



**HAL**  
open science

# Phenotypic plasticity trade-offs in an age-structured model of bacterial growth under stress

Meriem El Karoui, Ignacio Madrid, Sylvie Méléard

► **To cite this version:**

Meriem El Karoui, Ignacio Madrid, Sylvie Méléard. Phenotypic plasticity trade-offs in an age-structured model of bacterial growth under stress. 2024. hal-04511813

**HAL Id: hal-04511813**

**<https://polytechnique.hal.science/hal-04511813v1>**

Preprint submitted on 19 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Phenotypic plasticity trade-offs in an age-structured model of bacterial growth under stress

Meriem El Karoui\*

Ignacio Madrid Canales†

Sylvie Méléard‡

March 19, 2024

## Abstract

Under low concentrations of antibiotics causing DNA damage, *Escherichia coli* bacteria can trigger stochastically a stress response known as the SOS response. While the expression of this stress response can make individual cells transiently able to overcome antibiotic treatment, it can also delay cell division, thus impacting the whole population’s ability to grow and survive. In order to study the trade-offs that emerge from this phenomenon, we propose a bi-type age-structured population model that captures the phenotypic plasticity observed in the stress response. Individuals can belong to two types: either a fast-dividing but prone to death “vulnerable” type, or a slow-dividing but “tolerant” type. We study the survival probability of the population issued from a single cell as well as the population growth rate in constant and periodic environments. We show that the sensitivity of these two different notions of fitness with respect to the parameters describing the phenotypic plasticity differs between the stochastic approach (survival probability) and the deterministic approach (population growth rate). Moreover, under a more realistic configuration of periodic stress, our results indicate that optimal population growth can only be achieved through fine-tuning simultaneously both the induction of the stress response and the repair efficiency of the damage caused by the antibiotic.

Keywords: Multi-type branching process, Phenotypic plasticity, Stress response, Population growth rate, Long-time behaviour, Stochastic Individual-Based Model

Mathematics Subject Classification: 60J85, 45C05, 92D25, 95C70, 92D15

## 1 Introduction and stochastic model

Under the presence of antibiotics and other stress factors, bacteria can exhibit a dynamic and heterogeneous expression of stress-response genes. In the case of *Escherichia coli* growing under a sublethal concentration of an antibiotic causing DNA damage such as ciprofloxacin, the detection of DNA breaks on the chromosome triggers the initiation of the DNA damage response, which is

---

\*University of Edinburgh, School of Biological Sciences, Institute of Cell Biology; E-mail: meriem.elkaroui@ed.ac.uk

†Ecole Polytechnique, CNRS, Institut polytechnique de Paris, Inria, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: ignacio.madrid-canales@polytechnique.edu

‡Ecole Polytechnique et Institut Universitaire de France, CNRS, Institut polytechnique de Paris, Inria, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: sylvie.meleard@polytechnique.edu

called SOS response [29, 33, 42]. The intensity of this response depends on the amount of damage, but it also exhibits high heterogeneity among individuals, even in a isogenic population, because of the stochastic expression of several factors [2, 24, 25]. Observations at single-cell level of this stress response have shown that the heterogeneity of the SOS response is strongly dependent on the growth conditions [24]. Moreover, for each individual cell, the SOS response can fluctuate substantially in time, often as a sequence of pulses, transitioning from periods of apparent SOS inactivity to periods of strong SOS response [20, 27, 37]. A distinct signature of the high SOS expression is the highly perturbed division dynamics caused by the SOS-dependent expression of a cell division inhibitor that induces, in turn, inter-division times which are much longer than in non-SOS-inducing cells. At the same time, however, cell elongation is not repressed by the SOS response, leading to the production of cells several times longer than the normally observed ones, a phenomenon known as filamentation. Importantly, SOS-inducing cells are able to produce non-SOS-inducing offspring, which results in a subset of cells that divide normally and could be able to rapidly take over the population once the stress is stopped [26, 35, 41]. Interestingly, similar phenotypic variability has also been shown in other stress response systems, impacting cell elongation or division with important consequences for the survival of the population [36]. Collectively, these observations suggest strong links between single-cell phenotypic heterogeneity and population-level stress survival, which might be a key to explaining antibiotic tolerance [1, 2, 4].

Nonetheless, at least theoretically, the unconstrained heterogeneous expression of stress responses might lead to poor population-level performance, especially if the intensity of the stress response is anti-correlated to the mechanisms that usually contribute to the population fitness, such as fast division (inhibited by the SOS response) and volume increase (inhibited by other general stress responses). Furthermore, the expression of stress response genes has been shown to be generally within the noisiest in *E. coli*'s proteome [38]. Thus, expression of the stress response might allow cells to survive in the short term but may come at the expense of their capacity to grow and divide, giving rise to a trade-off between individual stress expression and population growth.

Here, to shed some conceptual and quantitative light on this question, we propose a minimal stochastic model in continuous time of a structured population that preserves the main elements that characterise the stress responses described above. First, each individual cell is characterised by their age  $a \geq 0$  and a type  $i \in \{0, 1\}$ . The type  $i = 0$  stands for the *vulnerable cells* which do not induce a stress response, and are therefore fast dividing but can die at division with probability  $p \in [0, 1]$ . Otherwise, with probability  $1 - p$ , such a cell will produce two new identical daughter cells. Although there is no biological reason for death to happen at division, it seems a reasonable approximation to account in a minimal and very simple way the fact that the considered stresses (i.e. antibiotics) target only proliferating bacteria. On the other hand, type  $i = 1$  are *tolerant cells* which react to the stress by inducing a response (e.g. by repairing their DNA in the case of the SOS response) which transiently protects them but leads them to divide after much longer times.

We include age as a variable for the following reasons. Firstly, we want to account for memory-effects for both normally dividing and slowly dividing cells. Indeed, the distributions of division times, even in the absence of antibiotic and under ideal conditions, are not exponential and are much better explained by probability laws with memory, thus requiring the inclusion of an age-like structure in the construction of the model [16]. Secondly, we use age as a proxy for other cellular characteristics. For example, in the case of SOS-inducing cells, filamentous cells correspond to old individuals of type 1. We, therefore, suppose that for each individual cell, the division mechanism is triggered by its age in a phenotype-dependent way. In particular, we suppose that the distribution

of the interdivision times of cells of type  $i$  are driven by an age-dependent division rate  $\beta_i$  in the sense that for all individuals we have

$$\mathbb{P}(\text{Divide at age } < a + \Delta a | \text{Type} = i, \text{Age} \geq a) = \beta_i(a) \cdot o(\Delta a).$$

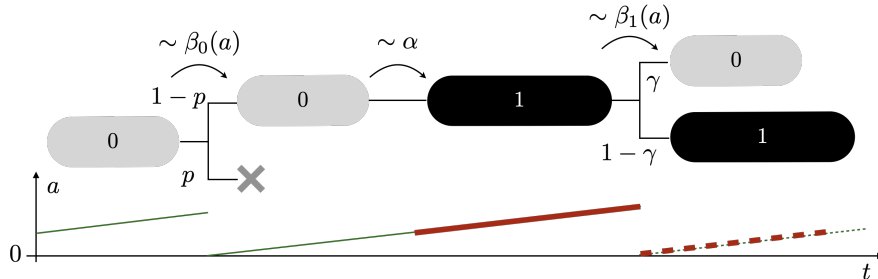


Figure 1: Schematic representation of the cell division mechanism and age.

Following the approach introduced by [19, 39], we represent the population as a measure-valued stochastic process  $(Z_t)_{t \geq 0}$  in continuous time, which at each time  $t \geq 0$  can be written as a point measure

$$Z_t = \sum_{k=1}^{N_t} \delta_{(A_k(t), I_k(t))} \quad (1)$$

where  $N_t = \langle Z_t, 1 \rangle$  is the total number of cells alive at time  $t$ , and  $(A_k(t), I_k(t)) \in \mathbb{R}_+ \times \{0, 1\}$  for  $k \in \{1, \dots, N_t\}$  is the age and phenotype at time  $t$  of cell number  $k$ , for any fixed arbitrary order (e.g. numbered using Neveu's notation, or by lexicographical order). In Appendix A.1 we define rigorously the paths of  $Z_t$  as the solution of a Stochastic Differential Equation, which describes the desired dynamics conditionally to a given initial population.

We suppose that a cell switch between type 0 and 1 occurs after a random exponential time with rate parameter  $\alpha \geq 0$ . In particular, this means that the  $0 \rightarrow 1$  switch is memoryless and can take place at any moment of the cell cycle, independently of the age of the cell. Importantly, the switch and division dynamics are supposed independent, so it is possible for a cell of type 0 to never switch, if its division time (determined by the value of  $\beta_0$  at its current age) occurs before the intended switching time (determined by  $\alpha$ ). On the other hand, we suppose that during the lifetime of a cell of type 1, switching back from type 1 to type 0 is impossible. However, when a cell of type 1 divides, its offspring can become of type 0 as consequence of a successful DNA repair. Specifically, we suppose that each daughter resulting by the division of a cell of type 1 can be born with type 0 with probability  $\gamma \in [0, 1]$ , or otherwise keep type 1 with probability  $1 - \gamma$ . Thus  $\gamma$  can be interpreted as a measure of the DNA repair efficiency from one generation to the next one.

Table 1 summarises all these parameters and Fig 1 represents the described dynamics.

Several models have been studied in the context of other similar phenomena relating to seed banks and others (see the review by Lennon et al. [28] and the references therein). In the specific case of microbiological populations, more recently, Blath et al. [6, 7] studied bi-type non-structured populations of *active* and *dormant* cells in a discrete-time setting and analysed the effect of different switching strategies in a randomly varying environment. In this work, we generalise these fundamental ideas to the case of stress response, where the division of tolerant individuals is not stopped but only affected, and give a detailed study of the fitness sensibility with respect to the parameters from both a probabilistic and deterministic approach.

Parameter	Definition
$\beta_i(a) \in \mathbb{R}_+$	Division rate at age $a$ and type $i$
$\alpha \in \mathbb{R}_+$	Switching rate from type 0 to 1 (phenotypic plasticity)
$p \in [0, 1]$	Probability of death at division for cells of type 0 (environmental effect)
$\gamma \in [0, 1]$	Probability that a type 1 divides and produces a type 0 (DNA repair effect)

Table 1: Summary of the model parameters.

Sections 3 to 6 give the main results of the paper and discuss their biological implications, whilst we postpone all the proofs and technical details to Sections 7 to 9. In Section 3 we derive the extinction probability of a population initiated by a single cell in an environment with a constant stress signal. Moreover, we obtain an explicit condition on the model parameters under which this population establishes with positive probability. We relate this result with the capacity of *evolutionary rescue* [5, 11] of phenotype-switching populations. In Section 4 we show that under the survival condition, we can observe a Malthusian behaviour for the first-moment semigroup of the stochastic process, characterised by an exponential population growth rate  $\lambda > 0$  and a stationary distribution of types and ages. Moreover, we show that there is an equivalence relation between the criterion for the establishment of the population obtained by stochastic and deterministic approaches, which is, in general, not trivial for infinite-dimensional branching processes. Indeed, the measure of *fitness* obtained from each approach, corresponds to different biological realities, as discussed in [15, 17, 31]. In the stochastic case it corresponds to the probability that the branching process initiated by a single individual survives forever, while in the macroscopic case it is the asymptotic rate of exponential growth of the population. Our results extend the parallel results of Campillo et al. in [9, 10] for a mono-type growth-fragmentation-death case.

However, in contrast with [10], we show in Section 5 that the extinction probability and the population growth rate do not always vary in the same direction with respect to variations in the parameter space. In particular, if the stress is low enough ( $p < \bar{p}$  for some critical value  $\bar{p} < 1/2$ ), increasing  $\gamma$  will lead to a decreased survival probability of the population, but at the same time, to a higher population growth rate. Our proofs show that the loss of the classical monotonic behaviour arises from the crucial Assumption 2.4 that vulnerable individuals divide faster than tolerant individuals. In a framework of size-structured cell growth, Calvez et al. in [8] have also shown such loss of monotonicity and Cloez et al. in [14], studied the variation of the population growth rate parameter with respect to an asymmetry factor.

Our results show that the optimal parameters in the sense of any of the two notions of fitness only correspond to extreme strategies  $\gamma = 0$  or  $\gamma = 1$ . From both the point of view of the population survival probability, the creation of prone-to-death type 1 individuals is always detrimental, favouring a null recovery probability  $\gamma = 0$ . The same occurs from both the point of view of the population growth rate under high stress ( $p > \bar{p}$ ). In that situation, an optimal population is expected to produce only type 1 “tolerant” individuals with  $\gamma = 0$ , since any creation of cells of type 0 is also detrimental for growth. Under low stress ( $p \leq \bar{p}$ ), from the point of view of the population growth rate, it is optimal to produce only fast-dividing type 0 individuals, favouring  $\gamma = 1$ .

A more realistic aspect of such switching strategies appears when the environment (i.e., the stress parameter  $p$ ) is allowed to fluctuate in time. In Section 6, using Floquet’s theory (cf. [12, 13]), we extend our results to the case where we consider a  $T$ -periodic signal  $p(t)$ . Although, in that case, our results cannot be equally quantitative, we show that the sign of the fitness variation with

respect to the model parameters is not constant. Numerical simulations show that from the point of view of population growth rate, the optimal parameters change drastically and might require the simultaneous fine-tuning of both the recovery ( $\gamma$ ) and switching ( $\alpha$ ). We conclude with an outlook on the biological consequences of general stress-response strategies, particularly the SOS response on survival probabilities and population growth rate.

## 2 Model assumptions and first properties

We will work under the following set of assumptions.

### Assumptions 2.1.

- (A1) Division times are almost surely finite: For all  $i \in \{0, 1\}$ ,  $\int_0^{+\infty} \beta_i(a) da = +\infty$ .
- (A2) Division rates are uniformly bounded:  $\exists \bar{b} > 0$  such that  $\forall i \in \{0, 1\}, a \geq 0, \beta_i(a) \leq \bar{b}$ .
- (A3) Division rates are uniformly bounded by below from a certain age:  $\exists a_0 > 0, \underline{b} > 0$  such that  $\forall i \in \{0, 1\}, \beta_i(a) \geq \underline{b}$  whenever  $a \geq a_0$ .

Under assumption (A1) we define the survival probabilities for an individual of type 0 or 1 not to experience any event between ages  $s$  and  $t$ , given that it has already survived until age  $s$ :

$$\psi_0(s, t) = \exp\left(-\int_s^t (\alpha + \beta_0(u)) du\right), \quad (2)$$

$$\psi_1(s, t) = \exp\left(-\int_s^t \beta_1(u) du\right), \quad (3)$$

and we define the  $\star$  composition between the two survival functions as

$$\psi_0 \star \psi_1(s, t) = \int_s^t \psi_0(s, u) \psi_1(u, t) du. \quad (4)$$

In particular,  $\alpha \psi_0 \star \psi_1(s, t)$  represents the probability for an individual of type 0 at initial age  $s$ , to switch to type 1 at some point between ages  $s$  and  $t$  and survive the remaining time until it has reached age  $t$ , given that it has already survived until age  $s$ .

We provide now some useful definitions that will be used in the next paragraphs.

**Definition 2.2.** 1. For any finite positive point measure  $\mu$ , set  $\mathbb{P}_\mu$  the probability under the initial condition  $Z_0 = \mu$ , and  $\mathbb{E}_\mu$  the respective expectation.

2. Let  $T$  be the first jump of  $Z$ .

3. Let  $(I_1, I_2)$  be the types of the two daughters obtained after the first jump, if it was a division.

Lemma 2.3 stated below characterises the probability laws of these random variables, and will be used to compute the probability that a population initiated by a single bacterium goes extinct.

**Lemma 2.3.** *Under Assumptions 2.1,*

1. The probability that one cell of initial state  $(a, i)$  dies before time  $t_0 \geq 0$ , instead of dividing or switching before that, is

$$\mathbb{P}_{\delta_{(a,i)}}(Z_T = 0, T \leq t_0) = \mathbb{1}_{i=0} p \int_0^{t_0} \beta_i(a+t) \exp\left(-\int_0^t \beta_i(a+u) du - \alpha t\right) dt. \quad (5)$$

2. For any bounded measurable function  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ , the conditional law of the switching time is characterised by

$$\mathbb{E}_{\delta_{(a,0)}} \left[ h(T) \mathbb{1}_{Z_T = \delta_{(a+T,1)}, T \leq t_0} \right] = \int_0^{t_0} h(t) \alpha \exp\left(-\int_0^t \beta_0(a+u) du - \alpha t\right) dt. \quad (6)$$

3. For any bounded measurable function  $h : \{0, 1\} \times \{0, 1\} \rightarrow \mathbb{R}$ , the conditional law of the daughter types after division is characterised by

$$\begin{aligned} \mathbb{E}_{\delta_{(a,i)}} \left[ h(I_1, I_2) \mathbb{1}_{Z_T = \delta_{(0,I_1)} + \delta_{(0,I_2)}, T \leq t_0} \right] = \\ \int_0^{t_0} \beta_i(a+t) \exp\left(-\int_0^t \beta_i(a+u) du - (1-i)\alpha t\right) dt \left\{ \mathbb{1}_{i=0} (1-p)h(0,0) \right. \\ \left. + \mathbb{1}_{i=1} \left( \gamma^2 h(0,0) + (1-\gamma)^2 h(1,1) + \gamma(1-\gamma)(h(0,1) + h(1,0)) \right) \right\} \quad (7) \end{aligned}$$

*Proof.* By the construction introduced in Definition A.1, if the initial population consists on only one individual,  $Z_0 = \delta_{(a,i)}$ , then  $T$  is the first jump time of the process

$$J_t = \int_0^t \int_{\{1\} \times \mathbb{R}_+ \times [0,1]^2} \mathbb{1}_{\{z \leq \alpha(1-i) + \beta_i(a+u)\}} \mathcal{N}(du, di, dz, d\omega),$$

which, by definition of  $\mathcal{N}$ , is a non-homogeneous Poisson process whose time-dependent rate is then given by  $t \mapsto ((1-i)\alpha + \beta_i(a+t))$ . Therefore  $T$  has the probability distribution

$$\mathbb{P}_{\delta_{(a,i)}}(T > t) = \mathbb{P}_{\delta_{(a,i)}}(J_t = 0) = \exp\left(-\int_0^t ((1-i)\alpha + \beta_i(a+u)) du\right). \quad (8)$$

From Assumptions 2.1 (A1) we can then deduce that  $\mathbb{P}_{\delta_{(a,i)}}(T < +\infty) = 1$ . Moreover, by differentiation we get that  $T$  admits the probability density function

$$\mathbb{P}_{\delta_{(a,i)}}(T \in [t + dt]) = ((1-i)\alpha + \beta_i(a+t)) \exp\left(-\int_0^t ((1-i)\alpha + \beta_i(a+u)) du\right) dt. \quad (9)$$

Now, the value of the process  $Z_t$  at time  $t = T$ , i.e. just after the first jump, is given by

$$Z_T = \begin{cases} 0 & \text{(death) with probability } \frac{(1-i)p\beta_i(a+T)}{(1-i)\alpha + \beta_i(a+T)} \\ \delta_{(0,I_1)} + \delta_{(0,I_2)} & \text{(division) with probability } \frac{((1-i)(1-p) + i)\beta_i(a+T)}{(1-i)\alpha + \beta_i(a+T)} \\ \delta_{(T,1)} & \text{(switch) with probability } \frac{(1-i)\alpha}{(1-i)\alpha + \beta_i(a+T)} \end{cases}. \quad (10)$$

The equations of Lemma 2.3 are obtained by computing joint probabilities and expectations using the marginal probability density of  $T$ , obtained from (9), and the conditional probability of  $Z_T$  given  $T$ , (10). For example, we have

$$\begin{aligned} \mathbb{P}_{\delta_{(a,i)}}(Z_T = 0, T \leq t_0) &= \mathbb{E}_{\delta_{(a,i)}}[\mathbb{P}(Z_T = 0|T)\mathbb{1}_{T \leq t_0}] \\ &= \mathbb{E}_{\delta_{(a,i)}}\left[\frac{(1-i)p\beta_i(a+T)}{(1-i)\alpha + \beta_i(a+T)}\mathbb{1}_{T \leq t_0}\right] \\ &= \mathbb{1}_{i=0} \int_0^{t_0} \frac{p\beta_i(a+t)}{\alpha + \beta_i(a+t)} (\alpha + \beta_i(a+t)) \exp\left(-\int_0^t (\alpha + \beta_i(a+u)) du\right) dt. \end{aligned}$$

We obtain analogously the conditional law of the switching time. For the conditional law of the daughter types, we note that, if  $i = 0$ , then  $I_1 = I_2 = 0$ , and that, if  $i = 1$ , then  $I_1$  and  $I_2$  are independent Bernoulli random variables of parameter  $1 - \gamma$ .  $\square$

We are interested by the case motivated in Section 1, in which individuals of type 1, while tolerant, divide slower than individuals of type 0. We will only use this assumption to study the sensitivity of the population fitness with respect to the model parameters, starting in Section 5.

**Assumptions 2.4.** Let  $T_{div}$  be the time of division of a non-switching cell, this is,

$$\mathbb{P}_{\delta_{(0,i)}}(T_{div} \geq a) = \exp\left(-\int_0^a \beta_i(s) ds\right).$$

We suppose that the type 0 division time is stochastically dominated (in first order) by the type 1 division time:

$$\mathbb{P}_{\delta_{(0,0)}}(T_{div} \geq a) < \mathbb{P}_{\delta_{(0,1)}}(T_{div} \geq a).$$

**Remark 2.5.** For individuals of type 0,  $T_1 = \min(T_{div}, T_{switch})$ , where  $T_{switch}$  is a Exponential random variable of parameter  $\alpha$ . Therefore, for all  $\alpha \geq 0$ , the first event time  $T_1$  is stochastically larger for individuals of type 1 than for individuals of type 0, which implies  $\psi_0(0, a) < \psi_1(0, a)$ . Moreover, for all non-decreasing function  $g$  we we have  $\mathbb{E}_{\delta_{(0,0)}}[g(T_1)] < \mathbb{E}_{\delta_{(0,1)}}[g(T_1)]$ .

### 3 Conditions for population establishment

We now give conditions for the parameters  $\alpha, \gamma$  such that for two given division rates  $\beta_0, \beta_1$  and a death probability  $p$ , the population initiated by a single initial cell establishes indefinitely. We begin by computing the probability that the population initiated by a single cell goes extinct.

**Definition 3.1.** Let  $\pi(a, i)$  be the extinction probability of a population initiated by a single cell of type  $i \in \{0, 1\}$  and age  $a \geq 0$ . This is

$$\pi(a, i) := \mathbb{P}_{\delta_{(a,i)}}(\exists t > 0 : N_t = 0). \quad (11)$$

Set  $\pi_i = \pi(0, i)$  the extinction probability associated with a single initial cell of age 0.



We will focus now on what happens when the initial cell has age 0. We will see, nevertheless, that the extinction probabilities of a population issued from a single cell of any initial age  $a > 0$  can be obtained explicitly from the two extinction probabilities  $\pi_0$  and  $\pi_1$ .

