



Persistent and susceptible bacteria with individual deaths

Fabio Zucca*

Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy



HIGHLIGHTS

- We introduce two stochastic models for the evolution of a two state bacterial population subject to antibiotic treatment.
- Bacteria can switch from susceptible to persistent state and back.
- The efficiency of the antibiotic depends on a parameter p varying from 0 to 1.
- Critical time-interval lengths between antibiotic injections are studied and they might be highly dependent on p .
- We show when switching between susceptible and persistent states is a good strategy for bacteria and when it is not.

ARTICLE INFO

Article history:

Received 10 June 2013

Received in revised form

11 November 2013

Accepted 12 November 2013

Available online 21 November 2013

Keywords:

Bacteria persistence

Stochastic model

Multitype branching process

Random environment

ABSTRACT

The aim of this paper is to study two models for a bacterial population subject to antibiotic treatments. It is known that some bacteria are not sensitive to antibiotics. These bacteria, called *persisters*, are in a state called *persistence* and each bacterium can switch from this state to a non-persistent (or susceptible) state and back (with rates b and a respectively). Our models extend those introduced in [Garet et al. \(2012\)](#) by adding a random natural life cycle for each bacterium and by allowing bacteria in the susceptible state to escape the action of the antibiotic with a fixed probability $1 - p$ (while every bacterium in a persistent state survives with probability 1). This last mechanism of survival to the antibiotics differs from the persistent state one (where reproduction is forbidden) since in this case the bacterium can replicate. We study two different models. In the first model we “inject” the antibiotics in the system at fixed, deterministic times while in the second one the time intervals are random. We show that, in order to kill eventually the whole bacterial population, these time intervals cannot be “too large”. The maximum admissible length is increasing with respect to p ; we see that even when p is close to 1, this interval length can be significantly smaller than in the case $p = 1$. While in the case $p = 1$ switching back and forth to the persistent state is the only chance of surviving for bacteria, when $p < 1$ and the death rate in the persistent state, say d_r , is positive then the situation is more complex. In this case our model suggests that if d_r and b are positive (and fixed) then for higher values of p there is an interval for the rate a , say $(0, a_p)$ where switching to the persistent state is a good strategy while for $a > a_p$ the situation is less favorable than $a = 0$. On the other hand, for smaller values of p the best strategy is $a = 0$, that is, not switching. Finally, when $d_r = 0$, switching to the susceptible state is always a better strategy, from the bacterial point of view, than staying in the susceptible state all the times.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

It is well known that some bacteria are not sensitive to antibiotics (see [Bigger, 1944](#)). This state, called *persistence*, is not permanent and each bacterium can switch during its lifetime from persistent to susceptible and back to persistent many times (see for instance [Kussell et al., 2005](#); [Levin, 2004](#)). In the persistent state it does not reproduce, while in the susceptible state it breeds but it is also vulnerable to antibiotics. Antibiotic resistant bacteria are a major health concern. Studying the mechanism of switching to a persistent

state is the key to understanding how to fight efficiently certain diseases. Indeed, in many cases treatment failures may be explained by the presence of bacteria in a persistent state (see for instance [Levin, 2004](#); [Levin and Rozen, 2006](#)). Clearly the efficiency of the antibiotic treatment and the comparison of the natural death rates in the two states (persistent and susceptible) play a fundamental role here. In this paper we mainly investigate two questions. How is the treatment strategy going to be changed if the antibiotic is not perfectly efficient? Is switching to a persistent state always a good strategy, from the bacterial point of view, in order to increase the probability of survival as a population?

Two models for this phenomenon have been introduced in [Garet et al. \(2012\)](#). In those models, bacteria are immortal except when subject to an antibiotic while in the susceptible state: in that

* Tel.: +39 0223994546; fax: +39 0223994621.

E-mail address: fabio.zucca@polimi.it

case it is assumed that the antibiotic performs a perfect job and kills them all. Bacteria in the persistent state, on the other hand, are untouched. Without natural death rates (in both the persistent and susceptible states) it turns out that switching between the two states is a better strategy than staying in the susceptible state, from the bacterial point of view. Indeed, despite the fact that replication is forbidden in the persistent state, the protection from the action of the antibiotic (which kills all susceptible bacteria) represents an advantage for the colony. Hence it seems quite natural to introduce a death rate for bacteria for different causes other than antibiotics (such as immune system action and temporary local unfavorable environment conditions) and to suppose that, also in the persistent state, bacteria have a finite (random) life cycle. The addition of a death parameter in the persistent state, in particular, deeply modifies the evolution of the colony and it is no longer clear which strategy is the best (switching between states or remaining in the susceptible state).

Another simplification of the previous models, that we are going to remove in the present work, is the perfect efficiency of the antibiotic action. It seems unrealistic to assume that all susceptible bacteria are killed by a single antibiotic dose. It is plausible that, even in a susceptible state, a bacterium can survive some doses of the antibiotic and still being capable of reproducing. One reason may be that, since the action of the antibiotic depends on the direct interaction with the target, some (susceptible) bacteria might not be reached before the antibiotic decomposes and ceases to be effective; this prevents the completion of the mass-killing of all susceptible bacteria. Moreover, it is well known that not all bacteria are equally sensitive to antibiotics.

All these reasons suggest an opportunity for enhancing the models in [Garet et al. \(2012\)](#). In this paper we extend these models by adding (1) a life cycle for each bacterium (that is, individual deaths) and (2) a possibly positive probability $1-p$ for each bacterium in the susceptible state to survive the action of the antibiotic (p is the efficiency of the antibiotic). Roughly speaking, the parameter p takes into account all survival mechanisms for bacteria which do not turn the reproduction rate to zero. A similar parameter can be found, for instance, in a deterministic model studied in [Gardner et al. \(2007\)](#) where they call it *relative non-persist survival through catastrophes* and it is denoted by s .

To be precise, in our models each bacterium has an independent random lifetime represented by two exponentially distributed random variables with parameters d_n and d_r for the susceptible state and the persistent state respectively. As in [Garet et al. \(2012\)](#), bacteria in the susceptible state are allowed to reproduce (with rate λ) while they cannot replicate in the persistent state. Bacteria switch independently from the susceptible state to the persistent one and back at rates a and b respectively. At certain times, that we call *mass killing times* or simply *killing times*, an antibiotic is injected in the system; the time intervals are deterministic and equally spaced in the first model and random in the second one. The action of the antibiotic does not affect the persistent population but it kills each bacterium in the susceptible state independently with probability $p \in [0, 1]$; $p=0$ means that there is no target for the antibiotic in the bacterial genome, $p=1$ means that the antibiotic performs a “perfect” mass-killing action in the susceptible state population. The models in [Garet et al. \(2012\)](#) can be recovered by setting $d_n = d_r = 0$ and $p=1$. Let us emphasize here the difference between these two survival mechanisms in our models: in the persistent state, bacteria are not killed by the antibiotics and their reproduction rate is 0; in the susceptible state, some bacteria, selected independently with probability $1-p$, survive the antibiotic action and they can still reproduce at rate λ (clearly these bacteria and their offsprings could be killed by a subsequent antibiotic injection). Hence in our model we call *persisters* only those bacteria which are in the persistent state, which temporarily prevents them by being killed by the antibiotic and,

at the same time, does not allow any reproduction; besides, we have susceptible bacteria which can survive an antibiotic dose while retaining their ability to replicate (they are randomly selected with probability $1-p$, independently at each killing time).

For some values of the parameters (see [Section 2.1](#) for details), the system dies out almost surely even without the action of the antibiotics; thus, we need to study only the so-called supercritical case, which is the case when the natural evolution of the system allows survival with positive probability.

In the deterministic killing times case we suppose that the mass-killings occur at times $S_n := nT$ where $T > 0$. We show that if the interval T between each killing time is too large (strictly larger than a critical value $T_c \in (0, +\infty)$), the bacterial population has a positive probability of survival, while if $T \leq T_c$ there is almost sure extinction (see [Theorem 3.1](#)). We are also interested in the dependence of T_c from p : the critical time interval length T_c is a nondecreasing function of p (as expected, a more efficient medication can be administered less frequently). In particular, it may increase rapidly when p is close to one: this implies that the “perfectly efficient case $p=1$ ” might not be a good approximation for the case “ p close to 1”. Moreover as p converges to 0, T_c converges to 0 as well; thus there is not a positive minimum time interval which guarantees the extinction of the bacterial population for all $p \in (0, 1]$. When the death rate d_r is positive, for some set of parameters it might happen that switching from the susceptible to the resistant state is not a good strategy from the bacterial point of view, since it results in a longer critical time T_c . More precisely, if $d_r=0$ then switching to the persistent state is a good strategy for bacteria since we are forced to increase the treatment frequency in order to wipe them out. On the other hand, if $d_r > 0$ and the switching rate from susceptible to persistent is large (compared to the switching rate from persistent to susceptible) then eventually T_c becomes very large and the treatment frequency can be lowered (which means that switching is no longer a good strategy). In particular, our results suggest that if the efficiency p is sufficiently close to 1 then, in a finite interval of positive values for the switching rate a , T_c is smaller than in the case $a=0$ while, if the efficiency is sufficiently small, then there are no benefits whatsoever, from the bacterial point of view, in switching to the persistent state (since, in that case, T_c is increasing with respect to a).

In the random killing time case we suppose that the mass-killings are separated by a sequence of random time intervals $\{T_n\}_{n \geq 1}$; these variables are *independent and identically distributed* (from now on, i.i.d.) and the distribution is given by a probability measure μ_β , where $\{\mu_\beta\}_{\beta > 0}$ is a one-parameter stochastically increasing family of probability measures satisfying some mild conditions (see [Section 4](#) for details). The expected time interval is a nondecreasing function of β , hence β plays here the same role played by T in the deterministic killing times model. In this case we have two randomizations, so to speak: first we choose a realization ξ of the sequence $\{T_n\}_{n \geq 1}$ (we call ξ a *realization of the environment*) and then we have a random evolution of the system with killing times given by ξ . We show that if β is large enough ($\beta > \beta_c^1(p)$) the population survives with positive probability for almost every realization of the environment (see [Theorem 4.1](#)). On the other hand if β is small enough ($\beta > \beta_c^2(p)$) then the population dies out almost surely for almost every realization of the environment (see [Theorem 4.2](#)). As in the deterministic case, $\lim_{p \rightarrow 0} \beta_c^2(p) = 0$. Roughly speaking, since the expected time between two consecutive mass-killings is a non-decreasing function of β , we have that, in order to kill almost surely the bacterial population, the expected time between two injections of antibiotics in the system cannot be too large.

