Branching in random environment and application to an infection in a cell population

Chaire MMB, Aussois, June 2023

Contents

1 Discrete branching processes in random environment

1.1 Galton-Watson processes

These processes have their origin in the study of the extinction probabilities of family names in Great Britain in the nineteenth century. Let ξ be an integer-valued random variable with law

$$\mathbb{P}(\xi = k) = p_k, \quad k \geq 0,$$

and Y_n be the population size at time n. The generation $n+1$ is composed of the descendants of the individuals of the generation n, and conditionally to Y_n, the individual $(1 \leq i \leq Y_n)$ of the generation n generates $\xi(i,n)$ descendants where the variables $\xi(i,n)$ and $\xi(j,n)$ are independent if $i \neq j$, and of the same law as ξ, which can be written as follows:

$$Y_{n+1} = \sum_{i=1}^{Y_n} \xi_{(i,n)}.$$

This procedure is then iterated, the variables $\xi(i,n)$ being independent of the variables $\xi(j,p)$ for $i \neq j$ or $p \neq n$. The Markov chain $(Y_n, n \geq 0)$ is called Galton-Watson process. If $Y(x)$ is a Galton-Watson process with an initial number of individuals $Y_0 = x$, we have the branching property which follows directly from the definition of the process:

$$Y(x+y) \overset{\mathcal{L}}{=} Y^{(1)}(x) + Y^{(2)}(y)$$

where $Y^{(1)}$ and $Y^{(2)}$ are independent copies of Y and $\overset{\mathcal{L}}{=}$ stands for equality in law. In terms of modelling, this property implies the absence of interaction between individuals. It is therefore well suited when considering small populations or populations with access to a large amount of resources.

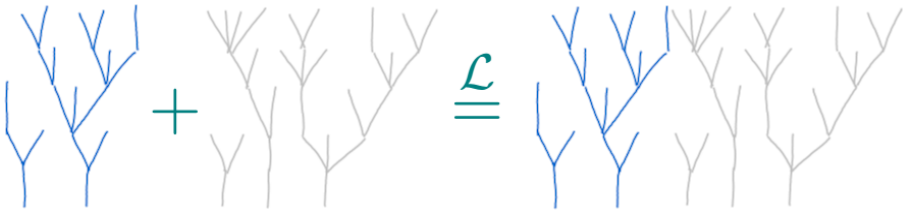

Figure 1: Illustration of the branching property

In the absence of precision, we will assume that $Y_0 = 1$. For $s \in [0,1]$, the generating function f of the random variable ξ is defined by:

$$f(s) := \mathbb{E}[s^\xi] = \sum_{i=0}^{\infty} p_i s^i, \quad s \in [0,1].$$

It characterises the law of ξ as

$$\mathbb{P}(\xi = k) = \frac{1}{k!} f^{(k)}(0).$$

We then notice that for any $(s,n) \in [0,1] \times \mathbb{N}$,

$$\mathbb{E}[s^{Y_{n+1}}] = \mathbb{E}[\mathbb{E}[s^{Y_{n+1}}|Y_n]] = \mathbb{E}[\mathbb{E}[s^{\xi_{(1,n)}+\xi_{(2,n)}+\ldots+\xi_{(Y_n,n)}}|Y_n]] = \mathbb{E}[f(s)^{Y_n}],$$

which entails, by iterating:

$$\mathbb{E}[s^{Y_n}] = f^{(n)}(s) = \underbrace{f \circ f \ldots \circ f}_{\text{n times}}, \quad (s,n) \in [0,1] \times \mathbb{N}.$$

The moments of the process, when they exist, can be expressed using the derivatives of f:

$$f'(1) \underset{s\to 1}{\leftarrow} f'(s) = \mathbb{E}[Y_1 s^{Y_1-1}] \underset{s\to 1}{\rightarrow} \mathbb{E}[Y_1],$$

$$(f'(1))^n = (f^{(n)})'(1) \underset{s\to 1}{\leftarrow} (f^{(n)})'(s) = \mathbb{E}[Y_n s^{Y_n-1}] \underset{s\to 1}{\rightarrow} \mathbb{E}[Y_n].$$

which leads to the following definition which will be justified later.

Definition 1.1. *If* $\ln f'(1)$ *is strictly less than* 0*, equal to* 0*, or strictly greater than* 0*, then the process is called subcritical, critical or supercritical respectively.*

This classification is equivalent to distinguishing processes that on average decrease, remain stable or increase. We restrict ourselves to the case $p_0 + p_1 < 1$; f is then strictly convex and increasing over $[0, 1]$.

Proposition 1.2. *The subcritical and critical processes almost surely die out. The supercritical process survives with positive probability.*

Knowing the probability of extinction, a natural question is to look at the extinction rates of processes that almost certainly die out, and at the long-time behaviour of processes that survive. Let us introduce the Markov chain

$$W_n := (f'(1))^{-n} Y_n, \quad n \in \mathbb{N}.$$

Then

$$\begin{aligned}\mathbb{E}[W_{n+1}|W_n] = \mathbb{E}[Y_{n+1}|Y_n](f'(1))^{-(n+1)} &= \mathbb{E}\left[\sum_{i=1}^{Y_n} \xi_{(i,n)} \Big| Y_n\right] (f'(1))^{-(n+1)} \\ &= Y_n f'(1)(f'(1))^{-(n+1)} = W_n.\end{aligned}$$

Such a process is called a (non-negative martingale) and has many properties. In particular it almost surely converges to a non-negative random variable W. This entails that in most cases Y_n roughly behaves as $(f'(1))^n$.

We have the following possibilities for the long time survival probability of the process (we exclude the trivial case when the process is deterministic).

Theorem 1.3.

- *Subcritical process: If* $\ln f'(1) < 0$ *and* $\mathbb{E}[Y_1 \ln^+ Y_1] < \infty$*, there exists a finite constant* c *such that*

$$\mathbb{P}(Y_n > 0) \sim c(f'(1))^n, \quad (n \to \infty).$$

- *Critical process: If* $\ln f'(1) = 0$*,*

$$\mathbb{P}(Y_n > 0) \sim \frac{2}{nVar(Y_1)}, \quad (n \to \infty).$$

- *Supercritical process: If* $\ln f'(1) > 0$ *and* $\mathbb{E}[Y_1 \ln^+ Y_1] < \infty$*, then* $\mathbb{E}[W] = 1$*, and*

$$\mathbb{P}(W = 0) < 1.$$

We may easily obtain an upper bound for the subcritical case using the Markov inequality:

$$\mathbb{P}(Y_n > 0) = \mathbb{P}(Y_n \geq 1) \leq \frac{\mathbb{E}[Y_n]}{1} = (f'(1))^n.$$

For the critical case, we notice that

$$\frac{1}{1-f(s)} - \frac{1}{1-s} = \frac{1}{(1-f(s))(1-s)}(f(s)-s).$$

In the vicinity of 1,

$$f(s) - s \sim \frac{\mathrm{Var}(Y_1)}{2}(1-s)^2$$

and

$$1 - f(s) = f(1) - f(s) \sim f'(1)(1-s) = 1-s.$$

Hence for s close to 1,

$$\frac{1}{1-f(s)} - \frac{1}{1-s} = \frac{1}{(1-f(s))(1-s)}\frac{\mathrm{Var}(Y_1)}{2}(1-s)^2 \sim \frac{\mathrm{Var}(Y_1)}{2}.$$

Moreover, as Y_n goes to 0 as n goes to inifinity, $f^{(n)}(s)$ goes to 1 uniformly for any $s \in [0,1]$ and

$$\frac{1}{1-f^{(n)}(s)} - \frac{1}{1-s} = \sum_{k=0}^{n-1}\left(\frac{1}{1-f^{(k+1)}(s)} - \frac{1}{1-f^{(k)}(s)}\right) \sim n\frac{\mathrm{Var}(Y_1)}{2}$$

for large n. Noticing that

$$\mathbb{P}(Y_n > 0) = 1 - f^{(n)}(0)$$

ends the proof.

1.2 Galton-Watson processes in variable environment

In the Galton-Watson process, the law of birth is the same for each generation, which means that the population lives in a constant environment. In the early 1970s, Smith and Wilkinson [SW69] and later Athreya and Karlin [AK71b, AK71a] sought to understand the effect of the environment on population dynamics. They introduced variability in birth laws reflecting environmental variability over generations. In the case of a population of annual plants, this variability may reflect the variability in sunshine, rainfall or pollination rate from one year to the next, variables that have a great influence on the quantity of offspring produced by the plants.

We thus assume now that the offspring distributions may depend on the generation n. For instance we can assume that there are dry and rainy years.

