Branching processes with reinforcement

Cécile Mailler, Peter Mörters and Anna Senkevich

University of Bath

Minicourse Berlin, 31.8.-1.9.2017

Contents

CHAPTER 1

Branching processes with reinforcement: Definition and Examples

This course is about a class of models for the growth of a stochastic system which accelerates through a mechanism of reinforcement. We call these models *branching processes with reinforcement* or simply *reinforced branching processes*. We shall see that, due to the reinforcement, these systems display rather complex behaviour and interesting phenomena such as

- phase transitions,
- self-organised criticality,
- condensation,
- travelling waves,...

may occur. Some of these phenomena can be explained by techniques from the classical theory of branching processes (and we shall explore this in the first half of the course) others require new ideas and several problems are unsolved today.

Although our models can describe a variety of objects, see the examples below, we shall describe them as a structured population. Parameters of our model are a probability distribution μ on the positive reals, and positive numbers $\beta, \gamma \leq 1$ with $\beta + \gamma \geq 1$. At any time t the population consists of a finite number $N(t)$ of immortal individuals. Each individual in the population has a fitness, and individuals are organised into families, such that all members of a family have the same fitness.

The process is started with one family of one individual, whose fitness is drawn from the distribution μ. Suppose, at time $t \geq 0$, the population consists of $M(t)$ families, and there are $Z_n(t)$ individuals of fitness F_n in the nth family, for $1 \leq n \leq M(t)$. Independently, every individual gives birth with a rate given by its fitness, or equivalently in every family birth events occur with a time-dependent rate $F_n Z_n(t)$. When a birth event occurs in the nth family, independently of everything else, one or both of the following happen,

- with probability β a new family is founded, initially consisting of one individual equipped with a fitness drawn, independently of everything else, from the distribution μ;
- with probability γ a new individual with fitness F_n is added to the nth family.

Note that both things happen simultaneously with probability $\beta+\gamma-1 \geq 0$, the probability that only a new family is founded is $1-\gamma$, and the probability that only a new individual is added to the family is $1-\beta$. Under mild conditions on μ the total number $N(t)$ of individuals in the population remains finite at all times (see [**19**] for details) and this is the case we will be interested in. It would be easy to generalise this to the case where more than two particles can be born at a time, but we prefer to focus on the basic case which already contains all relevant features.

The reinforced branching process is described by the following family of random variables. For $t \geq 0$ we denote by

- $N(t)$ the total size of the population at time t,
- $M(t)$ the number of different families at time t,

- σ_n the time of the n^{th} birth event,
- τ_n the time of the foundation of the n^{th} family,
- $Z_n(t)$ the size of the n^{th} family at time t (if $n > M(t)$ we set $Z_n(t) = 0$), and
- F_n the fitness of the n^{th} family.

The *empirical fitness distribution* at time t is defined as the measure

$$\Xi_t = \frac{1}{N(t)} \sum_{n=1}^{M(t)} Z_n(t)\, \delta_{F_n}. \tag{1.1}$$

We now describe our three main examples motivating this definition.

Example 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 [**18, 13**]). 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 [**16**] for a discussion of the relevance of these models in the theory of evolution.

Example 2: *Preferential attachment tree of Bianconi and Barabasi.*

This model is originally a discrete time network model. Putting it into our framework means embedding it into continuous time, a technique heavily advocated by Janson [**17**], who attributes the method to Athreya and Karlin [**1**], and by Bhamidi [**6**]. The network is constructed successively, starting with one vertex which is formally given degree one. The vertex is given a fitness, randomly chosen according to μ. At every time step a new vertex is introduced, equipped with a fitness, randomly chosen according to μ, and linked to one of the existing vertices. The probability of an existing vertex being chosen is proportional to the product of its fitness and its degree at the time when the new vertex is introduced. As new vertices prefer to attach to existing vertices of high degree and high fitness, this is called a *preferential attachment* model.

In our representation we choose $\beta = \gamma = 1$ and observe the system at the birth times of individuals. We think of every family as a vertex in the network, and of the size of a family as its degree. Note that when the nth birth event takes place, it arises in each of the existing families with a probability proportional to the product of its fitness and its degree. At the birth event a new family is founded, i.e. a new vertex is introduced, and simultaneously the family that has given birth is increased in size by one, meaning that the degree of the corresponding vertex is incremented by one. Our representation only keeps track of the vertices and their degrees, not of the actual edges. But this does not matter as the main object of interest for us is the long-term behaviour of the degree-weighted fitness distribution, which coincides with the empirical fitness distribution in our framework. This model was introduced by Bianconi and Barabasi in [**4**] and further analysed by Borgs et al. [**7**].

Example 3: *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 [**17**] is looking at the case where μ is finitely supported. A related model has been studied by Chung et al. [**8**] who draw balls depending in a non-linear way on the distribution of colours in the urn, and by Collevecchio et al. [**9**] 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.

We now give a construction of our model on an explicit probability space. Let

- F be a μ-distributed random variable,
- given F the process $Y = (Y(t)\colon t \geq 0)$ be an independent Yule process with rate $F\gamma$,
- given F, Y two independent point processes $\Pi^{(1)}$ and $\Pi^{(2)}$ where
  - $\Pi^{(1)}$ jumps at every jump of Y independently with probability $\frac{\beta+\gamma-1}{\gamma}$ and
  - $\Pi^{(2)}$ is an inhomogeneous Poisson process with intensity measure $(1-\gamma)FY(t)\,dt$.

  Denote $\Pi = \Pi^{(1)} + \Pi^{(2)}$.

We let $(\Omega, \mathscr{F}, \mathbb{P})$ be the countable product of the joint law of (F, Y, Π) and denote the coordinate process by (F_n, Y_n, Π_n), for $n \in \mathbb{N}$. We let $\tau_1 = 0$ and $Z_1(t) = Y_1(t)$ and iteratively define

$$\tau_n = \inf\{t > \tau_{n-1} : \exists m \in \{1, \dots, n-1\} \text{ with } \Delta\Pi_m(t - \tau_m) = 1\}, \tag{1.2}$$

for $n \in \{2, 3, \dots\}$, and

$$Z_n(t) = \begin{cases} Y_n(t - \tau_n), & \text{if } t \geq \tau_n \\ 0, & \text{otherwise.} \end{cases}$$

We let $M(t) = \max\{n\colon \tau_n \leq t\}$, set

$$N(t) = \sum_{n=1}^{M(t)} Z_n(t),$$

and denote by $\sigma_1, \sigma_2, \dots$ the jump times of $(N(t)\colon t \geq 0)$. It is obvious that this construction defines the reinforced branching process as described in the introduction.

***Tutorial*: The Yule process**

The Yule process $(Y_t\colon t \geq 0)$ with rate η is a process of immortal particles starting with one particle. At any time every particle independently gives birth to a new particle with rate η. Y_t is the number of particles alive at time t. 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.

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

(a) Let $a > 0$ and show that $(Y_{at}\colon t \geq 0)$ is a Yule process with rate $a\eta$.
(b) Show that $(e^{-\eta t} Y_t\colon t \geq 0)$ is a martingale.
(c) Infer that there exists a random variable ξ such that, almost surely,
$$\lim_{t\uparrow\infty} e^{-\eta t} Y_t = \xi.$$
(d) Show that ξ is exponentially distributed with parameter one.
(e) Show that $\sup_{t\geq 0} \mathbb{E} e^{-2\eta t} Y_t^2 < \infty$.

Solution:

(a) Fix $a > 0$ and let $\widehat{Y}_t := Y_{at}$. Then, for all $t < \tau/a$, we have $\widehat{Y}_t = 1$, and for all $t \geq \tau/a$,
$$\widehat{Y}_t = Y_{at} = Y^{(1)}_{at-\tau} + Y^{(2)}_{at-\tau} = \widehat{Y}^{(1)}_{a(t-\tau/a)} + \widehat{Y}^{(2)}_{a(t-\tau/a)},$$
where $\widehat{Y}^{(1)}$ and $\widehat{Y}^{(2)}$ are two independent copies of $\widehat{Y}$. Note that the random variable $\hat{\tau} = \tau/a$ is exponentially distributed of parameter $a\eta$, implying that $\widehat{Y}$ is indeed a Yule process of parameter $a\eta$.