**Theorem 3.2.** *Let  $T_1$  be the time of the first jump event. For all  $a \geq 0$ , let  $q_a$  the probability that an individual of type 0 switches before dividing, conditionally to have already survived until age  $a$ :*

$$q_a = \mathbb{P}_{\delta_{(0,0)}}(Z_{T_1} = \delta_{(T_1,1)} | T_1 \geq a) = \mathbb{P}_{\delta_{(0,a)}}(Z_{T_1} = \delta_{(T_1,1)}) = \int_0^{+\infty} \alpha \psi_0(a, t) dt$$

and let in particular

$$q = q_0 = \mathbb{P}_{\delta_{(0,0)}}(Z_{T_1} = \delta_{(T_1,1)}) = \int_0^{+\infty} \alpha \psi_0(0, t) dt. \quad (12)$$

The vector of extinction probabilities  $(\pi_0, \pi_1)$  is the smallest solution on  $[0, 1]$  of the quadratic system

$$\begin{cases} \pi_0 = (1 - q)p + (1 - q)(1 - p)\pi_0^2 + q\pi_1, & (13a) \\ \pi_1 = (\gamma\pi_0 + (1 - \gamma)\pi_1)^2, & (13b) \end{cases}$$

in the sense that for any other admissible solution  $\tilde{\pi}$ , we have  $\pi_i \leq \tilde{\pi}_i$  for both  $i \in \{0, 1\}$ . Moreover, for any  $a \geq 0$  we can obtain  $\pi(a, i)$  as explicit functions of  $\pi_0$  and  $\pi_1$  (which justifies our analysis focused in the initial condition 0), as given by

$$\pi(a, 0) = (1 - q_a)p + (1 - q_a)(1 - p)\pi_0^2 + q_a\pi_1 \quad (14)$$

$$\pi(a, 1) = \pi_1. \quad (15)$$

The proof of this theorem is postponed to Section 7.

**Remark 3.3.** Since cells of type 1 do not die in their generation, the division rate  $\beta_1$  does not play any role in the extinction probability.

**Remark 3.4.** Thanks to the integrability assumption (A1), we can differentiate the integral (12) to obtain  $\partial_\alpha q|_{\alpha=0} = \mathbb{E}_{\delta_{(0,0)}}[T_{\text{div}}]$ , where  $T_{\text{div}}$  is the division time of cells of type 0 (which depends only on  $\beta_0$ ). Therefore, at least for small values of  $\alpha$ , the sensitivity of  $q$  with respect to  $\alpha$  is proportional to the mean division time of vulnerable cells. We give a more practical example below.

**Example 3.5** (Gamma distributed inter-division times). Suppose that  $\beta_0$  is such that for some  $a_0, b_0 > 0$ , we can write for all  $t \geq 0$

$$\beta_0(t) \exp\left(-\int_0^t \beta_0(u) du\right) = \frac{b_0^{a_0}}{\Gamma(a_0)} t^{a_0-1} e^{-b_0 t}.$$

Then, the inter-division times are Gamma random variables of shape parameter  $a_0$  and rate parameter  $b_0$ . This has been shown to be a good parametric model to explain the distributions of division ages [22]. An integration by parts of (12) shows that under this assumption,

$$q = 1 - \left(1 + \frac{\alpha}{b_0}\right)^{-a_0}.$$

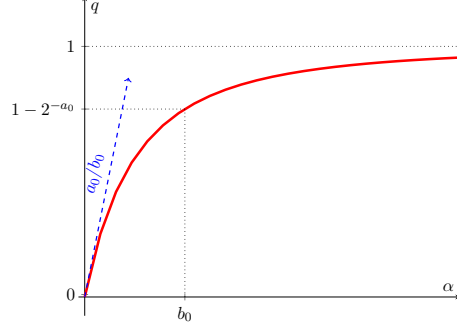


Figure 2: Form of  $q$  as function of  $\alpha$  in the case of division times following a Gamma distribution of parameters  $(a_0, b_0)$  (Example 3.5). Note that  $q$  is always an increasing function of  $\alpha$ .

The general shape for  $q$  as function of  $\alpha$  is given in Fig. 2. Calibrating the values of  $a_0$  and  $b_0$  can modify the sensibility of  $q$  with respect to  $\alpha$ . In particular, choosing  $a_0 = 1$  reduces to the memoryless case  $\beta_0 \equiv b_0$ , where age does not affect the division times, which are then identically distributed exponential random variables of rate parameter  $b_0$ . Generally, notice that the derivative at the origin of  $q$  with respect to  $\alpha$  is equal to  $a_0/b_0$ , which is the expected division time in the Gamma case.

In particular, we can characterise explicitly the subcritical region in which extinction happens almost surely as a function of  $\gamma$ ,  $p$  and the probability  $q$  introduced above. The proof of Theorem 3.6 is postponed to Section 7.

**Theorem 3.6.** *The population initiated by a single cell of age 0 survives with positive probability if and only if*

$$\left\{ p \leq \frac{1}{2} \right\} \text{ or } \left\{ p > \frac{1}{2} \text{ and } \gamma < \frac{1}{2} \left( 1 + \frac{q}{(2p-1)(1-q)} \right) \right\}. \quad (16)$$

**Corollary 3.7.** *Let  $p > 1/2$ . Then the extinction probability of the population is at most  $\frac{1}{2}(1 - \log 2)$ , a value which is attained in the extreme case  $p = 1$  where individuals of type 0 die almost surely at each division attempt.*

*Proof.* The proof consists simply in calculating the area of the extinction region delimited by the complement of condition (16). Indeed, the probability that the population traits belong to this region (so that the establishment probability is 0) is equal to

$$\begin{aligned} \iint_{[0,1]^2} \mathbb{1}_{\left\{ \gamma > \frac{1}{2} \left( 1 + \frac{q}{(2p-1)(1-q)} \right) \right\}} dq d\gamma &= \int_0^{\frac{2p-1}{2p}} dq \int_{\frac{1}{2} \left( 1 + \frac{q}{(2p-1)(1-q)} \right)}^1 d\gamma \\ &= \frac{1}{2} \left( p^2 + \frac{1}{2p-1} \log \left( 1 - \frac{p(2p-1)}{2} \right) \right). \end{aligned}$$

We can see easily that it is an increasing function of  $p$  for  $p > 1/2$  and thus its maximum is attained for  $p = 1$  which gives the value  $\frac{1}{2}(1 - \log 2)$ .  $\square$

**Remark 3.8.** A maximum *evolutionary risk* of  $\frac{1}{2}(1 - \log 2) \approx 15.34\%$  might seem not very restrictive. However, one could argue that keeping a high value of  $q$ , i.e. a high phenotypic plasticity, could be associated with a high energetic cost and other constraints not included in our model, which would enlarge the zone of non viability.

**Remark 3.9.** Note that the previous computation can also be interesting in an evolutionary framework, since it gives a measure of the capacity of random evolutionary rescue. If a mutant cell appears with  $p > 1/2$  and new characteristics  $(\gamma', q')$  that are independently and uniformly distributed on the square  $[0, 1]$ , the survival probability of its subpopulation will be at least  $\frac{1}{2}(1 + \log 2)$  and minimum for  $p = 1$ .

## 4 Long-time behaviour of the population and links with the population establishment condition

In the following we establish some fundamental links between necessary and sufficient conditions for the establishment of a population issued from a single cell from both a probabilistic approach based on the trajectories of  $Z_t$ , and for a deterministic approach based on the long-time behaviour of the semigroup  $M_t$ , which describes the expected value of the population dynamics, as defined below. Dichotomy properties linking the survival probability with the behaviour of  $M_t$  have been studied in size-structured models by [10]. Numerical studies have been performed in the same spirit by [21]. We show in particular that under the survival conditions there is a positive Malthusian parameter  $\lambda > 0$  such that the rescaled dynamics  $e^{-\lambda t} M_t$  converge to a non-zero stationary measure, and that this convergence is at exponential rate.

**Definition 4.1** (First-moment semigroup and vector representations). Let us define over the space of bounded Borel functions  $\mathcal{B}_b(\mathbb{R}_+ \times \{0, 1\})$ , the first-moment semigroup  $M_t : \mathcal{B}_b(\mathbb{R}_+ \times \{0, 1\}) \rightarrow \mathcal{B}_b(\mathbb{R}_+ \times \{0, 1\})$  by

$$M_t f(a, i) = \mathbb{E}_{\delta_{(a,i)}} [\langle Z_t, f \rangle], \quad \forall (a, i) \in \mathbb{R}_+ \times \{0, 1\} \quad (17)$$

and for all signed Borel measure  $\mu \in \mathcal{M}(\mathbb{R}_+ \times \{0, 1\})$  we define  $\mu M_t \in \mathcal{M}(\mathbb{R}_+ \times \{0, 1\})$  as the measure which, for all  $f \in \mathcal{B}_b(\mathbb{R}_+ \times \{0, 1\})$ , verifies the duality relation

$$\langle \mu M_t, f \rangle = \langle \mu, M_t f \rangle = \int_{\mathbb{R}_+ \times \{0, 1\}} M_t f(a, i) \mu(da, di). \quad (18)$$

We give also a vector representation, which will be useful in the sequel. Let us write  $\mathbf{f} = (f(\cdot, 0), f(\cdot, 1)) \in (\mathcal{B}_b(\mathbb{R}_+))^2$  and define for all  $a \geq 0$  the matrix semigroup  $\mathbf{M}_t : (\mathcal{B}_b(\mathbb{R}_+))^2 \rightarrow (\mathcal{B}_b(\mathbb{R}_+))^2$  as

$$\mathbf{M}_t \mathbf{f}(a) = (M_t f(a, 0), M_t f(a, 1)) \in \mathbb{R}^2.$$

Analogously, for all Borel set  $A$  of  $\mathbb{R}_+$ , we let  $\boldsymbol{\mu}(A) = (\mu(A \times \{0\}), \mu(A \times \{1\})) \in (\mathcal{M}(\mathbb{R}_+))^2$  and define

$$\boldsymbol{\mu} \mathbf{M}_t(A) = (\mu M_t(A \times \{0\}), \mu M_t(A \times \{1\})) \in \mathbb{R}^2.$$

Taking expectations in the semi-martingale decomposition (55) associated with  $Z_t$ , we easily show that the infinitesimal generator  $\mathcal{Q}$  associated to  $\mathbf{M}_t$  such that for all  $a \geq 0$

$$\frac{d}{dt} \mathbf{M}_t \mathbf{f}(a) = \mathbf{M}_t (\mathcal{Q} \mathbf{f})(a) = \mathcal{Q} (\mathbf{M}_t \mathbf{f})(a),$$

is given by

$$\mathcal{Q}\mathbf{f}(a) = \mathbf{f}'(a) - \mathbf{D}(a)\mathbf{f}(a) + 2\mathbf{B}(a)\mathbf{f}(0) \quad (19)$$

where  $\mathbf{f}'(a) := (\partial_a f(a, 0), \partial_a f(a, 1))$  and

$$\mathbf{B}(a) = \begin{bmatrix} (1-p)\beta_0(a) & 0 \\ \gamma\beta_1(a) & (1-\gamma)\beta_1(a) \end{bmatrix} \text{ and } \mathbf{D}(a) = \begin{bmatrix} \alpha + \beta_0(a) & -\alpha \\ 0 & \beta_1(a) \end{bmatrix}. \quad (20)$$

An useful alternative approach is the following representation of  $\mathbf{M}_t\mathbf{f}$  as the mild solution to a renewal equation:

**Proposition 4.2** (Forward Equation). *For all test function  $\mathbf{f} \in (\mathcal{B}_b(\mathbb{R}_+))^2$  in the form introduced above, the right action of the semigroup  $\mathbf{M}_t\mathbf{f}$  is solution to the renewal equation, for all  $a \geq 0$*

$$\mathbf{M}_t\mathbf{f}(a) = \Psi(a, a+t)\mathbf{f}(a+t) + 2 \int_0^t \mathbf{K}(a, a+s)\mathbf{M}_{t-s}\mathbf{f}(0)ds, \quad (21)$$

where the matrix  $\Psi(s, t)$  is given by

$$\Psi(s, t) = \begin{bmatrix} \psi_0(s, t) & \alpha\psi_0 \star \psi_1(s, t) \\ 0 & \psi_1(s, t) \end{bmatrix} \quad (22)$$

and the kernel  $\mathbf{K}$  is given by

$$\mathbf{K}(s, t) = \begin{bmatrix} (1-p)\beta_0(t)\psi_0(s, t) + \gamma\alpha\beta_1(t)\psi_0 \star \psi_1(s, t) & (1-\gamma)\alpha\beta_1(t)\psi_0 \star \psi_1(s, t) \\ \gamma\beta_1(t)\psi_1(s, t) & (1-\gamma)\beta_1(t)\psi_1(s, t) \end{bmatrix}. \quad (23)$$

Recall that  $\psi_0$  and  $\psi_1$  have been defined respectively in (2) and (3). The first column of the matrix kernel  $\mathbf{K}$  corresponds to the possible outcomes for individuals of type 0, which can persist with probability  $(1-p)$  or switch at rate  $\alpha$  and then give offspring of type 0 or 1 with probabilities  $\gamma$  and  $1-\gamma$  respectively. The second column corresponds to the individuals of type 1, whose contributions are pondered by  $\gamma$  or  $1-\gamma$  as recalled above.

**Remark 4.3.** In particular, by setting  $a = 0$ ,  $t \mapsto \mathbf{M}_t\mathbf{f}(0)$  is the unique fixed point of

$$\mathbf{M}_t\mathbf{f}(0) = \Psi(0, t)\mathbf{f}(t) + 2 \int_0^t \mathbf{K}(0, s)\mathbf{M}_{t-s}\mathbf{f}(0)ds, \quad (24)$$

and then we obtain  $\mathbf{M}_t\mathbf{f}(a)$  for all  $a \geq 0$  injecting this fixed point into the integral term of (21).

The proof of Proposition 4.2 is postponed to Section 8. Analogously, the left action of the semigroup can immediately be identified to the measure solution to the following PDE.

**Proposition 4.4** (Multitype renewal PDE). *For all initial vector measure  $\mu \in (\mathcal{M}(\mathbb{R}_+))^2$ , the vector measure  $\mu\mathbf{M}_t$  is equal to the measure-valued solution  $\mathbf{n}(t, \cdot)$  of the multitype McKendrick-von Foerster Equation*

$$\begin{cases} \partial_t \mathbf{n}(t, a) &= -\partial_a \mathbf{n}(t, a) - \mathbf{D}^\top(a)\mathbf{n}(t, a) \\ \mathbf{n}(t, 0) &= 2 \int_0^{+\infty} \mathbf{B}^\top(a)\mathbf{n}(t, da) \\ \mathbf{n}(0, \cdot) &= \mu \end{cases} \quad (25)$$

with  $\mathbf{B}^\top$  and  $\mathbf{D}^\top$  the transposed matrices of the ones defined by (20).

Moreover, if  $\mu$  are absolutely continuous with respect to the Lebesgue measure, then  $\mu\mathbf{M}_t$  is a strong solution to (25).

Then, solving (25) by variation of parameters we obtain that the vector measure  $\boldsymbol{\mu}\mathbf{M}_t$  admits, for any test function  $f \in \mathcal{B}_b(\mathbb{R}_+)$ , the representation

$$\langle \boldsymbol{\mu}\mathbf{M}_t, f \rangle = \int_0^t f(a)\boldsymbol{\Psi}^\top(0, a)\boldsymbol{\eta}(t-a)da + \int_t^{+\infty} f(a)\boldsymbol{\Psi}^\top(a-t, a)\boldsymbol{\mu}(a-t)da. \quad (26)$$

The function  $\boldsymbol{\Psi}(s, t)$  appears in this new context as the fundamental matrix solution to the ODE

$$\begin{cases} \partial_t \boldsymbol{\Psi}^\top(s, t) = -\mathbf{D}^\top(t)\boldsymbol{\Psi}^\top(s, t), & t > s \\ \boldsymbol{\Psi}(s, s) = \mathbf{I} \end{cases} \quad (27a)$$

$$(27b)$$

with  $\mathbf{I}$  the  $2 \times 2$  identity matrix, and  $\boldsymbol{\eta}$  is defined by the boundary condition of (25), giving

$$\boldsymbol{\eta}(t) = 2 \int_0^t \mathbf{B}^\top(a)\boldsymbol{\Psi}(0, a)\boldsymbol{\eta}(t-a)da + 2 \int_t^{+\infty} \mathbf{B}^\top(a)\boldsymbol{\Psi}(a-t, a)\boldsymbol{\mu}(a-t)da. \quad (28)$$

In particular, for an initial condition  $\boldsymbol{\mu} = (c_0\delta_0, c_1\delta_0)$  consisting on  $c_0$  initial individuals of type 0 with age 0, and  $c_1$  initial individuals of type 1 with age 0, we have that  $\boldsymbol{\eta}$  is solution to the linear Volterra equation of the second kind

$$\boldsymbol{\eta}(t) = g(t) + 2 \int_0^t \mathbf{K}^\top(0, a)\boldsymbol{\eta}(t-a)da. \quad (29)$$

since indeed  $\mathbf{B}^\top(a)\boldsymbol{\Psi}(0, a) = \mathbf{K}^\top(0, a)$  for  $\mathbf{K}$  introduced in (23). Meanwhile, the inhomogeneous term is given by

$$g(t) = 2\mathbf{B}^\top(t)\boldsymbol{\Psi}(0, t) \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = 2\mathbf{K}^\top(0, t)\mathbf{c} \quad (30)$$

We can interpret  $\boldsymbol{\eta}(t)$  as the instantaneous number of offspring produced at time  $t$ . The term  $g(t)$  gives the contribution of the initial individuals that have survived during  $[0, t]$  before dividing. The integral term counts the contribution at time  $t$  of the individuals of age  $a$  (born  $t-a$  ago).

We are interested in the long-time behaviour of the semigroup  $\mathbf{M}_t$  in the case when we have survival of the population, using some classical ideas from the spectral theory of  $C_0$ -semigroups, adapting the approach followed by Webb in [40] to age-structured population dynamics, and more recently applied in [30] to study the equilibrium of a birth-death model of ageing, also formulated as an individual-based stochastic model. To this end, we set ourselves on the Banach space  $(L^1(\mathbb{R}_+))^2$  equipped with the norm  $\|\mathbf{f}\|_1 = \int_0^{+\infty} (|f(a, 0)| + |f(a, 1)|)da$ . We also write  $\|\cdot\|_1$  for vectors and matrices, meaning, as usually:  $\|\mathbf{x}\|_1 = \sum_i |x_i|$ , and  $\|\mathbf{A}\|_1 = \max_j \sum_i |A_{ij}|$ . We then consider  $\mathbf{M}_t : (L^1(\mathbb{R}_+))^2 \rightarrow (L^1(\mathbb{R}_+))^2$ , which is the mild solution of (21) on  $(L^1(\mathbb{R}_+))^2$ . The existence of such semigroup is a direct consequence of the well-posedness of the measure-valued process  $Z_t$  and the control of its first moment as stated in Prop. A.2. We then obtain:

**Theorem 4.5.** *Under Assumptions 2.1 and if the survival conditions established by (16) are verified, there is a unique triplet of a positive function  $\mathbf{h} \in (L^1(\mathbb{R}_+))^2$ , a positive Radon measure  $\boldsymbol{\nu} \in (\mathcal{M}(\mathbb{R}_+))^2$ , normalised such that  $\langle \boldsymbol{\nu}, 1 \rangle = 1$  and  $\langle \boldsymbol{\nu}, \mathbf{h} \rangle = 1$ , and a positive constant  $\lambda > 0$  such that for all  $\mathbf{f} \in (L^1(\mathbb{R}_+))^2$*

$$\left\| e^{-\lambda t} \mathbf{M}_t \mathbf{f} - \langle \boldsymbol{\nu}, \mathbf{f} \rangle \mathbf{h} \right\|_1 \leq ce^{(\omega-\lambda)t} \|\mathbf{f} - \langle \boldsymbol{\nu}, \mathbf{f} \rangle \mathbf{h}\|_1. \quad (31)$$

The positive number  $\lambda$  is the population growth rate and is often called the Malthusian parameter or the population fitness; it is the largest real root of the characteristic equation

$$\det(\mathbf{F}(\lambda) - \mathbf{I}) = 0, \quad (32)$$

where

$$\mathbf{F}(\lambda) := 2 \int_0^{+\infty} e^{-\lambda a} \mathbf{K}(0, a) da.$$

Moreover, both coordinates of  $\nu$  admit a density with respect to the Lebesgue measure.

An important role is played by the matrix

$$\mathbf{K}_\infty := \mathbf{F}(0) = 2 \int_0^{+\infty} \mathbf{K}(0, a) da,$$

whose spectral properties determine the long-time behaviour of  $\mathbf{M}_t$ . Lemmas 4.6 and 4.7 will be useful to prove Theorem 4.5 in Section 8. Furthermore, they show how the conditions for survival with positive probability derived in Theorem 3.6 and the existence of a positive eigenvalue  $\lambda > 0$  are linked through the spectral properties of  $\mathbf{K}_\infty$ .

**Lemma 4.6.** *The survival condition (16) is equivalent to have to  $\rho(\mathbf{K}_\infty) > 1$*

**Lemma 4.7.** *Under Assumptions 2.1, there is a unique  $\lambda > 0$  such that  $\rho(\mathbf{F}(\lambda)) = 1$  (in particular,  $\lambda$  is solution to the characteristic equation (32)) if and only if the survival conditions established by (16) are verified.*

We conclude this section by noticing a useful bound for the value of  $\lambda$ , which is natural to obtain when the division rates are uniformly bounded.

**Remark 4.8.** We have  $\lambda \leq \bar{b}$ , with  $\bar{b}$  the bound on the division rate of (A1) of Assumptions 2.1.

## 5 Sensitivity of the population growth rate and the survival probability with respect to phenotypic switching strategies

In the following we denote by  $\mathcal{Q}_{\alpha, \gamma}$  the generator (19) and  $(\lambda_{\alpha, \gamma}, \nu_{\alpha, \gamma}, \mathbf{h}_{\alpha, \gamma})$  the triplet of elements verifying Theorem 4.5 for a given pair of parameters  $(\alpha, \gamma)$  in the survival region defined by Proposition 3.6. First, we show that the eigenfunction  $(\alpha, \gamma) \in \mathbb{R}_+ \times [0, 1] \mapsto \mathbf{h}_{\alpha, \gamma} \in (L^1(\mathbb{R}_+))^2$  is indeed continuous in  $\alpha$  and  $\gamma$ . This will allow us to study the variations of the population growth rate with respect to  $\alpha$  and  $\gamma$ .

The proofs of next lemmas and propositions are postponed to Section 9

**Lemma 5.1.** *Under Assumptions 2.1 and if  $\beta_0, \beta_1 \in C(\mathbb{R}_+)$ , then  $\mathbf{h} \in C^1(\mathbb{R}_+, \mathbb{R}_+^2)$ .*

**Lemma 5.2.** *Under Assumptions 2.1, for all fixed  $a \geq 0$ , the map  $(\alpha, \gamma) \mapsto \mathbf{h}_{\alpha, \gamma}(a)$  is continuous for the uniform norm.*

Proposition 5.3 characterises the partial variations of the population growth rate  $\lambda_{\alpha, \gamma}$  with respect to  $\alpha$  and  $\gamma$ .

**Proposition 5.3.** For fixed  $(\alpha, \gamma)$  and  $(\lambda_{\alpha, \gamma}, \nu_{\alpha, \gamma}, \mathbf{h}_{\alpha, \gamma})$  the triplet of eigenelements associated to  $\mathcal{Q}_{\alpha, \gamma}$  we have that both  $\alpha \mapsto \lambda_{\alpha, \gamma}$  and  $\gamma \mapsto \lambda_{\alpha, \gamma}$  are continuously differentiable functions such that

$$\partial_{\alpha} \lambda_{\alpha, \gamma} = \int_0^{+\infty} (h_{\alpha, \gamma}(a, 1) - h_{\alpha, \gamma}(a, 0)) \nu_{\alpha, \gamma}(da, 0), \quad (33)$$

$$\partial_{\gamma} \lambda_{\alpha, \gamma} = 2 (h_{\alpha, \gamma}(0, 0) - h_{\alpha, \gamma}(0, 1)) \int_0^{+\infty} \beta_1(a) \nu_{\alpha, \gamma}(da, 1). \quad (34)$$

Let us recall that the eigenfunction  $\mathbf{h}$  corresponds to Fisher's reproductive value [18]:  $h(a, i)$  is a measure of the contribution of an individual of age  $a$  and type  $i$  to the future growth of the population. Indeed, the long time behaviour of the expected total number of individuals issued from an individual of age  $a$  and type  $i$  is  $M_t \mathbf{1}(a, i)$  and by Theorem 4.5 is given by  $e^{\lambda t} h(a, i)$ . Thus, Equations (33) and (34) show that the value of the fitness response to variations in  $\alpha$  and  $\gamma$  depends on the difference in the reproductive values of type 1 and type 0.