According to [Example 4.4](#), it might happen that $\beta_c^2(p) < \beta_c^1(p)$, that is, there is almost sure extinction of the bacteria for $\beta = \hat{\beta}$ and a positive probability of survival for the bacteria for $\beta = \hat{\beta}$ for some $\beta_c^2(p) < \hat{\beta} < \beta_c^1(p)$. Heuristically, from a mathematical point of view, this is due to the fact that the system is not stochastically monotone due to the switching between the states (see for instance [Liggett, 1985, 1999](#) for stochastic monotonicity and stochastic coupling). More details are given in [Section 4](#) before [Example 4.4](#). This example relies on a fairly particular distribution for the time intervals $\{T_n\}_{n \geq 1}$. One can reasonably expect that for most regular and single-peaked distribution $\beta_c^2(p) = \beta_c^1(p)$. This is the case, for instance, when $\mu_\beta \sim \exp(1/\beta)$ (where $\exp(1/\beta)$ is the exponential distribution with expected value β), that is, the interval between two consecutive injections is exponentially distributed. In that case, according to [Theorem 4.3](#), $\beta_c^1(p) = \beta_c^2(p) = \beta_c(p)$ and $\lim_{p \rightarrow 0} \beta_c(p) = 0$.

2. The dynamics

This is a modification of the model described in [Garet et al. \(2012\)](#) with the introduction of individual deaths for each type of bacteria; indeed, it is quite natural to assume that each bacterium has its own life cycle in the absence of an antibiotic treatment. Another addition to the dynamics is the possibility for each susceptible bacterium (independently from the others) to survive the action of the antibiotics with a fixed probability $1-p$ (where $p \in [0, 1]$). We denote by N_t and R_t the number of susceptible and persistent bacteria respectively. This is a 2-type process in continuous time, with the following (nonnegative) rates:

$$\begin{aligned} (N_t, R_t) &\rightarrow (N_t + 1, R_t) \text{ at rate } \lambda N_t \\ (N_t, R_t) &\rightarrow (N_t - 1, R_t + 1) \text{ at rate } a N_t \\ (N_t, R_t) &\rightarrow (N_t + 1, R_t - 1) \text{ at rate } b R_t \\ (N_t, R_t) &\rightarrow (N_t - 1, R_t) \text{ at rate } d_n N_t \\ (N_t, R_t) &\rightarrow (N_t, R_t - 1) \text{ at rate } d_r R_t. \end{aligned} \quad (2.1)$$

We recall that a change of state takes place at rate α if it takes place after a random exponentially distributed time intervals $T \sim \exp(\alpha)$: due to the lack of memory of the exponential distribution, this means that whenever we start looking at the system, the random time to wait before the change of state is a $\exp(\alpha)$ -distributed random variable. In particular, the probability of the change of state in an interval of time $[t, t + \Delta t]$ is asymptotic to $\alpha \cdot \Delta t$ as Δt goes to 0. A more precise construction of the model is given in the proof of [Theorem 4.3](#). Roughly speaking, we can imagine that each particle has five clocks which ring at exponentially distributed time intervals with parameters λ , a , b , d_n and d_r (the clocks are independent). When a particle is in a susceptible state we have different possibilities: if its $\exp(\lambda)$ -clock rings it breeds, if its $\exp(d_n)$ -clock rings it dies and if its $\exp(a)$ -clock rings it changes into a persistent state (it is not affected by the other clocks). On the other hand when a particle is in a persistent state we observe the following behaviors: if its $\exp(d_r)$ -clock rings it dies and if its $\exp(b)$ -clock rings it moves to a susceptible state (and, again, it is not affected by the other clocks).

When $a=0$ and $b=0$ the two populations are completely separated, the N -population is a branching process with mass killing and the R -population is stable if $d_r=0$ or dying out if $d_r > 0$ (see [Bertacchi and Zucca, 2008, 2009a, 2009b; Bertacchi et al., 2007](#) for some results on continuous-time branching-like processes). If $a > 0$ and $b=0$ then the N -population is a branching process with mass killing and individual death rate $a+d_n$, while the R -population survives if and only if either $d_r=0$ or the N -population survives. The interesting case is $b > 0$. We note that

this process is not monotone with respect to the parameters a and b ; on the other hand, it is monotone with respect to the other parameters and to the initial condition.

Without the mass deaths caused by the antibiotics, the process has a discrete-time branching random walk counterpart (similar to the one described in [Zucca, 2011](#)). When the antibiotic is injected in the system the dynamics is the following:

$$(N_t, R_t) \rightarrow (B(N_t, 1-p), R_t)$$

which means that the number of surviving susceptible bacteria is a binomial-distributed random variable; thus, at a killing time each susceptible bacterium is killed (independent from the others) with probability $p \in [0, 1]$. After a mass killing the system performs a new evolution starting from the survivors. If we consider just the surviving population at these mass killing times, we have a discrete-time process; it turns out to be a 2-type branching process or a 2-type branching process in random environment depending on our choice of the killing times (deterministic or random). Our choice will be either an increasing sequence of killing times $\{S_n\}_{n \geq 1}$ where $S_n = nT$ (for a fixed $T > 0$) or $S_n = \sum_{i=1}^n T_i$ where $\{T_i\}_{i \geq 1}$ is an i.i.d. sequence ($S_0 := 0$).

According to [Athreya and Ney \(1972\)](#), [Harris \(1963\)](#) and [Zucca \(2011\)](#) the long-term behavior of this discrete-time branching process depends only on its first-moment matrix $M = (m_{ij})_{i,j=1,2}$, where m_{ij} is the expected number of offsprings of type j from a particle of type i (see for instance [Zucca, 2011](#)). In order to compute M we need to consider the *mean field model* (this is done in [Section 2.1](#)). The main results on the deterministic case and the random case are in [Sections 3 and 4](#) respectively. We note that all these results hold for any finite (non-void) initial condition. All the proofs and technical lemmas can be found in [Section Appendix A](#).

2.1. Mean field model

This section is a useful exercise which allows us to obtain some explicit expressions we need in the sequel. The linear system of equations for the expected values $(n_t, r_t) := \mathbb{E}[(N_t, R_t)]$ is

$$\begin{cases} \frac{d}{dt} n_t = (\lambda - a - d_n) n_t + b r_t \\ \frac{d}{dt} r_t = a n_t - (b + d_r) r_t, \end{cases} \quad (2.2)$$

where $b > 0$ and $\lambda, a, d_n, d_r \geq 0$. The eigenvalues x^+ , x^- (where $x^+ \geq x^-$) of the corresponding matrix

$$A := \begin{pmatrix} \lambda - a - d_n & b \\ a & -(b + d_r) \end{pmatrix}$$

are the solutions of the equation

$$h(x) := x^2 + x(b + d_r - \lambda + a + d_n) - ((b + d_r)(\lambda - d_n) - ad_r) = 0. \quad (2.3)$$

We note immediately that, since $h(-(b + d_r)) = h(\lambda - a - d_n) = -ab \leq 0$, the eigenvalues are always real numbers and $x^- \leq \min(-(b + d_r), \lambda - a - d_n)$, $x^+ \geq \max(-(b + d_r), \lambda - a - d_n)$ (when $ab=0$ the previous inequalities become equalities, otherwise they are both strict inequalities). Moreover, the basic branching process theory tells us that if the maximum eigenvalue $x^+ \leq 0$ then we have almost sure (spontaneous) extinction. Hence if the determinant of the matrix $h(0) = -(b + d_r)(\lambda - d_n) + ad_r \geq 0$ we have extinction for all p and for any choice of $\{T_n\}_{n \geq 1}$ (even when $T_1 = +\infty$). From now on we assume

$$(b + d_r)(\lambda - d_n) - ad_r > 0, \quad (2.4)$$

that implies immediately $x^+ > 0$; hence $x^+ > x^-$. A corresponding pair of eigenvectors is $Z = (1, a/(b + d_r + x^+))$, $C = (1, a/(b + d_r + x^-))$.

The generic solution can be written as

$$\begin{pmatrix} n(t) \\ r(t) \end{pmatrix} = e^{At} \begin{pmatrix} n(0) \\ r(0) \end{pmatrix}$$

where $e^B := \sum_{i=0}^{\infty} B^i / i!$ for every matrix B and $(n(0), r(0))$ is the initial state. The explicit computations of e^{At} are easy: one simply needs to evaluate the solution of the system starting from $(1, 0)$ and $(0, 1)$. Note that $x^+ + x^- = \lambda - b - d_r - d_n - a$ and that $x^+ x^- = ad_r - (b + d_r)(\lambda - d_n)$. We have

$$\begin{pmatrix} \tilde{n}(t) \\ \tilde{r}(t) \end{pmatrix} := \begin{pmatrix} \frac{b + d_r + x^+}{x^+ - x^-} \\ \frac{a}{x^+ - x^-} \end{pmatrix} e^{bx^+} - \begin{pmatrix} \frac{b + d_r + x^-}{x^+ - x^-} \\ \frac{a}{x^+ - x^-} \end{pmatrix} e^{bx^-},$$

$$\begin{pmatrix} \bar{n}(t) \\ \bar{r}(t) \end{pmatrix} := \begin{pmatrix} \frac{b}{x^+ - x^-} \\ -\frac{b + d_r + x^-}{x^+ - x^-} \end{pmatrix} e^{bx^+} + \begin{pmatrix} \frac{b}{x^+ - x^-} \\ \frac{b + d_r + x^+}{x^+ - x^-} \end{pmatrix} e^{bx^-} \quad (2.5)$$

(remember that $b + d_r + x^- < 0$). We note that $\lim_{t \rightarrow \infty} \bar{n}(t) = \lim_{t \rightarrow \infty} \tilde{n}(t) = +\infty$; if, in addition, $a > 0$ then $\lim_{t \rightarrow \infty} \bar{r}(t) = \lim_{t \rightarrow \infty} \tilde{r}(t) = +\infty$. Hence

$$e^{At} = \begin{pmatrix} \tilde{n}(t) & \bar{n}(t) \\ \tilde{r}(t) & \bar{r}(t) \end{pmatrix} \quad (2.6)$$

note that $e^{At} e^{As} = e^{As} e^{At} = e^{A(t+s)}$.