- When the year is rainy, the environment is denoted $\mathcal{E}_r$, and $p_2^{(r)} = p_3^{(r)} = 1/2$. Hence individuals have two or three children with equal probabilities.
- When the year is dry, the environment is denoted $\mathcal{E}_d$, and $p_0^{(d)} = p_1^{(d)} = 1/2$. Hence individuals have zero or one child with equal probabilities.
- A possible sequence of environment is for instance $(\mathcal{E}_r, \mathcal{E}_d, \mathcal{E}_d, \mathcal{E}_r, \mathcal{E}_d, \mathcal{E}_d, ...)$

More generally, The state of the environment at time $n-1$ is denoted $\mathcal{E}_n$. Associated with this state is a sequence of real numbers $(p_k^{(n)}, k \in \mathbb{N})$ where $p_k^{(n)}$ is the probability that an individual of the generation $n-1$ produces k descendants in the generation n. The generating function f_n at generation n is thus

$$f_n(s) := \mathbb{E}[s^{Z_n}] = \sum_{k=0}^{\infty} p_k^{(n)} s^k, \quad (s,n) \in [0,1] \times \mathbb{N}.$$

The generating functions of the inhomogeneous Markov chain $(Z_n, n \in \mathbb{N})$ corresponding to the sequence of population sizes in the environment $(\mathcal{E}_1, ..., \mathcal{E}_n, ...)$ are then obtained by iterating the successive generating functions. If we denote again $\xi_{(i,n)}$ the number of offspring of the ith individual of generation $n-1$:

$$\begin{aligned}
\mathbb{E}[s^{Z_n}] &= \mathbb{E}[s^{\xi_{(1,n-1)}+\xi_{(2,n-1)}+...+\xi_{(Z_{n-1},n-1)}}] \\
&= \mathbb{E}[\mathbb{E}[s^{\xi_{(1,n-1)}}]^{Z_{n-1}}] \\
&= \mathbb{E}[f_n(s)^{Z_{n-1}}] \\
&= \mathbb{E}[f_{n-1}(f_n(s))^{Z_{n-2}}] \\
&= ... \\
&= f_1 \circ f_2 \circ ... \circ f_n(s), \quad (s,n) \in [0,1] \times \mathbb{N}.
\end{aligned}$$

Therefore, the average of the population size Z_n in the environment $(\mathcal{E}_1, ..., \mathcal{E}_n)$ is:

$$\mathbb{E}[Z_n] = f_1'(1)...f_n'(1) =: e^{S_n}, \quad n \in N, \tag{1}$$

where S_n is defined by:

$$S_0 = 0, \quad S_n = \ln f_1'(1) + ... + \ln f_n'(1), \quad n \in \mathbb{N}. \tag{2}$$

1.3 Galton-Watson processes in iid random environments

We will now consider that the environments are themselves random variables, and will focus on the case where they are chosen randomly with the same law at each generation. In particular, their law does not depend on the population size Z_n. We refer to [BGK05, VDS13] or the recent monograph [KV17] for known results on critical and sub-critical processes (see Definition 1.4). We also refer to [BB11, BB13, Böi14] for results concerning the supercritical case. We will limit our presentation to the monotype case and to iid environments.

A Galton-Watson process $(Z_n, n \in \mathbb{N})$ in a random environment (GWRE) can thus be described in terms of its generating functions $(f_n, n \in \mathbb{N})$. As the law of the sequence of environments $(\mathcal{E}_1, ..., \mathcal{E}_n, ...)$ does not depend on anything else we can first draw the sequence, then study the population Z in this particular environment, and then take the average on the possible sequences of environments. This is called the quenched approach and has proven to be very useful to study the questions of extinction of GWRE (see [AGK+05] for instance).

The process $S := (S_n, n \in \mathbb{N})$ defined in (2) is thus a random walk. We will see that this random walk is intimately linked to the long time behaviour of the process $(Z_n, n \in \mathbb{N})$. But before

specifying more precisely this relation, we may already notice that:

$$\begin{aligned}\mathbb{P}(Z_n > 0|\mathcal{E}_1, ..., \mathcal{E}_n) &= \min_{0\le i\le n} \mathbb{P}(Z_i > 0|\mathcal{E}_1, ..., \mathcal{E}_n) \\ &\le \min_{0\le i\le n} \mathbb{E}[Z_i|\mathcal{E}_1, ..., \mathcal{E}_n] = \exp(\min(S_0, S_1, ..., S_n)), \end{aligned} \tag{3}$$

where we applied Markov inequality, Equality (1) and Definition (2). If the random walk $(S_n, n \in \mathbb{N})$ is not degenerate (i.e. we exclude the case $S \equiv 0$), it has only three possible a.s. behaviours at infinity [Fel71]: $\lim_{n\to\infty} S_n = +\infty$, $\lim_{n\to\infty} S_n = -\infty$, and $\limsup_{n\to\infty} S_n = -\liminf_{n\to\infty} S_n = +\infty$, which correspond respectively to cases $\mathbb{E}[\ln f_1'(1)] < 0$, > 0, and $= 0$ when $\mathbb{E}[|\ln f_1'(1)|] < \infty$. In the rest of this presentation and for the sake of simplicity we will restrict ourselves to the case where $\ln f_1'(1)$ is integrable. It is possible to define a classification for GWRE's similar to the Definition 1.1:

Definition 1.4. *If $\mathbb{E}[\ln f_1'(1)]$ is strictly less than 0, equal to 0, or strictly greater than 0, the Galton-Watson process in a random environment is called subcritical, critical or supercritical.*

Inequality (3) is sufficient to show that the process $(Z_n, n \in \mathbb{N})$ almost certainly gets extinct in subcritical and critical cases. We have in fact the equivalent of Theorem 1.3:

Proposition 1.5 (Theorem 3.1 [SW69])**.** *Subcritical and critical processes get extinct almost surely. The supercritical process survives with a positive probability if $\mathbb{E}[|\ln(1 - f_1(0))|] < \infty$.*

The condition $\mathbb{E}[|\ln(1 - f_1(0))|] < \infty$ excludes the possibility that a sequence of "catastrophic" environments leads to the death of the population within a few generations.

Finally, we have the following equivalent of Theorem 1.3 concerning the extinction rate of the critical case and the long time behaviour of the supercritical process in the case of survival:

Theorem 1.6 ([AK71a, Kap74, Koz76, GK01])**.**

- *Critical process: If $\mathbb{E}[\ln f_1'(1)] = 0$, $0 < Var[\ln f_1'(1)] < \infty$ and $f_1'(1)$ has a non-lattice distribution, there are positive finite constants c_1 and c_2 such that*

$$\mathbb{P}(Z_n > 0) \sim c_1 \mathbb{P}(\min(S_0, ..., S_n) \ge 0) \sim \frac{c_2}{\sqrt{n}}, \quad (n \to \infty).$$

- *Supercritical process: If $\mathbb{E}[\ln f_1'(1)] > 0$ and $\mathbb{E}[Z_1 \ln Z_1 / f_1'(1)] < \infty$, then the martingale $(Z_n \exp(-S_n), n \in \mathbb{N})$ has a non-zero finite limit on the non-extinction event of the process:*

$$\lim_{n\to\infty} Z_n e^{-S_n} = W \quad a.s., \quad \mathbb{P}(W > 0) = \mathbb{P}(\forall n \in \mathbb{N}, Z_n > 0|Z_0 = 1) > 0.$$

The behaviour of the subcritical process is more complex in the case of GWREs than in the case of conventional Galton-Watson processes. Knowing that $\mathbb{E}[\ln f_1'(1)] < 0$ is not sufficient to deduce the behaviour of the process in long time. The extinction rate in the subcritical case has been studied, in more and more general cases by Dekking [Dek87], D'Souza and Hambly [DH97], Guivarch and Liu [GL01], and finally by Geiger, Kersting and Vatutin [GKV03]. We have the following asymptotics for the probability of survival in the subcritical case (recall that $S_1 = \ln f'(1)$).

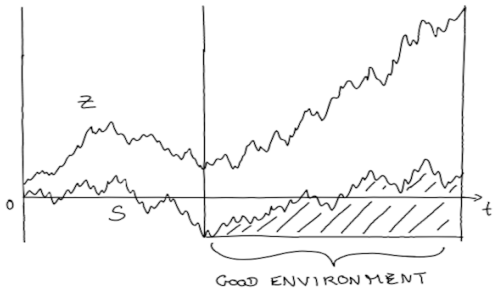

Figure 2: Idea (Afanasyev et al 2005): Decomposition with respect to the minimum of the auxiliary random walk S. In the critical and weakly subcritical cases, surviving $\simeq$ experiencing particularly good environments

Theorem 1.7 ([GKV03]). *Under additional moment assumptions we have the following asymptotics, where c denotes a positive finite constant whose value can change from one line to another:*

- *Strong subcritical case: If $\mathbb{E}[f_1'(1)\ln f_1'(1)] < 0$, then*
$$\mathbb{P}(Z_n > 0) \sim c(\mathbb{E}[f_1'(1)])^n, \quad (n \to \infty).$$
- *Intermediate subcritical case: If $\mathbb{E}[f_1'(1)\ln f_1'(1)] = 0$, then*
$$\mathbb{P}(Z_n > 0) \sim c(\mathbb{E}[f_1'(1)])^n n^{-1/2}, \quad (n \to \infty).$$
- *Weak subcritical case: If $0 < \mathbb{E}[f_1'(1)\ln f_1'(1)] < \infty$, then*
$$\mathbb{P}(Z_n > 0) \sim Cc_1\mathbb{P}(\min(S_0, ..., S_n) \geq 0) \sim c\gamma^n n^{-3/2}, \quad (n \to \infty)$$
where $\gamma := \inf_{\tau\in[0,1]} \mathbb{E}[f_1'(1)^\tau]$.