In particular, in the case of the variations with respect to parameter  $\gamma$ , we see that the sign of  $\partial_{\gamma} \lambda_{\alpha, \gamma}$  depends only on the sign of the difference between the newborn's reproductive values of type 0 and type 1, which are given by the vector  $\mathbf{h}_{\alpha, \gamma}(0)$ . Moreover, Proposition 5.4 below shows that increasing the probability  $\gamma$  is beneficial for the growth of the population if and only if the population growth rate is already larger than the population growth rate associated with the subpopulation of type 1. Then, exploiting the equivalency of Lemma 4.6, we can link the variations of the population growth rate with the variations of the establishment probability. This will allow to show that increasing  $\gamma$  is detrimental from a Malthusian (population growth) point of view only if the death probability  $p$  is greater than some critical value  $\bar{p}$ , i.e. if the environmental stress is high enough. In contrast, increasing  $\gamma$  is always detrimental from the point of view of the establishment probability. This is also shown by the numerical simulations presented in Fig. 7. We will see later that it is not true in the more general case where the environment changes in time.

**Proposition 5.4.** For  $(\alpha, \gamma)$  in the survival region defined by Proposition 3.6, we have the following implicit equivalence

$$\begin{aligned} \partial_{\gamma} \lambda_{\alpha, \gamma} > 0 &\iff \lambda_{\alpha, \gamma} > \lambda_1^* \\ \partial_{\gamma} \lambda_{\alpha, \gamma} < 0 &\iff \lambda_{\alpha, \gamma} < \lambda_1^* \end{aligned}$$

where  $\lambda_1^*$  is the population growth rate of subpopulation 1 alone, i.e., the unique solution to

$$1 = 2 \int_0^{+\infty} e^{-\lambda_1^* a} \beta_1(a) \psi_1(0, a) da.$$

**Corollary 5.5.** If  $\lambda_{\alpha, \gamma=0} > \lambda_1^*$  then for all  $\gamma \in [0, 1[$ ,  $\partial_{\gamma} \lambda_{\alpha, \gamma} > 0$ .

*Proof.* Let us recall that we denote  $\xi_i(\lambda) = \mathbb{E}_{\delta_{(0, i)}} [e^{-\lambda T_{div}}]$  the Laplace transform associated to the division time of type  $i \in \{0, 1\}$ . In the case  $\gamma = 0$ , the matrix  $\mathbf{F}(\lambda)$  becomes triangular, and thus the characteristic equation  $\det(\mathbf{I} - \mathbf{F}(\lambda_{\alpha, \gamma})) = 0$  reduces to

$$(2(1-p)\xi_0(\alpha + \lambda) - 1)(2\xi_1(\lambda) - 1) = 0,$$

which admits as solutions  $\lambda_1^*$  and some  $\tilde{\lambda}_0 \in \mathbb{R}$  such that  $2(1-p)\xi_0(\alpha + \tilde{\lambda}_0) - 1 = 0$ . By Theorem 4.5, the fitness  $\lambda_{\alpha, \gamma=0}$  corresponds then to the maximum value between  $\lambda_1^*$  and  $\tilde{\lambda}_0$ . If  $\lambda_1^* < \tilde{\lambda}_0$  (or

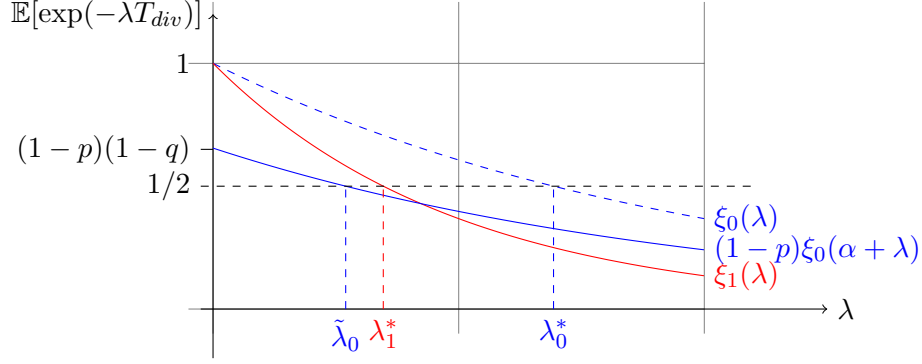


Figure 3: Laplace transforms of division times of type 0 ( $\xi_0$ , blue dashed line) and type 1 ( $\xi_1$ , red solid line), and the translation  $(1-p)\xi_0(\cdot + \alpha)$  (blue solid line). When  $\gamma = 0$ , the population growth rate is the largest value of  $\lambda$  at which one of the two latter functions (solid lines) pass through  $1/2$ .

equivalently,  $\lambda_{\alpha, \gamma=0} > \lambda_1^*$ ), by Proposition 5.4,  $\partial_\gamma \lambda_{\alpha, \gamma}|_{\gamma=0} > 0$ . And then, by the continuity of  $\gamma \mapsto \partial_\lambda \lambda_{\alpha, \gamma}$ , which follows easily from the continuity properties exhibited in the proof of Proposition 5.3,  $\partial_\gamma \lambda_{\alpha, \gamma} > 0$  for all  $\gamma \in [0, 1[$ .  $\square$

Notice however that if  $\lambda_1^* \geq \tilde{\lambda}_0$ , then  $\lambda_{\alpha, \gamma=0} = \lambda_1^*$  and thus by Proposition 5.4,  $\partial_\gamma \lambda_{\alpha, \gamma}|_{\gamma=0} = 0$ . Therefore, we cannot conclude about the sign of  $\partial_\gamma \lambda_{\alpha, \gamma}$  for  $\gamma \in (0, 1)$ . As shown by Fig. 3, this can happen if  $(1-p)(1-q)$  is small enough, for example, if  $p > 1/2$  or  $q > 1/2$  (but also in more general cases, as the one illustrated). Indeed, as proven above, when  $\gamma = 0$  the population growth rate  $\lambda$  is the largest solution of  $(2(1-p)\xi_0(\alpha + \lambda) - 1)(2\xi_1(\lambda) - 1) = 0$ , i.e., the largest value at which  $\lambda \mapsto (1-p)\xi_0(\alpha + \lambda)$  (solid blue curve in Fig. 3) or  $\lambda \mapsto \xi_1(\lambda)$  (solid red curve in Fig. 3) pass through  $1/2$ . Moreover, at  $\lambda = 0$ , the first function equals  $(1-p)\xi_0(\alpha) = \int_0^\infty \beta_0(a)\psi_0(0, a)da = (1-p)(1-q)$ , by the definition of  $q$  given in (12). In the case represented in Fig. 3,  $(1-p)(1-q)$  is small enough to impose  $\lambda_1^* > \tilde{\lambda}_0$ , so the largest solution to the characteristic equation (32) is the type 1 growth rate  $\lambda_1^*$ . In order to study  $\partial_\gamma \lambda_{\alpha, \gamma}$  in this more general case, we use the results derived for the establishment probability in Section 3 to obtain the following result.

**Proposition 5.6.** *Under the supplementary Assumption 2.4, for all  $\alpha \geq 0$  there exists a unique critical value  $\bar{p} \leq 1/2$  such that for all  $\gamma \in (0, 1)$*

$$\begin{aligned} \partial_\gamma \lambda_{\alpha, \gamma} > 0 &\iff p < \bar{p}, \\ \partial_\gamma \lambda_{\alpha, \gamma} < 0 &\iff p > \bar{p}. \end{aligned}$$

The proof are postponed to Section 9. Figure 4 summarises the results of Propositions 5.4 and 5.6. It illustrates the dependence of  $\lambda_{\alpha, \gamma}$  with respect to  $\gamma$ , for some fixed value of  $\alpha \in (0, 1)$ . The two basins delimited by  $\lambda_1^*$  identified in Proposition 5.4 are shown with arrows:  $\gamma \mapsto \lambda_{\alpha, \gamma}$  is increasing in the region  $\lambda > \lambda_1^*$ , and decreasing in the region in the region  $\lambda < \lambda_1^*$ . Whenever  $p < \bar{p}$  (blue curve),  $\lambda_{\alpha, \gamma=0} > \lambda_1^*$  and therefore  $\gamma \mapsto \lambda_{\alpha, \gamma}$  is increasing. This is the key idea of Proposition 5.6's proof.

Surprisingly, if we look at the establishment probability, we do not necessarily observe the same variations with respect to  $\gamma$ . Proposition 5.7 below shows that increasing  $\gamma$  will increase the extinction probability, for all  $p \in [0, 1]$  (and thus even for  $p < \bar{p}$ ):



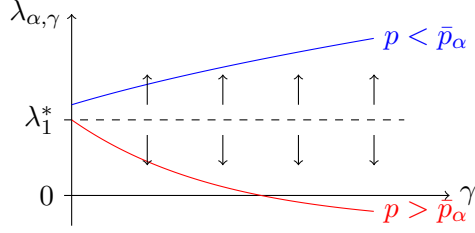


Figure 4: Illustration of Propositions 5.4 and 5.6.

**Proposition 5.7.** For all  $p \in [0, 1]$ ,  $\alpha > 0$ ,  $\gamma \in [0, 1[$ , let  $(\pi_0^{\alpha, \gamma}, \pi_1^{\alpha, \gamma}) \in [0, 1]^2$  the minimal solution to (13a)-(13b). Then

$$\partial_\gamma \pi_0^{\alpha, \gamma} > 0 \text{ and } \partial_\gamma \pi_1^{\alpha, \gamma} > 0.$$

*Proof.* Note that in the system (13a)-(13b) characterising the extinction probability, only (13b) depends on the value of  $\gamma$ . Moreover, for  $(\pi_0, \pi_1) \in (0, 1)^2$  verifying (13b) we have

$$\pi_1 = \frac{1 - 2\gamma(1 - \gamma)\pi_0 - \sqrt{1 - 4\gamma(1 - \gamma)\pi_0}}{2(1 - \gamma)^2},$$

and therefore, for all fixed value of  $\pi_0$  we have

$$\partial_\gamma \pi_1 = \frac{\sqrt{1 - 4\gamma(1 - \gamma)\pi_0} + (1 - \gamma)\pi_0 \left(1 + 2\gamma - \sqrt{1 - 4\gamma(1 - \gamma)\pi_0}\right) - 1}{(1 - \gamma)^3 \sqrt{1 - 4\gamma(1 - \gamma)\pi_0}},$$

which is always non-negative for  $\gamma \in [0, 1]$  and  $\pi_0 \in [0, 1]$ , since in that case  $\pi_0 \leq 1/(4\gamma(1 - \gamma))$ . Moreover, from (13a) we obtain that  $\pi_1$  is an increasing function of  $\pi_0$ . Thus, for  $(\pi_0, \pi_1)$  solution of the system, we have both  $\partial_\gamma \pi_0 > 0$  and  $\partial_\gamma \pi_1 > 0$ .  $\square$

In particular, only in the case  $p > \bar{p}$ , both the survival probability  $1 - \pi_i^{\alpha, \gamma}$  and the population growth rate  $\lambda_{\alpha, \gamma}$  decrease with respect to the recovery probability  $\gamma$ . This result is in contrast to [21], where the monotonic dependence observed in the survival probability translated to a monotonic dependence also in the population growth rate. As discussed by [21], this can be explained by a loss of monotonicity in the survival probabilities. In our case, albeit the survival probability varies monotonically with respect to  $\gamma$  (Proposition 5.7), the monotonicity is lost by the presence of two types. In Appendix A.4 we exhibit this loss of monotonicity using probabilistic arguments.

**Corollary 5.8.** For all  $\alpha > 0$  and  $\gamma \in (0, 1)$ , we have  $h_{\alpha, \gamma}(0, 1) > h_{\alpha, \gamma}(0, 0)$  if  $p > \bar{p}$ , and  $h_{\alpha, \gamma}(0, 1) \leq h_{\alpha, \gamma}(0, 0)$  otherwise.

*Proof.* It follows directly from Proposition 5.6 and the explicit form of  $\partial_\gamma \lambda_{\alpha, \gamma}$  obtained in Proposition 5.3.  $\square$

Corollary 5.8 shows that under high stress, the reproductive value  $h_{\alpha, \gamma}(0, 0)$  of newborn individuals of type 0, i.e., the contribution of type 0 individuals to the asymptotic size of the population, is always less important than the reproductive value of individuals of type 1. Indeed, even if they take a long time to divide and if the switching events are rare, the fact that type 1 individuals are

*perfectly adapted* to the stress environment (in the sense that they reproduce without dying), make that they end up contributing statistically more to the population size. It is not hard to imagine that this is not generally true when individuals of type 1 happen to be less adapted. For example, under a changing environment, the subpopulation of type 0 can be allowed to proliferate if the stress is reduced ( $p < \bar{p}$ ) at some intervals of time. This could be sufficient to make the contribution of type 0 larger in the asymptotic population. In the next section we explore an extension of the model where this happens. We will consider the case where the death probability  $p$  is allowed to vary periodically in time.

## 6 Sensitivity of the population growth rate and the survival probability under periodic stress

We now consider the case where the death probability  $p$  evolves periodically in time. Let  $T > 0$  a time period and  $t \in [0, T[ \mapsto p(t) \in [0, 1]$  a  $T$ -periodic function. Only this term is allowed to fluctuate in time;  $\alpha$  and  $\gamma$  are considered fixed traits of the population and remain constant. Since deaths occur at birth, only the birth matrix  $\mathbf{B}$  is affected and we define thereby

$$\mathbf{B}(t, a) := \begin{bmatrix} (1 - p(t))\beta_0(a) & 0 \\ \gamma\beta_1(a) & (1 - \gamma)\beta_1(a) \end{bmatrix}.$$

Analogously, we define the time-inhomogeneous generator

$$\mathcal{Q}(t)\mathbf{f}(a) := \mathbf{f}'(a) + 2\mathbf{B}(t, a)\mathbf{f}(0) - \mathbf{D}(a)\mathbf{f}(a), \quad (35)$$

whose adjoint operator is for every  $t \geq 0$

$$\mathcal{Q}^*(t)\mathbf{n}(a) := -\mathbf{n}'(a) - \mathbf{D}^\top(a)\mathbf{n}(a),$$

and has time-dependent domain

$$\mathcal{D}(\mathcal{Q}^*(t)) = \left\{ \mathbf{n} \in W^{1,\infty}(\mathbb{R}_+) : \mathbf{n}(0) = 2 \int_0^{+\infty} \mathbf{B}^\top(t, a)\mathbf{n}(a)da \right\}.$$

We consider as well the time-inhomogeneous associated matrices

$$\mathbf{F}(t, \lambda) := 2 \int_0^{+\infty} e^{-\lambda a} \Psi(0, a)\mathbf{B}(t, a)da.$$

Similarly to the use of Perron-Frobenius Theorem in the previous case, Floquet's Theorem (see for example [34], p.163) now allows us to construct the eigenelements that will drive the long-time behaviour of the periodic dynamics.

**Proposition 6.1.** *Let  $p : \mathbb{R}_+ \rightarrow [0, 1]$  a  $T$ -periodic continuous function. Under the same set of Assumptions 2.1 there exists a unique triplet of eigenelements  $(\lambda_T, \boldsymbol{\nu}, \mathbf{h})$  where  $\lambda_T \in \mathbb{R}$  and  $\boldsymbol{\nu} = \boldsymbol{\nu}(t, da)$  and  $\mathbf{h} = \mathbf{h}(t, da)$  are time-dependent  $T$ -periodic positive continuous functions such that*

$$\begin{cases} -\partial_t \mathbf{h}(t, a) = \mathcal{Q}(t)\mathbf{h}(t, a) - \lambda_T \mathbf{h}(t, a), & (36a) \\ \partial_t \boldsymbol{\nu}(t, a) = \mathcal{Q}^*(t)\boldsymbol{\nu}(t, a) - \lambda_T \boldsymbol{\nu}(t, a), \boldsymbol{\nu}(t, \cdot) \in \mathcal{D}(\mathcal{Q}^*(t)), & (36b) \\ \int_0^T \int_0^{+\infty} \boldsymbol{\nu}(t, a)dadt = 1, \int_0^T \int_0^{+\infty} \mathbf{h}(t, a)\boldsymbol{\nu}(t, a)dadt = 1. & (36c) \end{cases}$$

*Proof.* We adapt the general ideas developed on Section 5 of [32] and in Appendix A of [12], extending Section 4 to the case where  $t \mapsto p(t)$  is a continuous  $T$ -periodic function. If we wish to allow discontinuities (for example to account for piecewise constant treatments) we might follow the approach of [12], that requires to estimate supplementary controls.

By variation of parameters, if  $(\lambda_T, \boldsymbol{\nu}, \mathbf{h})$  are solutions to (36a)-(36c) then

$$\mathbf{h}(t, a) = 2 \int_a^{+\infty} e^{-\lambda_T(s-a)} \boldsymbol{\Psi}(a, s) \mathbf{B}(t+s-a, s) \mathbf{h}(t+s-a, 0) ds, \quad (37)$$

$$\boldsymbol{\nu}(t, a) = e^{-\lambda_T a} \boldsymbol{\Psi}^\top(0, a) \boldsymbol{\nu}(t-a, 0), \quad (38)$$

For the first equation, valuated the integral at  $a = 0$  we obtain that  $\mathbf{h}(t, 0)$  is solution to the integral fixed point problem

$$\mathbf{h}(t, 0) = \mathcal{F}_{\lambda_T}(t) \mathbf{h}(\cdot, 0) = 2 \int_0^{+\infty} e^{-\lambda_T a} \boldsymbol{\Psi}(0, a) \mathbf{B}(t+a, a) \mathbf{h}(t+a, 0) da. \quad (39)$$

Analogously, since  $\boldsymbol{\nu}$  must verify the boundary condition imposed by  $\boldsymbol{\nu}(t, \cdot) \in \mathcal{D}(\mathcal{Q}^*(t))$ , at  $a = 0$  we obtain that  $\mathbf{h}(t, 0)$  is solution to the integral fixed point problem

$$\boldsymbol{\nu}(t, 0) = \mathcal{G}_{\lambda_T}(t) \boldsymbol{\nu}(\cdot, 0) = 2 \int_0^{+\infty} e^{-\lambda_T a} \mathbf{B}^\top(t, a) \boldsymbol{\Psi}^\top(0, a) \boldsymbol{\nu}(t-a, 0) da. \quad (40)$$

The sequel is classical. By Arzela-Ascoli Theorem, one shows that under Assumptions 2.1, for all  $\lambda > 0$  and  $t \geq 0$  the operators  $\mathcal{F}_\lambda(t)$  and  $\mathcal{G}_\lambda(t)$  are continuous, strictly positive and compact. Thus, by Krein-Rutman Theorem there is a simple dominant eigenvalue  $\mu_\lambda > 0$  associated to eigenfunctions  $\boldsymbol{\nu}_\lambda(t, 0)$  and  $\mathbf{h}_\lambda(t, 0)$ . Then, by the maximum principle and analogously as done in the proof of Lemma 4.7, we obtain that  $\lambda \mapsto \mu_\lambda$  is a continuous and strictly decreasing map. Hence there is a unique  $\lambda_T$  such that  $\mu_{\lambda_T} = 1$  with associated eigenfunctions  $\boldsymbol{\nu}_{\lambda_T}(t, 0)$  and  $\mathbf{h}_{\lambda_T}(t, 0)$ . We can finally recover  $\boldsymbol{\nu}(t, a)$  and  $\mathbf{h}(t, a)$  using the reconstruction formulae (37) and (38).  $\square$

**Remark 6.2.** We recall that the Floquet dominant eigenvalue  $\lambda_T$  gives indeed the growth rate of the population. Set

$$M_{s,t} f(a, i) = \mathbb{E} [\langle Z_t, f \rangle | Z_s = \delta_{(a,i)}]$$

the time-inhomogeneous semigroup associated to  $\mathcal{Q}(t)$  and denote  $\mathbf{M}_{s,t}$  its vectorial form, as in the previous sections. Then, for  $\mathbf{h}(t, \cdot)$  solving (36a), we have immediately that

$$\mathbf{M}_{s,t} \mathbf{h}(s, a) = e^{\lambda_T(t-s)} \mathbf{h}(t, a).$$

Now, as in Section 5, we study the variations of  $\lambda_T$  with respect to the model parameters. Using the normalisation conditions (36c) and repeating the same calculations as in the proof of Lemma 5.1 and Proposition 5.3 we obtain the Proposition 6.3 below.

**Proposition 6.3.** *Let  $(\lambda_{T,\alpha,\gamma}, \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha,\gamma})$  the triplet of Floquet eigenelements associated to  $T$ -periodic  $\mathcal{Q}_{\alpha,\gamma}(t)$ . We have*

$$\partial_\alpha \lambda_{\alpha,\gamma} = \int_0^T \int_0^{+\infty} (h_{\alpha,\gamma}(t, a, 1) - h_{\alpha,\gamma}(t, a, 0)) \nu_{\alpha,\gamma}(t, a, 0) da dt, \quad (41)$$

$$\partial_\gamma \lambda_{\alpha,\gamma} = 2 \int_0^T (h_{\alpha,\gamma}(t, 0, 0) - h_{\alpha,\gamma}(t, 0, 1)) \left( \int_0^{+\infty} \beta_1(a) \nu_{\alpha,\gamma}(t, a, 1) da \right) dt. \quad (42)$$

*Proof.* It follows directly from (36c) and repeating the same calculations as in the proof of Proposition 5.3. The continuity of  $(\alpha, \gamma) \in \mathbb{R}_+ \times [0, 1] \mapsto \mathbf{h}_{\alpha, \gamma}(t, \cdot)$  for all fixed  $t \geq 0$ , follows also directly from Lemma 5.1.  $\square$

We compare this result with the one obtained in the constant environment case. We focus in the variations with respect to parameter  $\gamma$ . In the constant environment case we saw that if  $p > \bar{p}$  increasing  $\gamma$  is always detrimental from the point of view of the population growth rate (see Proposition 5.6, and the discussion of a numerical example in Section 10.1). This occurs since the reproductive value  $h_{\alpha, \gamma}(0, 1)$  of type 1 is always larger than the reproductive value  $h_{\alpha, \gamma}(0, 0)$  of type 0 when the level of stress is constant. However, we see now that in the fluctuating case the sign of  $\partial_\gamma \lambda_{\alpha, \gamma}$  depends on some time-average of the reproductive value difference. In particular, this difference is weighted proportionally to the mean division rate of type 1,  $\bar{\beta}_1(t) := \int_0^{+\infty} \beta_1(a) \nu_{\alpha, \gamma}(t, a, 1) da$ , observed at that time. Hence, if there are times  $t$  at which  $p(t) \leq 1/2$ , such that the reproductive value of type 0 is able to be larger than the reproductive value of type 1, it is possible for the average difference to be positive. In particular, the coincidence of these times with times at which the mean type 1 division rate  $\bar{\beta}_1(t)$  is large, can lead to positive values of  $\partial_\gamma \lambda_{\alpha, \gamma}$ . This comes off naturally from a heuristic reasoning. Indeed, big values of  $\bar{\beta}_1(t)$  and  $\gamma$  would lead to a burst in the creation of individuals of type 0 at time  $t$ , which possesses the biggest relative advantage at that time. We discuss a numerical illustration of this phenomenon in Section 10.3.

Finally, we can also compute the extinction probabilities. The time-dependent value of  $p$  breaks the Markovian property at the division stopping times, preventing the problem from being reduced to a simple algebraic system as in the constant environment case. The new system, however, can be solved numerically.