3. Deterministic mass killing times

Between killing times, the bacterial population evolves randomly according to the rates (2.1), each time starting from the set of survivors of the previous killing time. We choose fixed time intervals $T_n = T$, where $T > 0$; hence mass killings occur at $S_n = nT$. We follow the strategy of [Garet et al. \(2012\)](#). For all $n \geq 0$, we let the system evolve between S_{n-1} and S_n and we count the number of survivors of each type at time S_n . This is a 2-type branching process, whence we have survival if and only if the Perron–Frobenius eigenvalue γ_T^+ of its first-moment matrix

$$M(T) := \begin{pmatrix} (1-p)\tilde{n}(T) & (1-p)\bar{n}(T) \\ \tilde{r}(T) & \bar{r}(T) \end{pmatrix} = \begin{pmatrix} 1-p & 0 \\ 0 & 1 \end{pmatrix} e^{AT} \quad (3.7)$$

satisfies $\gamma_T^+ > 1$ (see [Bertacchi and Zucca, 2012](#)). Note that the entries of the j -th column of the matrix are the average number of survivors after a mass killing at time T starting from one particle of type j ($j=1$ being a susceptible particle, $j=2$ being a persistent particle).

The following theorem holds for any (non-void) finite initial condition and the critical time $T_c(p)$ does not depend on the initial condition (clearly it depends on all the parameters of the system, even though, here, we emphasized only the dependence on p).

Theorem 3.1. Let $\lambda, a, d_n, d_r \geq 0$, $b > 0$ such that Eq. (2.4) holds and $a + 1 - p > 0$. For any $p > 0$ there exists $T_c(p) \in (0, +\infty)$ such that the process dies out almost surely if and only if $T \leq T_c(p)$. Moreover, $p \mapsto T_c(p)$ is a continuous, strictly increasing function such that $\lim_{p \rightarrow 0} T_c(p) = 0$, $\lim_{p \rightarrow 1} T_c(p) = T_c(1) < +\infty$ (when $a > 0$) and $\lim_{p \rightarrow 1} T_c(p) = +\infty$ (when $a = 0$).

Requiring the inequality (2.4) is quite natural, since if it does not hold, the bacterial population would become extinct almost surely even without the action of the antibiotic. Analogously, if $a + 1 - p = 0$ (that is, $p = 1$ and $a = 0$) there cannot be survival since at the first killing time the whole susceptible bacterial population is killed and the persistent population decreases (since they

cannot reproduce without switching to susceptible state and there is no switching back from susceptible to persistent). If $p < 1$ there can be survival even when switching from susceptible to persistent state is forbidden (that is, $a = 0$).

Since $p \mapsto T_c(p)$ is increasing we have that if $t > T_c(1)$ then there is survival with positive probability for all p , while if $t \in (0, T_c(1))$ there is a critical value $p_c(t) \in (0, 1)$ such that there is almost sure extinction if and only if $p \geq p_c(t)$.

From the bacterial point of view, a good situation would be when T_c is as small as possible. Here are three plots of T_c where $b = 0.01$ and $d_n = d_r = 0.025$. [Fig. 1](#) is the function $(\lambda, p) \mapsto T_c$ ($a = 0.01$). [Fig. 2](#) represents the functions $p \mapsto T_c$ (where $\lambda = 3.7$) for $a = 0$ (solid line), $a = 0.01$ (dot-dashed line), $a = 0.1$ (dotted line) and $a = 1$ (dashed line); note that on the x -axis we put $1 - p$ and we are using a logarithmic scale. This figure shows, for small values of a , a fast increase of T_c with respect to p as p is close to 1. This means that, for these values of a , if the antibiotic is slightly less than perfectly efficient (that is, $p < 1$) then the maximum admissible time interval T_c to kill the bacterial population is rapidly decreasing as p decreases. The same is shown also in [Fig. 3](#), which represents the functions $\lambda \mapsto T_c$ for $p = 1$ (solid line), $p = 0.95$ (dot-dashed line) and $p = 0.9$ (dashed line).

When $a = 0$ (which implies $\tilde{r} = 0$) the results of [Theorem 3.1](#) are straightforward. Indeed, we can just consider starting with one susceptible particle (if the initial particle is resistant the result is analogous). In this case it is sufficient to consider only the susceptible population and T_c can be computed explicitly as $T_c = (\lambda - d_n)^{-1} \log(1/(1-p))$ (note that the inequality (2.4) is equivalent to $\lambda > d_n$). Indeed the expected size of the susceptible population at time t (before the antibiotic injection) is $\tilde{n}(t) = \exp((\lambda - d_n)t)$ hence there is survival if and only if $(1-p) \exp((\lambda - d_n)t) > 1$.

Let us discuss briefly the behavior of T_c with respect to a , in order to understand if $a > 0$ is a better strategy than $a = 0$ from the bacterial point of view. First of all, when $p = 1$ the only hope for survival for the bacterial population is when $a > 0$. On the other hand, when $p < 1$ there might be a positive probability of survival even if $a = 0$, since a small fraction of susceptible bacteria may survive the action of the antibiotics (indeed, when $p < 1$ we have that $T_c < +\infty$ even if $a = 0$). Hence in this case, it is not trivial to decide whether $a > 0$ is a better strategy than $a = 0$ or not. We note

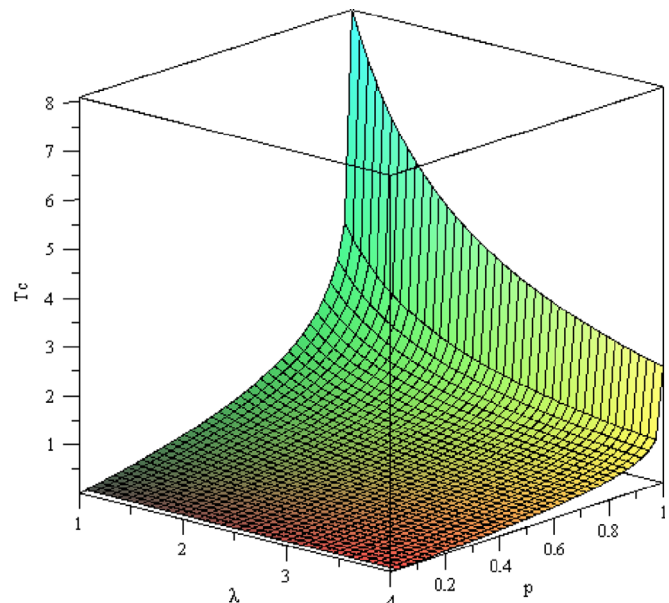
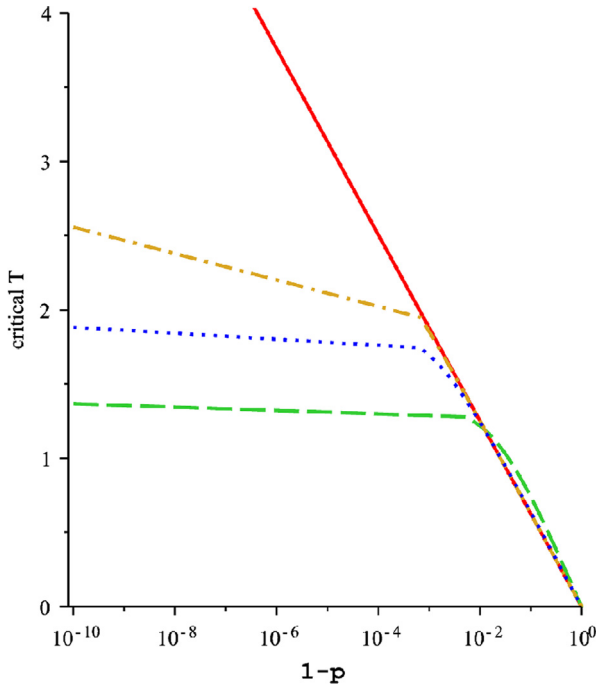
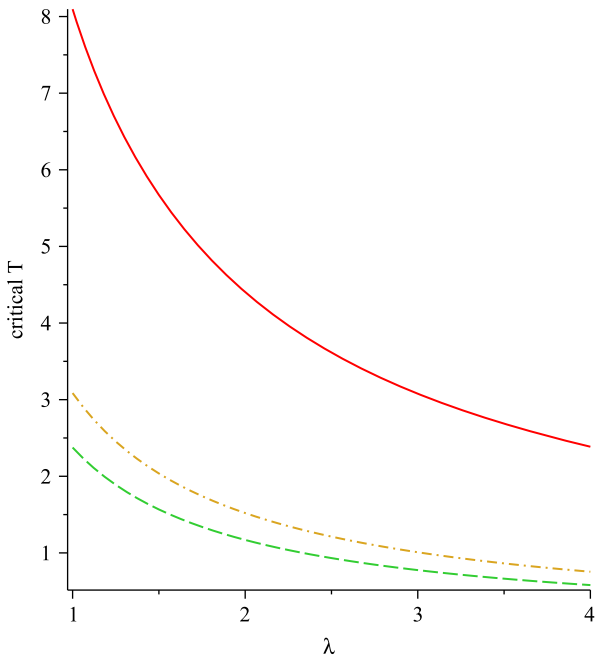
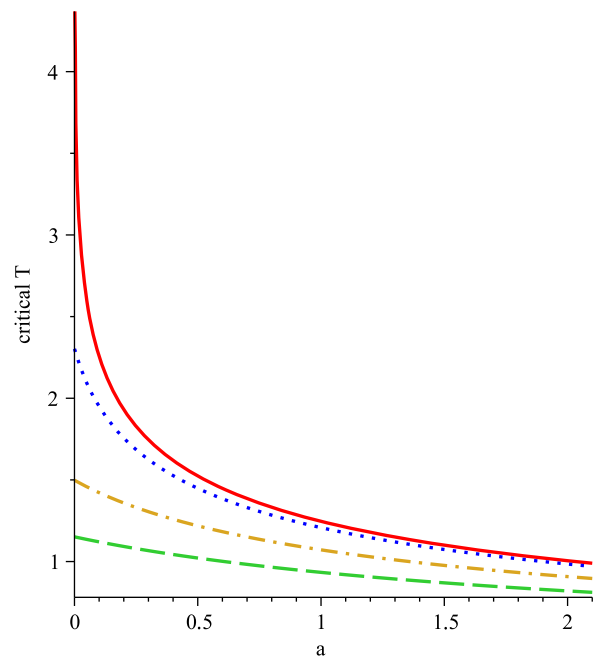


Fig. 1. $(\lambda, p) \mapsto T_c$.

