Mathematical models of within-host and population-level infection dynamics

Jamie Paterson

Department of Applied Mathematics

University of Leeds

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

August 2024

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below.

The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Joint publications

Almost all of the work in Chapter 3 has been refereed and published, as follows:

- **Paterson J**, López-García, M., Gillard, J., Laws, T.R., Lythe, G., Molina-París, C. (2021). Analysis of Single Bacterium Dynamics in a Stochastic Model of Toxin-Producing Bacteria. Performance Engineering and Stochastic Modeling. EPEW ASMTA 2021. Lecture Notes in Computer Science, vol 13104. Springer, Cham. https://doi.org/10.1007/978-3-030-91825-5_13

Additionally, almost all of the work in Section 4.2, within Chapter 4, has been refereed and published, as follows:

- Williams, B.*, **Paterson, J.***, Rawsthorne-Manning, H.J.* Jeffrey, P., Gillard, J., Laws, T.R., López-García, M. Quantifying in vitro B. anthracis growth and PA production and decay: a mathematical modelling approach. npj Syst Biol Appl 10, 33 (2024). (*: first co-authorship) https://doi.org/10.1038/s41540-024-00357-1

Acknowledgements

This research was funded jointly by the Engineering and Physical Science, Research Council (EPSRC) and the Defence Science and Technology Laboratory (Dstl) through a CASE studentship (project reference 2274495). My thanks to everyone involved in the funding, without whom the research would not have been possible.

I would like to thank my supervisors for all of their advice, encouragement and patience throughout this thesis; Grant Lythe, Carmen Molina-París, Bevelynn Williams and Martín López-García. I would like to particularly thank Bevelynn and Martín for their support over the last few years, my appreciation for their understanding and kindness as I struggled with my physical and mental health cannot be understated, and the completion of the thesis would not have been possible without it.

Next, I would like to thank Joseph Gillard and Thomas Laws from Dstl for all of their guidance and support. Their perspective in discussions always provided me with a better understanding and clearer picture of the work to be done, and for that I am truly grateful. I would also like to thank all of those that have shared experimental data in the course of this research.

I would also like to thank the current and past members of the Mathematical Biology and Medicine group, who I have enjoyed working with throughout the past five years. The University of Leeds Student Counselling and Wellbeing service also has my thanks for their support when I needed it.

Finally, my thanks go to my friends and family for all their support and encouragement. In particular, I would like to give special mention my parents, their belief in me has been the basis of everything I've achieved,

and to Alex, who was always there with love, support and Snowball-related nonsense when it was needed.

Abstract

In this thesis, mathematical models for the *in vitro* and *in vivo* dynamics of the bacterium *Bacillus anthracis*, the causative agent of the disease anthrax, are considered in Chapters 3 and 4. Additionally, the dynamics of infection transmission for a pathogenic wild-type strain of a given virus and defective interfering strain of the same virus within a closed population are also considered in Chapter 5.

In Chapter 3, a two-compartment stochastic model is presented which describes the behaviour of toxin-producing bacteria and the corresponding dynamics are analysed. Using a continuous-time Markov chain to describe the model allows several summary statistics to be considered, such as the number of toxin molecules produced during the lifespan of a bacterium. This two-compartment model is applied to *Bacillus anthracis* and antibiotic treatment. An attempt is made to quantify, for the first time, bacterial toxin production by making use of data from an *in vitro* assay for a particular strain of *B. anthracis*.

Then, in Chapter 4, the stochastic analog of a previously published model of within-host anthrax infection is considered, allowing the computation of the dose-response probabilities of this model. Furthermore, we then propose a single model, in terms of delay differential equations, to explain the *in vitro* dynamics of published experimental data of two strains of *B. anthracis*, making use of a Bayesian approach for parameter calibration.

Within the last chapter of this thesis, Chapter 5, a compartmental epidemic model of viral infection is introduced. This model explores the protection afforded by the presence of a strain of virus composed of defective interfering particles (DIPs) on an outbreak of the wild-type virus in a closed population. A number of summary statistics are introduced and described for this model,

followed by an investigation into their distributions and expectations for a range of parameter regimes.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this thesis, two types of infection shall be considered. Chapters 3 and 4 are motivated by the dynamics of anthrax infection, a disease caused by toxin-producing bacteria. Chapter 5 considers the dynamics of a small, closed population being exposed to a virus that exists in two strains, a pathogenic wild-type strain and a defective strain (composed of defective interfering particles (DIPs)) that offers an individual some protection from the wild-type strain.

1.1 Biological introduction

In the immune system, there exists multiple mechanisms that protect the individual from infection via exposure to a pathogen. The first such line of defence is made up of physical barriers, such as the skin and mucous membranes. However, if a pathogen manages to enter the body it will be confronted with an innate immune response, involving recruitment and activation of a wide range of cells to attempt to remove the pathogen. One such example of the innate immune response will be discussed in Chapter 4 as part of the model by Day *et al.* (2011). If the innate immune response is not sufficient to clear the infection, there is another level of response known as the adaptive immune response, which provides a specific response for the pathogen that has infected the body (Sompayrac (2022)).

1.1.1 Anthrax

Anthrax is an infectious disease with a high mortality rate in the absence of treatment. It is caused by the infection of a host with the gram-positive bacteria, *Bacillus anthracis*. These bacteria are rod-shaped and usually form chains of bacterial cells as they multiply in a suitable environment. However, if *Bacillus anthracis* finds itself within an adverse environment it forms a dormant spore which can remain viable in soil, air and water with a half-life of around 100 years (Goel (2015)). These spores wait until favourable conditions are detected, usually within a host where the bacteria can sustain itself, and begin to germinate into vegetative bacteria. The bacteria then replicate within the host both intracellularly and extracellularly.

There are three main routes in which *Bacillus anthracis* usually enters a human host. These routes will lead to different clinical forms of the disease, known as cutaneous, gastrointestinal and inhalational anthrax. Cutaneous anthrax infection occurs when *Bacillus anthracis* enters the body via an entry in the skin, such as a cut. This is by far the most common type of anthrax, making up approximately 95% of cases (Chambers *et al.* (2018)), but has the best prognosis as it can usually be resolved positively with timely treatment and only approximately 20% of untreated cases are fatal. Cutaneous anthrax often results in black lesions on the infection site which gives the disease its name, taken from anthrakis, a Greek word for coal. Gastrointestinal anthrax generally occurs when a host has ingested meat from an animal that was infected with the bacteria; this has overall fatality rates of between 25% and 60%. Inhalational anthrax occurs when a host breathes in spores of *Bacillus anthracis*. This thesis will exclusively consider anthrax infections that occur inhalationally, for a few reasons. Firstly, this type of infection is extremely dangerous and is usually fatal if not rapidly detected and treated (Cote *et al.* (2011)). Even for cases where treatment is administered, the fatality rates can be high if treatment is not delivered early enough after exposure and specifically before symptoms onset. This is because the symptoms of anthrax are caused by toxins produced by bacteria within the host and as such antibiotic treatments will not affect these if they have already been produced. A complication of this need for early and effective treatment is that the initial symptoms of inhalational anthrax are extremely similar to those of flu, so early diagnoses are often difficult (Goel (2015)). Another reason that this thesis focuses on inhalational anthrax is that *Bacillus anthracis* is viewed as potential biological terrorism

threat as spores can be produced, preserved and then deliberately distributed via release in the air, which could have a significant impact on both individuals infected with the bacteria and economies of affected areas due to the high cost of decontamination. The first confirmed case of anthrax as a bioterror agent occurred in the United States in 2001, in which envelopes containing *Bacillus anthracis* spores were sent to members of the media and politicians in the aftermath of the 2001 9/11 attacks. This led to 22 confirmed cases of anthrax, 5 of which resulted in death and a cost of $320 million dollars in decontamination costs (Jernigan *et al.* (2002), Schmitt & Zacchia (2012)).

1.1.1.1 Pathogenesis of inhalational anthrax

An inhalational anthrax infection *Bacillus anthracis* enters the host by being breathed into the alveolar region of the lungs as spores. It is generally assumed that these spores have no effect on the lung region itself (Hodges *et al.* (1965)). Within the lungs, these spores are ingested by alveolar phagocytes attracted by the innate immune response of the host. Once the phagocytes have ingested the spores, they migrate to the lymph nodes in the mediastinum; however, within the phagocytes, the spores germinate and begin producing new bacteria. At this stage the phagocytes can either kill all bacteria and recover (Kang *et al.* (2005)), or enough bacteria will be produced that the phagocytes rupture and the bacteria is released extracellularly into the lymph node region. Some models such as the one proposed by Day *et al.* (2011), discussed in Chapter 4 of this thesis, assume that in all cases the phagocytes are overwhelmed and release the bacteria. Once *Bacillus anthracis* is exposed to the lymph node region via this rupturing process it continues to multiply, which can cause oedema and haemorrhage of the mediastinal lymph nodes. This causes symptoms in the host to begin to show, typically respiratory problems due to fluid in the pleural cavity. As this bacteria is now extracellular it is also possible for it to spread through the body via the bloodstream.

1.1.1.2 Anthrax toxins

B. anthracis produces virulence factors, which are, in this case, additional molecules produced by the bacteria that can add to its effectiveness in propagating the disease by inhibiting the immune system, entering host cells and then damaging or killing the host cells. The genes in control of encoding these virulence factors are found on two plasmids

in the cytosol of the bacteria, pXO1 and pXO2. pXO1 carries genes that encode the production of three proteins which are commonly known as the anthrax toxin components. These proteins are the protective antigen (PA), lethal factor (LF) and oedema factor (EF). LF and EF are individually not toxic as they lack a mechanism to enter cells in isolation, however in combination with PA they form toxins that can enter cells. PA and LF combine to form lethal toxin, which causes cell death. PA and EF combine to form edema toxin which has a suppressive effect on the immune system by preventing bacteria to be phagocytosed by neutrophils (Banks *et al.* (2005)). PA molecules bind to receptors on host cells, undergo cleaving and form a heptamer. This heptamer provides suitable conditions for the LF and EF proteins to bind to, forming a complex. Once formed, the complex is then taken into the cell via endocytosis where the toxins are released into the host cell. This process is shown in Figure 1.1. The plasmid pXO2 is responsible for encoding an anti-phagocytic capsule, which is a layer around the outside of the cell wall which covers molecules on the bacteria's cell surface that would cause an immune response and as such, allows the bacteria to continue replicating and producing toxins without being phagocytosed (Sharma *et al.* (2020)). The capsule produced by pXO2 is crucial in the pathogenesis of anthrax and without it, the host is much more capable of resisting the disease. Due to this, strains of *Bacillus anthracis* that lack the pXO2 plasmid are used in vaccinations around the world. The UK and US use the Sterne strain of *Bacillus anthracis* to produce their Anthrax Vaccine Precipitated (AVP) and Biothrax (formerly known as Anthrax Vaccine Absorbed) vaccines respectively. In China the vaccine used contains the A16R strain, which also produces anthrax toxins via plasmid pXO1 but cannot encode the capsule. Zai *et al.* (2016) carried out experiments using both the Sterne and A16R strain; we will discuss the results of Zai *et al.* (2016) in more detail in Chapter 4.

1.1.2 Viruses and defective interfering particles

Viruses are infectious agents that cannot reproduce unless contained within a cell, but once within a cell a virus can use the host cell machinery to replicate. The extracellular virus particle, before entry into a host cell, is called a virion. Virions consist of a genome (which can be DNA or RNA), surrounded by a protein coating called a capsid. Together, the capsid and the nucleic acid core are called the nucleocapsid. Some viruses have their

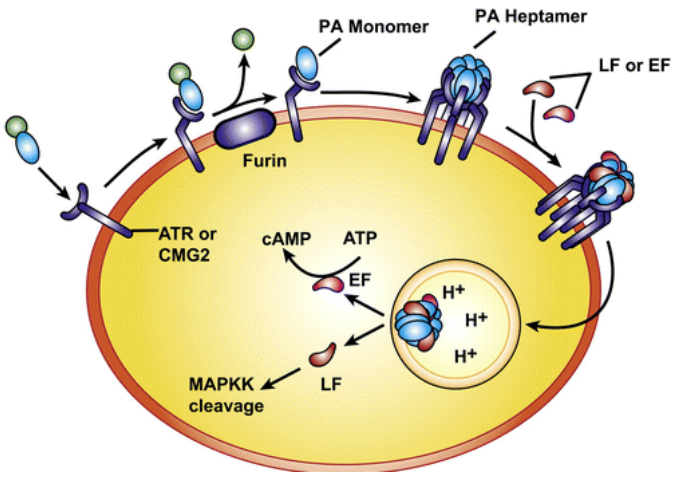

Figure 1.1: The process of PA molecules binding to a receptor, undergoing furin cleaving and forming a heptamer. This heptamer then binds LF and/or EF proteins (Hardenbrook *et al.* (2020)) to form toxin complexes and undergoes endocytosis to enter the cell. Figure taken from Sweeney *et al.* (2011).

nucleocapsid surrounded by a lipid envelope. The function of the capsid or envelope is to protect the viral genome whilst it is extracellular and allow it to enter a susceptible cell (Burrell *et al.* (2017), Strauss & Strauss (2008)). The proteins encoded within viral genomes can be categorised into three major classes. The first class encoded are enzymes required for replication of the genome. The second class are proteins that viruses must encode for the assembly of progeny virions. The third and final major class of viral protein, describes the case where viruses encode proteins that provide some protection from the defense mechanisms of the host. For example, this type of protein may interfere with the host's immune response. Generally, in a viral infection, the virion will attach to a host cell and become internalised into it; the viral nucleocapsid is then released into the cytoplasm of the cell. Depending on the type of virus, the viral genome replication process can begin in the cytoplasm; whereas some types of virus need to enter the nucleus for replication to begin (Strauss & Strauss (2008)).

During this replication process viral particles can be produced that lack some components of the viral genome, making them defective. However, not all of these defective particles interfere with the virus. The incomplete particles which do interfere with the complete virus are known as defective interfering particles (DIPs) if they satisfy the following criteria (Reta (2017)):

- The particles must contain the same structural proteins as the standard (wild-type (WT)) virus they are derived from and are antigenically identical. This requires that the DIP is simply a copy of the standard virus with some part of the viral genome missing.
- The missing part of the viral genome in the DIP causes it to be defective; it cannot replicate on its own as it lacks the capacity to produce all necessary viral proteins.
- However, DIPs can replicate during co-infections with the WT virus, also called under these circumstances a 'helper' virus, which provides the missing protein(s) needed for replication.
- During co-infections (i.e. WT virus and DIPs infecting the same cell), DIPs interfere with the reproduction of their helper virus, which leads to there being more DIP progeny than 'helper' virus progeny.

As these DIPs cannot replicate in the absence of the WT virus an individual infected solely with the DIP will have the virus simply decay and they will revert to a healthy state after some time. Due to the fourth criterion in the definition of a DIP, DIPs will have a suppressive effect on a disease if co-infection occurs with the WT strain. This is explained by the DIPs competing for intracellular resources with the WT particles, possibly reducing the total WT viral load within the host, and therefore they have the potential to mitigate the propagation of the disease (Reta (2017), Marriott & Dimmock (2010)). Therefore a DIP strain being present within a population will engender some level of protection from a WT strain; these dynamics will be investigated in Chapter 5. Within this chapter no specific virus is considered; we consider a generic virus with a WT and DIP strain as DIPs can be found for nearly all virus families (Rezelj *et al.* (2018)).

Chapter 2

Mathematical background

This section introduces the mathematical definitions and methods that will be used in the later chapters of this thesis are presented.

2.1 Probability theory

In this section, some of the core concepts from probability theory are introduced and defined, as some of the work in analysing stochastic processes will be underpinned by these definitions. These definitions can also be found in Chapter 1 of Allen (2010).

2.1.1 Conditional probability

In some cases in this thesis, it will be necessary to consider the effect that one event has upon another; this is related to the definition of conditional probability, which is given below.

Definition 2.1.1. Let A and B be events on a sample space, S, with $\mathbb{P}(B) > 0$. The **conditional probability** that event A occurs given that event B has occurred is given by the formula

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

It may be that knowing that event B occurs has no effect on the probability of event A; in that case events A and B are said to be **independent**. If that is the case then $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, and thus $\mathbb{P}(A|B) = \mathbb{P}(A)$.

2.1.2 Random variables

A random variable is an unknown quantity, the value of which can be determined by the outcome of some random event. Assume one such random variable, X. X is a real-valued function defined on the relevant sample space, Ω:

$$X : \Omega \to \mathbb{R} = (-\infty, \infty).$$

The set of values that X can take is known as the **state space** or **support** of X. If X has a finite or countably infinite state space then it is known as a **discrete** random variable. However, if X has a continuous interval for a state space then X is a **continuous** random variable.

Definition 2.1.2. Suppose X is a real-valued random variable. The **cumulative distribution function (cdf)** of X is the function $F_X : \mathbb{R} \to [0, 1]$, defined by

$$F_X(x) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R}.$$

Definition 2.1.3. Suppose X is a discrete random variable, with support S. The **probability mass function (pmf)** of X is defined by

$$f_X(x) = \mathbb{P}(X = x), \quad x \in S.$$

Definition 2.1.4. Suppose X is a continuous random variable, and there exists a

non-negative integrable function $f_X : \mathbb{R} \to [0, \infty)$ such that,

$$\mathbb{P}(a \leq x \leq b) = \int_a^b f_X(x)dx, \quad a < b.$$

Then the function f_X is called the **probability density function (pdf)** of X.

Definition 2.1.5. If X is a continuous random variable with pdf $f_X(x)$ and support S, then the **expectation** of X is defined as

$$\mathbb{E}[X] = \int_S x f_X(x)dx.$$

If X is a discrete random variable with pmf f_X, and support S, then the **expectation** or **mean** of X is defined as

$$\mu_X = \mathbb{E}[X] = \sum_{x \in S} x f_X(x).$$

The definition of the expectation of a random variable can be extended to the expectation of a function of a random variable, $g(X)$. If X is a continuous random variable, with support S, then the expectation of $g(X)$ is

$$\mathbb{E}[g(X)] = \int_S g(x) f_X(x)dx.$$

If X is a discrete random variable, with support S, then the expectation of $g(X)$ is

$$\mathbb{E}[g(X)] = \sum_{x \in S} g(x) f_X(x).$$

There are also other summary statistics of a random variable, X, that will be needed for analysis later in the thesis; namely the variance and moments of X. These statistics can all be defined in terms of the expectation of X.

The **variance** of X, denoted by σ_X^2 or $\text{Var}(X)$ is given by

$$\text{Var}(X) = \mathbb{E}\left[(X - \mu_X)^2\right] = \mathbb{E}[X^2] - \mu_X^2.$$

The **standard deviation** of X is the square root of the variance,

$$\sigma = \sqrt{\text{Var}(X)}.$$

The n**th moment of** X **about the point** a is

$$\mathbb{E}\left[(X - a)^n\right].$$

It will also be necessary to make use of **probability generating functions (pgfs)** to analyse stochastic processes. These are a specific sub-section of a wider class of **generating functions**. Generating functions are defined for a discrete random variable, X, and their definitions are extended to continuous random variables.

Definition 2.1.6. Suppose X is a discrete random variable, with state space $S = \{0, 1, 2, \dots\}$. The **probability generating function (pgf)** of X is defined by

$$\mathcal{P}_X(z) = \mathbb{E}[z^X] = \sum_{x=0}^{\infty} \mathbb{P}(X = x) z^x,$$

for some $z \in \mathbb{C}, |z| \leq 1$.

The pgf generates the probabilities associated with the distribution by differentiating and setting $z = 0$, to obtain,

$$\mathbb{P}(X = k) = \frac{1}{k!} \frac{d^{(k)}}{dz^{(k)}} \mathcal{P}_X(z) \bigg|_{z=0}, \quad k \geq 0.$$

The mean and variance can also be found using the pgf:

$$\mathcal{P}'_X(z)|_{z=1} = \sum_{x=0}^{\infty} x \mathbb{P}(X = x) = \mathbb{E}[X].$$

$$\mathcal{P}''_X(1) + \mathcal{P}'_X(1) - |\mathcal{P}'_X(1)|^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \text{Var}(X).$$

2.1.3 Discrete probability distributions

There are several well-known discrete probability distributions, some of which will be used throughout the course of this thesis. Their pmfs are given in this section.

2.1.3.1 Bernoulli distribution

A random variable that follows the Bernoulli distribution (a Bernoulli random variable) describes a single trial where the probability of success is p and the probability of failure is $1 - p$, where $0 < p < 1$. Its pmf is

$$f_X(x) = \begin{cases} p, & x = 1, \\ 1 - p, & x = 0, \\ 0, & \text{otherwise.} \end{cases}$$

2.1.3.2 Binomial distribution

A random variable that follows the binomial distribution (a binomial random variable) is obtained by taking the sum of n independent random variables that follow the Bernoulli distribution. In effect, a binomial random variable can be thought of as the number of successes, x, in n independent Bernoulli trials with a probability of success p in each. Its pmf is

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, \ldots, n, \\ 0, & \text{otherwise,} \end{cases}$$

where n is a positive integer.

2.1.3.3 Geometric distribution

A random variable that follows the geometric distribution (a geometric random variable), can be interpreted as the number of failed independent Bernoulli trials before the first success, where p is the probability of success. This distribution is biologically useful as it can be used to describe the behaviour of bacteria within a host; such as when counting

the number of times it will proliferate (a failure in this case) before dying (a success), as we investigate in Section 3.3.3. Its pmf is

$$f_X(x) = p(1-p)^x, \quad x = 0, 1, \dots .$$

2.1.4 The exponential distribution

It is not always possible to use discrete distributions to represent the behaviours of biological agents; for example, the times involved in biological processes related to an infection will lead to the study of continuous distributions. One such continuous distribution, which will be useful for analysing continuous-time Markov chain models, is the exponential distribution. In such models, the exponential distribution is associated with the time between consecutive events.

Definition 2.1.7. The pdf of a random variable that follows the exponential distribution, an **exponentially distributed random variable**, $X \sim Exp(\lambda)$, is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

for some rate parameter $\lambda > 0$. The cdf of an exponentially distributed random variable is given by

$$F_X(x) = 1 - e^{-\lambda x}, \quad x \geq 0.$$

The mean and variance of an exponentially distributed random variable are

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

The exponential distribution has some key properties that make it extremely useful for analysing stochastic processes. The first of which is the fact that it is memoryless.

2.1.4.1 The memoryless property

Unique to the exponential distribution among continuous probability distributions, the memoryless property states that the probability of an event occurring depends only on the current state of the system, independent of what has come before. Namely, let us consider an exponential random variable, X, with rate parameter, λ, which represents the time until some event occurs. If one knows the event has not occurred up until time t_1, the probability that it does not happen during $[t_1, t_2]$ is just the probability of the event not occurring in a time period of length t_2, (e.g., $[0, t_2]$). Mathematically:

$$\mathbb{P}\left(X \geq t_1 + t_2 \mid X \geq t_1\right) = \mathbb{P}(X \geq t_2), \quad \forall t_1, t_2 > 0.$$

This can be shown using the definition of conditional probability and the formula for the cdf of an exponential random variable:

$$\begin{aligned}
\mathbb{P}(X \geq t_1 + t_2 \mid X \geq t_1) &= \frac{\mathbb{P}\left(X \geq t_1 + t_2 \cap X \geq t_1\right)}{\mathbb{P}(X \geq t_1)} \\
&= \frac{\mathbb{P}\left(X \geq t_1 + t_2\right)}{\mathbb{P}(X \geq t_1)} \\
&= \frac{1 - F_X(t_1 + t_2)}{1 - F_X(t_1)} \\
&= \frac{e^{-\lambda(t_1+t_2)}}{e^{-\lambda t_1}} \\
&= e^{-\lambda t_2} \\
&= \mathbb{P}(X \geq t_2).
\end{aligned}$$

2.1.4.2 Competition between exponential random variables

The behaviour between two competing exponential random variables is also well-defined and provides the basis for the Gillespie algorithm, which will be discussed in more detail in Section 2.2.2.

Let us consider we are waiting for two independent exponentially distributed events to occur. One event happens after $X \sim Exp(\lambda)$ and the other after $Y \sim Exp(\mu)$. The probability that event X occurs before event Y, $\mathbb{P}(X < Y)$, can be found using the definition of the pdf of a continuous random variable (Definition 2.1.4). As we are

interested in $\mathbb{P}(X < Y)$, this can be thought of as the scenario that $\mathbb{P}(0 \leq X < Y)$ and $\mathbb{P}(0 \leq Y \leq \infty)$. Using this in conjunction with the specific pdf of an exponentially distributed random variable (Definition 2.1.4) we obtain

$$\begin{aligned}
\mathbb{P}(X < Y) &= \int_0^\infty \left(\int_0^y f(x)dx \right) f(y)dy \\
&= \int_0^\infty \int_0^y \lambda e^{-\lambda x} dx \cdot \mu e^{-\mu y} dy \\
&= \lambda\mu \int_0^\infty \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^y e^{-\mu y} dy \\
&= \lambda\mu \left(-\frac{1}{\lambda} \right) \int_0^\infty \left(e^{-\lambda y} - 1 \right) e^{-\mu y} dy \\
&= -\mu \left(\int_0^\infty e^{-(\lambda+\mu)y} dy - \int_0^\infty e^{-\mu y} dy \right) \\
&= -\mu \left(\left[-\frac{1}{\lambda+\mu} e^{-(\lambda+\mu)y} \right]_0^\infty - \frac{1}{\mu} \right) \\
&= -\mu \left(\frac{1}{\lambda+\mu} - \frac{1}{\mu} \right) \\
&= -\frac{\mu}{\lambda+\mu} + 1 \\
&= \frac{\lambda}{\lambda+\mu}.
\end{aligned} \tag{2.1}$$

It can be shown similarly that $\mathbb{P}(Y < X) = \frac{\mu}{\lambda+\mu}$.

2.1.4.3 Time for an event to occur

One can also consider the time taken for any event to occur. If again, two independent exponentially distributed variables, $X \sim Exp(\lambda)$ and $Y \sim Exp(\mu)$ are defined but now we are solely interested in the time it takes for either event to occur, we can define $Z = \min\{X, Y\}$ as the time until either event X or Y occurs. It can be shown that Z is also an exponentially distributed random variable by considering its cdf, $\mathbb{P}(Z \leq z)$ (from Definition 2.1.2).

$$\mathbb{P}(Z \leq z) = 1 - \mathbb{P}(Z > z) = 1 - \mathbb{P}(\min\{X, Y\} > z) = 1 - \mathbb{P}(X > z, Y > z).$$

As events X and Y are independent we know that $\mathbb{P}(X > z, Y > z) = \mathbb{P}(X > z)\mathbb{P}(Y > z)$. This means we can use Definition 2.1.4 to show that

$$\begin{aligned}\mathbb{P}(Z \leq z) &= 1 - \mathbb{P}(X > z)\mathbb{P}(Y > z) = 1 - e^{-\lambda z}e^{-\mu z} \\ &= 1 - e^{-(\lambda+\mu)z}.\end{aligned}$$

Notice from Definition 2.1.4 that this is itself a cdf of an exponentially distributed random variable with rate parameter $\lambda + \mu$. We can therefore say that

$$Z = \min\{X, Y\} \sim Exp(\lambda + \mu). \tag{2.2}$$

Equations (2.1) and (2.2) are the basis of the Gillespie algorithm, which will be discussed in further detail in Section 2.2.2.

Let us now consider the case where it is known which event happens first and investigate the effect on the distribution of the time until an event occurs. Assume for now that it is known that event X occurs first, $X < Y$. This means the distribution of interest is $Z|X < Y$. Therefore the pdf of the random variable $Z|X < Y$ can be considered as the case that $\mathbb{P}(X < Y, Z > z)$.

$$\mathbb{P}(X < Y, Z > z) = \mathbb{P}(Y > X > z) = \int_z^\infty \mathbb{P}(Y > x)\lambda e^{-\lambda x} dx.$$

Using the fact that $\mathbb{P}(Y > x) = 1 - \mathbb{P}(Y \leq x)$ and the definition of the cdf of an exponential random variable (Definition 2.1.4), it can be obtained that

$$\begin{aligned}\mathbb{P}(X < Y, Z > z) &= \int_z^\infty e^{-\mu x}\lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{\lambda + \mu}\left[-e^{-(\lambda+\mu)x}\right]_z^\infty \\ &= \frac{\lambda}{\lambda + \mu}e^{-(\lambda+\mu)z}.\end{aligned}$$

Note that, $\frac{\lambda}{\lambda+\mu} = \mathbb{P}(X < Y)$ and $e^{-(\lambda+\mu)z} = \mathbb{P}(Z > z)$, showing that the probability of event X happening before event Y and time taken for one of the events to occur are independent. In other words, the time until the next event occurring is not impacted by which event actually happens. The same logic follows if it is known that event Y occurs

before event X.

2.2 Stochastic processes

Previous sections have covered the concept of a single random variable; for the biological questions this thesis aims to investigate it is often required to consider a random process that evolves through time. This leads to the definition of a **stochastic process**. The work within this section is mainly based upon Chapters 5 and 6 of Allen (2010) and Chapter 6 of Kulkarni (2016).

Definition 2.2.1. A **stochastic process** is a collection of random variables, $\{X(t), t \in T\}$, indexed by parameter t which takes values in the parameter set T. The random variables, $X(t)$, take values in the set S, called the state space of the stochastic process.

The parameter set T within this thesis will be a set of times, typically $[0, \infty)$, so that $X(t)$ represents the state (e.g. size) of a population at time t, making this a continuous-time stochastic process. Also note, that the random variables within the stochastic process can either represent one population $X(t)$, or multiple $\mathbf{X}(t) = (X_1(t), X_2(t), \dots)$, in either case following the same definition.

Within this thesis, the main stochastic processes under analysis will be **continuous-time Markov chains**. Continuous-time is defined as above for a stochastic process. **Markov** in this sense implies that the future of the system is dependant solely on the current state of the system, independent of the past. This is known as the **Markov property**. Note that the Markov property is facilitated by the memoryless property of the exponential distribution that will be used for inter-event times. The term **chain** implies that the random variables, $X(t)$, take values in a discrete state-space; this is necessary as within the thesis the random variables of concern $X(t)$ will typically represent population sizes.

Definition 2.2.2. Let $\mathcal{X} = \{X(t) : t \in [0, \infty)\}$ be a collection of discrete random variables with values in a finite or infinite state space S, (for example

$S = \{0, 1, 2, ..., N\}$ or $S = \{0, 1, 2, ...\}$). Then the stochastic process $\mathcal{X}$ is called a **continuous-time Markov chain (CTMC)** if it satisfies

$$\begin{aligned}\mathbb{P}\left(X(t_{n+1}) = i_{n+1} | X(t_0) = i_0, X(t_1) = i_1, ..., X(t_n) = i_n\right) \\ = \mathbb{P}\left(X(t_{n+1}) = i_{n+1} | X(t_n) = i_n\right),\end{aligned}$$

for any set of real numbers such that $0 \leq t_0 \leq t_1 \leq ... \leq t_{n+1}$ and $i_0, \ldots, i_n, i_{n+1} \in S$.

For every $t \in [0, \infty)$, $X(t)$ is a random variable with probability mass function $\{p_i(t), i \in S\}$, where

$$p_i(t) = \mathbb{P}(X(t) = i),$$

for any $i \in S$. The probability that the chain moves from state i at time s, to state j at time t, with $i, j \in S$ and $t, s \in [0, \infty)$ is defined as follows.

Definition 2.2.3. The **transition probability** for a CTMC, $\mathcal{X}$, to get from state $X(s) = i$ to state $X(t) = j$ at time t, with $s, t \in [0, \infty)$, $s < t$ and $i, j \in S$, is defined as

$$p_{ij}(s, t) = \mathbb{P}(X(t) = j | X(s) = i).$$

These probabilities are referred to as **homogeneous** transition probabilities if they only depend on the length of the time interval $t - s$ and not particular values of s and t. Hence,

$$p_{ij}(s, t) = p_{ij}(t - s) = \mathbb{P}(X(t) = j | X(s) = i) = \mathbb{P}(X(t - s) = j | X(0) = i).$$

The **transition matrix**, $\mathbf{P}(t)$, contains all transition probabilities from a state i to a state j and is defined as

$$\mathbf{P}(t) = (p_{ij}(t))_{i,j \in S},$$

where $t \in [0, \infty)$.

For general CTMCs, and for all CTMCs found within this thesis, the following property

holds

$$\sum_{j \in S} p_{ij}(t) = 1, \quad t \geq 0, \quad i \in S. \tag{2.3}$$

This holds as the CTMC is a closed process, so if the process is at state i at time 0, it will have to be at some state, $j \in S$, within the process at time t; therefore the summation of transition probabilities over all possible states it can transition to will be 1.

The transition probabilities $p_{ij}(t)$ of a CTMC can be used to define **transition rates** q_{ij}. The transition probabilities $p_{ij}(t)$ are assumed to be continuous and differentiable for $t \geq 0$ and at $t = 0$ where

$$p_{ij}(0) = 0, \forall i \neq j, \text{ and } p_{ii}(0) = 1.$$

The transition rate, q_{ij}, is defined as

$$q_{ij} = \lim_{\Delta t \to 0^+} \frac{p_{ij}(\Delta t) - p_{ij}(0)}{\Delta t} = \lim_{\Delta t \to 0^+} \frac{p_{ij}(\Delta t)}{\Delta t}, \quad \text{for } i \neq j, \tag{2.4}$$

$$q_{ii} = \lim_{\Delta t \to 0^+} \frac{p_{ii}(\Delta t) - p_{ii}(0)}{\Delta t} = \lim_{\Delta t \to 0^+} \frac{p_{ii}(\Delta t) - 1}{\Delta t}\,. \tag{2.5}$$

Note that by definition, $q_{ij} \geq 0$. It holds from Equation (2.3) that

$$1 - p_{ii}(\Delta t) = \sum_{j \in S, j \neq i} p_{ij}(\Delta t) = \sum_{j \in S, j \neq i} [q_{ij} \cdot \Delta t + o(\Delta t)],$$

which leads directly to the following expression for q_{ii},

$$\begin{aligned} q_{ii} &= -\lim_{\Delta t \to 0^+} \frac{\sum_{j \in S, j \neq i} q_{ij} \cdot \Delta t + o(\Delta t)}{\Delta t} \\ &= -\sum_{j \in S, j \neq i} q_{ij}, \end{aligned}$$

where $\sum_{j \in S, j \neq i} o(\Delta t) = o(\Delta t)$. Note that for any given $i \in S$,

$$\sum_{j \in S} q_{ij} = 0.$$

Let $\mathbf{P}(t)$ be the transition matrix of a CTMC and $\mathbf{I}$ be the identity matrix of the same size.

Then the **infinitesimal generator matrix** $\mathbf{Q} = (q_{ij})_{i,j\in S}$ is equal to

$$\mathbf{Q} = \lim_{\Delta t \to 0^+} \frac{\mathbf{P}(\Delta t) - \mathbf{I}}{\Delta t}. \tag{2.6}$$

Definition 2.2.4. The matrix of transition rates for a CTMC $\mathcal{X}$ with state space S, $\mathbf{Q} = (q_{ij})_{i,j\in S}$ defined from Equations (2.4), (2.5) and (2.6), is known as the **infinitesimal generator matrix**. For example for S ={0, 1, 2, . . . }, $\mathbf{Q}$ is equal to

$$\begin{aligned}\mathbf{Q} &= \begin{pmatrix} q_{00} & q_{01} & q_{02} & \cdots \\ q_{10} & q_{11} & q_{12} & \cdots \\ q_{20} & q_{21} & q_{22} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \\ &= \begin{pmatrix} -\sum_{j=1}^{\infty} q_{0j} & q_{01} & q_{02} & \cdots \\ q_{10} & -\sum_{j=0,j\neq 1}^{\infty} q_{1j} & q_{12} & \cdots \\ q_{20} & q_{21} & -\sum_{j=0,j\neq 2}^{\infty} q_{2j} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}.\end{aligned}$$

The transition probabilities will change over time; the forward Kolmogorov differential equations describe the rate at which they change.

Definition 2.2.5. The **forward Kolmogorov differential equations** are a set of equations that describe the rate of change of the transition probabilities

$$\frac{dp_{ij}(t)}{dt} = \sum_{k\in S} q_{kj} p_{ik}(t), \quad i, j \in S.$$

This can be written in matrix form

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{Q}\mathbf{P}(t).$$

This equation is also known as the master equation of the CTMC.

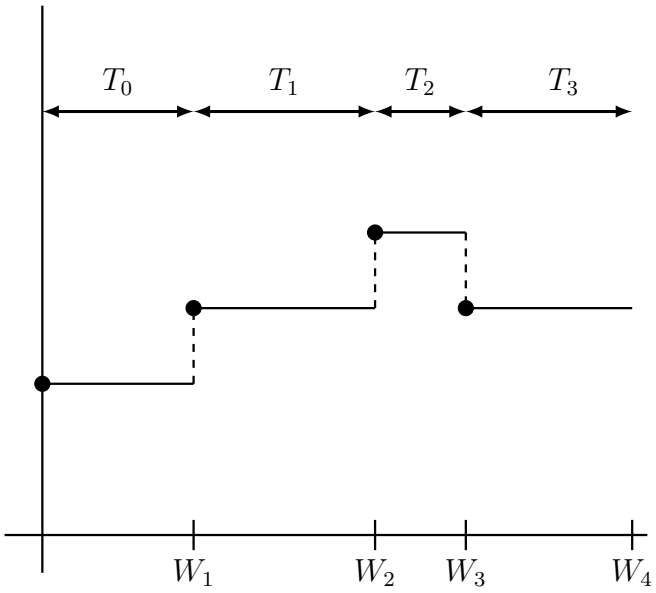

Figure 2.1: One possible realisation of $\mathcal{X}$ that shows the jump times, W_i, and the inter-event times T_i. Figure reproduced from Allen (2010).

It is also important when analysing stochastic processes to track the time taken between each step or jump of the process. This is known as the inter-event time.

Definition 2.2.6. For a given CTMC, $\mathcal{X} = \{X(t) : t \geq 0\}$, with state space S, denote the time of the nth event of $\mathcal{X}$ as W_n, with the assumption that $W_0 = 0$. The **inter-event time**, T_n, is the time between the nth event and the $(n+1)$th event of $\mathcal{X}$, or the time that the system stays in the state $X(W_n)$:

$$T_n = W_{n+1} - W_n.$$

Figure 2.1 shows one possible realisation of $\mathcal{X}$ and the corresponding inter-event times, T_n. In order to calculate such sample paths of continuous-time Markov chains, the distribution of T_n is needed. Let $X(W_n) = i$ be the state that the CTMC jumps to at the n-th jump. Assume that at jump n we have state i such that $X(W_n) = i$ where $i \in S$. Then the probability of leaving state i during $[W_n, W_n + \Delta t]$ is defined as $\alpha_i \Delta t + o(\Delta t)$ for some Δt, where α_i is the sum of the rates of all possible state changes out of state i,

$\alpha_i = \sum_{j \neq i} q_{ij}$. This can be written as

$$\sum_{j \in S, j \neq i} p_{ij}(\Delta t) = \alpha_i \Delta t + o(\Delta t).$$

Therefore the probability that no jump has occurred during time $[W_n, W_n + \Delta t]$, meaning that the probability of remaining in state i, is

$$p_{ii}(\Delta t) = 1 - \alpha_i \Delta t + o(\Delta t). \tag{2.7}$$

Let $G_n(t)$ be the probability that the process remains in state i for a time of at least $t \geq 0$, meaning that $T_n > t$. Therefore

$$G_n(t) = \mathbb{P}(T_n > t).$$

If $\alpha_i = 0$, then state i is called **absorbing**, a state where if the process enters, it will not leave. The process remains indefinitely in state i and $G_n(t) = \mathbb{P}(T_n > t) = 1$ for all $t \geq 0$. Otherwise, for a sufficiently small Δt, one has

$$\begin{aligned} G_n(t + \Delta t) &= \mathbb{P}(T_n \geq t + \Delta t) \\ &= \mathbb{P}(T_n > t + \Delta t | T_n > t) \mathbb{P}(T_n > t) \\ &= p_{ii}(\Delta t) G_n(t) \\ &= G_n(t)(1 - \alpha_i \Delta t + o(\Delta t)), \end{aligned}$$

by applying Equation (2.7) in the final step. Subtracting $G_n(t)$, dividing by Δt and then letting $\Delta t \to 0$ leads to a differential equation for $G_n(t)$,

$$\frac{d}{dt} G_n(t) = -\alpha_i G_n(t),$$

with initial condition $G_n(0) = 1$, as the first jump is guaranteed to occur after time $t = 0$. This can be solved using separation of variables to obtain

$$G_n(t) = e^{-\alpha_i t}, \quad t \geq 0.$$

This means that $\mathbb{P}(T_n \leq t)$, the cumulative distribution function of T_n, is equal to

$$\mathbb{P}(T_n \leq t) = 1 - G_n(t) = 1 - e^{-\alpha_i t}, t \geq 0.$$

This corresponds to the cdf of an exponentially distributed random variable with parameter α_i, meaning that the inter-event time, T_n, is an exponentially distributed random variable, $T_n \sim Exp(\alpha_i)$.

Theorem 1.2.1 Let $\mathcal{X}$ be a CTMC with state space S, $\mathcal{X} = \{X(t) : t \in [0, \infty)\}$, and with a transition matrix $\mathbf{P}(t) = (p_{ij}(t))_{i,j \in S}$ such that

$$\sum_{j \in S, j \neq i} p_{ij}(\Delta t) = \alpha_i \Delta t + o(\Delta t),$$

and

$$p_{ii}(\Delta t) = 1 - \alpha_i \Delta t - o(\Delta t),$$

for all $i \in S$ and Δt sufficiently small. Let us define W_n as the time of the nth jump resulting in a change in the state of $\mathcal{X}$. We can define the inter-event time as $T_n = W_{n+1} - W_n$. Given that $X(W_n) = i$, T_n is an exponential random variable with parameter α_i. The probability density function (pdf) for T_n is $f_{T_n}(t) = \alpha_i e^{-\alpha_i t}$ and its cumulative distribution function (cdf) is $F_{T_n}(t) = 1 - e^{-\alpha_i t}$ for $t \geq 0$, where $\alpha_i = \sum_{j \in S, j \neq i} q_{ij} = -q_{ii}$. Therefore the mean and variance of T_n are

$$E(T_n) = \frac{1}{\alpha_i}, \quad Var(T_n) = \frac{1}{\alpha_i^2}, \quad \alpha_i > 0.$$

2.2.1 The linear birth-and-death process

A well-known example of a CTMC widely used in mathematical biology is the birth-and-death process. A birth-and-death process is defined as a CTMC, $\mathcal{X} = \{X(t) : t \geq 0\}$, where $X(t)$ represents the size of some population at time t and has a state space $S = \{0, 1, \dots\}$. S could be bounded for a finite state space but here S is considered to be unbounded and therefore the state space is infinite. Generally a birth-and-death process

restricts the possible transitions from the state i to the adjacent states, the interpretation of which is a population which can either have a death (which occurs at rate μ_i) or a birth (which occurs at rate λ_i). For a linear birth-and-death process, the rates are defined as

$$\lambda_i = i\lambda, \quad \mu_i = i\mu,$$

for some $\lambda, \mu > 0$. Here, 0 is an absorbing state as both λ_0 and μ_0 are equal to 0, representing population extinction. The transition probabilities for this CTMC can be defined for sufficiently small Δt:

$$\begin{aligned} p_{ij}(\Delta t) &= \mathbb{P}(X(t+\Delta t) = j \mid X(t) = i) \\ &= \begin{cases} i\lambda\Delta t + o(\Delta t), & \text{if } j = i+1, \\ i\mu\Delta t + o(\Delta t), & \text{if } j = i-1, \\ 1 - i(\lambda+\mu)\Delta t + o(\Delta t), & \text{if } j = i, \\ o(\Delta t)^2, & \text{otherwise.} \end{cases} \end{aligned}$$

The transition rates are given by

$$q_{ij} = \begin{cases} i\lambda, & \text{if } j = i+1, \\ i\mu, & \text{if } j = i-1, \\ -i(\lambda+\mu), & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

leading to the infinitesimal generator matrix

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & 2\mu & -2(\lambda+\mu) & 2\lambda & \dots \\ \vdots & \vdots & \vdots & & \end{pmatrix}.$$

2.2.2 The Gillespie algorithm

This section is based upon the work by Gillespie (1977).

The Gillespie algorithm allows one to get stochastic realisations of a CTMC, $\mathcal{X}$, in terms of simulations, which is useful when analytical work (e.g. computing probabilities $p_{ij}(t)$

in an exact way) cannot be carried out. The algorithm works by exploiting the fact that the inter-event time is exponentially distributed and independent of which state the process changes to. Therefore the algorithm samples an inter-event time from an exponential distribution, and then independently randomly selects a transition event from all possible transitions from the current state according to their probabilities; then it updates the state of the CTMC, the simulation time and then repeats the process. Sampling the inter-event time from an exponentially distribution makes use of the following theorem, which is given in Chapter 5 of Allen (2010).

Theorem 1.2.2 Let U be a uniform random variable defined on $[0, 1]$ and T be a continuous random variable defined on $[0, \infty)$. Then $T = F^{-1}(U)$, where F is the cumulative distribution function (cdf) of the random variable T.

This theorem can be proved as follows: Since $\mathbb{P}(T \leq t) = F(t)$, from the definition of the cdf, we want to show that $\mathbb{P}(F^{-1}(U) \leq t) = F(t)$. Note, F^{-1} exists as $F : [0, \infty) \to [0, 1)$ is strictly increasing. Also, as U is a uniform random variable on $[0, 1]$, it is known that $\mathbb{P}(U \leq x) = x$ for $x \in [0, 1]$. Therefore, for $t \in [0, \infty)$:

$$\begin{aligned}
\mathbb{P}(F^{-1}(U) \leq t) &= \mathbb{P}(F\left(F^{-1}(U)\right) \leq F(t)) \quad \text{(as } F \text{ is strictly increasing)} \\
&= \mathbb{P}(U \leq F(t)) \\
&= F(t).
\end{aligned}$$

The Gillespie algorithm samples the inter-event time by making use of this transformation at each jump. Here an example is given to show how it works for the general inter-event time, T, in a CTMC currently at state i. As defined in Theorem 1.2.1, T_n is an exponential random variable and has cdf $F_{T_n}(t) = 1 - e^{-\alpha_i t}$. Therefore, by Theorem 1.2.2:

$$\begin{aligned}
T &= F^{-1}(U[0, 1]) \\
&= -\frac{1}{\alpha_i}\log(1 - U[0, 1]) \\
&= -\frac{1}{\alpha_i}\log(U[0, 1]),
\end{aligned}$$

where we have used in the last step the fact that $1 - U[0, 1]$ is itself a random uniform variable on $[0, 1]$.

Algorithm 2.1 shows how to obtain stochastic simulations using the Gillespie algorithm.

An explanation of each step of the Gillespie algorithm is as follows:

- Input initial conditions.
  These are $t = 0$ and $X(0) = i_0$ where i_0 is the initial state of the CTMC.
- Evaluate the rates.
  For each possible state $j \in S$, the jump from state i to state j occurs with rate q_{ij}. The total propensity of the system at state i is given by the formula $\alpha_i = \sum_{j \in S, j \neq i} q_{ij}$. Some of these q_{ij} may be equal to 0 if state j is unreachable from state i in a single jump. For example, in a birth-and-death process if $i = 3$ and $j = 5$, $q_{ij} = 0$ as only one-jump transitions between adjacent states are allowed.
- Choose the inter-event time.
  The inter-event time, Δt, is sampled from an exponential distribution with rate α_i. To do this, one can sample a random number, u_1, from the uniform distribution on the interval (0,1) and compute

  $$\Delta t = -\frac{\log(u_1)}{\alpha_i}.$$

  This ensures that the inter-event is being sampled from $Exp(\alpha_i)$, as according to Theorem 2.2.2.
- Choose a reaction to occur.
  Assume there is a list of n possible reactions $\{1, \ldots, n\}$ with corresponding rates $\{r_1, \ldots, r_n\}$, where if reaction k represents the jump $i \to j$ then $r_k = q_{ij}$. Each reaction k has a chance of being selected that is equal to its rate as a proportion of the total sum of all rates; this utilises the property discussed in Section 2.1.4.3 (namely Equation (2.1)). The way this is carried out is to create sub-intervals θ_k of $[0, 1]$, for $k \in \{1, \ldots, n\}$, where

  $$\theta_k = \left(\frac{\sum_{j=0}^{k-1} r_j}{\alpha_i}, \frac{\sum_{j=0}^{k} r_j}{\alpha_i} \right],$$

with $r_0 = 0$, and i is the current state of the process. Then a random number, u_2, is sampled from the uniform distribution in (0,1). u_2 will be in one of these sub-intervals θ_k and therefore, reaction k is chosen to occur. Reaction k moves the process to state j.

- Update the system.

$$t \to t + \Delta t,$$

$$X(t) \to j.$$

- If certain conditions are met (e.g the CTMC reaches a predetermined state or a certain amount of time has passed) the algorithm stops, if not return to the second step and repeat.

Algorithm 2.1: The Gillespie algorithm

Choose t_{max}, the maximum time we want to simulate to.
Set the initial time to be $t = t_0$.
Set the initial state of the process to be $\mathbf{x} = \mathbf{x}(t_0) = \mathbf{x}_0$.
Create an array that will retain all stages of the stochastic process, $L = [[\mathbf{x}_0, t_0]]$.
While $t < t_{max}$:
- Assign a number from $1, \ldots, n$ to each possible state that the process can jump to, $j_1, \ldots, j_n$.
- Compute the propensity function $\alpha(\mathbf{x}) = \sum_{i=1}^{n} q_{\mathbf{x},j_i}$, where n is the number of possible one-jump transitions from state $\mathbf{x}$.
- Sample a random number from the uniform distribution over the unit interval, $u_1 \sim (0, 1)$, and set $\Delta t = -\frac{\log(u_1)}{\alpha(\mathbf{x})}$.
- Sample another uniform random variable over the unit interval, $u_2 \sim (0, 1)$. The state will move to state j_k such that $\sum_{i=1}^{k-1} \frac{q_{\mathbf{x},j_i}}{\alpha(\mathbf{x})} \leq u_2 \leq \sum_{i=1}^{k} \frac{q_{\mathbf{x},j_i}}{\alpha(\mathbf{x})}$.
- Update the current state of the process to $\mathbf{x} = j_k$.
- Update the time to $t = t + \Delta t$.
- Add an entry to L with the new values for $\mathbf{x}$ and t.

Return L.

2.3 Approximate Bayesian Computation

A mathematical model representing a biological system would typically depend on some parameters, $\boldsymbol{\theta}$. This section deals with the problem of estimating parameters, $\boldsymbol{\theta}$, from some experimental dataset. In particular, we will explain the Approximate Bayesian Computation (ABC) rejection algorithm and its more complex alternative Approximate Bayesian Computation - Sequential Monte Carlo (ABC-SMC). For a given set of data and a model, both ABC and ABC-SMC techniques aim to find a posterior distribution for the model parameters by randomly sampling parameter values from a prior distribution, comparing the model prediction for those parameters to given data according to some distance measure, and accepting only the parameter choices that lie within a certain distance threshold.

The ABC rejection algorithm samples parameters from the prior distribution for the model parameters until a given number, N_1, of parameter sets have been accepted; whereas in ABC-SMC this is done as an iterative process, where both the sampling distribution for the parameters and the accepted distance threshold are updated upon each iteration. In more detail, the ABC-SMC algorithm starts with a predetermined prior distribution, $\pi(\boldsymbol{\theta})$, and accepted distance threshold, ε, which runs until a certain number, N_2, of parameter choices are accepted, at which point the next iteration can be carried out. Before starting the next iteration both the sampling distribution and the acceptance threshold are updated. The accepted parameter sets in this iteration become the sampling distribution of the next iteration. This process repeats until a number of iterations, T, are completed. Throughout this thesis, upon sampling a parameter set from the accepted parameter sets of the previous iteration, each parameter in the set is perturbed by adding noise proportional to the range of accepted values for that parameter and the accepted distance threshold for the next iteration is taken from the median returned distances of the accepted parameters sets in its current iteration. However, in general, there are many options for the perturbation method and choosing distance thresholds.

The ABC rejection algorithm returns one list of parameter sets of size N_1, whereas the ABC-SMC algorithm returns T lists of accepted parameter sets of size N_2 where each successive list improves the fit of the model to the data. This section is based on the work by Turner & Van Zandt (2012) and Toni *et al.* (2009).

2.3.1 The ABC rejection algorithm

The ABC rejection algorithm consists of a model, M, a sample size, N, an acceptance threshold, ε, a parameter vector, $\boldsymbol{\theta}$, consisting of all individual parameters within M and an associated prior distribution, $\pi(\boldsymbol{\theta})$.

A potential parameter set, $\boldsymbol{\theta}^*$, is obtained by randomly sampling from $\pi(\boldsymbol{\theta})$. This parameter set $\boldsymbol{\theta}^*$ is used to simulate M which produces a prediction D^* which is then compared to the empirical dataset D we are trying to find a model for, using some distance function $\rho(D, D^*)$ which measures the difference between the dataset and the model prediction. If $\rho(D, D^*) < \varepsilon$ then the parameter set $\boldsymbol{\theta}^*$ is accepted. However, if $\rho(D, D^*) \geq \varepsilon$ then $\boldsymbol{\theta}^*$ is rejected. The algorithm runs until the number of accepted parameter sets reaches N, which leads to a sample of size N that represents the posterior distribution of $\boldsymbol{\theta}$. This process is represented in Algorithm 2.2.