(b) Let us first calculate the expectation of Y_t for all $t \geq 0$. Using that, by definition,
$$Y_t = \mathbf{1}_{t<\tau} + \big(Y^{(1)}_{t-\tau} + Y^{(2)}_{t-\tau}\big)\,\mathbf{1}_{t>\tau},$$
where τ is exponentially distributed of parameter η, we get that
$$\mathbb{E}Y_t = e^{-\eta t} + \int_0^t 2\mathbb{E}Y_{t-u}\,\eta e^{-\eta u}\,\mathrm{d}u = e^{-\eta t} + 2\eta e^{-\eta t}\int_0^t \mathbb{E}Y_s\, e^{\eta s}\,\mathrm{d}s.$$
Thus, if we denote by $y(t) = e^{\eta t}\mathbb{E}Y_t$, we get, for all $t \geq 0$, $y'(t) = 2\eta y(t)$, implying that, since $y(0) = 1$, $y(t) = e^{2\eta t}$ for all $t \geq 0$. Thus, for all $t \geq 0$, $\mathbb{E}Y_t = e^{\eta t}$.

For all $s, t \geq 0$, using the Markov property,
$$\mathbb{E}\left[Y_{t+s}\,|\,\mathcal{F}_s\right] = \mathbb{E}\left[\sum_{i=1}^{Y_s} Y^{(i)}_t \,\Big|\, \mathcal{F}_s\right],$$
where the $Y^{(i)}$ are independent copies of Y, independent of Y_s. Thus,
$$\mathbb{E}\left[Y_{t+s}\,|\,\mathcal{F}_s\right] = Y_s\mathbb{E}Y_t = Y_s e^{\eta t},$$
implying that $(e^{-\eta t}Y_t)_{t\geq 0}$ is indeed a martingale.

(c) The martingale $(e^{-\eta t}Y_t)_{t\geq 0}$ is non-negative and thus converges almost surely to a random variable ξ.

(d) We define the random variables T_i as the succesive distances between successive jump times of the Yule process $(Y_t)_{t\geq 0}$. Then at time $T_1 + \cdots + T_{i-1}$, the Yule process is the sum of i independent copies of itself and each of them thus jumps after a random time of exponential law of parameter η. Thus, the time to wait before the next jump time is the minimum of these i random variables. T_i is thus exponentially distributed of parameter $i\eta$.

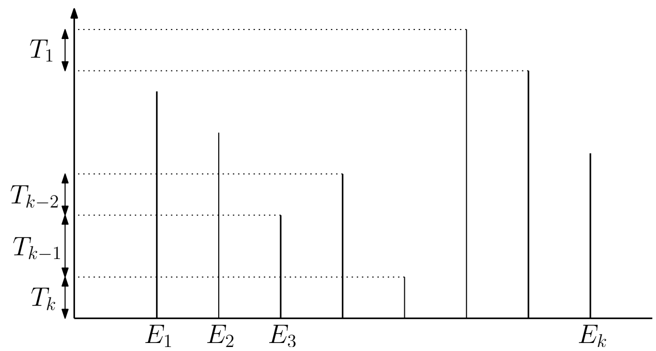

FIGURE 1. The k i.i.d. random variables $E_1, \dots, E_k$ are represented by the length of the vertical sticks. The T_i are independent random variables exponentially distributed, of respective parameters $i\eta$.

For all $t \geq 0$ and $k \in \mathbb{N}$, we have

$$\mathbb{P}(Y_t \geq k) = \mathbb{P}(T_1 + T_2 + \cdots + T_k \leq t) = \mathbb{P}\left(\max(E_1, \dots, E_k) \leq t\right),$$

where $(E_i)_{i\geq 1}$ is a sequence of i.i.d. exponential random variables of parameter η (see Figure 1 for an explanation of the last equality). Thus, $\mathbb{P}(Y_t \geq k) = (1 - e^{-\eta t})^k$ for all integers k and for all $t \geq 0$. It implies that

$$\mathbb{P}(e^{-\eta t} Y_t \geq x) = \mathbb{P}(Y_t \geq x e^{\eta t}) = (1 - e^{-\eta t})^{\lfloor x e^{\eta t} \rfloor} \to e^{-x},$$

when t goes to infinity, which concludes the proof.

(e) Using again the fact that $Y_t = \mathbf{1}_{t<\tau} + \left(Y^{(1)}_{t-\tau} + Y^{(2)}_{t-\tau}\right)\mathbf{1}_{t>\tau}$, we get that (we skip the details since it is very similar to the calculation of Y_t in the solution of (b))

$$y_2(t) := e^{\eta t}\mathbb{E}Y_t^2 = 1 + \int_0^t \left(2\mathbb{E}Y_s^2 + 2\big(\mathbb{E}Y_s\big)^2\right)\eta e^{\eta s}\mathrm{d}s = \frac{1}{3} + \frac{2e^{3\eta t}}{3} + 2\eta\int_0^t y_2(s)\,\mathrm{d}s,$$

because $\mathbb{E}Y_s = e^{\eta s}$ for all $s \geq 0$. We thus get that

$$y_2'(t) = 2\eta y_2(t) + 2\eta e^{3\eta t}, \qquad \text{and } y_2(0) = 1.$$

Solving this equation gives

$$y_2(t) = e^{\eta t}\mathbb{E}Y_t^2 = e^{2\eta t}\left(1 + 2\Big(e^{\eta t} - 1\Big)\right),$$

and thus

$$e^{-2\eta t}\mathbb{E}Y_t^2 = e^{-\eta t}\left(1 + 2\Big(e^{\eta t} - 1\Big)\right) \to 2,$$

when t goes to infinity, which implies the result.

CHAPTER 2

How fast does the system grow?

We let

$$w(\mu) := \sup\{x\colon \mu(-\infty, x) < 1\}$$

be the upper end point of μ. This value can be finite or infinite. To focus on the interesting cases we assume from now on that $\mu(\{w(\mu)\}) = 0$, i.e. there is no atom at the upper end point of μ.

There are two plausible scenarios for the growth of a reinforced branching process:

- **Growth driven by extremal behaviour**
  The size of a family with fitness F_n grows like a Yule process with rate γF_n. Hence
  $$\lim_{t\to\infty} \frac{1}{t} \log Z_n(t) = \gamma F_n,$$
  and the growth of the overall system is bounded from below by
  $$\lim_{t\to\infty} \frac{1}{t} \log N(t) \ge \sup_n \lim_{t\to\infty} \frac{1}{t} \log Z_n(t) = \sup_n \gamma F_n = \gamma w(\mu).$$
  In particular, if $w(\mu) = \infty$ the system grows superexponentially fast. But even if $w(\mu) < \infty$ the lower bound may be accurate and the growth of the system hence determined by the families with record fitness. Note that it is unclear at which time a family born at τ_n with a record fitness will have grown to contribute the largest proportion in the overall population.

- **Growth driven by bulk behaviour**
  The alternative scenario is that the empirical fitness distribution Ξ_t stabilizes near some probability distribution $f d\mu$. For this to happen we need $w(\mu) < \infty$ and also f to satisfy an eigenvalue equation characterising how the fitness distribution is maintained in an infinitesimal step of the Markov process. We now give a heuristic calculation which shows how this could happen. We assume without loss of generality that $w(\mu) = 1$.

  The operator $A\colon C(0,1) \to C(0,1)$ which describes how the fitness distribution is moved in an infinitesimal step is given by
  $$Af(x) = x\big(\gamma f(x) + \beta \int f(y)\mu(dy)\big).$$
  We have the eigenvalue equation
  $$\begin{aligned} Af = \lambda^* f &\Leftrightarrow f(x) = \frac{\beta x}{\lambda^* - \gamma x} \int f\, d\mu \\ &\Rightarrow \lambda^* \ge \gamma \text{ with } 1 = \beta \int \frac{x}{\lambda^* - \gamma x}\, \mu(dx) \\ &\Leftrightarrow \frac{\beta}{\gamma} \int \frac{x}{1-x}\, \mu(dx) \ge 1 \Leftrightarrow \frac{\beta}{\beta+\gamma} \int \frac{1}{1-x}\, \mu(dx) \ge 1. \end{aligned}$$

Hence we expect growth driven by bulk behaviour only under the condition

$$\frac{\beta}{\beta+\gamma}\int \frac{1}{1-x}\,\mu(dx) \geq 1,$$

and in this case the process grows with rate λ^* given as the unique solution of the equation

$$1 = \beta \int \frac{x}{\lambda^* - \gamma x}\,\mu(dx).$$

The eigenfunction is

$$f(x) = \frac{\beta x}{\lambda^* - \gamma x},$$

but the density of the asymptotic fitness distribution is given by the eigenmeasure

$$\nu(dx) = \frac{\beta}{\beta+\gamma}\frac{\lambda^*}{\lambda^* - \gamma x}\,\mu(dx)$$

of the dual operator A^* turns out to be the asymptotic fitness distribution.

We will later discuss the two scenarios in detail. Let us start with a couple of soft results that hold independently of the growth scenario.