Fig. 2. $p \mapsto T_c$.Fig. 3. $\lambda \mapsto T_c$.

that, if $d_r > 0$, then by Eq. (2.4) we have that $a \geq (\lambda - d_n)(b + d_r)/d_r$ implies a.s. extinction. More precisely one can prove that as $a \rightarrow (\lambda - d_n)(b + d_r)/d_r$ then $x^+ \rightarrow 0$ which implies $T_c \rightarrow \infty$ (see the proof of Theorem 3.1 for details) and eventually the situation becomes less favorable for the bacteria.

On the other hand, if we rewrite Eq. (2.3) as $h(x) = (x + b + d_r)(x - \lambda + d_n) + a(x + d_r)$ we see that, when $d_r = 0$, for every fixed $x > 0$ (resp. $x < 0$) h is strictly increasing (resp. decreasing) with respect to $a > 0$ and $h(x) \rightarrow +\infty$ (resp. $h(x) \rightarrow -\infty$) as $a \rightarrow \infty$. This implies that x^- and x^+ are strictly decreasing with respect to a and that $x^- \rightarrow -\infty$ and $x^+ \rightarrow 0$ as $a \rightarrow \infty$. Hence, when $d_r = 0$, it is easy to prove that $T_c \rightarrow 0$ as $a \rightarrow \infty$ (this can be done by checking that, for

Fig. 4. $d_r = d_n = 0$.

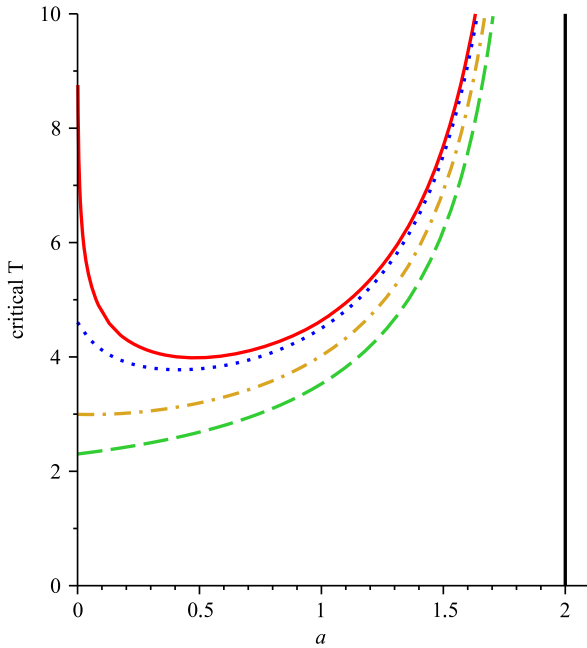
every fixed $t > 0$, the function $F_{t,p}(1)$, introduced in the proof of Theorem 3.1, is negative for all a sufficiently large). Here are two series of plots: Fig. 4 represents the case $d_r = d_n = 0$ and Fig. 5 the case $d_r = d_n = 1$ while the other parameters, with the exception of p , are fixed ($\lambda = 2$, $b = 1$): p equals 0.9 (dashed line), 0.95 (dot-dashed line), 0.99 (dotted line) and 1 (solid line).

We see that if $d_r = 0$ then the best strategy for the bacterial population is to have a as large as possible (which means having a large fraction of the population in the persistent state). If $d_r > 0$ there seems to be a critical value for p , above which, increasing a is a good strategy, up to a suitable value which minimizes T_c . Below the critical value for p the best situation for the bacterial population is $a = 0$. As $a \rightarrow (\lambda - d_n)(b + d_r)/d_r$ we have $x^+ \rightarrow 0$ which implies $T_c \rightarrow \infty$ and eventually the situation becomes worse for the survival of the bacteria for all values of p (in the previous figure $(\lambda - d_n)(b + d_r)/d_r = 2$ and it is represented by the black vertical asymptote). This difference between the cases $d_r = 0$ and $d_r > 0$ has a simple intuitive explanation: when $d_r = 0$ it is better to enhance the fraction of population which cannot be killed (even if this slows down the reproduction which is performed only by susceptible bacteria); indeed, in the limit $a \rightarrow \infty$, we have a stationary (immortal) population of persistent bacteria. On the other hand, if $d_r > 0$ and a is too large we have a population where the fraction of persistent bacteria is very high, hence the total reproduction rate is too low to counterbalance the death rate (which affects also persistent bacteria now). From another point of view, when $d_r > 0$ and $p < 1$, then bacteria find more convenient small values of a since they can make use of other survival mechanisms (represented by the probability $1 - p$ of surviving a single antibiotic dose) which are available in the susceptible state and, at the same time, they can take advantage of the reproduction process.

A more rigorous study of the behavior of T_c with respect to p , a and b is possible but it would exceed the aim of this paper.

4. Random mass killing times

We consider a sequence of i.i.d. positive random times $\{T_n\}_{n \geq 1}$. According to Lemma A.1, $\max\{n : S_n \leq t\} < +\infty$ a.s. for all $t > 0$, which means that there are a finite number of killing times in each

Fig. 5. $d_r = d_n = 1$.

finite interval almost surely. Moreover, suppose that the law of T_n is μ_β , where $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ is a family of probability measures on $(0, +\infty)$ (stochastically nondecreasing¹ w.r.t. β) satisfying

- (1) $\forall t_0 > 0, \lim_{\beta \rightarrow +\infty} \mu_\beta((0, t_0]) = 0$;
 - (2) $\forall \beta > 0, \mathbb{E}_\beta \int_0^\infty t \mu_\beta(dt) < +\infty$;
 - (3) $\forall t_0 > 0, \lim_{\beta \rightarrow 0} \int_{(t_0, \infty)} t \mu_\beta(dt) / \mathbb{E}_\beta = 0$.
- (4.8)

Clearly, since the family $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ is stochastically nondecreasing, we have that $\beta \mapsto \int_{(t_0, \infty)} t \mu_\beta(dt)$ is a nondecreasing function for every $t_0 \geq 0$. Moreover, (3) implies

- (4) $\forall t_0 > 0, \lim_{\beta \rightarrow 0} \mu_\beta((0, t_0]) = 1$.

The expected value of the length of the time intervals is a nondecreasing function of β . Roughly speaking, in Eq. (4.8), (1) implies that as β goes to infinity the probability of small time intervals is negligible. On the other hand, (3) tells us that, when β goes to 0, the contribution of large times to the expected length of the time intervals is negligible. As an example, consider the family of exponential laws $\exp(1/\beta)$.

In this case, we have two randomizations, first we choose a realization of the random sequence of times $\{T_n\}_{n \geq 1}$ (we call it, the *environment*) and then we consider the random evolution of the population with the chosen killing times. More precisely, the sequence of snapshots of the system taken at the random times $\{S_n\}_{n \geq 0}$ is a multitype branching process in random environment (see Tanny, 1981 for the definition). For each fixed β we call this the β -process and each realization ξ of the random time sequence $\{T_i\}_{i \geq 1}$ is our environment. Henceforth, when we say that some event \mathcal{A} (extinction or survival) has probability 0 (resp. > 0) for almost all realizations of the environment, we mean that the conditional probability of the event with respect to the realization ξ of the sequence of killing times is 0 (resp. > 0) for almost all

realizations ξ , that is, $\mathbb{P}(\xi : \mathbb{P}(\mathcal{A} | T_i = \xi_i, \forall i \geq 1) = 0) = 1$ (resp. $\mathbb{P}(\xi : \mathbb{P}(\mathcal{A} | T_i = \xi_i, \forall i \geq 1) > 0) = 1$).

Clearly if $\bar{q}(\xi) = (\bar{q}_1(\xi), \bar{q}_2(\xi))$ is the vector of extinction probabilities (starting from one susceptible bacterium or from one persistent bacterium respectively), we have that $\mathbb{P}(q(\xi) = \mathbf{1})$ is either 0 or 1. This means that there is a.s. extinction for almost all realizations of the environment or for almost no realizations of the environment.

Our results hold for any finite (non-void) initial condition (and, again, the critical values depend on all the parameters of the system but not on the initial condition). We assume again the inequality (2.4) to avoid spontaneous extinction of the bacterial population without the action of the antibiotic.

The first theorem states that if β exceeds some finite critical value β_c^1 there is survival almost surely, that is, for almost every realization of the environment. Roughly speaking, since the expected time is nondecreasing with respect to β , it means that if the expected time is too large, the action of the antibiotic might not be sufficient to kill the whole bacterial population.

Theorem 4.1. Let $\lambda, a, d_n, d_r \geq 0, b > 0$ such that Eq. (2.4) holds. Let $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ satisfy Eq. (4.8) and $a + 1 - p > 0$. If $\beta_c^1(p) := \sup \{\beta \in (0, +\infty) : \text{the } \beta\text{-process dies out a.s.}\}$ then $\beta_c^1(p) < +\infty$ and for all $\beta > \beta_c^1(p)$ we have survival with positive probability for almost all realizations of the environment. Moreover, $p \mapsto \beta_c^1(p)$ is nondecreasing.