Algorithm 2.2: The ABC algorithm (Turner & Van Zandt (2012))

Create an empty list, L. Load model, $M(\boldsymbol{\theta})$, prior distribution, $\pi(\boldsymbol{\theta})$, desired posterior sample size, N, empirical dataset, D, acceptance threshold, ε, and distance function, $\rho(D, D^*)$;
While len$(L) < N$:
    Sample test parameter set $\boldsymbol{\theta}^*$ from the associated prior, $\pi(\boldsymbol{\theta})$;
    Compute model prediction, D^*, using $M(\boldsymbol{\theta}^*)$;
    If $\rho(D, D^*) < \varepsilon$:
        Accept $\boldsymbol{\theta}^*$ and add it to L;
    Else:
        Continue;
Return L;

2.3.2 The ABC-SMC algorithm

The ABC-SMC algorithm consists of a model, M, a sample size, N, a desired number of iterations, T, a set of acceptance thresholds, $\boldsymbol{\varepsilon} = \{\varepsilon_1, \ldots, \varepsilon_T\}$ with $\varepsilon_1 \geq \cdots \geq \varepsilon_T \geq 0$, a parameter vector, $\boldsymbol{\theta}$, consisting of all parameters within M and associated prior distribution, $\pi(\boldsymbol{\theta})$. Within each iteration, parameter values are sampled from the accepted parameters of the previous iteration and are then perturbed with a kernel function. For example, a component-wise uniform perturbation kernel can be used, so

that each component of the parameter set is perturbed independently in a uniform interval. This perturbed parameter set is then used to calculate a prediction from the model D^* and is accepted if $\rho(D, D^*) < \varepsilon_t$, where t is the current iteration. This generates a set of intermediate samples of accepted parameters, θ_t, that gradually converge towards the target posterior, θ_T. The ABC-SMC algorithm returns T accepted parameter sets of size N each. This process is described in Algorithm 2.3

Algorithm 2.3: The ABC-SMC algorithm (Toni *et al.* (2009)

Load model, $M(\boldsymbol{\theta})$, prior distribution, $\pi(\boldsymbol{\theta})$, desired posterior sample size, N, desired number of iterations, T, empirical dataset, D, set of perturbation kernels, $\{K_2, \ldots, K_T\}$, set of acceptance thresholds, $\boldsymbol{\varepsilon} = \{\varepsilon_1, \ldots, \varepsilon_T\}$ with $\varepsilon_1 \geq \ldots \varepsilon_T \geq 0$, and distance function, $\rho(D, D^*)$;
Set the population indicator, $t = 1$;
While $t \leq T$:
  Set the particle indicator $i = 1$;
  While $i \leq N$:
    If $t = 1$:
      Sample $\boldsymbol{\theta}^{**}$ from $\pi(\boldsymbol{\theta})$;
    Else:
      Sample $\boldsymbol{\theta}^*$ from the previous population $\{\boldsymbol{\theta}_{t-1}^{(i)}\}$, with weights w_{t-1} and perturb the particle to obtain $\boldsymbol{\theta}^{**} \sim K_t(\boldsymbol{\theta}|\boldsymbol{\theta}^*)$;
    If $\pi(\boldsymbol{\theta}^{**}) > 0$:
      Simulate model prediction, $D^* = M(\boldsymbol{\theta}^{**})$;
      If $\rho(D, D^*) < \varepsilon_t$:
        Set $\boldsymbol{\theta}_t^{(i)} = \boldsymbol{\theta}^{**}$ and calculate the weight for particle $\boldsymbol{\theta}_t^{(i)}$ as

$$w_t^{(i)} = \begin{cases} 1, & \text{if } t = 1, \\ \frac{\pi(\boldsymbol{\theta}_t^{(i)})}{\sum_{j=1}^{N} w_{t-1}^{(j)} K_t(\boldsymbol{\theta}_t^{(i)}|\boldsymbol{\theta}_{t-1}^{(j)})}, & \text{if } t > 1; \end{cases}$$

        Set $i = i + 1$;
  Normalise the weights.;
  Set $t = t + 1$;
Return $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_T$;

2.3.3 Comparing the two algorithms

Let us look at a simple model with a simulated data-set to see how these two algorithms work. Consider a logistic growth model (e.g. for a bacterial population) where $B(t)$ represents the number of bacteria (e.g. CFU) at time $t \geq 0$. The dynamics of this population can be modelled as

$$\frac{dB(t)}{dt} = rB(t)\left(1 - \frac{B(t)}{K}\right),$$

where $r > 0$ is the growth rate and $K > 0$ is the carrying capacity. The known solution to this is (Tsoularis & Wallace (2002))

$$B(t) = \frac{KB(0)e^{rt}}{K + B(0)(e^{rt} - 1)}, \quad t \geq 0. \tag{2.8}$$

Consider a given data set $\{D(t),\ t \in T = \{0h, 4h, 8h, 12h, 16h, 20h, 24h\}\}$, which represents bacterial growth in an experimental set up. We aim to calibrate our model and estimate unknown parameters $(B(0), K, r)$ by leveraging this data set, and by using the ABC and the ABC-SMC algorithms. To show how to do this, a synthetic data set was produced by simulating Equation (2.8) at times $t \in T$, for illustrative parameter values, $B(0) = 100$ CFU, $K = 10^4$ CFU, $r = 0.5$ hours^{-1} and adding noise proportional to the size of $B(t)$,

$$B_d(t) = B(t) + \varepsilon(t),$$

where $\varepsilon(t) \sim U(-0.1B(t), 0.1B(t))$. As is often the case in experimental conditions, we obtain replicates of the data at each time point. Here we take 5 replicates at each time point and take the mean,

$$D(t) = \frac{1}{5}\sum_{i=1}^{5}(B_d^i(t)),$$

and standard deviation, $\sigma(t)$, at each $t \in T$ and use this as our final dataset, $D(t)$; see Table 2.1. Here $B_d^i(t)$ represents the i-th replicate of $B_d(t)$. In order to provide the best comparison between ABC and ABC-SMC we keep every step of the process the same, other than the key differences detailed in the section above. This means that we use the

Data	0h	4h	8h	12h	16h	20h	24h
$B_d^1(t)$	100.1	625.2	3.71×10^3	7.68×10^3	9.29×10^3	1.08×10^4	9.92×10^3
$B_d^2(t)$	109.8	646.7	3.89×10^3	7.68×10^3	9.76×10^3	9.99×10^3	9.24×10^3
$B_d^3(t)$	101.0	704.6	3.29×10^3	7.40×10^3	9.34×10^3	9.95×10^3	9.97×10^3
$B_d^4(t)$	99.8	685.0	3.31×10^3	8.17×10^3	9.13×10^3	1.07×10^4	9.20×10^3
$B_d^5(t)$	96.1	702.0	3.35×10^3	7.83×10^3	1.04×10^4	9.86×10^3	9.39×10^3
$D(t)$	101.4	672.7	3.51×10^3	7.75×10^3	9.28×10^3	1.03×10^4	9.54×10^3
$\sigma(t)$	4.54	31.48	244.7	249.3	457.67	406.28	331.94

Table 2.1: The replicates of the simulated data, $B_d^i(t)$, their mean, $D(t)$, and their standard deviation, $\sigma(t)$, at each time point $t \in \{0, 4, \ldots, 24\}$.

Parameter	True Value	Prior
$B(0)$	10^2	$\log_{10}(B(0)) \sim U(0, 4)$
K	10^4	$\log_{10}(K) \sim U(3, 7)$
r	0.5	$r \sim U(0, 2)$

Table 2.2: True values and prior distributions considered for parameters $B(0), K$ and r.

same distance function for both techniques, namely

$$\rho\left(D^*, D\right)^2 = \sum_{t \in T} \left(\log_{10}(D^*(t)) - \log_{10}(D(t)\right)^2,$$

where D^* is the model prediction for any given parameter set $\boldsymbol{\theta}$, given by Equation (2.8). We choose a distance function involving taking the log of the data as we aim to limit the extent that the distance function will more heavily weight the later datapoints due to their much greater magnitude. We also use the same priors for each parameter in both algorithms. For the purposes of this example we aimed to choose ranges that were not excessively large for speed but not so small that they were trivial for the algorithm to find the optimal parameters. The chosen priors are detailed in Table 2.2.

In Figure 2.2 the results of the ABC rejection algorithm and the ABC-SMC algorithm are shown. For the ABC-SMC algorithim we ran $T = 10$ iterations, accepting 500 parameter sets within each iteration and an initial distance threshold of $\varepsilon_0 = 10$, where each subsequent distance threshold, $\varepsilon_1, \ldots, \varepsilon_9$, was chosen as the median of the distances of the accepted parameter sets in the previous iteration. For the ABC rejection algorithm we accepted 500 parameter sets with an acceptance threshold of $\varepsilon = 0.276$, which was

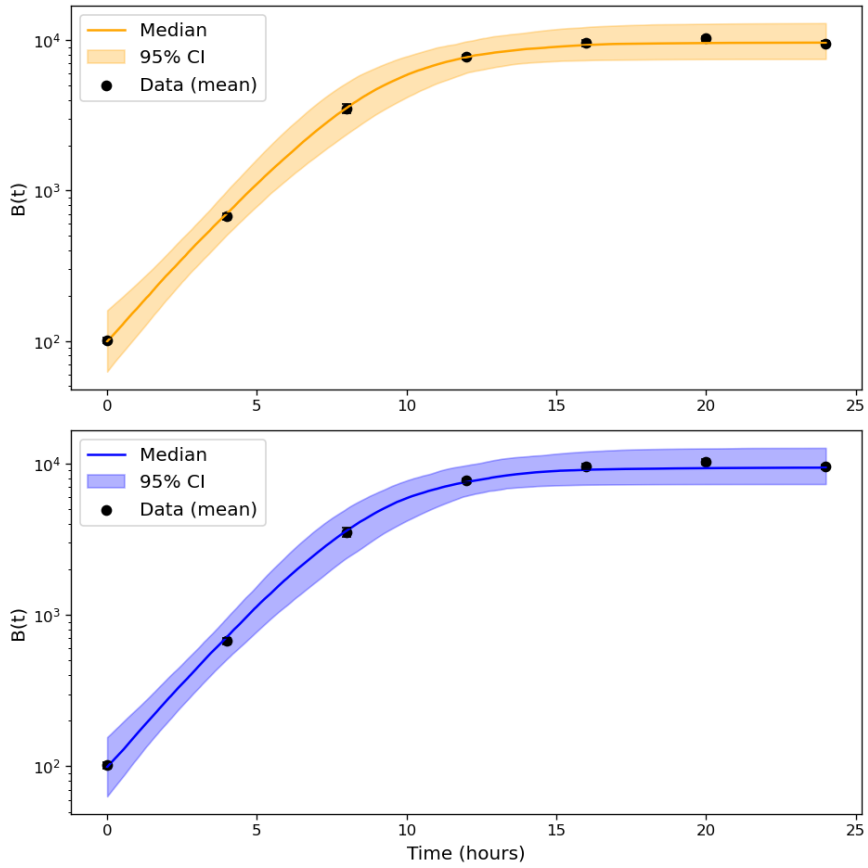

Figure 2.2: A comparison of the median simulation predictions and 95% confidence intervals (CI) for the accepted parameter values against the data (bars represent the standard deviation) for both the ABC rejection algorithm (top) and for the ABC-SMC algorithm (bottom).

the median distance of the accepted parameter sets of the 10th iteration of the ABC-SMC algorithm. From Figure 2.2, we can see that both the ABC rejection algorithm and ABC-SMC algorithm manage to capture the behaviour of the bacterial growth dynamics well with a tight confidence interval. However, the ABC-SMC algorithm worked much faster than the ABC algorithm (in the order of seconds and minutes, respectively); this is due to the ABC-SMC algorithm exploring the parameter space iteratively which is more efficient (as can be seen from the accepted parameter values region tightening with subsequent iterations in Figure 2.3), which leads to fewer rejected simulations and therefore a lower computational cost. To be precise, the 10th iteration of the ABC-SMC

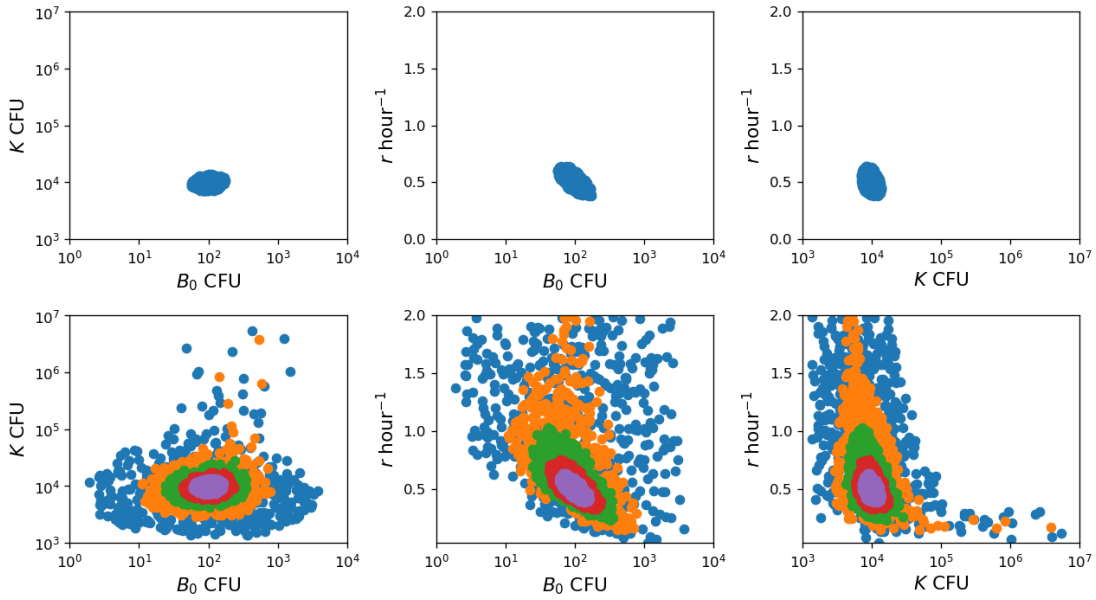

Figure 2.3: Posterior scatter plots of paired parameter values for the ABC rejection algorithm (top row) and 5 selected iterations of the ABC-SMC algorithm (bottom row). For the ABC-SMC plot, we have plotted the accepted parameter sets for iterations 2 (blue), 4 (orange), 6 (green), 8 (brown) and 10 (purple).

algorithm tested 1284 potential parameter sets to produce the posterior distribution whereas the ABC rejection algorithm tested 2.08×10^6 parameter sets to obtain its size 500 posterior distribution; this is a differential that can grow even wider for higher dimensional parameter spaces which explains why we use ABC-SMC algorithms within this thesis.

Figure 2.3 shows the pairwise scatter plots of the posterior distributions for the ABC rejection algorithm and for 5 iterations of the ABC-SMC algorithm. We again see similar behaviours to those we have just discussed, that both the ABC rejection algorithm and the ABC-SMC algorithm are able to find suitable posterior distributions for our parameter values (see Table 2.2 for the true values used to simulate the data). As mentioned above, we can see the ABC-SMC algorithm finding more accurate posterior distributions with each subsequent iteration. This allows the ABC-SMC algorithm to reach higher levels of accuracy much faster than the ABC rejection algorithm. Another benefit of ABC-SMC is that it does not require prior knowledge of an appropriate acceptance threshold to

perform well; whereas to obtain a posterior distribution as accurate as shown in Figure 2.3, we needed to be aware of the acceptance threshold $\varepsilon = 0.276$.

Chapter 3

Analysis of single-bacterium dynamics in a stochastic model of toxin-producing bacteria

In this chapter we stochastically model two bacterial populations which can produce toxins. We propose to analyse this biological system by following the dynamics of a single bacterium during its lifetime, as well as its progeny. The lifespan of a single bacterium, the number of divisions that a bacterium undergoes, and the number of toxin molecules that a bacterium produces during its lifespan are studied. The mean number of bacteria in the genealogy of the original bacterium and the number of toxin molecules produced by its genealogy are also computed. We illustrate the applicability of our methods by considering the bacteria *Bacillus anthracis* under antibiotic treatment. We quantify, for the first time, bacterial toxin production by exploiting an *in vitro* assay for the $A16R$ strain, and make use of the resulting parameterised model to illustrate our techniques.

3.1 Introduction

Mathematical modelling has proven to be a robust tool to analyse biological systems of relevance in infection and immunity at different scales, such as the molecular (López-

García *et al.* (2018)), intracellular (Carruthers *et al.* (2018)), within-host (Carruthers *et al.* (2020)) and population levels (Britton (2010)). While deterministic models are usually more amenable for mathematical analysis (Allen (2007)), stochastic approaches are generally better suited for representing biological systems involving few individuals (López-García (2016)) or cells (Carruthers *et al.* (2020)), or when extinction events play a crucial role (Brockwell (1986)). Markov processes, either in discrete or continuous time, have been particularly exploited in these areas given their mathematical convenience (Allen (2010)). While non-Markovian dynamics are typically more difficult to analyse (Castro *et al.* (2018); Gómez-Corral & López-García (2017)), the Markovian or memoryless property usually allows for mathematical tractability and efficient numerical simulations (Gillespie (1977)).

When considering a population of cells during an immune response, or bacteria during infection, competition for resources is usually represented in terms of logistic growth models (Allen (2007)). On the other hand, when individuals behave independently (*e.g.* they do not compete for common resources), the theory of branching processes (Kimmel & Axelrod (2002)) has been widely applied to follow these populations (of cells or bacteria) over time. Multi-type branching processes (Kyprianou & Palau (2018)) allow one to consider different types of bacteria, which might represent different phenotypes (Choi *et al.* (2008)) or different spatial locations (*e.g.,* tissues or organs) within the body during an infection (Carruthers *et al.* (2020)). The complexity of these processes, and their mathematical tractability, typically depends on the number of compartments considered, and the number of potential events that can occur in the system (*e.g.,* division or death of bacteria, or bacterial movement across compartments) (Thakur *et al.* (1973)).

Novel technological developments have recently allowed for single cells to be precisely followed, together with their progeny (Herzenberg *et al.* (2002); Johnson *et al.* (2009); Krutzik & Nolan (2006); Westera *et al.* (2013)). This motivates the idea of mathematically tracking single individuals in these stochastic systems, and to quantify summary statistics related to the lifetime of a single individual (or bacterium in our case), and its progeny or genealogy. In this chapter, we consider the genealogy of a bacterium to refer to a bacterium itself along with all of its descendants. Analysing the dynamics of the system by tracking a single individual has already been proposed in related areas such as population dynamics (Gómez-Corral & López-García (2015)) and, more recently,

when analysing the stochastic journey of T lymphocytes in lymph nodes and blood (de la Higuera *et al.* (2019)).

Although bacterial systems have been widely studied with stochastic methods in the past (Carruthers *et al.* (2018, 2020)), less attention has been paid to the study of toxin-producing bacteria. The production of toxins over time can be especially relevant for specific bacteria for which the secreted toxins can cause suppression of the host's immune system, being a key component of pathogenesis *in vivo* (Banks *et al.* (2005)). In this chapter, we illustrate our single-cell approach in a stochastic model of two types of toxin-producing bacteria. Two types are considered to allow flexibility of the model; bacteria could be changed due to differentiation during the replication process or by exposure to an external agent; in our case study in Section 3.4.2 we consider the change to be antibiotic binding of a *Bacillus anthracis* bacterium. We focus on computing the expected lifespan of a single bacterium in this system, as well as the number of toxin molecules secreted and the number of divisions undergone during its lifetime. We also compute summary statistics that are directly related to the progeny of a single bacterium, namely the number of bacteria in the progeny of a given bacterium and the number of toxin molecules produced by its genealogy, where we define the genealogy of a bacterium as the bacterium itself along with its progeny. We illustrate our results by focusing on the bacteria *Bacillus anthracis* and the anthrax toxins. For the $A16R$ *B. anthracis* strain, we quantify for the first time the rate of protective antigen (PA) production by exploiting data from an *in vitro* assay by Zai *et al.* (2016). We make use of the resulting parametrised mathematical model to illustrate our techniques, and we also consider antibiotic treatment.

The structure of this chapter is as follows. In Section 3.2 we introduce the mathematical model. The single-bacterium representation is discussed in Section 3.3. A number of summary statistics of interest related to a single bacterium and its progeny are studied analytically in Section 3.3 and numerically in Section 3.4.1. Model calibration for the $A16R$ *B. anthracis* strain is carried out in Section 3.4.2 using data from an *in vitro* assay, and the parameterised model is used in this section to illustrate our methodology. Concluding remarks are provided in Section 3.5.

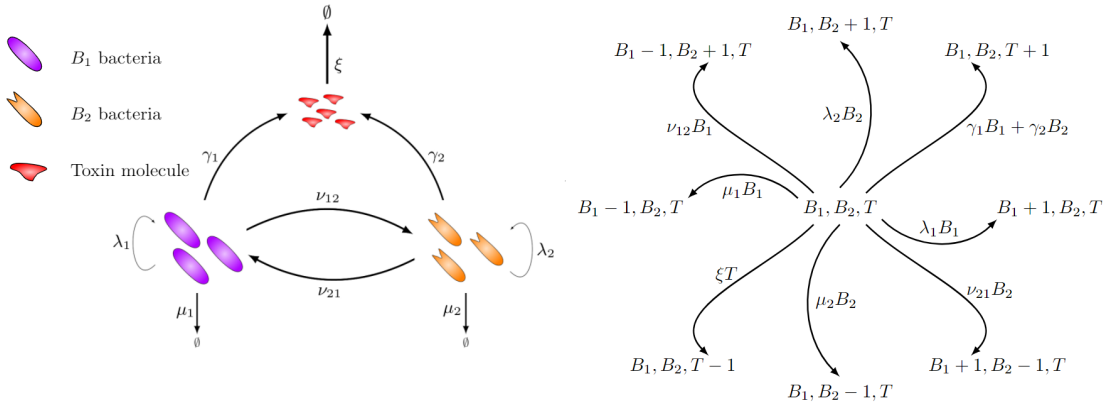

Figure 3.1: Left. Diagram showing the dynamics of the two toxin-producing bacterial populations. **Right.** Allowed transitions between states in $\mathcal{X}$ and their rates.

3.2 The mathematical model

Our interest is in modelling a system with two toxin-producing bacterial populations (see Figure 3.1). Type i bacteria, $i \in \{1, 2\}$, can divide with rate λ_i, produce toxins with rate γ_i, die with rate μ_i, or become type j bacteria, $j \in \{1, 2\}$ $j \neq i$, with rate ν_{ij}. We can thus develop a stochastic model of these events, as a continuous-time Markov chain (CTMC) $\mathcal{X} = \{(B_1(t), B_2(t), T(t)) : t \geq 0\}$, where $B_i(t)$ denotes the number of type i bacteria at time $t \geq 0$, $i \in \{1, 2\}$, and $T(t)$ represents the number of toxin molecules at time $t \geq 0$. We assume that bacteria and toxins behave independently of each other, and that toxins are degraded at rate ξ. The space of states of $\mathcal{X}$ is given by $\mathcal{S} = \mathbb{N}_0^3$, where we denote $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, and the possible one-step transitions between states in $\mathcal{X}$ are depicted in Figure 3.1.

Since each bacterium behaves independently, one can analyse the dynamics of a single bacterium without explicitly modelling the dynamics of the rest of the population. In Section 3.3, we propose a methodology that focuses on analysing the dynamics of a single bacterium and its progeny. In particular, and by means of first-step arguments, we compute the lifespan of a single bacterium, the number of divisions that a bacterium undergoes during its lifespan, and the number of toxin molecules that a bacterium produces during its lifetime. We also compute the mean number of cells within the progeny of the original bacterium and the number of toxin molecules produced by this progeny. We note that a particular advantage of this single-bacterium approach is that

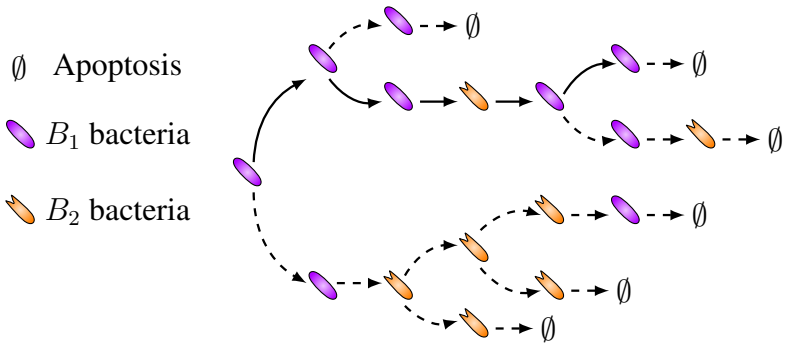

Figure 3.2: Example of a stochastic realisation of the population process, starting with one type 1 bacterium. Solid arrows indicate the single bacterium being tracked in process $\mathcal{Y}$. In this realisation, the stochastic process $\mathcal{Y}$ visits states $B_1 \to B_1 \to B_1 \to B_2 \to B_1 \to B_1 \to \emptyset$. Consecutive visits to the same state are due to bacterial division. Toxin production is not explicitly depicted here but can occur during the process.

it can be implemented regardless of the complexity of the model, *i.e.,* regardless of the number of compartments in the model, or the number of events governing the toxin and bacterial dynamics across compartments, as long as the dynamics of each bacterium is independent of the rest of the population.

3.3 Dynamics of a single bacterium and its progeny

Instead of focusing on the population CTMC $\mathcal{X}$, our interest in this section is in following a single bacterium of type i during its lifespan. In particular, we consider a single bacterium (either of type 1 or type 2) at time $t = 0$, and follow its dynamics during its lifetime by considering the individual-level continuous-time Markov chain $\mathcal{Y} = \{Y(t) : t > 0\}$, where $Y(t)$ represents the "*state*" of the bacterium at time $t \geq 0$. By state, we mean that the bacterium can be of type 1, type 2 or dead at any given time. Thus, $\mathcal{Y}$ is defined on the state space $S = \{1, 2, \emptyset\}$, where i here represents the bacterium being of type i at any given instant, and $\emptyset$ indicates the bacterium is dead. If the bacterium is of type i at a given instant, meaning that $\mathcal{Y}$ is in state i, production of a toxin molecule does not change its state, and $\mathcal{Y}$ remains in state i. If a division occurs, we randomly choose one of the daughter cells and consider it to be our bacterium of interest, which remains in state i.

Figure 3.2 shows one realisation of the population dynamics, while the state of the stochastic process $\mathcal{Y}$ only depends on tracking the original bacterium throughout its lifetime, which is depicted via solid arrows. When a division occurs, a daughter is randomly chosen to represent the tracked bacterium of interest. In the following subsections, we investigate a number of stochastic descriptors or summary statistics that relate to the single bacterium, as well as some descriptors that refer to a given bacterium's progeny or genealogy.

3.3.1 Lifespan of a bacterium

For an initial bacterium of type $i \in \{1, 2\}$, we define its lifespan as $T_i = \inf\{t \geq 0 : Y(t) = \emptyset | Y(0) = B_i\}$. We can ascertain various information about this descriptor T_i; in this section we will use the Laplace-Stieltjes transform to find a probability generating function for T_i, and then use this transform to find moments of T_i, such as the first order moment which gives us the expectation of T_i.

3.3.1.1 The Laplace-Stieltjes transform

We consider the Laplace-Stieltjes transform of T_i given by

$$\phi_i(s) = \mathbb{E}[e^{-sT_i}], \quad \text{Re}(s) \geq 0, \tag{3.1}$$

which one can compute with first step arguments. We denote $\phi_1(s) = \mathbb{E}[e^{-sT_1}|Y(0) = 1]$ and similarly $\phi_2(s) = \mathbb{E}[e^{-sT_2}|Y(0) = 2]$. If $\mathcal{Y}$ is in state 1 at any given time, let $1 \to 2$ represent the event transitioning $\mathcal{Y}$ from state 1 to state 2 in its next jump, $1 \circlearrowleft$ represent a division event in its next jump, $1 \mapsto 1$ represent a toxin being produced in its next jump and $1 \to \emptyset$ represent the bacterium dying in its next jump. Also let t_i denote the time taken for this jump to occur.

By examining the events that could occur next and using the total expectation law, $\phi_1(s)$

can be written as

$$\begin{aligned}\phi_1(s) = \mathbb{E}[e^{-s(t_1+T_1)}|Y(0) = 1, 1 \circlearrowleft] \cdot \mathbb{P}(1 \circlearrowleft) \\ + \mathbb{E}[e^{-s(t_1+T_1)}|Y(0) = 1, 1 \mapsto 1] \cdot \mathbb{P}(1 \mapsto 1) \\ + \mathbb{E}[e^{-s(t_1+T_2)}|Y(0) = 1, 1 \to 2] \cdot \mathbb{P}(1 \to 2) \\ + \mathbb{E}[e^{-st_1}|Y(0) = 1, 1 \to \emptyset] \cdot \mathbb{P}(1 \to \emptyset).\end{aligned}$$

Therefore the above equation is equivalent to

$$\begin{aligned}\phi_1(s) = \mathbb{E}[e^{-s(t_1+T_1)}|Y(0) = 1, 1 \circlearrowleft] \cdot \left(\frac{\lambda_1}{\lambda_1 + \gamma_1 + \nu_{12} + \mu_1}\right) \\ + \mathbb{E}[e^{-s(t_1+T_1)}|Y(0) = 1, 1 \mapsto 1] \cdot \left(\frac{\gamma_1}{\lambda_1 + \gamma_1 + \nu_{12} + \mu_1}\right) \\ + \mathbb{E}[e^{-s(t_1+T_2)}|Y(0) = 1, 1 \to 2] \cdot \left(\frac{\nu_{12}}{\lambda_1 + \gamma_1 + \nu_{12} + \mu_1}\right) \\ + \mathbb{E}[e^{-st_1}|Y(0) = 1, 1 \to \emptyset] \cdot \left(\frac{\mu_1}{\lambda_1 + \gamma_1 + \nu_{12} + \mu_1}\right).\end{aligned} \tag{3.2}$$

The first term in this equation represents the case in which a division event takes place in the next jump. In this case the lifespan of the bacterium can be expressed as the time for this event to occur plus the remaining lifespan of the bacterium. This occurs with probability $\frac{\lambda_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}$. In the second term of the equation, where a toxin is produced in the next jump, the lifespan can be expressed as the time for this event to occur plus the remaining lifespan of the bacterium after this jump. In the third term of the equation, where the bacterium changes to type 2 in the next jump, the lifespan of the bacterium can be represented as the time taken for this event plus the remaining lifespan of the new type 2 bacterium. In the case in which death occurs in the next jump, the lifespan of this bacterium is solely the time taken for this event to occur.

Consider the term $\mathbb{E}[e^{-s(t_1+T_2)}|Y(0) = 1, 1 \to 2]$; this is the expectation of the product of two functions of two random variables. The component $t_1|1 \to 2$ considers the event that has just happened, whereas $T_2|1 \to 2$ refers to the remaining lifespan of the bacterium just after $1 \to 2$ occurred. By the Markovian property of a CTMC this means that these two events must be independent of one another, leading to $\mathbb{E}[e^{-s(t_1+T_2)}|Y(0) = 1, 1 \to 2] = \mathbb{E}[e^{-st_1}|Y(0) = 1, 1 \to 2] \cdot \mathbb{E}[e^{-sT_2}|Y(0) = 1, 1 \to 2]$.

This means that Equation (3.2) can be rewritten as

$$\begin{aligned}\phi_1(s) = \mathbb{E}[e^{-st_1}|1 \circlearrowleft] \cdot \mathbb{E}[e^{-sT_1}] \cdot \left(\frac{\lambda_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right) \\ + \mathbb{E}[e^{-st_1}|1 \mapsto 1] \cdot \mathbb{E}[e^{-sT_1}] \cdot \left(\frac{\gamma_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right) \\ + \mathbb{E}[e^{-st_1}|1 \to 2] \cdot \mathbb{E}[e^{-sT_2}] \cdot \left(\frac{\nu_{12}}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right) \\ + \mathbb{E}[e^{-st_1}|1 \to \emptyset] \cdot \left(\frac{\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right).\end{aligned} \tag{3.3}$$

As the first jump occurs after an exponentially distributed time, we have $(t_1|1 \circlearrowleft) \equiv (t_1|1 \mapsto 1) \equiv (t_1|1 \to 2) \equiv (t_1|1 \to \emptyset) \sim \text{Exp}(\lambda_1+\gamma_1+\nu_{12}+\mu_1)$.

This means that $\mathbb{E}[e^{-st_1}|1 \to 2]$ is the Laplace-Stieltjes transform of $\text{Exp}(\lambda_1+\gamma_1+\nu_{12}+\mu_1)$ (and similarly for the other cases), which is equal to $\frac{\lambda_1+\gamma_1+\nu_{12}+\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1+s}$; as the Laplace-Stieltjes transform of an exponential random variable, $X \sim \text{Exp}(q)$ is known to be $\mathbb{E}[e^{-sX}] = \frac{q}{q+s}$. Therefore Equation (3.3) is equivalent to

$$\begin{aligned}\phi_1(s) = \left(\frac{\lambda_1+\gamma_1+\nu_{12}+\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1+s}\right)\phi_1(s)\left(\frac{\lambda_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right) \\ + \left(\frac{\lambda_1+\gamma_1+\nu_{12}+\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1+s}\right)\phi_1(s)\left(\frac{\gamma_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right) \\ + \left(\frac{\lambda_1+\gamma_1+\nu_{12}+\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1+s}\right)\phi_2(s)\left(\frac{\nu_{12}}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right) \\ + \left(\frac{\lambda_1+\gamma_1+\nu_{12}+\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1+s}\right)\left(\frac{\mu_1}{\lambda_1+\gamma_1+\nu_{12}+\mu_1}\right).\end{aligned} \tag{3.4}$$

By cancelling terms and defining $\Delta_i = \lambda_i+\mu_i+\gamma_i+\nu_{ij}$, $i,j \in \{1,2\}$, $j \neq i$, Equation (3.4) leads to the following equation

$$\phi_1(s) = \frac{\lambda_1}{\Delta_1+s}\phi_1(s) + \frac{\mu_1}{\Delta_1+s} + \frac{\gamma_1}{\Delta_1+s}\phi_1(s) + \frac{\nu_{12}}{\Delta_1+s}\phi_2(s).$$

This equation simplifies to

$$(\mu_1 + \nu_{12} + s)\,\phi_1(s) \;=\; \nu_{12}\phi_2(s) + \mu_1. \tag{3.5}$$

Symmetric arguments to the ones above allow one to find an equation for $\phi_2(s)$,

$$(\mu_2 + \nu_{21} + s)\,\phi_2(s) = \nu_{21}\phi_1(s) + \mu_2.$$

Interestingly, already we can see that these equations do not depend on the parameters λ_i (division rate) or γ_i (toxin production rate). This is to be expected since the division and toxin production events do not affect the lifespan of a bacterium, something we would expect from inspecting the dynamics in Figures 3.1 and 3.2. By substituting one equation into the other, we can find solutions for these $\phi_i(s)$, as shown below for $\phi_1(s)$.

We start by rearranging the equation for $\phi_2(s)$

$$\begin{aligned}(\mu_2 + \nu_{21} + s)\,\phi_2(s) &= \nu_{21}\phi_1(s) + \mu_2,\\ \Rightarrow \phi_2(s) &= \frac{\nu_{21}\phi_1(s) + \mu_2}{\mu_2 + \nu_{21} + s}.\end{aligned}$$

Substituting this into Equation (3.5) gives

$$(\mu_1 + \nu_{12} + s)\,\phi_1(s) = \nu_{12}\left(\frac{\nu_{21}\phi_1(s) + \mu_2}{\mu_2 + \nu_{21} + s}\right) + \mu_1.$$

Grouping the $\phi_1(s)$ terms gives

$$\begin{aligned}(\mu_1 + \nu_{12} + s)\,\phi_1(s) - \frac{\nu_{12}\nu_{21}\phi_1(s)}{\mu_2 + \nu_{21} + s} &= \frac{\nu_{12}\mu_2}{\mu_2 + \nu_{21} + s} + \mu_1\\ \Rightarrow \left(\mu_1 + \nu_{12} + s - \frac{\nu_{12}\nu_{21}}{\mu_2 + \nu_{21} + s}\right)\phi_1(s) &= \frac{\nu_{12}\mu_2}{\mu_2 + \nu_{21} + s} + \mu_1.\end{aligned}$$

To simplify this slightly, we introduce the notation

$$a(s) = 1 - \frac{\nu_{12}\nu_{21}}{(\mu_1 + \nu_{12} + s)(\mu_2 + \nu_{21} + s)}, \tag{3.6}$$

which makes the equation become

$$(\mu_1 + \nu_{12} + s)\, a(s)\phi_1(s) = \left(\frac{\nu_{12}\mu_2}{\mu_2 + \nu_{21} + s}\right) + \mu_1,$$

leading to a solution of

$$\phi_1(s) = \; a(s)^{-1}\frac{1}{\mu_1 + \nu_{12} + s}\left(\frac{\nu_{12}\mu_2}{\mu_2 + \nu_{21} + s} + \mu_1\right). \tag{3.7}$$

Once again, we can use a symmetric process to arrive at a solution for $\phi_2(s)$:

$$\phi_2(s) \;\; = \;\; a(s)^{-1}\frac{1}{\mu_2 + \nu_{21} + s}\left(\frac{\nu_{21}\mu_1}{\mu_1 + \nu_{12} + s} + \mu_2\right). \tag{3.8}$$

Before we use these solutions, it is worth noting how these complex solutions can simplify greatly even under a single constraint for the bacterial system. For instance, if the change from type 1 to type 2 bacterium was irreversible, so that $\nu_{21} = 0$, one obtains

$$\begin{aligned}\phi_1(s) &= \frac{1}{\mu_1 + \nu_{12} + s}\left(\frac{\nu_{12}\mu_2}{\mu_2 + s} + \mu_1\right),\\ \phi_2(s) &= \frac{\mu_2}{\mu_2 + s}.\end{aligned}$$

Note that in this case $T_2 \sim \mathrm{Exp}(\mu_2)$, which is consistent with the expression for $\phi_2(s)$ above. This is an interesting case to consider since the bacterial conversion with rate ν_{12} and reversion rate $\nu_{21} = 0$ represents the irreversible antibiotic treatment we consider in Section 3.4.2.

3.3.1.2 Moments of the random variable T_i

In this section the moments of T_1 will be computed in detail and the corresponding results for T_2 given, as the methodology is the same for both.

A key property of the Laplace-Stieltjes transform is that:

$$(-1)^k \frac{d^k}{ds^k}\phi_i(s)\bigg|_{s=0} = \mathbb{E}[T_i^k]. \tag{3.9}$$

Differentiating Equation (3.7) at $s = 0$ gives

$$\begin{aligned}\phi_1'(0) = &- a(0)^{-2}\left(\frac{\nu_{12}\nu_{21}}{(\mu_1+\nu_{12})(\mu_2+\nu_{21})^2}+\frac{\nu_{12}\nu_{21}}{(\mu_1+\nu_{12})^2(\mu_2+\nu_{21})}\right)\frac{1}{\mu_1+\nu_{12}}\times\\ &\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}}+\mu_1\right)-a(0)^{-1}\frac{1}{(\mu_1+\nu_{12})^2}\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}}+\mu_1\right)\\ &-a(0)^{-1}\frac{1}{\mu_1+\nu_{12}}\frac{\mu_2\nu_{12}}{(\mu_2+\nu_{21})^2}.\end{aligned}$$

We can take a factor of $-a(0)^{-1}$ and $\frac{1}{\mu_1+\nu_{12}}$ from each term leaving us with

$$\begin{aligned}\phi_1'(0) = &- a(0)^{-1}\frac{1}{\mu_1+\nu_{12}}\left[\frac{1}{(\mu_1+\nu_{12})}\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}}+\mu_1\right)+\frac{\mu_2\nu_{12}}{(\mu_2+\nu_{21})^2}\right.\\ &\left.+\,a(0)^{-1}\left(\frac{\nu_{12}\nu_{21}}{(\mu_1+\nu_{12})(\mu_2+\nu_{21})^2}+\frac{\nu_{12}\nu_{21}}{(\mu_1+\nu_{12})^2(\mu_2+\nu_{21})}\right)\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}}+\mu_1\right)\right].\end{aligned}$$

Inside the square bracket we can substitute in the formula for $a(0)^{-1}$ to get

$$\begin{aligned}\phi_1'(0) = &- a(0)^{-1}\frac{1}{\mu_1+\nu_{12}}\left[\frac{1}{(\mu_1+\nu_{12})}\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}}+\mu_1\right)+\frac{\mu_2\nu_{12}}{(\mu_2+\nu_{21})^2}\right.\\ &\left.+\,\frac{1}{(\mu_1+\nu_{12})(\mu_2+\nu_{21})-\nu_{12}\nu_{21}}\left(\frac{\nu_{12}\nu_{21}}{\mu_2+\nu_{21}}+\frac{\nu_{12}\nu_{21}}{\mu_1+\nu_{12}}\right)\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}}+\mu_1\right)\right].\end{aligned}$$

This simplifies to

$$\phi_1'(0) = -a(0)^{-1}\frac{1}{\mu_1+\nu_{12}}\left[\frac{\nu_{12}}{\mu_2+\nu_{21}}+1\right],$$

multiplying by -1 allows us to invoke Equation (3.9), leading to a solution for $\mathbb{E}[T_1]$,

$$\mathbb{E}[T_1] = a(0)^{-1}\frac{1}{\mu_1+\nu_{12}}\left(\frac{\nu_{12}}{\mu_2+\nu_{21}}+1\right).$$

Similarly we can find that

$$\mathbb{E}[T_2] = a(0)^{-1}\frac{1}{\mu_2+\nu_{21}}\left(\frac{\nu_{21}}{\mu_1+\nu_{12}}+1\right).$$

Alternatively, we can instead return to a simpler form of the equation, namely Equation (3.5). From this point, by multiplying Equation (3.5) by -1 and differentiating once with

respect to s at $s = 0$, it is possible to calculate the expectation of T_1. The result of this follows.

$$-1 \cdot \frac{d}{ds}\left((\mu_1 + \nu_{12} + s)\phi_1(s)\right)\Big|_{s=0} = -1 \cdot \frac{d}{ds}\left(\nu_{12}\phi_2(s) + \mu_1\right)\Big|_{s=0},$$

$$\Rightarrow -1 \cdot \phi_1'(0) \cdot (\mu_1 + \nu_{12}) - 1 \cdot \phi_1(0) = -1 \cdot \phi_2'(0) \cdot \nu_{12}.$$

By implementing Equation (3.9) we obtain the equation

$$\mathbb{E}[T_1] \cdot (\mu_1 + \nu_{12}) - \phi_1(0) = \nu_{12} \cdot \mathbb{E}[T_2]. \tag{3.10}$$

It can be shown using a similar process that

$$\mathbb{E}[T_2] \cdot (\mu_2 + \nu_{21}) - \phi_2(0) = \nu_{21} \cdot \mathbb{E}[T_1]. \tag{3.11}$$

Rearranging Equation (3.11) gives

$$\mathbb{E}[T_2] = (\mu_2 + \nu_{21})^{-1}\left(\phi_2(0) + \nu_{21}\mathbb{E}[T_1]\right),$$

and substituting this into Equation (3.10) gives

$$\mathbb{E}[T_1] \cdot (\mu_1 + \nu_{12}) - \phi_1(0) = \nu_{12} \cdot (\mu_2 + \nu_{21})^{-1}\left(\phi_2(0) + \nu_{21}\mathbb{E}[T_1]\right).$$

Grouping the $\mathbb{E}[T_1]$ terms gives

$$\mathbb{E}[T_1]\left((\mu_1 + \nu_{12}) - \frac{\nu_{12}\nu_{21}}{\mu_2 + \nu_{21}}\right) = \nu_{12} \cdot (\mu_2 + \nu_{21})^{-1}\phi_2(0) + \phi_1(0). \tag{3.12}$$

Recall the definition of $\phi_i(s)$ from Equation (3.1); therefore $\phi_i(0) = \mathbb{E}[e^0] = 1$. Substituting this into Equation (3.12) gives

$$\mathbb{E}[T_1]\left((\mu_1 + \nu_{12}) - \frac{\nu_{12}\nu_{21}}{\mu_2 + \nu_{21}}\right) = \nu_{12} \cdot (\mu_2 + \nu_{21})^{-1} + 1.$$

Making $\mathbb{E}[T_1]$ the subject gives

$$\begin{aligned}\mathbb{E}[T_1] &= \left((\mu_1+\nu_{12}) - \frac{\nu_{12}\nu_{21}}{\mu_2+\nu_{21}}\right)^{-1}\left(\frac{\nu_{12}}{(\mu_2+\nu_{21})}+1\right)\\ &= \left(\frac{\mu_2+\nu_{21}}{\mu_1\mu_2+\mu_1\nu_{21}+\mu_2\nu_{12}}\right)\left(\frac{\nu_{12}}{(\mu_2+\nu_{21})}+1\right).\end{aligned}$$

From Equation (3.6) we can find that

$$\begin{aligned}a(0)^{-1} &= \frac{(\mu_1+\nu_{12})(\mu_2+\nu_{21})}{\mu_1\mu_2+\mu_1\nu_{21}+\mu_2\nu_{12}}\\ \Rightarrow a(0)^{-1}\left(\frac{1}{\mu_1+\nu_{12}}\right) &= \left(\frac{\mu_2+\nu_{21}}{\mu_1\mu_2+\mu_1\nu_{21}+\mu_2\nu_{12}}\right).\end{aligned}$$

And thus we have re-obtained the solution

$$\mathbb{E}[T_1] = a(0)^{-1}\frac{1}{\mu_1+\nu_{12}}\left(\frac{\nu_{12}}{\mu_2+\nu_{21}}+1\right).$$

Note that, if $\nu_{21}=0$ as in our case study in Section 3.4.2,

$$\mathbb{E}[T_1] = \frac{\nu_{12}+\mu_2}{\mu_1\mu_2+\mu_2\nu_{12}}.$$

We note that one can also use the Laplace-Stieltjes transform to find higher order moments, one such example is the second order moment of a bacterium starting in state 1. We can find this by using Equation (3.9), and differentiating Equation (3.5) twice. This leads to

$$\begin{aligned}(\mu_1+\nu_{12})\phi_1''(0)+2\phi_1'(0) &= \nu_{12}\phi_2''(0)\\ \Rightarrow (\mu_1+\nu_{12})\mathbb{E}[T_1^2]-2\mathbb{E}[T_1] &= \nu_{12}\mathbb{E}[T_2^2].\end{aligned}$$

We can follow a similar process to obtain

$$(\mu_2+\nu_{21})\mathbb{E}[T_2^2]-2\mathbb{E}[T_2] = \nu_{21}\mathbb{E}[T_1^2].$$

By following a similar process for the first order moment: combining these into one equation for $\mathbb{E}[T_1^2]$ and then substituting in the formulae for $\mathbb{E}[T_1]$ and $\mathbb{E}[T_2]$, it can be

shown that

$$\mathbb{E}[T_1^2] = \frac{2a(0)^{-2}}{(\mu_1+\nu_{12})(\mu_2+\nu_{21})}\left(\frac{\nu_{12}}{\mu_2+\nu_{21}}\left(\frac{\nu_{21}}{\mu_1+\nu_{12}}+1\right)+\frac{\mu_2+\nu_{21}}{\mu_1+\nu_{12}}\left(\frac{\nu_{12}}{\mu_2+\nu_{21}}+1\right)\right).$$

We can also use this equation to find the second order moment of the lifespan of a bacterium starting in state 1 when $\nu_{21}=0$ which is given by

$$\mathbb{E}[T_1^2] = \frac{2}{(\mu_1+\nu_{12})\mu_2}\left(\frac{\nu_{12}}{\mu_2}+\frac{\mu_2}{\mu_1+\nu_{12}}\left(\frac{\nu_{12}}{\mu_2}+1\right)\right).$$

3.3.2 Number of toxin molecules produced by a bacterium in its lifetime

The random variable that describes the number of toxin molecules produced by the tracked bacterium during its lifetime is denoted by ω_i, where the tracked bacterium is initially of type $i \in \{1,2\}$. This is an important descriptor for the model since toxins can be one of the main causes of severe symptoms within a host. Within this section we shall consider the probability generating function and use it to find moments of this descriptor ω_i. The probability distribution of ω_i will also be considered.

3.3.2.1 Probability generating function of ω_i

In the previous section, we use the Laplace-Stieltjes transform to find the moments of T_i, as it was a continuous random variable. In this section, and moving forward, we will be considering discrete random variables. Therefore, we can consider the probability generating function of ω_i to be defined as follows

$$\psi_i(z) = \mathbb{E}[z^{\omega_i}],$$

for $|z| \leq 1$. By means of a first-step argument (similar to the ones above), one can show that

$$(\mu_1+\gamma_1(1-z)+\nu_{12})\psi_1(z) = \nu_{12}\psi_2(z)+\mu_1, \tag{3.13}$$

$$(\mu_2+\gamma_2(1-z)+\nu_{21})\psi_2(z) = \nu_{21}\psi_1(z)+\mu_2. \tag{3.14}$$

Using a similar technique as described when finding the Laplace-Stieltjes transform of

T_i, it can be shown that Equations (3.13)-(3.14) have the following solutions

$$\begin{aligned}\psi_1(z) &= b(z)^{-1}\frac{1}{\mu_1+\gamma_1(1-z)+\nu_{12}}\left(\frac{\nu_{12}\mu_2}{\mu_2+\nu_{21}+\gamma_2(1-z)}+\mu_1\right),\\ \psi_2(z) &= b(z)^{-1}\frac{1}{\mu_2+\gamma_2(1-z)+\nu_{21}}\left(\frac{\nu_{21}\mu_1}{\mu_1+\nu_{12}+\gamma_1(1-z)}+\mu_2\right),\end{aligned}$$

with $b(z)=1-\frac{\nu_{12}\nu_{21}}{(\mu_1+\gamma_1(1-z)+\nu_{12})(\mu_2+\gamma_2(1-z)+\nu_{21})}$. Once again, we can consider the scenario where $\nu_{21}=0$, a relevant consideration for applying the model to a *B. anthracis* model in Section 3.4.2, which leads to simplified solutions. Namely,

$$\begin{aligned}\psi_1(z) &= \frac{1}{\mu_1+\gamma_1(1-z)+\nu_{12}}\left(\frac{\nu_{12}\mu_2}{\mu_2+\gamma_2(1-z)}+\mu_1\right),\\ \psi_2(z) &= \frac{\mu_2}{\mu_2+\gamma_2(1-z)}.\end{aligned}$$

We note that in this case $\omega_2 \sim \text{Geo}(\frac{\mu_2}{\mu_2+\gamma_2})$, which is consistent with the solution of $\psi_2(z)$.

The mean number of toxin molecules produced by a single bacterium can be computed from direct differentiation of $\psi_i(z)$ with respect to z using the formula

$$\frac{d^k}{dz^k}\psi_i(z)\Big|_{z=1} = \mathbb{E}[\omega_i^k].$$

This is very similar to the techniques described when finding the first-order moment of T_i in Section 3.3.1. In particular, one can show that

$$\begin{aligned}\mathbb{E}[\omega_1] &= b(1)^{-1}\frac{1}{\mu_1+\nu_{12}}\left(\frac{\gamma_2\nu_{12}}{\mu_2+\nu_{21}}+\gamma_1\right),\\ \mathbb{E}[\omega_2] &= b(1)^{-1}\frac{1}{\mu_2+\nu_{21}}\left(\frac{\gamma_1\nu_{21}}{\mu_1+\nu_{12}}+\gamma_2\right).\end{aligned}$$

We can also find the second-order moment of this descriptor using a similar technique. As an example, under the constraint $\nu_{21}=0$, the second order moment of ω_1 is

$$\mathbb{E}[\omega_1^2] = \frac{2\gamma_1^2\left(\mu_1+\frac{\mu_2\nu_{12}}{\gamma_2+\mu_2}\right)}{(\gamma_1+\mu_1+\nu_{12})^3}+\frac{2\gamma_1\gamma_2\mu_2\nu_{12}}{(\gamma_2+\mu_2)^2(\gamma_1+\mu_1+\nu_{12})^2}+\frac{2\gamma_2^2\mu_2\nu_{12}}{(\gamma_2+\mu_2)^3(\gamma_1+\mu_1+\nu_{12})}.$$

3.3.2.2 Probability distribution of ω_i

We can also attempt to find the probability distribution of ω_1 by taking first-step arguments. Let us begin by defining the following notation:

$$\mathbb{P}(\omega_i = j) = \omega_i(j), \quad j = 0, 1, 2, \ldots,$$

and taking first-step arguments for a bacterium starting in state 1.

$$\omega_1(0) = \frac{\lambda_1}{\lambda_1 + \nu_{12} + \gamma_1 + \mu_1}\omega_1(0) + \frac{\nu_{12}}{\lambda_1 + \nu_{12} + \gamma_1 + \mu_1}\omega_2(0) + \frac{\mu_1}{\lambda_1 + \nu_{12} + \gamma_1 + \mu_1},$$

multiplying by $(\lambda_1 + \nu_{12} + \gamma_1 + \mu_1)$ and cancelling terms we get

$$\omega_1(0) = \frac{\nu_{12} \cdot \omega_2(0) + \mu_1}{\nu_{12} + \gamma_1 + \mu_1}. \tag{3.15}$$

By a similar process it can be shown that

$$\omega_2(0) = \frac{\nu_{21} \cdot \omega_1(0) + \mu_2}{\nu_{21} + \gamma_2 + \mu_2}. \tag{3.16}$$

By substituting Equation (3.16) into Equation (3.15) we obtain

$$\omega_1(0) = \frac{\nu_{12}\frac{\nu_{21}\cdot\omega_1(0)+\mu_2}{\nu_{21}+\gamma_2+\mu_2} + \mu_1}{\nu_{12} + \gamma_1 + \mu_1},$$

which becomes

$$\omega_1(0)(\nu_{12} + \gamma_1 + \mu_1)(\nu_{21} + \gamma_2 + \mu_2) = \nu_{12}\nu_{21}\omega_1(0) + \nu_{12}\mu_2 + \mu_1\nu_{21} + \mu_1\gamma_2 + \mu_1\mu_2.$$

As the $\nu_{12}\nu_{21}\omega_1(0)$ terms on each side will cancel, by factorising each side of the equation we can obtain a final equation for $\omega_1(0)$.

$$\omega_1(0) = \frac{\nu_{12}\mu_2 + \mu_1(\nu_{21} + \gamma_2 + \mu_2)}{\nu_{12}(\gamma_2 + \mu_2) + (\gamma_1 + \mu_1)(\nu_{21} + \gamma_2 + \mu_2)}, \tag{3.17}$$

whilst $\omega_2(0)$ is, similarly,

$$\omega_2(0) = \frac{\nu_{21}\mu_1 + \mu_2(\nu_{12} + \gamma_1 + \mu_1)}{\nu_{21}(\gamma_1 + \mu_1) + (\gamma_2 + \mu_2)(\nu_{12} + \gamma_1 + \mu_1)}. \tag{3.18}$$

A similar process can be used to get a formula for $\omega_1(1)$, but first recall the notation $\Delta_1 = \lambda_1 + \nu_{12} + \gamma_1 + \mu_1$, therefore by taking first-step arguments

$$\omega_1(1) = \frac{\lambda_1}{\Delta_1}\omega_1(1) + \frac{\nu_{12}}{\Delta_1}\omega_2(1) + \frac{\gamma_1}{\Delta_1}\omega_1(0) \Rightarrow \omega_1(1) = \frac{\nu_{12}\omega_2(1) + \gamma_1\omega_1(0)}{\nu_{12} + \gamma_1 + \mu_1}.$$

Using the symmetry between $\omega_2(1)$ and $\omega_1(1)$ and substituting in we get

$$\omega_1(1) = \frac{\nu_{12}\frac{\nu_{21}\omega_1(1) + \gamma_2\omega_2(0)}{\nu_{21} + \gamma_2 + \mu_2} + \gamma_1\omega_1(0)}{\nu_{12} + \gamma_1 + \mu_1},$$

and rearranging leads to the solution

$$\omega_1(1) = \frac{\nu_{12}\gamma_2\omega_2(0) + \gamma_1\left(\nu_{21} + \gamma_2 + \mu_2\right)\omega_1(0)}{(\gamma_1 + \mu_1)(\nu_{21} + \gamma_2 + \mu_2) + \nu_{12}(\gamma_2 + \mu_2)}.$$

In fact, we can show that the recursive relationship for $\omega_1(n)$ is:

$$\omega_1(n) = \frac{\gamma_1(\nu_{21} + \gamma_2 + \mu_2)\omega_1(n-1) + \nu_{12}\gamma_2\omega_2(n-1)}{(\gamma_1 + \mu_1)(\nu_{21} + \gamma_2 + \mu_2) + \nu_{12}(\gamma_2 + \mu_2)}.$$

It is very difficult to obtain an analytic solution for $\omega_1(n)$ for the general model with no constraints on the parameters, but the recursive solution can be implemented numerically.

3.3.3 Number of division events in the lifespan of a bacterium

Let us consider now the number of times that the tracked bacterium divides during its lifetime, D_i, if this bacterium is originally of type $i \in \{1, 2\}$. We can define its probability generating function as $\Phi_i(z) = \mathbb{E}[z^{D_i}]$ for $|z| \leq 1$. Via first-step arguments, $\Phi_i(z)$ satisfies the following equations:

$$\begin{aligned}
\Delta_1\Phi_1(z) &= \lambda_1 z\Phi_1(z) + \mu_1 + \gamma_1\Phi_1(z) + \nu_{12}\Phi_2(z), \\
\Delta_2\Phi_2(z) &= \lambda_2 z\Phi_2(z) + \mu_2 + \gamma_2\Phi_2(z) + \nu_{21}\Phi_1(z).
\end{aligned}$$

It can be shown, following a similar process to that in the previous section, that these equations have solutions

$$\begin{aligned}\Phi_1(z) = &\quad c(z)^{-1}\frac{1}{\mu_1+\nu_{12}+\lambda_1(1-z)}\left(\frac{\mu_2\nu_{12}}{\mu_2+\nu_{21}+\lambda_2(1-z)}+\mu_1\right),\\ \Phi_2(z) = &\quad c(z)^{-1}\frac{1}{\mu_2+\nu_{21}+\lambda_2(1-z)}\left(\frac{\mu_1\nu_{21}}{\mu_1+\nu_{12}+\lambda_1(1-z)}+\mu_2\right),\end{aligned}$$

with $c(z) = 1 - \frac{\nu_{12}\nu_{21}}{(\mu_1+\nu_{12}+\lambda_1(1-z))(\mu_2+\nu_{21}+\lambda_2(1-z))}$. We note that these expressions, as one would expect, do not depend on the toxin production rate, γ_i. Following a similar process to the one above, and using the formula

$$\frac{d^k}{dz^k}\Phi_i(z)\bigg|_{z=1} = \mathbb{E}[D_i^k],$$

it can be shown that the average number of divisions are then given by

$$\begin{aligned}\mathbb{E}[D_1] = &\quad c(1)^{-1}\frac{1}{\mu_1+\nu_{12}}\left(\frac{\lambda_2\nu_{12}}{\mu_2+\nu_{21}}+\lambda_1\right),\\ \mathbb{E}[D_2] = &\quad c(1)^{-1}\frac{1}{\mu_2+\nu_{21}}\left(\frac{\lambda_1\nu_{21}}{\mu_1+\nu_{12}}+\lambda_2\right).\end{aligned}$$

Once again, scenarios of interest might lead to simplified expressions. If one sets $\nu_{21} = 0$, as in the case of ν_{12} representing a non-reversible antibiotic treatment as in Section 3.4.2, this leads to

$$\begin{aligned}\mathbb{E}[D_1] &= \frac{1}{\mu_1+\nu_{12}}\left(\frac{\lambda_2\nu_{12}}{\mu_2}+\lambda_1\right),\\ \mathbb{E}[D_2] &= \frac{\lambda_2}{\mu_2}.\end{aligned}$$

This choice implies that D_2 is a geometrically distributed random variable, where the probability of death (thought of as success in this case) is $\frac{\mu_2}{\lambda_2+\mu_2}$, $D_2 \sim \mathrm{Geo}(\frac{\mu_2}{\mu_2+\lambda_2})$. This is consistent with the expression of $\mathbb{E}[D_2]$.