THEOREM 2.1. *Almost surely, as $t \uparrow \infty$, we have*

(a) $\displaystyle\lim_{t\uparrow\infty} \frac{M(t)}{N(t)} = \frac{\beta}{\beta+\gamma}.$

(b) *If* $\displaystyle\int x\,\Xi_t(dx) \to m \in [0,\infty]$, *then* $\displaystyle\frac{1}{t}\log N(t) \to (\beta+\gamma)m.$

PROOF. (a) At every birth event the type of birth is chosen independently. Hence, by the law of large numbers almost surely $\lim_{n\to\infty} \frac{1}{n} M(\sigma_n) = \beta$, and $\lim_{n\to\infty} \frac{1}{n} N(\sigma_n) = \beta+\gamma$, where the right hand side is the expected number of families, resp. individuals, created at a birth event. Hence

$$\lim_{n\to\infty} \frac{M(\sigma_n)}{N(\sigma_n)} = \frac{\beta}{\beta+\gamma}.$$

The result follows as the process $(M(t)/N(t) \colon t > 0)$ is piecewise constant with jumps occuring only at the times σ_n, for $n \in \mathbb{N}$.

(b) Given the population at the time σ_n of the nth birth event, the waiting time $\sigma_{n+1} - \sigma_n$ until the next individual is born is exponentially distributed with rate given by the sum of the fitnesses in the population, i.e. $N(\sigma_n) \int x\,\Xi_{\sigma_n}(dx) \sim n(\beta+\gamma)m$. Hence

$$\sigma_n = \sum_{i=1}^{n} (\sigma_i - \sigma_{i-1}) \sim \frac{1}{(\beta+\gamma)m}\log n$$

and, in particular, we obtain, almost surely,

$$\lim_{t\uparrow\infty} \frac{1}{t}\log N(t) = \lim_{n\to\infty} \frac{1}{\sigma_n}\log N(\sigma_n) = (\beta+\gamma)m \lim_{n\to\infty} \frac{\log N(\sigma_n)}{\log n} = (\beta+\gamma)m.$$

□

In the next chapter we investigate the precise growth behaviour in the case of bulk driven behaviour using tools from the theory of general branching processes.

CHAPTER 3

Growth driven by bulk behaviour

In this chapter we study the bulk driven phase using tools from the theory of general branching processes. The results here are adapted from Nerman [**21**] who uses a more general framework.

Recall that our process is defined using a sequence (F_n, Y_n, Π_n) of independent random variables where

- F_n is a μ-distributed random variable,
- given F_n the process $Y_n = (Y_n(t) \colon t \geq 0)$ is an independent Yule process with rate γF_n,
- given F_n, Y_n the point process $\Pi_n = (\Pi_n(t) \colon t \geq 0)$ with intensity measure

$$\left(\tfrac{\beta+\gamma-1}{\gamma}\, \delta Y_n(t) + (1-\gamma) F_n Y_n(t)\, dt \colon t \geq 0\right).$$

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 new families created from this family. For greater generality we enrich this triple (F, Y, Π) by a fourth component $\phi = (\phi(t) \colon t \geq 0)$, a cadlag process taking values in $\mathbb{N}_0$ assigning some kind of score to the family t time units after its foundation. In all our examples (below) ϕ is a function of (F, Y, Π) but this does not have to be the case. We use the convention that $\phi(t) = 0$ if $t < 0$.

Denote by $\mathscr{G}_n$ the σ-algebra generated by $(F_1, Y_1, \Pi_1, \phi_1), \ldots, (F_n, Y_n, \Pi_n, \phi_n)$. 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 is $\mathscr{G}_{n-1}$-measurable. 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

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

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 Barabasi 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 $\lambda^* > \gamma$, called the* Malthusian parameter*, such that*

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

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*

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

We define

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

We can now formulate the main result on convergence of general branching processes in the case of reinforced branching processes.

Theorem 3.1. *Suppose Assumptions 1 and 2 hold. Then there exists a positive random variable W, not depending on ϕ, such that*

$$\lim_{t\uparrow\infty} e^{-\lambda^* t} Z_t^\phi = W\, m_\infty^\phi \qquad \textit{almost surely.}$$

Sketch of Theorem 3.1. The proof comes in three steps:

- **(i)** Show convergence for a special choice of ϕ using a martingale argument.
- **(ii)** Show that the resulting limit W is positive almost surely.
- **(iii)** Approximate general ϕ by our special choice.

(i) Let $\mathscr{I}(t)$ be the set of families born after time t from families born before time t, formally

$$\mathscr{I}(t) = \bigcup_{\tau_n < t} \{m\colon \tau_m = \tau_n + s \in [t,\infty)\colon \Delta\Pi_n(s) = 1\}.$$

We put

$$W_t = \sum_{m \in \mathscr{J}(t)} e^{-\lambda^* \tau_m} = \sum_{\tau_n < t} \sum_{\substack{\Delta\Pi_n(s)=1: \\ \tau_n + s \geq t}} e^{-\lambda^*(\tau_n + s)} = e^{-\lambda^* t} Z_t^{\phi},$$

where

$$\phi_n(t) = \sum_{\substack{\Delta\Pi_n(s)=1: \\ s \geq t}} e^{-\lambda^*(s-t)}.$$

We now show convergence for this particular score function. Let $R_0 = 1$ and, for $n \in \mathbb{N}$,

$$R_n = 1 + \sum_{m=1}^{n} e^{-\lambda^* \tau_m} \Big(\sum_{\Delta\Pi_m(s)=1} e^{-\lambda^* s} - 1 \Big).$$

We rewrite

$$R_n = 1 + \sum_{m=1}^{n} \sum_{\Delta\Pi_m(s)=1} e^{-\lambda^*(\tau_m + s)} - \sum_{m=1}^{n} e^{-\lambda^* \tau_m},$$

which shows that R_n is a weighted sum of the direct descendants of the first n families, excluding the families indexed $2, \ldots, n$. We hence have

$$(W_t \colon t \geq 0) = (R_{M(t)} \colon t \geq 0).$$

Lemma 3.2. *$(R_n \colon n \geq 0)$ is a martingale with respect to the filtration $(\mathscr{G}_n \colon n \geq 0)$.*

Proof. Recall that R_n and τ_{n+1} are $\mathscr{G}_n$-measurable. Moreover,

$$\sum_{\Delta\Pi_{n+1}(s)=1} e^{-\lambda^* s}$$

is independent of $\mathscr{G}_n$. Hence

$$\mathbb{E}\big[R_{n+1} - R_n \big| \mathscr{G}_n\big] = e^{-\lambda^* \tau_{n+1}} \, \mathbb{E}\Big[\sum_{\Delta\Pi_{n+1}(s)=1} e^{-\lambda^* s} - 1 \Big] = 0,$$

by definition of the Malthusian parameter. □

As $(R_n \colon n \geq 0)$ is a non-negative martingale it converges, almost surely, to a nonnegative limit variable W. We then have, almost surely,

$$\lim_{t \uparrow \infty} W_t = \lim_{t \uparrow \infty} R_{M(t)} = \lim_{n \to \infty} R_n = W.$$

(ii) To see that W is positive almost surely, one relies on a *Kesten-Stigum type* theorem, which shows

$$W > 0 \text{ almost surely } \Leftrightarrow \quad \mathbb{E}[X \log_+ X] < \infty,$$

for

$$X = \sum_{\Delta\Pi(s)=1} e^{-\lambda^* s}.$$

An elegant proof of such a result for general branching processes is given by Olofson in [**22**]. It is then easy (but lengthy) to check that in our model always $\mathbb{E}[X^2] < \infty$, which implies the result.

(iii) Given two scores ϕ and ψ we need to show that almost surely

$$\frac{Z_t^{\phi}}{Z_t^{\psi}} \longrightarrow \frac{\hat{\phi}(\lambda^*)}{\hat{\psi}(\lambda^*)},$$

where $\hat\phi(\lambda^*)=\int_0^\infty e^{-\lambda^* t}\mathbb{E}\phi(t)\,dt$. We drop the considerable technicalities of this step and only give an informal argument confirming that

$$\mathbb{E}Z_t^\phi \sim m_\infty^\phi e^{\lambda^* t},$$

and hence

$$\frac{\mathbb{E}Z_t^\phi}{\mathbb{E}Z_t^\psi} \longrightarrow \frac{\hat\phi(\lambda^*)}{\hat\psi(\lambda^*)},$$

making the form of the limit plausible.