The second result tells us that if β is smaller than some (strictly positive) critical value β_c^2 then, with probability 1, the antibiotic will eventually kill the bacterial population. We show that β_c^2 tends to 0 as p tends to 0. We also show (see Example 4.4) that there are situations where $\beta_c^2 \neq \beta_c^1$, that is, there is almost sure extinction of the bacteria for $\beta = \beta_c^2$ and a positive probability of survival for the bacteria for $\beta = \beta_c^1$ for some $\beta_c^2 < \beta_c^1$.

Theorem 4.2. Let $\lambda, a, d_n, d_r \geq 0, b > 0$ such that Eq. (2.4) holds. Let $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ satisfy Eq. (4.8). If $\beta_c^2(p) := \inf \{\beta \in (0, +\infty) : \text{the } \beta\text{-process survives with positive probability}\}$ then $\beta_c^2(p) > 0$ such that for all $\beta < \beta_c^2(p)$ we have a.s. extinction for almost all realizations of the environment. Moreover, $p \mapsto \beta_c^2(p)$ is nondecreasing and $\inf_{p \rightarrow 0} \beta_c^2(p) = 0$.

The case $a=0$ is very easy. Again, as in the case of deterministic time intervals, it is enough to consider only the susceptible population and to assume one susceptible bacterium at time 0. It is possible to show (see the end of Section Appendix A for details) that there is survival for the bacteria if and only if $\mathbb{E}_\beta[\log((1-p)\tilde{n}(T))] > 0$. In this case $\mathbb{E}_\beta[\log((1-p)\tilde{n}(T))] = \mathbb{E}_\beta[\log((1-p)\exp((\lambda - d_n)T))] = \log(1-p) + (\lambda - d_n)\mathbb{E}_\beta$. Since $\beta \mapsto \mathbb{E}_\beta$ is nondecreasing, $\beta_c^2 = \beta_c^1 = \inf \{\beta : \mathbb{E}_\beta > (\lambda - d_n)^{-1} \log(1/(1-p))\}$. Moreover, if $\beta \mapsto \mathbb{E}_\beta$ is continuous then there is almost sure extinction when $\beta = \beta_c^2$.

Sharper results can be obtained if we assume that the random times have a exponential distribution with expected value $1/\beta$. In this case there is a unique critical threshold β_c separating almost sure extinction from survival with positive probability.

Theorem 4.3. Let $\lambda, a, d_n, d_r \geq 0, b > 0$ such that Eq. (2.4) holds and $a + 1 - p > 0$. Let $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ be a sequence of exponential laws $\mu_\beta \sim \exp(1/\beta)$. There exists $\beta_c(p) \in (0, +\infty)$ such that for all $\beta > \beta_c(p)$ we have survival with positive probability for almost all realizations of the environment and for all $\beta < \beta_c(p)$ we have a.s. extinction for almost all realizations of the environment. Moreover, $p \mapsto \beta_c(p)$ is nondecreasing and $\lim_{p \rightarrow 0} \beta_c(p) = 0$.

We conjecture that in the critical case $\beta = \beta_c(p)$ there is a.s. extinction for almost all realizations of the environment; this can be easily proven when $p=1$ (see the remarks at the end of Section

¹ The stochastic order is the usual one, meaning that, if $\hat{\beta} \leq \tilde{\beta}$ then for all $x \in \mathbb{R}$ we have $\mu_{\hat{\beta}}((-\infty, x]) \leq \mu_{\tilde{\beta}}((-\infty, x])$.

Appendix A). When $p < 1$, the critical case is still open. Moreover, the main difference between [Theorems 4.1 and 4.2](#) and [4.3](#) is that in the last case $\beta_c^2(p) = \beta_c^1(p)$ while in general this is not true. The intuitive reason behind this is that the model is not stochastically monotone. In particular, this means that the probability of survival does not need to be monotone with respect to β . More details can be found at the end of the next example which shows that for a generic $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ satisfying our hypotheses, we cannot always expect $\beta_c^2(p) = \beta_c^1(p)$.

Example 4.4. Let us take $\lambda = (\sqrt{21} + 3)/4$, $b = (\sqrt{21} - 3)/4$, $a = d_n = d_r = 1/2$ and $p = 1$. Since $p = 1$ it is enough to consider the expected size of the persistent population (at each killing time, susceptible bacteria are killed). We note that $x^+ = 1$, $x^- = -1$ and $\bar{r}(t) = e^t(5 - \sqrt{21})/8 + e^{-t}(3 + \sqrt{21})/8$. From [Eq. \(A.1\)](#), the Perron–Frobenius eigenvalue of $M(t)$ is $\gamma_t^+ = \bar{r}(t)$; moreover, the only strictly positive solution of $\bar{r}(t) = 1$ is $T_c = \log(3 + \sqrt{21}) - \log(5 - \sqrt{21})$. We have $\bar{r}(t) < 1$ for all $t \in (0, T_c)$ and $\bar{r}(t) > 1$ for all $t > T_c$. Consider the following family of measures

$$\mu_{\beta^+} = \begin{cases} \frac{1}{2}\bar{\delta}_{\beta/10} + \frac{1}{2}\bar{\delta}_{3\min(\beta, 1)} & \beta \in (0, 15] \\ \frac{1}{2}\bar{\delta}_{\beta-13.5} + \frac{1}{2}\bar{\delta}_{\beta-12} & \beta \in (15, +\infty) \end{cases}$$

where $\bar{\delta}_\alpha$ is the Dirac measure at $\alpha \in \mathbb{R}$. Roughly speaking, for any fixed β , every time interval T_i is chosen independently between two values with probability $1/2$ each. This is a model of a patient which forgets to take his/her antibiotic dosage after a prescribed fixed time with probability $1/2$ and in that case he/she takes it after another fixed time interval. It is straightforward to see that the family $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ is stochastically nondecreasing and satisfies [Eq. \(4.8\)](#). According to [Smith and Wilkinson \(1969, Theorem 3.1\)](#) (see also [Remark A.2](#) for details) if $\mathbb{E}[\log(\bar{r}(T_1))] \leq 0$ there is a.s. extinction for almost every realization of the environment, while if $\mathbb{E}[\log(\bar{r}(T_1))] > 0$ there is positive probability of survival for almost every realization of the environment. Observe that $\mathbb{E}[\log(\bar{r}(T_1))]$ depends on β since μ_β represents the law of T_1 . Here we have extinction if β is close to 0 (take for instance, $\beta = 0.5$), we have survival if $\beta = 1$, we have extinction again if $\beta = 15$ and we have survival if β is large (take for instance, $\beta = 16.5$). Thus the probability of survival is not monotone and $\beta_c^2(p) \leq 1 < 15 \leq \beta_c^1(p)$.

The main reason behind this behavior is that the model is not monotone. This implies, in particular, that in general the map $t \mapsto \gamma_t^+$ is not monotone. Nevertheless it is continuous and there is only one strictly positive solution to the equation $\gamma_t^+ = 1$. This implies that in the deterministic time-interval case discussed in [Section 3](#), which can be retrieved by setting $\mu_{\beta^+} = \delta_\beta$, there is only one critical threshold (see [Theorem 3.1](#)) corresponding to the unique solution (with respect to β) of the equation $\mathbb{E}[\log(\gamma_{T_1}^+)] = 0$ (or, equivalently, of the equation $\gamma_\beta^+ = 1$). One can expect to obtain a similar behavior in most single-peaked distribution families $\{\mu_\beta\}_{\beta \in (0, +\infty)}$ (such as the exponential family considered in [Theorem 4.3](#)). The heuristic explanation is that, in the general case, even when $p = 1$, there might be multiple solutions (with respect to β) of the equation $\mathbb{E}[\log(\gamma_{T_1}^+)] = 0$ due to the fact that, roughly speaking, the function $\mathbb{E}[\log(\gamma_{T_1}^+)]$ is a mixing of all possible values of $\gamma_{T_1}^+$.

5. Conclusions

The model shows important differences between the evolution of the bacterial population in the case $p = 1$ and in the case $p < 1$. First

of all, the maximum admissible time interval between antibiotic injections is strongly dependent on p . Indeed, on the one hand it can be significantly smaller in the second case even when p is close to 1 and, on the other hand, its length converges to 0 as p tends to 0. This means that we cannot find a strictly positive time interval which ensures the extinction of the bacteria for all values of p . This suggests that, in the applications, an estimate of p could be important.

In the case of random intervals of injection, our model shows that, for a generic choice of the laws of the time intervals, even when $p = 1$ it can happen that we succeed in killing the bacteria (with probability 1) when the expected time length between injections is large while the bacteria can survive with positive probability for some smaller expected times. This does not happen when the intervals are deterministic.

Finally this model suggests an answer to the question whether switching from a state to the other one is always a good strategy for the bacteria (a good strategy here means to reduce the maximum admissible time interval between injections). Switching is a winning strategy if the death rate d_r in the persistent state is 0. On the contrary, in the case $d_r > 0$ we show that, when the transition rate a from the susceptible to the persistent state is too large, eventually the situation becomes less favorable for the bacteria since the maximum time interval between injections tends to infinity when a tends to a suitable finite value (see [Section 3](#)). In particular, as expected, for small values of p the best strategy for the bacteria is to stay in the susceptible state ($a = 0$), while if p is close to 1 then the best choice is a strictly positive value for a which minimizes the time interval length. We conjecture that, when $d_r > 0$ there is a critical value p_c separating these two different situations. We recall that, in our models, in the persistent state cells are dormant and do not reproduce; hence when switching is not a good strategy, it means that bacteria can take advantage of other survival mechanisms (represented by the probability $1 - p$ of surviving a single antibiotic dose) which are “available” in the susceptible state.

Acknowledgments

The author is grateful to Rinaldo Schinazi, Roberto Ambrosini, Andrea Franzetti, Sergio Foresti and Davide Mangioni for helpful discussions. The author is also thankful to the anonymous reviewers for their useful suggestions.