Hence when the variability of the environment is small, the extinction probability is similar to the one of the subcritical Galton-Watson process. For an average growth rate $\mathbb{E}[Z_1|Z_0 = 1] = \mathbb{E}[\ln f_1'(1)] < 0$, the exponential extinction rate depends on the law of the environment only if the environment is sufficiently variable, more precisely when $\mathbb{E}[f_1'(1)\ln f_1'(1)] > 0$. In order to see that the variability of the environment is linked to this inequality it can be noted that for this inequality to be satisfied $f'(1)$ must take large values. Moreover, since $\mathbb{E}[\ln f'(1)]$ is negative in the subcritical case, $f'(1)$ must also take small values.

To understand better the reasons of surviving a long time for a population, Bansaye [Ban09b] has studied the following question: what can we say about environments when we condition on the surviving of the population during a large number of generations? In fact the answer is not the same if we consider the strong and intermediate subcritical cases or if we consider the weak subcritical case. In the first two cases, conditionally on population survival, the survival probability of the branching process in the selected environment is still zero. This means that the survival is due to a particularly high reproduction of an individual line in a subcritical environment. In the weak subcritical case, conditioning on the survival of the population selects only supercritical environments, which means that the sequence of selected environments has a.s. a positive survival probability.

2 (Stable) Continuous state branching processes in random environment

2.1 (Stable) Continuous state branching processes

Continuous state branching processes (CSBPs) are the analog in continuous time and space of Galton-Watson processes. They were introduced by Jirina [Jir58] and studied by many authors. Let us mention in particular [Lam67a, Lam67b, Gre74, Gri74, Bin76] for the first works on the subject, and [Kyp06, Li10] for recent reviews. A CSBP $Y = (Y_t, t \in \mathbb{R}_+)$ is a strong Markov process with values in $\mathbb{R}_+$ and which satisfies the branching property. Y admits 0 and ∞ as absorbing points, and $\mathbb{P}_x$ denotes the law of the process starting from x. Lamperti [Lam67b] has shown that CSBPs are the only possible scaling limits of Galton-Watson processes and that, conversely, any CSBP can be obtained as such a scaling limit.

CSBPs are spectrally positive, with jumps representing macroscopic birth events (an infinitesimal individual gives birth to a sufficiently large number of infinitesimal individuals to be visible macroscopically).

In this course, we will only focus on a particular class of CSBP's: the stable CSBP's. They are parametrized by a real number $\beta \in (-1, 0) \cup (0, 1]$. In our applications we will only need to take $\beta \in (-1, 0) \cup \{1\}$. For these parameters, the stable CSPB's have the following form:

- If $\beta = 1$ the process, called Feller diffusion [Fel71], is continuous. It may be realised as the unique strong solution to the stochastic differential equation

$$Y_t = Y_0 + g \int_0^t Y_s ds + \int_0^t \sqrt{2\sigma^2 Y_s} dB_s \tag{4}$$

  where B is a standard Brownian motion. In this case, we may prove that,

$$\mathbb{E}_x[Y_t] = xe^{gt}.$$

  g is thus the Malthusian parameter of the population dynamics. We then have the analog of Definitions 1.1 and 1.4.

  Definition 2.1. *If g is strictly greater than 0, equal to 0, or strictly less than 0, the CSBP Y is called supercritical, critical or subcritical, respectively.*

  In the case of the Feller diffusion, the events of extinction

$$Ext := \left\{ \lim_{t \to \infty} Y_t = 0 \right\}$$

  and the event of absorption

$$Abs := \{\exists t < \infty, Y_t = 0\}$$

  coincinde (which is not always the case; for example the CSBP $Y = (Y_0 e^{-t}, t \geq 0)$ satisfies for every $x \in \mathbb{R}_+^\star$, $\mathbb{P}_x(Ext) = 1$ and $\mathbb{P}_x(Abs) = 0$). We have the analog of Propositions 1.2 and 1.5.

  Proposition 2.2. *The subcritical and critical processes almost surely die out. The supercritical process survives with positive probability.*

Asymptotics for the survival probability at large time may be obtained and they are similar to the ones of Galton-Watson processes. This is a consequence of the fact that Feller diffusion is a scaling limit of Galton-Watson process sequences whose reproduction law has a finite variance. If

$$\mathfrak{Y}_t^{(n)} := \frac{1}{n} Y_{\lfloor nt \rfloor}^{(n)}, \quad t \geq 0$$

where $(Y^{(n)}, n \in \mathbb{N})$ is a sequence of GW processes with reproduction law ξ and satisfying $Y_0^{(n)} = n$ and $Var(Y_1^{(n)}) < \infty$ for all $n \in \mathbb{N}$, then $(\mathfrak{Y}_t^{(n)}, t \in [0,T])$ converges in law when n tends towards infinity (here T is a positive real) towards a Feller diffusion on the interval $[0,T]$.

- If $\beta \in (-1,0)$. In this case, the process may be realized as the unique strong solution to the SDE

$$Y_t = Y_0 + g\int_0^t Y_s ds + \int_0^t \int_0^\infty \int_0^{Y_{s^-}} zR(ds,dz,du), \tag{5}$$

where $R(ds,dz,du)$ is a Poisson random measure with intensity $ds\rho(dz)du$. The jump measure ρ is given by:

$$\rho(dx) = \frac{c_\beta \beta(\beta+1)}{\Gamma(1-\beta)x^{\beta+2}}, \quad x > 0 \tag{6}$$

(with $c_\beta < 0$). Notice that as $\beta < 0$,

$$\int^\infty x\rho(dx) = \infty.$$

These processes have thus an infinite mean, and as we will see they may explode (reach infinity) in finite time.

2.2 (Stable) Continuous state branching processes in Lévy random environment

The construction of CSBPs as scaling limits of Galton-Watson processes raises a natural question. Is it possible, like for the Galton-Watson processes, to immerse a CSBP in a random environment, in order to take into account the variability of the living and reproductive conditions of the (infinitesimal) individuals constituting the modelled populations? Here the question is more delicate. Indeed, to make the environment vary in a discrete framework we consider a new reproduction law at each discrete time step (or generation). In the continuous time case, the environment can vary in multiple ways: continuously, abruptly, more or less frequently, with a variable amplitude...

In the case where the environment can be described by a Lévy process, similarly as in the case of BPREs where the environment could be described by the random walk S, CSBPs in (Lévy) random environment can be obtained as scaling limits of BPREs. Historically, the existence of branching diffusions in Brownian environments constructed as scaling limits of Galton-Watson processes in random environment has been conjectured by Keiding [Kei75] and rigorously established by Kurtz [Kur78]. Recent work [BS15, BCM19] has focused on the scaling limits of a large class of branching processes in variable environments (in particular a reproduction law that could have infinite variance, bottlenecks, interactions between individuals,...), and the authors of these works obtained, among other processes, a large class of CSBPs in Lévy environments.

Like CSBPs, CSBPs in Lévy environment can be seen as solutions to SDEs. The realization of generalized branching processes (with interactions, immigration, bottlenecks,...) in variable environments as the only strong solution of SDEs is currently the subject of much work (see for example [PP18, HLX18, FL20]). The proofs of these results are generalizations of techniques introduced by Li and co-authors (see for example [FL10, DL12, LP12]). We will not need to apply these results in their full generality, and so we present only the case of interest for us, namely the case of a branching process (without additional interactions or immigration) in a random environment given by compouned Poisson process.

2.3 (Stable) Continuous state branching processes with catastrophes

For the applications we will consider, we are interested in modeling catastrophes which occur at random and kill each individual with some probability (depending on the catastrophe). In terms of the CSBP representing the scaling limit of the size of a large population, this amounts to letting the process make a negative jump, i.e. multiplying its current value by a random fraction. The process that we obtain is still Markovian whenever the catastrophes follow a time homogeneous Poisson Point Process.

The process Z that we consider in this section is then called a CSBP with catastrophes. Roughly speaking, it can be defined as follows: The process Z follows the SDE (4) or (5) between catastrophes, which are then given in terms of the jumps of a Lévy process with bounded variation paths. Thus the set of times at which catastrophes occur may have accumulation points, but the mean effect of the catastrophes has a finite first moment. When a catastrophe with effect θ_t occurs at time t, we have

$$Z_t = \theta_t Z_{t-}.$$

Let us now be more precise. The law of the catastrophes does not depend on the value of the process Z and are given by a Poisson random measure $N_1 = \sum_{i\in I} \delta_{t_i,\theta_{t_i}}$ on $[0,\infty) \times [0,\infty)$ with intensity $r\mathrm{d}t\kappa(\mathrm{d}\theta)$ with $r \in \mathbb{R}_+$ such that $\kappa(\{0\}) = 0$. The jump process

$$\Delta_t = \int_0^t \int_{(0,\infty)} \log(\theta) N_1(\mathrm{d}s, \mathrm{d}\theta) = \sum_{s\leq t} \log(\theta_s),$$

is thus a Lévy process with paths of bounded variation, which is non identically zero. The choice of this representation, which is equivalent to the measure N_1, stems from the following observation: when a catastrophe of intensity θ_t occurs, the population size is multiplied by a factor $\theta_t = e^{\ln\theta_t}$. Hence the introduction of the process $(\Delta_t, t \geq 0)$ will make it possible to explicitly express the conditions under which the population can survive with positive probability.

