,z) \sim \int z\, d\Pi_\zeta(s,x,z) \quad \text{as } t \to \infty,$$

where Π_ζ is the Poisson random measure with intensity measure ζ. We calculate

$$\zeta\left(\mathbb{R} \times (0,\infty) \times (a,b)\right) = \frac{\Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})}{(\lambda^\star)^\alpha}\left(a^{-\frac{\lambda^\star}{\gamma}} - b^{-\frac{\lambda^\star}{\gamma}}\right),$$

and hence, as $\lambda^\star \geq \gamma$, we get

$$\int z\, d\zeta(s,x,z) = \frac{\Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})}{(\lambda^\star)^\alpha} \int_0^\infty \frac{\lambda^\star}{\gamma} z^{-\lambda^\star/\gamma}\, dz = \infty.$$

From this we conclude that

$$\mathsf{e}^{-\gamma(t-T(t))} \sum_{n=1}^{M(t)} Z_n(t) \to \infty,$$

while $\mathsf{e}^{-\gamma(t-T(t))} \max_{n \leq M(t)} Z_n(t)$ converges in distribution and hence remains finite. □

5 Exercises

Exercise 5.1: The exponential law
Recall that a random variable X follows the exponential distribution of parameter $a > 0$ if and only if, for all $x \geq 0$,

$$\mathbb{P}(X \geq x) = \int_x^\infty a\mathsf{e}^{-ax}\, dx.$$

Let $X, X_1, \ldots, X_n$ be i.i.d. random variables exponentially distributed of parameter 1.

(a) What is the distribution of $\min_{i=1..n} X_i$?

(b) What is the probability that $\min_{i=1..n} X_i = X_1$?

(c) Show that for all $0 \leq x < y$,
$$\mathbb{P}(X \geq y \,|\, X \geq x) = \mathbb{P}(X \geq y - x).$$
(We say the the exponential distribution "lacks memory".)

Exercise 5.2: The Yule process
Recall that the Yule process of parameter η is characterised as follows: Let τ be an exponential random variable of parameter η, then $Y(t) = 1$ for all $t < \tau$, and for all $t \geq \tau$, $Y_t = Y^{(1)}_{t-\tau} + Y^{(2)}_{t-\tau}$ where $Y^{(1)}$ and $Y^{(2)}$ are two independent copies of Y.

Let $(Y_t : t \geq 0)$ be a Yule process with rate η.

(a) Let $a > 0$ and show that $(Y_{at} : t \geq 0)$ is a Yule process with rate $a\eta$.

(b) Show that $(\mathsf{e}^{-\eta t} Y_t : t \geq 0)$ is a martingale.

(c) Infer that there exists a random variable ξ such that, almost surely,

$$\lim_{t\to\infty} \mathsf{e}^{-\eta t} Y_t = \xi.$$

(d) Show that ξ is exponentially distributed with parameter one.

(e) Show that $\sup_{t\geq 0} \mathbb{E}\mathsf{e}^{-2\eta t} Y_t^2 < \infty$.

Exercise 5.3: Scale-free property of the BB tree.
Let us denote by

$$\Theta_t := \frac{1}{M(t)} \sum_{n=1}^{M(t)} \delta_{Z_n(t)}$$

is the empirical distribution of degrees in the Bianconi and Barabási continuous time tree at time t.

(a) Show that under Assumption 1 we have

$$\lim_{t\to\infty} \Theta_t = \nu \qquad \text{almost surely,}$$

where

$$\nu(k) = \int_0^1 \frac{\lambda^\star}{kx+\lambda^\star} \prod_{i=1}^{k-1} \frac{ix}{ix+\lambda^\star}\, d\mu(x).$$

(b) Show that $\lambda^\star \in (1,2)$ and that ν is a probability measure

(c) Show that $\nu(k) = k^{-(1+\lambda^\star)+o(1)}$ and hence the power law exponent ranges between the values 2 and 3, which is sometimes referred to as the supercritical regime.

Exercise 5.4: Size of the largest family.
Show that, in distribution as $t \to \infty$,

$$\mathsf{e}^{-\gamma(t-T(t))} \max_{n\in\mathbb{N}} Z_n(t) \to W^{-\frac{\gamma}{\lambda^\star}},$$

where W is exponentially distributed with parameter $\Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})(\lambda^\star)^{-\alpha}$

References

[BA99] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[BB01] Ginestra Bianconi and Albert-Laszlo Barabasi. Bose-Einstein condensation in complex networks. *Physical Review Letters*, 86:5632–5635, 2001.

[BCDR] Christian Borgs, Jennifer Chayes, Constantinos Daskalakis, and Sebastien Roch. First to market is not everything: An analysis of preferential attachment with fitness. In *STOC'07: Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 135–144. see `arXiv`:0710.4982 for a long version.

[BGT89] Nick Bingham, Charles Goldie, and Jozef Teugels. *Regular Variation.* Cambridge University Press, 1989.

[CCL13] Andrea Collevecchio, Codina Cotar, and Marco LiCalzi. On a preferential attachment and generalized Pólya's urn model. *Annals of Applied Probability*, 23:1219–1253, 2013.

[CHJ03] Fan Chung, Shirin Handjani, and Doug Jungreis. Generalizations of Pólya's urn problem. *Annals of Combinatorics*, 7:141–153, 2003.

[DM13] Steffen Dereich and Peter Mörters. Emergence of condensation in Kingman's model of selection and mutation. *Acta Applicandae Mathematica*, 127:17–26, 2013.

[DMM] Steffen Dereich, Cécile Mailler, and Peter Mörters. Non-extensive condensation in reinforced branching processes. `arXiv`:1601.08128 Preprint.

[GL10] Claude Godrèche and Jean-Marc Luck. On leaders and condensates in a growing network. *Journal of Statistical Mechanics: Theory and Experiment*, 07:P07031, 2010.

[HDRT15] Andrea Hodgins-Davis, Daniel P. Rice, and Jeffrey P. Townsend. Gene expression evolves under a house-of-cards model of stabilizing selection. *Molecular Biology and evolution*, 2015. doi: 10.1093/molbev/msv094.

[Jan04] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110:177–245, 2004.

[Kin78] John Kingman. A simple model for the balance between selection and mutation. *Journal of Applied Probability*, 15:1–12, 1978.

[Ner81] Olle Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 57:365–395, 1981.

[RTV07] Anna Rudas, Balint Toth, and Benedek Valko. Random trees and general branching processes. *Random Structures and Algorithms*, 31:186–202, 2007.

[vdBLP86] Michiela van den Berg, John T. Lewis, and Joe Pulé. A general theory of Bose-Einstein condensation. *Helvetica Physica Acta*, 59:1271–1288, 1986.