Appendix A. Proofs

Proof of Theorem 3.1. Define

$$\begin{aligned} F_{t,p}(x) &= x^2 - x((1-p)\tilde{n}(t) + \bar{r}(t)) \\ &\quad + (1-p)(\tilde{n}(t)\bar{r}(t) - \bar{n}(t)\tilde{r}(t)) \\ &= x^2 - x\left(e^{tx^+}\left(1 - p\frac{b+d_r+x^+}{x^+ - x^-}\right) + e^{tx^-}\left(1 + p\frac{b+d_r+x^-}{x^+ - x^-}\right)\right) \\ &\quad + (1-p)e^{t(x^+ + x^-)} \end{aligned} \quad (\text{A.1})$$

where $x^+ + x^- = \lambda - b - d_r - a - d_n$. Let $\gamma_t^+ \geq \gamma_t^-$ be the solutions of $F_{t,p}(x) = 0$. In order to check the inequality $\gamma_t^+ > 1$ we study the differentiable function $t \mapsto F_{t,p}(1)$ for every fixed $p \in (0, 1]$; in particular, we look for the solutions of the equation $F_{t,p}(1) = 0$ with respect to t . Clearly $\gamma_0^+ = 1$ and the other solution of $F_{0,p}(x) = 0$ is $\gamma_0^- = 1 - p < 1$. Using [Eqs. \(A.1\) and \(2.2\)](#) we have $(d/dt)F_{t,p}(1)|_{t=0} = p(d_r + b) > 0$. Hence there exists $\varepsilon > 0$ such that $F_{t,p}(1) > 0$ for all $t \in (0, \varepsilon)$; thus, by continuity and since $\gamma_0^- < 1$, we have that $\gamma_t^+ < 1$ for all $t \in (0, \varepsilon)$. Since $\lim_{t \rightarrow \infty} F_{t,p}(1) = -\infty$ for all $p \in (0, 1]$, there is at least one strictly positive solution to $F_{t,p}(1) = 0$ with respect to t . In order to show that it is unique, observe that,

since $x^- < 0$,

$$\frac{d}{dt}F_{t,p}(1) = e^{t(x^+ + x^-)}[(1-p)(x^+ + x^-) - (x^+ e^{t|x^-|}(1-ph) - |x^-|e^{-tx^+}(1+p(1-h)))]$$

where $h = (b + d_r + x^+)/ (x^+ - x^-)$. Clearly $\text{sgn}((d/dt)F_{t,p}(1)) = \text{sgn}(L(t, p))$ where $L(t, p) := F_{t,p}(1)e^{-t(x^+ + x^-)}$. Since $x^+ > 0$ we have that (for every fixed p) $t \mapsto L(t, p)$ is a strictly decreasing function such that $L(0, p) > 0$ and $\lim_{t \rightarrow +\infty} L(t, p) = -\infty$; thus there exists a unique $T_c = T_c(p) \in (0, +\infty)$ such that $F_{t,p}(1) > 0$ (resp. $F_{t,p}(1) < 0$) if $t \in (0, T_c)$ (resp. $t \in (T_c, \infty)$). This implies that $\gamma_t^+ < 1$ for all $t \in (0, T_c)$ and $\gamma_t^+ \geq 1$ for all $t \in [T_c, +\infty)$ (clearly $F_{T_c(p), p}(1) = 0$ and $\gamma_{T_c(p)}^+ = 1$).

By elementary analysis $p \mapsto T_p$ is a differentiable function (for every fixed $t \geq 0$). Moreover, by convexity, since $(b + d_r + x^+) / (x^+ - x^-) \in [0, 1]$,

$$\begin{aligned} \frac{d}{dp}F_{t,p}(1) &= e^{tx^+} \frac{b + d_r + x^+}{x^+ - x^-} + e^{tx^-} \left(-\frac{b + d_r + x^-}{x^+ - x^-} \right) - e^{t(x^+ + x^-)} \\ &\geq e^{t(x^+ + (b + d_r + x^+) / (x^+ - x^-))} - e^{t(x^+ + x^- + (b + d_r + x^-) / (x^+ - x^-))} \\ &\quad - e^{t(x^+ + x^-)} = e^{t(x^+ + x^- + b + d_r)} - e^{t(x^+ + x^-)} > 0 \end{aligned}$$

for all $t > 0$. Hence $p \mapsto F_{t,p}(1)$ is strictly increasing which implies that $p \mapsto T_c(p)$ is strictly increasing. Since $\lim_{p \rightarrow 0} F_{t,p}(1) = F_{t,0}(1) < 0$ for all $t > 0$ (indeed the process is supercritical in the absence of mass-killing), we have that $\lim_{p \rightarrow 0} T_c(p) = 0$. By continuity, if $a > 0$, $\lim_{p \rightarrow 1} T_c(p) = T_c(1)$ which is finite according to the first part of the proof. On the other hand, if $a = 0$ then $T_c = (\lambda - d_n)^{-1} \log(1/(1-p))$ and the conclusion follows.

Finally, if $a > 0$ then there is survival starting from 1 persistent particle if and only if there is survival starting from 1 susceptible particle; thus, since the process is monotone with respect to the initial state, the long-term behavior is the same as long as the initial state is finite. If $a = 0$ then $p < 1$ and the Perron–Frobenius eigenvalue $x^+ = m_{11}(t)$, hence our result holds starting from 1 susceptible particle; nevertheless, since $b > 0$, even if we start from 1 persistent particle there is a positive probability it becomes a susceptible one, hence there is a positive probability of survival starting from 1 susceptible particle if and only if there is a positive probability of survival starting from any finite initial state. \square

The proof of the following lemma is very easy, nevertheless we include it for completeness.

Lemma A.1. Let $\{T_i\}_{i \in \mathbb{N}}$ be nonnegative i.i.d. random variables. If $\mathbb{P}(T_1 > 0) > 0$ then $\mathbb{E}[N_t] < +\infty$ where $N_t := \max\{n : \sum_{i=0}^n T_i \leq t\}$.

Proof. Let $S_n := \sum_{i=0}^n T_i$ and suppose that $\mathbb{E}[T_i^4] < +\infty$: in this case define $\mathbb{E}[T_i] =: \mu > 0$, $\mathbb{E}[(T_i - \mu)^2] =: \sigma^2$ and $\mathbb{E}[(T_i - \mu)^4] =: r^4$. Clearly, eventually as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(N_t \geq n) &= \mathbb{P}(S_n \leq t) = \mathbb{P}(S_n/n - \mu \leq t/n - \mu) \\ &\leq \mathbb{P}(|S_n/n - \mu| \geq \mu/2) \\ &\leq \mathbb{E}[|S_n/n - \mu|^4] / (\mu/2)^4 \\ &= \frac{16}{n^4 \mu^4} \mathbb{E} \left[\left(\sum_{i=1}^n (T_i - \mu) \right)^4 \right] = (*). \end{aligned}$$

Now $(\sum_{i=1}^n (T_i - \mu))^4 = \sum_{i \in \{1, \dots, n\}^4} \prod_{j=1}^4 (T_{i_j} - \mu)$; moreover, the independence of $\{T_i\}_{i \in \mathbb{N}}$ yields $\mathbb{E}[\prod_{j=1}^4 (T_{i_j} - \mu)] = 0$ if there exists j such that $i_j \neq i_k$ for all $k \neq j$. Hence

$$(*) = \frac{16}{n^4 \mu^4} \mathbb{E} \left[\sum_{j=1}^n (T_{i_j} - \mu)^4 + \sum_{h,j,h \neq j} (T_{i_h} - \mu)^2 (T_{i_j} - \mu)^2 \right]$$

$$= \frac{16}{n^4 \mu^4} [nr^4 + n(n-1)\sigma^4] \leq C/n^2$$

thus $\mathbb{E}[N_t] = \sum_{n \in \mathbb{N}} \mathbb{P}(N_t \geq n) < +\infty$.

In the general case, define $\bar{T}_i := \min(T_i, 1)$. Then $\mathbb{E}[\bar{T}_i^4] < +\infty$ hence $\mathbb{E}[N_t] \leq \mathbb{E}[\bar{N}_t] < +\infty$ where $\bar{N}_t := \max\{n : \sum_{i=0}^n \bar{T}_i \leq t\}$. \square

Remark A.2. In the random killing time case we deal, in general, with a multitype branching process in random environment where the sequence of environments is i.i.d. hence, if we denote by $M_n := M(T_n)$ the first-moment matrix (3.7) with $T = T_n$, by Kingman Subadditive Theorem, we have (see for instance Athreya and Karlin, 1971; Furstenberg and Kesten, 1960; Kingman, 1973; Steele, 1989; Tanny, 1981)

$$\lim_{n \rightarrow \infty} n^{-1} \log \left(\left\| \prod_{i=1}^n M_i \right\| \right) = \delta_\beta = \mathbb{E}[\delta_\beta], \quad \text{a.s.}$$

and $\delta_\beta = \lim_{n \rightarrow \infty} n^{-1} \mathbb{E}[\log(\| \prod_{i=1}^n M_i \|)]$ where $\|M\| := \max_j \sum_i |M_{ij}|$ and $\prod_{i=1}^n M_i := M_n M_{n-1} \cdots M_1$. This plays the role of the Perron–Frobenius eigenvalue of the deterministic case and it will be useful in the next proofs (where we use Tanny, 1981, Theorems 9.6 and 9.10; in the case $p = 1$ one may use also Smith and Wilkinson, 1969, Theorem 3.1 instead).

Moreover, it is easy to check that the conditions of Tanny (1981, Theorems 9.10) are satisfied. Indeed, the entries of the first-moment matrix satisfy $m_{ij}(t) > 0$ for all $t > 0$ and for all $ij = 1, 2$. Moreover, since $\mu_\beta((0, +\infty)) = 1$ for all $\beta \in (0, +\infty)$, we have $\mathbb{P}(\min_{ij} (M_1)_{ij} > 0) = 1$. Finally, if we start with a susceptible particle then $\mathbb{P}(N(t) \geq 1) \geq e^{-(a+d_n)t}$ hence

$$\begin{aligned} \mathbb{E}[|\log(1 - \mathbb{P}(N(t) = 0))|] \\ \leq \int_{(0, +\infty)} (a + d_n)t \mu_\beta(dt) < +\infty, \quad \forall \beta \in (0, +\infty). \end{aligned}$$

On the other hand, if the initial condition is a persistent particle we proceed by using $R(t)$ instead of $N(t)$. One can check analogously that our branching process in random environment is strongly regular (see Tanny, 1981, Definition 9.1). Hence, according to Tanny (1981, Theorem 9.10), we have:

- (1) $\delta_\beta \leq 0$ implies a.s. extinction for almost all realizations of the environment,
- (2) $\delta_\beta > 0$ implies survival with positive probability for almost all realizations of the environment.