The CSBP with catastrophes κ is defined as the solution to the following stochastic differential equation:

$$\begin{aligned} Z_t =& Z_0 + g\int_0^t Z_s ds + \mathbf{1}_{\{\beta=1\}} \int_0^t \sqrt{2\sigma^2 Z_s} dB_s \\ &+ \mathbf{1}_{\{\beta\in(-1,0)\}} \int_0^t \int_0^\infty \int_0^{Z_{s-}} zR(ds,dz,du) + \int_0^t \int_{[0,\infty)} \Big(\theta - 1\Big) Z_{s-} N_1(\mathrm{d}s, \mathrm{d}\theta). \end{aligned} \tag{7}$$

The branching property of Z, conditionally on Δ, is inherited from the branching property of the CSBP and the fact that the additional jumps are multiplicative.

The key observation is that the process $\widetilde{Z}_t = Z_t \exp\{-gt - \Delta_t\}$ has a 'simplified' dynamics due to the compensations when a catastrophe occurs. Indeed, if a catastrophe of intensity θ_t occurs at time t,

$$\widetilde{Z}_t - \widetilde{Z}_{t-} = Z_{t-}\theta_t e^{-gt-\Delta_{t-}-\ln\theta_t} - Z_{t-}e^{-gt-\Delta_{t-}} = 0.$$

Moreover, as for $t \geq 0$, $e^{-gt-\Delta_t} \notin \{0, \infty\}$, we obtain the following equalities:

$$\{\widetilde{Z}_t = \infty\} = \{Z_t = \infty\} \quad \text{and} \quad \{\widetilde{Z}_t = 0\} = \{Z_t = 0\}, \quad t \geq 0.$$

We thus get that

$$\begin{aligned}\mathbb{E}\left[e^{-\lambda \widetilde{Z}_t}\right] &\underset{\lambda\to\infty}{\to} \mathbb{P}(\widetilde{Z}_t = 0) = \mathbb{P}(Z_t = 0)\\ &\underset{\lambda\to 0}{\to} \mathbb{P}(\widetilde{Z}_t < \infty) = \mathbb{P}(Z_t < \infty)\end{aligned}$$

It allows us prove the following proposition:

Proposition 2.3 ([BPS13, PPS16]). *For all $z > 0$ and $t \geq 0$, if $\beta = 1$:*

$$\mathbb{P}_z(Z_t > 0 \mid \Delta) = 1 - \exp\left\{-z\left(\sigma^2 \int_0^t e^{-(gs+\Delta_s)}\mathrm{d}s\right)^{-1}\right\} \qquad a.s \tag{8}$$

and if $\beta \in (-1, 0)$,

$$\mathbb{P}_z\Big(Z_t < \infty\Big|\Delta\Big) = \exp\left\{-z\left(\beta c_\beta \int_0^t e^{-\beta(gs+\Delta_s)}\mathrm{d}u\right)^{-1/\beta}\right\} \qquad a.s. \tag{9}$$

Both probabilities thus appear as functionals of the integral of the exponential of the Lévy process Δ. This allows one to obtain precise asymptotics at large time for these probabilities. The results are a bit technical and will not be presented in this course. Interested readers are referred to [BPS13, PPS16, PP17, LX18].

3 Parasite infection in a cell population

We are interested in the modelling of a parasite infection in a cell population, and, in particular, in the role of the stochasticity of repartition of the parasites at cell division. From the pioneering work of Kimmel [Kim97], several models and associated analysis have been proposed, both in discrete [Ban08, Ban09a, AG13, AG16] and continuous time [BT11, BPS13]. The asymmetric repartition of parasites is taken into account in all of those work: in [Ban08, OW20], each parasite chooses to go to one daughter cell with probability p (and to the other with probability $1 - p$), and in [Ban09a], a random environment is considered (the probability generating functions of the number of parasites at birth in the two daughter-cells of each cell in the population are i.i.d random variables). In branching-within-branching models, independently for each parasite and with the same distribution, the descendants are shared between the daughter cells. A different approach

has been proposed in [BT11], removing the independence property of the sharing of parasites descending from different lineages. Following the dynamics of the ($\mathbb{R}$-valued) quantity of parasites inside the cells (rather than a discrete count), this model assumes that a cell with x parasites, splits into two daughter cells with a quantity of parasites Θx and $(1-\Theta)x$ respectively, with Θ a random variable on $[0,1]$. In this section, we extend this approach and explore the role of the random variable Θ in the proliferation of the infection. We refer to [MS23] for the rigorous proofs.

3.1 Model

3.1.1 Parasites dynamics in a cell

Each cell contains parasites whose quantity evolves as a diffusion with positive jumps. More precisely, we consider the SDE

$$\mathfrak{X}_t = x + \int_0^t g(\mathfrak{X}_s)ds + \int_0^t \sqrt{2\sigma^2(\mathfrak{X}_s)}dB_s + \int_0^t \int_0^{\mathfrak{X}_{s^-}} \int_{\mathbb{R}_+} zR(ds,dx,dz), \tag{10}$$

where x is nonnegative, $g, \sigma \geq 0$ are real functions on $\overline{\mathbb{R}}_+$, B is a standard Brownian motion, and R has been defined in (5). Finally, B, Q and R are independent.

Compared to the branching processes considered until now, taking functions $g(x)$ and $\sigma^2(x)$ (and not necessarily gx and $\sigma^2 x$) allows to take into account interactions in the reproduction law of (infinitesimal) individuals.

We will provide later on conditions under which the SDE (10) has a unique nonnegative strong solution. In this case, it is a Markov process with infinitesimal generator $\mathcal{G}$, satisfying for all $f \in C_0^2(\mathbb{R}_+)$ (the set of twice continuously differentiable functions on $\mathbb{R}_+$ vanishing at 0 and infinity),

$$\mathcal{G}f(x) = g(x)f'(x) + \sigma^2(x)f''(x) + x\int_{\mathbb{R}_+} (f(x+z) - f(x))\,\rho(dz), \tag{11}$$

and 0 and $+\infty$ are two absorbing states. Following [Mar16], we denote by $(\Phi(x,s,t), s \leq t)$ the corresponding stochastic flow *i.e.* the unique strong solution to (10) satisfying $\mathfrak{X}_s = x$ and the dynamics of the trait between division events is well-defined.

3.1.2 Cell division

Each cell carrying a quantity x of parasites divides at rate r and is replaced by two daughter cells with quantity of parasites Θx and $(1-\Theta)x$, where Θ is a symmetrical random variable on $(0,1)$, with associated distribution $\kappa(\cdot)$. Finally, we assume that $\mathbb{E}[|\ln\Theta|] < \infty$.

3.1.3 Cell death

Cells can die because of two mechanisms. First, they have a natural death rate q. Second, cells die when the quantity of parasites they carry explodes (*i.e.*, reaches infinity in finite time), as a proper functioning of the cell is not possible anymore. Notice that to model this case, we do not 'kill the cell' strictly speaking. As infinity is an absorbing state for the quantity of parasites in a cell, and

as a cell with an infinite quantity of parasites transmits an infinite quantity of parasites to both its daughter cells, we let the process evolve and decide that a cell is dead if it contains an infinite quantity of parasites.

3.1.4 Existence and uniqueness

We use the classical Ulam-Harris-Neveu notation to identify each individual. Let us denote by

$$\mathcal{U} := \bigcup_{n\in\mathbb{N}} \{0,1\}^n$$

the set of possible labels, $\mathcal{M}_P(\overline{\mathbb{R}}_+)$ the set of point measures on $\overline{\mathbb{R}}_+$, and $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\overline{\mathbb{R}}_+))$, the set of càdlàg measure-valued processes. For any $Z \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\overline{\mathbb{R}}_+))$, $t \geq 0$, we write

$$Z_t = \sum_{u\in V_t} \delta_{X_t^u}, \tag{12}$$

where $V_t \subset \mathcal{U}$ denotes the set of individuals alive at time t and X_t^u the quantity of parasites at time t of cell u. By convention, Z_t is the null measure if $V_t = \emptyset$. By extension, for $u \in V_t$ and any $s \leq t$, X_s^u denotes the quantity of parasites in the ancestor of u in the population at time s. Thus, $(X_s^u, s \leq \sup(\mathfrak{t} \geq 0 \text{ s.t. } u \neq V_\mathfrak{t}))$ follows (10) between events of division impacting the lineage under consideration.

Let $\left(\Phi^u(x,s,t), u \in \mathcal{U}, x \in \overline{\mathbb{R}}_+, s \leq t\right)$ be a family of independent stochastic flows satisfying (10) describing the individual-based dynamics. Let $E = \mathcal{U} \times (0,1) \times \overline{\mathbb{R}}_+$ and $M(ds, du, d\theta, dz)$ be a PPM on $\mathbb{R}_+ \times E$ with intensity $ds \otimes n(du) \otimes \kappa(d\theta) \otimes dz$, where $n(du)$ denotes the counting measure on $\mathcal{U}$. We assume that M and $(\Phi^u, u \in \mathcal{U})$ are independent. We denote by $\mathcal{F}_t$ the filtration generated by the restriction of the PPM M to $[0,t] \times E$ and the family of stochastic processes $(\Phi^u(x,s,t), u \in \mathcal{U}, x \in \overline{\mathbb{R}}_+, s \leq t)$ up to time t.

We now introduce assumptions to ensure the strong existence and uniqueness of the process.