We let (t_j) be the sequence of birth times of mutant offspring of the first family (ie. the jump times of the point process Π_1) and write

$$Z_t^\phi = \phi_1(t) + \sum_{t_j\le t} Z_{t-t_j}^\phi .$$

Writing $m_t^\phi = \mathbb{E}Z_t^\phi$ and taking expectations we get the renewal equation

$$m_t^\phi = \mathbb{E}\phi(t) + \int_0^t m_{t-s}^\phi \,\mathbb{E}\Pi(ds).$$

Taking Laplace transforms gives

$$\hat m(\lambda) = \hat\phi(\lambda) + \hat m(\lambda)\,\mathbb{E}\hat\Pi(\lambda),$$

and hence informally

$$\hat m(\lambda) = \frac{\hat\phi(\lambda)}{1-\mathbb{E}\hat\Pi(\lambda)}.$$

As

$$\partial_\lambda \mathbb{E}\hat\Pi(\lambda^*) = -\int_0^\infty t e^{-\lambda^* t}\,\mathbb{E}\Pi(dt) < 0,$$

the function $1-\mathbb{E}\hat\Pi(\lambda)$ has a simple zero at λ^*. Hence we have a Laurent expansion

$$\hat m(\lambda) = \frac{\hat\phi(\lambda^*)}{-\partial_\lambda \mathbb{E}\hat\Pi(\lambda^*)}\,(\lambda-\lambda^*)^{-1} + o\big((\lambda-\lambda^*)^{-1}\big).$$

Inverting the Laplace transform termwise (as explained in [**24**, 21.14]) gives

$$m_t^\phi = \frac{\hat\phi(\lambda^*)}{-\partial_\lambda \mathbb{E}\hat\Pi(\lambda^*)}\,e^{\lambda^* t} + o(e^{\lambda^* t}),$$

and hence we get the required formula for $m_\infty^\phi = \lim e^{-\lambda^* t}\mathbb{E}Z_t^\phi$. $\square$

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 $\lambda^*\ge\gamma$,

$$\begin{aligned}\int_0^\infty e^{-\lambda^* s}\,\mathbb{E}\Pi(ds) &= \int \mu(df)\Big\{\tfrac{\beta+\gamma-1}{\gamma}\int_0^\infty e^{-\lambda^* s}\,de^{\gamma f s} + (1-\gamma)f\int_0^\infty e^{-\lambda^* s}e^{\gamma f s}\,ds\Big\}\\ &= \beta\int f\int_0^\infty e^{-\lambda^* s+\gamma f s}\,ds\,\mu(df)\\ &= \beta\int \frac{f}{\lambda^*-\gamma f}\,\mu(df).\end{aligned}$$

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

$$\frac{\beta}{\gamma}\int\frac{f}{1-f}\,\mu(df),$$

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

$$\frac{\beta}{\gamma}\int \frac{f}{1-f}\,\mu(df) > 1,$$

or, equivalently,

$$\frac{\beta}{\beta+\gamma}\int \frac{\mu(df)}{1-f} > 1. \tag{3.3}$$

This condition identifies the regime of bulk growth, as predicted in Chapter 2. The Malthusian parameter λ^* is then defined by the equation

$$\beta\int \frac{f}{\lambda^*-\gamma f}\,\mu(df) = 1. \tag{3.4}$$

Now we look at the examples of scores ϕ we listed earlier and harvest the results.

Under the condition of bulk growth (3.3) we get the following **results**.

(1) Almost surely,

$$\lim_{t\uparrow\infty} e^{-\lambda^* t}N(t) = W m_\infty^{\phi^{(1)}}.$$

To confirm this result we check that Assumption 2 holds for $\phi^{(1)}$. This follows using Doob's maximal inequality and the results of the tutorial for a standard Yule process $(Y_s\colon s>0)$ as

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

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

$$\int \mu(df)\int_0^\infty e^{-\lambda^* t}\mathbb{E}\phi^{(1)}(t)\,dt = \int \frac{\mu(df)}{\lambda^*-\gamma f} = \frac{\beta+\gamma}{\lambda^*\beta}.$$

(2) Almost surely,

$$\lim_{t\uparrow\infty} e^{-\lambda^* t}M(t) = W m_\infty^{\phi^{(2)}}.$$

To compare with (1) we calculate the score dependent numerator of $m_\infty^{\phi^{(2)}}$. We get $\int_0^\infty e^{-\lambda^* t}\,dt = \frac{1}{\lambda^*}$. This is in line with the result of Theorem 2 which shows that there is an asymptotic factor $(\beta+\gamma)/\beta$ between the number of individuals $N(t)$ and the number of families $M(t)$.

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

$$\frac{\lambda^*\beta}{\beta+\gamma}\int \mu(df)\int_0^a e^{-\lambda^* t+\gamma f t}\,dt.$$

This limit goes to one as $a\uparrow\infty$, which shows that most individuals come from recently established families, which confirms that in this phase the growth is indeed bulk driven. There is a good heuristic explanation for this formula. t time units before observation time the population had a proportion $e^{-\lambda^* t}$ of its current size. Mutants then generated with fitness f then grow families to size $e^{\gamma f t}$.

(4) Almost surely,

$$\lim_{t\uparrow\infty} \Xi_t[1-x,1] = \frac{\lambda^*\beta}{\beta+\gamma}\int_{1-x}^{1} \frac{1}{\lambda^*-\gamma f}\,\mu(df).$$

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 the density

$$\frac{\beta}{\beta+\gamma}\,\frac{\lambda^*}{\lambda^*-\gamma f}.$$

We look at Example (5) in the tutorial. 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. [**26**].

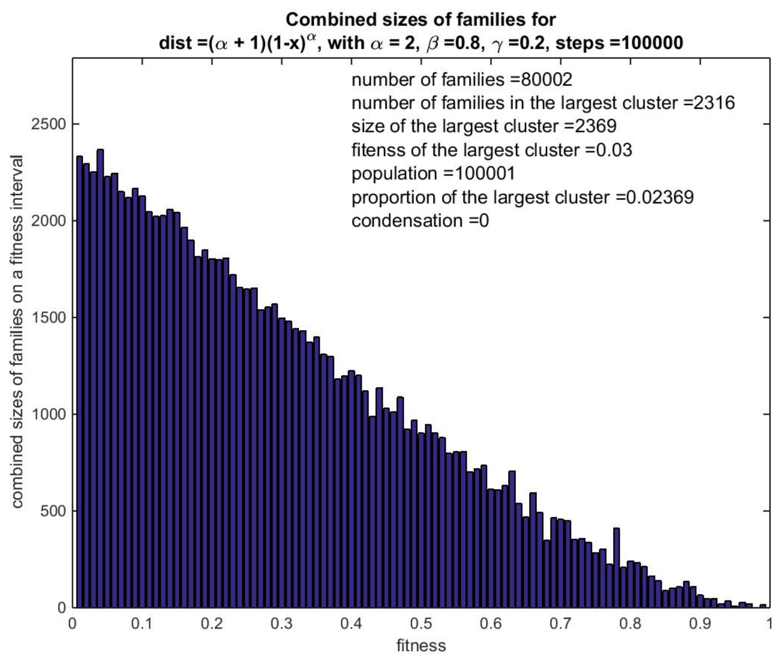

Figure 1. Empirical fitness distribution in a bulk-driven example. Simulation by Anna Senkevich.

***Tutorial*: The empirical degree distribution in the Bianconi-Barabasi Tree**

Assume we are in the situation of Example 2. Then

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

is the empirical distribution of degrees in the network at time t.

Problem:

(a) Show that under Assumption 1 we have

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

where

$$\nu(k) = \int_0^1 \frac{\lambda^*}{kf + \lambda^*} \prod_{i=1}^{k-1} \frac{if}{if + \lambda^*} \, \mu(df).$$

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

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

Solution:

(a) Combining examples (2) and (5) we get from Theorem 3.1 that

$$\lim_{t\uparrow\infty} \Theta_t(k) = \lambda^* \int_0^\infty e^{-\lambda^* t} \, \mathbb{P}(Y(t) = k) \, dt,$$

where $(Y(t) \colon t > 0)$ is a Yule process with random parameter F. We use the notation of the first tutorial to write

$$\mathbb{P}(Y(t) = k) = \mathbb{P}(T_1 + \cdots + T_{k-1} < t) - \mathbb{P}(T_1 + \cdots + T_k < t),$$

where T_j is exponential with parameter jF. Now

$$\begin{aligned} \int_0^\infty e^{-\lambda^* t} \, \mathbb{P}(T_1 + \cdots + T_k < t) \, dt &= \mathbb{E} \int_{T_1+\cdots+T_k}^\infty e^{-\lambda^* t} \, dt = \frac{1}{\lambda^*} \int \mu(df) \prod_{i=1}^{k} \mathbb{E}\big[e^{-\lambda^* T_i} \big| F = f\big] \\ &= \frac{1}{\lambda^*} \int \mu(df) \prod_{i=1}^{k} \frac{1}{1 + \frac{\lambda^*}{if}}. \end{aligned}$$

Hence,

$$\begin{aligned} \lambda^* \int_0^\infty e^{-\lambda^* t} \, \mathbb{P}(Y(t) = k) \, dt &= \int \mu(df) \Big(\prod_{i=1}^{k-1} \frac{if}{if + \lambda^*} - \prod_{i=1}^{k} \frac{if}{if + \lambda^*} \Big) \\ &= \int \mu(df) \, \frac{\lambda^*}{kf + \lambda^*} \prod_{i=1}^{k-1} \frac{if}{if + \lambda^*}. \end{aligned} \tag{3.5}$$

Observe that if ν is identified as a probability measure then the convergence holds automatically in the stronger total variation sense.