3.3.4 Number of bacteria in the progeny of a bacterium

We focus now on the random variable describing the number of bacteria in the progeny of the original bacterium. We denote this number as G_i, with i indicating the original bacterium type. To be precise, the initial type i bacterium is not included in G_i, we consider a change of state ($1 \to 2$ or vice versa) as a new bacterium being formed in the new state and thus do include these transitions as the creation of a new progeny. When a bacterium divides, we count both daughter bacteria as progeny. As an example, $G_1 = 15$ in the particular realisation depicted in Figure 3.2. We can find the expectation of this descriptor, again by using first-step arguments.

$$\mathbb{E}[G_1] = \frac{\lambda_1}{\Delta_1}\left(2 + 2\mathbb{E}[G_1]\right) + \frac{\nu_{21}}{\Delta_1}\left(\mathbb{E}[G_2] + 1\right) + \frac{\gamma_1}{\Delta_1}\mathbb{E}[G_1],$$

where once again $\Delta_i = \lambda_i + \nu_{ij} + \gamma_i + \mu_i$, $i, j \in \{1, 2\}, j \neq i$. This rearranges to

$$\mathbb{E}[G_1](\mu_1 + \nu_{12} - \lambda_1) = 2\lambda_1 + \nu_{12}(\mathbb{E}[G_2] + 1). \tag{3.19}$$

We can also find a similar equation for $\mathbb{E}[G_2]$

$$\mathbb{E}[G_2](\mu_2 + \nu_{21} - \lambda_2) = 2\lambda_2 + \nu_{21}(\mathbb{E}[G_1] + 1). \tag{3.20}$$

These quantities will be positive and finite only if $\mu_1 + \nu_{12} > \lambda_1$ and $\mu_2 + \nu_{21} > \lambda_2$, which are conditions that must be satisfied for the number of cells in the progeny to be finite. We can find analytic solutions for $\mathbb{E}[G_1]$ and $\mathbb{E}[G_2]$ by substituting these equations into one another, leading to

$$\begin{aligned}
\mathbb{E}[G_1] &= g^{-1}\frac{1}{\mu_1 + \nu_{12} - \lambda_1}\left(2\lambda_1 + \nu_{12}\frac{\lambda_2 + 2\nu_{21} + \mu_2}{\mu_2 + \nu_{21} - \lambda_2}\right), \\
\mathbb{E}[G_2] &= g^{-1}\frac{1}{\mu_2 + \nu_{21} - \lambda_2}\left(2\lambda_2 + \nu_{21}\frac{\lambda_1 + 2\nu_{12} + \mu_1}{\mu_1 + \nu_{12} - \lambda_1}\right),
\end{aligned}$$

with

$$g = 1 - \frac{\nu_{12}\nu_{21}}{(\mu_1 + \nu_{12} - \lambda_1)(\mu_2 + \nu_{21} - \lambda_2)}.$$

In order for these averages to be positive, we also require $g > 0$. This leads to a third condition for a positive and finite expected number of bacteria in the progeny (i.e. to

avoid unlimited bacterial population growth)

$$\frac{\nu_{12}\nu_{21}}{(\mu_1+\nu_{12}-\lambda_1)(\mu_2+\nu_{21}-\lambda_2)}<1.$$

For the specific case when $\nu_{21}=0$, one gets

$$\mathbb{E}[G_1]=\hat{G}_1=\frac{1}{\mu_1+\nu_{12}-\lambda_1}\left(2\lambda_1+\nu_{12}\frac{\lambda_2+\mu_2}{\mu_2-\lambda_2}\right),$$
$$\mathbb{E}[G_2]=\hat{G}_2=\frac{2\lambda_2}{\mu_2-\lambda_2}.$$

3.3.5 Number of toxin molecules produced by the genealogy of a bacterium

Our interest is to mathematically describe a system of toxin-producing bacteria, thus, we now compute the number of toxin molecules produced by the genealogy of the original bacterium. To be precise, we consider the genealogy of a bacterium to consist of the initial bacterium and all of its progeny. We then introduce Ω_i as the number of toxin molecules produced by the genealogy of an initial type i bacterium. We denote its expectation value by $\hat{\Omega}_i=\mathbb{E}[\Omega_i]$, for $i\in\{1,2\}$. We note that the number of toxin molecules produced by the genealogy of the single bacterium will be finite if and only if the number of bacteria within the genealogy is finite. This condition is equivalent to the number of progeny of a bacterium being finite. Therefore the conditions for a finite number of bacteria in the progeny discussed in the previous section also need to be satisfied for this descriptor. The expected values, $\hat{\Omega}_1$ and $\hat{\Omega}_2$, satisfy

$$\begin{aligned}(\mu_1+\nu_{12}-\lambda_1)\hat{\Omega}_1 &= \gamma_1+\nu_{12}\hat{\Omega}_2,\\ (\mu_2+\nu_{21}-\lambda_2)\hat{\Omega}_2 &= \gamma_2+\nu_{21}\hat{\Omega}_1,\end{aligned}$$

with solutions

$$\begin{aligned}\hat{\Omega}_1 &= g^{-1}\frac{1}{\mu_1+\nu_{12}-\lambda_1}\left(\gamma_1+\nu_{12}\frac{\gamma_2}{\mu_2+\nu_{21}-\lambda_2}\right),\\ \hat{\Omega}_2 &= g^{-1}\frac{1}{\mu_2+\nu_{21}-\lambda_2}\left(\gamma_2+\nu_{21}\frac{\gamma_1}{\mu_1+\nu_{12}-\lambda_1}\right).\end{aligned}$$

When $\nu_{21} = 0$, the equations simplify to

$$\hat{\Omega}_1 = \frac{1}{\mu_1 + \nu_{12} - \lambda_1}\left(\gamma_1 + \nu_{12}\frac{\gamma_2}{\mu_2 - \lambda_2}\right),$$
$$\hat{\Omega}_2 = \frac{\gamma_2}{\mu_2 - \lambda_2}.$$

We note that there exists a link between the expected number of toxin molecules produced by a bacterium, the expected number of bacteria in the genealogy of a bacterium (the number of progeny and the initial bacterium itself) and the number of toxins produced by the genealogy. For instance, if we take the case when $\nu_{21} = 0$, the average number of bacteria in the genealogy of an original type 2 bacterium including this original bacterium, is

$$\mathbb{E}[G_2] + 1 = \frac{2\lambda_2}{\mu_2 - \lambda_2} + 1 = \frac{\mu_2 + \lambda_2}{\mu_2 - \lambda_2},$$

see Section 3.3.4 for further detail. It is clear that, in this case, every bacterium in the genealogy will be type 2 bacteria since $\nu_{21} = 0$. Each of these type 2 bacteria will produce, on average

$$\frac{\gamma_2}{\lambda_2 + \mu_2},$$

toxins before they decide their fate, which could be division or death. Note this is not equivalent to $\mathbb{E}[\omega_2]$ when $\nu_{21} = 0$ from Section 3.3.2; as here we consider the number of toxin molecules produced before division or death whereas ω_2 counts the number of toxin molecules before death, since after a division event one of the daughter cells was considered to be the original bacterium. Thus, the mean number of toxin molecules produced by the genealogy of the original type 2 bacterium is

$$\frac{\mu_2 + \lambda_2}{\mu_2 - \lambda_2} \times \frac{\gamma_2}{\lambda_2 + \mu_2} = \frac{\gamma_2}{\mu_2 - \lambda_2},$$

which matches the above expression for $\hat{\Omega}_2$.

3.4 Numerical results

In this section we shall consider the effect that relevant parameters have on the stochastic descriptors described above. We will first carry out a global sensitivty analysis, to

Fixed Parameter	Parameter Value
λ_1	0.8
μ_1	1
γ_1	1
λ_2	0.8
μ_2	1
γ_2	1

Table 3.1: The baseline values for the parameters concerning type 1 and type 2 bacteria within this model. All rates have units hour^{-1}.

investigate the effect that the relationship between certain parameters can have on the summary statistics analysed in Section 3.3. Once we have investigated this, we will look at a case study of *B. anthracis* using *in-vitro* data from Zai *et al.* (2016) to try to find estimates for the parameters relevant to state 1 (untreated bacteria) and then investigate the effect of varying the parameters relevant to state 2 (antibiotic-bound bacteria).

3.4.1 Global sensitivity analysis

Within this section we shall impose no constraints on our parameters, and investigate the effect that changing the relationships between our parameters can have on the summary statistics analysed in Section 3.3. The conclusions we draw here will inform the parameter regimes we will investigate when applying our model to the *B. anthracis* case. In Table 3.1 the baseline values for parameters in the model are shown. Parameter values for a type 1 bacterium were chosen, and we assume that unless being investigated, the differentiation from a type 1 bacterium to a type 2 bacterium has no effect on the corresponding parameters. The value of $\lambda_1 = 0.8$ hour^{-1}, was chosen as it is the given extracellular bacterial growth rate of *B. anthracis* in Day *et al.* (2011). We chose here $\mu_1 = 1$ hour^{-1} for illustrative purposes. $\gamma_1 = 1$ hour^{-1} was chosen for illustrative purposes so that on average a type 1 bacterium would produce 1 toxin within its lifespan. As the rates of differentiation, ν_{12} and ν_{21}, are varied in most plots, no baseline value for these parameters are set.

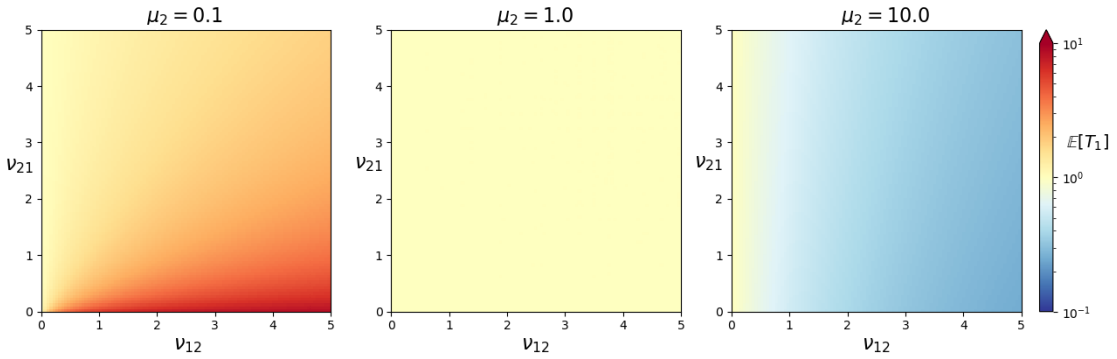

Figure 3.3: Average lifespan of a type 1 bacterium, $\mathbb{E}[T_1]$, for varying values of ν_{12} and ν_{21} with different values of μ_2. Parameter units are hour^{-1}.

3.4.1.1 Lifespan of a bacterium

We will investigate the effect of parameters on $\mathbb{E}[T_1]$. Firstly, we note that the division rates, λ_i, and the toxin production rates, γ_i, for $i \in \{1, 2\}$, have no effect on $\mathbb{E}[T_1]$. This makes sense intuitively; if we are interested in one particular bacterium it is of no consequence for its average lifespan if it makes many daughter bacteria or none. The same logic applies to the number of toxins it produces. However, it is important for us to investigate the effect that the death rates in each state and the differentiation rates have. Figure 3.3 shows the effect that varying ν_{12}, ν_{21} has for $\mu_2 < \mu_1$, $\mu_2 = \mu_1$ and $\mu_2 > \mu_1$. As is to be expected, $\mu_2 > \mu_1$ shows a much shorter lifespan for the bacterium as it is more likely to die as a type 2 bacterium. It is important to note that in the middle plot, when $\mu_2 = \mu_1$, the average lifespan of the bacterium is fixed at $\mathbb{E}[T_1] = 1$. This is to be expected as under those parameter values the bacterium dies at the same rate whether it is a type 1 or type 2 bacterium. It is also important to note that when $\nu_{12} = 0$, the bacterium will always stay as a type 1 bacterium and as such both ν_{21} and μ_2 have no effect on the system; this explains the behaviour shown where as ν_{12} gets close to 0 on all plots, $\mathbb{E}[T_1]$ gets close to 1. The largest values of $\mathbb{E}[T_1]$ are in the region where $\mu_2 = 0.1$, ν_{12} is large and ν_{21} is small (corresponding to the bottom-right of the left plot). This is the case where the bacterium dies slower as a type 2 bacterium, differentiates from a type 1 to a type 2 bacterium quickly but differentiates back from a type 2 to a type 1 bacterium slowly; this results in the bacterium spending more time as a type 2 bacterium where it expires at a slower rate so it makes sense this gives the largest expected lifespan.

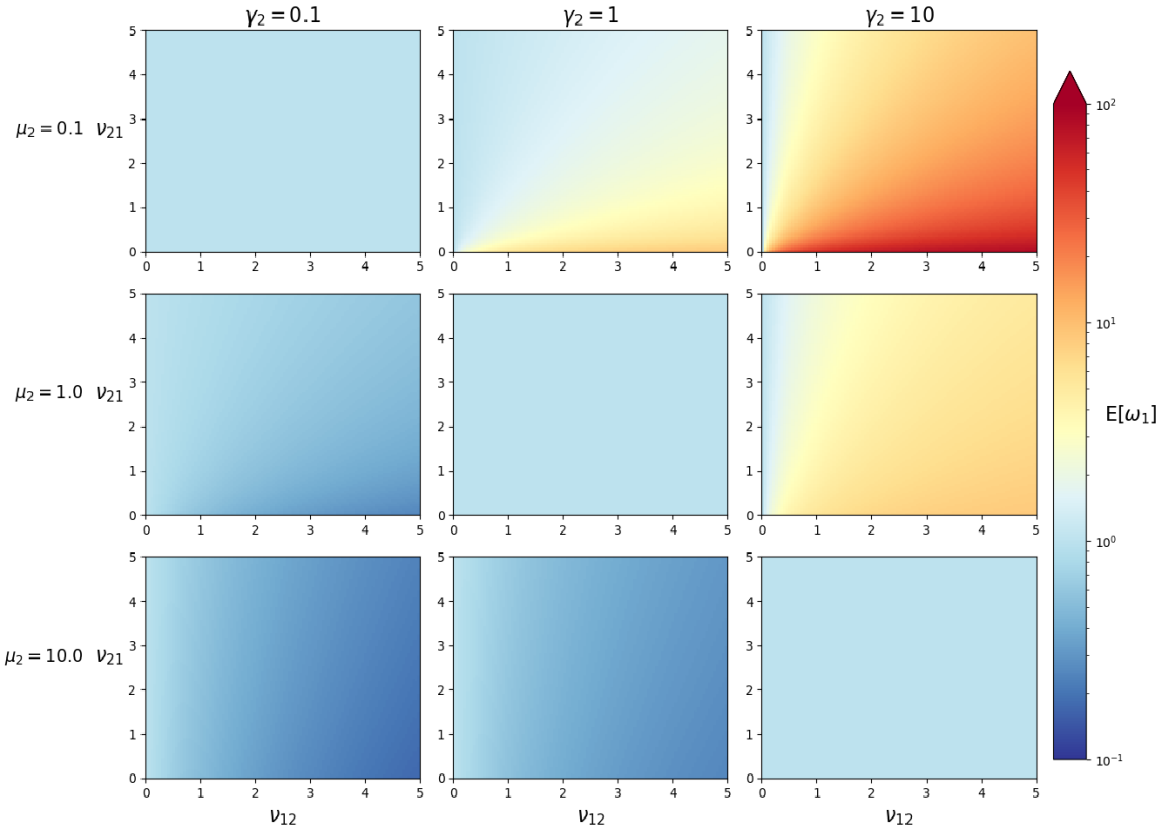

Figure 3.4: Average number of toxins produced in the lifespan of a type 1 bacterium, $\mathbb{E}[\omega_1]$, for varying values of $\gamma_2, \mu_2, \nu_{12}$ and ν_{21}. Parameter units are in hours^{-1}.

3.4.1.2 Toxins produced by a bacterium during its lifespan

Here, we focus on the average number of toxins produced by a type 1 bacterium during its lifespan, $\mathbb{E}[\omega_1]$. We retain our baseline values as shown in Table 3.1, with $\gamma_1 = 1$ chosen for illustrative purposes so that a type 1 bacterium produces an average of 1 toxin during its lifespan under no differentiation.

In Figure 3.4, we plot heatmaps of $\mathbb{E}[\omega_1]$ for varying values of $\gamma_2, \mu_2, \nu_{12}$ and ν_{21}. Here we can see the impact that changing our parameters has on the expectation of ω_1. We can see that increasing γ_2 has a dramatic effect on $\mathbb{E}[\omega_1]$, especially in the cases where ν_{12} is large, ν_{21} is small and μ_2 is small; this makes sense as those parametric regions represent the cases where the bacterium will spend a long time in state 2, and therefore γ_2 will be more important for $\mathbb{E}[\omega_1]$ than γ_1. This explains the interesting behaviour

where in the bottom-left plot the bottom-right region has the smallest values but in the top-right plot the bottom-right region has the largest values. In the middle plot we get the behaviour that if the behaviour of a type 1 bacterium and a type 2 bacterium is identical the rates of transition between these two types have no effect on the expected number of toxins. We also see this behaviour in the top-left and bottom-right figures; therefore we can note that as long as the ratio of γ_2 and μ_2 is the same as the ratio of γ_1 and μ_1 the rates of transition between type 1 and type 2 bacteria have no impact on the distribution of toxins. This makes sense as the descriptor ω_1 is timeless; it concentrates only on the number of toxins produced, regardless of how quickly this happens and therefore the number of toxins expected to be produced will not change if toxins are produced 100 times quicker but if the bacterium is expected to die 100 times quicker. In all three of these plots (in the top-left to bottom-right diagonal of Figure 3.4), the expected number of toxins $\mathbb{E}[\omega_1] = \frac{\gamma_1}{\mu_1} = \frac{\gamma_2}{\mu_2} = 1$. This is also the case when $\nu_{12} = 0$ and the bacterium will never differentiate to a type 2 bacterium, explaining the behaviour at $\nu_{12} = 0$ in all nine plots.

One can instead focus on the probability distribution of ω_1. In Figure 3.5 we plot the probability distribution, $\{\omega_1(n), n = 0, 1, 2, \dots\}$ for values up to $n = 6$. We notice that as expected, when the proportion of $\frac{\gamma_2}{\mu_2} = 1$, the same as $\frac{\gamma_1}{\mu_1}$, the probability distribution is not affected by the rates of transition between the two types of bacteria. We notice that, as expected, when γ_2 is smaller than μ_2 the probability of producing 0 toxins increases even if ν_{21} is large. This is because even if the bacterium is type 2 for a very small time; within that small time it be more likely to die before producing any toxin molecules than if it was a type 1 bacterium for that time. Naturally the converse is true; if $\gamma_2 > \mu_2$ the probability of producing 0 toxins decreases. It is also clear that the impact of the transition rate of a type 2 bacterium becoming a type 1 bacterium, ν_{21}, on the probability of 0 toxins being produced depends on this proportion $\frac{\gamma_2}{\mu_2}$. Namely, when $\frac{\gamma_2}{\mu_2} < \frac{\gamma_1}{\mu_1}$, such as in the plot in the second row and first column, $\mu_2 = 1, \gamma_2 = 0.1$, increasing the transition rate ν_{21} makes it less likely that 0 toxins will be produced. This makes sense as an increase in ν_{21} means that a bacterium will, on average, spend less time as a type 2 bacterium; therefore if a type 2 bacterium produces less toxins than a type 1 bacterium, as in this case, the chance of producing 0 toxins decreases. The converse is true, when $\frac{\gamma_2}{\mu_2} > \frac{\gamma_1}{\mu_1}$, such as in the plot in the second row, third column.

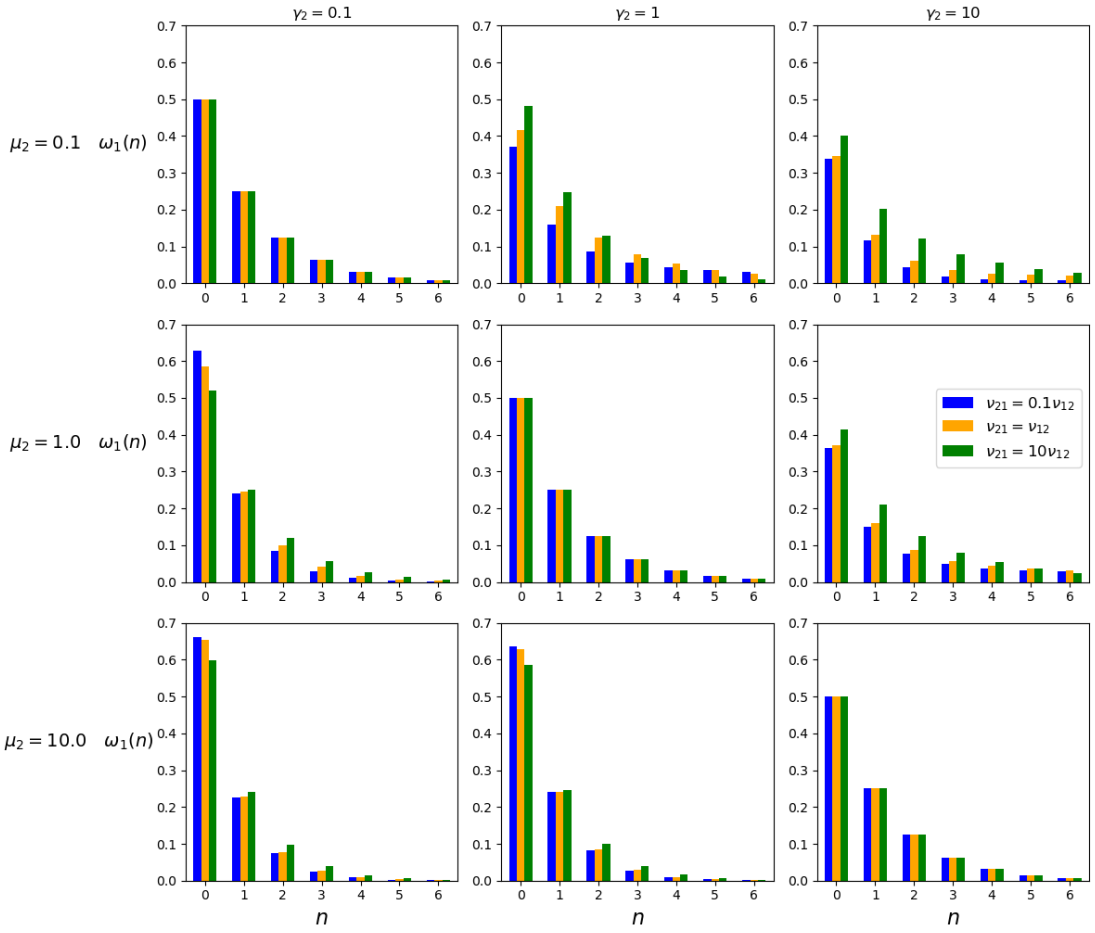

Figure 3.5: The probability distribution of the number of toxins in the lifespan of a bacterium, $\{\omega_1(n), n = 0, 1, 2, \dots\}$ for values up to $n = 6$, for varying values of γ_2, μ_2 and ν_{21}. In this plot ν_{12} is fixed to a value of 1 for illustrative purposes. Parameter units are in hours^{-1}.

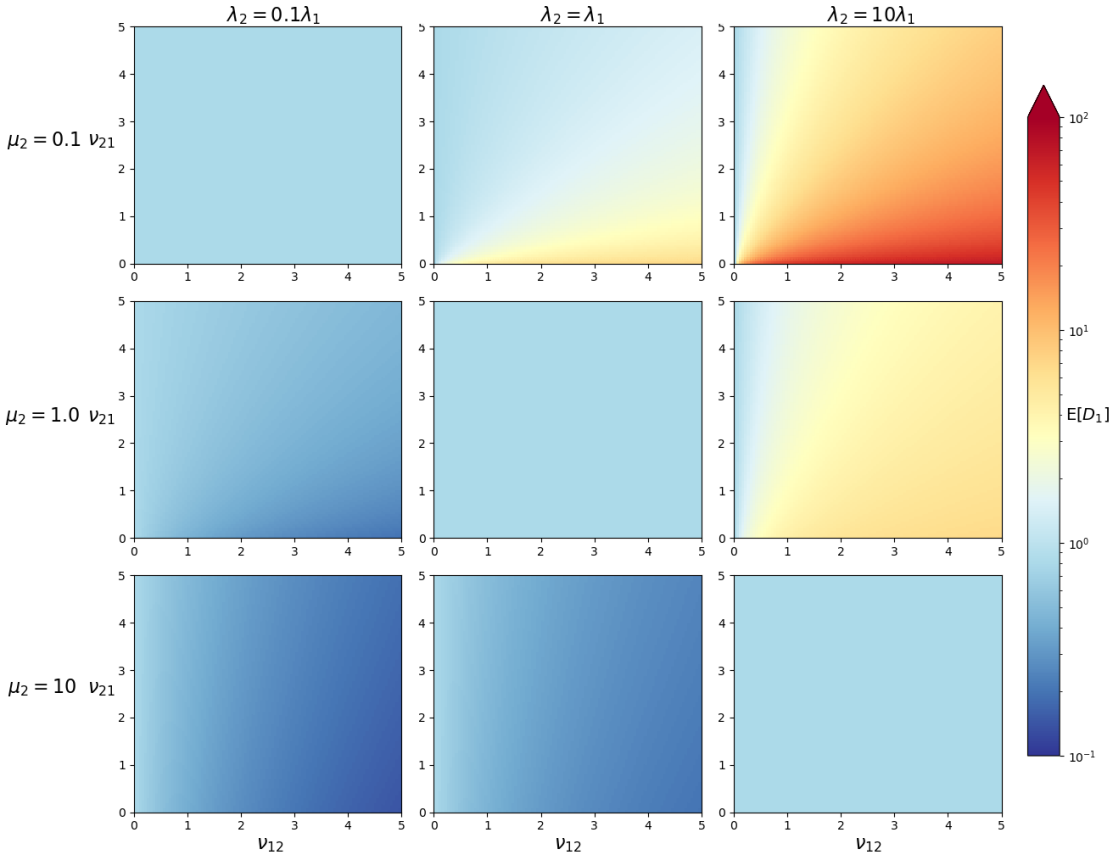

Figure 3.6: Average number of divisions undergone by a type 1 bacterium during its lifespan, $\mathbb{E}[D_1]$, for varying values of $\lambda_2, \mu_2, \nu_{12}$ and ν_{21}. Parameter units are in hours^{-1}.

3.4.1.3 Divisions undertaken by a bacterium during its lifespan

Within Figure 3.6 we plot heatmaps of $\mathbb{E}[D_1]$, the average number of divisions in a type 1 bacterium's lifespan, for different values of $\lambda_2, \mu_2, \nu_{12}$ and ν_{21}. We do not vary γ_2 as the number of toxins produced by a bacterium has no effect on the number of divisions it will undertake. Here we see behaviour we expect from analysis of previous summary statistics. For example, we see that if the ratio of the division rate and death rate for a type 2 bacterium, $\frac{\lambda_2}{\mu_2}$, is equal to the ratio of the division and death rates for a type 1 bacterium, $\frac{\lambda_1}{\mu_1}$, the transition rates, ν_{12} and ν_{21}, do not impact the number of divisions a bacterium would be expected to undergo during its lifespan. This makes sense as $\mathbb{E}[D_1]$ is a timeless quantity and as such, the time it takes for the bacterium to divide and die is not important, only the number of divisions before death is. It is also clear that as $\frac{\lambda_2}{\mu_2}$

increases in relation to $\frac{\lambda_1}{\mu_1}$, we expect to see more divisions in the bacterium's lifespan. In particular, for the case that $\frac{\lambda_2}{\mu_2} > \frac{\lambda_1}{\mu_1}$, increasing the type 1 bacterium to a type 2 bacterium transition rate ν_{12} causes more divisions as this will cause the bacterium to spend more time as a type 2 where more divisions occur. Again, the converse is true, when $\frac{\lambda_2}{\mu_2} < \frac{\lambda_1}{\mu_1}$, increasing the transition rate ν_{12} causes the expected number of divisions to decrease via a similar logic.

3.4.1.4 Number of bacteria in the progeny of a bacterium

Here, we consider the expected number of bacteria in the progeny of a type 1 bacterium, $\mathbb{E}[G_1] = \hat{G}_1$. Recall that the three criteria for a finite number of cells in the progeny are:

$$\begin{aligned} \mu_1 + \nu_{12} &> \lambda_1, \\ \mu_2 + \nu_{21} &> \lambda_2, \\ \frac{\nu_{12}\nu_{21}}{(\mu_1 + \nu_{12} - \lambda_1)(\mu_2 + \nu_{21} - \lambda_2)} &< 1. \end{aligned} \tag{3.21}$$

We can therefore only investigate $\hat{G}_1$ for parameter regions where these criteria are satisfied. Figure 3.7 shows the expected number of bacteria in the progeny of a type 1 bacterium, $\hat{G}_1$, for a variety of values for $\lambda_2, \mu_2, \nu_{12}$ and ν_{21}. γ_2 is not investigated as toxin production has no effect on the number of bacteria in the progeny. Firstly, we note that as parameter values approach the threshold of the criteria for a finite number of bacteria in the progeny the expected number of bacteria in the progeny becomes large; this is to be expected as at the threshold the expected number of bacteria in the progeny will be infinite. Secondly, as we would expect from our analysis of the expected number of divisions $\mathbb{E}[D_1]$, increasing the proportion $\frac{\lambda_2}{\mu_2}$, causes the expected number of bacteria in the progeny to increase. Likewise, the converse is true. In the regions where λ_2 is smaller than μ_2, increasing the transition rate of a type 1 bacterium becoming a type 2 bacterium, ν_{12}, compared to the rate at which a type 2 bacterium becomes a type 1 bacterium, ν_{21}, causes the average number of bacteria in the progeny to decrease. However, one thing to note for all of these plots is that the average number of bacteria in the progeny, $\hat{G}_1$, does not solely depend on the ratio of division rates against death rates; this is as we count each transition (e.g a type 1 bacterium becoming a type 2 bacterium) as one daughter bacterium being formed and as such this is added to the number of

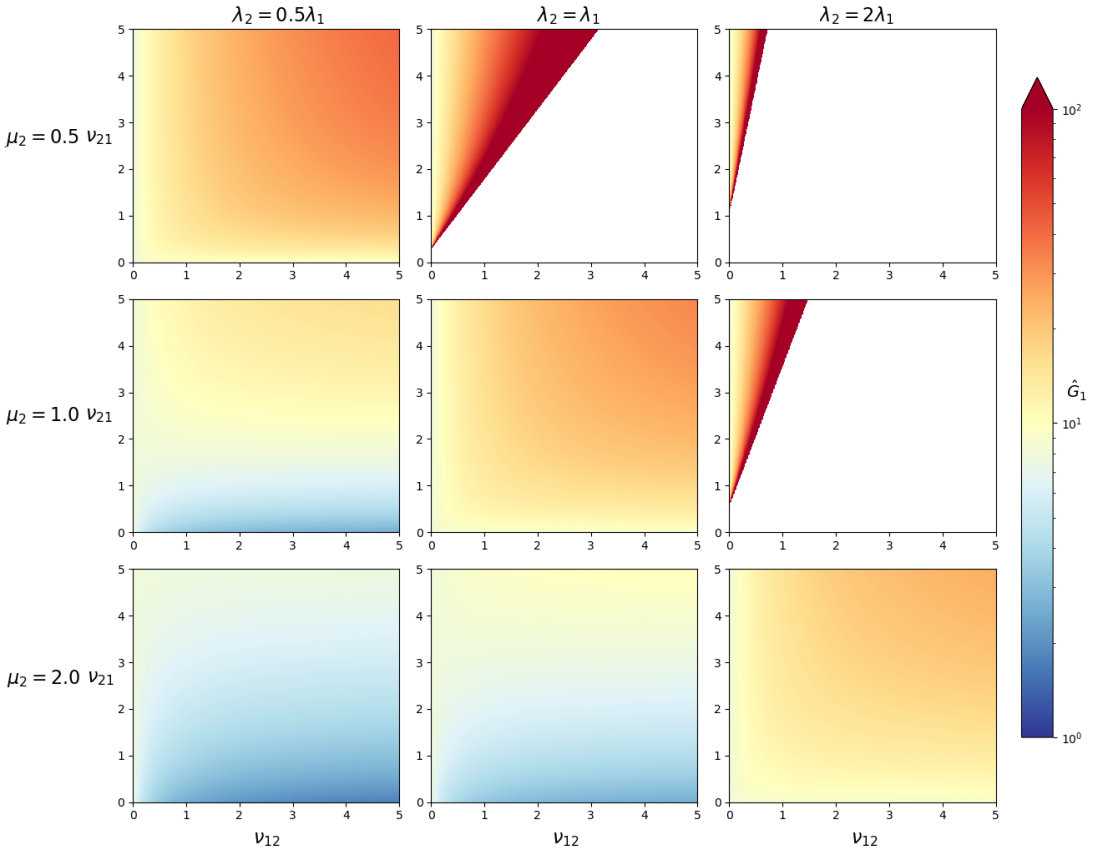

Figure 3.7: The expected value of the number of bacteria in the progeny for a given type 1 bacterium, $\hat{G}_1$, for various values of $\lambda_2, \mu_2, \nu_{12}$ and ν_{21}. Parameter units are in hours^{-1}. The unplotted regions are where the criteria for a positive and finite number of bacteria in the progeny are not satisfied.

bacteria in the progeny. This is particularly noticeable in the cases where the proportion of $\frac{\lambda_2}{\mu_2} = \frac{\lambda_1}{\mu_1}$ (the top-left to bottom-right diagonal in Figure 3.7) as in previous summary statistics we would expect these plots to be transition rate homogenous; however for this stochastic descriptor as both ν_{12} and ν_{21} increase so too does $\hat{G}_1$. This has a more pronounced effect when ν_{12} and ν_{21} are larger than the rates of division and death, which explains why the case where $\lambda_2 = 0.5\lambda_1$ and $\mu_2 = 0.5$ (top-left in Figure 3.7) has higher values for $\hat{G}_1$ for large values of ν_{12} and ν_{21} than the corresponding values of ν_{12}, ν_{21} for the case where $\lambda_2 = 2\lambda_1$ and $\mu_2 = 2$ (bottom-right in Figure 3.7).

3.4.1.5 Number of toxins produced by a bacterium and its progeny

Finally, here we consider the average number of toxins produced by a type 1 bacterium and all of its progeny, $\mathbb{E}[\Omega_1]$. Again, this quantity is only valid if the criteria in Equations (3.21) are satisfied. Figure 3.8 shows the expected number of toxins produced by a type 1 bacterium and its progeny, $\mathbb{E}[\Omega_1]$, for varying values of $\lambda_2, \mu_2, \nu_{12}$ and ν_{21}. This figure is the counterpart to Figure 3.7 and as such shows similar behaviour, where the larger the expected size of the progeny, the higher the expected number of toxins are and vice-versa. There is different behaviour seen across the top-left to bottom-right diagonal in Figure 3.8 compared to Figure 3.7 however. The number of toxins produced by the genealogy is not affected by the transition rates ν_{12} and ν_{21} in the middle plot; this is because the division and death rates of type 1 and type 2 bacteria are identical in this case so having more or less transitions between the two types of bacteria will not affect the number of toxins produced. In the top-left plot of Figure 3.8, Ω_1 increases as ν_{12} increases and as ν_{21} increases, Ω_1 decreases, whereas in the bottom-right plot the opposite occurs. This can be explained as the overall toxin production will be larger when the bacteria spend more time in a state where the toxin production rate is larger relative to the division and death rates, corresponding to $\nu_{12} > \nu_{21}$ in the top-left plot and $\nu_{21} > \nu_{12}$ in the bottom-right plot.

A limitation of this figure is that we are investigating the number of toxins produced, without varying the relevant parameters, γ_1 and γ_2. Therefore, in Figure 3.9 we instead consider varying γ_1 and γ_2 along the x-axis and y-axis of each plot. This requires choosing fixed values of ν_{12} and ν_{21}. So with a view to an application to *B. anthracis* we select $\nu_{12} = 0.675$ hour^{-1} for illustrative purposes (it is the value given to the antibiotic

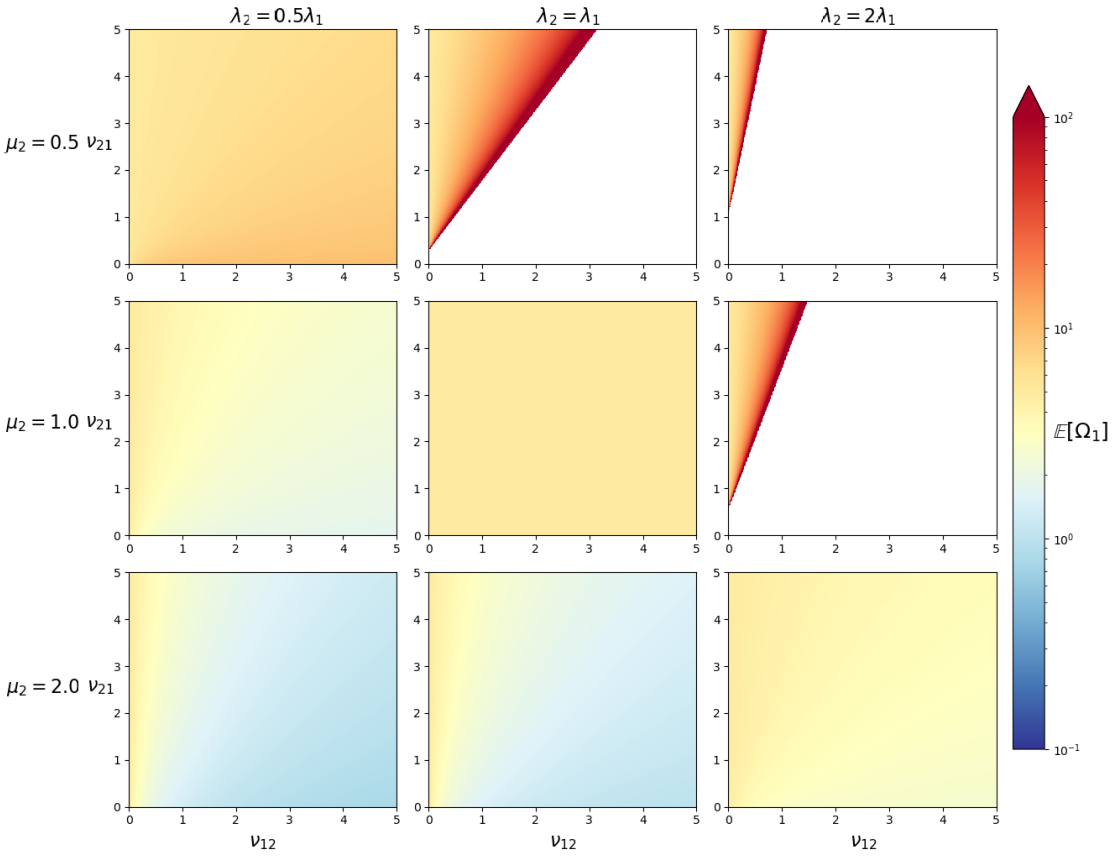

Figure 3.8: The expected value of the number of toxins produced by a type 1 bacterium and its progeny, $\mathbb{E}[\Omega_1]$, for various values of $\lambda_2, \mu_2, \nu_{12}$ and ν_{21}. The toxin production rate for a type 2 bacterium, γ_2, was chosen to be equal to the toxin production rate of a type 1 bacterium, $\gamma_1 = 1$. Parameter units are in hours^{-1}. The unplotted regions are where the criteria for a positive and finite number of bacteria in the progeny are not satisfied.

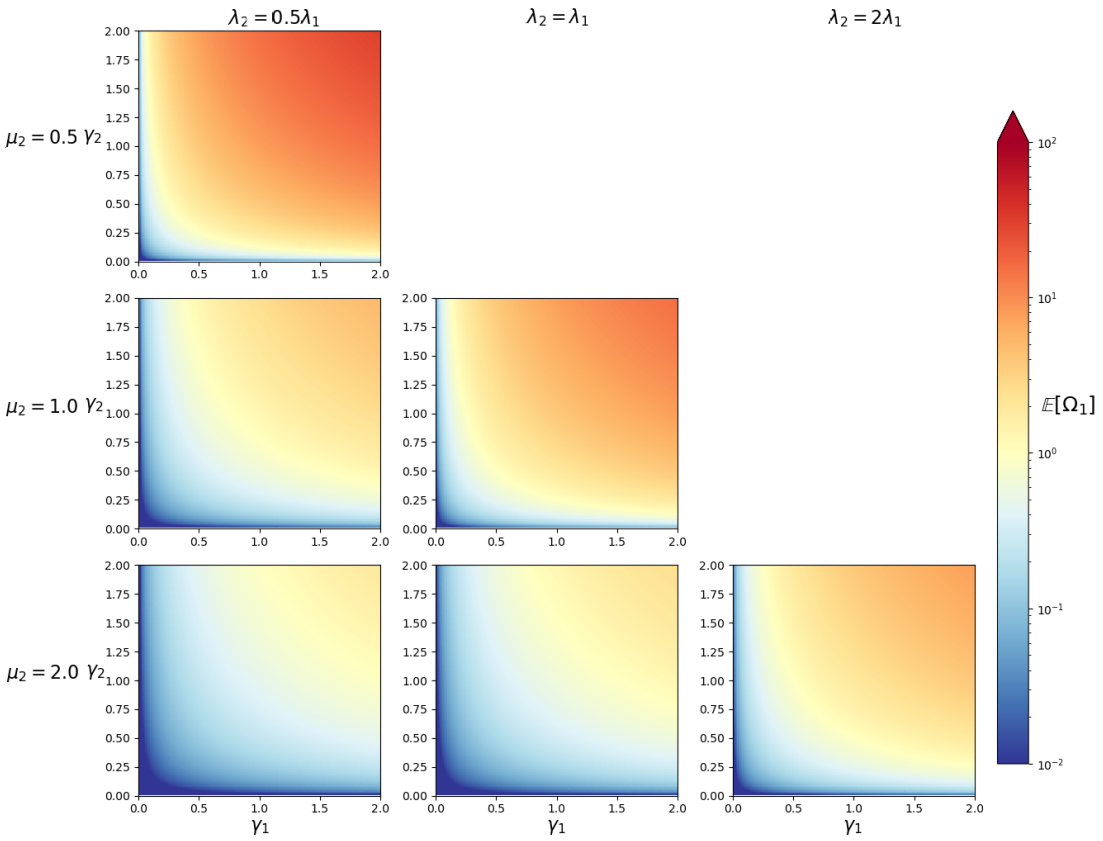

Figure 3.9: The expected value of the number of toxins produced by a type 1 bacterium and its progeny, $\mathbb{E}[\Omega_1]$, for various values of $\lambda_2, \mu_2, \gamma_1$ and γ_2. Parameter units are in hours^{-1}. Transition rates ν_{12}, and ν_{21} have fixed values of 0.675 and 0 respectively.

Ciprofloxacin contact rate by Day *et al.* (2011)) and $\nu_{21} = 0$, representing irreversible antibiotic-binding to the bacteria. Three axes in Figure 3.9 do not have plots as these parameter regions do not satisfy the second criteria of Equations (3.21). As is to be expected, as the parameters concerning toxin production, γ_1 and γ_2, increase we see a marked rise in $\mathbb{E}[\Omega_1]$. It is also significant that the rates of λ_2 and μ_2 have notable impact even if their ratio remains the same. This is explained by the toxin production rate becoming dominant as it increases if λ_2 and μ_2 are not sufficiently large. It is also of note how sensitive $\mathbb{E}[\Omega_1]$ is to the value of γ_1 and γ_2, ranging from 10^{-2} to 10^1 in each plot. This informs us that it will be crucial to find an acceptable value for γ_1 and γ_2 in our case study in Section 3.4.2.

3.4.2 Case study *B. anthracis*

We now make use of the previous results to analyse the behaviour of *Bacillus anthracis* bacteria, which causes anthrax infection, in the presence of antibiotic treatment. We assume non-treated fully vegetative *Bacillus anthracis* bacteria form the B_1 compartment in Figure 3.1, and the second compartment, B_2, will represent bacteria affected by the antibiotic. *B. anthracis* produces three anthrax exotoxin components (Liu *et al.* (2014)): protective antigen (PA), lethal factor (LF) and edema factor (EF). The effectiveness of the anthrax toxins in entering cells and causing symptoms is mainly due to the protective antigen (PA) (Leppla (2000)), with which the other toxin components can form complexes (Liu *et al.* (2014)). Therefore, we focus here on the production of PA when implementing our methods. We consider an antibiotic treatment, such as Ciprofloxacin, that inhibits bacterial division and triggers cellular death, so that $\mu_2 \geq \mu_1$ and $\lambda_2 = 0$. It is to be expected that the production rate of toxin molecules by antibiotic-treated cells would be at most equal to non-treated cells, and thus, we consider $\gamma_2 \leq \gamma_1$. Bacteria become treated at some rate ν_{12}, and we set $\nu_{21} = 0$ to indicate that the process is irreversible. In Section 3.4.2.1 we leverage data from an *in vitro* assay for the A16R strain of *B. anthracis* (Zai *et al.* (2016)) to inform our choice of parameters $(\lambda_1, \mu_1, \gamma_1)$. On the other hand, a global sensitivity analysis of model parameters $(\nu_{12}, \mu_2, \gamma_2)$ allows us in Section 3.4.2.2 to study the impact of treatment on the summary statistics introduced and analysed in Section 3.3, illustrating the applicability of our techniques.

3.4.2.1 Parameter calibration

Zai *et al.* (2016) examine the growth of the A16R *B. anthracis* strain by measuring the viable count of colony forming units (CFU) per mL in an *in vitro* assay for the following time points: $t \in \{4\text{h},8\text{h},12\text{h},16\text{h},20\text{h}\}$. They also develop a sandwich ELISA and cytotoxicity-based method to quantify the concentration of PA every two hours during the experiment, from $t = 4$h to $t = 26$h. In order to exploit this data set, and to estimate representative values for λ_1, μ_1 and γ_1, we consider its corresponding deterministic model (for the first compartment of non-treated bacteria)

$$\frac{dB}{dt} = (\lambda_1 - \mu_1)B, \quad \frac{dT}{dt} = \gamma_1 B - \xi T,$$

where $B(t)$ is the concentration [CFU/mL] of bacteria at time t, and $T(t)$ the concentration [ng/mL] of PA. Experimental results shown in Figure 1 of Zai *et al.* (2016) detected bacterial exponential growth during the first 12 hours of the experiment. The bacterial population reaches a carrying capacity after this point, representing competition for resources. Thus, since our interest (see Figure 3.1) is in the analysis of non-competing bacteria, we focus here on the first period of the experiment: $t \in [4\text{h}, 12\text{h}]$. In particular, we set $\lambda_1 = 0.8$ hour^{-1} from Day *et al.* (2011), and use bacterial counts shown in Figure 1 of Zai *et al.* (2016) and toxin concentration measurements from Figure 4 of Zai *et al.* (2016) to estimate the bacterial death rate, μ_1, and the toxin production rate, γ_1. Since the dynamics of the toxin population is likely to be dominated by the production of toxins from an exponentially growing bacterial population, we neglect toxin degradation and set $\xi = 0$ in what follows. We acknowledge that this might lead to underestimating the rate γ_1. Yet, we note that the rate ξ has no effect on any of the summary statistics analysed in Section 3.3.

Parameters μ_1 and γ_1 are estimated by using the `Python` programming language, specifically the `curve_fit` function from the `scipy.optimize`, which is based on a non-linear least squares method. This leads to point estimates $\mu_1 = 0.8 - 0.4277 = 0.3723$ hour^{-1} and $\gamma_1 = 4.6337 \times 10^{-6}$ ng $\cdot$ CFU$^{-1}\cdot$hour^{-1}. A comparison between model predictions and observed measurements is provided in Figure 3.10. Finally, in order to use our estimate for γ_1 in the stochastic model from Figure 3.1, one needs to convert units (from ng to number of molecules). To do this, we note that PA has a relative molecular mass of $83,000$ (Petosa *et al.* (1997)). This means that 7.2×10^9 PA molecules have an approximate weight of 1ng, so that $\gamma_1 = 3.3355 \times 10^4$ molecules$\cdot$ CFU$^{-1}\cdot$hour^{-1}.

3.4.2.2 Summary statistics

We now perform a global sensitivity analysis on a subset of the model parameters for the summary statistics of interest introduced in Section 3.3. We consider the stochastic model of Figure 3.1 with baseline parameter values: $\mu_1 = 0.3723$ hour^{-1}, $\lambda_1 = 0.8$ hour^{-1} and $\gamma_1 = 3.3355 \times 10^4$ molecules$\cdot$ CFU^{-1} $\cdot$hour^{-1}, according to the calibration carried out in the previous section. To analyse the role of antibiotic treatment (B_1 represents non-treated bacteria and B_2 antibiotic-treated bacteria), we explore parameter regimes with $\nu_{12} > 0$, $\nu_{21} = \lambda_2 = 0$, $\mu_2 \geq \mu_1$ and $\gamma_2 \leq \gamma_1$.

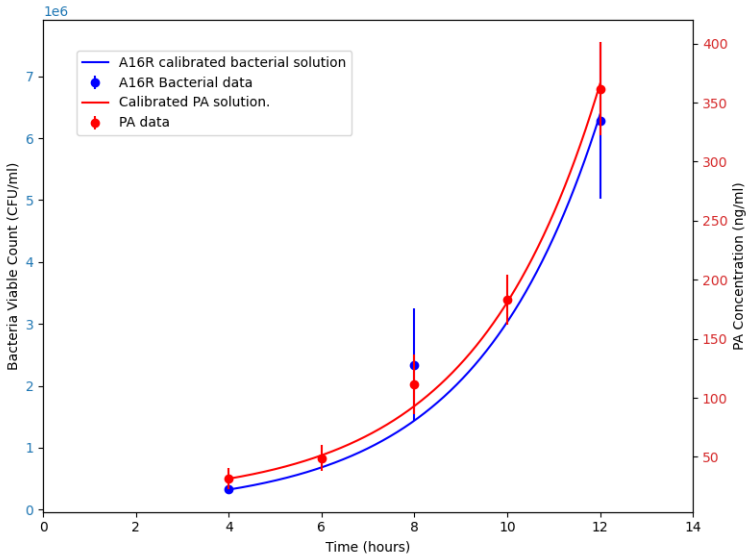

Figure 3.10: Model predictions compared to experimental observations by Zai *et al.* (2016).

In Figure 3.11, we look at summary statistics directly related to the lifetime of a single bacterium. We assume at time $t = 0$ we start with one non-treated bacterium. We first carry out a sensitivity analysis for parameters μ_2, ν_{12} and γ_2. This allows one to analyse the impact of treatment on the tracked bacterium during its lifespan. On the other hand, even though we have a baseline value for μ_1, we vary this parameter when considering the number of divisions undergone by the tracked bacterium, for illustrative purposes. The top-left plot in Figure 3.11 shows the impact of treatment on the mean lifespan of the bacterium, $\mathbb{E}[T_1]$, which varies between 1 and 3 hours for the parameter values considered. Increasing antibiotic efficiency (in terms of larger values of μ_2 and ν_{12}) leads to shorter lifespans. We note that if one assumes $\mu_2 = \mu_1 = 0.3723$ hour^{-1}, no effect of treatment on the lifespan is expected, and the value of ν_{12} (which is directly related to the rate at which antibiotic can affect bacteria, as well as the concentration of antibiotic present in the system) becomes irrelevant. Finally, increasing values of μ_2 make the value of ν_{12} increasingly relevant, as one would expect.

The top-right plot of Figure 3.11 shows the expected number of divisions undergone by the bacterium, $\mathbb{E}[D_1]$, for a range of μ_1 and ν_{12} values. We note here that since $\nu_{21} = \lambda_2 = 0$, μ_2 has no effect on D_1. Thus, we vary μ_1 instead. As one would expect,

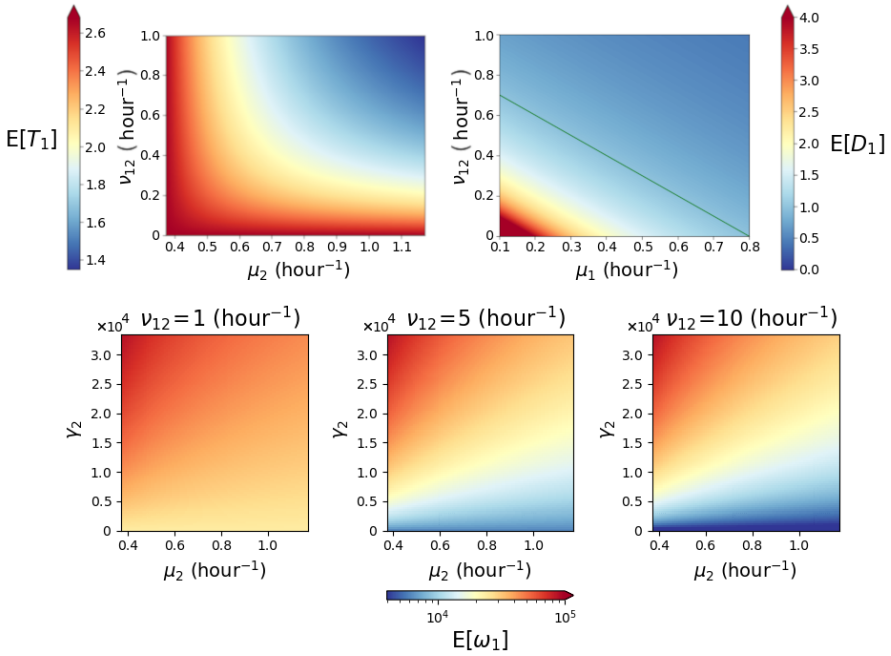

Figure 3.11: Top-left. Expected lifespan [$hours$] of a bacterium. **Top-right.** Expected number of divisions during its lifetime. **Bottom.** Expected number of toxin molecules produced during its lifetime for different values of $\nu_{12} \in \{1, 5, 10\}$ (left to right). Units for γ_2, not provided in the plot, are molecules $\cdot$ CFU$^{-1}\cdot$ hour^{-1}.

increasing values of ν_{12} and μ_1 lead to fewer divisions. We also note that in order for the bacterial population to grow as a function of time, each bacterium (on average) needs to undergo more than one division event. We highlight the value $\mathbb{E}[D_1] = 1$ by a green line in Figure 3.11, which is achieved when $\nu_{12} + \mu_1 = \lambda_1$. The bottom row in Figure 3.11 shows the effect of varying ν_{12}, μ_2 and γ_2 on the expected number of toxin molecules produced by a bacterium during its lifetime, $\mathbb{E}[\omega_1]$. Increasing values of μ_2 and ν_{12} can have a significant effect on the number of toxin molecules produced. The values $\gamma_2 = \gamma_1 = 3.3355 \times 10^4$ molecules $\cdot$ CFU^{-1} hour^{-1} and $\mu_2 = \mu_1 = 0.3723$ hour^{-1} represent no treatment effect for the tracked bacterium, and for these choices the value of ν_{12} has no effect on $\mathbb{E}[\omega_1]$. On the other hand, decreasing values of γ_2 have a significant effect on the predicted number of PA molecules produced, especially for increasing values of ν_{12}.