Assumption EU. *1. The function g is continuous on $\mathbb{R}_+$, $g(0) = 0$ and for any $n \in \mathbb{N}$ there exists a finite constant B_n such that for any $0 \leq x \leq y \leq n$*

$$|g(y) - g(x)| \leq B_n \phi(y-x), \text{ where } \phi(x) = \begin{cases} x\,(1 - \ln x) & \text{if } x \leq 1, \\ 1 & \text{if } x > 1. \end{cases}$$

2. The function σ is Hölder continuous with index $1/2$ on compact sets and $\sigma(0) = 0$.

Recall the definition of $\mathcal{G}$ in (11). Then, the structured population process may be defined as the strong solution to a SDE.

Proposition 3.1. *Under Assumption EU, there exists a strongly unique $\mathcal{F}_t$-adapted càdlàg process $(Z_t, t \geq 0)$ taking values in $\mathcal{M}_P(\overline{\mathbb{R}}_+)$ such that for all $f \in C_0^2(\overline{\mathbb{R}}_+)$ and $x_0, t \geq 0$,*

$$\begin{aligned}\langle Z_t, f\rangle =& f(x_0) + \int_0^t \int_{\mathbb{R}_+} \mathcal{G}f(x) Z_s(dx)\, ds + M_t^f(x_0) \\ &+ \int_0^t \int_E \mathbf{1}_{\{u\in V_{s^-}\}} \Big(\mathbf{1}_{\{z\leq r\}} \left(f(\theta X_{s^-}^u) + f((1-\theta)X_{s^-}^u) - f(X_{s^-}^u)\right) \\ &\qquad - \mathbf{1}_{\{0<z-r\leq q\}} f(X_{s^-}^u)\Big) M(ds, du, d\theta, dz),\end{aligned}$$

where for all $x \geq 0$, $M_t^f(x)$ is a $\mathcal{F}_t$-martingale.

For the sake of readability, we will assume that all the processes under consideration in the sequel satisfy Assumption **EU**, but we will not indicate it.

We will now investigate the long time behaviour of the infection in the cell population. The strategy to obtain information at the population level is to introduce an auxiliary process providing information on the behaviour of a 'typical individual'.

3.2 Many-to-One formula

We are interested in the long-term behaviour of the infection in the cell population. One possible approach is to select an individual uniformly at random at a given time t and look at the properties satisfied by this individual's trait (here, the quantity of parasites it contains, for example). We are therefore interested in quantities of the form

$$\frac{\sum_{u\in V_t} f(X_t^u)}{N_t}, \tag{13}$$

for bounded functions f, where N_t denotes the cell population size at time t. But as we shall see, it is simpler to look first at quantities of the form

$$\frac{\sum_{u\in V_t} f(X_t^u)}{\mathbb{E}[N_t]},$$

and then make the link with (13). Recall that in our case the size of the cell population follows a branching process of birth rate r and death rate q. It therefore has expectation $e^{(r-q)t}$ at time t (to simplify the proof, we do not actually 'kill' a cell with an infinite number of parasites, but leave it in the population with an infinite number of parasites, which it transmits to its daughter cells. It will therefore suffice to count the number of cells with an infinite number of parasites at time t to know the actual number of cells alive at time t). We are thus interested in the process

$$\mathfrak{K}_t := e^{-(r-q)t} \sum_{u\in V_t} f(X_t^u).$$

We will thus look at the dynamics of this process

$$\begin{aligned} d\mathfrak{K}_t = &-(r-q)dt\mathfrak{K}_t + e^{-(r-q)t} \sum_{u\in V_t} \mathcal{G}f(X_t^u)dt + dM_t \\ &- e^{-(r-q)t} \sum_{u\in V_t} \mathbf{1}_{\{\text{u dies during [t,t+dt]}\}} f(X_t^u) \\ &+ e^{-(r-q)t} \sum_{u\in V_t} \mathbf{1}_{\{\text{u div during [t,t+dt]}\}} (f(\theta_t^u X_{t-}^u) + f((1-\theta_t^u)X_{t-}^u) - f(X_{t-}^u)) \end{aligned}$$

where M is a $\mathcal{F}_t$-martingale and by abuse of notation θ_t^u and $1-\theta_t^u$ are the fractions of parasites transmited in the daugthers of the cell u dividing at time t. Recall that the θ_t^u are iid of law κ and

that the times of division for each cell follows an exponential law of parameter r. In particular it implies that

$$\mathbb{E}\left[\sum_{u\in V_t} \mathbf{1}_{\{\text{u div during [t,t+dt]}\}}\left(f(\theta_t^u X_{t-}^u) + f((1-\theta_t^u)X_{t-}^u) - f(X_{t-}^u)\right)\right]$$
$$= rdt\mathbb{E}\left[\sum_{u\in V_t}\int_0^1 (f(\theta X_{t-}^u) + f((1-\theta)X_{t-}^u) - f(X_{t-}^u))\kappa(d\theta)\right]$$
$$= rdt\mathbb{E}\left[\sum_{u\in V_t}\int_0^1 (2f(\theta X_{t-}^u) - f(X_{t-}^u))\kappa(d\theta)\right]$$

as Θ has a symmetrical distribution. Similarly,

$$\mathbb{E}\left[\sum_{u\in V_t} \mathbf{1}_{\{\text{u dies during [t,t+dt]}\}} f(X_t^u)\right] = qdt\mathbb{E}\left[\sum_{u\in V_t} f(X_t^u)\right].$$

We thus obtain

$$d\mathbb{E}[\mathfrak{K}_t] = -(r-q)dt\mathbb{E}[\mathfrak{K}_t] + e^{-(r-q)t}\mathbb{E}\left[\sum_{u\in V_t}\mathcal{G}f(X_t^u)\right]dt + qdt\mathbb{E}[\mathfrak{K}_t]$$
$$+ rdte^{-(r-q)t}\mathbb{E}\left[\sum_{u\in V_t}\int_0^1 (2f(\theta X_{t-}^u) - f(X_{t-}^u))\kappa(d\theta)\right]$$
$$= e^{-(r-q)t}\mathbb{E}\left[\sum_{u\in V_t}\mathcal{G}f(X_t^u)\right]dt + 2rdte^{-(r-q)t}\mathbb{E}\left[\sum_{u\in V_t}\int_0^1 (f(\theta X_{t-}^u) - f(X_{t-}^u))\kappa(d\theta)\right]$$
$$= \mathbb{E}\left[\mathcal{G}\left(e^{-(r-q)t}\sum_{u\in V_t} f(X_t^u)\right)\right]dt + 2rdt\int_0^1 \mathbb{E}\left[\sum_{u\in V_t} e^{-(r-q)t}f(\theta X_{t-}^u) - \sum_{u\in V_t} e^{-(r-q)t}f(X_{t-}^u)\right]\kappa(d\theta).$$

Now let us introduce the process $\mathfrak{Y}$ solution to the SDE:

$$\mathfrak{Y}_t = x + \int_0^t g(\mathfrak{Y}_s)ds + \int_0^t \sqrt{2\sigma^2(\mathfrak{Y}_s)}dB_s + \int_0^t\int_0^{\mathfrak{Y}_{s^-}}\int_{\mathbb{R}_+} zR(ds,dx,dz) + \int_0^t\int_0^1 (\theta-1)\mathfrak{Y}_{s^-}N_a(ds,d\theta),$$

where g σ, B and R have been defined in (10) and N_a is a PPP with intensity $2rdsd\theta$, independent of B and R. Hence for twice derivable and bounded functions, an application of Itô formula gives:

$$\frac{d\mathbb{E}[f(\mathfrak{Y}_t)]}{dt} = \mathbb{E}\left[\mathcal{G}\mathfrak{Y}_t\right] + 2r\int_0^1 \mathbb{E}\left[f(\theta\mathfrak{Y}_{t-}) - f(\mathfrak{Y}_{t-})\right]\kappa(d\theta).$$

We thus may conclude that

$$\mathbb{E}_{\delta_x}\left[\sum_{u\in V_t} f(X_t^u)\right] = e^{(r-q)t}\mathbb{E}_x\left[f(\mathfrak{Y}_t)\right]. \tag{14}$$

The process $\mathfrak{Y}$ is called auxiliary process and may be exhibited in more complex situations (division rate depending on the quantity of parasites in the cell for instance. However in most cases it is inhomogeneous in time and its study may be out of range. We deduce that a 'typical' lineage behaves as a typical lineage between division events and divides twice more often that a normal lineage. This bias is due to the fact that picking an individual uniformly at random among alive cells at a given time will favor lineages which uderwent more divisions than the other ones. We refer the reader to [BDMT11] for more details on these questions.

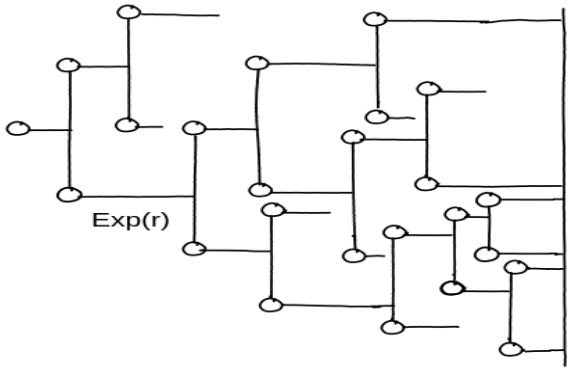

Figure 3: Illustration of the bias phenomenon for a typical lineage.