Clearly the probability of survival is 0 if and only if the conditional probability of survival is 0 for almost all realizations of the environment. On the other hand, since $\mathbb{P}(q(\xi) = 1)$ is either 0 or 1 (see Section 4), then the probability of survival is strictly positive if and only if the conditional probability of survival is strictly positive for almost all realizations of the environment.

Proof of Theorem 4.1. First of all we check the integrability condition, that is, for all $n \geq 1$,

$$\begin{aligned} \mathbb{E} \left[n^{-1} \log \left(\left\| \prod_{i=1}^n M_i \right\| \right) \right] \\ = \int n^{-1} \log \left(\left\| \prod_{i=1}^n M(t_i) \right\| \right) \prod_{i=1}^n \mu_\beta(dt_i) < +\infty, \end{aligned}$$

where $\prod_{i=1}^n \mu_\beta$ is a probability product measure on \mathbb{R}^n . Below we show that if $p < 1$ then $\| \prod_{i=1}^n M_i \| \geq (1-p)^n \varepsilon^n$ (for some $\varepsilon > 0$) while if $a > 0$ then $\| \prod_{i=1}^n M_i \| \geq \varepsilon^n$ (for some $\varepsilon > 0$). Hence if

$a+1-p > 0$, for some $\varepsilon' > 0$,

$$\int n^{-1} \log^{-} \left(\left\| \prod_{i=1}^n M(t_i) \right\| \right) \prod_{i=1}^n \mu_{\beta}(dt_i) \leq \log^{-}(\varepsilon') < +\infty$$

since \log^{-} is nonincreasing (where $\log^{-}(\cdot) := \max(0, -\log(\cdot))$). Thus we just need to prove that $\int n^{-1} \log^{+}(\|\prod_{i=1}^n M(t_i)\|) \prod_{i=1}^n \mu_{\beta}(dt_i) < +\infty$ (where $\log^{+}(\cdot) := \max(0, \log(\cdot))$). From Eq. (2.5) we have $\|M(t)\| \leq Ke^{\alpha t}$; hence, since \log^{+} is nondecreasing, $\|\prod_{i=1}^n M_i\| \leq \prod_{i=1}^n \|M_i\|$ and the expected value of μ_{β} is finite (for all β) we have $\int n^{-1} \log^{+}(\|\prod_{i=1}^n M(t_i)\|) \prod_{i=1}^n \mu_{\beta}(dt_i) < +\infty$.

Suppose $p < 1$ and $a > 0$. For all $t, \tau > 0$ there exists $\beta_0(\tau, t)$ such that $\mu_{\beta}([t, +\infty)) > 1 - \tau$ for all $\beta > \beta_0(\tau, t)$. By continuity and compactness we have that, for some $\varepsilon > 0$,

$$M(t) \geq \begin{pmatrix} (1-p)\varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} =: \bar{M}_0, \quad \forall t \geq 0$$

where, by definition, $A \geq B$ if and only if $A_{ij} \geq B_{ij}$. It is easy to show, by using Eq. (2.5), that there exists $t_p \in [0, +\infty)$ such that

$$M(t) \geq \begin{pmatrix} \frac{4}{\varepsilon(1-p)} & 0 \\ 0 & \frac{4}{\varepsilon} \end{pmatrix} =: \bar{M}_1, \quad \forall t \geq t_p.$$

Let $\beta > \beta_0(1/2, t_p)$ and $\{M_i\}_{i \geq 1}$ be the corresponding sequence of random first-moment matrices ($M_i := M(T_i)$); thus, according to the Law of Large Numbers, with probability 1, as $n \rightarrow \infty$, $\#\{i \leq n : M_i \geq \bar{M}_1\} \geq n/2$ which implies that $\prod_{i=1}^n M_i \geq \bar{M}_0^{n/2} \bar{M}_1^{n/2} = 2^{n/4}$ almost surely. Hence $\liminf_{n \rightarrow \infty} n^{-1} \log(\|\prod_{i=1}^n M_i\|) \geq \log 2 > 0$ a.s. which, according to Tanny (1981, Theorems 9.10), implies survival with positive probability for almost every realization of the environment. Hence, by definition, $\beta_1(p) \leq \beta_0(1/4, t_p)$.

The usual coupling technique shows that for any fixed choice of the parameters λ, a, b, d_n, d_r and for any realization of the environment, the probability of survival is nonincreasing with respect to p . Hence $p \mapsto \beta_1(p)$ is nondecreasing.

If $p=1$ then $a > 0$ and we are dealing essentially with a single population (the persistent bacteria as in Garet et al., 2012), since after each killing time we have just persistent bacteria left. The first moment, starting with a susceptible bacterium, is $\bar{r}(t)$. The proof is essentially the same since $\bar{r}(t) \geq \varepsilon > 0$ for all $t \geq 0$ and $\bar{r}(t) \geq 4/\varepsilon$ for all $t \geq t_p$. The result follows from Smith and Wilkinson (1969, Theorem 3.1). Since $a > 0$ then there is survival starting with a persistent bacterium if and only if there is survival starting with a susceptible one.

Finally if $a=0$ (hence $p < 1$) again we are dealing essentially with a single population: the susceptible bacteria. The first moment, starting with a susceptible bacterium, is \bar{n} and the result follows (from Smith and Wilkinson, 1969, Theorem 3.1 as before) from the inequalities $\bar{n}(t) \geq (1-p)\varepsilon$ for all $t \geq 0$ and $\bar{n}(t) \geq 4/((1-p)\varepsilon)$ for all $t \geq t_p$. \square

Proof of Theorem 4.2. If $p=1$ and $a=0$ the process becomes extinct almost surely. We suppose henceforth that $a+1-p > 0$. Since \log is increasing and $\{M_i\}_{i \geq 1}$ are identically distributed, we have

$$\mathbb{E} \left[n^{-1} \log \left(\left\| \prod_{i=1}^n M_i \right\| \right) \right] \leq n^{-1} \mathbb{E} \left[\log \left(\prod_{i=1}^n \|M_i\| \right) \right] \\ \leq \log(\mathbb{E}[\|M_1\|]).$$

Thus, if we prove that $\log(\mathbb{E}[\|M_1\|]) < 0$ for every sufficiently small β then Tanny (1981, Theorem 9.6) guarantees a.s. extinction for

almost all configurations (if $p=1$ one can also use Smith and Wilkinson (1969, Theorem 3.1) instead).

It is straightforward to show that $t \mapsto \|M(t)\|$ is differentiable from the right at 0. By elementary computations, since $M(0) = 1$, $(d/dt) \log \|M(t)\| |_{t=0} = (d/dt) \|M(t)\| |_{t=0} =: -m < 0$. Hence, there exists $t_0 > 0$ such that $\log(\|M(t)\|) \leq -tm/2$ for all $t \in [0, t_0]$. On the other hand, $\log(\|M(t)\|) \leq Cx^+ t$ for all $t > 0$ and some $C > 0$. Finally note that in Eq. (4.8), condition (3) is equivalent to

$$\forall t_0 > 0, \quad \lim_{\beta \rightarrow 0} \frac{\int_{(t_0, \infty)} t \mu_{\beta}(dt)}{\int_{(0, t_0]} t \mu_{\beta}(dt)} = 0,$$

thus

$$\begin{aligned} & \int \log(\|M(t)\|) \mu_{\beta}(dt) \\ &= \int_{(0, t_0]} \log(\|M(t)\|) \mu_{\beta}(dt) + \int_{(t_0, \infty)} \log(\|M(t)\|) \mu_{\beta}(dt) \\ &\leq - \int_{(0, t_0]} \frac{tm}{2} \mu_{\beta}(dt) + \int_{(t_0, \infty)} Cx^+ t \mu_{\beta}(dt) \\ &\leq - \int_{(0, t_0]} \frac{tm}{2} \mu_{\beta}(dt) \left(1 - \frac{\int_{(t_0, \infty)} t \mu_{\beta}(dt) 2Cx^+}{\int_{(0, t_0]} t \mu_{\beta}(dt) m} \right) \\ &< 0 \end{aligned}$$

for every sufficiently small β . Hence $\beta_2(p) > 0$.

As in the proof of Theorem 4.2, for every realization of the environment, the probability of survival is nonincreasing with respect to p . Hence $p \mapsto \beta_2(p)$ is nondecreasing.

Let us fix $\beta > 0$. It is well-known that

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{-1} \log \left(\left\| \prod_{i=1}^n e^{A T_i} \right\| \right) &= \lim_{n \rightarrow \infty} n^{-1} \log \left(\|e^{A \sum_{i=1}^n T_i}\| \right) \\ &= x^+ \mathbb{E}_{\beta} > 0, \quad \text{a.s.} \end{aligned}$$

(e^{x^+} is the maximum eigenvalue of e^A). Moreover, $M(t) \geq (1-p)e^{At}$ thus $\prod_{i=1}^n M_i \geq (1-p)^n e^{A \sum_{i=1}^n T_i}$ and

$$\lim_{n \rightarrow \infty} n^{-1} \log \left(\left\| \prod_{i=1}^n M_i \right\| \right) \geq \log(1-p) + x^+ \mathbb{E}_{\beta} \quad \text{a.s.}$$

which is eventually strictly positive, as $p \rightarrow 0$. Hence, according to Tanny (1981, Theorem 9.10), the process eventually survives as $p \rightarrow 0$; thus, by definition, $\beta_2(p) \leq \beta$ eventually as $p \rightarrow 0$. \square

Proof of Theorem 4.3. We use a modification of the construction shown in Garet et al. (2012). Given $\beta_1 \geq \beta_2$, it is possible to construct two sequences $\{\bar{T}_i^1\}_{i \geq 1}$ and $\{\bar{T}_i^2\}_{i \geq 1}$ in such a way that, for every trajectory, $\{\bar{T}_i^1(\omega) : i \geq 1\} \subseteq \{\bar{T}_i^2(\omega) : i \geq 1\}$. This can be done by using the classical decimation procedure: take the sequence of arrival times $\{T_i^2\}_{i \geq 1}$ and a sequence $\{B_i\}_{i \geq 1}$ of independent Bernoulli variables with parameter β_2/β_1 . Now, for every trajectory, remove the i -th arrival time $T_i^2(\omega)$ if and only if $B_i(\omega) = 0$; the surviving arrival times are a realization of $\{\bar{T}_i^1\}_{i \geq 1}$.