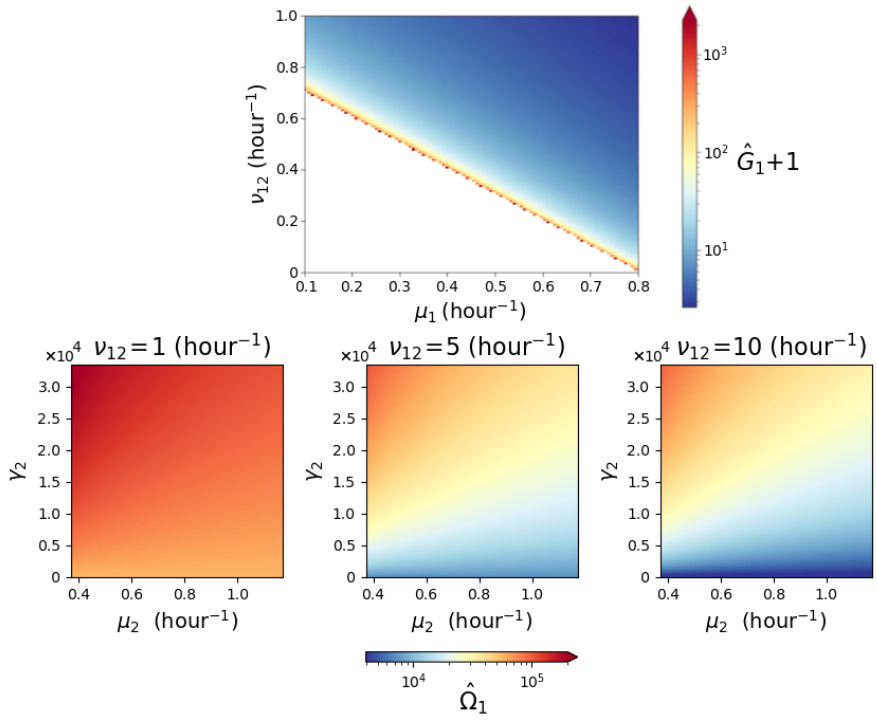

Figure 3.12: Top. Mean number of bacteria in the genealogy of a single bacterium, where the genealogy is the bacterium and all its progeny. **Bottom.** Mean number of toxin molecules produced by the genealogy of a single bacterium for different values of $\nu_{12} \in \{1, 5, 10\}$ (from left to right). Units for γ_2, not provided in the plot, are molecules $\cdot$ CFU$^{-1}\cdot$ hour^{-1}.

In Figure 3.12 we present summary statistics of relevance to the genealogy of a B_1 bacterium. The top plot of Figure 3.12 shows the effect that parameters ν_{12} and μ_1 have on the mean number of cells in the genealogy of a single bacterium, $1 + \hat{G}_1$. We note that, in this plot, the white area corresponds to parameter combinations for which the mean number of cells in the genealogy is not finite. This happens when $\lambda_1 \geq \mu_1 + \nu_{12}$. Values of $\mu_1 + \nu_{12}$ larger but close to λ_1 lead to increasing mean number of cells in the genealogy, as one would expect. On the other hand, the bottom row of Figure 3.12 shows the effect on the number of toxin molecules secreted by the genealogy of a single bacterium of varying μ_2 and γ_2. We investigate these parameter values for three different choices of $\nu_{12} \in \{1, 5, 10\}$. It is clear that γ_2 has a large impact on the expected value $\hat{\Omega}_1$, which mimics the similar effect that γ_2 has on its single-bacterium counterpart, $\mathbb{E}[\omega_1]$ (see Figure 3.11). Figure 3.11 and Figure 3.12 show the significance of ν_{12} on the expected number of toxin molecules produced. Interestingly, as ν_{12} becomes much larger than λ_1, we observe that $\mathbb{E}[\omega_1]$ approaches $\hat{\Omega}_1$, since in this case $1 + \hat{G}_1 \approx 1$ represents the single bacterium of interest.

3.5 Discussion

We have defined and analysed a two-compartment stochastic model for toxin-producing bacteria. Our focus has been a number of summary statistics that relate to the lifetime of a single bacterium (tracked over time) and its progeny. In particular, we have studied the lifespan of the bacterium, the number of divisions undergone and the number of toxin molecules produced during its lifetime, as well as the number of bacteria in its progeny, and the number of toxin molecules produced by this progeny and the original bacterium itself. We illustrated in Section 3.4 our methods by focusing on the growth of *B. anthracis* bacteria in the presence of antibiotic treatment. To the best of our knowledge, this is the first approach to quantify the PA production rate in this system. It is important to note that as mentioned in Section 3.4.2.1, we have neglected the effect of PA degradation, ξ, which will likely lead to an underestimate for our rate of PA production, γ_1. We will use the dataset considered in Section 3.4.2.1, by Zai *et al.* (2016), and other *in vitro* datasets to better quantify PA production and degradation in Section 4.2.

We point out that, although the model considered in Figure 3.1 is relatively simple, consisting only of two compartments of bacteria, our single-bacterium approach can be applied to any network *topology* of compartments, as long as the bacteria behave independently, so that the dynamics of a single bacterium can be effectively followed. Implementing our techniques in more complex systems, such as those representing *in vivo* infection and bacterial dissemination between different organs, remains the aim of future work. We also note that, while we have analysed probability generating functions and Laplace-Stieltjes transforms in Section 3.3, we have focused in practice, for simplicity and brevity, on computing the first order moments for the summary statistics of interest. However, this approach can be easily generalised to compute higher order moments or probability mass functions.

Some potential areas of further investigation based on this work could be to incorporate the behaviour of a toxin population, that is toxins being produced by the bacteria and decaying at some rate, by making use of the PA production and degradation estimates calculated in Section 4.2 to this model. This would be useful if we wanted to, for example, investigate the probability that a bacterium produced enough toxins to reach some given threshold value of toxins during the lifespan of its genealogy, a quantity that may be of interest if it is known for a given pathogen the toxin concentration within the system upon which symptoms are likely to start or when a hosts immune system may be overwhelmed.

Chapter 4

Mathematical models of *Bacillus anthracis* infection dynamics *in vivo* and *in vitro*

There has been a wide range of research in modelling anthrax infection. Anthrax is a disease, caused by exposure to the bacteria *Bacillus anthracis*, that encourages this broad scope of research as it can be induced by a number of entry points. This chapter focuses on inhalational anthrax infections, where the spores are breathed into the lungs by an individual. This type of anthrax infection is extremely rare in modern life, but has become a necessary topic of research due to its potential use as a biowarfare agent. As anthrax is a disease with high lethality, *in vivo* data in humans is scarce; there have been experiments where *in vivo* animal data has been obtained, for example in macaques (Vasconcelos *et al.* (2003)) or rabbits (Pitt *et al.* (2001)). This animal data is useful for understanding key mechanisms of the pathogenesis of an infection as well as obtaining key data (such as LD_{50}, the average lethal dose) for these animals. Biological knowledge gained from animal studies and *in vitro* data will therefore be used to calibrate our model parameters in this chapter. Two published sets of data are used in Section 4.2 and these results are compared with novel *in vitro* data from our partners at Dstl in order to draw conclusions about parameter values affecting population dynamics in an anthrax infection.

4.1 A within-host model of anthrax infection

In this section solely human infections will be considered, in which there has been a substantial amount of research. This includes models of the infection throughout the body (Albrink (1961)), the efficacy of different treatments (Plotkin & Grabenstein (2008)) and some investigation into the best practical response to a release of *B. anthracis* spores into a population (Inglesby *et al.* (2002)). One such deterministic model has been proposed by Day *et al.* (2011), which incorporates a host inhaling a number of spores into the lungs and then the spores migrating into the transmediastinal lymph nodes. In the following subsections, this model will be introduced (Section 4.1.1), the results found in the paper recreated (Section 4.1.2), and then this model will be used to develop a stochastic counterpart to each stage of the model (Sections 4.1.3 and 4.1.4).

4.1.1 Deterministic model by Day *et al.* (2011)

Here, we present the deterministic model proposed by Day *et al.* (2011). This model considers within-host anthrax infection dynamics in which spores of *B. anthracis* are inhaled into the alveolar region. Once the spores arrive into the body, the alveolar cells respond and phagocytose spores, attempting to destroy them. However, due to the robustness of *B. anthracis*, the spores survive and thrive in these conditions; the authors denote the new construct with the spore inside the alveolar cell as a "host cell". These host cells migrate to the transmediastinal lymph nodes (TMLN) region, while the spore(s) inside the host cell germinates, leading to intracellular vegetative bacteria that can then proliferate. Eventually these bacteria lyse the host cells and once they rupture, extracellular bacteria are released into the lymph nodes.

Once the extracellular bacteria have been released into the lymph nodes, multiple processes begin. The extracellular bacteria undergoes mitosis which enhances the infection. This is inhibited by the host's immune cells, including neutrophils that have been attracted from the blood stream to the infected area. If bacteria remain in the system they can begin to produce anthrax toxins which cause the symptoms and potentially death of the host. If the infection is left untreated it is extremely likely this will result in death of the host; treatment can greatly increase the chance of survival, with one study showing a 0% survival rate without treatment up to an average of above 90% survival rate with

treatment (Huang *et al.* (2015); Inglesby *et al.* (2002)). Although biologically anthrax toxins are made up of at least three virulence factors, such as PA, EF and LF (see Section 1.1.1.2), this model represents them as one dimensionless quantity within the system.

In the first part of the model, depicted in Figure 4.1, the authors assume that an amount of spores, S_0, is inhaled. Once within the host, these spores are phagocytosed by the alveolar cells, and these infected host cells, H, migrate to the TMLNs. Within these cells, the spores germinate and become vegetative bacteria. These bacteria, once released into the TMLNs, produce anthrax toxins.

The number of alveolar phagocytes is assumed to be in homeostasis before the infection, at a value of $\frac{s_A}{\mu_A}$, where s_A is the number of alveolar phagocytes that arrive within the system per hour and μ_A is the death rate of these phagocytes. Once the infection begins, the uptake of spores into the cells happens rapidly. Due to this, it is assumed that the number of alveolar phagocytes approaches a quasi-steady state given by the equation

$$A(t) = \frac{s_A}{k_2 S(t) + \mu_A},$$

where k_2 is the rate at which alveolar phagocytes phagocytose spores and $S(t)$ is the number of spores in the system.

Figure 4.1 depicts the different reactions (or events) that are considered within the model. This figure also highlights two of the limitations of the model: namely that the intracellular process within the host cell is not modelled and that anthrax toxins are modelled as a single dimensionless quantity.

The first two equations within this model represent the first stage of the infection process, where the host cells migrate from the lungs to the TMLN. This involves the populations of the spores, S, the alveolar cells, A, and the combination of the two forming a host cell, H. The evolution of the spores throughout time is represented by the equation

$$\frac{dS}{dt} = \underbrace{-k_2 S A}_{\substack{\text{Decrease caused by spores being} \\ \text{phagocytosed by resident cells}}}. \tag{4.1}$$

As spores are phagocytosed, the number of host cells within the lungs varies according

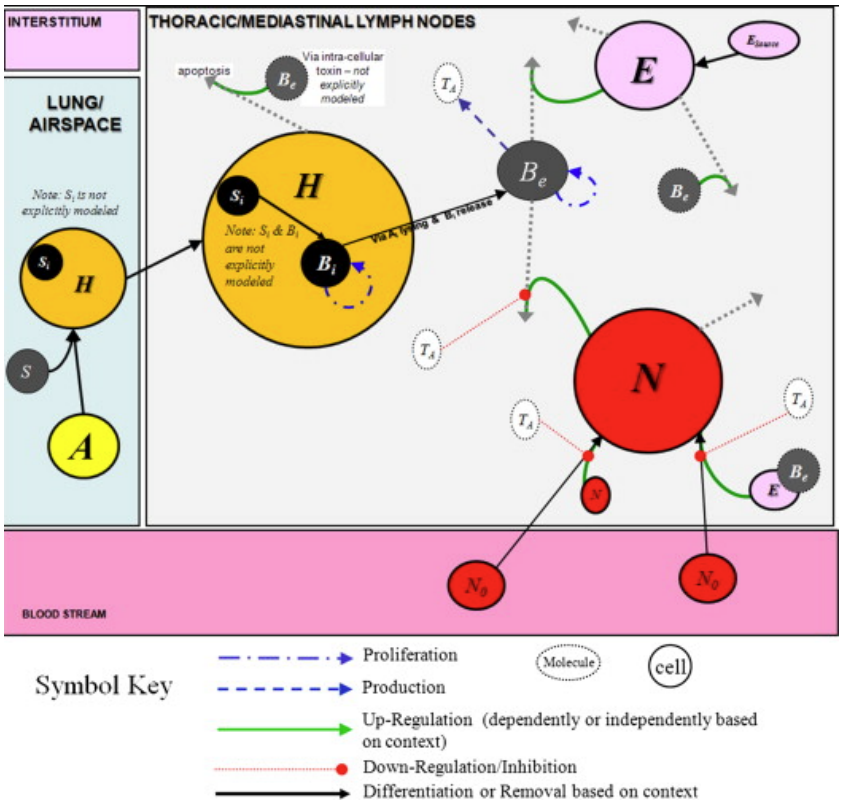

Figure 4.1: A pictoral representation taken from Day *et al.* (2011) of the model. In the lung region, S represents spores, H are host cells and A represents the alveolar phagocytes. In the TMLN compartment, B_e represents extracellular bacteria, T_A are anthrax toxins, N represents neutrophils and E shows lymphocyte mediators.

to the following equation

$$\frac{dH}{dt} = \underbrace{\frac{k_2 SA}{n_s}}_{\substack{\text{Increase as spores} \\ \text{are phagocytosed}}} \underbrace{-k_3 H}_{\text{Migration of host cells}} . \tag{4.2}$$

The first term on the right hand side in this equation describes alveolar cells becoming host cells by phagocytosing spores, with the assumption that each cell phagocytoses n_s spores. The second term describes the host cells leaving the alveoli by migrating to the TMLNs, where k_3 is the migration rate. For the purposes of this model, Day *et al.* (2011) assumed that this migration term includes the process of bacteria germinating and lysing the cells upon arrival into the TMLNs, releasing exactly n_B bacteria into the system.

Let us consider the second stage of infection where host cells have arrived into the TMLN region at rate k_3. Although once released into the lymph nodes extracellularly, bacteria and toxins can spread through the body via the blood stream, in this case it is assumed that the dynamics of the infection within the TMLN system is representative of the overall condition of the host. Namely, if the infection is sustained in this region it will progress though the body and the converse is also true. This assumption has been backed up by bodies of victims who have succumbed to anthrax; autopsies have consistently shown inflammation of this region in fatal cases (Day *et al.* (2011); Shafazand *et al.* (1999)). This section of the model represents the rate of change of four cell populations: early lymphocyte mediators, E; extracellular anthrax bacteria, B_e; Neutrophils, N; and anthrax toxins, T_A.

The first population considered within the TMLN system is the early lymphocyte mediators which have an inflammatory effect on the cells in the region which serves to attract neutrophils. The population of these lymphocytes evolves according to

$$\frac{dE}{dt} = s_E \underbrace{-\mu_E E}_{\substack{\text{Natural biological death} \\ \text{of lymphocytes}}} \underbrace{-k_1 B_e E}_{\text{Death induced by bacteria}} , \tag{4.3}$$

where s_E is the source of lymphocytes, μ_E is the natural death rate of lymphocytes and k_1 is the rate of apoptosis of lymphocytes caused by B_e. Of particular interest in terms of anthrax pathogenesis is the parameter k_1, as the biological processes it describes are

dependant on lethal toxin, a combination of two secreted toxins by the vegetative bacteria, namely protective antigen (PA) and lethal factor (LF). For example, it was shown by Popov *et al.* (2002) that the rate at which macrophages are killed by *Bacillus anthracis* is related to the level of lethal toxin in the system, whilst Friedlander (1986) found that this suppression on the immune response was inhibited by a high pH environment. The populations of protective antigen (PA) and lethal factor (LF), which combine to make lethal toxin, are not explicitly modelled here, and the model takes k_1 as implied by the effect of some intracellular lethal toxin dynamics.

The model described by Day *et al.* (2011) assumes that when host cells migrate to the TMLN system they rupture and release n_B bacteria, which can then proliferate logistically. Within this model, these extracellular bacteria can be killed by both the early immune response (population E) and the neutrophil population (N). Therefore, the change in the population of B_e is given by

$$\frac{dB_e}{dt} = \underbrace{k_3 n_B H}_{\substack{\text{Bacteria being}\\ \text{released from rupturing}\\ \text{host cells}}} \underbrace{+ k_5 B_e \left(1 - \frac{B_e}{B_{eMAX}}\right)}_{\text{Extracellular bacterial growth}} \underbrace{- k_6 E B_e}_{\substack{\text{Killing of bacteria by}\\ \text{early immune response}}} \underbrace{- \frac{k_8 N B_e}{1 + \frac{T_A}{c_{t_1}}}}_{\text{Killing of bacteria by neutrophils}}, \tag{4.4}$$

where k_5 is the growth rate of bacteria, B_{eMAX} is the carrying capacity of the bacteria population, k_6 is the rate at which resident immune cells cause death of bacteria, k_8 is the maximal rate at which neutrophils kill bacteria and c_{t_1} is a measure of the inhibition of neutrophil processes. Note that when $T_A(t) = c_{t_1}$ the rate at which neutrophils kill bacteria is half the maximal rate, therefore a smaller value of c_{t_1} would indicate the toxins have a stronger inhibitive effect.

Equation (4.3) shows that lymphocyte mediators in the TMLNs die of natural causes and due to the presence of extracellular bacteria (which produce toxins). Day *et al.* (2011) assume that the deaths of these lymphocytes attract neutrophils from the blood stream which are effective at killing anthrax bacteria, therefore they assume that the population of neutrophils evolves according to

$$\frac{dN}{dt} = \underbrace{\frac{k_9 B_e E N_0}{1 + \frac{T_A}{c_{t_2}}}}_{\substack{\text{Recruitment of neutrophils}\\ \text{by lymphocytes dying}}} + \underbrace{\frac{k_{10} N N_0}{1 + \frac{T_A}{c_{t_3}}}}_{\substack{\text{Recruitment of neutrophils}\\ \text{by other primed neutrophils}}} \underbrace{- \mu_N N}_{\text{Natural death rate}}. \tag{4.5}$$

In Equation (4.5), k_9 is the rate at which neutrophils are activated by B_e-caused death, N_0 is a constant source of resting neutrophils (approximately 5500) in the TMLN system, c_{t_2} and c_{t_3} are both measures of the extent that anthrax toxins inhibit neutrophil responses, k_{10} is the rate at which neutrophils are recruited by other activated neutrophils and μ_N is the natural death rate of neutrophils. The first two terms show neutrophils arriving in the TMLN due to signals sent by dying lymphocytes or by other neutrophils bringing in reinforcements; both of these processes however are inhibited by the anthrax toxins produced by the bacteria as shown in the denominator of each term.

Finally, the dynamics of the toxins is represented by

$$\frac{dT_A}{dt} = \underbrace{k_4 \frac{B_e}{c_{t_b} + B_e}}_{\substack{\text{Production of toxin} \\ \text{by extracellular bacteria}}} \underbrace{-\mu_{T_A} T_A}_{\text{Natural decay of toxins}}, \tag{4.6}$$

where k_4 is the maximal rate that extracellular bacteria produce toxins, c_{t_b} is the level at which bacteria produce toxins at half the maximal rate and μ_{T_A} is the natural decay rate of the anthrax toxin.

The parameter values proposed by Day *et al.* (2011) for this model are detailed in Table 4.1.

4.1.2 Results of the model

In this section, some of the numerical results presented by Day *et al.* (2011) are reproduced, along with some additional numerical results that are not discussed within the scope of the paper.

Survival and death outcomes for the host are not explicitly factored into the model. Instead, it is assumed that if the level of bacteria and neutrophils within the system remains constantly high then the disease will overwhelm the host, resulting in host death, but if these two quantities begin to decrease then it assumed that the host has begun to deal with the infection and will survive. Due to the deterministic nature of the model, there exists a threshold value for the initial dose of spores such that the system will switch from a survival case to a death case if it is surpassed. This threshold was discussed but the value was not explicitly reported by Day *et al.* (2011); it is shown here in Figure 4.2.

Table 4.1: The values for the parameters within this model, justification for which are found in Day *et al.* (2011, Table 1). The baseline initial conditions for this model were $S(0) = S_0, H(0) = 0, E(0) = 2 \times 10^8, B_e(0) = 0, N(0) = 0$ and $T_A(0) = 0$.

Fixed Parameter	Parameter Value	units
B_{eMAX}	5×10^{11}	B_e
$c_{t_1}, c_{t_2}, c_{t_3}$	1	T_A-units
c_{t_b}	1000	B_e
k_1	10^{-5}	$B_e^{-1}\ hour^{-1}$
k_2	10^{-10}	$H^{-1}\ hour^{-1}$
k_3	0.05	$hour^{-1}$
k_4	2	T_A-units $hour^{-1}$
k_5	0.8	$hour^{-1}$
k_6	5×10^{-10}	$E^{-1}\ hour^{-1}$
k_8	6×10^{-10}	$N^{-1}\ hour^{-1}$
k_9	5×10^{-10}	$B_e^{-1}E^{-1}\ hour^{-1}$
k_{10}	1×10^{-5}	$N^{-1}\ hour^{-1}$
N_0	5500	N
n_b	5	B_e
n_s	3	S
s_A	3×10^8	$Hhour^{-1}$
s_E	1×10^8	$Ehour^{-1}$
μ_A	0.05	$hour^{-1}$
μ_E	0.05	$hour^{-1}$
μ_N	0.06	$hour^{-1}$
μ_{T_A}	2	$hour^{-1}$

Within Figure 4.2 we have an initial exposure of $S(0) = 4270$ and $S(0) = 4271$ spores in blue and orange respectively. Therefore, for this deterministic model, the initial dose threshold for a death case is 4271 spores. Note that within Figure 4.2 is that a limitation of this deterministic model is shown; as $S(0) = 4270$ is given as a survival case but with a sustained infection on a time scale of over 6 days, empirically we know that as anthrax is a serious disease someone with a sustained infection like this does not have survival guaranteed. This limitation of the deterministic model, where near the threshold values we cannot quantify the probability of an outcome (survival or death), is motivation for the stochastic version proposed in Sections 4.1.3 and 4.1.4.

It is clear that for low values of S_0 a host will recover and for high values they will die; however, within the model there is a key parameter, k_3, which describes the rate at which host cells migrate to the TMLN system. This is a key parameter as there are a lot of implicit behaviours within this rate, such as the cell rupturing and releasing n_B bacteria, and k_3 determines how quickly the inhaled spores develop into bacteria within the host cell. Our knowledge of the disease anthrax is that it is extremely fast acting once a host begins to show symptoms, therefore this rate k_3 may determine the period of time where treatment (not modelled within this section, but Day *et al.* (2011) do consider it in their paper) will be most effective. Thus, in Figures 4.3 and 4.4 the impact of varying k_3 is explored for both a low dose of spores, $S_0 = 10^2$, and a high dose, $S_0 = 10^4$. For the low dose of spores, in Figure 4.3, there is a peak for both the bacterial population and the neutrophil population, after which the bacteria eventually returns to 0 and the neutrophil count decreases. This is interpreted as a survival case, as the bacteria and immune response have both returned to their baseline values. It is logical to expect that there is some threshold value for spore exposure where an infected host will survive, in fact it has been estimated that the dose of spores that one would expect to kill 50% of humans would be between 8000-10,000 spores (Passalacqua & Bergman (2006)). Alternatively, in Figure 4.4 an exposure to a high dose of spores is considered. This corresponds to a mortality case where both neutrophil and bacteria levels tend towards a limit where they remain, the neutrophil level reaching a cap shows the immune system is exerting its maximum effect on the bacterial population and is not controlling it, therefore we interpret this as a death case. Overall, it is clear that increasing k_3, meaning that bacteria arrive in the TMLNs at a faster rate, increases the rate at which the infection

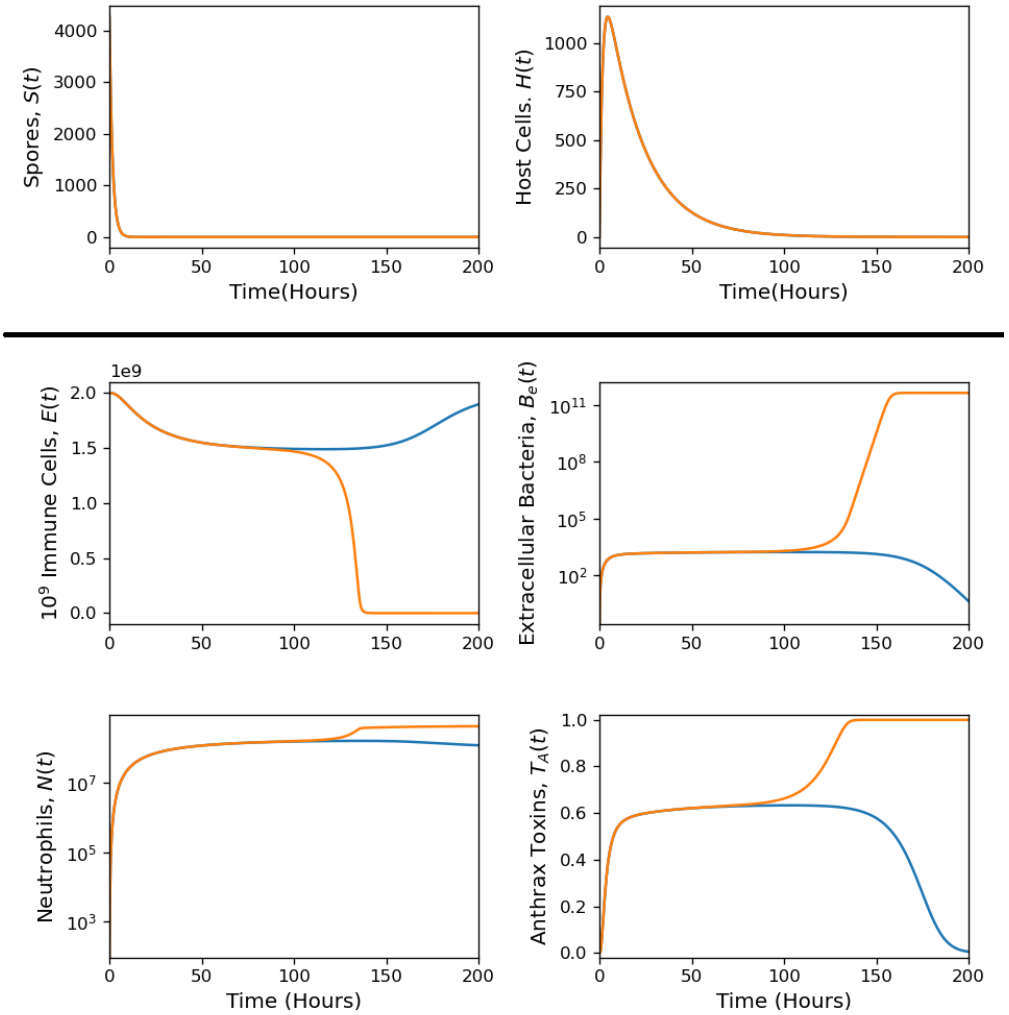

Figure 4.2: Case with initial spore exposure $S(0) = 4270$ (blue) and $S(0) = 4271$ (orange). The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section. Solutions for Equations (4.1)-(4.6).

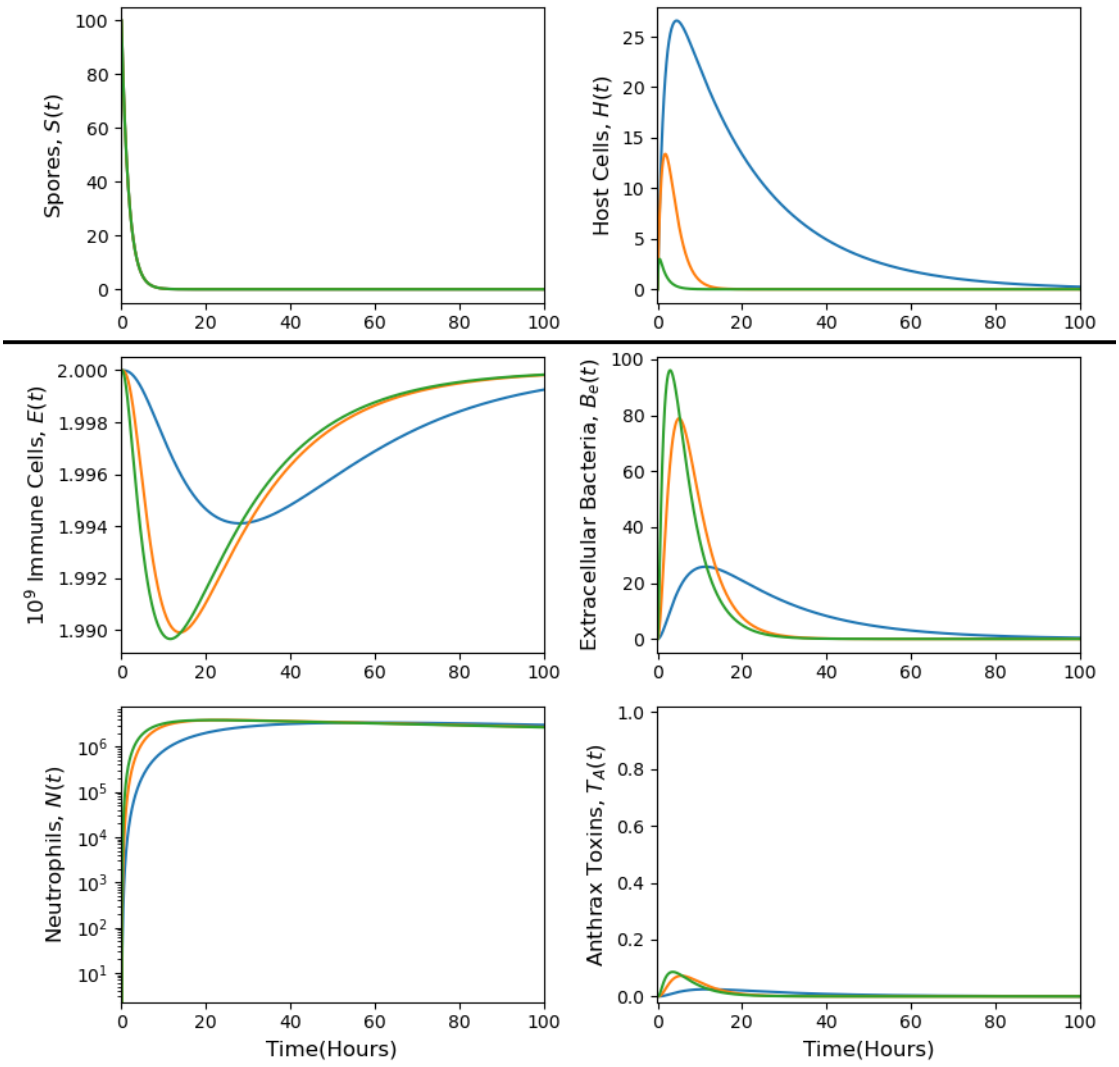

Figure 4.3: Case with initial spore exposure $S(0) = 10^2$. The three plots show different values of k_3, $k_3 = 0.05$/hour (blue), $k_3 = 0.5$/hour (orange) and $k_3 = 5$/hour (green), leading to survival. The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section. Solutions for Equations (4.1)-(4.6).

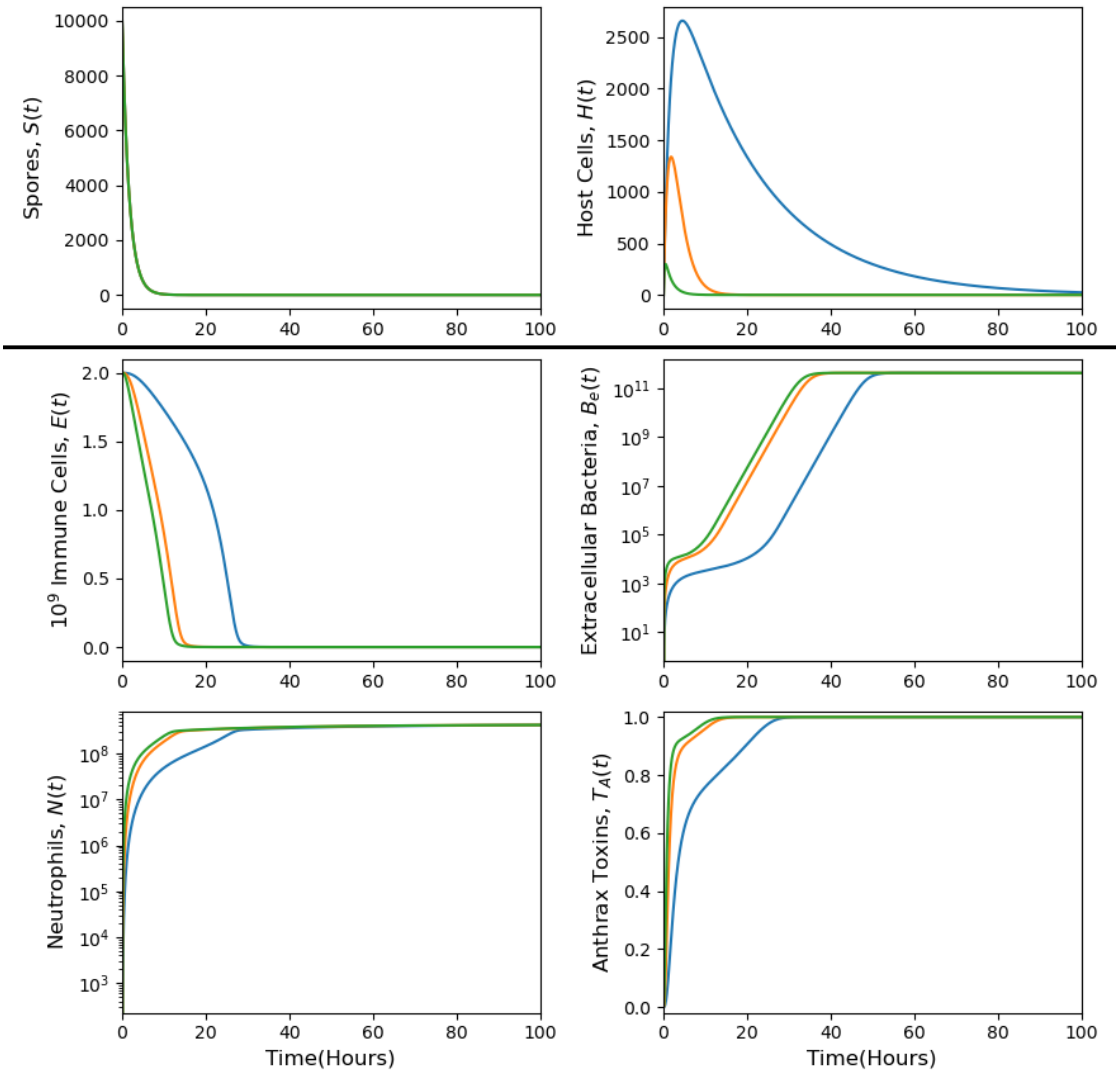

Figure 4.4: Case with initial spore exposure $S(0) = 10^4$. The three plots show different values of k_3, $k_3 = 0.05$/hour (blue), $k_3 = 0.5$/hour (orange) and $k_3 = 5$/hour (green), leading to death. The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section. Solutions for Equations (4.1)-(4.6).

progresses as well as increases the maximum number of bacteria within the TMLNs. For example, in Figure 4.3 we see that the effect of increasing k_3 from 0.05 to 5 increases the maximum number of bacteria in the TMLNs from approximately 25 to just under 100. This may have the effect of lowering the threshold number of spores inhaled to cause a death case within this system, which is something that will be demonstrated in Figure 4.6.

As it is a key parameter, we further investigate the effect of this transition rate, k_3, as was done in Day *et al.* (2011, Figure 5). Firstly, Day *et al.* (2011) define the phagocytosis-migration-germination (PMG) period of the infection within this model as the length of time that it takes for the population of the host cells to drop below 50, which the authors consider to be a negligible amount of cells within the system. Day *et al.* (2011) take $4, 30, 60, 90$ days as the PMG periods of interest, and report the values of k_3 these PMG periods correspond to for an initial dose $S(0) = 9000$ in Day *et al.* (2011, Figure 5). Using these published values as a reference point to validate our methods, we can find the minimum value of k_3 needed to obtain specific PMG periods for a variety of doses, in order to replicate Day *et al.* (2011, Figure 5). The method for obtaining a value of k_3 is to use a grid-search method. Namely, for some given initial dose of spores, $S(0)$, we start with a small value of $k_3 = 0.0001$ and calculate the deterministic solution for $S(t), H(t)$; we accept the value of k_3 if the population of host cells, $H(t)$, is less than the threshold at the given time point (e.g for a PMG length of 4 days, after 96 hours). If we do not accept the value of k_3 we increase it by 0.0001 and repeat the process until we accept a value of k_3. In comparing the output of this grid-search to the values of k_3 for $S(0) = 9000$, we hypothesise that the authors may have considered a threshold of 10 instead of 50. If we use 10 host cells as a threshold, we can show that the values of k_3 we find using our method correspond to the published values in Day *et al.* (2011) by plotting the solutions for $H(t)$ for the calculated k_3, which we do in Figure 4.5. In any case, this is a minor technical comment, since the difference between 10 host cells and 50 host cells is not major when considering a minimum initial exposure of 4500 spores. Figure 4.5 shows the population of the host cells for an initial dose of 9000 spores with the accepted parameter values of k_3 for PMG periods of 4, 30, 60 and 90 days.

The previously described process can be repeated for the other initial spore doses in Day *et al.* (2011, Figure 5), namely $S(0) = 4500, 2 \times 10^4, 2 \times 10^9$, to reproduce the figure

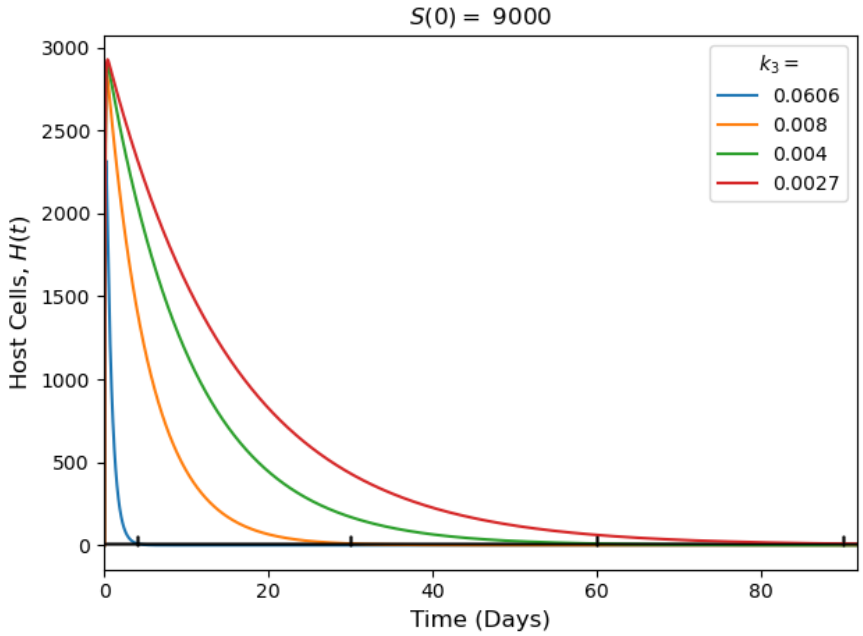

Figure 4.5: Four deterministic solutions of the host cell population, $H(t)$, after an initial exposure of $S(0) = 9000$ spores, with corresponding values of k_3, threshold value of 10 and relevant time points (4, 30, 60 and 90 days).

		PMG in days			
		4	30	60	90
$S(0)$	4500	0.0532	0.0070	0.0035	0.0024
	9000	0.0606	0.0080	0.0040	0.0027
	2×10^4	0.0691	0.0091	0.0046	0.0031
	2×10^9	0.1426	0.0187	0.0094	0.0063

Table 4.2: Computed values of k_3 (hours^{-1}) for various PMG lengths and initial conditions, S(0).

showing the outcomes for the given initial conditions and PMG periods. The values for k_3 we calculate for each initial condition and PMG length are given in Table 4.2. Figure 4.6 shows that, as expected from Figures 4.3, 4.4 and 4.5, decreasing k_3 (which increases the PMG period) will have a notable effect on the survival threshold. This is significant as it introduces another mechanism that treatments could seek to exploit to improve mortality rates within sufferers of anthrax infections.

We can show the effect that the change in k_3 will have on the deterministic simulations for the full model. In Figure 4.7, we set $S(0) = 4500$ and use the values of k_3 from the first row in Table 4.2. We plot the dynamics of the populations, $S(t), H(t), E(t), B_e(t), N(t)$ and $T_A(t)$ for an initial condition of $S_0 = 4500$ spores. We chose these populations as $H(t)$ determines the PMG length, and the others are illustrative in showing the difference between a survival case and a death case. We can see for the largest value of k_3 we calculated in Table 4.2 for 4500 spores ($k_3 = 0.0532$ in blue) we get a death case where bacteria reach their carrying capacity; whereas for the slower transition phases we obtain survival scenarios. It is important to observe from the figures that the reasons for these results are that the fast migration of host cells which rupture to produce extracellular bacteria results in a rapid expansion in the bacterial population which causes a suppressive effect on the immune cells; this suppression allows the bacteria to maintain its population through the plateau before growing again exponentially when the immune response has been defeated. In this section we have reproduced the model proposed by Day *et al.* (2011) and determined that, due to its deterministic nature, there is a threshold for the initial number of spores inhaled in which the system switches from a death case to a survival case. We have shown the effect that k_3 has on this survivability threshold and on the model dynamics in Figures 4.6 and 4.7, respectively. However, a limitation of this

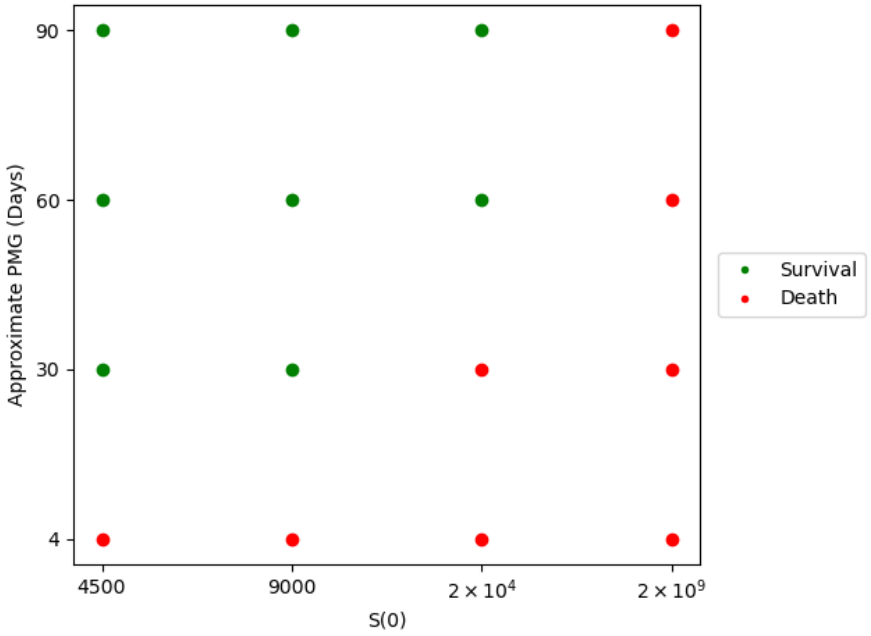

Figure 4.6: Outcomes of the system for varying phagocytosis-migration-germination (PMG) lengths and initial spore exposure, $S(0)$.

type of deterministic model is that it is unrealistic to assume that each infection will have the same dynamics; and therefore it will not be the case that survival is always achieved if the threshold is not reached and that death occurs 100% of the time if it is. Therefore, in the next sections we introduce a stochastic counterpart to the model proposed by Day *et al.* (2011) to introduce this element of randomness and allow us to investigate the probability of a survival outcome for a number of initial dose of inhaled spores.

4.1.3 A stochastic model of *Bacillus anthracis* spores in the lungs

The focus within this section is in proposing a stochastic counterpart of the within-host deterministic model in Section 4.1.1 . To this aim, I consider first only the biological processes occurring in the lungs (represented by Equations (4.1) and (4.2)). This limits the two events that could occur to a spore being ingested by a phagocyte to create a host cell, or a host cell migrating to the TMLN system.

We represent this stochastically in terms of a bidimensional CTMC $\mathcal{X} = \{(S(t), H(t)) : t \geq 0\}$ where $S(t)$ represents the population of spores and $H(t)$ represents the population

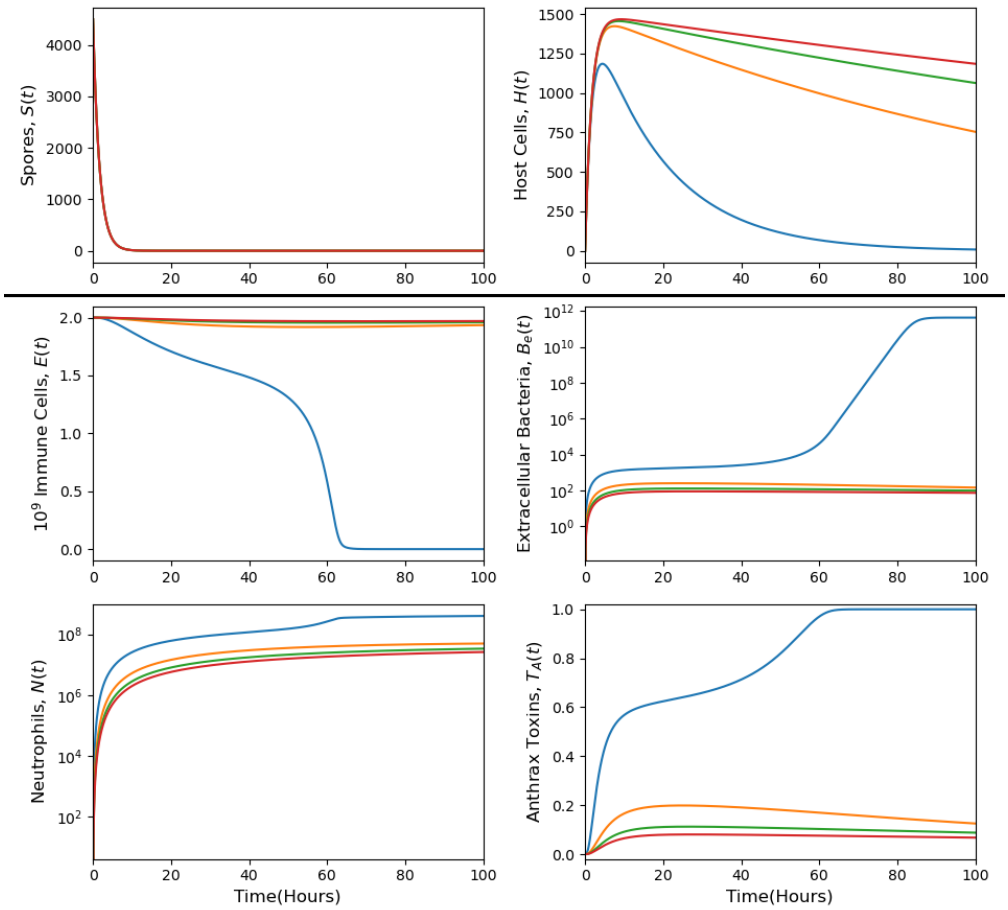

Figure 4.7: Deterministic solutions of the model for selected populations with $k_3 = 0.0532$ (blue), 0.007 (green), 0.0035 (orange), and 0.0024 (red), for initial dose of spores $S(0) = 4500$. Dynamics of cells in the lungs (spores and host cells) are shown in the plots above the line, and in the TMLNs below the line.

of infected host cells. $S(t)$ and $H(t)$ are defined over the space of states $\mathcal{S} = \{(s,h) \in \mathbb{N}_0^2 : \ 0 \leq s \leq S_0, \ 0 \leq h \leq S_0 - s\}$, where $S_0 = S(0)$ is the initial dose of spores inhaled.

The transition probabilities can be defined as follows:

$$\mathbb{P}\{S(t+\Delta t) = s', H(t+\Delta t) = h' | S(t) = s, H(t) = h\} =$$

$$\begin{cases} k_2 A s \Delta t + o(\Delta t), & \text{if } (s' = s-1, h' = h+\delta), \\ k_3 h \Delta t + o(\Delta t), & \text{if } (s' = s, h' = h-1), \\ 1 - (k_2 A s + k_3 h)\Delta t + o(\Delta t), & \text{if } (s' = s, h' = h), \\ o(\Delta t), & \text{otherwise}, \end{cases}$$

where

$$A = \frac{s_A}{k_2 s + \mu_A}, \quad \text{and } \delta \sim \text{Bin}\left(1, \frac{1}{n_s}\right).$$

This means that the possible one-jump transitions that $\mathcal{X}$ can make are

$$\begin{aligned} (s,h) &\xrightarrow{k_2 s A} (s-1, h+\delta) \\ (s,h) &\xrightarrow{k_3 h} (s, h-1) \end{aligned},$$

We note that when a spore is phagocytosed with rate $k_2 s A$, a spore is always lost but a host cell is not always created. A host cell is created with probability $\frac{1}{n_s}$, as δ is a Bernoulli (as number of trials, $n = 1$) random variable with probability of success $\frac{1}{n_s}$. The result of this is to retain the concept of n_s being the average number of spores within a host cell from Day *et al.* (2011). From our method of sampling, on average every n_s spores being absorbed will produce a new host cell. These host cells then migrate to the TMLN system with rate k_3. The time for the next event to occur, if the process is at a given state $(s,h) \in \mathcal{S}$ at any given time, is sampled from $t \sim \text{Exp}(k_2 s A + k_3 h)$. In practice, we consider

$$t = -\frac{\log(u_1)}{k_2 s A + k_3 h},$$

where $u_1 \sim U(0,1)$ from the Gillespie algorithm (see Section 2.2.2). Figure 4.8 shows a comparison of the deterministic model and the proposed stochastic model. From this figure we can see that the stochastic model accurately reflects the analytic solution of the deterministic model of the lungs. This figure also effectively shows the fact that there

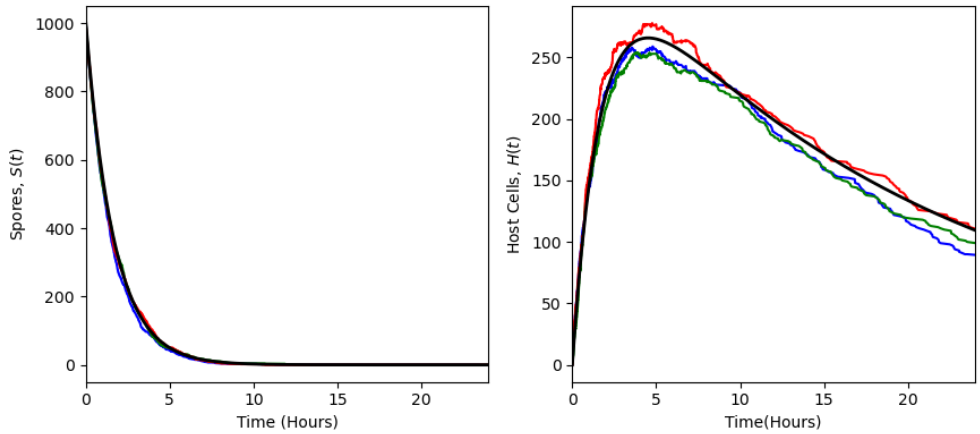

Figure 4.8: This figure shows the deterministic model (black line) based upon Equations (4.1) and (4.2) and three stochastic realisations (red, green and blue) that run until $t = 24$ hours, with initial conditions $S(0) = 1000, H(0) = 0$ and parameter values given in Table 4.1.

is no randomness within the deterministic model, whereas the stochastic model does change with each realisation.

The stochasticity of the CTMC allows one to capture the uncertainty in some of the model outputs. For example, the time until the spores population goes to extinction, $T = inf\{t \geq 0 : S(t) = 0\}$, is a random variable in the stochastic process (since in each stochastic realisation, T takes a different value). One way of approximating the distribution of T is by sampling values of T by means of many stochastic simulations (with the Gillespie algorithm). In Figure 4.9, we plot the distribution of T approximated by means of 10^4 simulations of the stochastic process. 10^4 simulations were chosen as a balance between a high number of simulations to appropriately capture the shape of the distribution and the increase in computational time as a consequence of the increase in the number of simulations. Figure 4.9 shows that while the simulations record a variety of extinction times, the results are centred around the time that the deterministic solution approaches zero (starting after 10 hours from Figure 4.8).

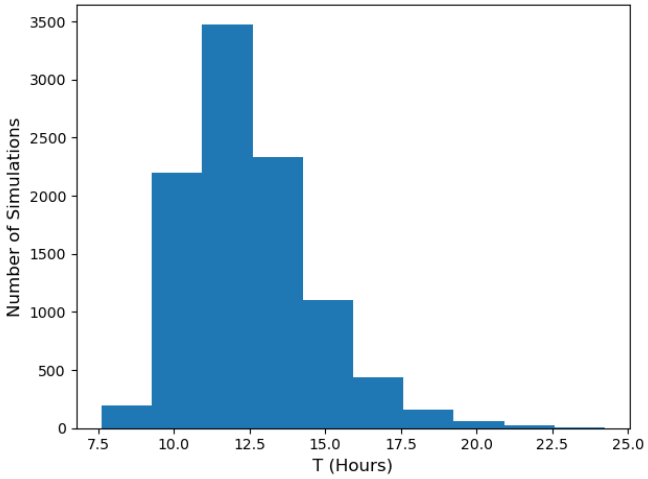

Figure 4.9: Histogram of the probability distribution of T approximated by 10^4 stochastic simulations, with $S(0) = 1000$.

4.1.4 A stochastic model of anthrax infection in the lungs and TMLN

Within this section, the aim is to expand the stochastic model to include the dynamics within the TMLN region. To expand the system it is necessary to consider a continuous-time stochastic process $\mathcal{X} = \{(S(t), H(t), E(t), B_e(t), N(t), T_A(t)) : t \geq 0\}$ where $S(t)$ and $H(t)$ are, as before, the spores and host cells in the lungs, and the other four populations represent populations in the TMLN region; $E(t)$ the lymphocyte mediators, $B_e(t)$ represents extracellular bacteria, $N(t)$ the neutrophils and $T_A(t)$ represents the anthrax toxins. In this model anthrax toxins are included for their suppressive effect on the immune response, but are modelled in an abstract way such that the units of this population do not correspond to a physical measurement of the amount of toxins. Instead, anthrax toxins within the model, $T_A(t)$, are a continuous quantity on a scale from 0 to 1, representing the proportion of the maximal suppressive effect toxins have on the immune system. As this population is not restricted to integer values, $\mathcal{X}$ is not currently a CTMC even though it has the Markovian property. To work around this limitation, let us consider $\mathcal{Y} = \{(S(t), H(t), E(t), B_e(t), N(t)) : t \geq 0\}$, where the populations S, H, E, B_e, N are defined as before, and we evaluate the toxin population $T_A(t)$ separately. As $T_A(t)$ is

a continuous quantity, we update it deterministically using equation

$$T_A(t+\Delta t) = \left(k_4 \frac{B_e(t)}{c_{t_b} + B_e(t)} - \mu_{T_A} T_A(t)\right)\Delta t + T_A(t), \tag{4.7}$$

based on Equation (4.6), where Δt is a sufficiently small inter-event time for any given jump in the Gillespie algorithm. If the returned inter-event time Δt is not sufficiently small this may not track the behaviour of the toxins accurately; this problem can be avoided by splitting the time interval into multiple time intervals and using Equation (4.7) multiple times to more accurately track its behaviour.