3.3 Quantity of parasites in the cells

We now consider that the dynamics of the parasites in a cell follows the SDE (10) without the stable positive jumps, that is to say

$$\mathfrak{X}_t = x + \int_0^t g(\mathfrak{X}_s)ds + \int_0^t \sqrt{2\sigma^2(\mathfrak{X}_s)}dB_s. \tag{15}$$

In this case we can observe moderate infections, extinctions of the parasites in the cell population, but also cases where the quantity of parasites goes to infinity with an exponential growth in a positive fraction of the cells.

In order to state the next result, we need to introduce two assumptions. The first one provides a condition under which the quantity of parasites may reach 0. It is almost a necessary and sufficient condition (see [MS21, Remark 3.2 and Theorem 3.3]).

(LN0) There exist $0 < a < 1$, $\eta > 0$ and $x_0 > 0$ such that for all $x \leq x_0$

$$\frac{g(x)}{x} - a\frac{\sigma^2(x)}{x^2} \leq -\ln(x^{-1})\left(\ln\ln(x^{-1})\right)^{1+\eta}.$$

The second assumption ensures that the process does not explode in finite time almost surely (see [MS21, Theorem 4.1]).

(SN∞) There exist $0 < a < 1$, and a nonnegative function f on $\mathbb{R}_+$ such that

$$\frac{g(x)}{x} - a\frac{\sigma^2(x)}{x^2} = -f(x) + o(\ln x), \quad (x \to +\infty).$$

Recall that the total number of cells is given by a continuous-time birth and death process with individual birth rate r and individual death rate q. From classical results on branching processes (see for instance [AN72]), we know that the cell population survives with probability $0 \vee (1 - q/r)$. The long time behaviour for the quantity of parasites in the cells is described in the next proposition. Recall that N_t is the cardinality of V_t.

Proposition 3.2. *Assume that the quantity of parasites in a cell follows the SDE (15) and that $r > q \geq 0$.*

1. *If $\mathbb{E}[\ln^2 \Theta] < \infty$, and there exists $\eta > 0$ such that for $x \geq 0$,*

$$\frac{g(x)}{x} + 2r\mathbb{E}[\ln \Theta(x)] > \eta,$$

and if the function $x \mapsto \sigma^2(x)/x$ is bounded, then for $\varepsilon > 0$,

$$\liminf_{t\to\infty} \mathbb{E}\left[\mathbf{1}_{\{N_t \geq 1\}} \frac{\#\{u \in V_t : X_t^u > e^{(\eta-\varepsilon)t}\}}{N_t}\right] > 0.$$

2. *If there exists $\eta > 0$ such that for all $x \geq 0$,*

$$\frac{g(x)}{x} + 2r\mathbb{E}[\ln \Theta] < -\eta,$$

then for $\varepsilon > 0$

$$\lim_{t\to\infty} \mathbf{1}_{\{N_t \geq 1\}} \frac{\#\{u \in V_t : X_t^u > \varepsilon\}}{N_t} = 0 \quad \textit{in probability.}$$

3. *If Assumption* **(LN0)** *and* **(SN∞)** *hold and there exist $\eta > 0$ and $x_0 \geq 0$ such that for $x \geq x_0$,*

$$\frac{g(x)}{x} - \frac{\sigma^2(x)}{x^2} + 2r\mathbb{E}[\ln \Theta] < -\eta,$$

then

$$\lim_{t\to\infty} \mathbf{1}_{\{N_t \geq 1\}} \frac{\#\{u \in V_t : X_t^u > 0\}}{N_t} = 0 \quad a.s.$$

Let us give an idea of the proof. Let us take a non-negative real function a. From the Many-to-one formula (14) we get that

$$\mathbb{E}_{\delta_x}\left[\sum_{u\in V_t} \mathbf{1}_{\{X_t^u > a(t)\}}\right] = e^{(r-q)t}\mathbb{E}_x\left[\mathbf{1}_{\{\mathfrak{Y}_t > a(t)\}}\right] = e^{(r-q)t}\mathbb{P}_x\left(\mathfrak{Y}_t > a(t)\right).$$

We may thus focus on the process $\mathfrak{Y}$. If we introduce the compound Poisson Process $\mathcal{D}$ describing the log of the jump intensities as follows

$$\mathcal{D}_t = \int_0^t \int_0^1 \ln \theta N_a(ds, d\theta) = \sum_{i \leq N_a(t)} \ln \theta_i$$

(where we used the notation $N_a(t) = \int_0^t \int_0^1 N_a(ds, d\theta)$), we get that

$$\mathfrak{Y}_t \exp\left(-\int_0^t \frac{g(\mathfrak{Y}_s)}{\mathfrak{Y}_s} ds - \mathcal{D}_t\right)$$

is a local martingale. Once again this means that we may focus on the behaviour of

$$\exp\left(\int_0^t \frac{g(\mathfrak{Y}_s)}{\mathfrak{Y}_s} ds + \mathcal{D}_t\right)$$

to study the process $\mathfrak{Y}$. The end of the proof consists in comparing $\int_0^t \frac{g(\mathfrak{Y}_s)}{\mathfrak{Y}_s} ds + \mathcal{D}_t$ with a Lévy process with positive or negative drift.

From Proposition 3.2 we see that in some sense an equal sharing is the worst strategy at the population level. Indeed, from the concavity of the function $x \mapsto \ln x$, we can prove that if the proportion of highly infected cells is positive for large time with a given partitioning strategy, then the equal sharing strategy would have led to the same result. Conversely, if the equal sharing strategy guarantees the healing of the cell population for large time, then it would have been the case for any partitioning strategy.

Lemma 3.3. *Under the assumptions of Proposition 3.2, we have the following:*

1. *If there exists a sharing kernel κ such that the assumptions of point 1 of Proposition 3.2 hold, then they also hold for the equal sharing (corresponding to $\Theta \equiv 1/2$).*
2. *If the assumptions of point 2 (resp. 3) of Proposition 3.2 hold for the equal sharing ($\Theta \equiv 1/2$), then they also hold for any a sharing kernel κ such that $\mathbb{E}[\Theta] = 1/2$ and $|\mathbb{E}[\ln \Theta]| < \infty$.*

It is a consequence of the concavity of the function $x \mapsto \ln x$ entailing via Jensen inequality that

$$\ln(1/2) = \ln \mathbb{E}[\Theta] \geq \mathbb{E}[\ln \Theta].$$

3.4 Mean number of cells alive: General results

We denote by $\mathfrak{C}_t$ the number of cells alive at time t. We consider that the quantity of parasites in a cell follows the SDE:

$$\mathfrak{X}_t = x + g \int_0^t \mathfrak{X}_s ds + \int_0^t \int_0^{\mathfrak{X}_{s^-}} \int_{\mathbb{R}_+} z R(ds, dx, dz), \tag{16}$$

where $g \geq 0$, $x \geq 0$, B is a standard Brownian motion and the Poisson measure R has been defined in (10). In this case, we are able to obtain an equivalent of the mean number of cells alive at a large time t. It emphasizes how crucial is the choice of repartition of the parasites at division between daughter cells.

Let

$$\mu := g + 2r\mathbb{E}\left[\ln \Theta\right] \tag{17}$$

and

$$\mathfrak{d} := \hat{\tau} g + r\left[2\mathbb{E}[\Theta^{\hat{\tau}}] - 1\right] - q, \tag{18}$$

where $\hat{\tau}$ satisfies

$$\frac{g}{2r} = \mathbb{E}[\Theta^{\hat{\tau}} \ln(1/\Theta)].$$

Proposition 3.4. *Assume that the dynamics of the quantity of parasites in a cell follows (16), and that $q \neq r$.*

1. *If $\mu < 0$, then for every $x > 0$ there exists $0 < c_1(x) < 1$ such that*

$$\lim_{t \to \infty} e^{(q-r)t} \mathbb{E}_{\delta_x}\left[\mathfrak{C}_t\right] = c_1(x).$$

2. *If $\mu = 0$, then for every $x > 0$ there exists $c_2(x) > 0$ such that*

$$\lim_{t\to\infty} \sqrt{t} e^{(q-r)t} \mathbb{E}_{\delta_x} [\mathfrak{C}_t] = c_2(x).$$

3. *If $\mu > 0$, then for every $x > 0$ there exists $c_3(x) > 0$ such that*

$$\lim_{t\to\infty} t^{\frac{3}{2}} e^{-\mathfrak{d}t} \mathbb{E}_{\delta_x} [\mathfrak{C}_t] = c_3(x).$$

Note that the dependency on β (parameter of the law of positive jumps for the parasites) is hidden in the limiting functions c_1, c_2, c_3.

Hence, the cell population survives the infection with positive probability if the strategy for repartition of the parasites at division is well-chosen. The sign of μ indicates if the quantity of parasites stays finite with a positive probability in a typical cell line. If it is the case ($\mu \leq 0$), then the expected number of cells alive goes to infinity as time goes to infinity because the cell population grows exponentially at a rate larger than $r - q > 0$. If $\mu > 0$, then the probability that the quantity of parasites is infinite in a typical cell line goes to 1 as time goes to infinity. In that case, the speed of convergence of this probability has to be compared with the growth of the population. And if the growth of the population is strong enough ($\mathfrak{d} > 0$), the expected number of cells that are alive still goes to infinity as time goes to infinity.