(b) λ^* is the unique solution of the equation

$$\int \frac{f}{\lambda - f}\,\mu(df) = 1.$$

The left hand side is monotonically decreasing in λ and takes a value > 1 for $\lambda = 1$ and a value < 1 for $\lambda = 2$. Hence the solution lies in the interval $(1, 2)$. Summing over $k = 1, 2, \ldots$ in (3.5) shows that ν is a probability measure.

(c) Note that, for $k > n$,

$$\log \prod_{i=n}^{k-1} \frac{if}{if + \lambda^*} = \sum_{i=n}^{k-1} \log \frac{1}{1 + \frac{\lambda^*}{if}} = -(1 + o_n(1)) \sum_{i=n}^{k-1} \frac{\lambda^*}{if} = -(1 + o_n(1)) \frac{\lambda^*}{f} \Big(\log \big(\tfrac{k}{n}\big) + o_n(1) \Big).$$

We infer (without spelling out all details here) that for large k the main contribution to the integral comes from values of f close to one and that therefore

$$\int \mu(df)\, \frac{\lambda^*}{kf + \lambda^*} \prod_{i=1}^{k-1} \frac{if}{if + \lambda^*} = k^{-(1+\lambda^*)+o_k(1)}.$$

Remark: This result holds (for a suitable choice of λ^*) without Assumption 1 and also extends, with a completely different proof, to a class of Bianconi-Barabasi networks which are typically not trees, see Corollary 2.8 in Dereich and Ortgiese [**14**].

CHAPTER 4

Growth driven by extremal behaviour

We have seen in Chapter 3 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 form (1.1),

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

THEOREM 4.1. *Assume that*

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

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

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

(ii) $\Xi_t \to \pi$ *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 4.1 Combining (i) with Theorem 2.1(b) shows that

$$\lim_{t\uparrow\infty} \frac{1}{t} \log N(t) = \gamma,$$

which is indeed the scenario of growth driven by extremal behaviour. 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. We will investigate further properties of this behaviour in Chapter 5 and state some open problems related to it in Chapter 6.

The proof we develop here uses the results proved in Chapter 3: 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 process, 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 $F_1 \mathbf{1}_{F_1 < 1-\varepsilon} + \mathbf{1}_{F_1 \geq 1-\varepsilon}$ 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 $f\mathbf{1}\{f < 1-\varepsilon\} + \mathbf{1}\{f \geq 1-\varepsilon\}$ 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 $f < 1-\varepsilon$ is born at time $n+1$ in the original process, then if
$$U_{n+1} \leq \left(\frac{\widehat{\Xi}_n^{(\varepsilon)}(\{f\})}{\int_0^1 x \,\mathrm{d}\widehat{\Xi}_n^{(\varepsilon)}(x)}\right)\left(\frac{\widehat{\Xi}_n(\{f\})}{\int_0^1 x \,\mathrm{d}\widehat{\Xi}_n(x)}\right)^{-1},$$
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 Chapter 3. 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 $\varepsilon \downarrow 0$. Using the techniques developed in Chapter 3, we deduce that, for all $0 \leq a, b < 1-\varepsilon$, 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 $0 \leq a, b < 1$ and $0 < \varepsilon < 1-b$, we thus have

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

Letting $\varepsilon \downarrow 0$ concludes the proof of the lower bound.

The coupling of the processes (upper bound).

Fix $\varepsilon > 0$, and let $(\Xi_t^{[\varepsilon]})_{t\geq 0}$ 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 $F_n^{[\varepsilon]}$ 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 $(V_n)_{n\geq 1}$ be a sequence of i.i.d. random variables uniformly distributed on $[0,1]$ and $(W_n, W_n')_{n\geq 1}$ 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, $\widehat{\Xi}_0 = \delta_{F_1}$, where $F_1 = F_1^{[\varepsilon]}\mathbf{1}\{F_1^{[\varepsilon]} < 1-\varepsilon\} + W_1\mathbf{1}\{F_1^{[\varepsilon]} = 1-\varepsilon\}$ 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 f is born at time $n+1$ in the ε-truncated process, then
$$\widehat{\Xi}_{n+1} = \widehat{\Xi}_n + \delta_f \mathbf{1}\{f < 1-\varepsilon\} + W_{n+1}\mathbf{1}\{f = 1-\varepsilon\};$$
- 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 $f < 1-\varepsilon$ is born at time $n+1$ in the ε-truncated process, then if
$$V_{n+1} \leq \frac{\Xi_n\big(\{f\}\big)}{\int_0^1 x \,\mathrm{d}\widehat{\Xi}_n(x)} \left(\frac{\widehat{\Xi}_n^{[\varepsilon]}\big(\{f\}\big)}{\int_0^1 x\,\mathrm{d}\widehat{\Xi}_n^{[\varepsilon]}(x)}\right)^{-1},$$
then $\Xi_{n+1} = \Xi_n + \delta_f$, 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 $(\widehat{\Xi}_n)_{n\geq 0}$ has indeed the same law as the empirical fitness distribution of the original reinforced branching process.

We get that, for all $0 < a < b < 1-\varepsilon$,
$$\limsup_{n\to\infty} \widehat{\Xi}_n(a,b) \leq \lim_{n\to\infty} \widehat{\Xi}_n^{[\varepsilon]}(a,b) = \frac{\beta}{\beta+\gamma}\int_a^b \frac{\lambda^{[\varepsilon]}}{\lambda^{[\varepsilon]} - \gamma x}\, d\mu^{[\varepsilon]}(x).$$

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 4.1.

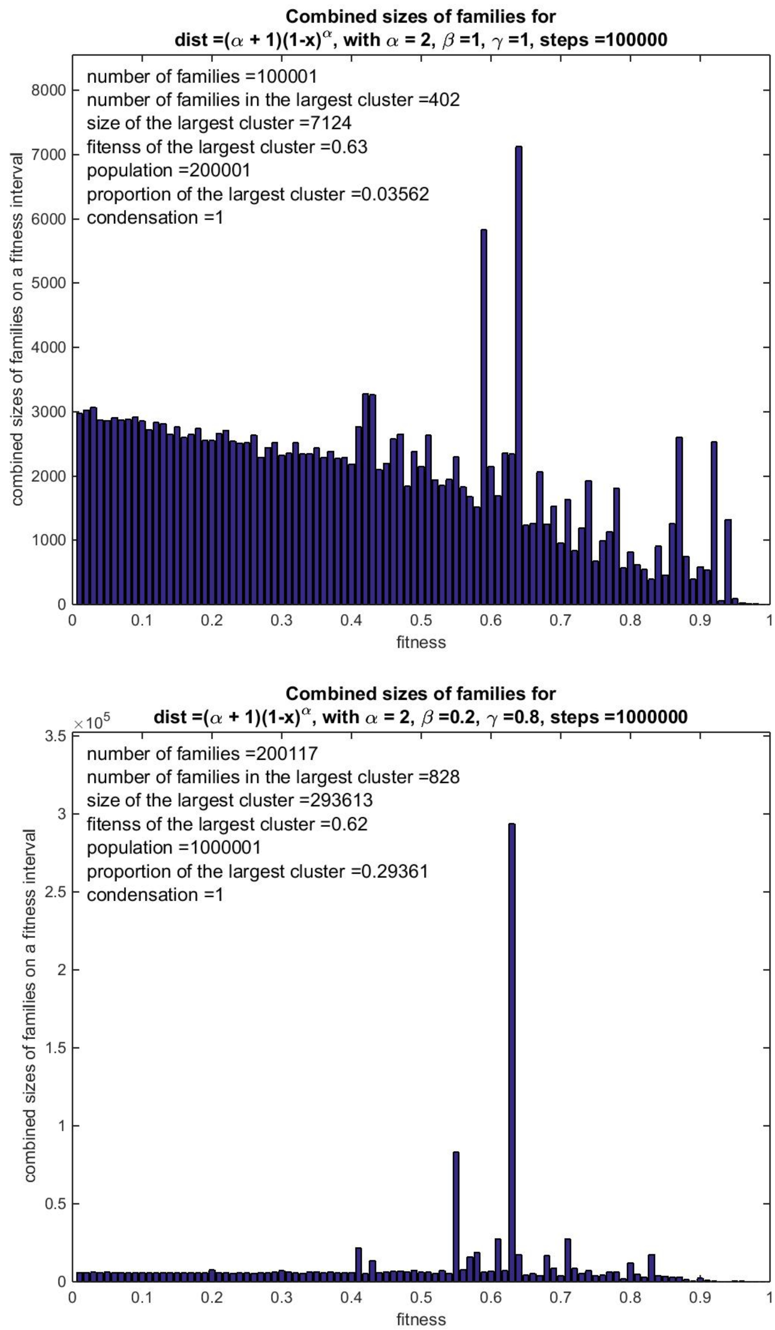

FIGURE 1. Empirical fitness distribution in two examples with condensation. Simulations by Anna Senkevich.