**Proposition 6.4.** *Let  $\pi_i(s, a) := \mathbb{P}(\exists t \geq s : N_t = 0 | Z_s = \delta_{(i, a)})$ . Then, we have that  $(\pi_0, \pi_1)$  is the minimal solution on  $[0, 1]^2$  of the system*

$$\begin{aligned} \pi_0(s, a) &= \int_0^{+\infty} p(s+t) \beta_0(a+t) \psi_0(a, a+t) dt + \int_0^{+\infty} \alpha \pi_1(s+t, a+t) \psi(a, a+t) dt \\ &\quad + \int_0^{+\infty} (\pi_0(s+t, 0))^2 (1-p(s+t)) \beta_0(a+t) \psi_0(a, a+t) dt; \\ \pi_1(s, a) &= \int_0^{+\infty} (\gamma \pi_0(s+t, 0) + (1-\gamma) \pi_1(s+t, 0))^2 \beta_1(a+t) \psi_1(a, a+t) dt. \end{aligned}$$

Moreover  $t \mapsto \pi_i(t, \cdot)$  is  $T$ -periodic.

*Proof.* It follows directly repeating the steps in Prop. 3.6  $\square$

## 7 Proofs of Section 3

### 7.1 Proof of Theorem 3.2

We compute the extinction probability in a general way that will be useful when time inhomogeneity is included in Sections 6 and 9, this is, when  $p$  is a periodic function instead of a constant. Note however that since for now  $p$  is fixed, we could have shown that the extinction probability of  $Z_t$  equals the extinction probability of some embedded discrete-time branching process giving the number of particles at the  $n$ -th generation, which is no other than a multitype Galton-Watson process [3].

*Proof.* Conditioning with respect to the possible outcomes of the first jump, we have

$$\begin{aligned}\pi(a, i) &= \mathbb{P}_{\delta_{(a,i)}}(Z_T = 0) + \mathbb{P}_{\delta_{(a,i)}}(\exists t > T : N_t = 0, Z_T \neq 0) \\ &= \mathbb{P}_{\delta_{(a,i)}}(Z_T = 0) + \mathbb{P}_{\delta_{(a,i)}}(\exists t > T : N_t = 0, Z_T = \delta_{a+T,1}) \\ &\quad + \mathbb{P}_{\delta_{(a,i)}}(\exists t > T : N_t = 0, Z_T = \delta_{(0,I_1)} + \delta_{(0,I_2)}).\end{aligned}$$

Applying the strong Markov property on the stopping time  $T$  gives

$$\begin{aligned}\pi(a, i) &= \mathbb{P}_{\delta_{(a,i)}}(Z_T = 0) + \mathbb{E}_{\delta_{(a,i)}} \left[ \mathbb{P}_{\delta_{(a+T,1)}}(\exists t > 0 : N_t = 0) \mathbf{1}_{Z_T = \delta_{a+T,1}} \right] \\ &\quad + \mathbb{E}_{\delta_{(a,i)}} \left[ \mathbb{P}_{\delta_{(0,I_1)} + \delta_{(0,I_2)}}(\exists t > 0 : N_t = 0) \mathbf{1}_{Z_T = \delta_{(0,I_1)} + \delta_{(0,I_2)}} \right],\end{aligned}$$

which by the independence of the processes starting from  $\delta_{(0,I_1)}$  and  $\delta_{(0,I_2)}$ , gives

$$\begin{aligned}\pi(a, i) &= \mathbb{P}_{\delta_{(a,i)}}(Z_T = 0) + \mathbb{E}_{\delta_{(a,i)}} \left[ \mathbb{P}_{\delta_{(a+T,1)}}(\exists t > 0 : N_t = 0) \mathbf{1}_{Z_T = \delta_{a+T,1}} \right] \\ &\quad + \mathbb{E}_{\delta_{(a,i)}} \left[ \mathbb{P}_{\delta_{(0,I_1)}}(\exists t > 0 : N_t = 0) \mathbb{P}_{\delta_{(0,I_2)}}(\exists t > 0 : N_t = 0) \mathbf{1}_{Z_T = \delta_{(0,I_1)} + \delta_{(0,I_2)}} \right] \\ &= \mathbb{P}_{\delta_{(a,i)}}(Z_T = 0) + \mathbb{E}_{\delta_{(a,i)}} \left[ \pi(a+T, 1) \mathbf{1}_{Z_T = \delta_{a+T,1}} \right] + \mathbb{E}_{\delta_{(a,i)}} \left[ \pi(0, I_1) \pi(0, I_2) \mathbf{1}_{Z_T = \delta_{(0,I_1)} + \delta_{(0,I_2)}} \right].\end{aligned}$$

Now, using Lemma 2.3, we obtain that

$$\begin{aligned}\pi(a, i) &= \mathbf{1}_{i=0} p \int_0^{+\infty} \beta_0(a+t) \exp\left(-\int_0^t \beta_0(a+u) du - \alpha t\right) dt \\ &\quad + \mathbf{1}_{i=0} \int_0^{+\infty} \pi(a+t, 1) \alpha \exp\left(-\int_0^t \beta_0(a+u) du - \alpha t\right) dt \\ &\quad + \int_0^{+\infty} \beta_i(a+t) \exp\left(-\int_0^t \beta_i(a+u) du - (1-i)\alpha t\right) dt \left\{ \mathbf{1}_{i=0}(1-p)\pi(0,0)^2 \right. \\ &\quad \left. + \mathbf{1}_{i=1}(\gamma\pi(0,0) + (1-\gamma)\pi(0,1))^2 \right\}.\end{aligned}\tag{43}$$

In particular, doing  $a = 0$ , we obtain

$$\begin{cases} \pi_0 = \pi(0,0) = pq' + (1-p)q'\pi_0^2 + \int_0^{+\infty} \pi(t,1) \alpha \exp\left(-\int_0^t \beta_0(u) du - \alpha t\right) dt, \\ \pi_1 = \pi(0,1) = (\gamma\pi_0 + (1-\gamma)\pi_1)^2, \end{cases}\tag{44}$$

$$\tag{45}$$

with

$$q' = \int_0^{+\infty} \beta_0(t) \exp\left(-\int_0^t \beta_0(u) du - \alpha t\right) dt = 1 - q,$$

for  $q$  defined by (12), since

$$q' + q = \int_0^{+\infty} (\alpha + \beta_0(t)) \exp\left(-\int_0^t \beta_0(u) du - \alpha t\right) dt = \mathbb{P}_{\delta_{(0,0)}}(T < +\infty) = 1.$$

Moreover, from (43) we have that for all  $t \geq 0$

$$\begin{aligned}\pi(t, 1) &= (\gamma\pi_0 + (1-\gamma)\pi_1)^2 \int_0^{+\infty} \beta_1(t+u) \exp\left(-\int_0^u \beta_1(t+w) dw\right) du \\ &= (\gamma\pi_0 + (1-\gamma)\pi_1)^2 \left(1 - \exp\left(-\int_0^{+\infty} \beta_1(t+u) du\right)\right).\end{aligned}$$

Hence using the integrability Assumptions 2.1 (A1), we obtain for all  $t \geq 0$

$$\pi(t, 1) = (\gamma\pi_0 + (1 - \gamma)\pi_1)^2 = \pi_1,$$

which gives immediately (15). Finally, after injecting these results back in (44) we get the system (13a)-(13b):

$$\begin{cases} \pi_0 = (1 - q)p + (1 - q)(1 - p)\pi_0^2 + q\pi_1 \\ \pi_1 = (\gamma\pi_0 + (1 - \gamma)\pi_1)^2. \end{cases}$$

Analogously, injecting these results back in (43) with  $i = 0$  we get (14) for all  $a \geq 0$ .

Next, we prove that  $(\pi_0, \pi_1)$  is the minimal solution of this system. Set  $(T_n)_{n \in \mathbb{N}}$  the jump times of  $Z$ , with  $T_0 = 0$ , so in our previous notation  $T_1 = T$ . Define the extinction probabilities at the  $n$ -th jump by

$$\pi^{(n)}(a, i) = \mathbb{P}_{\delta_{(a, i)}}(N_{T_n} = 0).$$

Therefore

$$\lim_{n \rightarrow +\infty} \pi^{(n)}(a, i) = \mathbb{P}_{\delta_{(a, i)}}\left(\bigcup_{n \in \mathbb{N}} \{N_{T_n} = 0\}\right) = \mathbb{P}_{\delta_{(a, i)}}(\exists t > 0 : N_t = 0) = \pi(a, i).$$

Now, suppose that we have some positive real solution  $\tilde{\pi}$  of (13a)-(13b). Then for both  $i \in \{0, 1\}$ ,  $\tilde{\pi}_i \geq \pi^{(0)}(0, i) = 0$ . We now show inductively that the same is verified for each  $n \in \mathbb{N}_*$  and in the limit  $n \rightarrow +\infty$ . As before, we condition with respect to the first jump and use the strong Markov property to obtain the following recursive equation

$$\begin{aligned} \pi^{(n)}(a, i) &= \mathbb{P}_{\delta_{(a, i)}}(Z_T = 0) + \mathbb{E}_{\delta_{(a, i)}}\left[\mathbb{P}_{\delta_{(a+T, 1)}}(N_{T_{n-1}} = 0) \mathbf{1}_{Z_T = \delta_{a+T, 1}}\right] \\ &\quad + \mathbb{E}_{\delta_{(a, i)}}\left[\mathbb{P}_{\delta_{(0, I_1) + \delta_{(0, I_2)}}}(N_{T_{n-1}} = 0) \mathbf{1}_{Z_T = \delta_{(0, I_1) + \delta_{(0, I_2)}}}\right] \\ &= \pi^{(1)}(a, i) + \mathbb{E}_{\delta_{(a, i)}}\left[\pi^{(n-1)}(a + T, 1) \mathbf{1}_{Z_T = \delta_{a+T, 1}}\right] \\ &\quad + \mathbb{E}_{\delta_{(a, i)}}\left[\pi^{(n-1)}(0, I_1) \pi^{(n-1)}(0, I_2) \mathbf{1}_{Z_T = \delta_{(0, I_1) + \delta_{(0, I_2)}}}\right], \quad n \in \mathbb{N}_* \end{aligned}$$

where we have again

$$\pi^{(1)}(a, i) = \mathbf{1}_{i=0} p \int_0^{+\infty} \beta_0(a + t) \exp\left(-\int_0^t \beta_0(a + u) du - ct\right) dt.$$

Suppose that  $\pi^{(n-1)} \leq \tilde{\pi}$ . Then

$$\begin{aligned} \pi^{(n)}(a, i) &\leq \pi^{(1)}(a, i) + \mathbb{E}_{\delta_{(a, i)}}\left[\tilde{\pi}(T, 1) \mathbf{1}_{Z_T = \delta_{T, 1}}\right] + \mathbb{E}_{\delta_{(a, i)}}\left[\tilde{\pi}(0, I_1) \tilde{\pi}(0, I_2) \mathbf{1}_{Z_T = \delta_{(0, I_1) + \delta_{(0, I_2)}}}\right] \\ &= \tilde{\pi}(a, i), \end{aligned}$$

since  $\tilde{\pi}$  is a solution of (43). Therefore, by induction, for all  $n \in \mathbb{N}$ ,  $\pi^{(n)} \leq \tilde{\pi}$ . Moreover, since  $\{\pi^{(n)}(a, i)\}_{n \in \mathbb{N}}$  is a sequence of probabilities of monotonic increasing events, we pass to the limit and conclude that  $\pi \leq \tilde{\pi}$ . Finally, notice that  $\pi_0 = \pi_1 = 1$  is always an admissible solution, therefore the extinction probabilities are well contained in  $[0, 1]$ .  $\square$

## 7.2 Proof of Theorem 3.6

*Proof.* First, note that (13a) and (13b) define two parabolic curves in the plane  $(\pi_0, \pi_1)$  which intersect at least at the point  $(1, 1)$ , since  $\pi_0 = \pi_1 = 1$  is always solution of the system. The proof consists in showing that other intersection occurs in the unit square  $[0, 1[ \times ]0, 1[$  if and only if Condition (16) is verified. As in the classical characterisation of the extinction probability in Galton-Watson branching processes, this property can be obtained as a consequence of the value of the derivatives of the curves at the intersection point  $(1, 1)$ .

Note that the parametric curve defined by (13a) is a concave parabola whose intercept is located at  $\pi_1 = -p \frac{1-q}{q} < 0$  and whose derivative is given by

$$\frac{d\pi_1}{d\pi_0}(\pi_0) = \frac{1 - 2(1-p)(1-q)\pi_0}{q}.$$

In particular, the derivative in  $(1, 1)$  equals

$$\frac{d\pi_1}{d\pi_0}(1) = \frac{1 - 2(1-p)(1-q)}{q}.$$

Note that the curve defined by (13a) admits two solutions at  $\pi_0 = 1$ . However, using the implicit function theorem around  $(1, 1)$ , we obtain a locally well defined function such that, by the implicit differentiation of (13b), it has derivative

$$\frac{d\pi_1}{d\pi_0}(\pi_0) = 2(\gamma\pi_0 + (1-\gamma)\pi_1(\pi_0)) \left( \gamma + (1-\gamma) \frac{d\pi_1}{d\pi_0}(\pi_0) \right),$$

and therefore at  $(1, 1)$  we have

$$\frac{d\pi_1}{d\pi_0}(1) = \frac{2\gamma}{2\gamma - 1}.$$

We can show that there is a second solution  $\bar{\pi}_1$  of (13b) at  $\pi_0 = 1$  comprised strictly between 0 and 1 if and only if  $0 < \gamma < 1/2$ . Moreover, (13b) also admits  $(0, 0)$  as solution. Thus, the trace of the curve described by (13b) connects  $(0, 0)$  to  $(1, 1)$  if  $\gamma \geq 1/2$ , or to  $(1, \bar{\pi}_1)$ , if  $\gamma < 1/2$ . Meanwhile, the curve of (13a) connects the negative ordinate  $(0, -p \frac{1-q}{q})$  with  $(1, 1)$ . Therefore, no intersection other than  $(1, 1)$  can occur inside the unit square if and only if  $\gamma \geq 1/2$ , and (13a) arrives at  $(1, 1)$  with non-negative derivative and whose value is at least as much as the value of the derivative of (13b). Otherwise, by the continuity and strict monotonicity of the curves we would have some other intersection point below  $(1, 1)$  (see Fig. 5). This is then:

$$\gamma \geq 1/2 \quad , \quad 1 - 2(1-p)(1-q) \geq 0 \quad \text{and} \quad \frac{1 - 2(1-p)(1-q)}{q} \geq \frac{2\gamma}{2\gamma - 1},$$

which gives finally, for  $p \neq 1/2$

$$\gamma \geq \frac{1}{2} \left( 1 + \frac{q}{(2p-1)(1-q)} \right).$$

In particular, the condition cannot be verified if  $p < 1/2$ . In the case  $p = 1/2$ , the previous conditions cannot be verified either. Thus finally, extinction occurs almost surely if and only if  $p > 1/2$  and condition (16) is verified, which gives the result.  $\square$

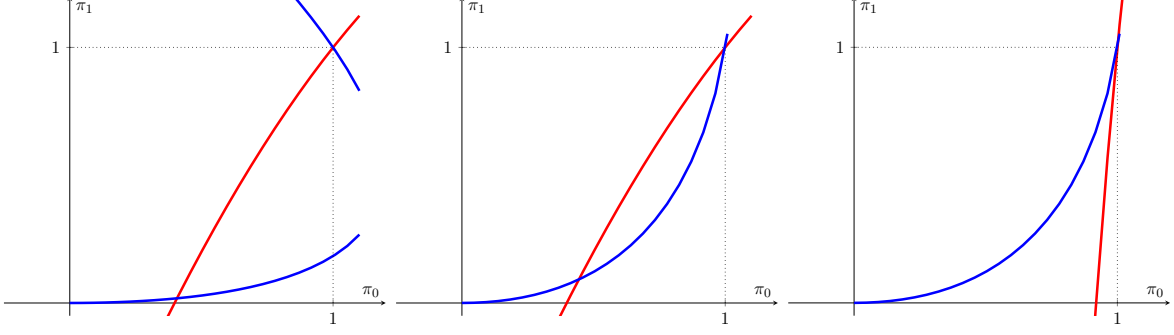


Figure 5: Parabolic curves defined by (13a) (red) and (13b) (blue) for  $p = 0.6$ . In the first case we have  $\gamma = 0.3 < 1/2$  and  $q = 0.4$ . In the second case we have  $\gamma = 0.6 > 1/2$  and  $q = 0.4$ . In the third case we have  $\gamma = 0.6$  and  $q = 0.02$ , so the condition (16) is violated and the only intersection in the unit square is  $(1,1)$ .

## 8 Proofs of Section 4

### 8.1 Proof of Proposition 4.2

*Proof.* Let  $(a, i) \in \mathbb{R}_+ \times \{0, 1\}$ ,  $f \in \mathcal{B}_b(\mathbb{R}_+ \times \{0, 1\})$ . Conditioning on the first jump event, we develop  $M_t f(a, i)$  using Lemma 2.3 to compute the expectations at each jump case, and the strong Markov property, similarly as we did in the Proof of Proposition 3.5. We obtain the following Duhamel's representation:

$$\begin{aligned}
M_t f(a, i) = & f(a + t, i) \psi_i(a, a + t) + (1 - i) \alpha \int_0^t M_{t-s} f(a + s, 1) \psi_i(a, a + s) ds \\
& + 2 \int_0^t \beta_i(a + s) \psi_i(a, a + s) \left\{ (i - 1)(1 - p) M_{t-s} f(0, 0) \right. \\
& \left. i (\gamma M_{t-s} f(0, 0) + (1 - \gamma) M_{t-s} f(0, 1)) \right\} ds,
\end{aligned} \tag{46}$$

where the first term of the RHS corresponds to the deterministic evolution when there are no events before time  $t$ , the second term corresponds to the case when the first jump is a type switch, and the third one to the case when the first jump is a division.

We iterate Duhamel's formula once more for the second line of the RHS, using (46) with  $i = 1$  and  $a = a + s$ , and then re-injecting the obtained result. We obtain therefore a representation that uses only the semigroup valued at initial age 0. Rearranging the terms for  $i = 0$  and  $i = 1$  in a vector, we obtain (21).  $\square$

### 8.2 Proof of Lemma 4.6

*Proof.* We start by the direct integration of (23) to obtain  $\mathbf{K}_\infty$ . The integration of the composition  $\psi_0 \star \psi_1$  requires some attention. First, we can easily remark that for all  $s < a$ ,  $\psi_1(s, a) =$



$\psi_1(0, a)/\psi_1(0, s)$ , and then using Fubini's theorem we obtain that

$$\begin{aligned} \int_0^{+\infty} \alpha\beta_1(a)\psi_0 \star \psi_1(0, a)da &= \int_0^{+\infty} \alpha\beta_1(a) \int_0^a \psi_0(0, s)\psi_1(s, a)dsda \\ &= \int_0^{+\infty} \alpha \frac{\psi_0(0, s)}{\psi_1(0, s)} \left( \int_s^{+\infty} \beta_1(a)\psi_1(0, a)da \right) ds. \end{aligned}$$

But, by definition of the survival function, we know that  $\psi_1(0, s) = \int_s^{+\infty} \beta_1(a)\psi_1(0, a)da$ . Thus, using the definition of  $q$  introduced by (12) we have that

$$\int_0^{+\infty} \alpha\beta_1(a)\psi_0 \star \psi_1(0, a)da = \int_0^{+\infty} \alpha\psi_0(0, s)ds = q.$$

Thus, we obtain finally

$$\mathbf{K}_\infty = 2 \begin{bmatrix} (1-p)(1-q) + \gamma q & (1-\gamma)q \\ \gamma & 1-\gamma \end{bmatrix},$$

which is a  $2 \times 2$  matrix of non-negative terms. In particular, it has non-negative discriminant and non-negative eigenvalues which are given by

$$\frac{1}{2} \left( \text{tr}(\mathbf{K}_\infty) \pm \sqrt{\text{tr}(\mathbf{K}_\infty)^2 - 4 \det(\mathbf{K}_\infty)} \right).$$

with  $\text{tr}(\mathbf{K}_\infty) = 2(1 + (1-p-\gamma)(1-q))$  and  $\det(\mathbf{K}_\infty) = 4(1-p)(1-q)(1-\gamma)$ . In particular, the largest eigenvalue is larger than 1 if and only if

$$\sqrt{\text{tr}(\mathbf{K}_\infty)^2 - 4 \det(\mathbf{K}_\infty)} > 2 - \text{tr}(\mathbf{K}_\infty) = (\gamma - (1-p))(1-q). \quad (47)$$

1) Case  $\gamma \leq 1-p$  : Since the RHS of (47) is non-positive, (47) is trivially verified and we have immediately  $\rho(\mathbf{K}_\infty) > 1$ .

2) Case  $\gamma > 1-p$  : Since the RHS of (47) is positive, by taking squares we have that (47) is equivalent to the following inequalities:

$$\begin{aligned} \text{tr}(\mathbf{K}_\infty)^2 - 4 \det(\mathbf{K}_\infty) &> 4 + \text{tr}(\mathbf{K}_\infty)^2 - 4\text{tr}(\mathbf{K}_\infty) \\ \iff \text{tr}(\mathbf{K}_\infty) - \det(\mathbf{K}_\infty) - 1 &> 0 \\ \iff 2 + 2(1-p-\gamma)(1-q) - 4(1-p)(1-q)(1-\gamma) - 1 &> 0 \\ \iff 2(1-q)((1-\gamma)(1-2(1-p)) - p) + 1 &> 0. \end{aligned}$$

Then, isolating the value of  $\gamma$  we obtain

$$(2p-1)\gamma < \frac{(2p-1)(1-q) + q}{2(1-q)}. \quad (48)$$

We study the case  $p > 1/2$  and  $p \leq 1/2$  separately. If  $p > 1/2$ , the factor  $2p-1$  is positive, and then dividing (48) by  $2p-1$  we obtain directly condition (16):

$$\gamma < \frac{1}{2} \left( 1 + \frac{q}{(2p-1)(1-q)} \right).$$

If  $p = 1/2$ , then (48) is trivially verified. Finally, if  $p < 1/2$ , the factor  $2p - 1$  is negative and dividing (48) by  $2p - 1$  we obtain

$$\gamma > \frac{1}{2} \left( 1 + \frac{q}{(2p-1)(1-q)} \right).$$

However, since we are under the assumption  $\gamma > 1 - p$  and  $p < 1/2$ , we have  $\gamma > 1/2$  and the inequality above is verified a fortiori.

Summarising,

$$\begin{aligned} \rho(\mathbf{K}_\infty) > 1 &\iff \{\gamma \leq 1 - p\} \cup (\{\gamma > 1 - p\} \cap \{p > 1/2\} \cap \{(16) \text{ is true}\}) \\ &\quad \cup (\{\gamma > 1 - p\} \cap \{p \leq 1/2\}) \\ &\iff \{p \leq 1/2\} \cup (\{p > 1/2\} \cap \{(16) \text{ is true}\}), \end{aligned}$$

which is exactly the condition ensuring survival with positive probability in Theorem 3.6.  $\square$

### 8.3 Proof of Lemma 4.7

*Proof.* Notice that for all  $\lambda \in \mathbb{R}$  the matrix  $\mathbf{F}(\lambda)$  is a  $2 \times 2$  matrix with all strictly positive entries. In particular, this ensures irreducibility and by Perron-Frobenius Theorem we have the existence, for all fixed  $\lambda \in \mathbb{R}$ , of a unique triplet of eigenelements  $(\mu(\lambda), \mathbf{n}(\lambda), \mathbf{h}(\lambda))$  such that

$$\begin{cases} \mathbf{F}(\lambda)\mathbf{h}(\lambda) = \mu(\lambda)\mathbf{h}(\lambda) \\ \mathbf{n}(\lambda)^\top \mathbf{F}(\lambda) = \mu(\lambda)\mathbf{n}(\lambda)^\top \\ \mathbf{n}(\lambda)^\top \mathbf{h}(\lambda) = 1 \\ \mathbf{n}(\lambda)^\top (1, 1) = 1. \end{cases}$$

We start by the sufficiency direction of the equivalence. To show that there is a unique  $\lambda > 0$  such that  $\rho(\mathbf{F}(\lambda)) = 1$  we will use a classical monotonicity argument. First, we notice that  $\mathbf{F}(0) = \mathbf{K}_\infty$ , which under the survival conditions and thanks to Lemma 4.6 has spectral radius  $\rho(\mathbf{F}(0)) = \mu(0) > 1$ . Second, notice that as  $\lambda \rightarrow +\infty$ ,  $e^{-\lambda s} \mathbf{K}(0, s)$  decreases coordinate by coordinate to the null matrix. Thus, we have by monotone convergence that  $\mathbf{F}(+\infty) = \mathbf{0}_{2 \times 2}$  and therefore  $\lim_{\lambda \rightarrow +\infty} \mu(\lambda) = 0$ . It remains to show that  $\lambda \mapsto \mu(\lambda)$  is a decreasing continuous function. This is a classical property that comes from Lemma A.5. The monotonicity is given by the assertion (i) of the latter, whilst the continuity comes from the estimation (ii) and the continuity of  $\lambda \mapsto e^{-\lambda s}$ . Finally, this implies that there exists a unique  $\lambda^* > 0$  such that  $\mu(\lambda^*) = 1$ .