Again, we see that the mean of the jump process ($\mathbb{E}[\ln \Theta]$) is not enough to obtain the long term behaviour of the process, but that the variability is also key via the parameter $\mathfrak{d}$ as it was the case for the survival probability of subcritical branching processes in random environment.

Focusing on the role of the partitioning kernel, Proposition 3.4 shows that if the cell population manages to adapt its partitioning strategy to make it more asymmetrical, it can save the cell population (in the sense of making the mean number of cells alive tend to infinity for large time). Indeed, for any choice of the pair (g, r), we can find a kernel κ such that

$$\mathbb{E}[\ln \Theta] \leq -\frac{g}{2r},$$

and thus $\mu \leq 0$.

In absence of parasites, if $r > q$, the cell population evolves as a supercritical Galton-Watson process and survives with probability $1 - q/r$ [AN72]. In the presence of parasites, the condition $r > q$ does not ensure that the cell population survives with a positive probability, as it goes extinct almost surely if $\mathfrak{m} > 0$ and $\mathfrak{d} \leq 0$. More generally, from the previous proposition, we deduce the following corollary on the asymptotic behaviour of $\mathfrak{C}_t$.

Corollary 3.5. *Under the assumptions of Proposition 3.4, for any $x > 0$,*

1. *If $q \geq r$ or if ($\mu > 0$ and $\mathfrak{d} \leq 0$), then $\lim_{t\to\infty} \mathbb{E}_{\delta_x}[\mathfrak{C}_t] = 0$.*

2. *If ($\mu \leq 0$ and $r > q$) or if ($\mu > 0$ and $\mathfrak{d} > 0$), then $\lim_{t\to\infty} \mathbb{E}_{\delta_x}[\mathfrak{C}_t] = \infty$.*

Notice that the case $r = q$ is not taken into account in Proposition 3.4 for the simplicity of its statement as it corresponds to a critical birth and death process for the cell population dynamics. However, we know that in this case the number of cells (with a finite or infinite quantity of parasites) reaches 0 in finite time, hence the number of cells alive also reaches 0 in finite time.

3.5 Mean number of cells alive: Role of the partitioning kernel

The most simple examples of partitioning strategies are the uniform law and the symmetrical sharing. For those laws, we can explicit the bounds of Corollary 3.5.

Corollary 3.6. *Assume that the quantity of parasites in a cell follows the SDE (16) and that $r > q \geq 0$.*

- *For any r, q there exists $g_{\lim}(r,q)$ such that $\lim_{t\to\infty} \mathbb{E}_{\delta_x}[\mathfrak{C}_t] = 0$ if $g \geq g_{\lim}(r,q)$ and $\lim_{t\to\infty} \mathbb{E}_{\delta_x}[\mathfrak{C}_t] = \infty$ if $g < g_{\lim}(r,q)$*
- *If $\kappa(d\theta) = d\theta$,*
$$g_{\lim}(r,q) = 3r - q + 2\sqrt{2r(r-q)}.$$
- *If $\kappa(d\theta) = \delta_{1/2}(d\theta)$,*
$$g_{\lim}(r,q) = r x_0(q/r) \ln 2$$
where $x_0(q/r) > 2$ is the unique value such that
$$x_0(q/r) = \left(1 + \frac{q}{r}\right)\left(1 + \ln 2 - \ln\left(x_0(q/r)\right)\right)^{-1}. \tag{19}$$

From this result, one can prove with a few more computations that the 'uniform sharing' strategy is always better than the 'equal sharing' strategy in terms of survival of the cell population. In fact, the symmetrical sharing is the worst strategy, as stated in the next proposition.

Proposition 3.7. *Assume that the quantity of parasites in a cell follows the SDE (16) and that $r > q \geq 0$. For any partitioning kernel κ, if $g/r < x_0(q/r)\ln 2$, where $x_0(q/r)$ is defined in (19), we have for all $x \geq 0$,*
$$\lim_{t\to+\infty} \mathbb{E}_{\delta_x}[\mathfrak{C}_t] = \infty.$$

As $x_0(q/r)\ln 2$ is the limiting value corresponding to the case of an equal sharing, Proposition 3.7 proves that any other sharing strategy is better than the symmetric partitioning.

More generally, we expect that a more unequal strategy is beneficial for the cell population: it amounts to 'sacrificing' some lineages in order to save the other ones. We were not able to prove such a general statement, but we will try to understand better the effect of unequal sharing in the next two propositions.

Let us introduce the expectation of the minimal fraction of parasites transmited to daughters at division:
$$\vartheta = \mathbb{E}\left[\min(\Theta, 1-\Theta)\right].$$
We will see that this criterium allows us to compare what we call 'deterministic partitioning'
$$\kappa_z(d\theta) = \frac{1}{2}(\delta_z(d\theta) + \delta_{1-z}(d\theta))$$
to any other partinioning law with the same value for ϑ.

We will also wonder if there exists, for any level of infection and for a fixed value of ϑ, a partitioning distribution that leads to survival of the cell population.

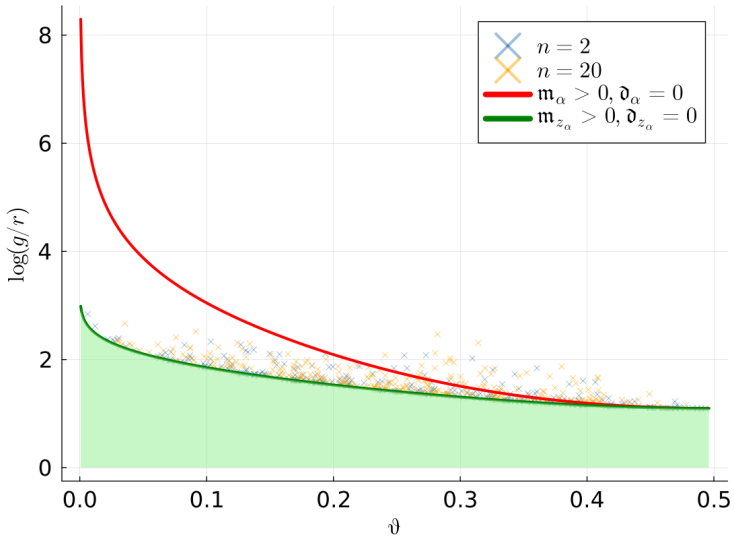

Figure 4: Illustration of the value of $g_{\lim}(r, 0)$ as a function of ϑ (figure from [MS23]). Parameters in the green area lead to survival of the cell population for any finite point partitioning kernel. Each cross corresponds to the limit for a finite point distribution with a given value of ϑ above which the cell population goes to extinction. The orange (resp. blue) crosses correspond to distributions with 20 (resp. 2) modes below $1/2$. The red curve corresponds to the limit above which a cell population with a partitioning kernel of the form $\kappa_\alpha(d\theta) = c_\alpha \theta^\alpha (1-\theta)^\alpha d\theta$ goes to extinction.

In Figure 4, we plot the logarithm of the limiting value g/r at which $\mathfrak{d} = 0$ for various multimodal distributions as a function of ϑ (in the case $q = 0$). We observe that for a fixed value of ϑ, the worst scenario seems to be the case of a deterministic partitioning κ_z. It indeed holds and is stated in the following proposition.

Proposition 3.8. *Assume that the quantity of parasites in a cell follows the SDE (16), and that $r > q \geq 0$. Let $\vartheta \in (0, 1/2]$ and let κ be a symmetric distribution on $[0,1]$ such that*

$$\mathbb{E}\left[\min(\Theta, 1-\Theta)\right] = \vartheta.$$

Finally, let

$$\kappa_\vartheta(d\theta) = 1/2\left(\delta_\vartheta + \delta_{1-\vartheta}\right),$$

be the associated deterministic partitioning kernel. Then, for any $x > 0$,

$$\lim_{t\to+\infty} \mathbb{E}_{\delta_x}^{(\kappa)}[\mathfrak{C}_t] \geq \lim_{t\to+\infty} \mathbb{E}_{\delta_x}^{(\kappa_\vartheta)}[\mathfrak{C}_t],$$

where $\mathbb{E}^{(\kappa)}$ (resp. $\mathbb{E}^{(\kappa_\vartheta)}$) denotes the expectation for the population process with partitioning kernel κ (resp. κ_ϑ).

On the other hand, there is no upper bound: for any value of $\vartheta \in (0, 1/2)$, and any value of $y \geq 0$, one can find a finite point measure (with $n = 2$ for example) such that for all $g/r < y$, the mean number of cells alive goes to infinity when time goes to infinity. This can be achieved by

taking very small values for z_1, which is the smallest atom of the partitioning distribution. This is formally stated in the following proposition.

Proposition 3.9. *Let $g, r, q \in \mathbb{R}_+$ with $q < r$ and $\vartheta \in (0, 1/2)$. Then, there exists a multimodal distribution*

$$\kappa_2(d\theta) = \sum_{i=1}^{2} \left(\delta_{z_i}(d\theta) + \delta_{1-z_i}(d\theta)\right) p_i,$$

with $2(p_1 + p_2) = 1$, and $(z_1, z_2) \in (0, 1/2)^2$, such that if $\Theta \sim \kappa_2$,

$$\mathbb{E}\left[\min(\Theta, 1-\Theta)\right] = \vartheta \quad \textit{and} \quad \lim_{t\to\infty} \mathbb{E}_{\delta_x}[\mathfrak{C}_t] = \infty \quad \textit{for any} \quad x > 0.$$

References

[AG13] Gerold Alsmeyer and Sören Gröttrup. A host-parasite model for a two-type cell population. *Adv. in Appl. Probab.*, 45(3):719–741, 09 2013.