CHAPTER 5

Largest families and related problems

In this chapter we follow [**11**] 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

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

or equivalently

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

for a slowly varying function ℓ and some $\alpha > 1$, see [**5**]. This corresponds to the most common type of behaviour of μ at its tip that allows a condensation phase. Other cases can be studied and will be covered in a forthcoming paper [**20**].

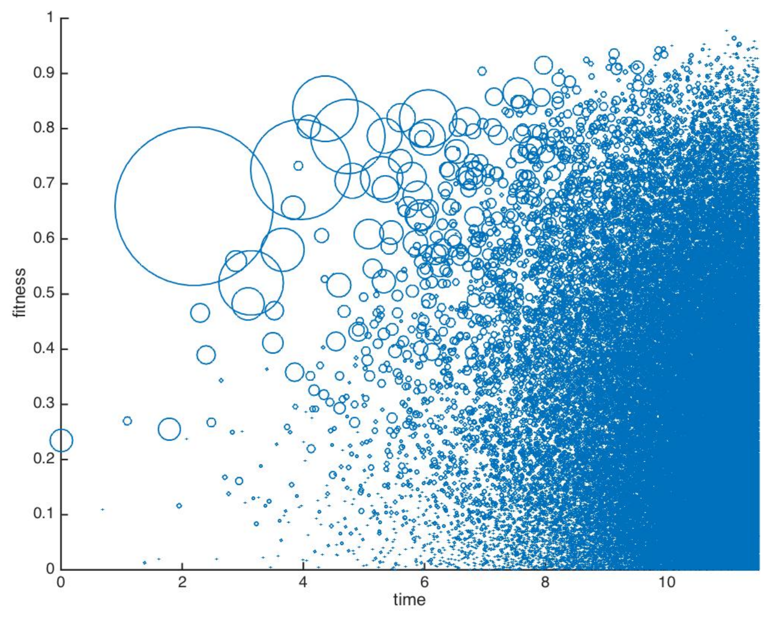

FIGURE 1. A simulation of a reinforced branching process in the condensation case. Parameters are $\mu(dx) = 3(1-x)^2\,dx$ and $\beta = \gamma = 1$. Each family is represented by a circle with area proportional to its size at time $t = 12$ and centred at its time of birth (horizontal axis) and its fitness (vertical axis). Simulation courtesy of Anna Senkevich.

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 ${}^1\!/\!{}_{\lambda^\star} \log n(t) \sim \frac{\alpha}{\lambda^*} \log t$;
- the largest families at t have fitness F_n with $1 - F_n$ of order $1/t$ and size of order $e^{\gamma(t-T(t))}$.

To confirm our intuition we consider the point process

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

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

Theorem 5.1 (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*

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

Remark 5.1 Note the compactifications at $\pm\infty$ in Theorem 5.1. As the limiting point process has a continuous density, Theorem 5.1 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 5.2 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, see also Chapter 6.

Corollary 5.2 (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, as $t \to \infty$,*

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

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

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

Remark 5.3 The birth time of the family of maximal size at time t is of asymptotic order $T(t)+O(1)$ and hence of leading order ${}^\alpha\!/\!{}_{\lambda^\star} \log t$. This answers a question of Borgs et al. [**7**] about the rate at which new nodes with higher fitness become the leading influence in the population. We prove Corollary 5.2 and give further details of the limit laws in the tutorial.

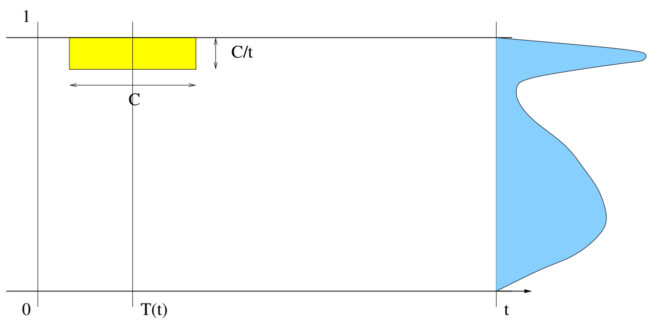

Figure 2. The largest family at time t is born at the time and with a fitness from the yellow window, with high probability as C becomes large.

A further problem that can be solved using Theorem 5.1 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. [**3**] 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 [**3**] .
- 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. [**3**] have shown that this occurs in the free Bose gas for an intermediate temperature range.

We shall see in Theorem 5.3 that in our model under a natural assumption on μ the second scenario prevails. To show this we need to investigate the behaviour of the largest family in our system.

Theorem 5.3 (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 5.4 Theorem 5.3 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 [**4**], 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 [**15**] who use a nonrigorous 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 5.3. Subject to a cut-off argument we have in view of Theorem 5.1,

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

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 $\lambda*\geq\gamma$, we get

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

From this we conclude that

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

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

Sketch of Theorem 5.1. We start by looking at the first component of Γ_t. The main difficulty of our model is that the time τ_n of birth of the nth family is not known with good accuracy. It is therefore important the yellow box has only constant width, as from our logarithmic growth rates we get a rough bound for the births occuring around the stopping times $T(t)=\tau_{n(t)}$, which is sufficiently accurate on a constant time scale.

Lemma 5.4. *For all $\varepsilon>0$, we have with high probability as $t\uparrow\infty$, for all $n\in\mathbb{N}$,*

$$\frac{1}{\lambda^\star\pm\varepsilon}\log\frac{n}{n(t)}-\varepsilon\leq\tau_n-T(t)\leq\frac{1}{\lambda^\star\mp\varepsilon}\log\frac{n}{n(t)}+\varepsilon,$$

where the sign in $\pm$ is the same as that of the logarithm.

How do we get a Poisson limit with the given density?

Recall that $n(t)=\lceil\mu(1-\frac{1}{t},1)^{-1}\rceil$, then

- by Lemma 5.4 points are spaced at $\tau_n-T(t)\sim\frac{1}{\lambda^*}\log\left(\frac{n}{n(t)}\right)$,
- by regular variation

$$n(t)\mu[1-\tfrac{x}{t},1]\to x^\alpha=\int_0^x\alpha f^{\alpha-1}\,df.$$

An extreme value calculation gives, for $a_0<a_1$ and $0<b_0<b_1$ and $B=(a_0,a_1)\times(b_0,b_1)$ that

$$\begin{aligned}\mathbb{P}(\Gamma_t(B)=0)&=\mathbb{P}\big(t(1-F_n)\notin(b_0,b_1)\ \forall n\text{ with }\tau_n-T(t)\in(a_0,a_1)\big)\\&=\prod_{n:\ \frac{1+o(1)}{\lambda^*}\log\frac{n}{n(t)}\in(a_0,a_1)}\mathbb{P}\Big(F_n\notin\big(1-\tfrac{b_1}{t},1-\tfrac{b_0}{t}\big)\Big)\\&=\Big(1-\mu\big(1-\tfrac{b_1}{t},1-\tfrac{b_0}{t}\big)\Big)^{n(t)(e^{\lambda^*(a_1+o(1))}-e^{\lambda^*(a_0+o(1))})}\\&=\Big(1-\frac{b_1^\alpha-b_0^\alpha}{n(t)}\Big)^{n(t)(e^{\lambda^*(a_1+o(1))}-e^{\lambda^*(a_0+o(1))})}\\&\longrightarrow\exp(-(b_1^\alpha-b_0^\alpha)(e^{\lambda^*a_1}-e^{\lambda^*a_0}))=e^{-\zeta_0(B)},\end{aligned}$$

where $\zeta_0(ds\,df)=\lambda^*e^{\lambda^*s}\,ds\otimes\alpha f^{\alpha-1}\,df$. Similarly,

$$\begin{aligned}\mathbb{E}\Gamma_t(B) &= \sum_{n(t)(e^{\lambda^*(a_0+o(1))}}^{n(t)(e^{\lambda^*(a_1+o(1))})} \mathbb{P}\big(t(1-F_n)\in(b_0,b_1)\big)\\ &= n(t)(e^{\lambda^*(a_1+o(1))}-e^{\lambda^*(a_0+o(1))})\,\mu\big(1-\tfrac{b_1}{t},1-\tfrac{b_0}{t}\big)\\ &\longrightarrow (b_1^\alpha-b_0^\alpha)(e^{\lambda^* a_1}-e^{\lambda^* a_0})=\zeta_0(B).\end{aligned}$$