The necessity direction of the equivalence follows easily from the previous remarks. Since  $\lambda \mapsto \mu(\lambda)$  is a continuous decreasing function, if  $\lambda^* \in \mathbb{R}$  such that  $\mu(\lambda^*) = 1$  is strictly positive, then  $\rho(\mathbf{F}(0)) = \mu(0) > 1$ . Finally, by Lemma 4.6, this is equivalent to verify the survival condition (16).  $\square$

### 8.4 Proof of Proposition 4.5

*Proof.* The proof follows classical arguments, presented for example in Section 4.3 of [40]. The first part of the proof consists in showing the existence of a spectral gap, this is, of a positive constant  $\omega$  such that for the value of  $\lambda$  defined by (32) we have

$$\max \left\{ -b, \sup_{z \in \sigma(\mathcal{Q}) \setminus \sigma_{ess}(\mathcal{Q}) \setminus \{\lambda\}} \operatorname{Re}(z) \right\} < \omega < \lambda.$$

To do so, we show in Appendix A.3 that the *growth bound*  $\omega_1(\mathbf{M}_t)$  associated with the measure of non-compactness of the semigroup  $\mathbf{M}_t$  (see Definition A.7) is bounded by  $-\underline{b}$ . Then, by Theorem 4.6 of [40] we have that the essential spectrum of  $\mathcal{Q}$  must be contained within  $\{z \in \mathbb{C} : \operatorname{Re}(z) \leq -\underline{b}\}$ . Therefore, for any  $\lambda \in \mathbb{C}$  such that  $\operatorname{Re}(\lambda) > -\underline{b}$  which is a root of the characteristic equation (32), we have that  $\lambda$  is an eigenvalue of  $\mathcal{Q}$ . This is, there exists a non-zero  $\mathbf{h} \in (L^1(\mathbb{R}_+))^2$  such that  $\mathcal{Q}\mathbf{h} = \lambda\mathbf{h}$ .

Indeed, suppose that the pair  $(\lambda, \mathbf{h})$  is solution to  $\mathcal{Q}\mathbf{h} = \lambda\mathbf{h}$ , then  $\mathbf{h}' \in (L^1(\mathbb{R}_+))^2$  and is given almost everywhere by

$$\mathbf{h}'(a) = (\lambda\mathbf{I} + \mathbf{D}(a))\mathbf{h}(a) - 2\mathbf{B}(a)\mathbf{h}(0).$$

Then, we notice that

$$\begin{aligned} \partial_a \left( e^{-\lambda a} \Psi(0, a) \mathbf{h}(a) \right) &= \partial_a \left( e^{-\lambda a} \Psi(0, a) \right) \mathbf{h}(a) + e^{-\lambda a} \Psi(0, a) \mathbf{h}'(a) \\ &= -e^{-\lambda a} \Psi(0, a) (\lambda\mathbf{I} + \mathbf{D}(a)) \mathbf{h}(a) + e^{-\lambda a} \Psi(0, a) (\lambda\mathbf{I} + \mathbf{D}(a)) \mathbf{h}(a) \\ &\quad - 2e^{-\lambda a} \Psi(0, a) \mathbf{B}(a) \mathbf{h}(0) \\ &= -2e^{-\lambda a} \mathbf{K}(0, a) \mathbf{h}(0). \end{aligned} \tag{49}$$

Therefore, since  $\mathbf{h} \in (L^1(\mathbb{R}_+))^2$  and  $\lim_{a \rightarrow +\infty} \Psi(0, a) = \mathbf{0}$ , we have

$$\mathbf{h}(0) = - \int_0^{+\infty} \partial_a \left( e^{-\lambda a} \Psi(0, a) \mathbf{h}(a) \right) da = 2 \int_0^{+\infty} e^{-\lambda a} \mathbf{K}(0, a) \mathbf{h}(0) da = \mathbf{F}(\lambda) \mathbf{h}(0).$$

This linear equation has a non-trivial solution  $\mathbf{h}(0)$  if and only if  $\det(\mathbf{F}(\lambda) - \mathbf{I}) = 0$ . Conversely if  $\lambda$  is the largest root of  $\det(\mathbf{F}(\lambda) - \mathbf{I}) = 0$  and is simple and associated with some eigenvector  $\mathbf{h}(0) \in \mathbb{R}_+^2$ , we have that

$$\mathbf{h}(a) = e^{\lambda a} \Psi^{-1}(0, a) \left( \mathbf{I} - 2 \int_0^a e^{-\lambda s} \Psi(0, s) \mathbf{B}(s) ds \right) \mathbf{h}(0) \tag{50}$$

$$= 2e^{\lambda a} \int_a^{+\infty} e^{-\lambda s} \Psi(a, s) \mathbf{B}(s) ds \mathbf{h}(0) \tag{51}$$

is a  $(L^1(\mathbb{R}_+))^2$  solution to  $\mathcal{Q}\mathbf{h} = \lambda\mathbf{h}$  (both representations will be useful in the sequel). Finally, Lemma 4.7 allows us to conclude for the simplicity of the eigenvalue  $\lambda$ .

Then, thanks to Proposition 4.65 of [40], we can identify  $\ker(\lambda\mathbf{I} - \mathcal{Q})$  to the image of the projection operator  $\mathbf{P} : (L^1(\mathbb{R}_+))^2 \rightarrow (L^1(\mathbb{R}_+))^2$  given by the resolvent

$$\mathbf{P}\mathbf{f}(a) := \frac{1}{2\pi i} \oint_{C_\lambda} (z\mathbf{I} - \mathcal{Q})^{-1} \mathbf{f}(a) dz,$$

where  $C_\lambda$  is a closed counterclockwise oriented curve of the complex plane enclosing  $\lambda$  but no other point of the spectrum of  $\mathcal{Q}$ .

Then, since  $\lambda$  is a simple root of (32), we have a unique pair  $(\lambda, \mathbf{h})$  (up to normalisation of  $\mathbf{h}$ ) such that  $\mathcal{Q}\mathbf{h} = \lambda\mathbf{h}$ , and so  $\mathbf{P}$  is of rank 1. Therefore, for all  $\mathbf{f}$  we can write  $\mathbf{P}\mathbf{f}(a) = \bar{\nu}(\mathbf{f})\mathbf{h}(a)$ , where  $\bar{\nu}(\mathbf{f})$  is a normalisation constant depending on  $\mathbf{f}$ .

Moreover, by the linearity of  $\mathbf{P}$ , we have that  $\boldsymbol{\nu} : (C_c(\mathbb{R}_+))^2 \rightarrow \mathbb{R}_+^2$  given for all  $\mathbf{f}$  of the form  $\mathbf{f} = (f(\cdot, 0), f(\cdot, 1))$  by

$$\boldsymbol{\nu}(\mathbf{f}) = (\bar{\nu}(f(\cdot, 0), 0), \bar{\nu}(0, f(\cdot, 1)))$$

is a linear application. Therefore, Riesz–Markov–Kakutani representation Theorem allows us to write  $\bar{\nu}(\mathbf{f}) = \int_0^{+\infty} \mathbf{f}(a)^\top \boldsymbol{\nu}(da)$  where  $\boldsymbol{\nu} \in (\mathcal{M}(\mathbb{R}_+))^2$  is a (vector) positive Radon measure. Furthermore, by duality we can conclude that this measure  $\boldsymbol{\nu}$  is the limiting distribution of the population ages in the sense that for all initial distribution  $\boldsymbol{\mu}$  and all  $\mathbf{f} \in (L^1(\mathbb{R}_+))^2$  we have

$$\begin{aligned} e^{-\lambda t} \langle \boldsymbol{\mu} \mathbf{M}_t, \mathbf{f} \rangle &= \int_0^{+\infty} \mathbf{f}(a) e^{-\lambda t} \boldsymbol{\mu} \mathbf{M}_t(da) = \int_0^{+\infty} e^{-\lambda t} \mathbf{M}_t \mathbf{f}(a) \boldsymbol{\mu}(da) \\ &\xrightarrow{t \rightarrow +\infty} \left( \int_0^{+\infty} \mathbf{h}(a) \boldsymbol{\mu}(da) \right) \left( \int_0^{+\infty} \mathbf{f}(a) \boldsymbol{\nu}(da) \right). \end{aligned}$$

Finally, by Proposition 4.15 of [40] we have

$$\|\mathbf{M}_t \mathbf{f} - \mathbf{P} \mathbf{f}\|_1 \leq c e^{\omega t} \|\mathbf{f} - \mathbf{P} \mathbf{f}\|_1,$$

from where we conclude the exponential rate of convergence.  $\square$

## 8.5 Proof of Remark 4.8

*Proof.* Notice that taking as test function  $\mathbf{f} \equiv \mathbf{1} = (1, 1)$ , we get

$$\begin{aligned} \mathbf{M}_t \mathbf{1}(a) &= \mathbf{1} + \int_0^t \mathbf{M}_s(\mathcal{Q}\mathbf{1})(a) ds = \mathbf{1} + \int_0^t \mathbf{M}_s((2\mathbf{B}(\cdot) - \mathbf{D}(\cdot))\mathbf{1})(a) ds \\ &= \mathbf{1} + \int_0^t \begin{bmatrix} \mathbf{M}_s(\beta_0, 0)(a)^\top \\ \mathbf{M}_s(0, \beta_1)(a)^\top \end{bmatrix} ds \begin{pmatrix} 1 - 2p \\ 1 \end{pmatrix} ds. \end{aligned}$$

Therefore for both coordinates  $i \in \{0, 1\}$  we have

$$M_t \mathbf{1}(a, i) \leq 1 + \bar{b} \int_0^t \mathbf{M}_s \mathbf{1}(a, i) ds.$$

Hence, by Grönwall's inequality, for all  $a \geq 0$ ,  $M_t \mathbf{1}(a, u) \leq e^{\bar{b}t}$ . Therefore, by duality, and since  $\boldsymbol{\nu} \mathbf{M}_t(da) = e^{\lambda t} \boldsymbol{\nu}(da)$ ,

$$e^{\lambda t} \langle \boldsymbol{\nu}, \mathbf{1} \rangle = \langle \boldsymbol{\nu} \mathbf{M}_t, \mathbf{1} \rangle = \langle \boldsymbol{\nu}, \mathbf{M}_t \mathbf{1} \rangle \leq e^{\bar{b}t} \langle \boldsymbol{\nu}, \mathbf{1} \rangle.$$

Finally,  $\langle \boldsymbol{\nu}, \mathbf{1} \rangle = 1 \neq 0$  and the result follows.  $\square$

## 9 Proofs of Section 5

### 9.1 Proof of Lemma 5.1

*Proof.* Since  $\mathbf{h}$  is solution to the eigenproblem  $\mathcal{Q}\mathbf{h} = \lambda \mathbf{h}$ , representation (51) gives us

$$\mathbf{h}(a) = 2 \int_0^{+\infty} e^{-\lambda s} \boldsymbol{\Psi}(a, a+s) \mathbf{B}(a+s) \mathbf{h}(0) ds.$$

We deduce the result by Dominated Convergence Theorem in each coordinate of the expression above. Set  $\ell(a, s) := 2e^{-\lambda s} \boldsymbol{\Psi}(a, a+s) \mathbf{B}(s) \mathbf{h}(0)$ . Note that only matrix  $\boldsymbol{\Psi}$  (defined in (22), and with

$\psi_0, \psi_1, \psi_0 \star \psi_1$  in (2)-(4) depends on  $a$ . First, we know by construction of the Perron eigenlements that for all  $a \geq 0$ ,  $\ell(a, \cdot) \in (L^1(\mathbb{R}_+))^2$ . Indeed:

$$\|\ell(a, s)\|_\infty \leq 2\bar{b} \|\mathbf{h}(0)\|_\infty e^{-\lambda s} \in L^1(\mathbb{R}_+, ds).$$

Moreover, since we assume additionally that  $\beta_0, \beta_1$  are continuous, we deduce that for all  $s \geq 0$ ,  $\ell(\cdot, s) \in C^1(\mathbb{R}_+, \mathbb{R}_+^2)$ . Indeed, for all  $a \geq 0, s \geq 0$ , we have explicitly

$$\begin{aligned} \partial_a \psi_0(a, a+s) &= (\beta_0(a) - \beta_0(a+s))\psi_0(a, a+s), \\ \partial_a \psi_1(a, a+s) &= (\beta_1(a) - \beta_1(a+s))\psi_1(a, a+s), \\ \partial_a(\psi_0 \star \psi_1)(a, a+s) &= (\alpha + \beta_0(a) - \beta_1(a+s))\psi_0 \star \psi_1(a, a+s) + \psi_0(a, a+s) - \psi_1(a, a+s), \end{aligned}$$

where all the terms at the RHS are well defined and continuous for all  $a \geq 0$ . Finally, all these terms are uniformly bounded by  $\bar{b} + \alpha + 1$ , and therefore

$$\|\partial_a \ell(a, s)\|_\infty \leq 4\bar{b}(\bar{b} + \alpha + 1) \|\mathbf{h}(0)\|_\infty e^{-\lambda s} \in L^1(\mathbb{R}_+, ds).$$

Hence, by the Dominated Convergence Theorem,  $\mathbf{h} \in C^1(\mathbb{R}_+, \mathbb{R}_+^2)$ .  $\square$

## 9.2 Proof of Lemma 5.2

*Proof.* Writing the dependencies on  $\alpha$  and  $\gamma$  explicitly, we express  $\mathbf{h}_{\alpha, \gamma}(a)$  for all  $a \geq 0$  as

$$\mathbf{h}_{\alpha, \gamma}(a) = 2 \int_0^{+\infty} e^{-\lambda_{\alpha, \gamma} s} \Psi_\alpha(a, a+s) \mathbf{B}_\gamma(a+s) \mathbf{h}_{\alpha, \gamma}(0) ds. \quad (52)$$

Again by Dominated Convergence Theorem, it suffices to prove the continuity and domination by an integrable function of both coordinates of

$$(\alpha, \gamma) \mapsto \ell_a(\alpha, \gamma) := 2e^{-\lambda_{\alpha, \gamma} s} \Psi_\alpha(a, a+s) \mathbf{B}_\gamma(a+s) \mathbf{h}_{\alpha, \gamma}(0) \in \mathbb{R}_+^2.$$

Hence, we show first that  $\|\mathbf{h}_{\alpha, \gamma}(0)\|_\infty$  is uniformly bounded in every open neighborhood of  $(\alpha, \gamma) \in \mathbb{R}_+ \times (0, 1)$ . Using the normalisation condition  $\langle \nu_{\alpha, \gamma}, \mathbf{h}_{\alpha, \gamma} \rangle = 1$  with  $\nu_{\alpha, \gamma}$  the left eigenmeasure, we have that for all  $a \geq 0$ ,

$$1 = \langle \nu_{\alpha, \gamma}, \mathbf{h}_{\alpha, \gamma} \rangle = \mathbf{u}_{\alpha, \gamma}^\top \mathbf{h}_{\alpha, \gamma}(0),$$

where

$$\mathbf{u}_{\alpha, \gamma} = 2 \int_0^{+\infty} \left( \int_0^{+\infty} e^{-\lambda_{\alpha, \gamma} \tau} \mathbf{B}_\gamma(a+\tau)^\top \Psi_\alpha(a, a+\tau)^\top d\tau \right) \nu_{\alpha, \gamma}(a) da.$$

Therefore, to bound each coordinate of  $\mathbf{h}_{\alpha, \gamma}(0) \in \mathbb{R}_+^2$  it suffices to bound by below both coordinates of  $\mathbf{u}_{\alpha, \gamma}$ . Set  $\mathbf{u}_{\alpha, \gamma} = (u_{\alpha, \gamma}^0, u_{\alpha, \gamma}^1)$ , such that

$$\begin{aligned} u_{\alpha, \gamma}^0 &= 2 \iint_{\mathbb{R}_+^2} e^{-\lambda_{\alpha, \gamma} \tau} \nu_{\alpha, \gamma}(a)^\top \begin{pmatrix} (1-p)\beta_0(a+\tau)\psi_0(a+\tau) + \gamma\beta_1(a+\tau)\alpha\psi_0 \star \psi_1(a, a+\tau) \\ \gamma\beta_1(a+\tau)\psi_1(a+\tau) \end{pmatrix} d\tau da, \\ u_{\alpha, \gamma}^1 &= 2 \iint_{\mathbb{R}_+^2} e^{-\lambda_{\alpha, \gamma} \tau} \nu_{\alpha, \gamma}(a)^\top \begin{pmatrix} (1-\gamma)\beta_1(a+\tau)\alpha\psi_0 \star \psi_1(a, a+\tau) \\ (1-\gamma)\beta_1(a+\tau)\psi_1(a+\tau) \end{pmatrix} d\tau da. \end{aligned}$$

By Remark 4.8, the value of  $\lambda_{\alpha,\gamma}$  is bounded by  $\bar{b}$  for all  $(\alpha, \gamma)$ . Then, using Assumptions 2.1 to bound the division rates in matrix  $\mathbf{B}$  and the survival functions in matrix  $\Psi$  we obtain :

$$\begin{aligned} u_{\alpha,\gamma}^0 &\geq 2\gamma \int_0^{+\infty} \left( \int_{a_0}^A \underline{b} e^{-2\bar{b}\tau} (1 - e^{-\alpha\tau}) d\tau \right) \boldsymbol{\nu}_{\alpha,\gamma}(a)^\top \mathbf{1} da, \\ u_{\alpha,\gamma}^1 &\geq 2(1 - \gamma) \int_0^{+\infty} \left( \int_{a_0}^A \underline{b} e^{-2\bar{b}\tau} (1 - e^{-\alpha\tau}) d\tau \right) \boldsymbol{\nu}_{\alpha,\gamma}(a)^\top \mathbf{1} da, \end{aligned}$$

for some fixed arbitrary quantity  $A > a_0$  with  $a_0$  being given by (A3) in Assumptions 2.1. Then, the integral with respect to  $\tau$  can be uniformly bounded by below by some positive constant  $\tilde{c}_{\alpha,A}$  dependent on  $\alpha$  and the choice of  $A$ . Finally, by the normalisation condition  $\langle \boldsymbol{\nu}, \mathbf{1} \rangle = 1$ , we obtain

$$u_{\alpha,\gamma}^0 \geq 2\gamma \tilde{c}_{\alpha,A}, \quad u_{\alpha,\gamma}^1 \geq 2(1 - \gamma) \tilde{c}_{\alpha,A},$$

and therefore,

$$h_{\alpha,\gamma}(0, 0) \leq \frac{1}{2\gamma \tilde{c}_{\alpha,A}}, \quad h_{\alpha,\gamma}(0, 1) \leq \frac{1}{2(1 - \gamma) \tilde{c}_{\alpha,A}}.$$

Thus, for all fixed  $a \geq 0$

$$\begin{aligned} \ell_{\alpha,\gamma}^0(a, s) &\leq 2\tilde{c}_{\alpha,A}^{-1} \left( \frac{1-p}{\gamma} \beta_0(a+s) \psi_0(a, a+s) + \beta_1(a+s) \alpha \psi_0 \star \psi_1(a, a+s) \right) \in L^1(\mathbb{R}_+, ds), \\ \ell_{\alpha,\gamma}^1(a, s) &\leq 2\tilde{c}_{\alpha,A}^{-1} \beta_1(a+s) \psi_1(a, a+s) \in L^1(\mathbb{R}_+, ds), \end{aligned}$$

which can be bound uniformly in any neighbourhood around  $(\alpha, \gamma)$ , whenever  $\alpha, \gamma \notin \{0, 1\}$ . Finally, we show that  $(\alpha, \gamma) \mapsto \ell_a(\alpha, \gamma)$  is continuous for the uniform norm. Fix some couple  $(\alpha, \gamma) \in \mathbb{R}_+ \times (0, 1)$  and let  $(\alpha_n, \gamma_n)_{n \in \mathbb{N}}$  be some sequence converging to  $(\alpha, \gamma)$  as  $n \rightarrow +\infty$ . We know already that  $\mathbf{h}_{\alpha_n, \gamma_n}(0)$  and  $\lambda_{\alpha_n, \gamma_n}$  are bounded for all  $n$ . Therefore we can extract some convergent subsequences with adherence values

$$\lambda_{\alpha_{n_k}, \gamma_{n_k}} \rightarrow \lambda_\infty \geq 0, \quad \mathbf{h}_{\alpha_{n_k}, \gamma_{n_k}}(0) \rightarrow \boldsymbol{\eta}_\infty \in \mathbb{R}_+^2.$$

Moreover, note that for all fixed vector  $\mathbf{x} \in \mathbb{R}_+^2$ , the linear application  $(\alpha, \gamma) \mapsto \mathbf{K}_{\alpha,\gamma}(a, a+s)\mathbf{x} = \Psi_\alpha(a, a+s)\mathbf{B}_\gamma(a+s)\mathbf{x}$  is a continuous function of  $(\alpha, \gamma)$  for the uniform norm. Therefore, we have that entry-wise

$$[\mathbf{K}_{\alpha_n, \gamma_n}(a, a+s)]_{i,j} \rightarrow [\mathbf{K}_{\alpha,\gamma}(a, a+s)]_{i,j}.$$

To conclude, we identify  $\lambda_\infty$  and  $\boldsymbol{\eta}_\infty$  to the Perron eigenlements of the associated limit matrix  $\mathbf{K}_{\alpha,\gamma}(a, a+s)$ . Indeed, let

$$\mathbf{h}_\infty(a) = e^{\lambda_\infty a} \Psi_\alpha^{-1}(0, a) \left( \mathbf{I} - 2 \int_0^a e^{-\lambda_\infty s} \Psi_\alpha(0, s) \mathbf{B}_\gamma(s) ds \right) \boldsymbol{\eta}_\infty.$$

Therefore  $\mathbf{h}_\infty(0) = \boldsymbol{\eta}_\infty$ , and differentiating  $\mathbf{h}_\infty(a)$  we obtain that

$$\mathbf{h}'_\infty(a) = \lambda_\infty \mathbf{h}_\infty(a) - 2\mathbf{B}_\gamma(a)\boldsymbol{\eta}_\infty + \mathbf{D}_\alpha(a)\mathbf{h}_\infty(a),$$

or equivalently

$$\mathcal{Q}_{\alpha,\gamma} \mathbf{h}_\infty(a) = \lambda_\infty \mathbf{h}_\infty(a).$$

By uniqueness of the solution to the eigenvalue problem associated to  $\mathcal{Q}_{\alpha,\gamma}$  (see previous sections) we conclude that

$$\mathbf{h}_\infty(a) = \mathbf{h}_{\alpha,\gamma}(a), \quad \lambda_\infty = \lambda_{\alpha,\gamma}.$$

This allow us to conclude that for all fixed values of  $a, s \geq 0$ , both coordinates of  $\ell_{\alpha_n, \gamma_n}(a, s)$  converge to  $\ell_{\alpha,\gamma}(a, s)$ . Thus, by Dominated Convergence Theorem, we can conclude the continuity of  $(\alpha, \gamma) \mapsto \int_0^{+\infty} \ell_{\alpha,\gamma}(a, s) ds = \mathbf{h}_{\alpha,\gamma}(a)$ .  $\square$