The other populations are slightly simpler as they take integer values. By considering the terms within the differential equations in the deterministic version of the model the following transition probabilities can be obtained:

$$\mathbb{P}\{S(t+\Delta t) = s', H(t+\Delta t) = h', E(t+\Delta t) = e', B_e(t+\Delta t) = b', N(t+\Delta t) = n' | S(t) = s, H(t) = h, E(t) = e, B_e(t) = b, N(t) = n\} =$$

$$\begin{cases}
k_2 A s \Delta t + o(\Delta t), & (s', h', e', b', n') = (s-1, h+\delta, e, b, n), \\
k_3 h \Delta t + o(\Delta t), & (s', h', e', b', n') = (s, h-1, e, b+n_B, n), \\
s_E \Delta t + o(\Delta t), & (s', h', e', b', n') = (s, h, e+1, b, n), \\
(\mu_E + k_1 b) e \Delta t + o(\Delta t), & (s', h', e', b', n') = (s, h, e-1, b, n), \\
k_5 b \Delta t + o(\Delta t), & (s', h', e', b', n') = (s, h, e, b+1, n), \\
\left(k_6 b e + k_5 b \frac{b}{B_{eMAX}} + \frac{k_8 n b}{1+\frac{T_A(t)}{c_{t_1}}}\right)\Delta t + & (s', h', e', b', n') = (s, h, e, b-1, n), \\
o(\Delta t), & \\
\left(\frac{k_9 b e N_0}{1+\frac{T_A(t)}{c_{t_2}}} + \frac{k_{10} n N_0}{1+\frac{T_A(t)}{c_{t_3}}}\right)\Delta t + o(\Delta t), & (s', h', e', b', n') = (s, h, e, b, n+1), \\
\mu_N n \Delta t + o(\Delta t), & (s', h', e', b', n') = (s, h, e, b, n-1), \\
1 - (k_2 A s + k_3 h + s_E + (\mu_E + k_1 b) e + & \\
k_5 b + \left(k_6 b e + k_5 b \frac{b}{B_{eMAX}} + \frac{k_8 n b}{1+\frac{T_A(t)}{c_{t_1}}}\right) & (s', h', e', b', n') = (s, h, e, b, n), \\
+ \left(\frac{k_9 b e N_0}{1+\frac{T_A(t)}{c_{t_2}}} + \frac{k_{10} n N_0}{1+\frac{T_A(t)}{c_{t_3}}}\right) + \mu_N n\Big)\Delta t + & \\
o(\Delta t), & \\
o(\Delta t), & \text{otherwise.}
\end{cases}$$

4.1.4.1 Limitations of the stochastic model

This model, even with the toxin population being modelled deterministically, is clearly much more complicated than the model we developed solely for the lung parts; as such there are some problems that arise. The first and foremost is that a system of this size is extremely computationally expensive to run; the reason for this is twofold. Firstly, there are now 8 possible state changes instead of 2 within the lung model; this in itself would not be a great problem, but it is still worth noting that the added complexity does have a computational cost. The second and more difficult problem to get around is the magnitude of some of the rates of these reactions, and therefore how quickly a reaction is likely to take place. Recall that the inter-event time is given by

$$\Delta t = -\frac{\log(u_1)}{\sum_i \rho_i},$$

where u_1 is a random number between 0 and 1 and ρ_i is the rate at which each reaction i occurs. Therefore the time between events, Δt, is inversely proportional to the sum of the rates of reactions that can occur. This causes an issue for events with high rates, such as the third and fourth possible state changes, corresponding to the arrival or death of an immune cell. For the parameter values being considered, these are both of orders of magnitude 10^8 per hour at time $t = 0$, so the time for an event to occur just from these two events is likely to be of order of 10^{-8} hours. This means that there is likely to be tens of millions of events concerning solely the immune cell population each hour from the first hour, which when we want to be able to consider days of exposure to the disease has an extremely high computational demand. One solution is proposed where we approach populations with higher orders of magnitude in a deterministic way. This would especially lower the computational cost in the early time stages. Another solution is proposed, known as the tau-leaping method (Gillespie (2001)), which carries out approximate stochastic simulations, by allowing several reactions to occur in a single time step.

However, before moving on to these solutions to the limitations discussed, for illustrative purposes, we can compare the dynamics of this stochastic model with its deterministic counterpart when some parameter values are adjusted so that the computational

limitations discussed above are mitigated. In Figure 4.10, we show the results of a deterministic simulation against three stochastic simulations for adjusted parameter values. For these results, we set the parameter concerning the source of the immune cells to $s_E = 10^3$, instead of 10^8 from Table 4.1 and the initial condition for the immune cells, $E(0) = 2 \times 10^4$. This means that the ratio between s_E and $E(0)$ stays the same but on a lower magnitude. This also means that the rates concerning the immune cell population (arrival and death) are of order of magnitude 10^3.

From Figure 4.10, we can clearly see that the stochastic simulations are reflecting the behaviour we see in the deterministic solution, which illustrates how the stochastic model proposed is able to mimic its deterministic counterpart.

4.1.4.2 A hybrid approach for large size populations

As briefly discussed in the previous section, one possible solution to the high computational cost of simulating the model stochastically could be to solve some populations of a high order of magnitude deterministically as we did previously with $T_A(t)$ (Equation 4.7). The justification for this is that if these populations are of a high order of magnitude they are very likely to behave mostly deterministically.

This means that for each iteration of the Gillespie algorithm we could consider

$$\begin{aligned} T_A(t+\Delta t) &= \left(k_4 \frac{B_e(t)}{c_{t_b} + B_e(t)} - \mu_{T_A} T_A(t)\right) \Delta t + T_A(t), \\ E(t+\Delta t) &= \left(s_E - (\mu_E + k_1 B_e(t) E(t))\right) \Delta t + E(t), \\ N(t+\Delta t) &= \left(\frac{k_9 B_e(t) E(t) N_0}{1 + \frac{T_A(t)}{c_{t_2}}} + \frac{k_{10} N(t) N_0}{1 + \frac{T_A(t)}{c_{t_3}}} - \mu_N N(t)\right) \Delta t + N(t), \end{aligned} \tag{4.8}$$

and the possible state changes for the stochastic variables are given by the transition probabilities

$$\begin{aligned} &\mathbb{P}\{S(t+\Delta t) = s', H(t+\Delta t) = h', B_e(t+\Delta t) = b' | \\ &\quad S(t) = s, H(t) = h, B_e(t) = b\} = \end{aligned}$$

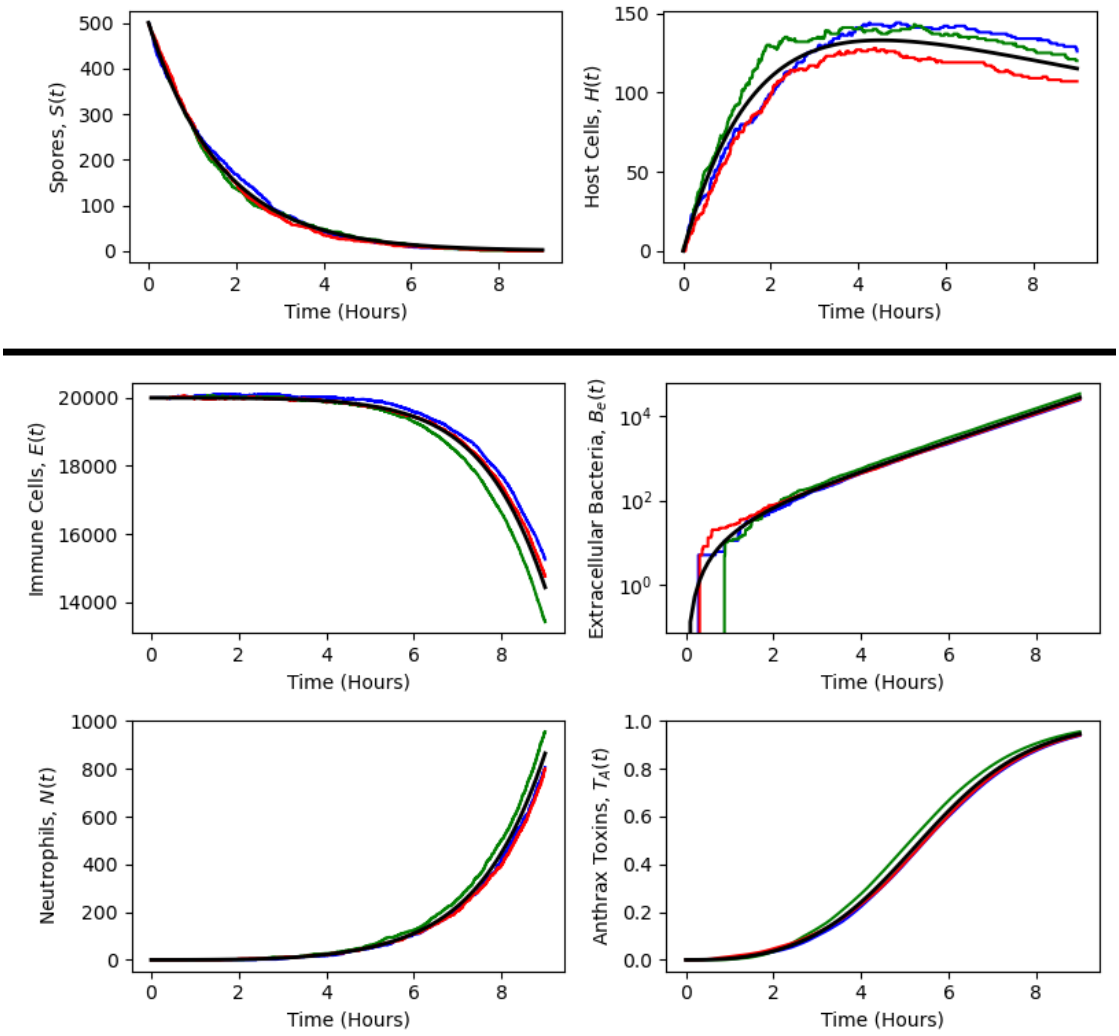

Figure 4.10: Deterministic solutions (black) against stochastic realisations for $S(0) = 500, E(0) = 2 \times 10^4$, $s_E = 10^3$ and other parameters as shown in Table 4.1. The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section.

$$\begin{cases} k_2 As\Delta t + o(\Delta t), & (s', h', b') = (s-1, h+\delta, b), \\ k_3 h\Delta t + o(\Delta t), & (s', h', b') = (s, h-1, b+n_B), \\ k_5 b\Delta t + o(\Delta t), & (s', h', b') = (s, h, b+1), \\ \left(k_6 bE(t) + k_5 b\frac{b}{B_{eMAX}} + \frac{k_8 N(t) b}{1+\frac{T_A(t)}{c_{t_1}}}\right)\Delta t + o(\Delta t), & (s', h', b') = (s, h, b-1), \\ 1 - (k_2 As + k_3 h + k_5 b \\ + \left(k_6 bE(t) + k_5 b\frac{b}{B_{eMAX}} + \frac{k_8 N(t) b}{1+\frac{T_A(t)}{c_{t_1}}}\right))\Delta t + o(\Delta t), & (s', h', b') = (s, h, b), \\ o(\Delta t), & \text{otherwise.} \end{cases}$$

In Figures 4.11 and 4.12, we set $S(0) = 4000$ and $S(0) = 8000$ respectively and use the rest of the parameters as in Table 4.1. We then compare the dynamics of all simulated variables over time between this hybrid approach and the fully deterministic counterpart. We can see that the behaviour is captured relatively well for each of the populations in both the survival and death case.

4.1.4.3 Tau-leaping simulations

The Gillespie algorithm is extremely useful but does have limitations; for example in our case some of the populations are simply so large that events happen so often that the time-step is extremely small, making it extremely computationally expensive to calculate each step exactly over a period of hours or days. Gillespie himself proposed an approximation algorithm (Gillespie (2001)) along with subsequent improvements; the method that we will use (due to efficacy in tau selection) was proposed by Cao *et al.* (2006). We will first explain the methodology used, alongside some of the justifications for these adaptations before moving onto our results of applying the method.

Tau-leaping methodology

Consider a CTMC $\mathcal{X} = \{\mathbf{X}(t) = (X_1(t), \ldots, X_M(t)); t \geq 0\}$, representing the evolution of M populations over time. $\mathcal{X}$ is defined over the state space $\mathcal{S} = \{0, 1, 2, ...\}^M$, and a set of reactions $R = \{1, \ldots, n\}$, which change the state of $\mathbf{X}(t)$, and associated rates $\{\alpha_1, \ldots, \alpha_n\}$. The magnitude of these rates depend on the state of $\mathbf{X}(t)$. Tau-leaping allows one to simulate a stochastic realisation using an approximation of the Gillespie algorithm to reduce the number of computations needed. The approximation consists of considering a short jump in time, τ, in which it is assumed the rates do not change (in

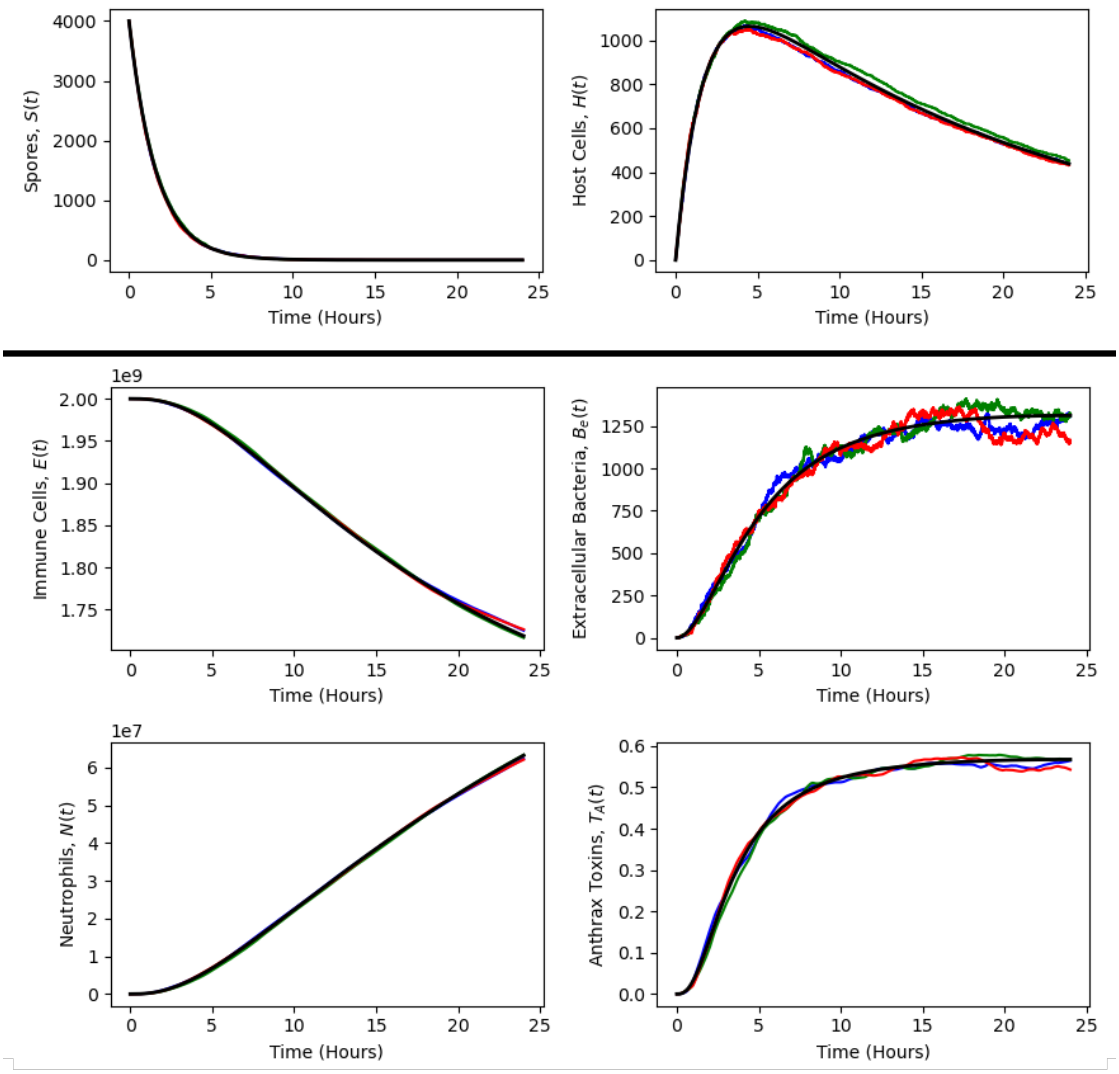

Figure 4.11: Deterministic solutions (black) against stochastic realisations for $S(0) = 4000$, the parameter values as in Table 4.1 with the immune cell, neutrophil populations and level of toxins calculated at each time step of the Gillespie algorithm using Equations (4.8). The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section.

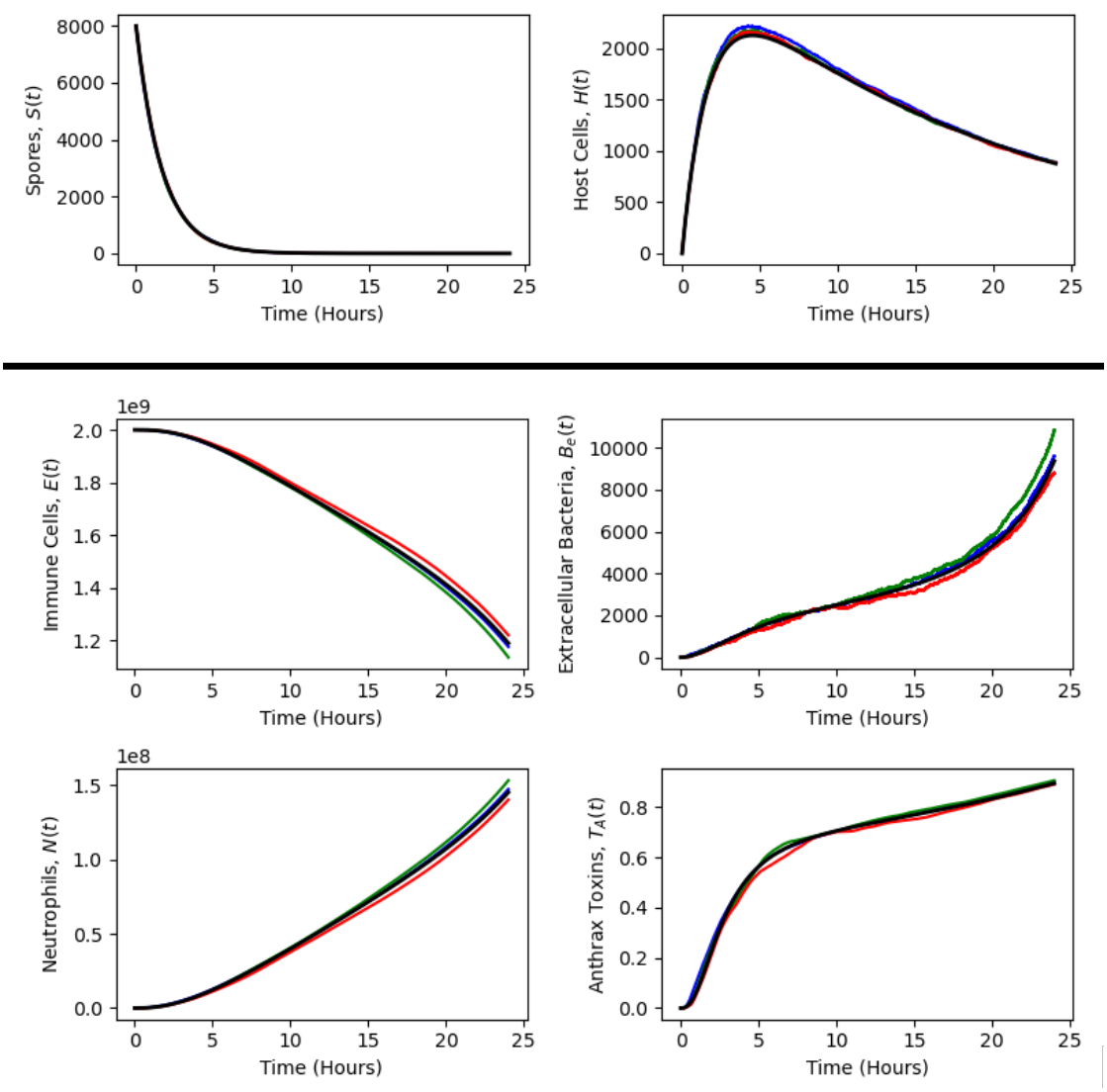

Figure 4.12: Deterministic solutions (black) against stochastic realisations for $S(0) = 8000$, the parameter values as in Table 4.1 with the immune cell, neutrophil populations and level of toxins calculated at each time step of the Gillespie algorithm using Equations (4.8). The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section.

reality, they do change but size of τ is limited so that these changes are not significant). Within this time period, τ, the number of times each reaction, j, in R occurs is sampled from a Poisson random variable with mean and variance equal to $\alpha_j(\mathbf{X}(t))\tau$, where $\alpha_j(\mathbf{X}(t))$ is the value of α_j at state $\mathbf{X}(t)$.

A bound on the time period, τ, is defined such that

$$|X_i(t+\tau) - X_i(t)| \leq \max\{\varepsilon_i X_i(t), 1\}, \quad i = 1, \ldots, M.$$

where $X_i(t)$ corresponds to a given population in the process and ε_i corresponds to a given bound for the propensity of population i. This ensures that within the time period τ, each population, X_i, will not be forced to change by an amount less than one; if $\varepsilon_i X_i(t) < 1$ it is still possible that the population $X_i(t)$ can move to an adjacent state within the time period τ. The size of change in each population is otherwise bounded upwards by the parameter ε_i. The parameter ε_i is given by the formula

$$\varepsilon_i = \frac{\varepsilon}{g_i}.$$

Cao *et al.* (2006) suggest that the standard chosen value for ε is 0.03 so we shall use this value going forward. g_i is depends on the highest order of reaction in which species i appears as a reactant. The order of the reaction is obtained from the number of reactants in the event. For example, in the reaction in which a lymphocyte dies due to a reaction with bacteria the rate is $k_1 B_e E$. Therefore this would be a second order reaction, as two molecules are required for this reaction, and $g_E = 2$. Each of the reactions in the lungs section of the model are first order (for $S(t)$ and $H(t)$, $g_i = 1$), and the toxin population is of first order, whereas the remaining populations after migration all have at least some interaction with another population (but not all at the same time) so for these populations $g_i = 2$. With this bound for τ established, we can express each step in the tau-leap algorithm using this equation:

$$\mathbf{X}(t+\tau) = \mathbf{X}(t) + \sum_{j=1}^{n} P_j(\alpha_j \mathbf{X}(t)\tau)\mathbf{V}_j,$$

where $\mathbf{X}(t) = (X_1(t), \ldots, X_M(t))$, n is the total number of possible state changes in the model, $P_j(\alpha_j(\mathbf{X}(t))\tau)$ is the aforementioned Poisson random variable with mean and

variance $\alpha_j(\mathbf{X}(t))\tau$ for state change j and $\mathbf{V}_j$ is the vector representing the change that state change j has on each variable.

In the case described in Section 4.1.4, $M = 5, \mathcal{X} = \{(S(t), H(t), E(t), B_e(t), N(t)) : t \geq 0\}$ and $n = 8$. If we first consider the reaction that describes a spore being phagocytosed and forming a host cell the associated vector, $\mathbf{V}_1$, would be $(-1, \delta, 0, 0, 0)$ as population $S(t)$ decreases by 1 and population $H(t)$ increases by $\delta \sim \text{Bin}\left(1, \frac{1}{n_s}\right)$, and all other populations are unchanged in the system for this reaction.

The benefit of tau-leaping is that it decreases the number of times that the rates of the system need to be reset, thus decreasing the computational power needed. However, this is only true if the right value of τ is chosen. If τ is too small there will not be a large enough reduction in computational time to make it worthwhile using an approximation of the Gillespie algorithm instead of an exact Gillespie algorithm; and if τ is too large the big time differences between each step of the simulation would cause the rates to alter a non-trivial amount between jumps, making the simulation lose a large degree of accuracy. Therefore, it is crucial that τ is chosen carefully and efficiently. As an aside, if τ is chosen such that only one jump is included in each time-step this recovers the Gillespie algorithm.

To select a suitable τ, we begin in each iteration of the tau-leaping algorithm by separating populations into one of two distinct categories, the non-critical and critical populations. A critical population is defined as one which is below a certain threshold (here defined as 10) such that care is needed in our selection of τ so as not to cause any population to drop below zero. A tau-leaping algorithm considers that in each leap a single critical reaction occurs exactly once and the non-critical reactions happen as many times as possible while this condition is met.

Then a tau is selected via the following formula:

$$\tau_1 = \min_{i \in \text{NCrit}} \left\{ \frac{\max\{\varepsilon_i X_i(t), 1\}}{|\mu_i(X_i(t))|}, \frac{\max\{\varepsilon_i X_i(t), 1\}^2}{\sigma_i^2(X_i(t))} \right\},$$

where NCrit is the set of indices of non-critical reactions, ε_i is given by $\frac{0.03}{g_i}$ as discussed above, and $\mu_i(X_i(t))$ and $\sigma_i^2(X_i(t))$ are the expectation and the variance of the number of transitions of species i at $\mathbf{X}(t)$ in time τ respectively.

From there, three things can happen. The first outcome depends on the value of tau being large enough that a non-trivial number of reactions will take place. This is monitored by checking that $\tau > 10/\sum_i^n(\alpha_i)$ where α_i is the rate at which reaction i occurs. If this condition is not fulfilled, an arbitrary number of iterations of a standard Gillespie algorithm are carried out. In our model this number was chosen to be 100. If the condition does hold then it is crucial that this choice of tau does not run the risk of sending a population below zero. A second tau is chosen such that τ_2 is a sample of an exponential random variable with mean

$$\frac{1}{\sum_{i\in\text{Crit}}\alpha_i},$$

where Crit is the set of indices of reactions that concern critical populations. This second tau is used to estimate the time taken for a critical reaction to occur.

As stated above, it is crucial to ensure that no more than one critical reaction occurs between each time step, so $\tau = \min\{\tau_1, \tau_2\}$ is used. If $\tau_1 < \tau_2$ it is assumed that no critical reactions occur during this leap and the number of reactions during τ_1, k_j, of each non-critical reaction is equal to $P_j(\alpha_j(\mathbf{X}(t))\tau_1)$. That is, a sample of a Poisson random variable with mean $\alpha_j(\mathbf{X}(t))\tau_1$ where α_j is the rate of reaction j at $\mathbf{X}(t)$. However, if $\tau_1 > \tau_2$ then it is assumed that exactly one critical reaction occurs in this leap. The choice of critical reaction is determined by dividing the interval $(0, 1)$ into segments, where the size of the j-th segment assigned to critical reaction α_j is given by

$$\frac{\alpha_j(\mathbf{X}(t))}{\sum_{i\in\text{Crit}}\alpha_i}.$$

We then choose a random variable by sampling from the uniform distribution between 0 and 1, find whichever interval this random variable lies in and then choose that corresponding reaction to occur once and all other critical reactions to not occur. In this case the number of times each non-critical reaction occurs during τ_2, k_j, is selected by setting $k_j = P_j(\alpha_j(\mathbf{X}(t))\tau_2)$, as described above.

For each jump the populations in $\mathcal{X}$ are updated using the equation $\mathbf{X}(t) = \mathbf{X}(t) + k_j\mathbf{V}_j + \mathbf{V}_c$ where k_j is the number of times non-critical reaction j occurs, $\mathbf{V}_j$ is the effect that non-critical reaction j has on each population and $\mathbf{V}_c$ is the effect that the selected critical reaction has on the populations.

As a final step it must be checked that no populations in $\mathbf{X}(t)$ have fallen below 0. If there is a population below 0 the state changes must be undone and the processes repeated with the selected tau being reduced by a half. This ensures that in the time it takes one critical reaction to occur a non-critical reaction does not occur enough times to exhaust its population.

4.1.4.4 Results of tau-leaping

The results of the tau-leaping algorithm generally approximate the output of the exact Gillespie algorithm but resolves much quicker for larger sizes of populations and can resolve issues at much longer time intervals. We define our code to run for time, T, defined as

$$\begin{aligned} t_1 =& t_{max} \\ t_2 =& \min\{t \geq 0 : B_e(t) \geq 10^8\} \\ T =& \min\{t_1, t_2\} \end{aligned} \tag{4.9}$$

where t_{max} is the maximum time we want the code to run over. The reason we introduce t_2 is that 10^8 is a sufficiently large population such that an infected individual is extremely unlikely to recover without treatment (and therefore can be considered as a 'death' case in this model), and this saves a high amount of computational cost whilst still allowing us to confidently describe the system as a survival or death case.

We can compare the tau-leaping simulations to the solutions of the deterministic model to validate its effectiveness. To illustrate this, we consider two marginal cases $S(0) = 4250$ and $S(0) = 4500$, chosen to be close to our deterministic critical threshold of $S(0) = 4271$ from Figure 4.2. In Figures 4.13 and 4.14 we plot the pointwise median values across 1000 tau-leaping stochastic simulations for each population in the model $(S(t), H(t), E(t), B_e(t), N(t), T_A(t))$, for initial conditions $S(0) = 4250$ and $S(0) = 4500$ respectively. We keep all parameter values as they are in Table 4.1. From these figures we can see that the tau-leaping simulations are able to closely match the deterministic behaviour of the system. It is important to note that in Figure 4.14, by our definition of T, the code ends if the bacterial population reaches 10^8, which is the cause for the plateau and divergence once the bacteria reaches this limit. We once again see in Figure 4.13 the behaviour we originally noticed in Figure 4.3, that a marginal survival

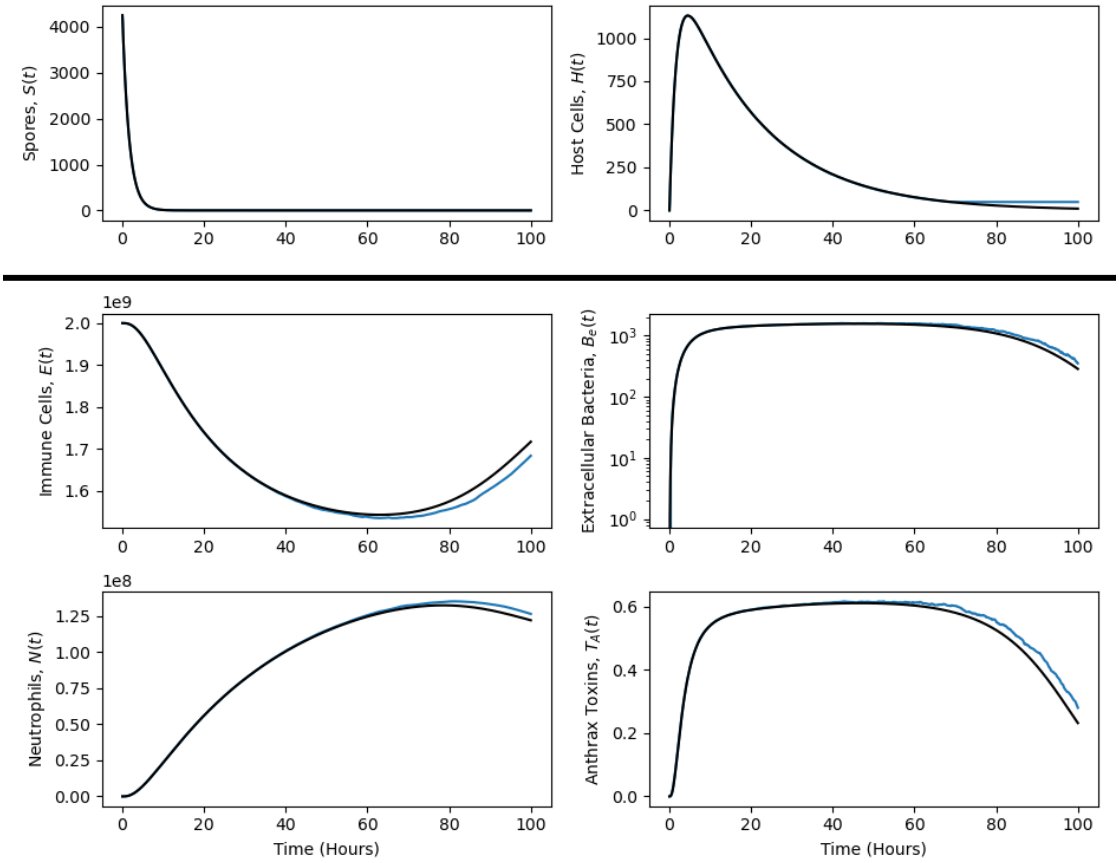

Figure 4.13: The pointwise median of 1000 tau-leaping simulations (blue) and the deterministic solutions (black) for a marginal survival case $S(0) = 4250$ and parameter values as in Table 4.1. The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section.

case will resolve to a survival case but only after an extended period of extracellular bacteria being present in the system.

We can look at individual simulations for these initial conditions too, showing what is happening to each population in detail. In Figure 4.15 we plot the individual simulations from the tau-leaping algorithm for each population with an initial dose of spores $S(0) = 4500$. The populations $S(t)$ and $H(t)$ are extremely consistent across all simulations which is reassuring as these populations have relatively low orders of magnitude and few parameters describing their behaviour. We can see from the dynamics of the extracellular bacteria population that the majority of these simulations result in an explosion in the

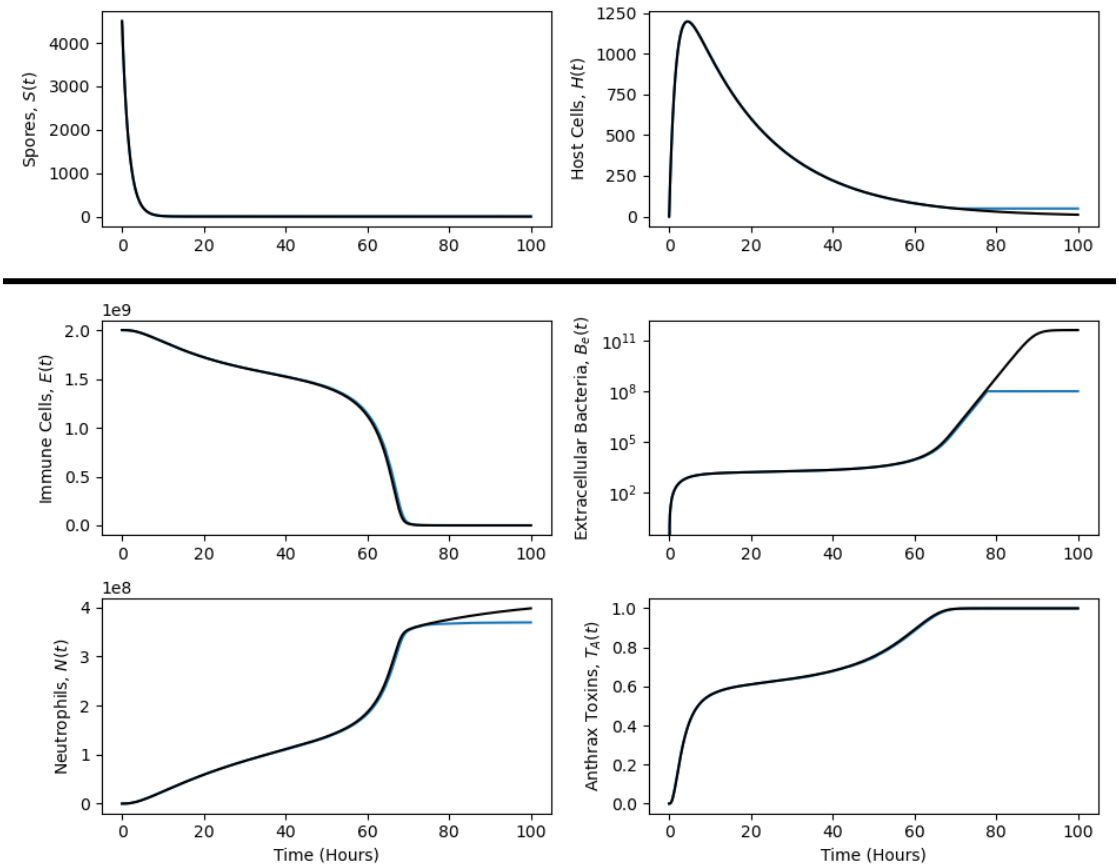

Figure 4.14: The median of 1000 tau-leaping simulations (blue) and the deterministic solutions (black) for a marginal death case $S(0) = 4500$ and parameter values as in Table 4.1. The bacterial population $B_e(t)$ is capped at 10^8; if the population reaches this limit we assume that a death case has occurred. The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section.

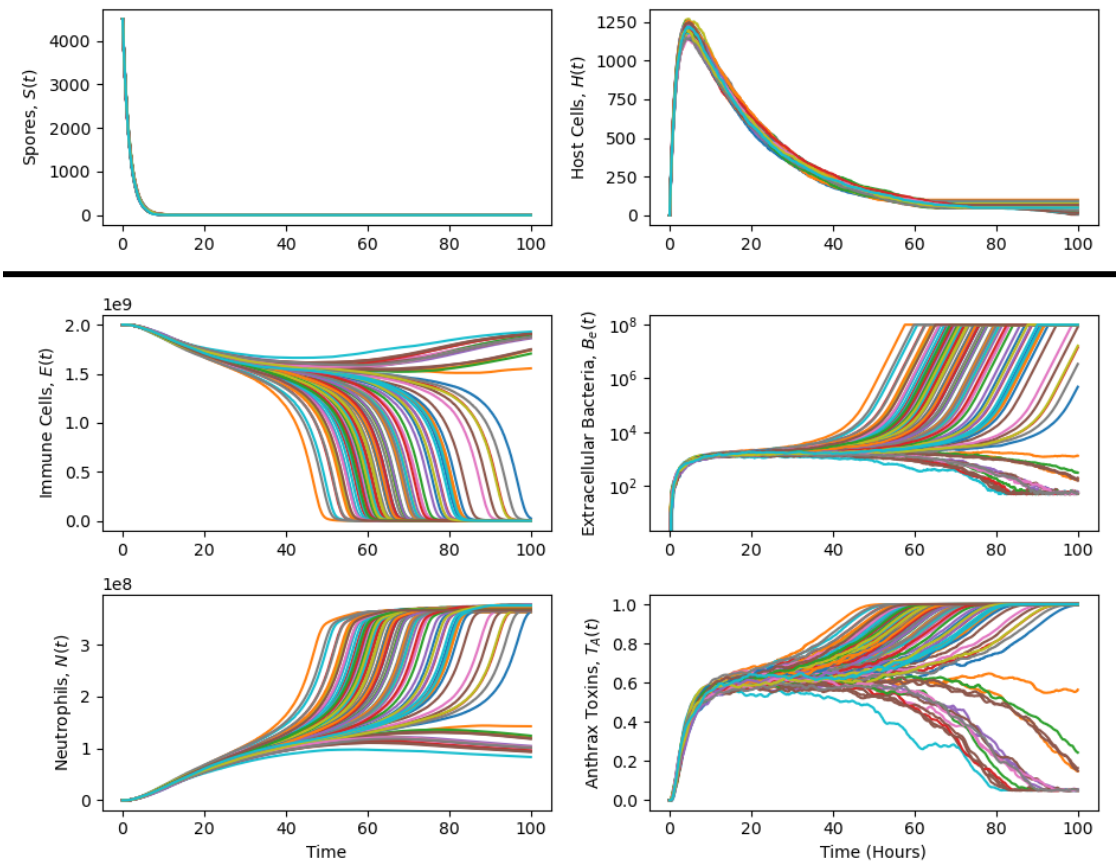

Figure 4.15: 1000 Tau-leaping simulations for a marginal death case $S(0) = 4500$. The bacterial population $B_e(t)$ is capped at 10^8; if the population reaches this limit we assume that a death case has occurred. The dynamics of cells in the lungs (spores and host cells) are shown in the top section, and the dynamics in the TMLNs in the lower section.

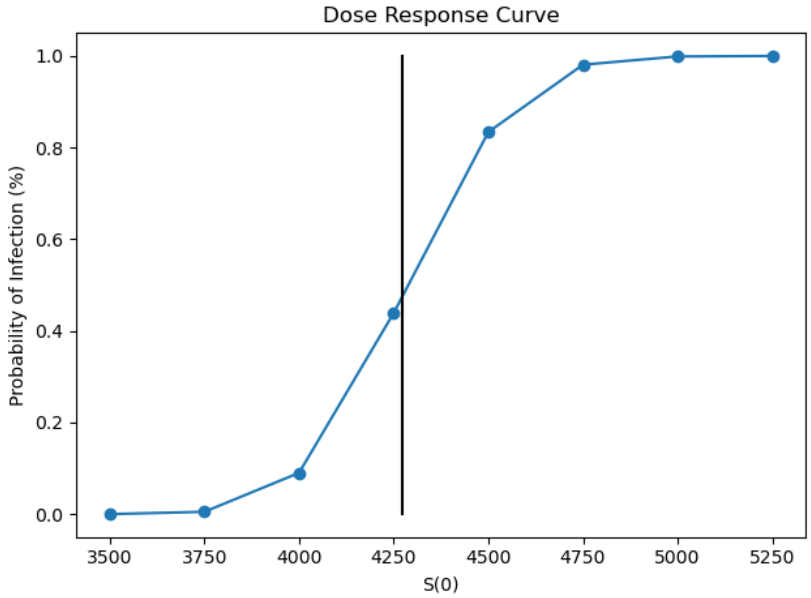

Figure 4.16: Probability of infection from 10^3 tau-leaping simulations for each $S(0)$. The threshold for a death case for the deterministic model in Section 4.1.2 of $S(0) = 4271$ is shown via the black line.

population leading to what we consider as a death case, but not all the simulations do. This allows us to begin to answer the question we posed when discussing Figure 4.2; what is the probability that a simulation will result in a death case for the model for a given initial dose? Answering this question is the reason this stochastic model can be more informative than the deterministic model by Day *et al.* (2011), as if we had an individual exposed to a initial dose of $S(0) = 4000$ spores, we have a probability of them succumbing to the disease (we call this the response probability for a dose of $S(0) = 4000$) rather than the model always predicting a survival scenario. This response probability can also be informative when considering a variety of treatment options.

In Figure 4.16 we calculate the response probability for a number of initial spore doses, centred around the threshold value for a death case from the deterministic model in Section 4.1.2 of $S(0) = 4271$. We obtain this probability, p, by running 10^3 tau-leaping simulations of the system. We calculate the proportion of simulations where the time that the model ran for, T, is less than the maximum time we would allow it to run for, namely $T < t_{max}$. This corresponds to the bacterial population, $B_e(t)$, reaching the threshold

value, 10^8. Clearly, it is important that we choose a reasonable value for t_{max} to attempt to capture all reasonable scenarios that would end as a death case; here t_{max} was chosen to be 100 hours. Our justification for this is that if the infection has not progressed to the threshold value at this time (slightly more than 4 days after infection) the host would be reasonably expected to seek treatment and become much more likely to survive. We can see from Figure 4.16 that the deterministic threshold value of initial spores, $S(0) = 4271$ gives just under a 50% chance of death. It is important to notice that the probability of infection changes from 0 to 1 over a very small magnitude of spores (from $S(0) = 3500$ to $S(0) = 5250$) meaning that a precise quantification of an individual's spore exposure is extremely important.

One limitation is that we are still differentiating between a survival and death case by simply looking at the size of the bacterial population, whereas we know that one of the key drivers of symptoms and eventual death in the pathogenesis of anthrax infection is the anthrax toxins, made up of PA, LF and EF. It is not possible to investigate these individual toxin components in this model as the anthrax toxins are thought of as a single dimensionless continuous quantity. This also makes it difficult to in future incorporate an anti-toxin treatment into the model.

It is of note, that even within this model, which includes many biological processes that describe the immune response to an anthrax infection, the behaviour of the anthrax toxins are not captured. Therefore, within the next section, we aim to quantify toxin production and decay by leveraging *in vitro* data with the aim of developing a model that can incorporate these processes.

4.2 Quantifying toxin production and decay *in vitro*

A particular ingredient that is missing in order to incorporate the anthrax toxin into within-host model, in a more quantitatively precise manner, is the quantification of toxin production and decay. Thus, in this section, the aim is to leverage several *in vitro* datasets in order to quantify this. To this end we will use previously published data as well as novel experimental data provided by our industrial partners at Dstl.

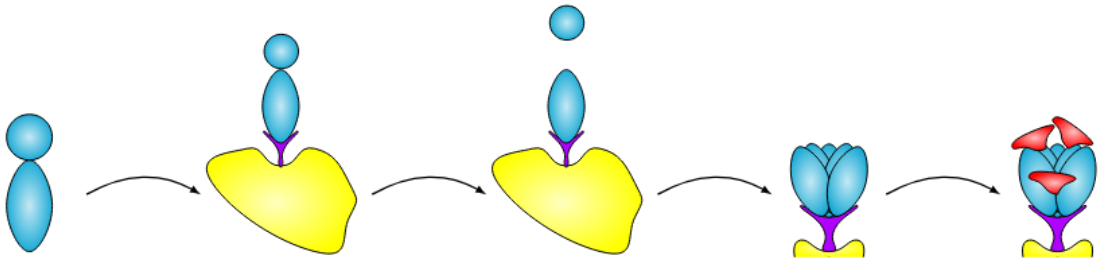

Figure 4.17: The mechanism by which PA83 proteins bind to cell receptors (purple) and undergo cleaving to form PA63 (blue ovals) and PA20 (blue circle) proteins. The bound PA63 proteins form a heptamer which LF (red) and/or EF (orange) proteins can bind to in order to enter cells (yellow) via endocytosis.

Recall that there are three main components of the anthrax toxins: protective antigen (PA), lethal factor (LF) and edema factor (EF). PA is responsible for binding to cells and forming a heptamer that provides an environment that the LF and EF proteins can bind to and then enter cells. In particular, we start with a PA83 protein which then undergoes cleaving to split into PA63 and PA20 proteins. The PA63 component can then bind to a CMG2 cell receptor, providing an environment where more PA63 proteins can bind together to form what is known as a PA heptamer. This heptamer allows LF and EF to bind to it at which point the heptamer can enter cells and cause toxic effects on the cell. We show a pictoral representation of this process in Figure 4.17.

Therefore, when considering the concentration of anthrax toxins, we will focus on studying the PA levels. This strategy is also borne out when thinking about potential applications of our modelling approach; if we wanted to analyse the effects of an anti-toxin treatment, some of the most common (such as Raxibacumab) operate by preventing PA from binding to cells. Therefore, the relevant behaviour of the anti-toxin treatment can still be captured if we only consider the PA component.

4.2.1 Experimental data and methods

We use experimental data from published studies by Zai *et al.* (2016) and Charlton *et al.* (2007), as well as from experiments carried out by our industrial partners at Dstl. We will discuss the experimental setup for the experiments carried out by our industrial partners and some of the similarities and differences between the experiments in the other two

datasets and explain the effect this may have on the data and results we obtain from each dataset.

4.2.1.1 Growth of bacteria and viable counts

Here we describe the method used by Dstl to grow the bacteria and how they calculated the number of bacteria present (the viable count CFU/ml).

The Dstl data was produced using the Sterne strain 34F2 of *B. anthracis* from the Porton Down strain collection. Sterne (pXO1+, pXO2-) is un-encapsulated but retains the ability to produce toxins and is used in the UK's anthrax vaccine AVP. Starting with a working stock containing 10^7 CFU/ml, 500 µl of Sterne spores was removed and inoculated into 50ml Brain Heart Infusion (BHI) broth in a 250ml Erlenmeyer flask to produce a culture containing 10^5 CFU/ml. Sodium bicarbonate (Sigma-Aldrich) was then added to result in a final concentration of 48 mM (Sirard *et al.* (1994)). The culture was contained in a Biojar and a CO_2 gas generator sachet (Scientific Laboratory Supplies) was added. The CO_2 sachet was replaced after 24 hours of use as per the manufacturer instructions. It is important to note that both sodium bicarbonate and CO_2 were added as both have been shown to encourage the production of PA *in vitro* (Koehler *et al.* (1994); Sirard *et al.* (1994)) by helping to simulate the *in vivo* environment (Koehler (2009)). The culture was incubated at 37°C with continuous shaking at approximately 182 rpm. In total there were three independent growth experiments performed in duplicate. The first experiment had the culture sampled at 0, 1.5, 3.5, 4.25, 5, 6, and 7 hours post-inoculation, in order to measure spore germination and bacterial growth at early stages of growth. In the other two experiments the aim was to understand the behaviour of the system after enough time had passed for the bacteria to begin to approach its carrying capacity for the experimental setup; therefore the culture was sampled at 16, 18, 20, 22, 24, 40 and 48 hours post-inoculation. Upon each sampling, total viable counts were obtained by plating serial 10-fold dilutions (100 µl aliquots) onto L agar in triplicate. In order to count both the number of spores and vegetative bacteria within the culture, at each time point another sample was taken, diluted and then heated to 70°C for 30 minutes with vigorous shaking to kill vegetative cells (Turnbull *et al.* (2007)) but crucially not spores that have not yet germinated. The number of vegetative cells in the original sample was then calculated by subtracting the number of spores (the count in the heat treated sample)

from the total CFU/ml (in the non-heat treated sample). In the experiments where we investigated the behaviour of the system at longer time scales, the culture was also filter sterilised at each time point by passing through a 0.22 µm syringe filter and stored at -20°C before analysis using the automated western blot system, Jess™ Simple Western (ProteinSimple, San Jose CA, USA).

4.2.1.2 Automated western blot

The process to quantify the production of PA during the bacterial growth was done using Jess™ Simple Western. The benefit of Jess™ is that the separation, probing and detection of protein is automated in a single hands-free assay. A combination of filter sterilised supernatant from the *B. anthracis* Sterne culture was mixed neat with Fluorescent 5× master mix at a ratio of 4:1; these samples were then denatured by heating statically to 95°C for 5 minutes. For all assays the 12-230 kDa Separation module (SM-W004, ProteinSimple) and Anti-Mouse Detection Module (DM-002, ProteinSimple) were used. Reagents were diluted and pipetted into the assay plate as per the manufacturer instructions. Jess™ aspirates the reagents into glass capillaries before separating the HRP-conjugated MW ladder and sample by size. The sample proteins become immobilised to the capillary wall before immunoprobing with 1 µg/ml monoclonal PA4 primary antibody (2D4J, produced at Dstl) and HRP-conjugated anti-mouse secondary antibody was carried out. Luminol-Peroxide was then added to the sample and the chemiluminescent signal intensity from the PA target protein is output graphically and as a virtual blot in the Compass Simple Western software (version 6.1.0, ProteinSimple). The signal intensity generated by the target PA protein was then interpolated against a standard curve of PA signal intensity. This standard curve is generated by taking recombinant PA (PharmAthene, Inc.) and diluting in 0.1× Sample Buffer and Fluorescent 5× master mix (ProteinSimple) to create samples with final concentrations ranging between 0.04 µg/ml - 5 µg/ml. Interpolating the target PA's result along this standard curve allows the concentration of PA at each time point of the Sterne growth experiment to be determined. Default assay conditions for a chemiluminescent 12-230 kDa size assay were chosen in the Compass Simple Western software; this included a sample separation time of 25 minutes and antibody incubation time of 30 minutes. Two negative system controls were

run per assay, one containing no sample and the other containing no primary antibody to check for cross-reactivity of reagents.

4.2.1.3 Previously published datasets

In addition to the experiments carried out at Dstl (described in Sections 4.2.1.1 and 4.2.1.2) we use in this section data from two previously published studies, by Zai *et al.* (2016) and Charlton *et al.* (2007).

The authors of the first dataset, Zai *et al.* (2016), conducted similar experiments to the ones carried out by Dstl (described above) and measured bacterial growth and PA concentration for both the A16R strain and the Sterne strain of *B. anthracis*. Both strains are un-encapsulated but retain the ability to produce the toxin proteins, which is why they are used in anthrax vaccines (A16R in China and Sterne in the UK). Specifically, Erlenmeyer flasks containing 100 ml of Luria-Bertani (LB) liquid medium were sterilized by autoclaving at 121°C for 15 min and then warmed to 37°C prior to inoculation with 1 ml of *B. anthracis* culture. Once inoculated the flasks were then incubated at 37°C with vigorous agitation for up to 28h. Viable counts and PA concentration were obtained by taking culture supernatant samples throughout the time course. The data for the viable counts (CFU/ml) and PA concentrations (ng/ml) of each strain are shown in Zai *et al.* (2016, Figure 1B) and Zai *et al.* (2016, Figure 4A) respectively. One minor difference between Zai *et al.* (2016) and the Dstl experiment is that Zai *et al.* (2016) used a traditional western blot technique to qualitatively detect PA and then used an ELISA to quantify the concentration of PA, whereas Dstl used a new automated western blot system (Jess^{TM}) to quantify PA concentration. Zai *et al.* (2016) also conducted an experiment where they obtained the viable counts of bacteria (CFU/ml) and PA concentrations (ng/ml) from solutions where protease inhibitors were added to the flasks.

Charlton *et al.* (2007) simulated the UK anthrax vaccine manufacturing process, which uses the Sterne 34F2 strain (the strain used in the Dstl experiments), and obtained *in vitro* data on bacterial growth and PA concentration. Specifically, Thompson bottles had 450 ml of basal medium added and then were sterilized by autoclaving at 121°C for 15 min. The Thompson bottles were then warmed to 37°C prior to inoculation with 50 ml of a spore suspension with a concentration of 2×10^4 CFU/ml. This results in an initial spore concentration in the Thompson bottles of 2×10^3 CFU/ml. These Thompson

bottles were then incubated statically at 37°C for up to 32 hours. At various time points, three Thompson bottles were sacrificed, with the numbers of spore and bacterial CFU in the culture being measured and the PA concentration of the culture supernatants of these sacrificed bottles determined by using an antigen-capture ELISA. Individual bottle sacrificing was used because repeated sampling from the same bottle was found to disturb the growing cultures. The viable counts (CFU/ml) and the PA concentrations (μg/ml) are shown in Charlton *et al.* (2007, Figure 1) and Charlton *et al.* (2007, Figure 4) respectively.

One of the main differences between the experimental methods of the Dstl experiment and the ones used by Zai *et al.* (2016) is that whereas Dstl directly inoculated spores into the assay culture, Zai *et al.* (2016) grew the *B. anthracis* bacteria for 24 hours prior to this inoculation into the assay culture. Charlton *et al.* (2007) also directly inoculated spores into the assay culture, however these spores had previously been heat activated by heating the spore suspension at 60°C for 60 minutes, whereas the spores used in the Dstl experiment had not been heat activated. Another key difference between the experimental methods of the three studies is the type of culture medium used. Zai *et al.* (2016) inoculated bacteria into 100ml of Luria-Bertani (LB) liquid medium, Charlton *et al.* (2007) used basal medium, and Dstl used 50ml of BHI broth and also added sodium bicarbonate and CO_2. Finally, it may be important when interpreting the data to note that Charlton *et al.* (2007) incubated statically, whereas both the Dstl and Zai *et al.* (2016) experiments incubated with vigorous agitation.

4.2.2 Mathematical model

We want to investigate the specific dynamics of toxin production and decay, as this is a key driver in the pathogenesis of anthrax infection. Our motivation is to develop a model that is flexible enough to capture the behaviour of (at least) these two anthrax strains across a variety of experimental setups and methods whilst remaining able to accurately show the behaviour in each individual case. Therefore, we need to develop a model that captures the behaviours of the populations shown in the data, whilst incorporating biological processes that affect these populations in *in vitro* conditions. As the Zai *et al.* (2016) experiments have data for both strains, as well as an additional dataset with the presence of protease inhibitors, we begin by looking at that data.

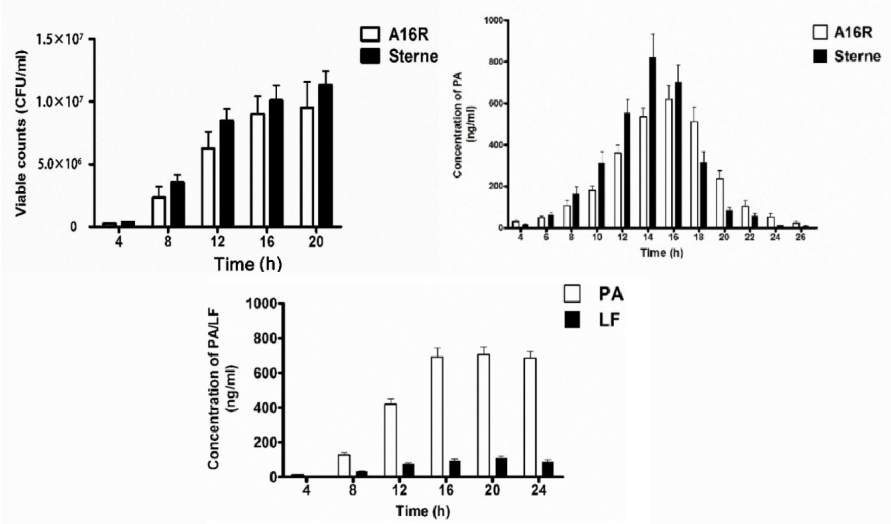

Figure 4.18: Figures taken from Zai *et al.* (2016). **Top-left:** Time course of the bacterial CFU for the A16R and Sterne strains. **Top-right:** Time course of PA concentration for the A16R and Sterne strains. **Bottom:** Time course of PA concentration for the A16R strain in the experiment with protease inhibitors.