[AG16] Gerold Alsmeyer and Sören Gröttrup. Branching within branching: A model for host,Äìparasite co-evolution. *Stoch. Proc. Appl.*, 126(6):1839 – 1883, 2016.

[AGK+05] Valery I Afanasyev, Jochen Geiger, Götz Kersting, Vladimir A Vatutin, et al. Criticality for branching processes in random environment. *The Annals of probability*, 33(2):645–673, 2005.

[AK71a] Krishna B Athreya and Samuel Karlin. Branching processes with random environments, ii: Limit theorems. *The Annals of Mathematical Statistics*, 42(6):1843–1858, 1971.

[AK71b] Krishna B Athreya and Samuel Karlin. On branching processes with random environments: I: Extinction probabilities. *The Annals of Mathematical Statistics*, 42(5):1499–1520, 1971.

[AN72] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag Berlin, Mineola, NY, 1972. Reprint of the 1972 original [Springer, New York; MR0373040].

[Ban08] Vincent Bansaye. Proliferating parasites in dividing cells: Kimmel's branching model revisited. *Ann. Appl. Probab.*, 18(3):967–996, 2008.

[Ban09a] Vincent Bansaye. Cell contamination and branching processes in a random environment with immigration. *Adv. Appl. Probab.*, 41(4):1059,Äì1081, 2009.

[Ban09b] Vincent Bansaye. Surviving particles for subcritical branching processes in random environment. *Stochastic processes and their applications*, 119(8):2436–2464, 2009.

[BB11] Vincent Bansaye and Christian Böinghoff. Upper large deviations for branching processes in random environment with heavy tails. *Electronic Journal of Probability*, 16:1900–1933, 2011.

[BB13] Vincent Bansaye and Christian Böinghoff. Lower large deviations for supercritical branching processes in random environment. *Proceedings of the Steklov Institute of Mathematics*, 282(1):15–34, 2013.

[BCM19] Vincent Bansaye, Maria-Emilia Caballero, and Sylvie Méléard. Scaling limits of population and evolution processes in random environment. *Electronic Journal of Probability*, 24, 2019.

[BDMT11] Vincent Bansaye, Jean-François Delmas, Laurence Marsalle, and Viet Chi Tran. Limit theorems for markov processes indexed by continuous time galton–watson trees. *The Annals of Applied Probability*, 21(6):2263–2314, 2011.

[BGK05] Matthias Birkner, Jochen Geiger, and Götz Kersting. Branching processes in random environment,Äîa view on critical and subcritical cases. In *Interacting stochastic systems*, pages 269–291. Springer, 2005.

[Bin76] Nicholas H Bingham. Continuous branching processes and spectral positivity. *Stochastic Processes and their Applications*, 4(3):217–242, 1976.

[Böi14] Christian Böinghoff. Limit theorems for strongly and intermediately supercritical branching processes in random environment with linear fractional offspring distributions. *Stochastic Processes and their Applications*, 124(11):3553–3577, 2014.

[BPS13] Vincent Bansaye, Juan Carlos Pardo, and Charline Smadi. On the extinction of continuous state branching processes with catastrophes. *Electronic Journal of Probability*, 18, 2013.

[BS15] Vincent Bansaye and Florian Simatos. On the scaling limits of galton-watson processes in varying environments. *Electronic Journal of Probability*, 20, 2015.

[BT11] Vincent Bansaye and Viet C Tran. Branching feller diffusion for cell division with parasite infection. alea-lat. *Am. J. Probab*, 8(81-127):241–242, 2011.

[Dek87] Frederik Michel Dekking. On the survival probability of a branching process in a finite state iid environment. *Stochastic processes and their applications*, 27:151–157, 1987.

[DH97] Juan Carlos D'souza and Ben M Hambly. On the survival probability of a branching process in a random environment. *Advances in Applied Probability*, 29(1):38–55, 1997.

[DL12] Donald A Dawson and Zenghu Li. Stochastic equations, flows and measure-valued processes. *The Annals of Probability*, 40(2):813–857, 2012.

[Fel71] William Feller. *An Introduction to Probability theory and its application Vol II*. John Wiley and Sons, 1971.

[FL10] Zongfei Fu and Zenghu Li. Stochastic equations of non-negative processes with jumps. *Stochastic Processes and their Applications*, 120(3):306–330, 2010.

[FL20] Rongjuan Fang and Zenghu Li. Construction of continuous-state branching processes in varying environments. *arXiv preprint arXiv:2002.09113*, 2020.

[GK01] Jochen Geiger and Götz Kersting. The survival probability of a critical branching process in a random environment. *Theory of Probability & Its Applications*, 45(3):517–525, 2001.

[GKV03] Jochen Geiger, Götz Kersting, and Vladimir A Vatutin. Limit theorems for subcritical branching processes in random environment. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 39, pages 593–620. Elsevier, 2003.

[GL01] Yves Guivarc'h and Quansheng Liu. Propriétés asymptotiques des processus de branchement en environnement aléatoire. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 332(4):339–344, 2001.

[Gre74] David R Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *Journal of Applied Probability*, 11(4):669–677, 1974.

[Gri74] Anders Grimvall. On the convergence of sequences of branching processes. *The Annals of Probability*, pages 1027–1045, 1974.

[HLX18] Hui He, Zenghu Li, and Wei Xu. Continuous-state branching processes in lévy random environments. *Journal of Theoretical Probability*, 31(4):1952–1974, 2018.

[Jir58] Miloslav Jirina. Stochastic branching processes with continuous state space. *Czechoslovak Mathematical Journal*, 8(2):292–313, 1958.

[Kap74] Norman Kaplan. A note on the supercritical branching processes with random environments. *The Annals of Probability*, 2(3):509–514, 1974.

[Kei75] Niels Keiding. Extinction and exponential growth in random environments. *Theoretical Population Biology*, 8(1):49–63, 1975.

[Kim97] Marek Kimmel. *Quasistationarity in a Branching Model of Division-Within-Division*, pages 157–164. Springer New York, 1997.

[Koz76] Mikhail Vasil'evich Kozlov. On asymptotics of the probability of nonextinction for a critical branching process in random environment. *Teoriya Veroyatnostei i ee Primeneniya*, 21(4):813–825, 1976.

[Kur78] Thomas G Kurtz. Diffusion approximations for branching processes. In *Branching processes (Conf., Saint Hippolyte, Que., 1976)*, volume 5, pages 269–292, 1978.

[KV17] Götz Kersting and Vladimir Alekseevich Vatutin. *Discrete time branching processes in random environment*. Wiley Online Library, 2017.

[Kyp06] Andreas E Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Springer Science & Business Media, 2006.

[Lam67a] John Lamperti. Continuous state branching processes. *Bulletin of the American Mathematical Society*, 73(3):382–386, 1967.

[Lam67b] John Lamperti. The limit of a sequence of branching processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 7(4):271–288, 1967.

[Li10] Zenghu Li. *Measure-valued branching Markov processes*. Springer Science & Business Media, 2010.

[LP12] Zenghu Li and Fei Pu. Strong solutions of jump-type stochastic equations. *Electronic Communications in Probability*, 17(2011):1–13, 2012.

[LX18] Zenghu Li and Wei Xu. Asymptotic results for exponential functionals of lévy processes. *Stochastic Processes and their Applications*, 128(1):108–131, 2018.

[Mar16] Aline Marguet. Uniform sampling in a structured branching population. *arXiv preprint arXiv:1609.05678*, 2016.

[MS21] Aline Marguet and Charline Smadi. Long time behaviour of continuous-state nonlinear branching processes with catastrophes. *Electronic Journal of Probability*, 26:1–32, 2021.

[MS23] Aline Marguet and Charline Smadi. Parasite infection in a cell population: role of the partitioning kernel. *arXiv preprint arXiv:2305.06962*, 2023.

[OW20] Luis Osorio and Anita Winter. Two level branching model for virus population under cell division. Preprint, arXiv:2004.14352, 2020.

[PP17] Sandra Palau and Juan Carlos Pardo. Continuous state branching processes in random environment: The brownian case. *Stochastic Processes and their Applications*, 127(3):957–994, 2017.

[PP18] Sandra Palau and Juan Carlos Pardo. Branching processes in a lévy random environment. *Acta Applicandae Mathematicae*, 153(1):55–79, 2018.

[PPS16] Sandra Palau, Juan Carlos Pardo, and Charline Smadi. Asymptotic behaviour of exponential functionals of l\'evy processes with applications to random processes in random environment. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2):1235,Äì1258, 2016.

[SW69] Walter L Smith and William E Wilkinson. On branching processes in random environments. *The Annals of Mathematical Statistics*, 40(3):814–827, 1969.

[VDS13] Vladimir A. Vatutin, Elena E. Dyakonova, and Serik Sagitov. Evolution of branching processes in a random environment. *Proceedings of the Steklov Institute of Mathematics*, 282(1):220–242, 2013.