By a celebrated theorem of Kallenberg [**25**, Proposition 3.22] this is enough to ensure convergence of the point process to the Poisson limit. More precisely, we obtain

$$\sum_{n=1}^{M(t)} \delta\big(\tau_n - T(t), t(1-F_n)\big) \Longrightarrow \Pi_{\zeta_0}.$$

We have

$$Z_n(t) = Y(F_n\gamma(t-\tau_n)) \quad \text{ for } t\geq \tau_n,$$

for a standard Yule process Y independent of F_n and τ_n. Hence, by Tutorial 1,

$$e^{-\gamma F_n(t-\tau_n)} Z_n(t) \longrightarrow X,$$

where X is standard exponential independent of F_n and τ_n. Therefore, if the nth point has initial components s, f, the third component becomes

$$\begin{aligned}e^{-\gamma(t-T(t))}Z_n(t) &= e^{-\gamma(\tau_n-T(t))}\,e^{-\gamma(t-\tau_n)}Z_n(t)\\ &= e^{-\gamma(\tau_n-T(t))}\,e^{-\gamma(t-\tau_n)(1-F_n)}\,e^{-\gamma(t-\tau_n)F_n}Z_n(t)\\ &\longrightarrow e^{-\gamma s}\,e^{-\gamma f}\,X,\end{aligned}$$

Hence the third component is exponentially distributed with mean $e^{-\gamma(s+f)}$. It therefore contributes a density of

$$e^{-ze^{\gamma(s+f)}}\,e^{\gamma(s+f)}\,dz.$$

Why does the biggest family originate from this box?

The yellow box is centred around

$$T(t) = \inf\{s>0\colon M(s)\geq n(t)\} = \inf\{s>0\colon \Xi_s(1-\tfrac{1}{t},1)>0\}+O(1).$$

If the nth family is born at time τ_n, its size is

$$Z_n(t)\sim e^{\gamma F_n(t-\tau_n)}.$$

Therefore, for every $\epsilon>0$ there exists $C>0$ such that

- if the family is born after time $T(t)+C$ it size is at most
$$e^{\gamma(t-T(t)-C)} \leq \epsilon e^{\gamma(1-\frac{1}{t})(t-T(t))}.$$
- if the family has fitness $1-\frac{C}{t}$ or lower, its size is at most
$$e^{\gamma(1-\frac{C}{t})(t-T(\frac{t}{C}))} \leq \epsilon\, e^{\gamma(1-\frac{1}{t})(t-T(t))}.$$

Hence it is plausible that families born outside the yellow box, for large C, are not the largest with high probability. The actual argument for this is much more technical, as we need to compare all families outside the box simultaneously with a small total probability of failure. □

***Tutorial*: The size and fitness of the largest family**

Problem:

(a) Show that, as $t \to \infty$,
$$e^{-\gamma(t-T(t))} \max_{n\in\mathbb{N}} Z_n(t) \Longrightarrow W^{-\frac{\gamma}{\lambda^\star}},$$
where W is exponentially distributed with parameter $\Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})(\lambda^\star)^{-\alpha}$.

(b) Let $V(t)$ the fitness of the family of maximal size at time t. Show that, under Assumption (cond), as $t \to \infty$, we have
$$t(1-V(t)) \Longrightarrow V,$$
where V is Gamma-distributed with scale parameter $\lambda^\star$ and shape parameter α.

Solution:

(a) We fix $x > 0$ and apply the vague convergence proved in Theorem 5.1 to the compact set
$$K := [-\infty, +\infty] \times [0, \infty] \times [x, \infty].$$
We get that
$$\sum_{n=1}^{M(t)} 1_K(\tau_n - T(t), t(1-F_n), e^{-\gamma(t-T(t))} Z_n(t)) \Rightarrow \texttt{Poisson}\big(\textstyle\int_K d\zeta\big).$$
Hence
$$\mathbb{P}\Big(e^{-\gamma(t-T(t))} \max_{n\in\{1,\dots,M(t)\}} Z_n(t) \ge x\Big) \to \mathbb{P}\Big(\texttt{Poisson}\big(\textstyle\int_K d\zeta\big) \ge 1\Big) = 1 - \exp\big(-\textstyle\int_K d\zeta\big). \tag{5.6}$$
Integrating out gives
$$\begin{aligned}\int_K d\zeta &= \int_{-\infty}^{+\infty} \int_0^\infty \int_x^\infty \alpha f^{\alpha-1} \lambda^\star e^{\lambda^\star s} e^{-z e^{\gamma(s+f)}} e^{\gamma(s+f)} dz\ df\ ds \\ &= \int_0^\infty e^{-w} \int_0^\infty \alpha f^{\alpha-1} \int_{-\infty}^{\frac{1}{\gamma}\log\frac{w}{x} - f} \lambda^\star e^{\lambda^\star s} ds\ df\ dw \\ &= \Big(\int_0^\infty e^{-w} \Big(\frac{w}{x}\Big)^{\frac{\lambda^\star}{\gamma}} dw\Big)\Big(\int_0^\infty \alpha f^{\alpha-1} e^{-\lambda^\star f} df\Big) \\ &= \tfrac{\Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})}{(\lambda^\star)^\alpha}\ x^{-\frac{\lambda^\star}{\gamma}}.\end{aligned}$$
Thus, the right hand side in (5.6) is $1 - \exp(-\Lambda x^{-\eta})$, for $\Lambda = \Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})(\lambda^\star)^{-\alpha}$ and $\eta = \frac{\lambda^\star}{\gamma}$, which proves the statement.

(b) The probability that $t(1-V(t))$ is in $[f, f+df]$ converges in the appropriate sense to
$$\int_{-\infty}^{+\infty} \int_0^\infty e^{-\zeta([-\infty,+\infty]\times[0,\infty]\times[z,\infty])}\, \zeta(ds, f, dz),$$
where the inner integration is with respect to z, and the outer with respect to s. We recall from above that
$$\zeta([-\infty, +\infty] \times [0, \infty] \times [z, \infty]) = \tfrac{\Gamma(\alpha+1)\Gamma(1+\frac{\lambda^\star}{\gamma})}{(\lambda^\star)^\alpha}\ z^{-\frac{\lambda^\star}{\gamma}}.$$

Under (cond), we have $\lambda^\star = \gamma$ and the right hand side becomes $\frac{\alpha\Gamma(\alpha,\lambda^\star)}{z}$, where

$$\Gamma(\alpha, \lambda^\star) := \int_0^\infty f^{\alpha-1} e^{-\lambda^\star f}\, df = \frac{\Gamma(\alpha)}{(\lambda^\star)^\alpha}.$$

We get, substituting $v = e^{\gamma(s+f)}$ and recalling that $\lambda^\star = \gamma$,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \int_0^\infty e^{-\zeta([-\infty,+\infty]\times[0,\infty]\times[z,\infty])} d\zeta(s,f,z) \\
&\qquad = \alpha f^{\alpha-1} e^{-\lambda^\star f} df \int_0^\infty \left(\int_0^\infty v e^{-zv} dv \right) e^{-\alpha\Gamma(\alpha,\lambda^\star)/z} dz \\
&\qquad = \alpha f^{\alpha-1} e^{-\lambda^\star f} df \int_0^\infty \frac{e^{-\alpha\Gamma(\alpha,\lambda^\star)/z}}{z^2}\, dz \\
&\qquad = \frac{f^{\alpha-1} e^{-\lambda^\star f} df}{\Gamma(\alpha,\lambda^\star)}.
\end{aligned}$$

CHAPTER 6

Open problems and ongoing work

- *Precise growth of the system*

  A question that remains open is about the precise rate of growth of the system in the condensation phase. We know from Theorem 2.1 that

  $$\frac{M(t)}{N(t)} \longrightarrow \frac{\beta}{\beta+\gamma} \qquad \text{almost surely.}$$

  Hence it is equivalent to ask for the absolute growth of either of the processes $(M(t)\colon t>0)$ or $(N(t)\colon t>0)$. In the bulk driven case we have, almost surely,

  $$\lim_{t\uparrow\infty} e^{-\lambda^* t} N(t) = \widetilde{W},$$

  for some positive random variable $\widetilde{W}$. The open problem is, whether in the condensation case there exists a deterministic function ψ such that $N(t)/\psi(t)$ converges, and to identify this function ψ. For regularly varying fitness distributions μ a lower bound follows by considering the growth of the largest family, which gives

  $$\frac{N(t)}{e^{\lambda^\star(t-T(t))}} \to \infty,$$

  from which an explicit bound follows by considering that $T(t) \sim {}^{\alpha}\!/_{\lambda^\star} \log t$.