We can clearly see the classic sigmoidal shape for logistic growth in the bacteria population so we propose a logistic growth differential equation for the vegetative bacterial population

$$\frac{dV(t)}{dt} = \lambda V(t)\left(1 - \frac{V(t)}{K}\right).$$

However, the behaviour of the PA population is less standard. We note from Figure 4.18 that the PA growth begins to slow at a similar time as to when the bacterial population begins to approach its carrying capacity. As bacteria stop producing PA as the bacterial population approaches its carrying capacity there must be some limiting factor causing this. We hypothesise that these bacteria may be unable to produce PA due to a lack of resources of some kind which causes the bacteria to stop production. It was discussed by Chiang *et al.* (2011) that a key factor in toxin production is the presence of glucose; they found where glucose was not present the production and virulence of toxin genes were inhibited greatly. Therefore we introduce a third population of "nutrients", $G(t)$, normalised to the experimental setup, meaning that the nutrients start the experiment with an initial concentration of $G(0) = 1$ and decays at a rate, α, proportional to the

interaction between this nutrient population and the bacterial population,

$$\frac{dG(t)}{dt} = -\alpha V(t)G(t).$$

We also note that from Figure 4.18 (bottom row) that if protease inhibitors are added to the system then the toxin population does not decay in the same way. It is believed some of the PA will actively be digested due to proteases produced by the bacteria (Pflughoeft *et al.* (2014)) and thus we need to introduce a mechanism of PA degradation in which the rate of PA removal, ν, is assumed to also be proportional to the number of vegetative bacterial CFU. We also need a second mechanism for the natural decay rate of PA, ν_0. Therefore our differential equation for the PA is

$$\frac{dP(t)}{dt} = \beta V(t-\tau_1)G(t-\tau_1) - \left(\nu_0 + \nu V(t-\tau_2)\right)P(t),$$

where τ_1 represents the delay in toxin production caused by the time taken for vegetative bacteria to produce PA, and τ_2 represents the delay in toxin degradation due to the vegetative bacteria taking time to produce the relevant proteases (not explicitly modelled) which are able to degrade PA.

However, we need this model to be flexible enough to represent not only multiple strains but various experimental conditions as well, so we also need to look at some of the other processes not incorporated in the Zai *et al.* (2016) data.

One key feature of the Dstl data not incorporated by the Zai *et al.* (2016) data is that Dstl has data for the spore population; we therefore need to introduce a time-dependent variable in our model for the spores, $S(t)$. It is known that *B. anthracis* spores cannot replicate, but first must undergo processes collectively known as germination to convert into a vegetative cell (Setlow (2003)). Therefore our differential equation for the spore population is

$$\frac{dS(t)}{dt} = -gS(t),$$

where g is the rate at which spores germinate. When a spore germinates, it produces a single newly germinated bacterium, which will grow into chains of rod-shaped cells, with each chain measured as 1 CFU in the experiments. However, it takes time for these newly germinated bacteria to grow into a chain of cells. Thus we incorporate this in the model

by splitting the bacterial population into two subpopulations, newly germinated bacteria, $N(t)$, and vegetative bacterial chains, $V(t)$. The newly desporulated bacteria grow into vegetative bacterial chains with rate m and the vegetative bacterial chains (measured in CFU) proliferate at a rate which is modelled by logistic growth, as explained above.

Thus, we propose a deterministic delay differential equation (DDE) model of the *in vitro* experiments described in Section 4.2.1, given by the following system of DDEs:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -gS(t), \\
\frac{dN(t)}{dt} &= gS(t) - mN(t), \\
\frac{dV(t)}{dt} &= mN(t) + \lambda V(t)\left(1 - \frac{V(t)}{K}\right), \\
\frac{dG(t)}{dt} &= -\alpha V(t)G(t), \\
\frac{dP(t)}{dt} &= \beta V(t-\tau_1)G(t-\tau_1) - (\nu_0 + \nu V(t-\tau_2))P(t).
\end{aligned} \tag{4.10}$$

We made the choice not to explicitly model the proteases as a population as there was no available data for this. We can also represent the model graphically as in Figure 4.19.

The initial conditions of the model variables are given by,

$$S(0) = fS_0^*,\ N(0) = \varepsilon B_0^*,\ V(0) = (1-\varepsilon)B_0^*,\ G(0) = 1,\ P(0) = 0,$$

where S_0^* and B_0^* are the number of spores and bacterial CFU measured (or inferred) at time zero, respectively. These depend on the experiment that is being modelled and will be discussed in more detail when we look at each dataset in turn. It has been observed in these type of experiments that often some small proportion of the initial spores never germinate. Therefore, the parameter f represents the fraction of initial spores that are able to desporulate, which are the ones represented by variable $S(t)$ in the model. The parameter ε represents the fraction of initial bacterial CFU that are in the newly germinated state at time 0.

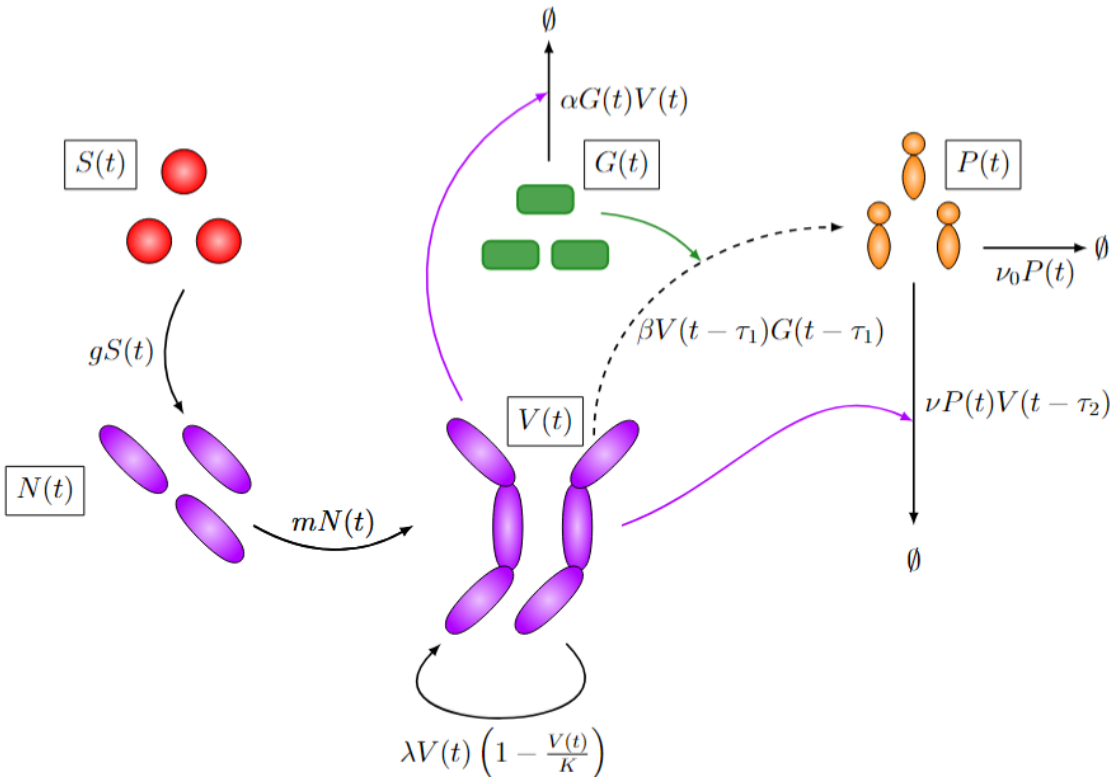

Figure 4.19: A schematic representation of the model in Equation (4.10). Black arrows represent species transitioning from one state to another, coloured arrows indicate that a population contributes to a particular reaction, and the dashed arrow represents toxin production.

The only non-trivial steady state of the model is $(S(\infty), N(\infty), V(\infty), G(\infty), P(\infty))= (0, 0, K, 0, 0)$. This state is stable, since all initial germinating spores and newly germinated bacteria will eventually become vegetative bacteria and the population of vegetative bacteria will approach the carrying capacity, K. Since the nutrients, $G(t)$, are being used up by the vegetative bacteria, this variable will approach zero, at which point no more PA can be produced and the PA will also decay to zero.

4.2.3 Parameter calibration

We calibrate parameters to the data using the approximate Bayesian computation-sequential Monte Carlo (ABC-SMC) algorithm described in Toni *et al.* (2009). The details of this algorithm are given in Section 2.3.2. We obtained 2000 accepted parameter sets for each dataset using 10 parallel versions of the ABC-SMC algorithm, each consisting of 20 iterations with an accepted sample size of 200 parameter sets and collating the last accepted set of parameter sets for each parallel version. We show the parameters and their interpretation in Table 4.3 and show their units and our selected priors for each parameter in Table 4.4.

Parameter	Description
f	Fraction of initial spores that are able to germinate
g	Germination rate of spores into newly desporulated bacteria
ε	Fraction of initial bacterial CFU that are newly desporulated
m	Maturation rate of desporulated spores into vegetative bacteria
λ	Rate of vegetative bacterial growth
K	Bacterial carrying capacity (per ml)
α	Rate that bacteria use up nutrients
β	PA production rate
ν_0	PA natural decay rate
ν	PA decay rate induced by bacteria (due to proteases)
τ_1	Time delay in PA production
τ_2	Time delay in protease production and toxin degradation

Table 4.3: Parameters in the mathematical model along with their biological interpretations.

Parameter	Unit	Prior
f	-	$f \sim U(0,1)$
g	h^{-1}	$\log_{10} g \sim U(-3,1)$
ε	-	$\varepsilon \sim U(0,1)$
m	h^{-1}	$\log_{10} m \sim U(-3,1)$
λ	h^{-1}	$\log_{10} \lambda \sim U(-1,1)$
K	CFU	$\log_{10} K \sim U(6,9)$
α	$(\text{CFU} \cdot \text{h})^{-1}$	$\log_{10} \alpha \sim U(-12,-3)$
β	$\text{ng} \cdot (\text{CFU} \cdot \text{h})^{-1}$	$\log_{10} \beta \sim U(-7,0)$
ν_0	h^{-1}	$\log_{10} \nu_0 \sim U(-6,0)$
ν	$(\text{CFU} \cdot \text{h})^{-1}$	$\log_{10} \nu \sim U(-15,0)$
τ_1	h	$\tau_1 \sim U(0,15)$
τ_2	h	$\tau_2 \sim U(0,24)$

Table 4.4: Parameters in the mathematical model along with their units, and prior distributions.

Uniform priors are used for each parameter (log-transformed in some cases) over the ranges in Table 4.4; we chose uniform priors over the parameter space so as to try to bias parameter selection as little as possible from our prior. Once our parameters have been selected and perturbed as described in Section 2.3.2, we move on to the model simulation step of the ABC-SMC algorithm; here once we have our model prediction we add noise to each simulated data point, to take into account measurement errors in the observed data (Alahmadi *et al.* (2020)). In the noise we choose added errors that are independent Gaussian with zero mean and standard deviation equal to the standard deviation of the experimental data at the corresponding time-point. It is important to note that this model is simultaneously being fitted to datasets of different types (bacterial CFU measurements and PA measurements) with different units (CFU vs ng/ml) and of different magnitudes (in the middle time points of our experiments bacteria is often of magnitude 10^7 and PA is of magnitude 10^3). Therefore, the distance function, if it is to include both populations, must consider quantities without units and still return values which are not adversely affected by the scale differences. This is possible, but it is difficult to ensure that the distance function does not give a higher priority to one of the populations. Therefore, we

use two different distance functions which will measure the distances in each population separately, meaning that we instead accept a parameter set if both distances fall below their own respective acceptance threshold. Namely, we define the following distances to compare model predictions with observed values.

$$D_1 = \sum_t \left(log_{10}\left(\frac{B_t^*}{B(t)}\right)\right)^2 + \left(log_{10}\left(\frac{S_t^*}{S(t)+(1-f)S_0^*}\right)\right)^2, \tag{4.11}$$

$$D_2 = \sum_t \left(P_t^* - P(t)\right)^2, \tag{4.12}$$

where $S(t) + (1 - f)S_0^*$, $B(t) = N(t) + V(t)$, and $P(t)$ are the model predictions from Eq. (4.10) (plus noise) for the amount of spores (desporulating and dormant), total bacterial CFU, and PA, respectively, at time t. B_t^* is the geometric mean of the observed number of bacterial CFU at time t, and P_t^* is the mean amount of PA observed at time t. S_t^* is the geometric mean of the observed number of spores at time t, which we only use for the Dstl dataset. We chose to use the log-transformed values for spore and bacterial CFU predictions as these populations span a large range of magnitude. As stated above, at each iteration of the ABC-SMC algorithm, two distance thresholds are generated from the accepted distances from the previous iteration and parameter sets are only accepted if D_1 and D_2 both fall below their respective distance thresholds.

4.2.3.1 Zai *et al.* (2016) dataset

Zai *et al.* (2016) conducted the experiments described in Section 4.2.1.3 using two different strains of *B. anthracis* - A16R and Sterne. For each strain separately, we fit the model in Eq. (4.10) to the data of viable counts and PA concentrations that they obtained, shown in Zai *et al.* (2016, Figures 1B and 4A, respectively).

The reported data is useful but has limitations. For example, for both strains the first data measurements given are at 4 hours and the initial conditions are not specified. Thus, we propose to calculate the initial condition for the bacteria. We do this by first fixing the bacterial population at $t = 4$ to the data point at that time; then we know that in the early time points the bacterial population will not be large enough that competition for resources will be much of a concern. Therefore, we assume at these time points

the bacteria are undergoing exponential growth; under this assumption we can work backwards to obtain the initial condition

$$B_0^* = \frac{B_4^*}{\left(1 - \frac{\lambda\varepsilon}{\lambda+m}\right) e^{4\lambda} + \frac{\lambda\varepsilon}{\lambda+m} e^{-4m}}, \tag{4.13}$$

where B_4^* is the data value at 4 hours. Furthermore, we have assumed that there are no spores present that are going to germinate, since it is reasonable to assume that all such desporulating spores would have already desporulated during the 24 hours of bacterial growth prior to inoculation into the assay culture. This means that we do not calibrate the parameters f and g for the Zai *et al.* (2016) datasets, since these parameters determine the dynamics of desporulating (or germinating) spores. We assume that there is no PA present at time 0. In summary, the initial conditions of the model are taken to be, $S(0) = 0$, $N(0) = \varepsilon B_0^*$ (where B_0^* is calculated from Eq. (4.13)), $V(0) = (1-\varepsilon)B_0^*$, $G(0) = 1$, and $P(0) = 0$.

A16R strain

For the A16R strain, an additional experiment was carried out by Zai *et al.* (2016), where a protease inhibitors cocktail was added to the LB culture medium. The measurements of PA concentration corresponding to this experiment are shown in the bottom row of Figure 4.18. In order to model the PA concentration in this experiment, we add the following equation to the system in Eq. (4.10),

$$\frac{dP_i(t)}{dt} = \beta V(t-\tau_1)G(t-\tau_1) - \nu_0 P_i(t), \tag{4.14}$$

with initial condition $P_i(0) = 0$. The variable P_i represents the PA concentration in an experiment in which protease inhibitors have been added. This has most of the same terms as the equation as for $P(t)$, but we set $\nu = 0$, as ν is the rate at which proteases decay PA and we assume that proteases will not be contributing to the degradation of PA due to the presence of protease inhibitors in the experiment. To include this additional model variable and data into the distance function, we adapt our second distance function to

$$D_2 = \sum_t \left(P_t^* - P(t)\right)^2 + \left(P_{i,t}^* - P_i(t)\right)^2, \tag{4.15}$$

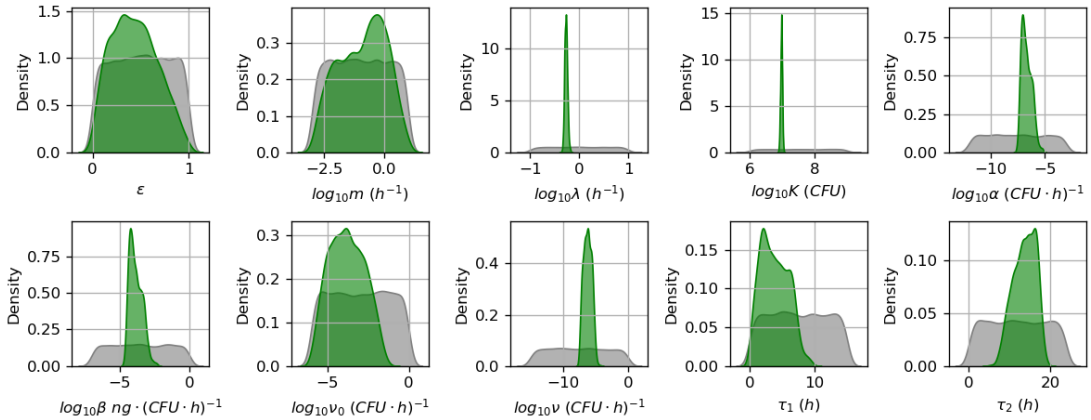

Figure 4.20: Prior distributions (grey) and kernel density estimates of the marginal posterior distributions (green) from fitting the model in Eq. (4.10) to the three datasets for the A16R strain from Zai *et al.* (2016).

where $P_i(t)$ is the model prediction (plus noise) for the amount of PA in an experiment with protease inhibitors at time t and $P^*_{i,t}$ is the mean amount of PA observed at time t in the experiment by Zai *et al.* (2016) using the A16R strain with protease inhibitors. In Figure 4.20 we show the marginal posterior distributions for each parameter in the accepted parameter sets of our ABC-SMC algorithm.

We can see clearly from Figure 4.20 that we can learn a significant amount about the regions in which the majority of parameters lie from these marginal posterior distributions of each parameter. For example, for this dataset we obtain very narrow distributions for λ and K. On the other hand, ε and m have relatively wide posterior distributions; these parameters determine the fraction of bacterial CFU that are initially newly desporulated bacteria, and the rate at which these progress into vegetative bacteria, respectively. It is perhaps unsurprising that we have not been able to learn significantly about these parameters, since the type of data used here does not allow one to distinguish between newly germinated and vegetative bacterial CFU. We can also compare these accepted parameter sets to the actual data by using the model predictions of the parameter sets. Figure 4.21 shows the predicted bacterial population and PA concentrations, both with and without protease inhibitors, versus the *in vitro* observations for the A16R strain; the solid lines represent the pointwise median of the model predictions from all parameter estimates in the posterior sample obtained via ABC-SMC, and the shaded regions represent the

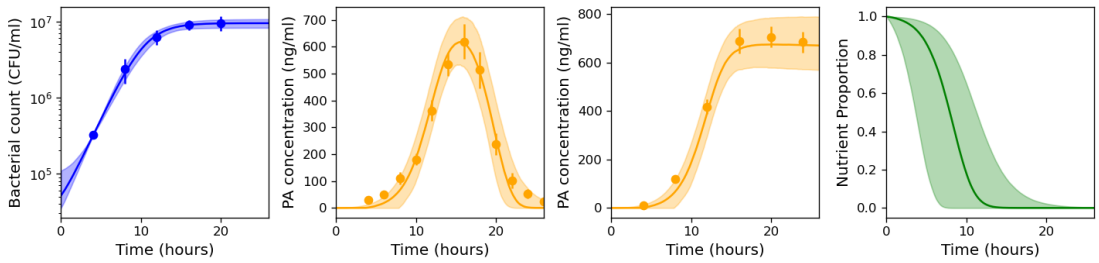

Figure 4.21: Pointwise medians and 95% credible intervals of the model posterior predictions for $B(t) = N(t) + V(t)$, $P(t)$, $P_i(t)$ and $G(t)$ (from left to right), using the parameter posterior distribution shown in Figure 4.20. The A16R strain experimental data used to fit the model are presented as mean $\pm$ standard error (SEM) from three independent experiment runs, extracted from Zai *et al.* (2016, Figure 1B (viable counts), Figure 4A (PA concentration), and Figure 7 (PA concentration in the presence of protease inhibitors)).

95% credible intervals (CI) of these model predictions. We can feel confident that our parameter calibration using ABC-SMC methods are working effectively as the model predictions agree well with data for all variables; our mathematical model is able to successfully explain the initial exponential bacterial growth reaching a carrying capacity, the increase and peak of PA concentration followed by rapid decay, and incorporate the impact of protease inhibitors in preventing a rapid decline in PA concentration when protease inhibitors are present.

We can look at the pairwise correlation between each of these parameters, as shown in Figure 4.22. This gives a good idea of which parameter pairs will be intrinsically linked when assessing parameter sets suitability to be accepted by the code. A lot of these correlations are unsurprising, for example the rate at which PA is produced in the system, β, is positively correlated with the time delay it takes bacteria to produce PA, τ_1. This makes logical sense as if the delay is longer, the PA will have to be produced quicker to reach the level observed in the data.

We can also look more closely at the pairs of parameters with a high absolute value of correlation coefficient; we chose the threshold value to be a magnitude of 0.35. There are therefore 5 parameter pairs that exceed the threshold for selection, and we look at all the accepted parameter values for these parameters in Figure 4.23. These figures show that some parameters depend on each other extremely closely, for example α and β; this is to

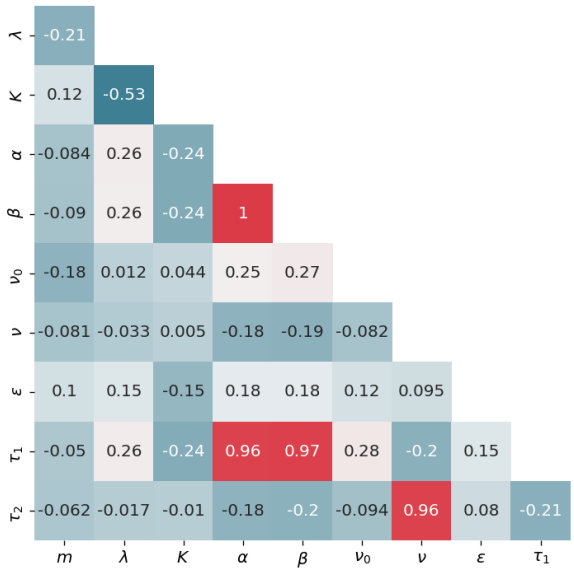

Figure 4.22: The pairwise correlation coefficients within the accepted parameter sets, the distributions for which are shown in Figure 4.20.

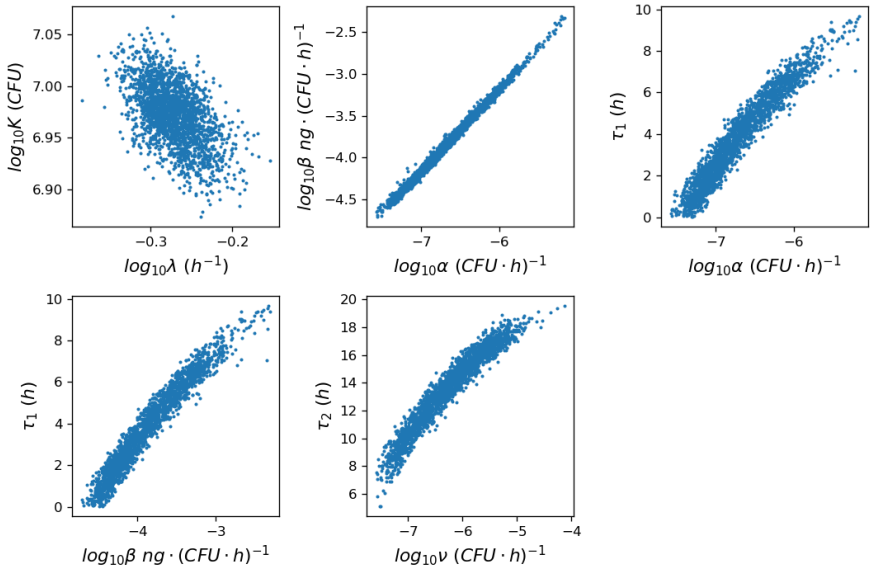

Figure 4.23: Selected pairwise parameter plots for the accepted parameter sets (those pairs with an absolute value of the correlation coefficient greater than 0.35 in Figure 4.22).

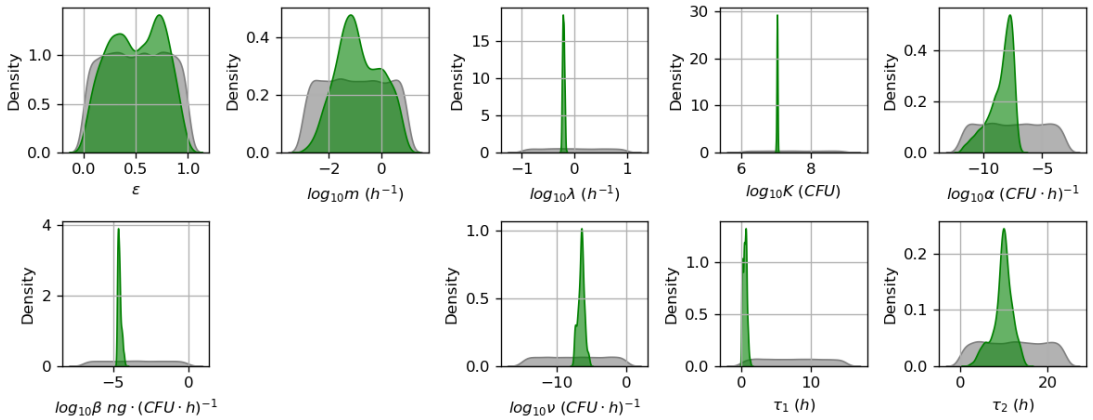

Figure 4.24: Prior distributions (grey) and kernel density estimates of the marginal posterior distributions (green) from fitting the model in Eq. (4.10) to the two datasets for the Sterne strain from Zai *et al.* (2016).

be expected as they govern the rate that bacteria use nutrients and the toxin production rate respectively.

Sterne strain

From the measurements in Figure 4.18, it is observed that the Sterne strain bacteria replicate faster and produce more PA than the A16R strain for these experiments. Hence we would expect most of the model parameter values to differ slightly between the two strains; it will be interesting how these differences manifest. The data acquired from the experiment with protease inhibitors was very useful in the A16R calibration, as having a dataset that removed the effect of protease decay on PA allowed us to more accurately estimate the production rate of PA. Unfortunately, this kind of data is not available for the Sterne strain. However, we believe that the value of the natural decay rate of PA, ν_0, which we were able to estimate using the A16R data, should be intrinsic to the PA protein itself. Therefore, in theory, ν_0 will not change due to the strain which produced the PA. We use this understanding to set a value for this parameter from the A16R calibration to be fixed when calibrating the model to the other datasets. In particular, from now on we set ν_0 to be equal to the median value from the posterior in Figure 4.20, giving $\nu_0 = 2 \times 10^{-4}\ h^{-1}$.

In Figure 4.24 we show the marginal posterior distributions of the parameters in green

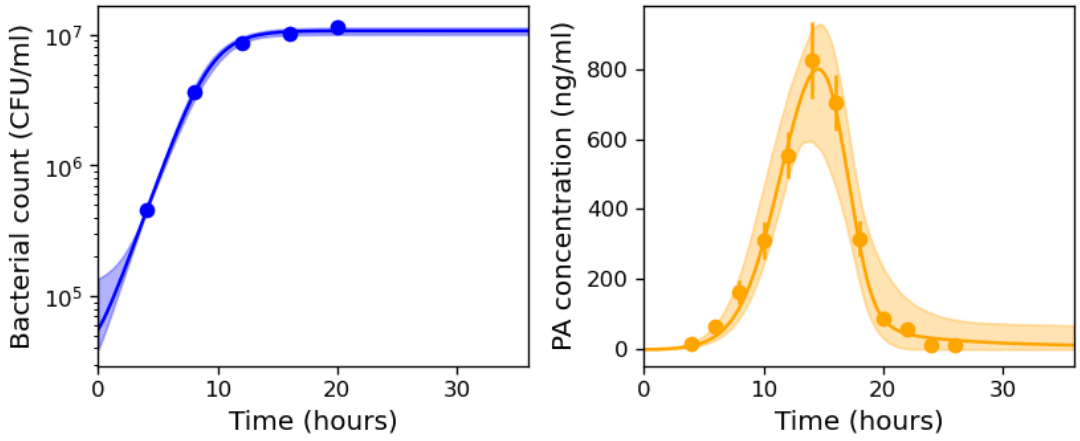

Figure 4.25: Pointwise medians and 95% credible intervals of the model posterior predictions for $B(t) = N(t) + V(t)$ (left) and $P(t)$ (right), using the parameter posterior distribution shown in Figure 4.24. The Sterne strain experimental data used to fit the model are presented as mean $\pm$ SEM from three independent experiment runs, extracted from Zai *et al.* (2016, Figure 1B (viable counts) and Figure 4A (PA concentration)).

against the prior distributions for these parameters. We endeavoured to layout these plots in a similar way to Figure 4.20 for ease of comparison, and therefore we have a gap in the second row and second column as we have a fixed value for ν_0 and therefore have neither a prior or posterior distribution for ν_0. In Figure 4.25 we show the pointwise median and 95% confidence interval for the 2000 simulations of the model, one for each accepted parameter set obtained from the ABC-SMC algorithm, against the data from Zai *et al.* (2016).

In general, the posterior distributions in Figure 4.24 for the Sterne strain calibration are narrower than the ones in Figure 4.20 for the A16R strain, and this is reflected in narrower 95% credible intervals for the posterior predictions in Figure 4.25. The exceptions to this are once again, ε and m, which, as discussed for the A16R strain, are difficult to learn significantly about as the $N(t)$ and $V(t)$ populations are not differentiated in the data. We have already discussed that the data for the Sterne strain results in higher bacterial and PA counts, therefore it is unsurprising that here the Sterne strain bacteria divide slightly faster and grow to a higher concentration, which is reflected in the posterior distributions of the corresponding parameters, λ and K. Furthermore, we also observed that the PA concentration data from Zai *et al.* (2016) shows a higher peak for the Sterne

strain than the A16R strain. This does not manifest in a higher PA production rate per CFU, as β is not estimated to be larger in the Sterne strain experiment. However, this behaviour is incorporated via the rate of depletion of nutrients, α, as well as the delay until PA production is initiated, τ_1, as these are estimated to be very small. These changes result in the model predictions being able to capture this increased PA yield. The peak of PA concentration also occurs slightly earlier for the Sterne strain, reflected in smaller estimated values of τ_2. On the other hand, the slope of the subsequent decline in PA concentration, and hence estimates of the parameter ν, are fairly similar between the two strains.

In Figure 4.26 we present the pairwise correlation coefficients between the parameters within our model for the accepted parameter sets. We can see that we have some very strongly linked parameters. As we saw for the A16R strain, the parameters β and α are once again strongly positively correlated (although to a much lower extent than for the A16R data), and the delay for PA degradation to occur, τ_2, is highly correlated with the rate at which the PA is degraded by proteases from the bacteria, ν. This seems logical as if the bacteria takes longer to degrade the proteases, they must do it faster once it has started to get the observed behaviour. Similarly, the delay in toxin degradation, τ_2, is negatively correlated with the rate at which toxins are produced by the bacteria, β. For these pairs of parameters, and other pairs with a high absolute value of the correlation coefficient, we plot the pairwise scatter plots of the accepted parameter values in Figure 4.27.

We can see from the plot concerning α and β in Figure 4.27, that there appears to be very little correlation between the two parameters until $\log_{10}(\alpha)$ becomes large enough, at which point they become strongly correlated. The threshold at which $\log_{10}(\alpha)$ becomes large enough to start this correlation effect is approximately -7.6, which was the smallest value accepted for $\log_{10}(\alpha)$ in the parameter calibration for the A16R strain. We therefore investigate the effect of splitting parameter sets into those which have a value of $\log_{10}(\alpha)$ below -7.6 and those that have a value above this threshold in Figure 4.28. In the top row of Figure 4.28, each plotted point represents an individual parameter set in the Sterne strain posterior sample coloured depending on whether α is smaller (purple) or larger (green) than the threshold value. As discussed, there is no significant correlation between the value of α and the value of β below this threshold but a positive correlation

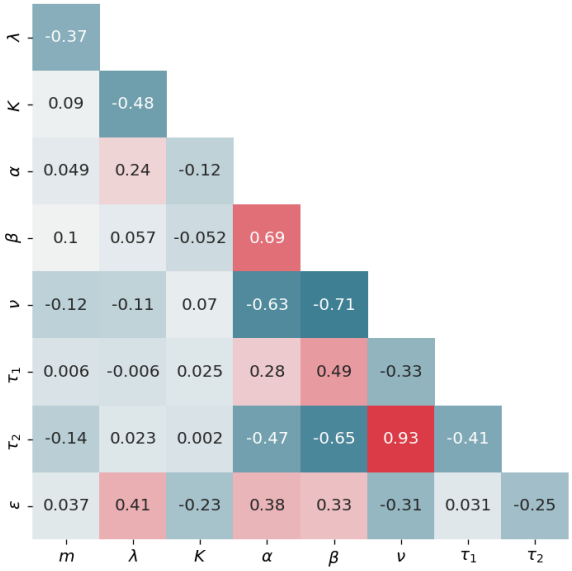

Figure 4.26: The pairwise correlation coefficients within the accepted parameter sets, the distributions for which are shown in Figure 4.24.

emerges above this threshold. A similar effect is seen between the value of α and ν; a negative correlation is observed once α becomes larger than the threshold value. This is because if nutrients are consumed faster (corresponding to a larger α), a larger maximal PA production rate (β) and a slower rate of PA degradation (ν) are needed to describe the data.

We can see from the middle row of plots in Figure 4.28, that whether the value of $\log_{10}(\alpha)$ is above or below -7.6 does not have a large effect on the model fit to the bacteria, whilst below the threshold value the fit to the PA data is slightly better than above the threshold. However, note that there is one less dataset to help us obtain the correct parameters in the Sterne set compared to the A16R set, namely there was no experiment done with protease inhibitors. From the A16R data we can intuit that if this experiment had been carried out it is possible that the PA population would remain constant instead of decaying to 0. We can simulate the hypothetical scenario in which protease inhibitors are added by removing the component of PA decay that involves bacterial proteases. Mathematically this is equivalent to fixing the parameter $\nu = 0$. In the bottom row we show the predictions for the variables we do not have data for: the simulated nutrients, and the hypothetical experiment where protease inhibitors were added to the experiment.

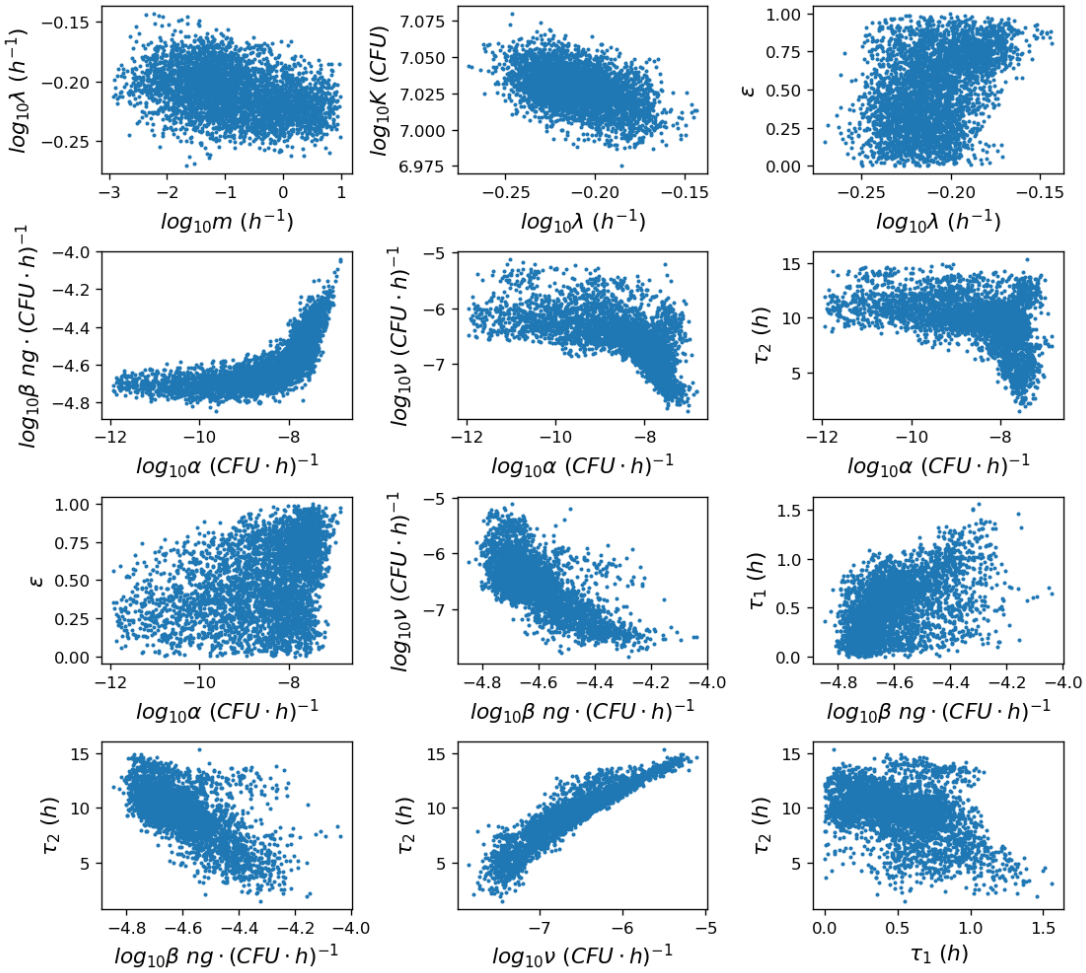

Figure 4.27: Selected pairwise parameter plots for the accepted parameter sets. Parameter pairs were selected if they had an absolute value of the correlation coefficient greater than 0.35.

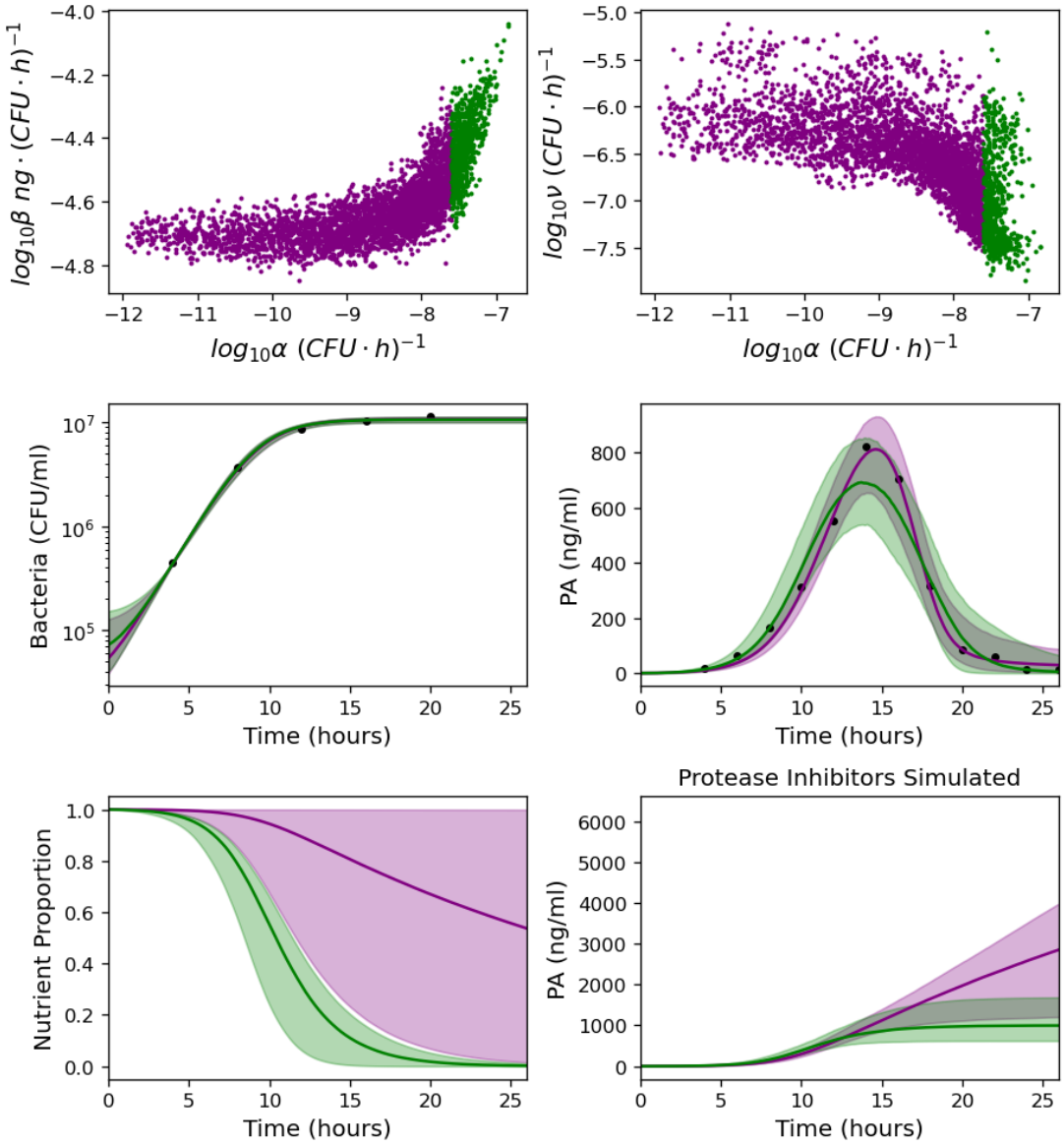

Figure 4.28: Selected pairwise parameter plots and population levels over time for parameter values of $\log_{10}(\alpha)$ below -7.6 (purple) and above -7.6 (green).

We observe that nutrients are used up much more rapidly when α is larger and that there will still be a relatively large proportion of nutrients left when α is smaller than this. However, for our simulated experiment with protease inhibitors it is observed that for the smaller α parameter region the PA population will continue to grow. This seems unlikely to be realistic given our observations from the experiment with protease inhibitors for the A16R strain. Therefore, an experiment using protease inhibitors would be crucial to determine which case is reality. However, without this experiment it seems more consistent with the A16R dataset that the rapid decay of toxins observed is due to the depletion of nutrients and therefore we believe that the parameter region where $\log_{10}(\alpha)$ is above -7.6 is the more realistic interpretation.

4.2.3.2 Charlton *et al.* (2007) dataset

In this section we fit our model (Equation (4.10)) to the data of bacterial counts and PA concentrations obtained by Charlton *et al.* (2007), presented in Charlton *et al.* (2007, Figures 1 and 4). One key experimental difference between the data presented within Charlton *et al.* (2007) and the data within Zai *et al.* (2016) is that there is data provided for the spore count. In Charlton *et al.* (2007, Figure 1) we can observe that the spore counts of each bottle remained fairly constant throughout the 32 hours, approximately at 30% of the number of spores used to inoculate each bottle. We hypothesise that this is likely to be because the spores were heat activated prior to inoculation, so would have germinated quickly on contact with the glucose and amino acids of the culture media. It is our understanding that by the time the first CFU measurements were obtained, the spores still within the system would not go on to germinate during the timescale of the experiment. Therefore, even though there is spore data presented in Charlton *et al.* (2007), we do not use it in our subsequent model calibration, as the data concerning the spores that germinate is not captured, similarly to Zai *et al.* (2016), and thus we do not calibrate f and g from this experiment.

As there is no data provided for the initial conditions of the bacteria, we follow a similar process to the one we carried out for the Zai *et al.* (2016) data. Here, we have fixed the number of bacteria at 2 hours (the first reported measurement) in the model to be equal to the data point at that time. We then set the initial conditions to $S(0) = 0$, $N(0) = \varepsilon B_0^*$,

$V(0) = (1-\varepsilon)B_0^*$, and $P(0) = 0$, where

$$B_0^* = \frac{B_2^*}{\left(1 - \frac{\lambda\varepsilon}{\lambda+m}\right) e^{2\lambda} + \frac{\lambda\varepsilon}{\lambda+m} e^{-2m}}.$$

Again, we fix the value of the natural PA decay rate to $\nu_0 = 2 \times 10^{-4}\ h^{-1}$, which is the median value obtained from the posterior in Figure 4.20.

We can compare the posterior distributions in Figure 4.29 to those in Figures 4.20 and 4.24, to see that the estimates for most parameters are fairly consistent between the experiments of Charlton *et al.* (2007) and Zai *et al.* (2016). However some parameters do differ in a significant way. For example, the rate of depletion of nutrients, α, and the maximal per CFU PA production rate, β, are estimated to be higher in the Charlton *et al.* (2007) experiment. It has been found that agitation can influence PA production by *B. anthracis*, possibly due to a change in the dissolved oxygen concentration of the assay culture (Mukhopadhyay (2008)). Therefore, a possible explanation for this increase in PA production rate could be the method of static incubation implemented by Charlton *et al.* (2007). Furthermore, within this data there is absence of an observed peak and subsequent decline in PA concentration which was crucial in calibrating certain parameters in the Zai *et al.* (2016) data. Therefore, calibration with the Charlton *et al.* (2007) dataset has not allowed us to learn significantly about the value of τ_2, representing the delay in the production of proteases by the bacteria. Another notable difference between the datasets is that the parameter ν, representing the rate of PA decay triggered by bacterial proteases, is estimated to be lower in the Charlton *et al.* (2007) experiment than the Zai *et al.* (2016) experiments, whilst having a relatively wide posterior distribution. This is expected; we do not capture the behaviour of a rapid decay of PA as it has not been observed in the Charlton *et al.* (2007) dataset, behaviour which is crucial for the estimation of ν. The median model predictions with a 95% confidence interval along with the data in Charlton *et al.* (2007) are presented in Figure 4.30. Within this figure it is clear the bacterial growth curve takes longer to reach the carrying capacity, compared with the Zai *et al.* (2016) experiments; this is to be expected since the initial amount of bacterial CFU is several orders of magnitude lower. It could be that the accumulation of proteases is also delayed in this experiment due to the lower level of bacteria in the initial conditions, which causes

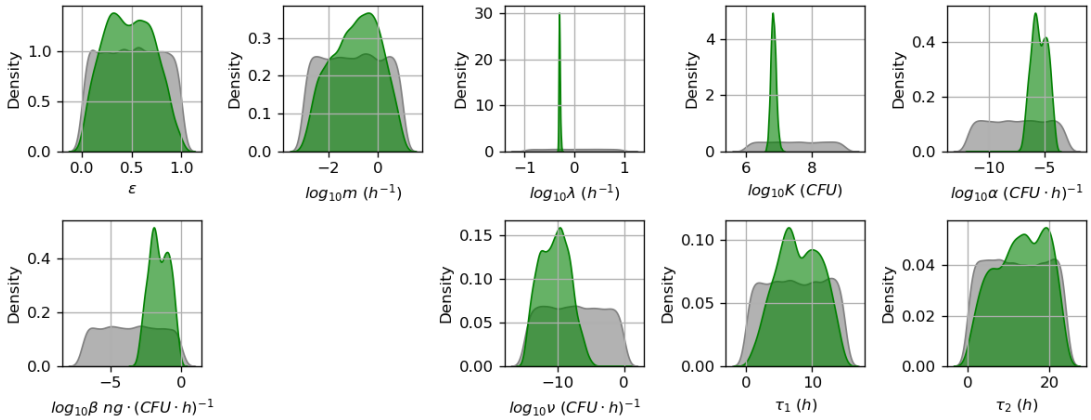

Figure 4.29: Prior distributions (grey) and kernel density estimates of the marginal posterior distributions (green) from fitting the model in Eq. (4.10) to the Charlton *et al.* (2007) datasets.

the lack of rapid decay in PA. Thus, it cannot be discounted that if measurements had been taken beyond $32h$, a PA decay might have been observed.

We can again see from Figure 4.31 that we once again have a strong relationship between α, β and in this case τ_1. The parameters τ_2 and ν are less strongly correlated than in Zai *et al.* (2016), although this is likely explained by the lack of significant learning about τ_2 for this dataset for the reasons previously discussed. We can further investigate these correlations by plotting scatter plots of accepted parameter sets by pairs of parameters in Figure 4.32. We can see that as $\log_{10}(\alpha)$ increases past the threshold value of -7.6 there is strong correlation between α and β, however not many parameter sets are accepted that have a value of $\log_{10}(\alpha)$ below the threshold value and as such it is not necessary to investigate the different profiles of behaviour either side of the threshold as was done in Figure 4.28.

4.2.3.3 Dstl dataset

In this section we fit the model in Equation (4.10) to the Dstl dataset provided by our collaborators, which is reported in Table 4.5. It is important to note that within this dataset we do include spore data within our model parameter calibration. The spores within this dataset were not heat treated prior to inoculation into the assay culture and they are observed to germinate much more slowly than in the Charlton *et al.* (2007)

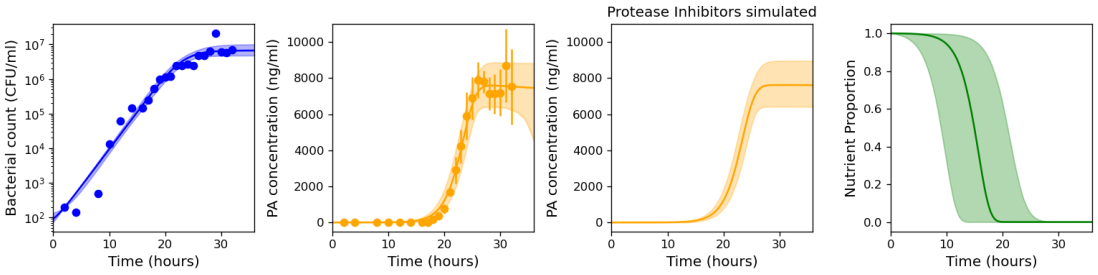

Figure 4.30: Pointwise medians and 95% credible intervals of the model posterior predictions for $B(t) = N(t) + V(t)$, $P(t)$, $P_i(t)$ and $G(t)$ (from left to right), using the parameter posterior distribution shown in Figure 4.29. The experimental data used to fit the model are presented as mean $\pm$ standard deviation from three independent Thompson bottles, obtained by Charlton *et al.* (2007). Numerical values were provided by Dr. Sue Charlton via private communication.

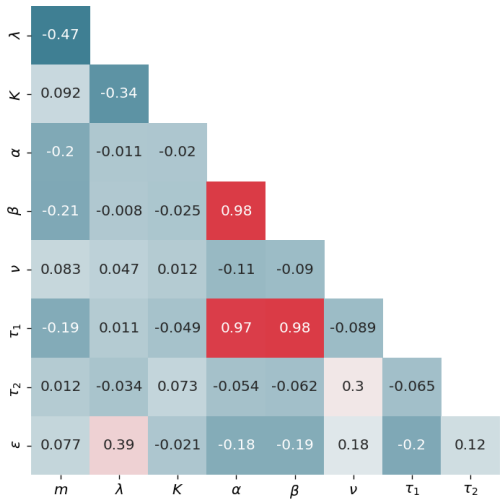

Figure 4.31: The pairwise correlation coefficients within the accepted parameter sets, the distributions for which are shown in Figure 4.29.

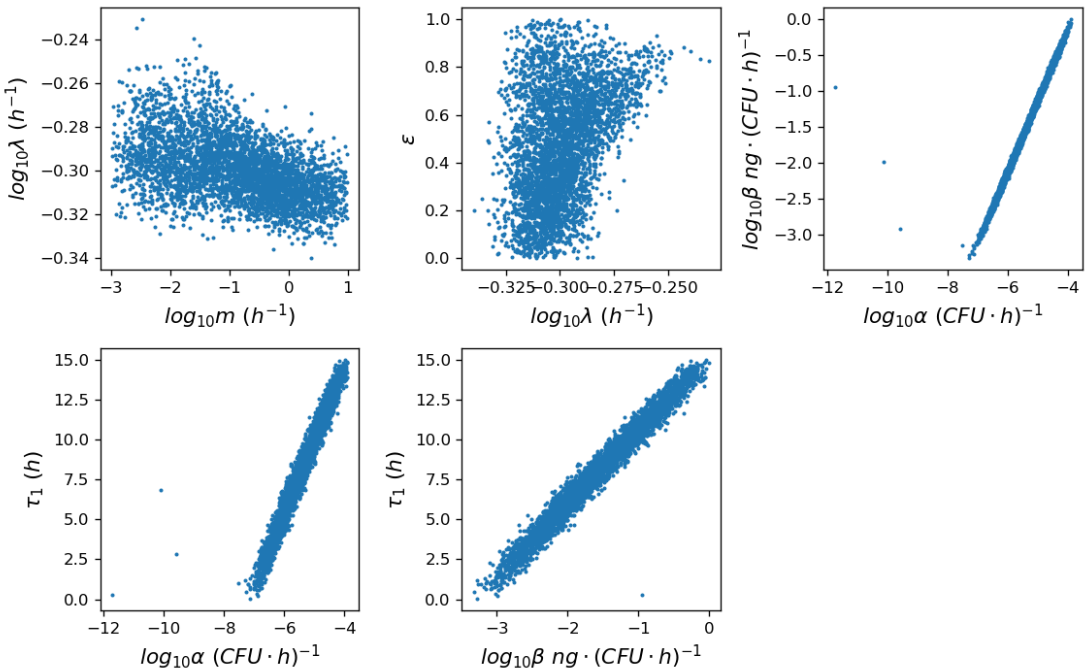

Figure 4.32: Pairwise parameter plots for the accepted parameter sets with correlation coefficient above the threshold magnitude of 0.35.

experiment. Therefore, this is the first dataset in which we calibrate the germination rate of spores, g, as well as the fraction of initial spores that are able to germinate, f. Furthermore, initial spore and bacterial counts were obtained as soon as possible after inoculation (1-2 minutes), so that we were able to fix the initial conditions of the model using experimental measurements, instead of inferring them by working backwards from a later time point. The initial conditions are therefore set to $S(0) = fS_0^*,\ N(0) = \varepsilon B_0^*,\ V(0) = (1-\varepsilon)B_0^*,\ G(0) = 1,\ P(0) = 0$, where S_0^* and B_0^* are the number of spores and bacterial CFU measured at time zero, respectively. Once again, we fix the value of the natural PA decay rate to $\nu_0 = 2 \times 10^{-4}\ h^{-1}$.

Time (hours)	$\log_{10}$(spores/ml)	$\log_{10}$(bacterial CFU/ml)	PA (ng/ml)
0	4.44 ± 0.09	4.34 ± 0.20	-
1.5	3.33 ± 0.14	4.73 ± 0.10	-
3.5	2.25 ± 0.18	5.48 ± 0.24	-
4.25	1.96 ± 0.59	5.74 ± 0.19	-
5	1.44 ± 0.12	6.20 ± 0.23	-
6	1.98 ± 0.64	7.05 ± 0.02	-
7	1.64 ± 0.66	7.45 ± 0.12	-
16	-	7.94 ± 0.06	194 ± 24
18	-	8.02 ± 0.03	386 ± 161
20	-	8.14 ± 0.02	1227 ± 511
22	-	8.18 ± 0.06	1950 ± 403
24	-	8.21 ± 0.14	2359 ± 289
40	-	7.96 ± 0.07	161 ± 0

Table 4.5: Data for the spore counts, bacterial counts, and PA concentration (mean $\pm$ standard deviation) obtained at Dstl following the experimental methods described in Section 4.2.1.

Figure 4.33 shows the prior and marginal posterior distributions of the parameters within our model. We can see that these posteriors are more consistent with those of the Zai *et al.* (2016) dataset than Charlton *et al.* (2007). This outcome is likely due to differences between the Dstl experimental setup and the experimental setup in Charlton *et al.* (2007), for example in the Dstl experiments spores were not heat activated prior to inoculation as

they were in Charlton *et al.* (2007). However, despite the consistency in most parameters between the Dstl data and Zai *et al.* (2016), we find that the delay in PA production by each bacterial CFU and the delay in production of proteases by the bacteria, given by τ_1 and τ_2, respectively, are estimated to be slightly longer in this experiment. This may be because spores were directly inoculated into the assay culture at the start of the experiment, whereas Zai *et al.* (2016) used a bacterial culture that had already been growing for 24 hours.