### 9.3 Proof of Proposition 5.3

*Proof.* First, since  $\mathbf{h}_{\alpha,\gamma}$  is normalised by  $\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha,\gamma} \rangle = 1$ , we can write

$$\lambda_{\alpha,\gamma} = \langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathcal{Q}_{\alpha,\gamma} \mathbf{h}_{\alpha,\gamma} \rangle = \langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathcal{Q}_{\alpha,\gamma} \mathbf{h}_{\alpha,\gamma} \rangle.$$

Then, for all  $\delta \in \mathbb{R}$  we have

$$\begin{aligned} \langle \boldsymbol{\nu}_{\alpha,\gamma}, (\mathcal{Q}_{\alpha+\delta,\gamma} - \mathcal{Q}_{\alpha,\gamma}) \mathbf{h}_{\alpha+\delta,\gamma} \rangle &= \langle \boldsymbol{\nu}_{\alpha,\gamma}, \lambda_{\alpha+\delta,\gamma} \mathbf{h}_{\alpha+\delta,\gamma} \rangle - \langle \lambda_{\alpha,\gamma} \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle \\ &= \langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle (\lambda_{\alpha+\delta,\gamma} - \lambda_{\alpha,\gamma}) \end{aligned}$$

and therefore

$$\begin{aligned} \lambda_{\alpha+\delta,\gamma} - \lambda_{\alpha,\gamma} &= \frac{\langle \boldsymbol{\nu}_{\alpha,\gamma}, (\mathcal{Q}_{\alpha+\delta,\gamma} - \mathcal{Q}_{\alpha,\gamma}) \mathbf{h}_{\alpha+\delta,\gamma} \rangle}{\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle} \\ &= \frac{1}{\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle} \int_0^{+\infty} \boldsymbol{\nu}_{\alpha,\gamma}(da)^\top \begin{bmatrix} -\delta & \delta \\ 0 & 0 \end{bmatrix} \mathbf{h}_{\alpha+\delta,\gamma}(a) \\ &= \frac{\delta}{\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle} \int_0^{+\infty} (h_{\alpha+\delta,\gamma}(a, 1) - h_{\alpha+\delta,\gamma}(a, 0)) \boldsymbol{\nu}_{\alpha,\gamma}(da, 0). \end{aligned}$$

Thanks to Lemma 5.1, and once more by Dominated Convergence Theorem, we have that  $\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle \rightarrow \langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha+\delta,\gamma} \rangle$  and  $\langle \boldsymbol{\nu}_{\alpha,\gamma}(\cdot, 0), h_{\alpha+\delta,\gamma}(\cdot, 1) - h_{\alpha+\delta,\gamma}(\cdot, 0) \rangle \rightarrow \langle \boldsymbol{\nu}_{\alpha,\gamma}(\cdot, 0), h_{\alpha,\gamma}(\cdot, 1) - h_{\alpha,\gamma}(\cdot, 0) \rangle$  as  $\delta \rightarrow 0$ . Then, dividing by  $\delta$  and making  $\delta \rightarrow 0$ , we obtain (33). Analogously for any  $\delta \in \mathbb{R}$  small enough we have

$$\begin{aligned} \lambda_{\alpha,\gamma+\delta} - \lambda_{\alpha,\gamma} &= \frac{\langle \boldsymbol{\nu}_{\alpha,\gamma}, (\mathcal{Q}_{\alpha,\gamma+\delta} - \mathcal{Q}_{\alpha,\gamma}) \mathbf{h}_{\alpha,\gamma+\delta} \rangle}{\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha,\gamma+\delta} \rangle} \\ &= \frac{1}{\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha,\gamma+\delta} \rangle} \int_0^{+\infty} \boldsymbol{\nu}_{\alpha,\gamma}(da)^\top 2 \begin{bmatrix} 0 & 0 \\ \delta \beta_1(a) & -\delta \beta_1(a) \end{bmatrix} \mathbf{h}_{\alpha,\gamma+\delta}(0) \\ &= \frac{\delta}{\langle \boldsymbol{\nu}_{\alpha,\gamma}, \mathbf{h}_{\alpha,\gamma+\delta} \rangle} (h_{\alpha,\gamma+\delta}(0, 0) - h_{\alpha,\gamma+\delta}(0, 1)) \int_0^{+\infty} \beta_1(a) \boldsymbol{\nu}_\alpha(da, 1), \end{aligned}$$

from where (34) is obtained by the same arguments.  $\square$

### 9.4 Proof of Proposition 5.4

*Proof.* From the proof of Theorem 4.5, if  $(\alpha, \gamma)$  is in the survival region, then  $\det(\mathbf{I} - \mathbf{F}(\lambda_{\alpha,\gamma})) = 0$ , and  $\mathbf{h}_{\alpha,\gamma}(0)$  is the unique non-trivial solution to the linear problem  $(\mathbf{I} - \mathbf{F}(\lambda_{\alpha,\gamma})) \mathbf{h}_{\alpha,\gamma}(0) = 0$ . From the expression of  $\mathbf{F}(\lambda)$  and the first equation, we have that for  $\lambda = \lambda_{\alpha,\gamma}$

$$\det \begin{bmatrix} 2(1-p)\xi_0(\alpha + \lambda) + 2\gamma\xi_{01}(\lambda) - 1 & 2(1-\gamma)\xi_{01}(\lambda) \\ 2\gamma\xi_1(\lambda) & 2(1-\gamma)\xi_1(\lambda) - 1 \end{bmatrix} = 0,$$

where  $\xi_0(\lambda) = \int_0^{+\infty} e^{-\lambda a} \beta_0(a) \exp(-\int_0^a \beta(s) ds) da$  and  $\xi_1(\lambda) = \int_0^{+\infty} e^{-\lambda a} \beta_1(a) \psi_1(0, a) da$  are the Laplace transforms associated to the division times of types 0 and 1, and  $\xi_{01}(\lambda) = \int_0^{+\infty} e^{-\lambda a} \psi_0 \star \psi_1(0, a) da$ . This implies the following implicit relation characterising  $\lambda_{\alpha, \gamma}$ :

$$\xi_{01}(\lambda_{\alpha, \gamma}) = \frac{(1 - 2(1-p)\xi_0(\alpha + \lambda_{\alpha, \gamma}))(1 - 2(1-\gamma)\xi_1(\lambda_{\alpha, \gamma}))}{2\gamma},$$

which allows to simplify the matrix  $\mathbf{F}(\lambda_{\alpha, \gamma})$  in order to obtain that

$$\mathbf{h}_{\alpha, \gamma}(0) \in \text{span} \left\{ \begin{pmatrix} 1 - 2\xi_1(\lambda_{\alpha, \gamma}) + 2\gamma\xi_1(\lambda_{\alpha, \gamma}) \\ 2\gamma\xi_1(\lambda_{\alpha, \gamma}) \end{pmatrix} \right\}.$$

Since  $\mathbf{h}_{\alpha, \gamma}(0)$  is a non-negative vector, we have finally that

$$\text{sign}(h_{\alpha, \gamma}(0, 0) - h_{\alpha, \gamma}(0, 1)) = \text{sign}(1 - 2\xi_1(\lambda_{\alpha, \gamma})).$$

Notice that  $\lambda \mapsto \xi_1(\lambda)$  is a continuous decreasing function, such that  $\xi_1(0) = 1$  and  $\xi_1 \rightarrow 0$  as  $\lambda \rightarrow +\infty$ . Moreover, the Malthusian parameter associated to  $\xi_1$ ,  $\lambda_1^*$ , is the unique solution to  $\xi_1(\lambda_1^*) = 1/2$  (this is classical, see for example [40]). Therefore, if  $\lambda_{\alpha, \gamma} > \lambda_1^*$ , then  $\xi_1(\lambda_{\alpha, \gamma}) < 1/2$  and by (34),  $\partial_\gamma \lambda_{\alpha, \gamma} > 0$ . Analogously, if  $\lambda_{\alpha, \gamma} < \lambda_1^*$ , then  $\partial_\gamma \lambda_{\alpha, \gamma} < 0$ .

□

## 9.5 Proof of Proposition 5.6

*Proof.* We study the sign of  $\partial_\gamma \lambda_{\alpha, \gamma}$  in the case where  $p$  is big or small enough, and then relate this two partial analyses by continuity. Assumption 2.4 plays a key role.

- (i) **Case  $p > 1/2$ .** From Proposition A.2 and the equivalence established in Lemmas 4.6 and 4.7, if  $p > 1/2$  we have that

$$\text{sign} \left( \frac{1}{2} \left( 1 + \frac{(2p-1)q}{1-q} \right) - \gamma \right) = \text{sign}(\rho(\mathbf{K}_\infty^{\alpha, \gamma}) - 1) = \text{sign}(\lambda_{\alpha, \gamma}).$$

In particular, for all  $\alpha \geq 0$  it exists

$$\bar{\gamma}_\alpha := \frac{1}{2} \left( 1 + \frac{(2p-1)q}{1-q} \right)$$

such that  $\lambda_{\alpha, \bar{\gamma}_\alpha} = 0$ . Recall from Corollary 5.5 and Fig. 3 that  $\lambda_{\alpha, 0} \geq \lambda_1^* > 0$ . Therefore, by the continuity of  $\gamma \mapsto \lambda_{\alpha, \gamma}$ , for some  $\hat{\gamma} > 0$  we must have  $\lambda_{\alpha, \hat{\gamma}} < \lambda_1^*$ , and by Proposition 5.4,  $\partial_\gamma \lambda_{\alpha, \gamma}|_{\gamma=\hat{\gamma}} < 0$ . This implies that for all  $\gamma \geq \hat{\gamma}$ ,  $\partial_\gamma \lambda_{\alpha, \gamma} < 0$ . and  $\partial_\gamma \lambda_{\alpha, \gamma} = 0$  for all  $\gamma < \hat{\gamma}$ . Thus, whenever  $p > 1/2$ , we have that for all  $\alpha, \gamma$ ,  $\partial_\gamma \lambda_{\alpha, \gamma} \leq 0$ .

- (ii) **Small  $p$  case.** Now we prove that for all  $\alpha$ , if  $p$  is small enough, then we have the opposite, namely: for all  $\gamma \geq 0$ ,  $\partial_\gamma \lambda_{\alpha, \gamma} \geq 0$ . Consider the limit case  $p = 0$  at  $\gamma = 1$ . The characteristic equation for  $\lambda_{\alpha, \gamma=1}$  becomes

$$\xi_0(\alpha + \lambda) + \xi_{01}(\lambda) = \frac{1}{2}.$$



Recall that  $T_{div}$  is the division time of a non-switching cell. By Jensen's inequality,

$$\xi_0(\alpha + \lambda) = \mathbb{E}_{\delta_{(0,0)}} [\exp(-(\lambda + \alpha)T_{div})] \geq \exp\left(-(\lambda + \alpha)\mathbb{E}_{\delta_{(0,0)}} [T_{div}]\right) = \xi_0(\lambda)\xi_0(\alpha).$$

In particular, for  $\lambda = \lambda_1^*$ , and since  $\xi_0(\lambda) > \xi_1(\lambda)$  by Remark 2.5,

$$\xi_0(\alpha + \lambda_1^*) \geq \xi_0(\alpha)\xi_0(\lambda_1^*) > \xi_0(\alpha)\xi_1(\lambda_1^*) = \frac{1-q}{2}.$$

On the other hand, letting  $\tau \sim \text{Exp}(\alpha)$  independent from  $T_{div}$  be the switching time of type 0, and recalling the definition of  $\psi_0 \star \psi_1$  from (4), by Fubini's theorem we can write

$$\begin{aligned} \xi_{01}(\lambda) &:= \int_0^{+\infty} e^{-\lambda a} \alpha \beta_1(a) \psi_0 \star \psi_1(0, a) da \\ &= \int_0^{+\infty} \left( \frac{1}{\psi_1(0, \tau)} \int_{\tau}^{+\infty} e^{-\lambda a} \beta_1(a) \psi_1(0, a) da \right) \alpha \psi_0(0, \tau) d\tau \\ &= q \mathbb{E}_{\delta_{(0,0)}} \left[ \mathbb{E}_{\delta_{(0,1)}} \left[ e^{-\lambda T_{div}} \mid T_{div} \geq \tau \right] \mid \tau < T_{div} \right]. \end{aligned}$$

By Assumption 2.4, consider a monotone coupling  $(\hat{T}_{div}^0, \hat{T}_{div}^1)$  on a common probability space such that for all  $i \in \{0, 1\}$ ,  $\hat{T}_{div}^i$  has the same distribution under  $\mathbb{P}$  as  $T_{div}$  under  $\mathbb{P}_{\delta_{(0,i)}}$ , and  $\mathbb{P}(\hat{T}_{div}^1 > \hat{T}_{div}^0) = 1$ . In particular  $\{\tau \leq \hat{T}_{div}^0\} \subseteq \{\tau \leq \hat{T}_{div}^1\}$ . Therefore

$$\xi_{01}(\lambda) = q \mathbb{E} \left[ \mathbb{E} \left[ e^{-\lambda \hat{T}_{div}^1} \mid \hat{T}_{div}^1 \geq \tau \right] \mid \tau < \hat{T}_{div}^0 \right] = q \mathbb{E} \left[ \mathbb{E} \left[ e^{-\lambda \hat{T}_{div}^1} \right] \mid \tau < \hat{T}_{div}^0 \right] = q \xi_1(\lambda).$$

In particular, at  $\lambda = \lambda_1^*$ ,  $\xi_{01}(\lambda_1^*) = q/2$ . Finally,

$$\xi_0(\alpha + \lambda_1^*) + \xi_{01}(\lambda_1^*) > \frac{1}{2}$$

and hence  $\lambda_{\alpha, \gamma=1} > \lambda_1^*$ . Then, analogously to Corollary 5.5, we have that  $\partial_{\gamma}^- \lambda_{\alpha, \gamma} \big|_{\gamma=1} > 0$ , where  $\partial_{\gamma}^-$  is the left derivative with respect to  $\gamma$  (whose values are restricted to  $[0, 1]$ ), and by Proposition 5.4, that for all  $\gamma \geq 0$ ,  $\partial_{\gamma} \lambda_{\alpha, \gamma} \geq 0$ .

- (iii) **Conclusion.** The continuity of  $p \mapsto \partial_{\gamma} \lambda_{\alpha, \gamma}$  can be exhibited following the same arguments presented in the proof of Lemma 5.2. Moreover, as we did in Proposition 5.4 we can further show that  $p \mapsto \partial_{\gamma} \lambda_{\alpha, \gamma}$  is continuously differentiable and that for all  $\alpha, \gamma, p \in (0, 1)$ ,

$$\partial_p \lambda_{\alpha, \gamma} = \langle \nu, (\partial_p \mathcal{Q}) \mathbf{h} \rangle = -2h(0, 0) \int \beta_0(a) \nu(a, 0) da < 0.$$

Hence, for all  $\alpha, \gamma$ , the function  $p \mapsto \partial_{\gamma} \lambda_{\alpha, \gamma}$  is continuous and strictly decreasing. Therefore, for all  $\alpha \geq 0$ , there is a unique  $\bar{p}_{\alpha, \gamma}$  such that  $\lambda_{\alpha, \gamma} = \lambda_1^*$  and  $\partial_{\gamma} \lambda_{\alpha, \gamma} = 0$  for all  $\gamma \in (0, 1)$ . In particular, this  $\bar{p}_{\alpha, \gamma}$  is then constant in  $\gamma$ , and uniquely determined by  $\alpha$ , which gives the result. The conclusion is presented in Figure 4.

□

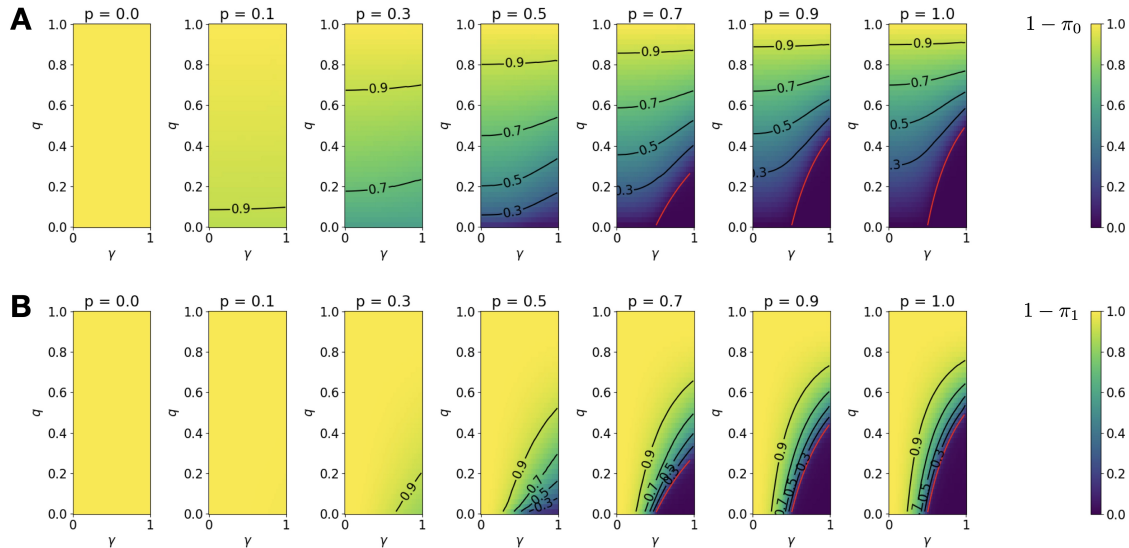


Figure 6: Value of the survival probability  $1 - \pi_i$  for a population issued from a single cell of type  $i = 0$  (A) and  $i = 1$  (B). The values were numerically computed as the minimal solution to the system (13a)-(13b). When  $p > 1/2$ , the red line represents the critical case of equality in (16).

## 10 Biological discussion and outlook

### 10.1 Under constant stress, high recovery might lead to increased extinction

Fig. 6 shows numerical solutions of the system (13a)-(13b), giving the values of the survival probabilities  $1 - \pi_i$  for a population initiated by a single vulnerable cell ( $i = 0$ , in row A) and by a single tolerant cell ( $i = 1$ , in row B). We can observe in row A that as expected, the survival of the population arising from a type 0 cell, increases as the probability of switching before dividing,  $q$ , increases. This is expected since we suppose that type 1 is not susceptible to death. On the other hand, as shown by Proposition 3.6, increasing the recovery probability  $\gamma$  introduces a significant risk of extinction, particularly when the antibiotic-induced death probability  $p$  is large. While this might appear biologically counter-intuitive, increasing the number of recovered individuals of type 0 introduces individuals who are prone to death and might lead, with positive probability, to the extinction of the whole population. This is also true, although for a higher level of stress, if we start with a type 1 individual as recovery can lead to the production of type 0 individuals who are susceptible to death (Figure 6B). When recovery ( $\gamma$ ) is high, having a higher switching rate is of interest for the survival of the population, in order to escape from the risk region. Thus, under fixed environmental stress, the optimal strategy from the point of view of the survival probability of the population is to switch as fast as possible.

### 10.2 Under constant low stress, high recovery leads to opposite effects on the survival probability and the population growth rate

We have shown in Propositions 5.6 and 5.7 that measuring the response of the extinction probability of a population issued from a single cell and of the population growth rate with respect to

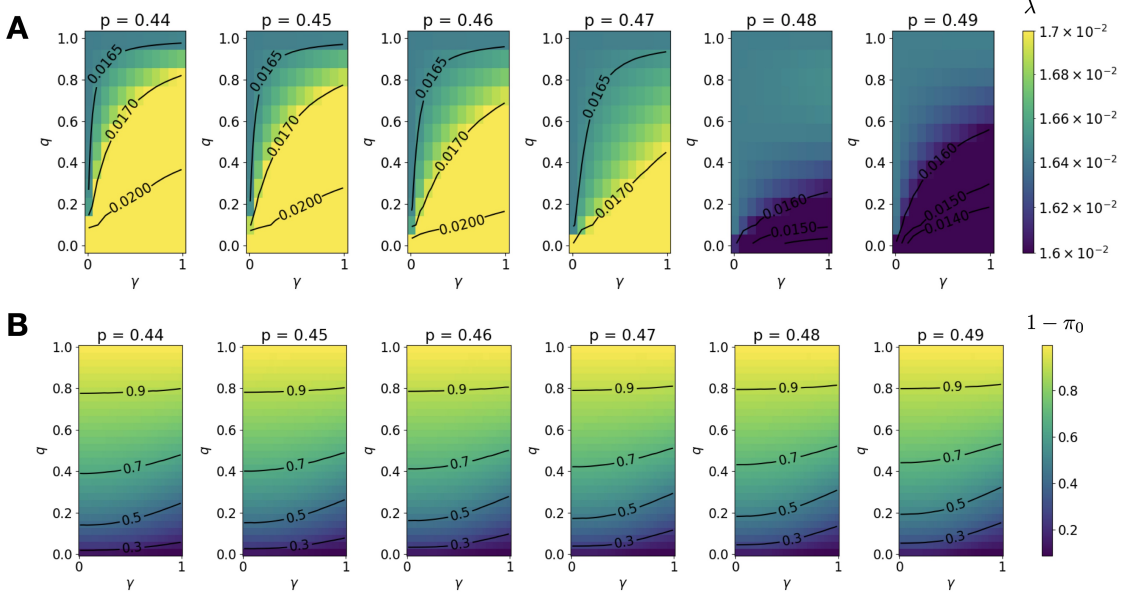


Figure 7: Variation of the population growth rate  $\lambda$  (row **A**) and of the survival probability  $1 - \pi_0$  starting from an initial cell of type 0 (row **B**), for some values of  $p < 1/2$ . The division rates  $\beta_0$  and  $\beta_1$  correspond to Gamma distributed division ages, as in Example 3.5. For type 0, the parameters are  $a_0 = 3, b_0 = 1$  (mean division time equals to 3), and for type 1, the parameters are  $a_0 = 3, b_0 = 0.1$  (mean division time equals to 30).

variations in the recovery probability  $\gamma$  can lead to seemingly opposite conclusions. On the one hand, as discussed in the paragraph above, increasing  $\gamma$  will always decrease the population survival probability. On the other hand, increasing  $\gamma$  can increase the population growth rate ( $\lambda$ ) if  $p$  is low enough. In that context, creating vulnerable but fast-proliferating type 0 individuals becomes advantageous.

Fig. 7 provides a quantitative representation of this behaviour. Row A shows numerical approximations of  $\lambda$  obtained from the numerical solution to PDE (25). We observe that for the chosen  $\beta_0, \beta_1$  (see caption), and for a stress weak enough such that  $p \leq 0.47$ ,  $\lambda$  increases with  $\gamma$  for any  $q$ . However, between  $p = 0.47$  and  $p = 0.48$ , the direction of the fitness gradient changes, such as increasing  $\gamma$  leads to a lower growth rate for the population. On the contrary, on row B we observe, as we already did in Fig. 6, that for all values of  $p$  (and  $q$ ), the survival probability decreases, though ever so slightly, with  $\gamma$ .

This surprising result raises some methodological questions about comparing the growth response using single-cell and population techniques. Indeed, parameter variations might not have the same effects on the population growth rate in large population experiments and on the viability curves of small colonies. In other words, the optimal parameters that characterise the single lineages that survive by the end of the experiment and the parameters that maximise the exponential growth of the population might not be the same.