- *Origin of the condensate*

  At which times are the families contributing to the condensate born? We have seen that the window of constant width around $T(t)$ from which the largest family originates does not contribute substantially to the condensate. Condensate particle must come from a different, larger, window.

- *Shape of the condensate*

  Our results offer only a partial answer to the question raised in Borgs et al. [**7**] how the links in the network are distributed among the highest fitnesses present in the system at any given time. The most interesting question that remains open here is whether the families that together form the condensate have a characteristic collective behaviour prior to condensation. The work on Kingman's model in Dereich and Mörters [**13**], and on a growth model without self-organisation in Dereich [**10**], suggests that this is indeed the case. More precisely, we believe that in a shrinking neighbourhood of the maximal fitness, the empirical fitness distribution takes an asymptotic shape of a universal nature.

  Conjecture 6.1 (Condensation wave). *Under assumption (RV) we have*

  $$\lim_{t\to\infty} \Xi_t(1-\tfrac{x}{t},1) = \frac{\omega(\beta,\gamma)}{\Gamma(\alpha+1)} \int_0^x y^\alpha e^{-y}\,dy,$$

in probability, i.e. the condensation wave has the shape of a Gamma distribution *with shape parameter* $1+\alpha$.

It is also plausible from calculations in the Kingman model that in cases of stronger decay of μ at $w(\mu)$ the condensation wave has a Gaussian shape. Such results are very hard to check in simulations as the simulation volume would have to be unrealistically large. They are also currently beyond the scope of rigorous mathematical analysis.

- *More general branching and Bianconi-Barabasi networks*

  There are two natural generalisations of our model:
  - In our model description as a branching process one could generalise the offspring rules to allow the birth of more than one mutant or selectant at a birth event. This will complicate notation but we do not expect any difficulties, or indeed significant new insights, from this generalisation.
  - In our Example 2 which describes our model as a growing network it would be natural to allow new vertices to connect to more than one existing vertex. The techniques of this paper are unsuitable for this generalisation. Dereich and Ortgiese [**14**] have developed a stochastic approximation technique which allows to generalise the results of Tutorial 2 (and some others) to this case.

- *Other classes of bounded fitness distributions*

  In this course we have investigated maximal families for a broad class of bounded fitness distributions, those of regular variation at the maximal fitness value, or equivalently where μ is in the maximal domain of attraction of the Weibull distribution. With a similar technique we are also able to understand fitness distributions with a faster decay at the maximal fitness value, ie. bounded μ in the maximal domain of attraction of the Gumbel distribution. For example, the class with

  $$\log \mu(1-\varepsilon, 1) \sim -\varepsilon^{-\gamma} \quad \text{for some } \gamma > 0.$$

  These results will be included in the forthcoming paper [**20**]. It is open whether there exist bounded fitness distributions where we experience condensation by macroscopic occupancy.

- *Unbounded fitness distributions and explosions*

  In the case of unbounded fitness distributions $w(\mu) = \infty$ explosions may occur, ie. the population can become infinitely large in finite time. When exactly do they occur and what can be said about the explosion time? There is some published work by Komjathy [**19**] that can be exploited.

- *Unbounded fitness distributions and travelling waves*

  In the case of unbounded fitness distributions if there is no explosion the empirical fitness distribution splits into a *bulk* part of asymptotic mass $\beta/(\beta+\gamma)$ and shape μ, and another part of mass $\gamma/(\beta+\gamma)$ going to infinity. The techniques of Chapter 4 are suitable to prove the mass and shape of the bulk, see also [**7**].

  The other part is called a *travelling wave* and is more interesting. Under suitable assumptions on μ one can identify the speed at which this wave travels to infinity [**12**]. But there are lots of open problems, in particular fine results on the spread and speed of the travelling wave, the size and fitness of the largest family, and possible the asymptotic shape of the wave. These problems are currently under investigation, see the forthcoming paper [**12**].

Acknowledgements: We are grateful to EPSRC for support through grant EP/K016075/1, and SFB 878 as well as the *Institut für mathematische Statistik* at the University of Münster for hosting PM during a sabbatical in 2015/16. We also thank Achim Klenke for useful discussions on branching processes with selection and mutation, which were the starting point of this project. Last but not least we would like the organisers of the workshop *Interplay of Probability and Analysis in Applied Mathematics* at the Mathematisches Forschungsinstitut Oberwolfach in 2015, where this work was extensively discussed.

References

1. Krishna B. Athreya and Samuel Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39 (1968) 1801-1817.
2. Krishna B. Athreya and Peter E. Ney. Branching Processes. *Springer-Verlag*, 1972.
3. Michiel van den Berg, John T. Lewis, and Joe Pulé. A general theory of Bose-Einstein condensation. *Helvetica Physica Acta* 59 (1986) 1271-1288.
4. Ginestra Bianconi and Albert-Laszlo Barabasi. Bose-Einstein condensation in complex networks. *Physical Review Letters*, 86 (2001) 5632-5635.
5. Nick Bingham, Charles Goldie, and Jozef Teugels. Regular Variation. *Cambridge University Press*, 1989.
6. Shankar Bhamidi. Universal techniques to analyze preferential attachment trees: Global and local analysis. Preprint available at http://www.unc.edu/∼bhamidi/.
7. 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, ACM, pp. 135-144, see arXiv:0710.4982 for a long version.
8. Fan Chung, Shirin Handjani and Doug Jungreis. Generalizations of Pólya's urn problem. *Annals of Combinatorics*, 7 (2003) 141-153.
9. Andrea Collevecchio, Codina Cotar, and Marco LiCalzi. On a preferential attachment and generalized Pólya's urn model. *Annals of Applied Probability*, 23 (2013) 12191253.
10. Steffen Dereich. Preferential attachment with fitness: Unfolding the condensate. *Electronic Journal of Probability*, 21 (2016) Paper 3.
11. Steffen Dereich, Cécile Mailler, and Peter Mörters. Non-extensive condensation in reinforced branching processes. *Ann. Appl. Probab.* to appear.
12. Steffen Dereich, Cécile Mailler, and Peter Mörters. Traveling waves in reinforced branching processes. *In preparation.*
13. Steffen Dereich and Peter Mörters. Emergence of condensation in Kingman's model of selection and mutation. *Acta Appl. Math.*, 127 (2013) 17-26.
14. Steffen Dereich and Marcel Ortgiese. Robust analysis of preferential attachment models with fitness. *Combin. Probab. Comput.* 23 (2014) 386-411.
15. Claude Godrèche and Jean Marc Luck. On leaders and condensates in a growing network. *Journal of Statistical Mechanics: Theory and Experiment* 07 (2010) P07031.
16. Andrea Hodgins-Davis, Daniel P. Rice and Jeffrey P. Townsend. Gene expression evolves under a house-of-cards model of stabilizing selection. *Mol. Biol. Evol.* (2015) doi: 10.1093/molbev/msv094.

17. Svante Janson. Functional limit theorems for multitype branching processes and generalized Polya urns. *Stochastic Processes and their Applications* 110 (2004) 177-245.
18. John Kingman, A simple model for the balance between selection and mutation. *Journal Appl. Prob.* 15 (1978) 1–12.
19. Júlia Komjáthy, Explosive Crump-Mode-Jagers branching processes. *arXiv:1602.01657 Preprint.*
20. Cécile Mailler, Peter Mörters, and Anna Senkevich . The universality classes in reinforced branching processes (working title). *In preparation.*
21. Olle Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 57 (1981) 365–395.
22. Peter Olofson The $x \log x$ condition for general branching processes. *J. Appl. Prob.* 35 (1998) 537–544.
23. Robin Pemantle. A survey of random processes with reinforcement. *Probability Surveys*, 4 (2007) 1-79.
24. Hilary A. Priestley. Complex Analysis. *Oxford University Press*, 2003.
25. Sidney Resnick. Extreme values, regular variation, and point processes. *Springer-Verlag*, 1987.
26. Anna Rudas, Balint Toth, and Benedek Valko. Random trees and general branching processes. *Random Structures and Algorithms*, 31 (2007) 186–202.