The model predictions and data in Figure 4.34 show quicker bacterial growth and a higher steady-state level of bacteria compared with the other experiments. Thus, while the estimates for the parameters concerning this behaviour, λ and K, are fairly consistent between the Zai *et al.* (2016) and Charlton *et al.* (2007) experiments, we estimate a faster bacterial growth rate, λ, and a higher carrying capacity, K, for the Dstl dataset. One possible explanation for this is that the BHI medium used in the Dstl experiment may have been richer than in the other experiments, which would have enabled the bacteria to divide faster and grow to a higher concentration.

We note that our estimates for the germination rate, g, are slightly higher than previous estimates of this rate obtained by Williams *et al.* (2021), but the order of magnitude is similar. Furthermore, our results predict that almost all initial spores will germinate, with f estimated to be very close to 1. This is in contrast to the observations from Charlton *et al.* (2007), where roughly 30% of spores did not germinate.

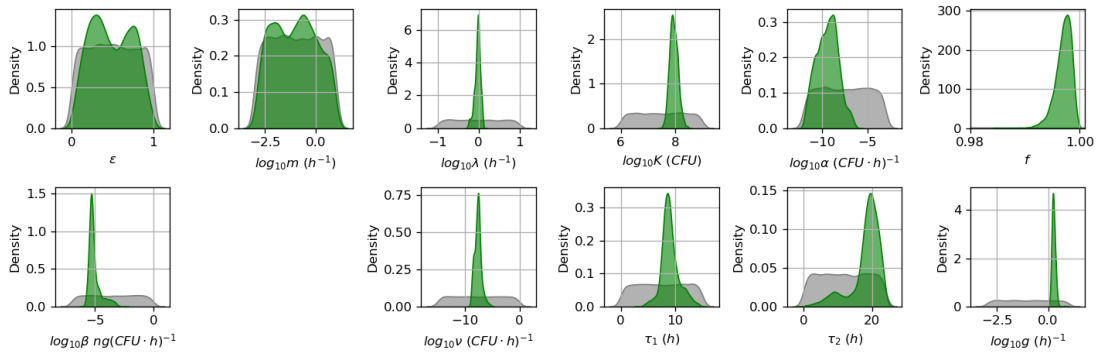

Figure 4.33: Prior distributions (grey) and kernel density estimates of the marginal posterior distributions (green) from fitting the model in Eq. (4.10) to the Dstl datasets in Table 4.5.

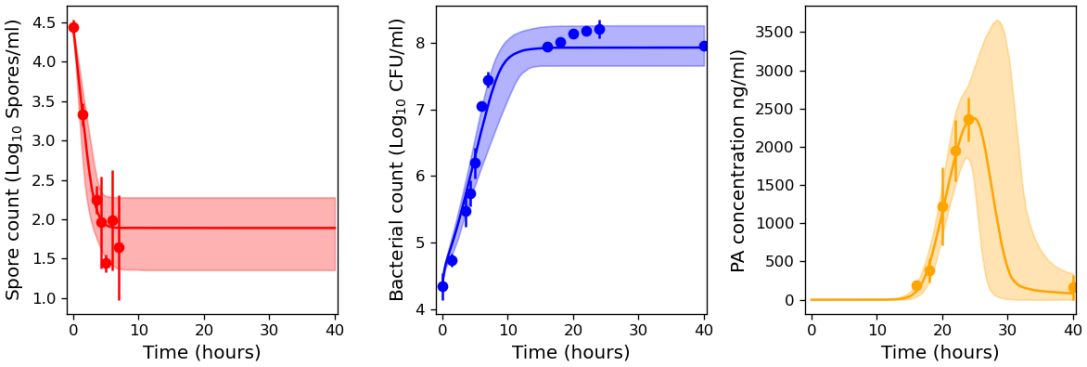

Figure 4.34: Pointwise medians and 95% credible intervals of the model posterior predictions for $S(t) + (1 - f)S_0^*$ (left), $B(t) = N(t) + V(t)$ (middle), and $P(t)$ (right), using the parameter posterior distribution shown in Figure 4.33. The experimental data used to fit the model are presented as mean $\pm$ standard deviation from three independent experiment runs, obtained from the Dstl experiment described in Section 4.2.1.

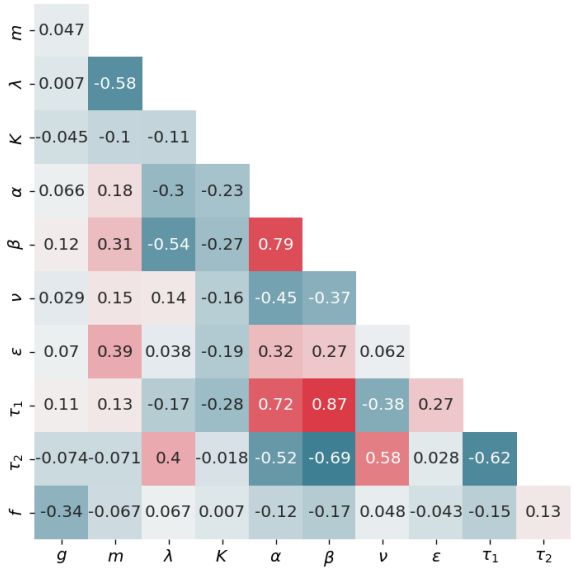

Figure 4.35: The pairwise correlation coefficients within the accepted parameter sets, the distributions for which are shown in Figure 4.33.

Figure 4.35 shows again the strong relationship between α and β and each of these with τ_1. The relationship between the τ_2 and ν is less strongly correlated than in the Zai *et al.* (2016) experiment but more strongly correlated than in the Charlton *et al.*

(2007) experiment; this is likely as the posterior for τ_2 in Figure 4.33 has a peak at 20 hours that mirrors the behaviour from the experiments in Zai *et al.* (2016) but also has a second small peak at approximately 10 hours that is not present in the Zai *et al.* (2016) calibration. We can further investigate these correlations by plotting scatter plots of accepted parameter sets by pairs of parameters in Figure 4.36, where again we focus on the parameter pairs which have a correlation magnitude of over 0.35. We can see that the relationships between these pairs of parameters are similar across the different datasets. There are more pairs of parameters which are above our chosen threshold value of a correlation magnitude above 0.35 in this dataset, one possible reason for which could be inclusion of the spore data within this experiment and the additional parameters f and g.

In Figure 4.36 it can be observed that a strengthening in the correlation occurs between α and β at approximately $\log_{10}(\alpha) = -7.6$, which is consistent with behaviour seen in calibrations for the Sterne dataset. Therefore, we investigate in Figure 4.37 the differences in the system between the parameter sets with α either side of this threshold. Values of α below this threshold show an extremely slow decay in the nutrient levels of the system (population $G(t)$); this corresponds to the PA not reaching a steady state in the simulated experiment with protease inhibitors. It is important to note that in this experiment the values of α above the threshold (in green) have a much tighter confidence interval in the behaviour of the PA than the values of α below the threshold (in purple). These values where $\log_{10}(\alpha) > -7.6$ appear to be more consistent across all datasets, likely due to the depletion of nutrients being the main mechanism behind PA decay.

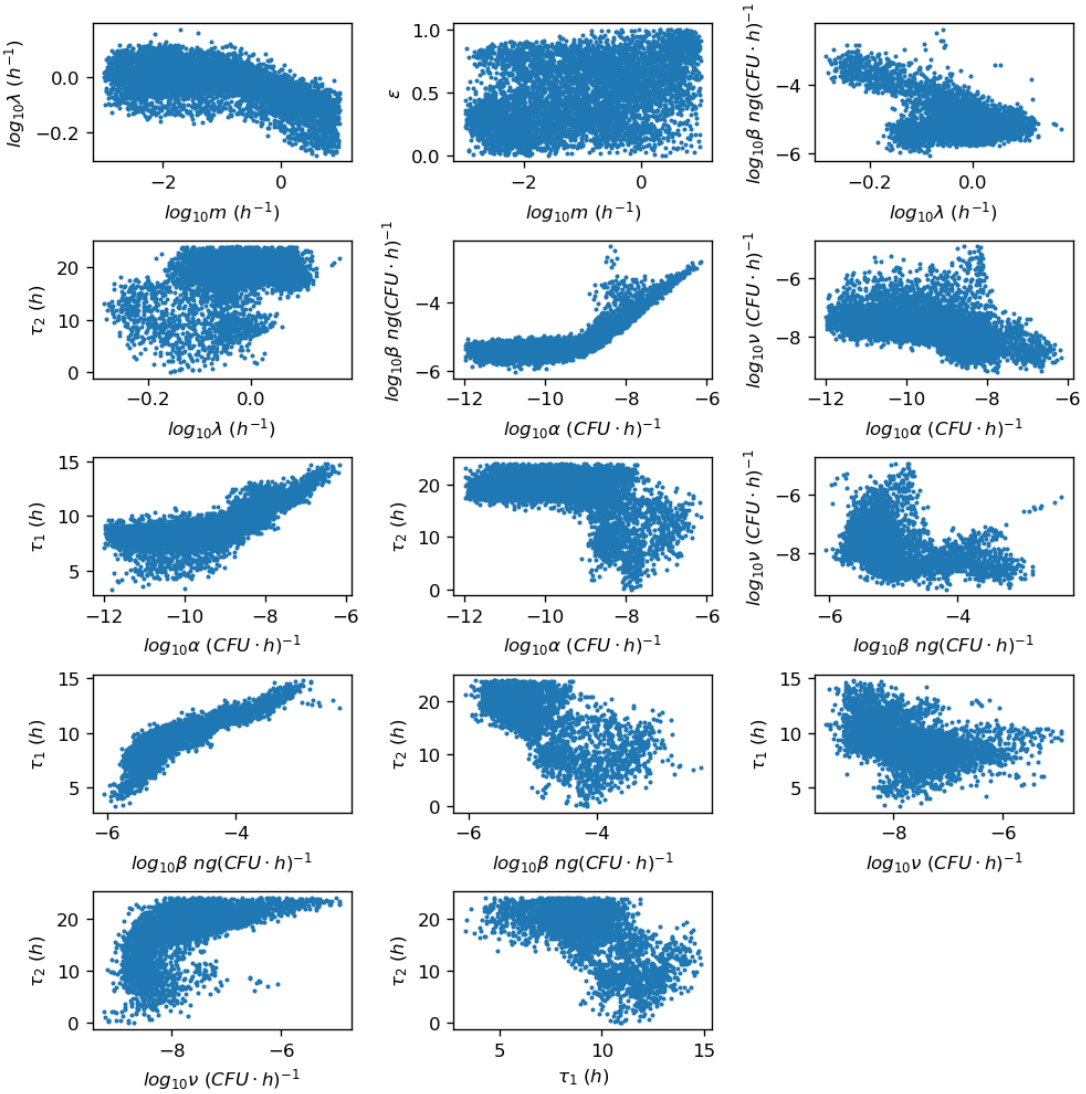

Figure 4.36: Selected pairwise parameter plots for the accepted parameter sets, with pairs of parameters selected if they had a pairwise correlation coefficient magnitude of greater than 0.35.

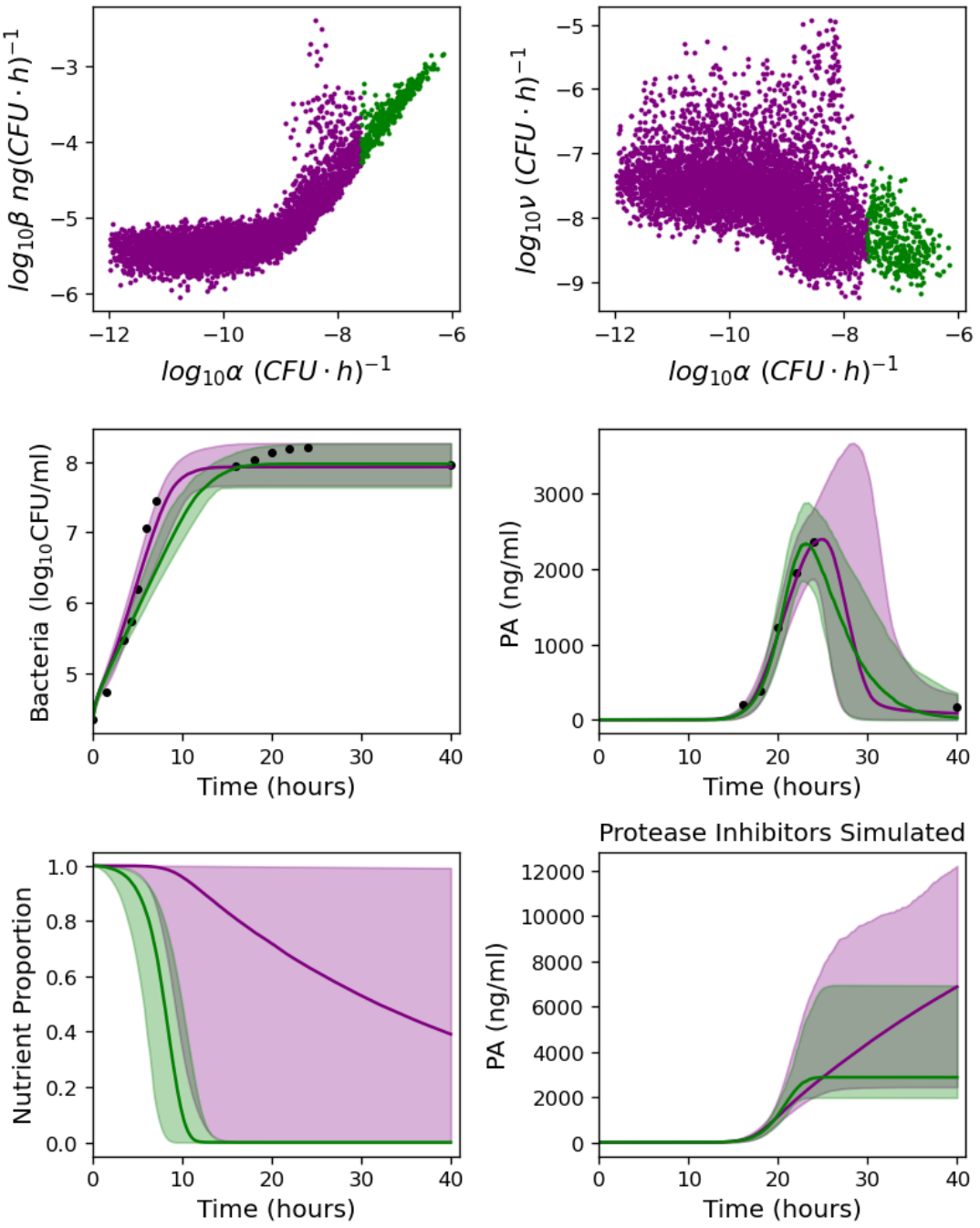

Figure 4.37: Selected pairwise parameter plots and population levels over time plots for parameter values of $\log_{10}(\alpha)$ below -7.6 (purple) and above -7.6 (green).

4.2.3.4 A summary comparison between datasets

Within this section we will summarise some of the key similarities and differences between our model predictions and the various datasets. Overall, the general consistency between most parameter estimates throughout the different experimental datasets, shows how the model we have proposed is flexible enough to capture behaviour of a variety of *Bacillus anthracis* strains or experimental conditions. In Table 4.6 we report summary

Parameter	Zai (A16R)	Zai (Sterne)
f	N/A	N/A
g	N/A	N/A
ε	$0.44\ (0.05, 0.89)$	$0.50\ (0.06, 0.92)$
m	$0.12\ (2\times10^{-3}, 5.0)$	$0.14\ (4\times10^{-3}, 4.9)$
λ	$0.53\ (0.47, 0.61)$	$0.62\ (0.57, 0.68)$
K	$9\times10^{6}\ (8\times10^{6}, 10^{7})$	$10^{7}\ (10^{7}, 1.1\times10^{7})$
α	$2\times10^{-7}\ (4\times10^{-8}, 2\times10^{-6})$	$2\times10^{-9}\ (6\times10^{-12}, 5\times10^{-8})$
β	$10^{-4}\ (3\times10^{-5}, 1\times10^{-3})$	$3\times10^{-5}\ (2\times10^{-5}, 5\times10^{-5})$
ν_0	$2\times10^{-4}\ (2\times10^{-6}, 10^{-2})$	$\mathbf{2\times10^{-4}}$
ν	$6\times10^{-6}\ (5\times10^{-8}, 8\times10^{-6})$	$3\times10^{-7}\ (3\times10^{-8}, 3\times10^{-6})$
τ_1	$3.8\ (0.46, 7.8)$	$3.3\ (1.14, 12.0)$
τ_2	$13.6\ (8.2, 17.9)$	$9.6\ (4.2, 13.6)$

Parameter	Charlton (Sterne)	Dstl (Sterne)
f	N/A	$0.997\ (0.993, 0.999)$
g	N/A	$1.84\ (1.29, 2.84)$
ε	$0.47\ (0.06, 0.90)$	$0.49\ (0.07, 0.92)$
m	$0.12\ (2\times10^{-3}, 5.4)$	$0.09\ (2\times10^{-3}, 6.6)$
λ	$0.50\ (0.48, 0.54)$	$0.93\ (0.61, 1.22)$
K	$7\times10^{6}\ (4\times10^{6}, 10^{7})$	$8\times10^{7}\ (5\times10^{7}, 2\times10^{8})$
α	$3\times10^{-6}\ (2\times10^{-7}, 6\times10^{-5})$	$4\times10^{-10}\ (3\times10^{-12}, 9\times10^{-8})$
β	$3\times10^{-2}\ (10^{-3}, 0.47)$	$9\times10^{-6}\ (2\times10^{-6}, 3\times10^{-4})$
ν_0	$\mathbf{2\times10^{-4}}$	$\mathbf{2\times10^{-4}}$
ν	$3\times10^{-11}\ (4\times10^{-15}, 5\times10^{-7})$	$2\times10^{-8}\ (2\times10^{-9}, 7\times10^{-7})$
τ_1	$8.0\ (1.97, 13.8)$	$9.0\ (5.88, 12.9)$
τ_2	$12.9\ (1.5, 22.8)$	$18.1\ (5.9, 23.3)$

Table 4.6: A comparison between the means and 95% credible intervals of the posterior distributions for each parameter after fitting the mathematical model to each dataset. The values of ν_0 in bold indicate fixed values that were used for calibration of the datasets.

statistics (mean and 95% credible intervals) for the posterior distributions obtained from the different datasets, in order to facilitate comparisons.

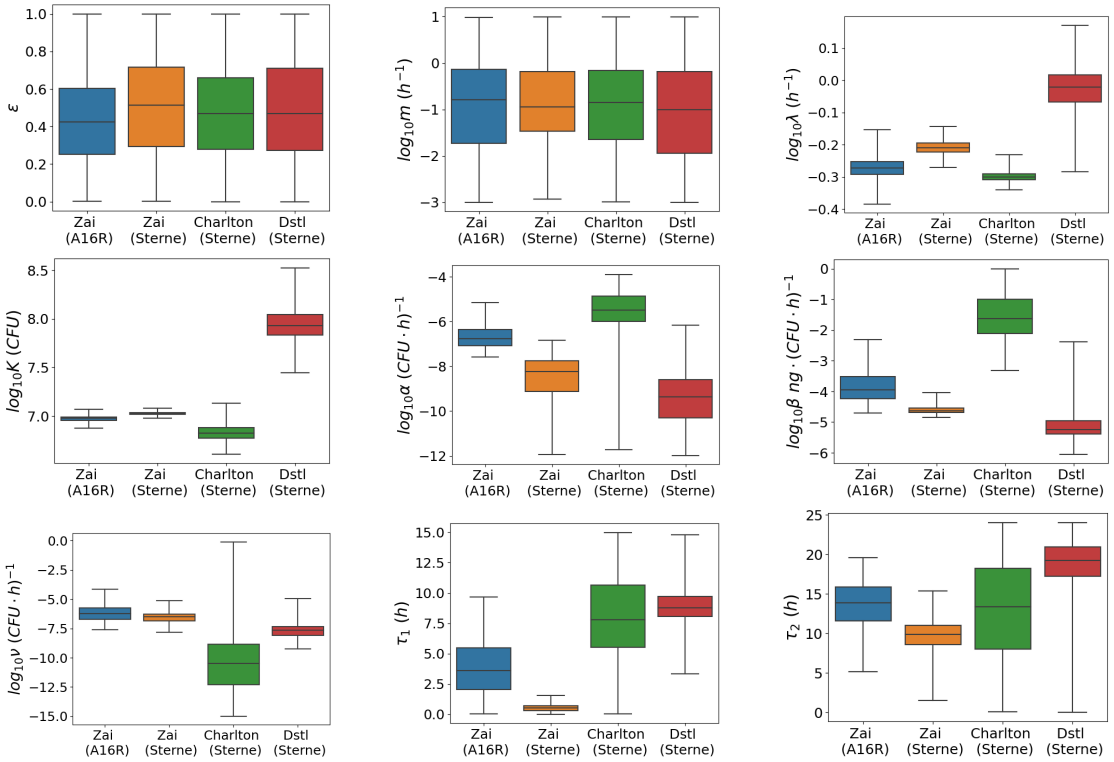

Figure 4.38: Box-plots showing the median, interquartile range, and range of each marginal posterior distribution, to illustrate how the posterior estimates of each parameter differ between datasets.

A visual comparison between the posterior distributions of each dataset is shown in Figure 4.38. For the parameters that determine the bacterial growth (the bacterial growth rate, λ, and the carrying capacity, K), the estimates from the Zai *et al.* (2016) and Charlton *et al.* (2007) datasets are fairly consistent, while the estimates from the Dstl data are slightly higher, possibly due to a richer BHI medium in the experimental setup. On the other hand, for parameters that influence the PA production and degradation (α, β, ν), estimates from the Zai *et al.* (2016) and Dstl data seem consistent, while estimates from the Charlton *et al.* (2007) data differ slightly. This is possibly due to the decay in the PA not being observed in the Charlton *et al.* (2007) data. The estimates for τ_1 are generally consistent between the Charlton *et al.* (2007) and Dstl datasets, whereas the delay until PA production is estimated to be much shorter for the Zai *et al.* (2016) data. This is likely due to the fact that in the Zai *et al.* (2016) experiments, the bacteria had already been growing for 24 hours prior to inoculation. The delay until protease production, τ_2, is estimated to be between 10-25 hours across all datasets, although Charlton *et al.* (2007) has a wider spread in the posteriors as the effect of the proteases is not captured at the time points in the data.

4.3 Discussion

In this chapter we have recreated the results of the model by Day *et al.* (2011), allowing us to investigate how significant the rate at which bacteria are released into the lymph nodes, k_3 in the model, is in the progression of an anthrax infection. We have also developed a stochastic analog of the model, using the Gillespie algorithm as a basis and some approximative techniques that allow this stochastic model to capture the behaviour for the population levels described by Day *et al.* (2011). This stochastic analog has allowed us to investigate the response probability for a given initial number of spores, allowing the basis for further investigation into this model and the capacity to expand its scope.

A possible extension of this stochastic model could be to carry out sensitivity analysis of the dose-response probabilities for parameters within the model, such as k_3, the rate at which bacteria is released into the lymph nodes. Another potential area of

investigation is to incorporate antibiotic treatment, as was done in Day *et al.* (2011) or via a pharmacokinetic-pharmocodynamic method, into the stochastic model.

We have developed and proposed a DDE model of *in vitro* dynamics of *B. anthracis* growth and PA production and degradation, using a new experimental dataset obtained by Dstl and making use of two other independent datasets by Zai *et al.* (2016) and Charlton *et al.* (2007). Our aim is that the model we propose is flexible enough to accurately capture the dynamics of various strains and culture conditions. To that end we have carried out parameter calibration for each dataset by means of ABC-SMC (Toni *et al.* (2009)). Many of the parameters are consistently estimated across different datasets, but there are a few notable differences. For each of these differences we have tried to explain possible causes for this, most likely due to the varying experimental conditions. For example, as discussed above, the data from the Dstl experiment shows quicker bacterial growth, reaching a higher steady-state level. We have hypothesised that this could be due to differences in the culture medium used in the experiments. Furthermore, a much higher PA yield was obtained in the Charlton *et al.* (2007) experiment, which is reflected in the corresponding parameter estimates. This may be due to the method of static incubation implemented by Charlton *et al.* (2007).

It is important to note that our proposed model distinguishes between natural PA decay and the decay caused by proteases secreted by the bacteria, such as InhA1. This decay due to proteases is implicitly included in the model via a term in which the rate of PA removal is assumed to be proportional to the number of vegetative bacteria that were present in the culture τ_2 hours ago, where τ_2 represents a delay taken for the bacteria to produce proteases. We investigated the parameter governing the natural decay rate of PA, ν_0, explicitly due to an experiment carried out by Zai *et al.* (2016). In this experiment, protease inhibitors were introduced to the system meaning that we assumed that the decay rate of PA due to proteases was neglible, $\nu = 0$. This allowed a more accurate estimation of ν_0. As this was the only experiment where we could isolate the rate ν_0 in this way, we used the median value of ν_0 for this experiment (with the A16R strain) as a fixed parameter value for the other datasets. In the future, additional experiments with protease inhibitors introduced to the system would help determine whether using ν_0 across experimental conditions and strains like this was valid. Additionally, protease concentration is not explicitly modelled as a variable within our model, as we are limited

in the available experimental data. However, if future experiments were able to obtain protease measurements, then the model could be adapted to include a more detailed description of the production of proteases, and their action on proteins produced by the bacteria (e.g. PA). This would enable us to move away from an implicit term of PA degradation in the model to a more explicit one. Computationally this may have a benefit too, as we would not need to calibrate τ_2 because this behaviour would be captured in the rate of protease production.

Future investigations to expand the scope of this work could be to discover whether these proteases that decay PA are also produced *in vivo*. The PA production rate predicted by the *in vitro* modelling results presented here may also be accurate *in vivo* as the bacteria should behave in a similar way. However, the degradation of PA is more uncertain within a host because there will likely be some degradation due to bacterial proteases, as well as proteases produced by the host.

The model variable that represents nutrient level is normalised by the initial nutrient level for each experiment, so that $G(0) = 1$. However, the type and amount of nutrients available to the bacteria will have varied significantly between the different studies. This implies that the estimated value of the maximal per CFU PA production rate, β, will depend on the nutrient availability in each individual experiment. In future, incorporating measurements of specific nutrients would help to unify the interpretation of parameters across studies. Furthermore, the model assumes that the per CFU PA production rate is proportional to the amount of nutrients. However, the relationship between nutrient level and PA expression by *B. anthracis* is likely to be more complicated than this. For instance, intermediate nutrient levels may provide the best environment for maximal PA production. Further experiments could be carried out with different specified nutrient levels to quantitatively investigate the impact of nutrient level on PA expression. This would assist in the calibration of parameters that describe the relationship between nutrient level and PA expression and would also help to incorporate a more realistic description of this relationship into a mathematical model.

We consider the behaviour of an explicit toxin population within this model as production of the anthrax toxin proteins is a key factor in the within-host survival of *B. anthracis*. Both lethal toxin and oedema toxin contribute to the severe symptoms suffered by a host

infected by *B. anthracis*, since they impact numerous functions of the immune system, for example by inhibiting the phagocytosis of bacteria by neutrophils. Quantifying PA production and degradation *in vitro* is an important step towards gaining a fuller understanding of *in vivo* toxin dynamics, since PA is the essential toxin component that facilitates binding of the other toxin proteins to cell surfaces. If we were to investigate the *in vivo* dynamics we may need to model the LF and EF populations too to monitor their suppressive effect on the immune system, similar to the way the toxin population worked in the model by Day *et al.* (2011), however for these *in vitro* experiments without an immune response, the PA population captures the behaviour of the anthrax toxin sufficiently.

A benefit of this mechanistic modelling approach is that we can extend and adapt the model as we learn more about the dynamics of the populations and more data is generated. This is evidenced in the development of our model to this date, how we were able to incorporate separate terms in the decay rate of PA due to the experiment with protease inhibitors in Zai *et al.* (2016). As this model can be adapted and extended, it would be possible to use it and the parameter estimates obtained across the four datasets to form a preliminary framework for future within-host modelling efforts for anthrax. For example, if we wanted to calibrate a within-host anthrax model with animal study data we could use the parameter estimates for bacterial growth rate and PA production rate as priors; in fact this has already been shown by a collaborator in Williams (2022).

The understanding of PA dynamics gained through this work will also be valuable for the development of future mechanistic within-host models that incorporate medical treatments for anthrax, such as anti-toxin treatments. These type of models could be developed through coupling pharmacokinetic (PK) data that describes how the within-host concentration of the treatment will change through time (Nagy *et al.* (2018); Subramanian *et al.* (2005)), with a pharmacodynamic (PD) description of the binding rate of PA as a function of anti-PA antibody concentration.

Chapter 5

A stochastic model of viral transmission in the presence of defective interfering particles

5.1 Introduction

In previous chapters, we have implemented stochastic processes to model *in vitro* and within-host infection dynamics. Here, we aim to show how similar methods can also be used to model infection transmission dynamics at the population level. In particular, we consider in this chapter how a wild-type (WT) virus spreads through a population of individuals in the presence of defective interfering particles (DIPs). As described in Section 1.1.2, DIPs lack some key components of the virus needed to replicate within the host, due to deletions or mutations in their genome (Frensing (2015)). They thus require co-infection with the WT virus, under these circumstances called a 'helper' virus, in order to replicate (Perrault (1981)). The process of the WT strain helping the DIP reproduce, however, interferes with the production of WT virus in a co-infected host, leading to reduced WT viral load (Frensing (2015)), which can impact on prognosis (i.e probability of recovery vs death, and timescales) as well as infectivity; this impact has led to the study of DIPs as a potential therapeutic agent (Marriott & Dimmock (2010)). DIPs can be studied at different macroscopic levels depending on the focus of the investigation.

For example, there has been research into how the populations of WT-virus and the DIP strain interact and compete for resources under *in vitro* experimental conditions (Thompson & Yin (2010)). It is also of interest to analyse these interactions between WT and DIP strains of the virus to investigate their impact as potential therapeutic agents, for example for the virus SARS-CoV-2 (Locke *et al.* (2024)). While the field of viral epidemic research has been investigated in great detail (Shil (2016)), there has been less research into these epidemic models in the presence of a DIP strain of virus. Therefore, within this chapter we will introduce a mathematical model at the population level to investigate how the DIP impacts infectivity and outcomes when present in a closed population with the WT virus.

The aim of this chapter is to introduce a compartmental epidemic model with susceptible individuals who can be infected by the DIP-strain, WT-strain or co-infected and describe summary statistics for the model. We are also interested in the impact of the DIP on the outcomes for individuals infected with the WT-strain, either on its own or as a co-infection, and as such track whether an individual recovers or dies. To be more precise, we will investigate the effect that the DIP-related parameters in this model have on these summary statistics, such as the number of individuals affected by the outbreak or the cumulative number of co-infected individuals within the length of the outbreak.

5.2 Compartmental epidemic model

We consider a closed population of N individuals and aim to model the spread of WT virus in this population in the presence of DIPs. In particular, we consider the following assumptions:

- Each individual in the population can either be susceptible, DIP-infected, WT-infected, co-infected or "removed" from the population (either through recovery or death).
- DIPs lack the key components of the virus needed to replicate within the host, therefore an individual only infected with a DIP would be asymptomatic and not infectious.

- Due to the previous assumption, DIPs in the absence of the WT virus decay, and a DIP-infected individual would become susceptible again after some random time.
- A co-infected individual is infected with both the DIP and the WT virus.
- A co-infected individual can transmit either the DIP, the WT or both strains of the virus during contact with another individual.
- WT-infected and co-infected individuals can either recover (gaining immunity) or die.

We represent the dynamics considered under these assumptions via a continuous-time Markov chain (CTMC) $\mathcal{X} = \{\mathbf{X}(t) = (S(t), I_D(t), I_W(t), I_C(t)) : \ t \geq 0\}$, where $S(t)$ represents the susceptible population and $I_W(t), I_D(t)$ and $I_C(t)$ represent the number of people infected with the WT-virus, the DIP or co-infected, respectively, at time t. We note that since infected individuals can either recover or die, we can define $R(t)$ and $D(t)$ as the number of recovered and dead individuals at time t respectively. Therefore,

$$S(t) + I_D(t) + I_W(t) + I_C(t) + R(t) + D(t) = N,$$

for all times $t \geq 0$. However, since removed individuals will not play any role in the infection dynamics, they do not need to be explicitly considered as random variables in the CTMC, $\mathcal{X}$, and since $R(t) + D(t) = N - S(t) - I_D(t) - I_W(t) - I_C(t), \forall t \geq 0$.

We consider that the following events can occur within the system:

- DIP-infected individuals become susceptible individuals at rate δ, $I_D \to S$.
- Infection events:
  - A susceptible individual becomes a WT-infected individual from contact with a WT-infected individual with rate β_W, $S + I_W \to I_W + I_W$.
  - A DIP-infected individual becomes a co-infected individual from contact with a WT-infected individual with rate β_W, $I_D + I_W \to I_C + I_W$.
  - A susceptible individual becomes a WT-infected individual from contact with a co-infected individual with rate $\phi_W \beta_C$, $S + I_C \to I_W + I_C$, where ϕ_W is

the probability that a co-infected individual only passes on the WT strain during infectious contact.

- A susceptible individual becomes a DIP-infected individual from contact with a co-infected individual with rate $\phi_D\beta_C, S + I_C \to I_D + I_C$, where ϕ_D is the probability that a co-infected individual only passes on the DIP strain during infectious contact.
- A susceptible individual becomes a co-infected individual from contact with a co-infected individual with rate $(1 - \phi_W - \phi_D)\,\beta_C, S + I_C \to I_C + I_C$.
- A DIP-infected individual becomes a co-infected individual from contact with a co-infected individual with rate $(1 - \phi_D)\beta_C, I_D + I_C \to I_C + I_C$.
- A WT-infected individual becomes a co-infected individual due to contact with a co-infected individual at rate $(1 - \phi_W)\beta_C, I_W + I_C \to I_C + I_C$.

- Removal events:
  - A WT-infected individual recovers with rate $\varepsilon_W\rho_W, I_W \to R$, where ρ_W is the removal rate of a WT-infected individual and ε_W is the probability that a WT-infected individual recovers.
  - A WT-infected individual dies with rate $(1 - \varepsilon_W)\rho_W, I_W \to D$.
  - A co-infected individual recovers with rate $\varepsilon_C\rho_C, I_C \to R$, where ρ_C is the removal rate of a co-infected individual and ε_C is the probability that a co-infected individual recovers.
  - A co-infected individual dies with rate $(1 - \varepsilon_C)\rho_C, I_C \to D$.

The CTMC, $\mathcal{X}$, evolves over the space of states

$$\mathcal{S} = \{(s, i_d, i_w, i_c) \in \mathbb{N}_0^4 \mid s + i_d + i_w + i_c \leq N\},$$

where N is the population size and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. All possible one-step transitions between states in $\mathcal{X}$, corresponding to the different events described above, are depicted in Figure 5.1.

These events, in particular, lead to the following non-null infinitesimal transition rates:

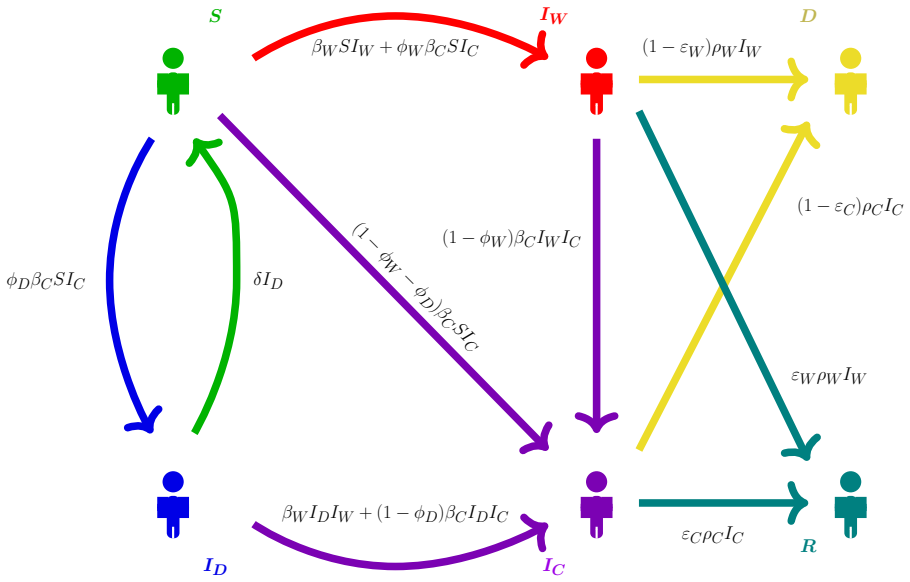

Figure 5.1: The single-step transition diagram for the model.

$$q_{(s,i_d,i_w,i_c),n'} = \begin{cases} \delta i_d, & \text{if } n' = (s+1, i_d-1, i_w, i_c), i_d > 0, \\ \phi_D \beta_C s i_c, & \text{if } n' = (s-1, i_d+1, i_w, i_c), s > 0, \\ (\beta_W i_w + \phi_W \beta_C i_c)s, & \text{if } n' = (s-1, i_d, i_w+1, i_c), s > 0, \\ (\beta_C(1-\phi_D)i_c + \beta_W i_w)i_d, & \text{if } n' = (s, i_d-1, i_w, i_c+1), i_d > 0, \\ \beta_C(1-\phi_W-\phi_D)s i_c, & \text{if } n' = (s-1, i_d, i_w, i_c+1), s > 0, \\ \beta_C(1-\phi_W)i_w i_c, & \text{if } n' = (s, i_d, i_w-1, i_c+1), i_w > 0, \\ (1-\varepsilon_W)\rho_W i_w + \varepsilon_W \rho_W i_w, & \text{if } n' = (s, i_d, i_w-1, i_c), i_w > 0, \\ (1-\varepsilon_C)\rho_C i_c + \varepsilon_C \rho_C i_c, & \text{if } n' = (s, i_d, i_w, i_c-1), i_c > 0, \end{cases}$$

for any $(s, i_d, i_w, i_c) \in \mathcal{S}$. We note that states of the form $(s, 0, 0, 0)$ are absorbing states, representing the end of the outbreak. When one of these absorbing states is reached, $N-s$ represents the number of recovered or dead individuals, which is typically referred to as the "size of the outbreak". We also note that although states $(s, i_d, 0, 0)$ are technically not absorbing (as transitions to $(s+1, i_d-1, 0, 0)$ can occur due to DIP decay), when we arrive in this state the outbreak has already finished, as the system is guaranteed to eventually arrive at the absorbing state $(s+i_d, 0, 0, 0)$ due to DIP decay. Due to this, we can consider these states as absorbing for the purpose of analysing the dynamics during an outbreak.

In the following sections stochastic descriptors of this model will be studied. In Section 5.2.1, the probability distribution of the number of infected individuals that recover and die respectively, throughout the course of an outbreak, is studied. We also provide the methodology of how an organisation of the space of states in $\mathcal{S}$ allows an algorithmic approach to calculating this stochastic descriptor exactly. In Section 5.2.1.6, the expected value of this descriptor is studied. Sections 5.2.2 and 5.2.3 describe the probability distribution and mean of the cumulative number of co-infected individuals in an outbreak and the exact reproduction number of a marked individual, respectively.

5.2.1 Size of the outbreak

If at any stage $I_W(t) + I_C(t) = 0$, then no individuals remain infected by the WT virus and as DIP-infected individuals are thought to be asymptomatic, we say the outbreak has finished. All individuals that have not been WT-infected or co-infected at the time the outbreak finishes will eventually become susceptible, and $\mathcal{X}$ will reach an absorbing state $(s, 0, 0, 0)$. We could use this value of s to calculate the number of people affected by the outbreak but the system may take some time to reach this state if there are a large number of DIP-infected individuals in I_D when the outbreak finishes. Still, we can define the length of the outbreak as the time

$$\tau = \inf\{t \geq 0 : I_W(t) + I_C(t) = 0\}. \tag{5.1}$$

Our interest is in the random vector

$$(D(\tau), R(\tau)),$$

representing the number of dead and recovered individuals at time τ when the outbreak finishes. The random variable that defines the size of an outbreak is

$$Y = R(\tau) + D(\tau) = N - S(\tau) - I_D(\tau). \tag{5.2}$$

However, it is more interesting to analyse the make up of this random variable Y, as it is made up of individuals that have either recovered or died from the outbreak. We can do this by considering a given initial state at time 0, meaning that our CTMC $\mathcal{X}$ is at state

$\mathbf{X}(0) = (S(0), I_D(0), I_W(0), I_C(0)) = (s, i_d, i_w, i_c)$. At this time, $t = 0$, we assume no recoveries or deaths have occurred previoulsy. We can then compute the joint probability mass function

$$\alpha_{(s,i_d,i_w,i_c)}(r,d) = \mathbb{P}(R(\tau) = r, D(\tau) = d | S(0) = s, I_D(0) = i_d, I_W(0) = i_w, I_C(0) = i_c), \tag{5.3}$$

for any $(s, i_d, i_w, i_c) \in \mathcal{S}$ and $i_w + i_c \leq r + d \leq N$. For any values of r and d that do not satisfy this inequality, $\alpha_{(s,i_d,i_w,i_c)}(r,d)$ is trivially 0. Since the outbreak ends when $I_W(t) + I_C(t) = 0$, we also have

$$\begin{aligned}
&\alpha_{(s,i_d,0,0)}(0,0) = 1, \\
&\alpha_{(s,i_d,0,0)}(r,d) = 0, \quad r + d > 0.
\end{aligned}$$

5.2.1.1 Organising the space of states $\mathcal{S}$

In order to efficiently compute probabilities $\alpha_{(s,i_d,i_w,i_c)}(r,d)$, we propose here to organise the state space $\mathcal{S}$ in terms of subsets called "levels" and "sub-levels", where levels, $L(i)$, are indexed by i, representing the total number of susceptible and DIP-infected individuals in the system, $s + i_d = i$. Sub-levels, $L(i,j)$, within $L(i)$ are indexed by j, representing the total number of individuals infected with the WT-strain, either on its own or via co-infection, $i_w + i_c = j$. Equivalently,

$$\mathcal{S} = \bigcup_{i=0}^{N} L(i), \quad L(i) = \{(s, i_d, i_w, i_c) \in \mathcal{S} : s + i_d = i\}, 0 \leq i \leq N,$$

$$L(i) = \bigcup_{j=0}^{N-i} L(i,j), \quad L(i,j) = \{(s, i_d, i_w, i_c) \in \mathcal{S} : s + i_d = i, i_w + i_c = j\},$$
$$0 \leq i \leq N, 0 \leq j \leq N - i.$$

It is important to note that the outbreak ends at the first time, τ, that the system arrives at $I_W(\tau) + I_C(\tau) = 0$. This is equivalent to the system arriving to any sub-level $L(i,0), 0 \leq i \leq N$. We note that states inside level $L(i)$ can be listed as in Figure 5.2. We can see that each sub-level $L(i,j)$ has size $\#L(i,j) = (j+1)(i+1)$. Therefore we

		$L(i)$		
$L(i,0)$	$L(i,1)$	$\cdots$	$L(i,N-i-1)$	$L(i,N-i)$
$(0,i,0,0)$	$(0,i,0,1)$	$\cdots$	$(0,i,0,N-i-1)$	$(0,i,0,N-i)$
$(1,i-1,0,0)$	$(0,i,1,0)$	$\cdots$	$(0,i,1,N-i-2)$	$(0,i,1,N-i-1)$
$(2,i-2,0,0)$	$(1,i-1,0,1)$	$\cdots$	$\vdots$	$\vdots$
$\vdots$	$(1,i-1,1,0)$	$\cdots$	$(0,i,N-i-1,0)$	$\vdots$
$(i,0,0,0)$	$(2,i-2,0,1)$	$\cdots$	$(1,i-1,0,N-i-1)$	$(0,i,N-i,0)$
$\Downarrow$	$(2,i-2,1,0)$	$\cdots$	$(1,i-1,1,N-i-2)$	$(1,i-1,0,N-i)$
$i+1$ states	$\vdots$	$\cdots$	$\vdots$	$\vdots$
	$(i,0,0,1)$	$\cdots$	$(i,0,0,N-i-1)$	$\vdots$
	$(i,0,1,0)$	$\cdots$	$(i,0,1,N-i-2)$	$(i,0,0,N-i)$
	$\Downarrow$		$\vdots$	$(i,0,1,N-i-1)$
	$2(i+1)$ states		$(i,0,N-i-1,0)$	$\vdots$
			$\Downarrow$	$(i,0,N-i,0)$
			$(N-i)(i+1)$ states	$\Downarrow$
				$(N-i+1)(i+1)$ states

Figure 5.2: Lexicographic ordering of the states within $L(i)$.

can calculate the size of each level, $\#L(i)$,

$$\begin{aligned}\#L(i) &= \sum_{j=0}^{N-i} \#L(i,j) = \sum_{j=0}^{N-i}(i+1)(j+1) = (i+1)\frac{(N-i+1)(N-i+2)}{2} \\ &= (i+1)\binom{N-i+2}{2}.\end{aligned}$$

We can also calculate the size of the space of states, $\#\mathcal{S}$.

$$\#\mathcal{S} = \sum_{i=0}^{N} \#L(i) = \sum_{i=0}^{N}(i+1)\binom{N-i+2}{2}.$$

It is worth noting that the number of states in $\mathcal{S}$ can be computed differently. In particular, it is clear that

$$\mathcal{S} = \bigcup_{n=0}^{N}\{(s,i_d,i_w,i_c) \in \mathbb{N}_0^4 : s+i_d+i_w+i_c = n\},$$

so that $\#\mathcal{S} = \sum_{n=0}^{N} \#\mathcal{S}(n)$ where $\#\mathcal{S}(n)$ is the number of compositions of n into 4 parts, which gives

$$\#\mathcal{S}(n) = \binom{n+4-1}{4-1} = \binom{n+3}{3} = \binom{n+3}{n}.$$

From this we can use the Hockey-stick identity to show that

$$\#\mathcal{S} = \sum_{n=0}^{N} \#\mathcal{S}(n) = \sum_{n=0}^{N} \binom{n+3}{n} = \sum_{n=0}^{(N+3)-3} \binom{n+3}{n} = \binom{N+3+1}{N+3-3} = \binom{N+4}{N}.$$

We can illustrate that these two methods of calculating the size of the space of states are equivalent by way of example, here lets take a population of $N = 3$. From the first method we have

$$\begin{aligned} \#\mathcal{S} = \sum_{i=0}^{3} (i+1)\binom{3+2-i}{2} &= 1 \cdot \binom{5}{2} + 2 \cdot \binom{4}{2} + 3 \cdot \binom{3}{2} + 4 \cdot \binom{2}{2} \\ &= 10 + 12 + 9 + 4 = 35, \end{aligned}$$

and by the second method we have $\#S = \binom{3+4}{3} = \binom{7}{3} = 35$.

5.2.1.2 Computation of probabilities $\alpha_{(s,i_d,i_w,i_c)}(r,d)$

In general, one can aim to compute probabilities $\alpha_{(s,i_d,i_w,i_c)}(r,d)$ (from Equation (5.3)) for any initial state (s, i_d, i_w, i_c) and values r and d via a first-step argument. In particular, by conditioning on the next even which can occur, one gets Equation (5.4).

$$\begin{aligned}\alpha_{(s,i_d,i_w,i_c)}(r,d) =& \frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s+1,i_d-1,i_w,i_c)}(r,d) + \frac{\phi_D\beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s-1,i_d+1,i_w,i_c)}(r,d)\\ &+ \frac{(\beta_W i_w + \phi_W\beta_C i_c)s}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s-1,i_d,i_w+1,i_c)}(r,d)\\ &+ \frac{\beta_C(1-\phi_W)i_w i_c}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s,i_d,i_w-1,i_c+1)}(r,d)\\ &+ \frac{\beta_C(1-\phi_W-\phi_D)s i_c}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s-1,i_d,i_w,i_c+1)}(r,d)\\ &+ \frac{(\beta_C(1-\phi_D)i_c + \beta_W i_w)\, i_d}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s,i_d-1,i_w,i_c+1)}(r,d)\\ &+ \frac{(1-\varepsilon_W)\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s,i_d,i_w-1,i_c)}(r,d-1)\\ &+ \frac{\varepsilon_W\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s,i_d,i_w-1,i_c)}(r-1,d)\\ &+ \frac{(1-\varepsilon_C)\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s,i_d,i_w,i_c-1)}(r,d-1)\\ &+ \frac{\varepsilon_C\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}}\alpha_{(s,i_d,i_w,i_c-1)}(r-1,d).\end{aligned} \tag{5.4}$$

Within Equation (5.4) and throughout the rest of this chapter, $\Delta_{(s,i_d,i_w,i_c)}$ is defined as

$$\begin{aligned}\Delta_{(s,i_d,i_w,i_c)} =& \delta i_d + \phi_D\beta_C s i_c + (\beta_W i_w + \phi_W\beta_C i_c)s + \beta_C(1-\phi_W)i_w i_c\\ &+ \beta_C(1-\phi_W-\phi_D)s i_c + (\beta_C(1-\phi_D)i_c + \beta_W i_w)\, i_d\\ &+ (1-\varepsilon_W)\rho_W i_w + \varepsilon_W\rho_W i_w + (1-\varepsilon_C)\rho_C i_c + \varepsilon_C\rho_C i_c.\end{aligned} \tag{5.5}$$

From Equation (5.4), we can already see that the probabilities, $\alpha_{(s,i_d,i_w,i_c)}(r,d)$, depend on each other. Therefore, we aim to algorithmically solve this system of equations, which consists of one equation for each possible initial state $(s,i_d,i_w,i_c) \in \mathcal{S}$ and pair of values (r,d).

5.2.1.3 States in level L(0)

Let us first consider the states $(s,i_d,i_w,i_c) \in L(0)$, so that $s = i_d = 0$. From our model in Figure 5.1, we note that each WT and co-infected individual gets removed

at rate ρ_W and ρ_C, respectively. When removal occurs, the individual either recovers with probability ε_W for WT infections and ε_C for co-infections, or dies with probability $(1-\varepsilon_W)$ for WT and $(1-\varepsilon_C)$ for co-infections. Since these events occur independently of each other, the number of recoveries arising from i_w WT-infected individuals is Binomial(i_w, ε_W) so that

$$\begin{aligned} \alpha_{(0,0,i_w,0)}(r, i_w - r) &= \binom{i_w}{r} \varepsilon_W^r (1-\varepsilon_W)^{i_w - r}, \quad 0 \le r \le i_w, \\ \alpha_{(0,0,i_w,0)}(r, d) &= 0, \qquad\qquad\qquad\qquad\quad r + d \neq i_w. \end{aligned} \tag{5.6}$$

Similarly, we can show that for i_c co-infected individuals the number of recovered individuals follows a Binomial(i_c, ε_C) so that

$$\begin{aligned} \alpha_{(0,0,0,i_c)}(r, i_c - r) &= \binom{i_c}{r} \varepsilon_C^r (1-\varepsilon_C)^{i_c - r}, \quad 0 \le r \le i_c, \\ \alpha_{(0,0,0,i_c)}(r, d) &= 0, \qquad\qquad\qquad\qquad\quad r + d \neq i_c. \end{aligned} \tag{5.7}$$

More generally, for other states $(0, 0, i_w, i_c) \in L(0)$, with $i_w + i_c > 0$, we can use the following recursive equation.

$$\begin{aligned} \alpha_{(0,0,i_w,i_c)}(r,d)\Delta_{(0,0,i_w,i_c)} = {} & \varepsilon_W \rho_W i_w \alpha_{(0,0,i_w-1,i_c)}(r-1,d) \\ & + (1-\varepsilon_W)\rho_W i_w \alpha_{(0,0,i_w-1,i_c)}(r,d-1) \\ & + (1-\phi_W)\beta_C i_w i_c \alpha_{(0,0,i_w-1,i_c+1)}(r,d) \\ & + \varepsilon_C \rho_C i_c \alpha_{(0,0,i_w,i_c-1)}(r-1,d) \\ & + (1-\varepsilon_C)\rho_C i_c \alpha_{(0,0,i_w,i_c-1)}(r,d-1). \end{aligned} \tag{5.8}$$

We can solve these hierarchically. First, consider the case where both populations are of size 1:

$$\begin{aligned} \alpha_{(0,0,1,1)}(r,d)\Delta_{(0,0,1,1)} = {} & \varepsilon_W \rho_W \alpha_{(0,0,0,1)}(r-1,d) + (1-\varepsilon_W)\rho_W \alpha_{(0,0,0,1)}(r,d-1) \\ & + (1-\phi_W)\beta_C \alpha_{(0,0,0,2)}(r,d) \\ & + \varepsilon_C \rho_C \alpha_{(0,0,1,0)}(r-1,d) + (1-\varepsilon_C)\rho_C \alpha_{(0,0,1,0)}(r,d-1), \end{aligned}$$

which we can solve analytically, as we have solutions for both only co-infected and WT infected populations in Equations (5.6) and (5.7). If we then look at the case where we

have 1 WT infected individual and 2 co-infected individuals we get

$$\begin{aligned}\alpha_{(0,0,1,2)}(r,d)\Delta_{(0,0,1,2)} &= \varepsilon_W\rho_W\alpha_{(0,0,0,2)}(r-1,d) + (1-\varepsilon_W)\rho_W\alpha_{(0,0,0,2)}(r,d-1)\\ &\quad + 2(1-\phi_W)\beta_C\alpha_{(0,0,0,3)}(r,d)\\ &\quad + 2\varepsilon_C\rho_C\alpha_{(0,0,1,1)}(r-1,d) + 2(1-\varepsilon_C)\rho_C\alpha_{(0,0,1,1)}(r,d-1).\end{aligned}$$

Again, we now have terms on the right hand side that we already have solutions for. We can repeat this procedure iteratively to find solutions for all cases where $i_w = 1$ and $i_c \in \{1,\dots,N-1\}$.

Now, let's look at the case where $i_w = 2$ and $i_c = 1$:

$$\begin{aligned}\alpha_{(0,0,2,1)}(r,d)\Delta_{(0,0,2,1)} &= 2\varepsilon_W\rho_W\alpha_{(0,0,1,1)}(r-1,d) + 2(1-\varepsilon_W)\rho_W\alpha_{(0,0,1,1)}(r,d-1)\\ &\quad + 2(1-\phi_W)\beta_C\alpha_{(0,0,1,2)}(r,d)\\ &\quad + \varepsilon_C\rho_C\alpha_{(0,0,2,0)}(r-1,d) + (1-\varepsilon_C)\rho_C\alpha_{(0,0,2,0)}(r,d-1).\end{aligned}$$

We have solutions to the first three terms from above and we have solutions to the final two terms from Equation (5.6). We can therefore follow this to get solutions for all cases where $i_w \in \{1,\dots,N-1\}$ and $i_c = 1$.

Now, let us consider an initial population of $i_w = 2, i_c = 2$. From first-step arguments, we obtain the equation

$$\begin{aligned}\alpha_{(0,0,2,2)}(r,d)\Delta_{(0,0,2,2)} &= 2\varepsilon_W\rho_W\alpha_{(0,0,1,2)}(r-1,d) + 2(1-\varepsilon_W)\rho_W\alpha_{(0,0,1,2)}(r,d-1)\\ &\quad + 4(1-\phi_W)\beta_C\alpha_{(0,0,1,3)}(r,d)\\ &\quad + 2\varepsilon_C\rho_C\alpha_{(0,0,2,1)}(r-1,d) + 2(1-\varepsilon_C)\rho_C\alpha_{(0,0,2,1)}(r,d-1),\end{aligned}$$

in which we have solutions to all terms. By direct inspection of these recursive equations, we can use Algorithm 5.1 to solve them.

Algorithm 5.1: An algorithm to calculate $\alpha_{(0,0,i_w,i_c)}(r,d)$, for all pairs r, d such that $r+d=i_w+i_c$.