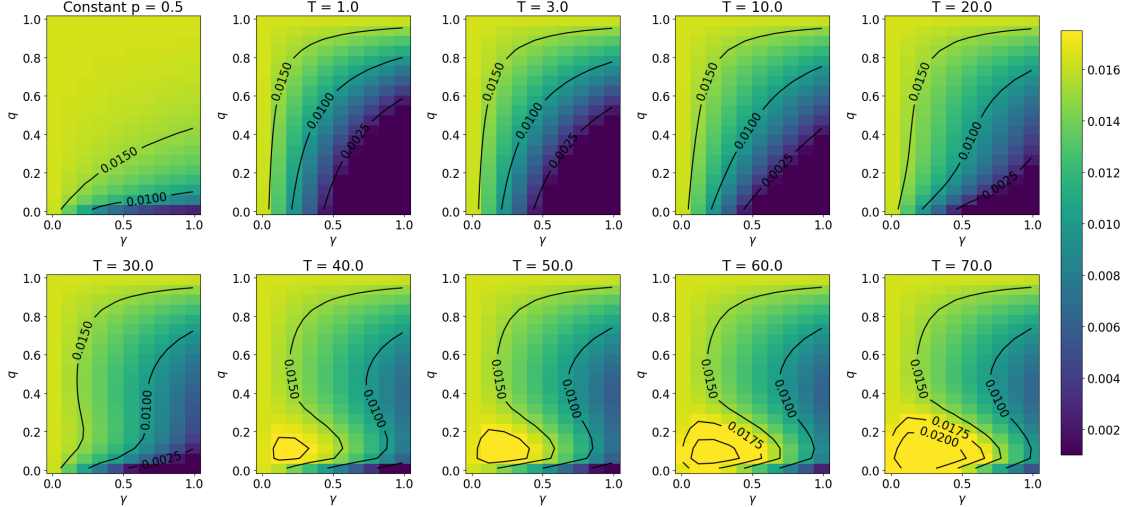


Figure 8: Value of the Floquet eigenvalue  $\lambda_T$  for increasing values of  $T$  as indicated in the head of the panel, as function of  $\gamma$  and  $q$ , obtained from simulations of (25). The death probability  $p$  varies between 0.25 and 0.75 spending a time  $T/2$  in each. Division times are Gamma distributed as in Fig. 7 (mean division time of type 0 is 3, and of type 1 is 30).

### 10.3 Under periodic stress, fine-tuning the switching rate and the recovery probability with respect to the period length can lead to higher growth

Proposition 6.3 shows that, under periodic stress, depending on the form of the eigenfunctions  $h_{\alpha,\gamma}(t, \cdot)$ , the gradient of the population growth rate  $\lambda_{\alpha,\gamma}$  admits non-trivial changes of sign. This implies that the fitness landscape changes when the environmental stress is allowed to fluctuate. In particular, it is possible that phenotypic plasticity strategies that would have led almost surely to extinction in a constant environment become advantageous in the periodic case.

As shown for illustrative purposes in Fig. 8, simulations of PDE (25) with a  $T$ -periodic stress (i.e.  $T$ -periodic  $p$ ) illustrate this result for different values of  $T$ . We use the same parameters as in the constant stress case showed in Fig. 7, but with a  $T$ -periodic stress signal  $p$  switching periodically between “good” and “bad” times, such that  $p(t) = 0.25$  for  $0 < t \leq T/2$  and  $p(t) = 0.75$  for  $T/2 < t \leq T$ . We compare the simulations with the case where  $p$  is fixed at its average value of  $p \equiv 0.5$  (top left panel).

We observe that under different values of  $T$ , the best stress response strategy changes. First, when the period  $T$  is short ( $T \leq 30$  in the illustrative case shown in Fig. 8), a non-growing region emerges in the fitness landscape, for a strategy combining low switching and high recovery. This suggests that this strategy, which leads to a high number of type 0 cells, is detrimental. As type 0 cells are prone to death under stress, and the “good” periods are too short to let them thrive, the overall population stops growing.

However, as the value of  $T$  increases, the fitness landscape changes appreciably. Surprisingly, we see the emergence of a “Goldilocks zone” where populations that recover and switch slowly (but still recover and switch, with  $\gamma > 0$  and  $q > 0$ ) grow even faster than in the constant case. At the same time, the previously well-adapted strategy of likely recovery ( $\gamma > 1/2$  verifying (16)) and fast switch ( $q > 1/2$ ) becomes less adapted as the period increases. In other words, under periodic stress,

bacteria can allow themselves to switch less (i.e. to reduce the induction of the stress response) as long as they do not repair very fast. This seems to lead to an optimal coexistence of type 0 and type 1, where the small switching rate allows the population to take advantage of the good times, while the slow repair allows them to keep a reserve of type 1 individuals that maintain the population during bad times. Overall, our results suggest that the population growth rate and the survival probability of a population issued from a single individual need to be analysed separately and can sometimes show opposite behaviour. Moreover, under the more realistic configuration of periodic stress, optimal population growth can only be achieved through fine-tuning both the repair and the switching rate simultaneously.

## Acknowledgments

This work has been supported by the Chair Modélisation Mathématique et Biodiversité of Veolia Environnement - École polytechnique - Muséum National d'Histoire Naturelle - Fondation X. This work has been supported in part by a Wellcome Trust Investigator Awards (Grant No. 205008/Z/16/Z) awarded to M.E.K. It is also funded by the European Union (ERC, SINGER, 101054787). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

## References

- [1] M. Acar, J. T. Mettetal, and A. Van Oudenaarden. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics*, 40(4):471–475, apr 2008.
- [2] R. N. Alnahhas and M. J. Dunlop. Advances in linking single-cell bacterial stress response to population-level survival. *Current Opinion in Biotechnology*, 79:102885, feb 2023.
- [3] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1972.
- [4] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, sep 2004.
- [5] G. Bell. Evolutionary Rescue. *Annual Review of Ecology, Evolution, and Systematics*, 48:605–627, nov 2017.
- [6] J. Blath, F. Hermann, and M. Slowik. A branching process model for dormancy and seed banks in randomly fluctuating environments. *Journal of Mathematical Biology*, 83(2):1–40, aug 2021.
- [7] J. Blath and A. Tóbiás. Invasion and fixation of microbial dormancy traits under competitive pressure. *Stochastic Processes and their Applications*, 130(12):7363–7395, dec 2020.
- [8] V. Calvez, M. Doumic, and P. Gabriel. Self-similarity in a general aggregation-fragmentation problem. Application to fitness analysis. *Journal des Mathématiques Pures et Appliquées*, 98(1):1–27, jul 2012.

- [9] F. Campillo, N. Champagnat, and C. Fritsch. Links between deterministic and stochastic approaches for invasion in growth-fragmentation-death models. *Journal of Mathematical Biology*, 73(6-7):1781–1821, dec 2016.
- [10] F. Campillo, N. Champagnat, and C. Fritsch. On the variations of the principal eigenvalue with respect to a parameter in growth-fragmentation models. *Communications in Mathematical Sciences*, 15(7):1801–1819, 2017.
- [11] O. Carja and J. B. Plotkin. Evolutionary rescue through partly heritable phenotypic variability. *Genetics*, 211(3):977–988, mar 2019.
- [12] J. Clairambault, S. Gaubert, and T. Lepoutre. Comparison of Perron and Floquet Eigenvalues in Age Structured Cell Division Cycle Models. *Mathematical Modelling of Natural Phenomena*, 4(3):183–209, 2009.
- [13] J. Clairambault, S. Gaubert, and B. Perthame. An inequality for the Perron and Floquet eigenvalues of monotone differential systems and age structured equations. *Comptes Rendus Mathématique*, 345(10):549–554, nov 2007.
- [14] B. Cloez, B. de Saporta, and T. Roget. Long-time behavior and Darwinian optimality for an asymmetric size-structured branching process. *Journal of Mathematical Biology*, 83(6-7):1–30, dec 2021.
- [15] U. Dieckmann and R. Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J Math Biol*, 34(5):579–612, 1996.
- [16] M. Doumic, M. Hoffmann, N. Krell, and L. Robert. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli*, 21(3):1760–1799, aug 2015.
- [17] M. Durinx and J. A. Metz. Multi-type Branching Processes and Adaptive Dynamics of Structured Populations. In *Branching Processes*, chapter 7, pages 266–277. Cambridge University Press, may 2005.
- [18] R. A. Fisher. *The genetical theory of natural selection*. Clarendon Press, Oxford, 1930.
- [19] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann Appl Probab*, 14(4):1880–1919, nov 2004.
- [20] N. Friedman, S. Vardi, M. Ronen, U. Alon, and J. Stavans. Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biology*, 3(7):1261–1268, 2005.
- [21] C. Fritsch, F. Campillo, and O. Ovaskainen. A numerical approach to determine mutant invasion fitness and evolutionary singular strategies. *Theoretical Population Biology*, 115:89–99, jun 2017.
- [22] A. Golubev. Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression. *Journal of Theoretical Biology*, 393:203–217, mar 2016.
- [23] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, 1981.

- [24] S. Jaramillo-Riveri, J. Broughton, A. McVey, T. Pilizota, M. Scott, and M. El Karoui. Growth-dependent heterogeneity in the DNA damage response in *Escherichia coli*. *Molecular Systems Biology*, 18(5):e10441, may 2022.
- [25] E. C. Jones and S. Uphoff. Single-molecule imaging of LexA degradation in *Escherichia coli* elucidates regulatory mechanisms and heterogeneity of the SOS response. *Nature Microbiology* 2021 6:8, 6(8):981–990, jun 2021.
- [26] D. C. Karasz, A. I. Weaver, D. H. Buckley, and R. C. Wilhelm. Conditional filamentation as an adaptive trait of bacteria and its ecological significance in soils. *Environmental Microbiology*, 24(1):1–17, jan 2022.
- [27] V. Lagage and S. Uphoff. Pulses and delays, anticipation and memory: seeing bacterial stress responses from a single-cell perspective. *FEMS Microbiology Reviews*, 44(5):565–571, sep 2020.
- [28] J. T. Lennon, F. den Hollander, M. Wilke-Berenguer, and J. Blath. Principles of seed banks and the emergence of complexity from dormancy. *Nature Communications*, 12(1):4807, aug 2021.
- [29] J. W. Little and D. W. Mount. The SOS regulatory system of *Escherichia coli*, 1982.
- [30] S. Méléard, M. Rera, and T. Roget. A birth–death model of ageing: from individual-based dynamics to evolutive differential inclusions. *Journal of Mathematical Biology*, 79(3):901–939, aug 2019.
- [31] J. A. Metz, R. M. Nisbet, and S. A. Geritz. How should we define ‘fitness’ for general ecological scenarios? *Trends Ecol Evol*, 7(6):198–202, 1992.
- [32] P. Michel. Fitness optimization in a cell division model. *Comptes Rendus Mathématique*, 341(12):731–736, dec 2005.
- [33] N. Ojkic, E. Lilja, S. Direito, A. Dawson, R. J. Allen, and B. Waclaw. A Roadblock-and-Kill Mechanism of Action Model for the DNA-Targeting Antibiotic Ciprofloxacin. *Antimicrobial Agents and Chemotherapy*, 64(9), aug 2020.
- [34] B. Perthame. *Transport Equations in Biology*. Frontiers in Mathematics. Birkhäuser Basel, Basel, 2007.
- [35] S. Raghunathan, A. Chimthanawala, S. Krishna, A. G. Vecchiarelli, and A. Badrinarayanan. Asymmetric chromosome segregation and cell division in DNA damage-induced bacterial filaments. *Molecular Biology of the Cell*, 31(26):2920–2931, dec 2020.
- [36] N. M. Sampaio, C. M. Blassick, V. Andreani, J. B. Lugagne, and M. J. Dunlop. Dynamic gene expression and growth underlie cell-to-cell heterogeneity in *Escherichia coli* stress response. *Proceedings of the National Academy of Sciences of the United States of America*, 119(14):e2115032119, apr 2022.
- [37] Y. Shimoni, S. Altuvia, H. Margalit, and O. Biham. Stochastic Analysis of the SOS Response in *Escherichia coli*. *PLOS ONE*, 4(5):e5363, may 2009.

- [38] O. K. Silander, N. Nikolic, A. Zaslaver, A. Bren, I. Kikoin, U. Alon, and M. Ackermann. A genome-wide analysis of promoter-mediated phenotypic noise in *Escherichia coli*. *PLoS Genetics*, 8(1), jan 2012.
- [39] V. C. Tran. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM*, 12:345–386, oct 2008.
- [40] G. Webb. *Theory of Nonlinear Age-Dependent Population Dynamics*. Chapman & Hall Pure and Applied Mathematics. Taylor & Francis, 1985.
- [41] M. Wehrens, D. Ershov, R. Rozendaal, N. Walker, D. Schultz, R. Kishony, P. A. Levin, and S. J. Tans. Size Laws and Division Ring Dynamics in Filamentous *Escherichia coli* cells. *Current Biology*, 28(6):972–979.e5, mar 2018.
- [42] E. M. Witkin. The radiation sensitivity of *Escherichia coli* B: a hypothesis relating filament formation and prophage induction. *Proceedings of the National Academy of Sciences of the United States of America*, 57(5):1275–1279, may 1967.



# A Appendix

## A.1 Construction and well-posedness of $Z_t$

**Definition A.1** (Pathwise representation of the population process). Let  $Z_0$  be a counting measure on  $\mathbb{R}_+ \times \{0, 1\}$ , of the form of (1), and  $\mathcal{N}(du, dk, dz, d\omega)$  an independent Poisson point measure over  $\mathbb{R}_+ \times \mathbb{N}_* \times \mathbb{R}_+ \times [0, 1]^2$  with intensity  $du n(dk) dz d\omega$ , with  $n$  the counting measure on  $\mathbb{N}_*$ .  $Z_0$  represents the initial population and  $\mathcal{N}$  clocks the division and switching times that occur in some time interval measured by the integrating variable  $u$ . These events are measured for each individual  $k$  independently and happen proportionally to their division and switching rates, which are measured by the integrating variable  $z$ . Finally, the two independent uniform random variables  $\omega = (\omega_1, \omega_2)$  determine, respectively, the outcome of the first and the second daughter.

Thus, under the canonical filtration  $(\mathcal{F}_t)_{t \geq 0}$  generated by  $(Z_0, \mathcal{N})$ , we define the process  $(Z_t)_{t \geq 0}$  as

$$\begin{aligned}
Z_t &= \sum_{k=1}^{N_0} \delta_{(A_k(0)+t, I_k(0))} \\
&+ \int_0^t \int_{\mathbb{N}^* \times \mathbb{R}_+ \times [0, 1]^2} \mathbb{1}_{\{k \leq N_{u^-}\}} \left\{ \mathbb{1}_{\{z \leq \alpha(1 - I_k(u^-))\}} \left[ \delta_{(A_k(u^-)+(t-u), 1)} - \delta_{(A_k(u^-)+(t-u), I_k(u^-))} \right] \right. \\
&+ \mathbb{1}_{\{0 < z - \alpha(1 - I_k(u^-)) \leq \beta_{I_k(u^-)}(A_k(u^-))\}} \left( - \delta_{(A_k(u^-)+(t-u), I_k(u^-))} \right. \\
&+ \mathbb{1}_{I_k(u^-)=0} \left[ \mathbb{1}_{\omega_2 > p} \delta_{(t-u, 0)} + \mathbb{1}_{\omega_3 > p} \delta_{(t-u, 0)} \right] \\
&\left. \left. + \mathbb{1}_{I_k(u^-)=1} \left[ \mathbb{1}_{\omega_2 \leq \gamma} \delta_{(t-u, 0)} + \mathbb{1}_{\omega_2 > \gamma} \delta_{(t-u, 1)} + \mathbb{1}_{\omega_3 \leq \gamma} \delta_{(t-u, 0)} + \mathbb{1}_{\omega_3 > \gamma} \delta_{(t-u, 1)} \right] \right) \right\} \mathcal{N}(du, dk, dz, d\omega).
\end{aligned} \tag{53}$$

Note that

$$Z_t(da, di) = Z_t(da, \{0\})\delta_0(di) + Z_t(da, \{1\})\delta_1(di).$$

We explain now each term of the RHS of (A.1). The first line represents the deterministic evolution of the population when no random events happen before time  $t$ . The second line represents the switching events, that occur only for individuals of type  $I_k = 0$  at rate  $\alpha$ , and gives a new individual of type 1 with identical age while removing the previous individual of type 0 from the population. The third and four lines represent the divisions, which remove the divided cell from the population. If the mother type is 0 (third line), we add independently two cells of type 0, but only with probability  $1 - p$  each. If the mother type is 1 (fourth line) we add independently two cells whose type is decided by a Bernoulli random variable of parameter  $1 - \gamma$ .

We show that the stochastic process  $(Z_t)_t$  is well-defined under Assumptions 2.1:

**Proposition A.2** (Well-posedness and first-moment control). *Under Assumptions 2.1, and if  $\mathbb{E} \left[ \int (Z_0(da, \{0\}) + Z_0(da, \{1\})) \right] < \infty$  and  $\mathbb{E} \left[ \int a(Z_0(da, \{0\}) + Z_0(da, \{1\})) \right] < \infty$ , then the SDE (A.1) has a well-posed solution  $(Z_t)_{t \geq 0} \in \mathbb{D}(\mathbb{R}_+, \mathcal{M}(\{0, 1\} \times \mathbb{R}_+))$  which verifies for every  $t > 0$*

$$\begin{aligned}
\mathbb{E} \left[ \sup_{s \in [0, t]} \int (1 + a)(Z_s(da, \{0\}) + Z_s(da, \{1\})) \right] &\leq \mathbb{E} \left[ \int (1 + a)(Z_0(da, \{0\}) + Z_0(da, \{1\})) \right] \exp((\bar{b} + 1)t) \\
&< +\infty.
\end{aligned} \tag{54}$$

*Proof.* The proof is classical for populations with uniformly bounded birth rates, a suitable sequence of stopping times and Gronwall inequality to conclude. See for example [39].  $\square$

Using the Compensated Poisson Point Measure associated to  $\mathcal{N}$  and Ito's formula for semi-martingales [23] we can derive the following representation for  $\langle Z_t, f \rangle$ .

**Proposition A.3** (Semi-martingale decomposition). *Under control assumptions for the moments of  $Z_0$  and for control assumptions for  $\beta$  2.1,  $Z_t$  is well-posed for  $t \in [0, T]$  for any  $T > 0$ . Moreover, for any  $f \in C^1(\mathbb{R}_+ \times \{0, 1\})$ , we can write for any  $t \geq 0$*

$$\begin{aligned} \int (f(a, 0)Z_t(da, \{0\}) + f(a, 1)Z_t(da, \{1\})) &= \int (f(a, 0)Z_0(da, \{0\}) + f(a, 1)Z_0(da, \{1\})) \\ &+ \int_0^t \int (\mathcal{Q}f(a, 0)Z_s(da, \{0\}) + \mathcal{Q}f(a, 1)Z_s(da, \{1\}))ds + \mathcal{M}_t^f, \end{aligned} \quad (55)$$

where  $\mathcal{M}_t^f$  is a squared-integrable  $\mathcal{F}_t$ -martingale and the infinitesimal generator  $\mathcal{Q}$  is defined as

$$\begin{aligned} \mathcal{Q}f(a, i) &= \partial_a f(a, i) + (1 - i)\alpha(f(a, 1) - f(a, 0)) - \beta_i(a)f(a, i) \\ &+ 2(1 - i)(1 - p)\beta_i(a)f(0, 0) + 2i\beta_i(a)(\gamma f(0, 0) + (1 - \gamma)f(0, 1)). \end{aligned} \quad (56)$$

We can write Proposition A.3 in an abbreviated vectorial form. For all  $f : \mathbb{R}_+ \times \{0, 1\} \rightarrow \mathbb{R}$  which is  $C^1$  in the first coordinate, we set  $\mathbf{f}(a) = (f(a, 0), f(a, 1))$  and analogously  $\mathbf{f}'(a) = (\partial_a f(a, 0), \partial_a f(a, 1))$ . Then, we can write for  $\mathbf{f} \in (C^1(\mathbb{R}_+))^2$ ,  $\mathcal{Q}\mathbf{f} = (\mathcal{Q}f(a, 0), \mathcal{Q}f(a, 1))$  with  $\mathcal{Q}$  given by (19).

## A.2 Technical lemmas

We recall some useful classical properties.

**Proposition A.4.** *Let  $S, T$  be two bounded linear operators in a Banach space  $\mathcal{X}$  and let  $\alpha(S)$  the measure of non-compactness of  $S$  as defined in Definition A.7. Then,*

- (i)  $\alpha(S) \leq \|S\|_{op}$ .
- (ii)  $\alpha(TS) \leq \alpha(T)\alpha(S)$ .
- (iii)  $\alpha(T + S) \leq \alpha(T) + \alpha(S)$ .
- (iv)  $\alpha(S) = 0$  if and only if  $S$  is compact.

**Lemma A.5.** *Let  $A$  and  $B$  two positive matrices such that  $A_{ij} \geq B_{ij}$  for all pairs  $(i, j)$ , then:*

- (i)  $\rho(A) \geq \rho(B)$ , and
- (ii)  $\rho(B) \geq \rho(A) \left(1 - \frac{\max_{i,j}(A_{i,j} - B_{i,j})}{\min_{i,j} A_{i,j}}\right)$ .

### A.3 Complements on the proof of Proposition 4.5

We start our analysis by studying some estimates about the boundness and compactness of  $\mathbf{M}_t$ . To that purpose, we write the decomposition

$$\mathbf{M}_t \mathbf{f}(a) = \mathbf{U}_t \mathbf{f}(a) + \mathbf{W}_t \mathbf{f}(a),$$

where  $\mathbf{U}_t$  is destined to be small and  $\mathbf{W}_t$  is destined to be compact, in the senses that will be detailed further below (see for example Section 3.4 of [40] for a detailed motivation of this decomposition). In our case, the natural decomposition is given by Lemma A.6.

**Lemma A.6.** *Consider the semigroup  $\mathbf{M}_t$  on  $(L^1(\mathbb{R}_+))^2$  introduced by Def. 4.1. Then, for all  $\mathbf{f} \in (L^1(\mathbb{R}_+))^2$  we can write  $\mathbf{M}_t \mathbf{f}(a) = \mathbf{U}_t \mathbf{f}(a) + \mathbf{W}_t \mathbf{f}(a)$  where*

$$\mathbf{U}_t \mathbf{f}(a) := \Psi(a, a+t) \mathbf{f}(a+t), \quad (57a)$$

$$\mathbf{W}_t \mathbf{f}(a) := 2 \int_0^t \mathbf{K}(a, a+t-s) (\mathbf{I} - \mathbf{S}_1)^{-1} \mathbf{S}_2 \mathbf{f}(s) ds, \quad (57b)$$

where  $\mathbf{S}_1 : (L^1(\mathbb{R}_+))^2 \rightarrow (L^1(\mathbb{R}_+))^2$  is a locally compact linear operator of spectral radius 0, and  $\mathbf{S}_2 : (L^1(\mathbb{R}_+))^2 \rightarrow (L^1(\mathbb{R}_+))^2$  is a bounded linear operator, both defined by

$$\begin{aligned} \mathbf{S}_1 \mathbf{f}(t) &:= 2 \int_0^t \mathbf{K}(0, t-s) \mathbf{f}(s) ds, \\ \mathbf{S}_2 \mathbf{f}(t) &:= \Psi(0, t) \mathbf{f}(t). \end{aligned}$$

Moreover, for all fixed  $t > 0$ ,  $\mathbf{W}_t : (L^1([0, t]))^2 \rightarrow (L^1([0, t]))^2$  is a compact operator.

*Proof.* Fix  $t > 0$  and let  $\mathbf{f} \in (L^1([0, t]))^2$ . When needed, we extend  $\mathbf{f}$  to  $(L^1(\mathbb{R}))^2$  as  $\mathbf{f}(s) = (0, 0)$  whenever  $s > t$  or  $s < 0$ . The Duhamel representation obtained in (21), along with (24) gives that for all  $a \geq 0$ ,

$$\mathbf{M}_t \mathbf{f}(a) = \Psi(a, a+t) \mathbf{f}(a+t) + 2 \int_0^t \mathbf{K}(a, a+t-s) \mathbf{g}(s) ds,$$

where  $\mathbf{g}$  solves the fixed point problem

$$\mathbf{g}(s) = \Psi(0, s) \mathbf{f}(s) + 2 \int_0^s \mathbf{K}(0, s-u) \mathbf{g}(u) du = \mathbf{S}_2 \mathbf{f}(s) + \mathbf{S}_1 \mathbf{g}(s).$$

Therefore, if  $(\mathbf{I} - \mathbf{S}_1)^{-1}$  is well defined, we have that  $\mathbf{g}$  can be obtained as  $\mathbf{g} = (\mathbf{I} - \mathbf{S}_1)^{-1} \mathbf{S}_2 \mathbf{f}$ , from where we get (57b).