Consider the binary tree \mathbb{T} whose vertices are the set V of finite words of the alphabet $[0, 1]$ and whose root is the empty word \emptyset . Every nonempty word $v_1 v_2 \dots v_n$ is connected to its parent $v_1 v_2 \dots v_{n-1}$ and its two children $v_1 v_2 \dots v_n 0$ and $v_1 v_2 \dots v_n 1$ (the root is connected to 0 and 1).

To each vertex v corresponds a variable $S_v \sim \exp(\lambda)$ which represents the time interval between its birth and its splitting (when it gives birth). We assume that $\{S_v\}_{v \in V}$ is an i.i.d. family of random variables. Define $T_{\emptyset} := 0$ and, for every nonempty word $v = v_1 \dots v_n$, $T_v = \sum_{i=1}^n S_{v_1 \dots v_i}$. Consider now the tree $\hat{\mathbb{T}}$ on $\mathbb{T} \times [0, +\infty)$ as follows: the set of vertices is $\hat{V} := \{(v, T_v), (v, T_v + S_v) : v \in \mathbb{T}\}$. We have vertical edges between (v, T_v) and $(v, T_v + S_v)$ (for all $v \in V$); we have horizontal edges between $(v, T_v + S_v)$ and each of its two children

$(vw, T_v + S_v)$ where $w \in \{0, 1\}$ (for all $v \in V$). The vertical edge between (v, T_v) and $(v, T_v + S_v)$ represents the time interval between the birth of the particle v and its splitting time. The horizontal edge between (v, T_v) and $(v1, T_v)$ represent the birth of a child of v while we consider the other particle, namely $v0$, as v itself after giving birth.

Independent of everything constructed so far, we consider four independent families of Poisson point processes $\{W_v^1\}_{v \in V}$, $\{W_v^2\}_{v \in V}$, $\{D_v^1\}_{v \in V}$ and $\{D_v^2\}_{v \in V}$ on $[0, +\infty)$ with intensities b , a , d_n and d_r respectively. We color the tree in white (susceptible state), red (persistent state) and black (dead particle) as follows: we start with a white vertex $(\emptyset, 0)$ and we extend the color to the branches along the timeline until we reach a point of one of the Poisson processes. If we meet a D_v^1 point and the current color is white we switch to black and there are not modifications anymore in that subtree along the timeline (death of the particle), the same happens if we meet a D_v^2 point and the current color is red. If the current color is not black, then everytime we meet a W_v^1 point we switch to red and everytime we meet a W_v^2 point we switch to white. Black color is not modified when we meet new points in the Poisson processes. At every split point if the current color is white or black then we use the same color for the horizontal edges and we continue starting from the two new vertices. If the color is red then we use the same color for the horizontal edge which connects to the child whose name ends with 0 (and we start again from there with a red vertex); we switch to black for the horizontal edge connecting to the other son (hence, the whole subtree branching from this vertex is black, since red (persistent) particles do not reproduce).

So far we modelled the natural evolution of the system; now we add the action of the antibiotics. To this aim, independently again, we add the coupled Poisson processes $\{\tilde{T}_i^j\}_{i \geq 1}$ ($j=1,2$) defined above and we consider two independent families of independent Bernoulli variables $\{B_e^1\}_{e \in V\hat{T}}$ with parameter p_1 and $\{B_e^2\}_{e \in V\hat{T}}$ with parameter p_2 where $p_1 \leq p_2$ and $V\hat{T}$ is the set of vertical edges of \hat{T} . There is an analogous decimation procedure which allows to couple these two Bernoulli processes in such a way that if $B_e^1 = 1$ then $B_e^2 = 1$. At each time T_i^1 (resp. T_i^2) we consider all the white vertical edges intersecting the horizontal plane with time coordinate T_i^1 (resp. T_i^2); for all such edges e , if $B_e^j = 1$ we switch to black in the corresponding j model (hence the whole subtree is black), otherwise nothing happens.

In each model, at any time $t \geq 0$, let N_t (resp. R_t) be the number of white (resp. red) vertical edges which intersects the horizontal plane with time coordinate t ; $\{(N_t, R_t)\}_{t \geq 0}$ is the formal definition of the process. It is clear that the white/red edges in the second model have the same color in the first one, hence the non-black portion of the tree in the second model is a subtree of the non-black portion of the tree in the first model. In this construction, the event of survival is the collection of all the trees which have at least a red/white branch intersecting the horizontal plane t for every $t > 0$. This implies easily that the probability of survival of a model (β_1, p_1) is larger or equal than the probability of survival of a model (β_2, p_2) (where $\beta_1 \geq \beta_2$ and $p_2 \geq p_1$). More precisely, we coupled the environments in such a way that the conditional (with respect to the environment) probabilities of survival of the model (β_1, p_1) are larger or equal than the conditional (with respect to the coupled environment) probabilities of survival of the model

(β_2, p_2) . In particular, for any fixed p , the probability of survival is nondecreasing with respect to β and, for any fixed β , is nonincreasing with respect to p .

If we define $\beta_c(p) := \inf \{\beta > 0 : \text{the process survives with positive probability}\}$ then $\beta_c(p) > 0$ (according to Theorem 4.2). Since the probability of survival is nondecreasing with respect to β , we have that for all $\beta > \beta_c(p)$ there is survival with positive probability and for all $\beta < \beta_c(p)$ we have almost sure extinction. Hence $\beta_c(p) = \sup \{\beta > 0 : \text{the process dies out almost surely}\}$, thus $\beta_c(p) < +\infty$ (according to Theorem 4.1). Clearly $\beta_c(p)$ is nondecreasing with respect to p and, according to Theorem 4.2, $\lim_{p \rightarrow 0} \beta_c(p) = 0$. \square

Theorem 4.3 does not cover the critical case $\beta = \beta_c(p)$. The exact computation of δ_β (see Remark A.2) is not trivial, unless $p=1$, where it is simply $\delta_\beta = \mathbb{E}_\beta[\log(\bar{r}(T))]$, or $a=0$, where it is $\delta_\beta = \mathbb{E}_\beta[\log((1-p)\hat{n}(T))]$. If one were able to prove that $\beta \rightarrow \delta_\beta$ is continuous from the left then, from Remark A.2, it would follow easily that $\delta_{\beta_c(p)} \leq 0$ and there would be a.s. extinction for almost every realization of the environment in the critical case. This is straightforward when $\mu_\beta \sim \exp(1/\beta)$ and $p=1$; on the other hand, proving this continuity when $p < 1$ does not seem to be an easy task.

References

- Athreya, K.B., Ney, P.E., 1972. Branching processes, Die Grundlehren der mathematischen Wissenschaften, vol. 196. Springer-Verlag.
- Athreya, K.B., Karlin, S., 1971. On branching processes with random environments I: extinction probabilities. *Ann. Math. Stat.* 42, 1499–1520.
- Bertacchi, D., Posta, G., Zucca, F., 2007. Ecological equilibrium for restrained random walks. *Ann. Appl. Probab.* 17 (4), 1117–1137.
- Bertacchi, D., Zucca, F., 2008. Critical behaviors and critical values of branching random walks on multigraphs. *J. Appl. Probab.* 45, 481–497.
- Bertacchi, D., Zucca, F., 2009a. Characterization of the critical values of branching random walks on weighted graphs through infinite-type branching processes. *J. Stat. Phys.* 134 (1), 53–65.
- Bertacchi, D., Zucca, F., 2009b. Approximating critical parameters of branching random walks. *J. Appl. Probab.* 46, 463–478.
- Bertacchi, D., Zucca, F., 2012. Recent Results on Branching Random Walks, Statistical Mechanics and Random Walks: Principles, Processes and Applications. Nova Science Publishers, pp. 289–340.
- Bigger, J.W., 1944. Treatment of staphylococcal infections with penicillin. *Lancet*, 497–500.
- Furstenberg, H., Kesten, H., 1960. Products of random matrices. *Ann. Math. Stat.* 31, 457–469.
- Gardner, A., West, S.A., Griffin, A.S., 2007. Is bacterial persistence a social trait? *PLoS ONE* 2, e752.
- Garet, O., Marchand, R., Schinazi, R.B., 2012. Bacterial persistence: a winning strategy? *Markov Process. Relat. Fields* 18 (4), 639–650.
- Harris, T.E., 1963. The Theory of Branching Processes. Springer-Verlag, Berlin.
- Kingman, J.F.C., 1973. Subadditive ergodic theory. *Ann. Probab.* 1, 883–909.
- Kussell, E., Kishony, R., Balaban, N.Q., Leibler, S., 2005. Bacterial persistence: a model of survival in changing environments. *Genetics* 169, 1807–1814.
- Levin, B.R., 2004. Noninherited resistance to antibiotics. *Science* 305, 1578–1579.
- Levin, B.R., Rozen, D.E., 2006. Non-inherited antibiotic resistance. *Nat. Rev. Microbiol.* 4, 556–562.
- Liggett, T.M., 1985. Interacting Particle Systems, Grundlehren der Mathematischen Wissenschaften, vol. 276. Springer-Verlag, New York.
- Liggett, T.M., 1999. Stochastic Interacting Systems: Contact, Voter and Exclusion Processes, Grundlehren der Mathematischen Wissenschaften, vol. 324. Springer-Verlag, Berlin.
- Smith, W.L., Wilkinson, W.E., 1969. On branching processes in random environments. *Ann. Math. Stat.* 40, 814–827.
- Steele, J.M., 1989. Kingman's subadditive ergodic theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* 25 (1), 93–98.
- Tanny, D., 1981. On multitype branching processes in a random environment. *Adv. Appl. Probab.* 13, 464–497.
- Zucca, F., 2011. Survival, extinction and approximation of discrete-time branching random walks. *J. Stat. Phys.* 142 (4), 726–753.