Hence, we start proving that  $\mathbf{S}_1 : (L^1([0, t]))^2 \rightarrow (L^1([0, t]))^2$  is a compact linear operator, to then conclude thanks to the Fredholm alternative.

Since functions are defined on  $[0, t]$ , the compactness of  $\mathbf{S}_1$  is deduced from the equicontinuity assumption of the Riesz-Fréchet-Kolmogorov Theorem. Consider some bounded sequence  $\{\mathbf{f}_n\}_n$  on the unit ball of  $(L^1([0, t]))^2$ . For all  $\mathbf{f} \in (L^1([0, t]))^2$  we have

$$\begin{aligned}
\int_0^t \|\mathbf{S}_1 \mathbf{f}(s+a) - \mathbf{S}_1 \mathbf{f}(s)\|_1 ds &= 2 \int_0^t \left\| \int_0^{s+a} \mathbf{K}(0, s+a-u) \mathbf{f}(u) du - \int_0^s \mathbf{K}(0, s-u) \mathbf{f}(u) du \right\|_1 ds \\
&\leq 2 \int_0^t \int_0^s \|\mathbf{K}(0, s-u+a) - \mathbf{K}(0, s-u)\|_1 \|\mathbf{f}(u)\|_1 duds \\
&\quad + 2 \int_0^t \int_s^{s+a} \|\mathbf{K}(0, s-u+a)\|_1 \|\mathbf{f}(u)\|_1 duds \\
&\leq 2 \int_0^t \int_0^s \|\mathbf{K}(0, s-u+a) - \mathbf{K}(0, s-u)\|_1 \|\mathbf{f}(u)\|_1 duds \\
&\quad + 2 \int_0^a \left( \int_0^u \|\mathbf{K}(0, s-u+a)\|_1 ds \right) \|\mathbf{f}(u)\|_1 du \\
&\quad + 2 \int_a^t \left( \int_{u-a}^u \|\mathbf{K}(0, s-u+a)\|_1 ds \right) \|\mathbf{f}(u)\|_1 du \\
&\leq 2 \int_0^t \int_0^s \|\mathbf{K}(0, s-u+a) - \mathbf{K}(0, s-u)\|_1 \|\mathbf{f}(u)\|_1 duds \\
&\quad + 2|a| \sup_{0 < s < t} \|\mathbf{K}(0, s)\|_1 \|\mathbf{f}\|_1.
\end{aligned}$$

The second term of the RHS can be controlled uniformly as  $|a| \rightarrow 0$ . We focus in the first term. By construction, as shown in (9), the jump times are absolutely continuous random variables. Therefore  $s \mapsto \mathbf{K}(0, s)$ , which contains the probability densities of the different jump events at time  $s$ , is a continuous application from  $\mathbb{R}_+$  to the vector space of positive  $2 \times 2$  matrices. In particular, it is uniformly continuous over the compact  $[0, t]$ , this is, for all  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$\|\mathbf{K}(0, s-a) - \mathbf{K}(0, s)\|_1 < \varepsilon \quad \text{for all } s \in [0, t[ \quad \text{whenever } a < \delta.$$

Therefore, for all  $a < \delta$  we have

$$\int_0^t \|\mathbf{S}_1 \mathbf{f}(s-a) - \mathbf{S}_1 \mathbf{f}(s)\|_1 ds \leq 2\varepsilon t + 2|a| \sup_{0 < s < t} \|\mathbf{K}(0, s)\|_1 =: \tilde{\varepsilon},$$

uniformly with respect to  $\mathbf{f}$ . Therefore  $\{\mathbf{S}_1 \mathbf{f}_n\}$  is relatively compact.

Since  $\mathbf{S}_1 : (L^1([0, t]))^2 \rightarrow (L^1([0, t]))^2$  is compact, any  $\lambda \neq 0$  is either an eigenvalue of  $\mathbf{S}_1$  or otherwise lies in the domain of the resolvent. Let's suppose by absurd that  $\lambda \neq 0$  is an eigenvalue associated to some nonzero eigenfunction  $\mathbf{u} \in (L^1([0, t]))^2$  such that  $\mathbf{S}_1 \mathbf{u} = \lambda \mathbf{u}$ , this is, for all  $a \in [0, t]$ ,

$$\mathbf{u}(a) = \frac{2}{\lambda} \int_0^a \mathbf{K}(0, a-s) \mathbf{u}(s) ds,$$

and thereby,

$$\|\mathbf{u}(a)\|_1 \leq \frac{2}{\lambda} \sup_{0 < s < t} \|\mathbf{K}(0, s)\|_1 \int_0^a \|\mathbf{u}(s)\|_1 ds.$$

Thus, by Grönwall's inequality  $\|\mathbf{u}(a)\|_1 = 0$  for all  $a \in [0, t]$ , and therefore  $\mathbf{u} \equiv 0$ , which is a contradiction. Therefore the spectrum of  $\mathbf{S}_1$  consists only in  $\{0\}$ . In particular the constant  $\lambda = 1$  is in the domain of the resolvent and  $(\mathbf{I} - \mathbf{S}_1)^{-1}$  is then well defined bounded linear operator.

We prove that  $\mathbf{S}_2 : (L^1([0, t]))^2 \rightarrow (L^1([0, t]))^2$  is bounded and then, rejoining all the previous results, we obtain the compactness of  $\mathbf{W}_t : (L^1([0, t]))^2 \rightarrow (L^1([0, t]))^2$  which allows us to conclude the Lemma.

1. **Boundness of  $\mathbf{S}_2$ .** For all  $\mathbf{f} \in (L^1([0, t]))^2$  we have

$$\int_0^t \|\mathbf{S}_2 \mathbf{f}(s)\|_1 ds \leq \int_0^t \|\Psi(0, s)\|_1 \|\mathbf{f}(s)\|_1 ds \leq \|\mathbf{f}\|_1,$$

since  $t \mapsto \Psi(0, t)$  is decreasing in each coordinate and therefore  $\sup_{t>0} \|\Psi(0, t)\|_1 = \|\Psi(0, 0)\|_1 = 1$ .

2. **Compactness of  $\mathbf{W}_t$ .** As we did for  $\mathbf{S}_1$ , for all  $\mathbf{f} \in (L^1([0, t]))^2$  we have

$$\begin{aligned} \int_0^t (\mathbf{W}_t \mathbf{f}(a+s) - \mathbf{W}_t \mathbf{f}(a)) da &= 2 \int_0^t \int_0^t (\mathbf{K}(a+s, a+s+t-u) - \mathbf{K}(a, a+t-u)) \\ &\quad (\mathbf{I} - \mathbf{S}_1)^{-1} \mathbf{S}_2 \mathbf{f}(u) du da, \end{aligned}$$

and therefore

$$\begin{aligned} \|\mathbf{W}_t \mathbf{f}(\cdot + s) - \mathbf{W}_t \mathbf{f}(\cdot)\|_1 &\leq 2 \int_0^t \left( \int_0^t \|\mathbf{K}(a+s, a+s+t-u) - \mathbf{K}(a, a+t-u)\|_1 da \right) \\ &\quad \|\mathbf{S}_2 \mathbf{f}(u)\|_1 du. \end{aligned}$$

By our previous calculations, we can bound  $\|(\mathbf{I} - \mathbf{S}_1)^{-1} \mathbf{S}_2 \mathbf{f}(u)\|_1$  uniformly for  $u \in [0, t]$  by some constant  $M > 0$  times  $\|\mathbf{f}\|_1$ . By the absolute continuity of the jump time densities,  $(a, s) \mapsto \mathbf{K}(a, s)$  is a continuous application in both coordinates, and therefore uniformly continuous on  $[0, t]$ . Thus, for all  $\varepsilon > 0$  there exists a  $\delta > 0$  such that whenever  $|s| < \delta$ , then

$$\|\mathbf{W}_t \mathbf{f}(\cdot + s) - \mathbf{W}_t \mathbf{f}(\cdot)\|_1 \leq 2Mt^2\varepsilon,$$

uniformly for all  $\mathbf{f} \in (L^1([0, t]))^2$  with  $\|\mathbf{f}\|_1 \leq 1$ . Therefore  $\mathbf{W}_t$  maps bounded sets to relatively compact sets and we conclude the proof. □

We recall now some useful definitions.

**Definition A.7** (Measure of non-compactness and growth bounds). We define the Kuratowski's measure of non-compactness of a bounded set  $A$  of a Banach space  $\mathcal{X}$  as

$$\alpha(A) := \inf \{ \varepsilon > 0 : A \text{ can be covered by a finite number of subsets of } \mathcal{X} \text{ of diameter } \leq \varepsilon \},$$

and the associated measure of non-compactness of a bounded operator  $S$  in  $\mathcal{X}$  as

$$\alpha(S) := \inf \{ \varepsilon > 0 : \alpha(S(A)) \leq \varepsilon \alpha(A) \text{ for all bounded sets } A \subset \mathcal{X} \}.$$

Finally we define the  $\alpha$ -growth bound of the semigroup  $\mathbf{M}_t$  as  $\omega_1(\mathbf{M}_t) := \lim_{t \rightarrow +\infty} \frac{1}{t} \log(\alpha(\mathbf{M}_t))$ .

Thanks to the decomposition proven in Lemma A.6, we obtain the following estimate on the non-compactness of  $\mathbf{M}_t$ :

**Proposition A.8.** *Under Assumptions 2.1 we have that  $\omega_1(\mathbf{M}_t) \leq -\bar{b} < 0$ .*

*Proof.* From Proposition A.4 applied to the decomposition proven on Lemma A.6, since  $\mathbf{W}_t$  is compact we have that

$$\alpha(\mathbf{M}_t) \leq \alpha(\mathbf{U}_t) \leq \|\mathbf{U}_t\|.$$

Now, for all fixed  $t \geq 0$ , and for all  $a \geq 0$ , we have term by term

$$\Psi(a, a+t) \leq \begin{bmatrix} e^{-(\alpha+b)t} & e^{-bt}(1-e^{-\alpha t}) \\ 0 & e^{-bt} \end{bmatrix},$$

so  $\|\Psi(a, a+t)\|_1 \leq e^{-bt}(2-e^{-\alpha t})$ . Therefore, for all  $\mathbf{f} \in (L^1(\mathbb{R}_+))^2$ , we have

$$\|\mathbf{U}_t \mathbf{f}\|_1 = \int_0^{+\infty} \|\Psi(a, a+t) \mathbf{f}(a+t)\|_1 da \leq e^{-bt}(2-e^{-\alpha t}) \|\mathbf{f}\|_1,$$

and therefore

$$\omega_1(\mathbf{M}_t) \leq \lim_{t \rightarrow +\infty} \frac{1}{t} \log \left( e^{-bt}(2-e^{-\alpha t}) \right) = -\bar{b}.$$

□

Then we conclude with the rest of the proof of Proposition 4.5 as stated in the main text.

#### A.4 Links between $\gamma \mapsto \lambda_{\alpha, \gamma}$ and the establishment probability conditions

Fix  $\gamma^* \in (0, 1)$ . We will introduce an auxiliary process  $\tilde{Z}_t$  that will allow to rederive Proposition 5.6 using the population establishment conditions. Let  $\tilde{Z}_t$  be the measure-valued process characterised under  $\mathbb{P}_\mu^{\alpha, \gamma}$  by an initial condition  $\tilde{Z}_0 = \mu$  and by the infinitesimal generator  $\tilde{\mathcal{Q}}_{\alpha, \gamma}$  defined as

$$\tilde{\mathcal{Q}}_{\alpha, \gamma} \mathbf{f}(a) := \mathbf{f}'(a) + 2\mathbf{B}_\gamma(a) \mathbf{f}(0) - (\mathbf{D}_\alpha(a) + \lambda_{\alpha, \gamma^*} \mathbf{I}) \mathbf{f}(a).$$

The only difference with respect to  $\mathcal{Q}_{\alpha, \gamma}$  defined in (19) is an additional death rate of value  $\lambda_{\alpha, \gamma^*}$  for both types. Consider the associated survival functions

$$\tilde{\psi}_i(s, t) := \psi_i(s, t) e^{-\lambda_{\alpha, \gamma^*}(t-s)}.$$

Then, following the proof of Theorem 3.2 we have that the extinction probabilities

$$\tilde{\pi}_i^{\alpha, \gamma} := \mathbb{P}_{\delta_{(0, i)}}^{\alpha, \gamma} \left( \exists t \geq 0 : \tilde{Z}_t = 0 \right)$$

are the minimal solutions in  $[0, 1]^2$  to the quadratic system

$$\begin{cases} \tilde{\pi}_0^{\alpha, \gamma} = \omega_0 + (1 - \tilde{q} - \omega_0)p + (1 - \tilde{q} - \omega_0)(1 - p)(\tilde{\pi}_0^{\alpha, \gamma})^2 + \tilde{q}\tilde{\pi}_1^{\alpha, \gamma}, & (58a) \end{cases}$$

$$\begin{cases} \tilde{\pi}_1^{\alpha, \gamma} = \omega_1 + (1 - \omega_1)(\gamma\tilde{\pi}_0^{\alpha, \gamma} + (1 - \gamma)\tilde{\pi}_1^{\alpha, \gamma})^2, & (58b) \end{cases}$$

where

$$\omega_i := \mathbb{P}_{\delta_{(0, i)}}^{\alpha, \gamma} (\text{Die by death rate}) = \lambda_{\alpha, \gamma^*} \int_0^{+\infty} \tilde{\psi}_i(0, s) ds, \quad (59)$$

$$\tilde{q} := \mathbb{P}_{\delta_{(0, 0)}}^{\alpha, \gamma} (\text{Switch before dividing}) = \alpha \int_0^{+\infty} \tilde{\psi}_0(0, s) ds. \quad (60)$$

We can further reduce (58a) to

$$\tilde{\pi}_0^{\alpha,\gamma} = (1 - \tilde{q})\tilde{p} + (1 - \tilde{q})(1 - \tilde{p})(\tilde{\pi}_0^{\alpha,\gamma})^2 + \tilde{q}\tilde{\pi}_1^{\alpha,\gamma}, \quad (61)$$

where  $\tilde{p} := p + (1 - p)\frac{\omega_0}{1 - q}$  is the *effective* death probability at division of type 0 induced by the additional death rate  $\lambda_{\alpha,\gamma^*}$ . Note that since  $0 \leq \omega_0 + q \leq 1$ , we have indeed  $0 \leq \tilde{p} \leq 1$ . This way, (61) has the same form as (13a) studied previously. Then, following the proof of Theorem 3.6, system (58a)-(58b) admits a minimal solution different than  $(1, 1)$  if and only if

$$\left\{ \tilde{p} \leq \frac{1}{2} \right\} \text{ or } \left\{ \tilde{p} > \frac{1}{2} \text{ and } \omega_1 \leq \frac{1}{2} \text{ and } \gamma < \frac{(1 - 2\omega_1)(\tilde{q} + (1 - \tilde{q})(2\tilde{p} - 1))}{2(\omega_1\tilde{q} + (1 + \omega_1)(1 - \tilde{q})(2\tilde{p} - 1))} \right\}. \quad (62)$$

We can recover the previous case (condition (16)) by making  $\omega_1 = 0$ .

We exhibit a monotonicity result analogous to Proposition 5.7, but for the system (58a)-(58b) that has an additional extinction probability  $\omega_1$ , absent in the previously studied system (13a)-(13b).

**Lemma A.9.** *For any solution of (58a)-(58b), for any  $i \in \{0, 1\}$ :*

- *If  $\omega_1 = \tilde{p}$ , then for all  $\gamma \in (0, 1)$ ,  $\partial_\gamma \tilde{\pi}_i^{\alpha,\gamma} = 0$ .*
- *If  $\omega_1 > \tilde{p}$ , then for all  $\gamma \in (0, 1)$ ,  $\partial_\gamma \tilde{\pi}_i^{\alpha,\gamma} \leq 0$ .*
- *If  $\omega_1 < \tilde{p}$ , then for all  $\gamma \in (0, 1)$ ,  $\partial_\gamma \tilde{\pi}_i^{\alpha,\gamma} \geq 0$ .*

*Proof.* Fix  $x \in (0, 1)$ . To study (58b), we consider the minimal solution to the quadratic equation

$$y = \omega + (1 - \omega)(\gamma x + (1 - \gamma)y)^2, \quad y \in [0, 1], \quad (E_1)$$

which is given by

$$y_x(\gamma, \omega) = \frac{1 - 2\gamma(1 - \gamma)(1 - \omega)x - \sqrt{1 - 4\gamma(1 - \gamma)(1 - \omega)x - 4(1 - \gamma)^2(1 - \omega)\omega}}{2(1 - \omega)(1 - \gamma)^2}. \quad (63)$$

We show first that there exists a unique  $\bar{\omega}_x \in (0, 1)$  such that for all  $\gamma \in (0, 1)$ ,  $\partial_\gamma y_x(\gamma, \bar{\omega}_x) = 0$ . By differentiation of (63) we have that, for  $x \neq 1$ ,  $\partial_\gamma y_x(\gamma, \bar{\omega}_x) = 0$  for all  $\gamma \in (0, 1)$  if and only if

$$\bar{\omega}_x = \frac{x}{1 + x}.$$

Now, to determine the value of  $p = p_x$  compatible with  $y_x(\gamma, \bar{\omega}_x) = x$ , we study the equation

$$x = (1 - q)p_x + (1 - q)(1 - p_x)x^2 + qy_x(\gamma, \omega), \quad (E_0)$$

which admits a solution  $x = y_x(\gamma, \omega)$  if and only if  $p_x = \frac{x}{1 + x}$ , from where we deduce  $\bar{\omega}_x = p_x$ . Then, by implicit differentiation of (E<sub>0</sub>), valuating under  $x = y_x(\gamma, \omega)$ , we have for all  $\gamma \in (0, 1)$ ,

$$\partial_\omega (\partial_\gamma y_x(\gamma, \bar{\omega}_x)) = -\frac{2(1 - x)x(1 + x)^2}{(1 - (1 - 2\gamma)x)^2} < 0.$$

In the final case, if  $x = 1$  then from (E<sub>0</sub>), for all  $\gamma \in (0, 1)$ ,  $y_x(\gamma, \omega) = 1$  too. In particular,  $\partial_\gamma y_x(\gamma, \omega) = 0$  for all  $\gamma \in (0, 1)$  and  $\omega \in (0, 1)$ .

Finally, since from (E<sub>0</sub>),  $x \in (0, 1) \mapsto y_x(\gamma, \omega) \in (0, 1)$  is increasing, if  $\partial_\omega \partial_\gamma y_x(\gamma, \bar{\omega}_x) \leq 0$  for all fixed  $x$ , by the implicit function theorem,  $\partial_\omega \partial_\gamma \tilde{\pi}_1^{\alpha,\gamma} \leq 0$ .  $\square$

Then, as in Lemma 4.6, we show now that there is an equivalence relation between condition (62) and the spectral properties of the matrix

$$\tilde{\mathbf{K}}_\infty^{\alpha,\gamma} := 2 \int_0^{+\infty} \tilde{\Psi}_\alpha(0, a) \mathbf{B}_\gamma(a) da.$$

As in Lemma 4.6 we obtain

$$\tilde{\mathbf{K}}_\infty^{\alpha,\gamma} = 2 \begin{bmatrix} (1 - \tilde{p})(1 - \tilde{q}) + \gamma\tilde{q} & (1 - \gamma)\tilde{q} \\ (1 - \omega_1)\gamma & (1 - \omega_1)(1 - \gamma) \end{bmatrix},$$

from which we can compute explicitly the eigenvalues, so that we are able to show that

$$\rho(\tilde{\mathbf{K}}_\infty^{\alpha,\gamma}) > 1 \iff (62) \text{ is verified} \iff \tilde{\pi}_0^{\alpha,\gamma^*}, \tilde{\pi}_1^{\alpha,\gamma^*} < 1. \quad (\star)$$

Then, following Lemma 4.7, there is a unique  $\tilde{\lambda}_{\alpha,\gamma} \in \mathbb{R}$  such that  $\rho(\tilde{\mathbf{F}}_{\alpha,\gamma}(\tilde{\lambda}_{\alpha,\gamma})) = 1$ . By construction,  $\tilde{\lambda}_{\alpha,\gamma} = \lambda_{\alpha,\gamma} - \lambda_{\alpha,\gamma^*}$ . Indeed,  $\lambda_{\alpha,\gamma}$  is the largest solution to (32) and then,

$$\det \left( 2 \int_0^{+\infty} e^{-(\lambda_{\alpha,\gamma} - \lambda_{\alpha,\gamma^*})s} \tilde{\mathbf{K}}_{\alpha,\gamma}(a, s) ds - \mathbf{I} \right) = \det \left( 2 \int_0^{+\infty} e^{-\lambda_{\alpha,\gamma}s} \mathbf{K}_{\alpha,\gamma}(a, s) ds - \mathbf{I} \right) = 0.$$

In particular, for  $\gamma = \gamma^*$ ,  $\tilde{\lambda}_{\alpha,\gamma^*} = 0$ , and therefore by  $(\star)$ ,

$$\tilde{\pi}_0^{\alpha,\gamma^*} = \tilde{\pi}_1^{\alpha,\gamma^*} = 1.$$

Now, suppose that  $\gamma^*$  is such that  $\omega_1 < \tilde{p}$ . Then, by Lemma A.9, for all  $\gamma \geq \gamma^*$ ,  $\tilde{\pi}_0^{\alpha,\gamma} = \tilde{\pi}_1^{\alpha,\gamma} = 1$ . Then, by  $(\star)$ ,  $\rho(\tilde{\mathbf{K}}_\infty^{\alpha,\gamma}) \leq 1$  for all  $\gamma \geq \gamma^*$ . Therefore, again by Lemma 4.7, we have  $\lambda_{\alpha,\gamma} \leq 0$  for all  $\gamma \geq \gamma^*$ , or equivalently,  $\lambda_{\alpha,\gamma} \leq \lambda_{\alpha,\gamma^*}$  for all  $\gamma \geq \gamma^*$ . We proceed analogously for the case  $\omega_1 > \tilde{p}$  and  $\omega_1 = \tilde{p}$ . Finally, recalling from Proposition 5.3 that  $\gamma \mapsto \lambda_{\alpha,\gamma}$  is continuously differentiable, we obtain that

$$\text{sign} \left( \partial_\gamma \lambda_{\alpha,\gamma} \Big|_{\gamma=\gamma^*} \right) = \text{sign}(\omega_1 - \tilde{p}) = \text{sign} \left( \mathbb{P}_{\delta_{(0,1)}}^{\alpha,\gamma^*} \left( \tilde{Z}_{T_1} = 0 \right) - \mathbb{P}_{\delta_{(0,0)}}^{\alpha,\gamma^*} \left( \tilde{Z}_{T_1} = 0 \mid \text{Do not switch} \right) \right).$$

Finally, from (59) and (60) and by an integration by parts we obtain

$$\omega_1 = \tilde{p} \iff p = \bar{p}(\alpha, \gamma) := \frac{\alpha \xi_0(\lambda_{\alpha,\gamma}) \xi_1(\lambda_{\alpha,\gamma}) + \lambda_{\alpha,\gamma} (\xi_0(\lambda_{\alpha,\gamma}) + \xi_1(\lambda_{\alpha,\gamma}) - 1)}{(\alpha + \lambda_{\alpha,\gamma}) \xi_0(\lambda_{\alpha,\gamma})},$$

where

$$\xi_i(\lambda) := \int_0^{+\infty} e^{-\lambda a} \beta_i(a) \exp \left( - \int_0^a \beta_i(s) ds \right) da$$

is the Laplace transform of the division time distribution of type  $i \in \{0, 1\}$ . Thus, for all  $\alpha, \gamma$ ,  $\bar{p}(\alpha, \gamma)$  is uniquely defined.