For $r = 0, 1, \ldots, N$:
  For $d = 0, 1, \ldots, N-r$:
    For $i_w = 0, \ldots, r+d$:
      $i_c = r + d - i_w$;
      If $r + d = 0$:
        $\alpha_{(0,0,i_w,i_c)}(r,d) = 1$;
      Else:
        If $i_w = 0$:
          Calculate $\alpha_{(0,0,0,i_c)}(r,d)$ using Equation (5.7);
        Elif $i_c = 0$:
          Calculate $\alpha_{(0,0,i_w,0)}(r,d)$ using Equation (5.6);
        Else:
          Calculate $\alpha_{(0,0,i_w,i_c)}(r,d)$ using Equation (5.8);

Algorithm 5.1 allows us to obtain all probabilities in level $L(0)$,

$$\alpha_{(0,0,i_w,i_c)}(r,d), (0,0,i_w,i_c) \in L(0), \quad r,d > 0.$$

5.2.1.4 States in level $L(i), i > 0$

We can now work to obtain probabilities, $\alpha_{(s,i_d,i_w,i_c)}(r,d)$, for all other levels, $L(i), i \in \{1,\ldots,N\}$. It is important to note that for all sub-levels $L(i,0)$, the outbreak ends and the probability of initial states in $L(i,0)$ correspond to the boundary conditions for the system of equations shown in Equation (5.4). In particular, for any $0 \leq i \leq N$,

$$\begin{aligned} \alpha_{(s,i_d,0,0)}(0,0) &= 1, \\ \alpha_{(s,i_d,0,0)}(r,d) &= 0, \text{ for } r > 0 \text{ or } d > 0, \end{aligned}$$

for any state $(s,i_d,0,0) \in L(i,0)$. Next, we need to consider the relationship that probabilities in each level $L(i)$ have to those in other levels. We consider a general state $(s,i_d,i_w,i_c) \in L(i,j), s+i_d = i, i_w+i_c = j$. By looking at Equation (5.4) we can depict the relationships in Figure 5.3. These dependencies, along with the absorbing states in $L(i,0)$, imply that once we have the probabilities for $L(0)$, we can recursively obtain the probabilities for the rest of the levels and sub-levels. In particular, we can proceed as in Algorithm 5.2. In order to do $(*)$ from Algorithm 5.2, we proceed to solve

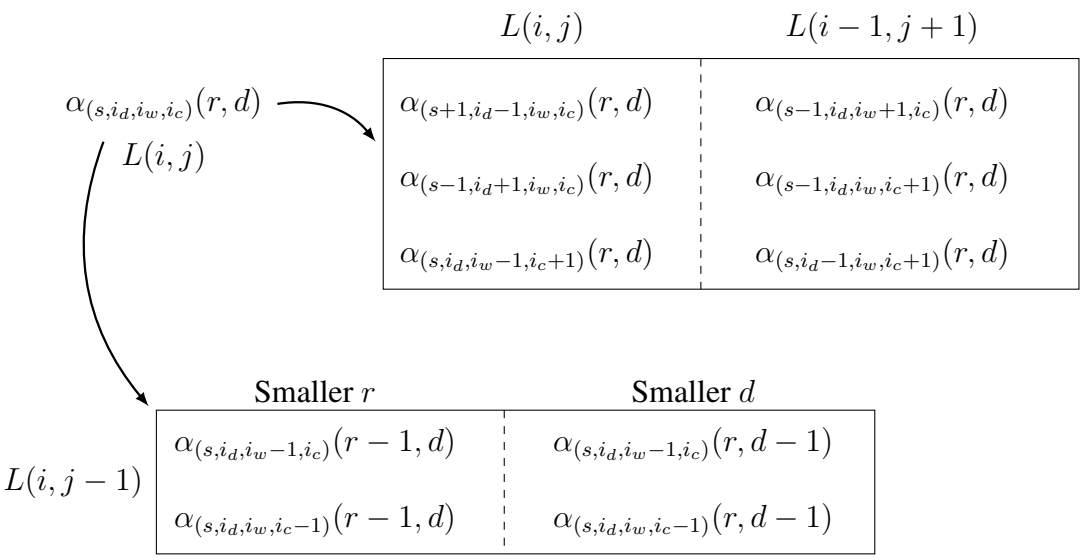

Figure 5.3: The dependencies of $\alpha_{(s,i_d,i_w,i_c)}(r,d)$ from Equation (5.4). The arrow $a \to b$ represents that probability a depends on probability b in the system of Equations (5.4).

Algorithm 5.2: An algorithm to calculate $\alpha_{(s,i_d,i_w,i_c)}(r,d)$.

For $s = 0, i_d = 0$:
    Use Algorithm 5.1 to calculate $\alpha_{(0,0,i_w,i_c)}(r,d)$, for all possible values of i_w, i_c, r and d;
For $i = 1, 2, \ldots, N$:
    For $j = 1, \ldots, N - i$:
        $(*)$ Compute $\alpha_{(s,i_d,i_w,i_c)}(r,d)$ for $(s, i_d, i_w, i_c) \in L(i,j)$ and $j \leq r + d \leq i + j$;

the corresponding system of equations in a matrix-oriented fashion. Matrix equations will be written at the sub-level. In particular, for states $(s, i_d, i_w, i_c) \in L(i, j)$ and values (r, d) such that $j \leq r + d \leq i + j$, we can define vectors

$$\boldsymbol{\alpha}_{i,j}(r,d) = \begin{bmatrix} \alpha_{(0,i,0,j)}(r,d) \\ \alpha_{(0,i,1,j-1)}(r,d) \\ \vdots \\ \alpha_{(0,i,j,0)}(r,d) \\ \alpha_{(1,i-1,0,j)}(r,d) \\ \vdots \\ \alpha_{(i,0,j,0)}(r,d) \end{bmatrix},$$

where $\boldsymbol{\alpha}_{i,j}(r,d)$ is a column vector that contains the probabilities $\alpha_{(s,i_d,i_w,i_c)}(r,d)$ for $s + i_d = i$ and $i_w + i_c = j$. From the dependencies between equations for sub-level $L(i,j)$ in Figure 5.3, we can write

$$\boldsymbol{\alpha}_{i,j}(r,d) = \mathbf{A}_{i,j}\boldsymbol{\alpha}_{i,j}(r,d) + \mathbf{b}_{i,j}(r,d) \tag{5.9}$$

where $\mathbf{A}_{i,j}$ is the matrix of coefficients corresponding to jumps between states in sub-level $L(i,j)$, as discussed previously. Vector $\mathbf{b}_{i,j}(r,d)$ contains independent terms in the equation, corresponding to probabilities which have already been calculated within the algorithm. In particular, the entries of $\mathbf{b}_{i,j}(r,d)$ are calculated by the sum of the terms in Equation (5.4) that depend on sub-levels $L(i-1, j+1)$ and $L(i, j-1)$ (which are previously calculated probabilities within Algorithm 5.2). We can therefore solve Equation (5.9) numerically by using the `linalg` function from the `numpy` package in Python.

In order to practically construct matrix $\mathbf{A}_{i,j}$ and column vector $\mathbf{b}_{i,j}(r,d)$, we need to be able to identify where each state lies within its sub-level. We can calculate the position of a state, (s, i_d, i_w, i_c), in $L(i,j)$, where $i = s + i_d$ and $j = i_w + i_c$, by defining a position function based upon our lexicographic ordering in Figure 5.2. Therefore, position in the list of states within the sub-level is given by

$$\text{pos}(s, i_d, i_w, i_c) = 1 + s(i_w + i_c + 1) + i_w. \tag{5.10}$$

This can be shown for generic state, $(s, i_d, i_w, i_c) \in L(i,j)$, by looking at our ordering of

sub-level $L(i, j)$ with $i = s + i_d$, $j = i_w + i_c$. The first state, which would have position 1, in the sub-level is $(0, i, 0, j)$ and there are $i_w + i_c + 1$ states between this state and the first time the first entry changes, $(1, i - 1, 0, j)$. Therefore, there are $s(i_w + i_j + 1)$ states between the first state and state $(s, i_d, 0, j)$. As the next state is found by adding one to the i_w entry and taking one from the i_c entry, it takes i_w steps to get from state $(s, i_d, 0, j)$ to state (s, i_d, i_w, i_c), leading to the formula in Equation (5.10). Now that we have the position function, a more detailed version of Algorithm 5.2 is given as Algorithm 5.3. We notice that, although $\boldsymbol{\alpha}_{i,j}(r, d)$ contains positive probabilities for $j \leq r + d \leq i + j$, other values of (r, d) just represent boundary conditions with $\boldsymbol{\alpha}_{i,j}(r, d) = \mathbf{0}$. Thus practically when initialising the algorithm, it is useful to set $\boldsymbol{\alpha}_{i,j}(r, d) = \mathbf{0}$ for all $r = 0, \ldots, N, d = 0, \ldots, N - r$.

5.2.1.5 Long-lasting DIPs

We can now consider the interesting situation where the timescales of DIP decay (the transition $I_D \to S$, controlled by parameter δ) are of a greater magnitude than the typical expected duration of the outbreak. Then, one would be able to approximate the dynamics of the system by setting $\delta = 0$. In this situation our diagram of possible transition events becomes the one in Figure 5.4.

We note that, within the S, I_D, I_W, I_C compartments of the model, movement can only occur to the right and down, with no "backward" reactions. This, together with Equation (5.4), suggests the following recursive scheme to solve for the probabilities of interest:

1. Compute probabilities of the form $\alpha_{(0,0,0,i_c)}(r, d)$.
2. Compute probabilities of the form $\alpha_{(0,0,i_w,i_c)}(r, d)$.
3. Compute probabilities of the form $\alpha_{(0,i_d,i_w,i_c)}(r, d)$.
4. Compute probabilities of the form $\alpha_{(s,i_d,i_w,i_c)}(r, d)$.

This can be computed in Python via Algoritm 5.4. As a result of the lack of backwards transitions this algorithm is much more computationally efficient than Algorithm 5.3, as we do not need to numerically solve a system or construct any matrices. Due to this it will be possible to consider larger values of N under the constraint $\delta = 0$.

Algorithm 5.3: An algorithm to calculate $\boldsymbol{\alpha}_{i,j}(r,d)$.

$i=0$;
Implement Algorithm 5.1 to compute vectors $\boldsymbol{\alpha}_{0,j}(r,d)$, where $(r,d)\in\{(j,0),\ldots,(0,j)\}$
For $i=1,\ldots,N$:
  $j=0$;
  For $s=0,\ldots,i$:
    $i_d=i-s$;
    $i_w=0$;
    $i_c=0$;
    $(\boldsymbol{\alpha}_{i,0}(0,0))_{\text{pos}(s,i_d,i_w,i_c)}=1$;
  For $j=1,\ldots,N-i$:
    For $s=0,\ldots,i$:
      $i_d=i-s$;
      For $i_w=0,\ldots,j$:
        $i_c=j-i_w$;
        If $i_d>0$:
          $(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s+1,i_d-1,i_w,i_c)}=\frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}$;
        If $s>0$:
          $(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s-1,i_d+1,i_w,i_c)}=\frac{\phi_D\beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
        If $i_w>0$:
          $(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s,i_d,i_w-1,i_c+1)}=\frac{(1-\phi_W)\beta_C i_w i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
        For all (r,d) such that $j\le r+d\le i+j$:

$$
\begin{aligned}
(\mathbf{b}_{i,j}(r,d))_{\text{pos}(s,i_d,i_w,i_c)} = \big[&(\beta_W i_w+\phi_W\beta_C i_c)s(\boldsymbol{\alpha}_{i-1,j+1}(r,d))_{\text{pos}(s-1,i_d,i_w+1,i_c)}\\
&+\beta_C(1-\phi_W-\phi_D)si_c(\boldsymbol{\alpha}_{i-1,j+1}(r,d))_{\text{pos}(s-1,i_d,i_w,i_c+1)}+\\
&(\beta_C(1-\phi_D)i_c+\beta_w i_w)i_d(\boldsymbol{\alpha}_{i-1,j+1}(r,d))_{\text{pos}(s,i_d-1,i_w,i_c+1)}\\
&+(1-\varepsilon_W)\rho_W i_w(\boldsymbol{\alpha}_{i,j-1}(r,d-1))_{\text{pos}(s,i_d,i_w-1,i_c)}\\
&+\varepsilon_W\rho_W i_w(\boldsymbol{\alpha}_{(i,j-1}(r-1,d))_{\text{pos}(s,i_d,i_w-1,i_c)}\\
&+(1-\varepsilon_C)\rho_C i_c(\boldsymbol{\alpha}_{i,j-1}(r,d-1))_{\text{pos}(s,i_d,i_w,i_c-1)}\\
&+\varepsilon_C\rho_C i_c(\boldsymbol{\alpha}_{i,j-1}(r-1,d))_{\text{pos}(s,i_d,i_w,i_c-1)}\big]\,\frac{1}{\Delta_{(s,i_d,i_w,i_c)}},
\end{aligned}
$$

        where terms in this equation are conditioned that all co-ordinates in (s,i_d,i_w,i_c) are non-negative; else the terms involving that particular state are 0;
  For (r,d) such that $j\le r+d\le i+j$:
    Numerically solve Equation (5.9) to compute $\boldsymbol{\alpha}_{i,j}(r,d)$;

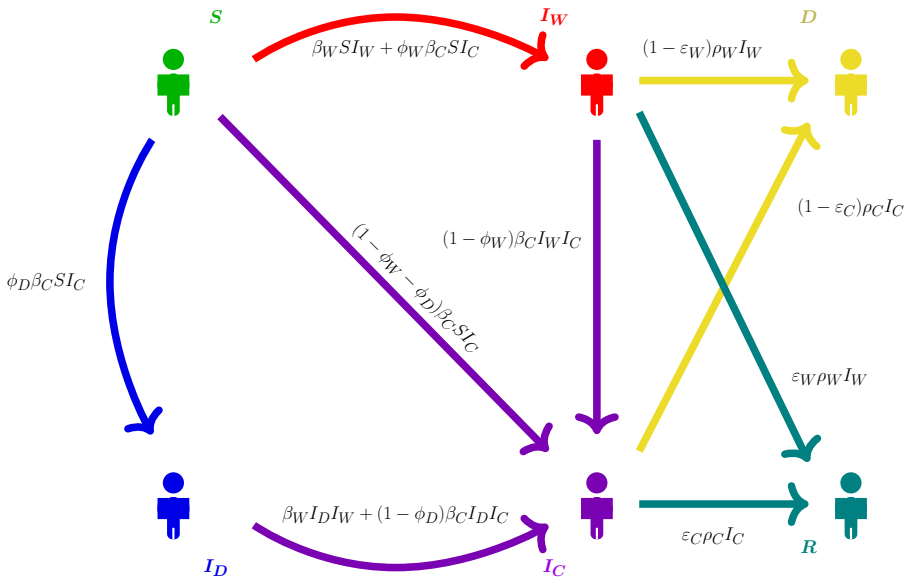

Figure 5.4: The single-step transition diagram for the model under the assumption that $\delta = 0$.

Algorithm 5.4: An algorithm to calculate $\boldsymbol{\alpha}_{i,j}(r,d)$ when $\delta = 0$.

$i = 0$;
Implement Algorithm 1.2 to compute vectors $\boldsymbol{\alpha}_{0,j}(r,d)$, where $(r,d) \in \{(j,0),\ldots,(0,j)\}$;
For $i = 1,\ldots,N$:
    $j = 0$;
    For $s = 0,\ldots,i$:
        $i_d = i - s$;
        $i_w = 0$;
        $i_c = 0$;
        $(\boldsymbol{\alpha}_{i,0}(0,0))_{\text{pos}(s,i_d,i_w,i_c)} = 1$;
    For $j = 1,\ldots,N-i$:
        For $s = 0,\ldots,i$:
            $i_d = i - s$;
            For $i_w = 0,\ldots,j$:
                $i_c = j - i_w$;
                For all (r,d) such that $j \le r + d \le i + j$:
                    Calculate $\alpha_{(s,i_d,i_w,i_c)}(r,d)$ using Equation (5.4);

5.2.1.6 Mean size of the outbreak

We are going to consider the mean values of our removed populations at the end of the outbreak

$$\begin{aligned}\bar{R}_{(s,i_d,i_w,i_c)} = &\quad \mathbb{E}[R(\tau)|S(0)=s, I_D(0)=i_d, I_W(0)=i_w, I_C(0)=i_c],\\ \bar{D}_{(s,i_d,i_w,i_c)} = &\quad \mathbb{E}[D(\tau)|S(0)=s, I_D(0)=i_d, I_W(0)=i_w, I_C(0)=i_c],\end{aligned}$$

for any initial state, $(s, i_d, i_w, i_c) \in \mathcal{S}$, and where τ is defined as in Equation (5.1). We note that these quantities can be computed from the probability mass function given by probabilities $\alpha_{(s,i_d,i_w,i_c)}(r,d)$. However, computing the probabilities is computationally expensive. If one is just interested in computing the mean values, they can be computed directly via first-step arguments.

In this section we will discuss how to compute $\bar{R}_{(s,i_d,i_w,i_c)}$ and $\bar{D}_{(s,i_d,i_w,i_c)}$ for a given initial state, $(s, i_d, i_w, i_c) \in \mathcal{S}$. Let us first focus on $\bar{R}_{(s,i_d,i_w,i_c)}$. One can follow a first-step argument to obtain an equivalent equation to Equation (5.4); in particular:

$$\begin{aligned}\bar{R}_{(s,i_d,i_w,i_c)} = &\frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s+1,i_d-1,i_w,i_c)} + \frac{\phi_D\beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s-1,i_d+1,i_w,i_c)}\\
&+ \frac{(\beta_W i_w + \phi_W\beta_C i_c)s}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s-1,i_d,i_w+1,i_c)} + \frac{\beta_C(1-\phi_W)i_w i_c}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s,i_d,i_w-1,i_c+1)}\\
&+ \frac{\beta_C(1-\phi_W-\phi_D)s i_c}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s-1,i_d,i_w,i_c+1)}\\
&+ \frac{(\beta_C(1-\phi_D)i_c + \beta_W i_w)\, i_d}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s,i_d-1,i_w,i_c+1)}\\
&+ \frac{(1-\varepsilon_W)\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s,i_d,i_w-1,i_c)} + \frac{\varepsilon_W\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}}(\bar{R}_{(s,i_d,i_w-1,i_c)}+1)\\
&+ \frac{(1-\varepsilon_C)\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}}\bar{R}_{(s,i_d,i_w,i_c-1)} + \frac{\varepsilon_C\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}}(\bar{R}_{(s,i_d,i_w,i_c-1)}+1).\end{aligned} \tag{5.11}$$

Note that the terms with $+1$ in them correspond to events where the recovery of an individual occurs.

5.2.1.7 Initial states in level $L(0)$

As we did before, consider first states of the form $(0, 0, i_w, i_c) \in L(0)$. When either i_w or i_c is 0, the mean values are given by a binomial distribution and are easy to calculate:

$$\begin{aligned} R(\tau)|\mathbf{X}(0) = (0, 0, i_w, 0) \sim \text{Binomial}(i_w, \varepsilon_W) \Rightarrow \quad & \bar{R}_{(0,0,i_w,0)} = \varepsilon_W i_w, \\ R(\tau)|\mathbf{X}(0) = (0, 0, 0, i_c) \sim \text{Binomial}(i_c, \varepsilon_C) \Rightarrow \quad & \bar{R}_{(0,0,0,i_c)} = \varepsilon_C i_c. \end{aligned}$$

More generally, for other states in level $L(0)$, $(0, 0, i_w, i_c)$, with $i_w, i_c > 0$, one can use the first-step equation

$$\begin{aligned} \Delta_{(0,0,i_w,i_c)}\bar{R}_{(0,0,i_w,i_c)} =& \beta_C(1-\phi_W)i_w i_c \bar{R}_{(0,0,i_w-1,i_c+1)} + (1-\varepsilon_W)\rho_W i_w \bar{R}_{(0,0,i_w-1,i_c)} \\ &+ \varepsilon_W \rho_W i_w(\bar{R}_{(0,0,i_w-1,i_c)} + 1) + (1-\varepsilon_C)\rho_C i_c \bar{R}_{(0,0,i_w,i_c-1)} \\ &+ \varepsilon_C \rho_C i_c(\bar{R}_{(0,0,i_w,i_c-1)} + 1), \end{aligned} \tag{5.12}$$

which leads to Algorithm 5.5.

Algorithm 5.5: An algorithm to calculate $\bar{R}_{(0,0,i_w,i_c)}$.

```
For i_w = 0, 1, ..., N:
    For i_c = 0, 1, ..., N − i_w:
        If i_w + i_c = 0:
            R̄_(0,0,i_w,i_c) = 0;
        Elif i_c = 0:
            R̄_(0,0,i_w,0) = ε_W i_w;
        Elif i_w = 0:
            R̄_(0,0,0,i_c) = ε_C i_c;
        Else:
            Calculate R̄_(0,0,i_w,i_c) using Equation (5.12);
```

5.2.1.8 States in level $L(i), i > 0$

In order to compute $\bar{R}_{(s,i_d,i_w,i_c)}$ for more general states $(s, i_d, i_w, i_c) \in \mathcal{S}$, for all other levels of $\mathcal{S}$, one can inspect Equation (5.11) to identify the dependencies between different mean values for different initial states, shown in Figure 5.5.

These dependencies, along with the absorbing states in $L(i, 0)$, imply that once we have the mean values for $L(0)$, we can recursively obtain the mean quantities for the rest of the

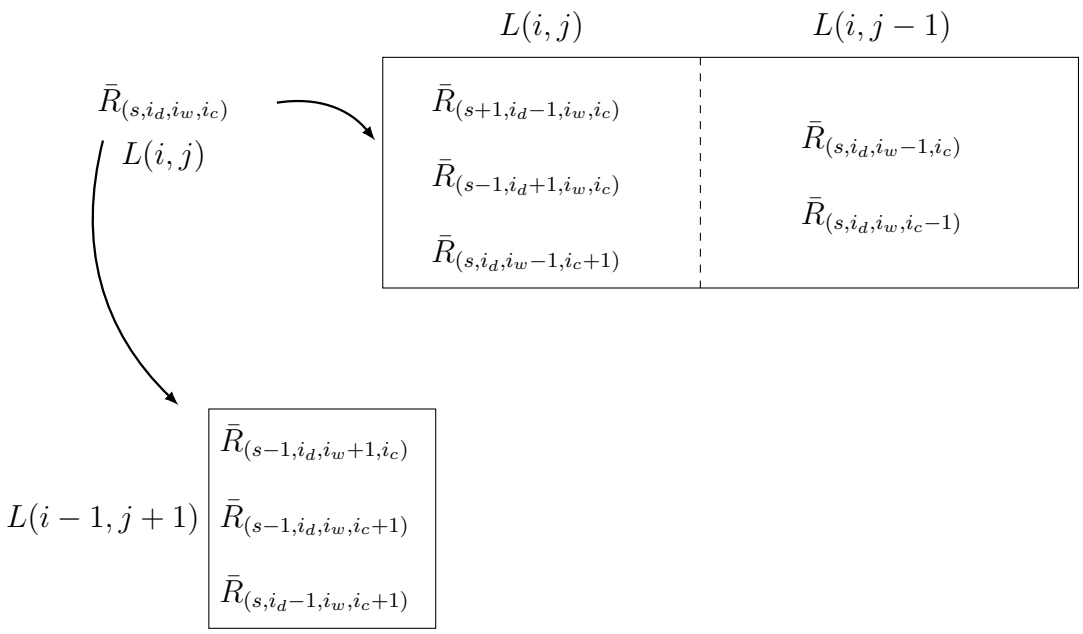

Figure 5.5: The dependencies of $\bar{R}_{(s,i_d,i_w,i_c)}$ from Equation (5.11).

levels and sub-levels. We can do this by defining column vectors for states in sub-levels $L(i,j)$ ordered lexicographically as before,

$$\bar{\mathbf{R}}_{i,j} = \begin{bmatrix} \bar{R}_{(0,i,0,j)} \\ \bar{R}_{(0,i,1,j-1)} \\ \vdots \\ \bar{R}_{(0,i,j,0)} \\ \bar{R}_{(1,i-1,0,j)} \\ \vdots \\ \bar{R}_{(i,0,j,0)} \end{bmatrix},$$

for $i = 1, \ldots, N, j = 0, \ldots, N - i$. Then we can rewrite Equation (5.11) in matrix form as

$$\bar{\mathbf{R}}_{i,j} = \mathbf{A}_{i,j}\bar{\mathbf{R}}_{i,j} + \mathbf{b}_{i,j}, \tag{5.13}$$

where matrix $\mathbf{A}_{i,j}$ is filled by first-step probabilities that remain in the same sub-level, $L(i,j)$, from Equation (5.11), and vector $\mathbf{b}_{i,j}$ contains independent terms in Equation (5.11), which correspond to values of $\bar{R}$ for other initial states, accessible (in one jump) from states in $L(i,j)$ as described in Figure 5.5, which will have already been calculated

within the algorithm. In particular,

$$\mathbf{b}_{i,j} = \begin{bmatrix} \vdots \\ \frac{(\beta_W i_w + \phi_W \beta_C i_c)s}{\Delta_{(s,i_d,i_w,i_c)}} \bar{R}_{(s-1,i_d,i_w+1,i_c)} + \frac{\beta_C(1-\phi_W-\phi_D)si_c}{\Delta_{(s,i_d,i_w,i_c)}} \\ \cdot \bar{R}_{(s-1,i_d,i_w,i_c+1)} + \frac{(\beta_C(1-\phi_D)i_c+\beta_W i_w)i_d}{\Delta_{(s,i_d,i_w,i_c)}} \bar{R}_{(s,i_d-1,i_w,i_c+1)} \\ + \frac{(1-\varepsilon_W)\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}} \bar{R}_{(s,i_d,i_w-1,i_c)} + \frac{\varepsilon_W \rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}} (\bar{R}_{(s,i_d,i_w-1,i_c)} + 1) \\ + \frac{(1-\varepsilon_C)\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}} \bar{R}_{(s,i_d,i_w,i_c-1)} + \frac{\varepsilon_C \rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}} (\bar{R}_{(s,i_d,i_w,i_c-1)} + 1) \\ \vdots \end{bmatrix}$$

for a given state $(s, i_d, i_w, i_c) \in \mathcal{S}$ with $s + i_d = i, i_w + i_c = j$. This is the sum of the terms in Equation (5.11) that depend on levels $L(i, j-1)$ and $L(i-1, j+1)$. This leads to Algorithm 5.6.

If we again consider the case where DIPs last a longer time period than we would expect the outbreak to occur over, we can simplify this algorithm in a similar manner to that which we carried out when calculating the probabilities $\alpha_{(s,i_d,i_w,i_c)}(r, d)$. In this situation, we get the more efficient Algorithm 5.7.

We can now turn our attention to $\bar{D}_{(s,i_d,i_w,i_c)}$, by following a first-step argument we can obtain an equivalent equation to Equation (5.4). In particular,

$$\begin{aligned} \bar{D}_{(s,i_d,i_w,i_c)} =& \frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s+1,i_d-1,i_w,i_c)} + \frac{\phi_D \beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s-1,i_d+1,i_w,i_c)} \\ &+ \frac{(\beta_W i_w + \phi_W \beta_C i_c)s}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s-1,i_d,i_w+1,i_c)} + \frac{\beta_C(1-\phi_W) i_w i_c}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s,i_d,i_w-1,i_c+1)} \\ &+ \frac{\beta_C(1-\phi_W-\phi_D)si_c}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s-1,i_d,i_w,i_c+1)} \\ &+ \frac{(\beta_C(1-\phi_D)i_c + \beta_W i_w)\, i_d}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s,i_d-1,i_w,i_c+1)} \\ &+ \frac{(1-\varepsilon_W)\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}} (\bar{D}_{(s,i_d,i_w-1,i_c)} + 1) + \frac{\varepsilon_W \rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s,i_d,i_w-1,i_c)} \\ &+ \frac{(1-\varepsilon_C)\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}} (\bar{D}_{(s,i_d,i_w,i_c-1)} + 1) + \frac{\varepsilon_C \rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}} \bar{D}_{(s,i_d,i_w,i_c-1)}, \end{aligned} \tag{5.14}$$

where the terms with $+1$ in Equation (5.14) correspond to events where the death of an individual occurs. Notice that this equation is analogous to Equation (5.11), therefore

Algorithm 5.6: An algorithm to calculate mean values $\bar{\mathbf{R}}_{i,j}$ for any $(s, i_d, i_w, i_c) \in \mathcal{S}$.

$i = 0$;
Implement Algorithm 5.5 to compute vectors $\bar{\mathbf{R}}_{0,j}$, where $j = 0, \ldots, N$;
For $i = 1, \ldots, N$:
$j = 0$;
For $s = 0, \ldots, i$:
$i_d = i - s$;
$i_w = 0$;
$i_c = 0$;
$\left(\bar{\mathbf{R}}_{i,0}\right)_{\text{pos}(s,i_d,i_w,i_c)} = 0$;
For $j = 1, \ldots, N - i$:
For $s = 0, \ldots, i$:
$i_d = i - s$;
For $i_w = 0, \ldots, j$:
$i_c = j - i_w$;
If $i_d > 0$:
$(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s+1,i_d-1,i_w,i_c)} = \frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}$;
If $s > 0$:
$(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s-1,i_d+1,i_w,i_c)} = \frac{\phi_D \beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
If $i_w > 0$:
$(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s,i_d,i_w-1,i_c+1)} = \frac{(1-\phi_W)\beta_C i_w i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
Fill vector $\mathbf{b}_{i,j}$ as follows:

$$
\begin{aligned}
(\mathbf{b}_{i,j})_{\text{pos}(s,i_d,i_w,i_c)} = \big[&(\beta_W i_w + \phi_W \beta_C i_c) s (\bar{\mathbf{R}}_{i-1,j+1})_{\text{pos}(s-1,i_d,i_w+1,i_c)} \\
&+ \beta_C (1 - \phi_W - \phi_D) s i_c (\bar{\mathbf{R}}_{i-1,j+1})_{\text{pos}(s-1,i_d,i_w,i_c+1)} \\
&+ (\beta_C (1 - \phi_D) i_c + \beta_w i_w) i_d (\bar{\mathbf{R}}_{i-1,j+1})_{\text{pos}(s,i_d-1,i_w,i_c+1)} \\
&+ (1 - \varepsilon_W) \rho_W i_w (\bar{\mathbf{R}}_{i,j-1})_{\text{pos}(s,i_d,i_w-1,i_c)} \\
&+ \varepsilon_W \rho_W i_w \left((\bar{\mathbf{R}}_{i,j-1})_{\text{pos}(s,i_d,i_w-1,i_c)} + 1 \right) \\
&+ (1 - \varepsilon_C) \rho_C i_c (\bar{\mathbf{R}}_{i,j-1})_{\text{pos}(s,i_d,i_w,i_c-1)} \\
&+ \varepsilon_C \rho_C i_c \left((\bar{\mathbf{R}}_{i,j-1})_{\text{pos}(s,i_d,i_w,i_c-1)} + 1 \right) \big] \frac{1}{\Delta_{(s,i_d,i_w,i_c)}},
\end{aligned}
$$

where terms in this equation are conditioned that all co-ordinates in (s, i_d, i_w, i_c) are non-negative; else the terms involving that particular state are 0;

Calculate $\bar{\mathbf{R}}_{i,j}$ by numerically solving Equation (5.13);

Algorithm 5.7: An algorithm to calculate vectors $\bar{\mathbf{R}}_{i,j}$, when $\delta = 0$.

$i = 0$;
Implement Algorithm 5.5 to compute vectors $\bar{\mathbf{R}}_{0,j}$, where $j = 0, \cdots, N$;
For $i = 1, \ldots, N$:
    $j = 0$;
    For $s = 0, \ldots, i$:
        $i_d = i - s$;
        $i_w = 0$;
        $i_c = 0$;
        $(\bar{\mathbf{R}}_{i,0})_{\text{pos}(s,i_d,i_w,i_c)} = 0$;
    For $j = 1, \ldots, N - i$:
        For $s = 0, \ldots, i$:
            $i_d = i - s$;
            For $i_w = 0, \ldots, j$:
                $i_c = j - i_w$;
                Calculate $\left(\bar{\mathbf{R}}_{i,j}\right)_{\text{pos}(s,i_d,i_w,i_c)}$ using Equation (5.11);

similar methods as used to find for $\bar{R}_{(s,i_d,i_w,i_c)}$ can be used to find $\bar{D}_{(s,i_d,i_w,i_c)}$, and equivalent algorithms to Algorithms 5.6 and 5.7 can be obtained.

5.2.2 Number of co-infected individuals

The efficacy of DIPs as a mitigation strategy can be evaluated from the descriptors previously analysed, for example by looking at how the presence of DIPs impacts on the number of recovered and dead individuals during an outbreak. However, it is clear that DIPs may be effective either because they prevent death very efficiently, or because they spread across the population very efficiently. To better understand how many individuals are affected by DIPs during an outbreak, one can study the random variable, C, where C is the cumulative number of co-infected individuals during the outbreak (that is, during $[0, \tau]$).

Our aim in this section is to analyse the probability mass function of C:

$$\xi_{(s,i_d,i_w,i_c)}(n) = \mathbb{P}\left(C = n \mid \mathbf{X}(0) = (s, i_d, i_w, i_c)\right), (s, i_d, i_w, i_c) \in \mathcal{S},$$

for $n \geq i_c$. It is important to note that if one starts the outbreak with i_c co-infected

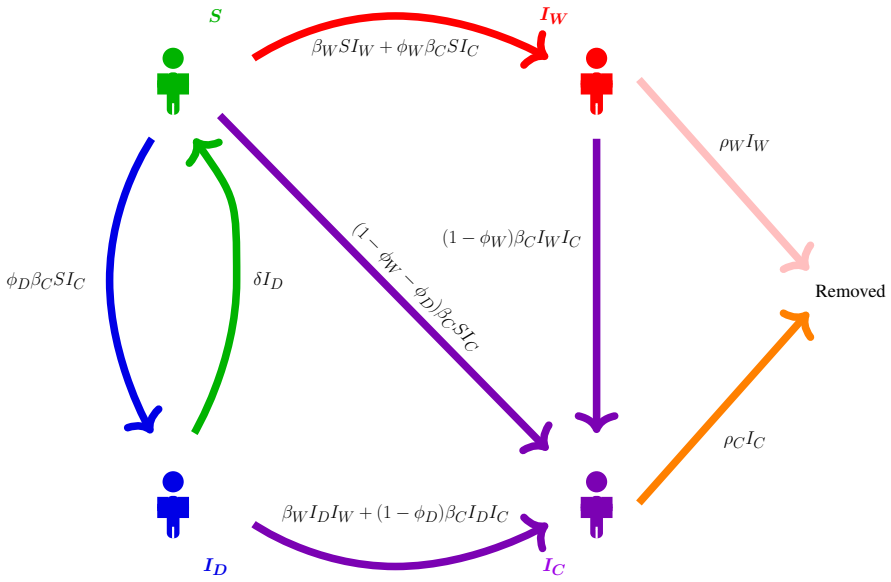

Figure 5.6: The single-step transition diagram for the model where the particular outcome of an infection (death or recovery) is not tracked.

individuals at state $(s, i_d, i_w, i_c) \in \mathcal{S}$, we count these within C, and the support of random variable C is $\{i_c, i_c + 1, \ldots, s + i_d + i_w + i_c\}$.

One can consider here the simplified diagram in Figure 5.6, since our interest is not on the outcome of those infections (i.e., we can consider "removed" individuals, who have either recovered or died). We also note that $\mathbb{P}(C = n)$ corresponds to the probability of observing exactly n events identified by the orange arrow (where an I_C infected individual is removed) or n events identified by the purple arrows (where an individual becomes an I_C infected individual). These are equivalent as all co-infected individuals are removed eventually; we will focus on the orange events to do the analysis of C, as by counting these events we capture all co-infected individuals, including those co-infected at the start of the outbreak, whereas if we counted purple events we would also need to add on the initial number of co-infected individuals to calculate C. The analysis could be done for the purple events though with this addition, and would produce the same results.

Thus, for a particular initial state $(s, i_d, i_w, i_c) \in \mathcal{S}$, the number of co-infected individuals observed during the outbreak, C, corresponds to the number of orange events (in Figure

5.6) that occur during the outbreak. A first-step argument yields the following equation:

$$\begin{aligned}\xi_{(s,i_d,i_w,i_c)}(n) = &\frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s+1,i_d-1,i_w,i_c)}(n) + \frac{\phi_D\beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s-1,i_d+1,i_w,i_c)}(n)\\ &+ \frac{(\beta_W i_w + \phi_W\beta_C i_c)s}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s-1,i_d,i_w+1,i_c)}(n)\\ &+ \frac{\beta_C(1-\phi_W)i_w i_c}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s,i_d,i_w-1,i_c+1)}(n)\\ &+ \frac{\beta_C(1-\phi_W-\phi_D)s i_c}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s-1,i_d,i_w,i_c+1)}(n)\\ &+ \frac{(\beta_C(1-\phi_D)i_c + \beta_W i_w)\, i_d}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s,i_d-1,i_w,i_c+1)}(n)\\ &+ \frac{\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s,i_d,i_w-1,i_c)}(n)\\ &+ \frac{\rho_C i_c}{\Delta_{(s,i_d,i_w,i_c)}}\xi_{(s,i_d,i_w,i_c-1)}(n-1),\end{aligned} \tag{5.15}$$

where the last term in this equation represents a co-infected individual being removed from the system.

We note that, for states in level $L(0)$, one gets

$$\begin{aligned}\Delta_{(0,0,i_w,i_c)}\xi_{(0,0,i_w,i_c)}(n) =&\beta_C(1-\phi_W)i_w i_c\xi_{(0,0,i_w-1,i_c+1)}(n) + \rho_W i_w\xi_{(0,0,i_w-1,i_c)}(n)\\ &+ \rho_C i_c\xi_{(0,0,i_w,i_c-1)}(n-1),\end{aligned} \tag{5.16}$$

which allows one to propose Algorithm 5.8 (which is similar to Algorithm 5.5).

For more general states in other levels $(s, i_d, i_w, i_c) \in L(i), i = 1, \ldots, N$, one needs to think about the dependencies in Equation (5.15). However, we first look at states in a given sub-level

$$L(i,j) = \{(0,i,0,j),(0,i,1,j-1),\ldots,(0,i,j,0),\ldots,(i,0,j,0)\}.$$

We note that, in order to study descriptor C, this organisation of states is not ideal, since when calculating $\xi_{(s,i_d,i_w,i_c)}(n)$, only states with $i_c \leq n$ should be considered, as if $i_c > n$ then we know trivially that $\xi_{(s,i_d,i_w,i_c)}(n) = 0$. It is therefore more convenient to define sub-levels in a different way for this descriptor, as co-infected individuals play a special

Algorithm 5.8: An algorithm to calculate probabilities $\xi_{(0,0,i_w,i_c)}(n)$.

For $n = 0, \dots, N$:
  For $i_c = n, n-1, \dots, 0$:
    For $i_w = 0, \dots, N - i_c$:
      If $i_c = 0$:
        If $n = 0$:
          $\xi_{(0,0,i_w,i_c)}(n) = 1$;
        If $n > 0$:
          $\xi_{(0,0,i_w,i_c)}(n) = 0$;
      Else:
        If $n = 0$:
          $\xi_{(0,0,i_w,i_c)}(n) = 0$;
        If $n > 0$:
          Calculate $\xi_{(0,0,i_w,i_c)}(n)$ using Equation (5.16);

role in C. Thus, we suggest to define $L(i)$ as before (all states in $\mathcal{S}$ where $s + i_d = i$), but consider new sub-levels $\tilde{L}(i,j)$, where

$$L(i) = \bigcup_{j=0}^{N-i} \tilde{L}(i,j), \quad \tilde{L}(i,j) = \{(s, i_d, i_w, i_c) \in \mathcal{S} : s + i_d = i, i_c = j\},$$

so that,

$$\begin{aligned}\tilde{L}(i,j) = \{&(0,i,0,j), (0,i,1,j), (0,i,2,j), \dots, (0,i,N-i-j,j), (1,i-1,0,j),\\ &(1,i-1,1,j), \dots, (1,i-1,N-i-j,j), \dots, (i,0,0,j),\\ &\dots, (i,0,N-i-j,j)\}.\end{aligned}$$

We note that

$$\#\tilde{L}(i,j) = (N-i-j+1)(i+1),$$

and

$$\begin{aligned}\#L(i) &= \sum_{j=0}^{N-i} \#\tilde{L}(i,j) = \sum_{j=0}^{N-i} (i+1)(N-i-j+1) = (i+1)\sum_{j=0}^{N-i}(N-i-j+1)\\ &= (i+1)\sum_{k=0}^{N-i}(k+1) = \sum_{k=0}^{N-i}(i+1)(k+1) = \sum_{k=0}^{N-i}\#L(i,k),\end{aligned}$$

using a transformation $k = N - i - j$. With this, we have shown that the new way of defining sub-levels, $\tilde{L}(i,j)$, is consistent with the previous way of defining sub-levels, $L(i,j)$, and level $L(i)$ still has the same number of states, as expected.

We also need to be able to find the position of a state (s, i_d, i_w, i_c) within its sub-level $\tilde{L}(i,j)$. We can do this by analysing how many states will precede state (s, i_d, i_w, i_c) within its sub-level. State $(s, i_d, i_w, i_c) \in \mathcal{S}$, will be preceded by all states within that sub-level with a lower value of s:

$$\begin{array}{rcl} (0, i, \dots, \dots) & \to & (N-i-j+1) \text{ states}, \\ (1, i-1, \dots, \dots) & \to & (N-i-j+1) \text{ states}, \\ \vdots & & \vdots \\ (s-1, i_d+1, \dots, \dots) & \to & (N-i-j+1) \text{ states}, \end{array}$$

which represents the $s(N-i-j+1)$ states that precede (s, i_d, i_w, i_c) with a lower s value. There will also be states within $\tilde{L}(i,j)$ which have the same s value as (s, i_d, i_w, i_c) but a lower i_w value. These are:

$$\begin{array}{c} (s, i_d, 0, i_c), \\ (s, i_d, 1, i_c), \\ \vdots \\ (s, i_d, i_w - 1, i_c), \end{array}$$

which represent i_w states preceding (s, i_d, i_w, i_c). Therefore the position of state (s, i_d, i_w, i_c) in sub-level $\tilde{L}(i,j)$ is given by the function:

$$\tilde{\text{pos}}(s, i_d, i_w, i_c) = s(N-i-j+1) + i_w + 1. \tag{5.17}$$

Finally, we can note that the last state in $\tilde{L}(i,j)$ is $(i, 0, N-i-j, j)$, which is in position

$$\begin{aligned} \tilde{\text{pos}}(i, 0, N-i-j, j) &= i(N-i-j+1) + N - i - j + 1 \\ &= (i+1)(N-i-j+1) = \#\tilde{L}(i,j), \end{aligned}$$

which is consistent with the size of sub-level $\tilde{L}(i,j)$ computed above.

Thus we propose to organise the probabilities, $\xi_{(s,i_d,i_w,i_c)}(n)$, in column vectors

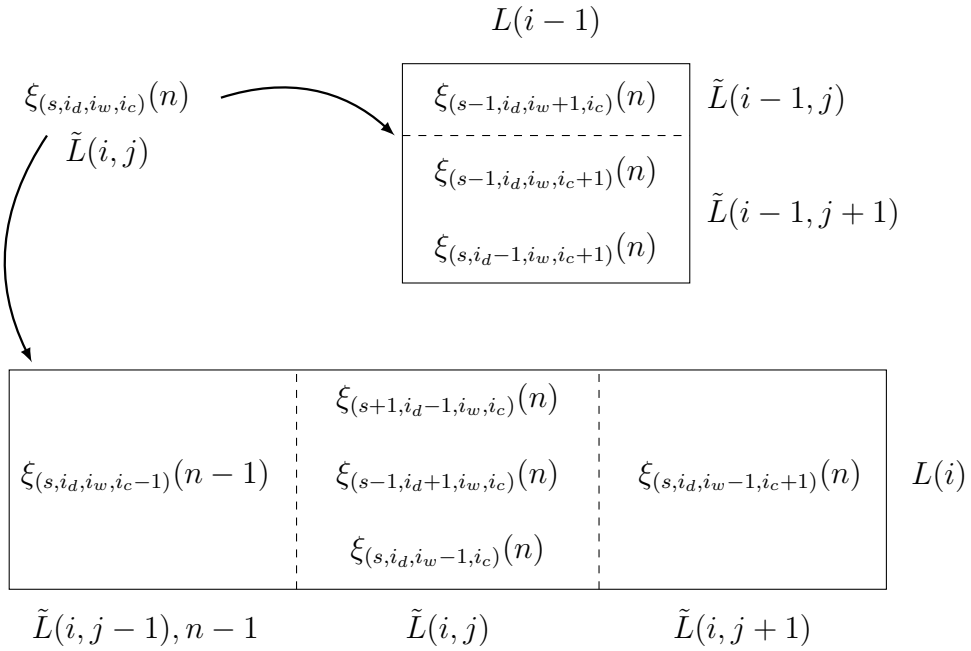

Figure 5.7: The dependencies of $\xi_{(s,i_d,i_w,i_c)}(n)$ from Equation (5.15).

$$\boldsymbol{\xi}_{i,j}(n) = \begin{bmatrix} \xi_{(0,i,0,j)}(n) \\ \xi_{(0,i,1,j)}(n) \\ \vdots \\ \xi_{(0,i,N-i-j,j)}(n) \\ \xi_{(1,i-1,0,j)}(n) \\ \vdots \\ \xi_{(i,0,N-i-j,j)}(n) \end{bmatrix}, \text{ for } 0 \leq i \leq N, \text{ and } 0 \leq j \leq \min(N-i, n).$$

We consider the relationship that each level $\tilde{L}(i,j)$ has to other levels, in Equation (5.15). This is shown for a general state $(s, i_d, i_w, i_c) \in \tilde{L}(i,j), s + i_d = i, i_c = j$, in Figure 5.7. The dependencies shown in this figure suggest an iterative solution is needed.

Practically, we design an algorithm by realising that due to these dependencies, one needs to ensure that smaller values of n, smaller values of i and larger values of j have already been completed in each iteration. By following this recursive approach, one solves Equation (5.15) for states within a given sub-level $\tilde{L}(i,j)$ while only needing to use probabilities for previous values of n, smaller i or larger j which have already been

computed. Equation (5.15) can be written in matrix form as

$$\boldsymbol{\xi}_{i,j}(n) = \tilde{\mathbf{A}}_{i,j}\boldsymbol{\xi}_{i,j}(n) + \tilde{\mathbf{b}}_{i,j}(n). \tag{5.18}$$

That is, matrix $\tilde{\mathbf{A}}_{i,j}$ contains one-step transition probabilities between states in $\tilde{L}(i,j)$. and $\tilde{\mathbf{b}}_{i,j}(n)$ contains independent terms related to previously computed probabilities, $\boldsymbol{\xi}_{i',j'}(n')$ with $(i',j',n') \in \{(i,j+1,n),(i,j-1,n-1),(i-1,j,n),(i-1,j+1,n)\}$. This leads to Algorithm 5.9.

Finally, once again a direct inspection of Equation (5.15) shows that if DIPs last for a longer timescale than the outbreak (using the approximation $\delta \approx 0$), one can solve these equations iteratively in a scalar way. This leads to Algorithm 5.10.

5.2.3 Exact reproduction number

Finally, one can analyse the efficacy of DIPs by focusing on exact reproduction numbers. This concept has been shown before in papers such as Artalejo & Lopez-Herrero (2013) and López-García (2016), which have defined this as the number (measured as a random variable, rather than as an average) of secondary infections caused by a marked individual during their infectious period. In particular, we define three key random variables:

R^W: The number of infections involving the WT-strain caused by an initially marked WT-infected individual during their infectious period (which could include some time as a co-infected individual, if the individual becomes co-infected during their infectious period).

$R^C(W)$: The number of infections involving the WT-strain caused by an initially marked co-infected individual during their infectious period.

$R^C(D)$: The number of infections involving the DIP-strain caused by an initially marked co-infected individual during their infectious period.

These quantities track different transitions within Figure 5.1. In particular, for R^W we track the events $S \to I_W$ and $I_D \to I_C$, and under the circumstance that our initially marked WT-infected individual eventually becomes co-infected, the event $S \to I_C$. For $R^C(W)$, we are interested in the events $S \to I_W, S \to I_C$ and $I_D \to I_C$. Whereas

Algorithm 5.9: An algorithm to calculate $\boldsymbol{\xi}_{i,j}(n)$.

$i = 0$;
Implement Algorithm 5.8 to compute vectors $\boldsymbol{\xi}_{0,j}(n)$, for all $0 \leq j \leq n \leq N$;
For $i = 1, \ldots, N$:
  For $j = 0, \ldots, N - i$;
    For $s = 0, \ldots, i$:
      $i_d = i - s$;
      For $i_w = 0, 1, \ldots, N - i - j$:
        $i_c = j$;
        If $i_d > 0$:
          $(\tilde{\mathbf{A}}_{i,j})_{\tilde{\text{pos}}(s,i_d,i_w,i_c),\tilde{\text{pos}}(s+1,i_d-1,i_w,i_c)} = \frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}$;
        If $s > 0$:
          $(\tilde{\mathbf{A}}_{i,j})_{\tilde{\text{pos}}(s,i_d,i_w,i_c),\tilde{\text{pos}}(s-1,i_d+1,i_w,i_c)} = \frac{\phi_D \beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
        If $i_w > 0$:
          $(\tilde{\mathbf{A}}_{i,j})_{\tilde{\text{pos}}(s,i_d,i_w,i_c),\tilde{\text{pos}}(s,i_d,i_w-1,i_c)} = \frac{\rho_W i_w}{\Delta_{(s,i_d,i_w,i_c)}}$;
For $n = 0, \ldots, N$:
  For $i = 1, \ldots, N$:
    For $j = \min(n, N - i), \min(n, N - i) - 1, \ldots, 1, 0$:
      If $i = N$:
        Note that j would be 0.
        If $n = 0$:
          $\boldsymbol{\xi}_{N,0}(n) = \mathbf{1}_{\#\tilde{L}(N,0)}$;
        Else:
          $\boldsymbol{\xi}_{N,0}(n) = \mathbf{0}_{\#\tilde{L}(N,0)}$;
      Else:
        For $s = 0, \ldots, i$:
          $i_d = i - s$;
          For $i_w = 0, \ldots, N - i - j$:
            $i_c = j$;

$$
\begin{aligned}
\tilde{\mathbf{b}}_{i,j}(n)_{\tilde{\text{pos}}(s,i_d,i_w,i_c)} = \Big[& \beta_C(1-\phi_W) i_w \left(\boldsymbol{\xi}_{i,j+1}(n)\right)_{\tilde{\text{pos}}(s,i_d,i_w-1,i_c+1)} \\
& + s(\beta_W i_w + \phi_W \beta_C i_c) \left(\boldsymbol{\xi}_{i-1,j}(n)\right)_{\tilde{\text{pos}}(s-1,i_d,i_w+1,i_c)} \\
& + \beta_C(1-\phi_W-\phi_D) s i_c \left(\boldsymbol{\xi}_{i-1,j+1}(n)\right)_{\tilde{\text{pos}}(s-1,i_d,i_w,i_c+1)} \\
& + i_d \left(\beta_C(1-\phi_D) i_c + \beta_W i_w\right) \left(\boldsymbol{\xi}_{i-1,j+1}(n)\right)_{\tilde{\text{pos}}(s,i_d-1,i_w,i_c+1)} \\
& + \rho_C i_c \left(\boldsymbol{\xi}_{i,j-1}(n-1)\right)_{\tilde{\text{pos}}(s,i_d,i_w,i_c-1)} \Big] \frac{1}{\Delta_{(s,i_d,i_w,i_c)}};
\end{aligned}
$$

        If $i_w = i_c = i_d = n = 0$:
          $\left(\tilde{\mathbf{b}}_{i,j}(n)\right)_{\tilde{\text{pos}}_{(s,0,0,0)}} = 1$;
        Calculate $\boldsymbol{\xi}_{i,j}(n)$ by numerically solving Equation (5.18);

Algorithm 5.10: An algorithm to calculate $\boldsymbol{\xi}_{i,j}(n)$, when $\delta = 0$.

$i = 0$;
Implement Algorithm 5.8 to compute vectors $\boldsymbol{\xi}_{0,j}(n)$, for all $0 \leq j \leq n \leq N$;
For $n = 0, \ldots, N$:
  For $i = 1, \ldots, N$;
    For $j =$min$(n, N-i)$,min$(n, N-1)-1, \ldots, 1, 0$:
      If $i = N$:
      Note that j would be 0.
        If $n = 0$:
          $\boldsymbol{\xi}_{N,0}(n) = \mathbf{1}_{\#\tilde{L}(N,0)}$;
        Else:
          $\boldsymbol{\xi}_{N,0}(n) = \mathbf{0}_{\#\tilde{L}(N,0)}$;
      Else:
        For $s = 0, \ldots, i$:
          $i_d = i - s$;
          For $i_w = 0, \ldots, N - i - j$:
            $i_c = j$;
            If $i_w + i_c = 0$ and $n = 0$:
              $(\boldsymbol{\xi}_{i,j}(n))_{\tilde{\text{pos}}_{(s,i_d,i_w,i_c)}} = 1$;
            Else:
              Calculate $(\boldsymbol{\xi}_{i,j}(n))_{\tilde{\text{pos}}_{(s,i_d,i_w,i_c)}}$ using Equation (5.15);

for $R^C(D)$, the events $S \to I_D, S \to I_C$ and $I_W \to I_C$ are tracked. It is important to note that although, in principle, we could consider $R^C = R^C(W) + R^C(D)$ as the exact reproduction number for a co-infected individual, this would in fact "double count" co-infection events (as both $R^C(W)$ and $R^C(D)$ track these events). This does not cause a problem, but is something to be aware of when investigating results.

The focus on exact reproduction numbers will be computing probability mass functions for these random variables; namely

$$\begin{aligned}\eta^W_{(s,i_d,i_w,i_c)}(w) =& \mathbb{P}(R^W = w \mid \mathbf{X}(0) = (s, i_d, i_w, i_c)),\\ & (s, i_d, i_w, i_c) \in \mathcal{S},\ 0 \le w \le s + i_d,\ i_w \ge 1.\\ \eta^C_{(s,i_d,i_w,i_c)}(w, d) =& \mathbb{P}(R^C(W) = w, R^C(D) = d \mid \mathbf{X}(0) = (s, i_d, i_w, i_c)),\\ & (s, i_d, i_w, i_c) \in \mathcal{S},\ 0 \le w \le s + i_d,\ d \ge 0, i_c \ge 1.\end{aligned} \tag{5.19}$$

We note that one should consider $d \ge 0$, where d is not bounded from above, due to the decay of the DIP, $I_D \to S$. This transition means that, in theory, one could observe as many DIP infections as possible during the infectious period and therefore d has no upper bound. One can follow a first-step argument to form an equation for $\eta^C_{(s,i_d,i_w,i_c)}(w, d)$, namely Equation (5.20).

It is important to note that each event involving a co-infected individual within this first-step argument is split into two scenarios: that it directly involves our marked individual with probability, $\frac{1}{i_c}$, and that it involves any other co-infected individual with probability $\frac{i_c-1}{i_c}$.

We note that the dependencies between probabilities arising from Equation (5.20) are very similar to those in Equation (5.15), with probabilities for values (w, d) depending on some of the ones for $w-1$ and/or $d-1$. This suggests that one can solve sequentially for $(w = 0, d = 0)$ and then for $n = 1, 2, \ldots$ with $w + d = n$, while exploiting the dependencies in Figure 5.8. We note that Figure 5.8 is structured in terms of the original sub-levels $L(i, j)$ (rather than $\tilde{L}(i, j)$). This is because sub-levels $\tilde{L}(i, j)$ do not lead to a clear recursive structure for the probabilities of interest. In particular, probabilities for states in $\tilde{L}(i, j)$ depend on probabilities for states in $\tilde{L}(i, j-1), \tilde{L}(i, j+1), \tilde{L}(i, j), \tilde{L}(i-1, j)$ and $\tilde{L}(i-1, j+1)$. Figure 5.8 suggests a recursive approach where one solves Equation

$$
\begin{aligned}
\Delta_{(s,i_d,i_w,i_c)}\eta^C_{(s,i_d,i_w,i_c)}(w,d) = {} & \delta i_d \eta^C_{(s+1,i_d-1,i_w,i_c)}(w,d) \\
& + \phi_D \beta_C s i_c \left(\frac{1}{i_c} \eta^C_{(s-1,i_d+1,i_w,i_c)}(w,d-1) \right. \\
& \left. + \frac{i_c - 1}{i_c} \eta^C_{(s-1,i_d+1,i_w,i_c)}(w,d) \right) \\
& + \left(\beta_W i_w + \phi_W \beta_C i_c \frac{i_c - 1}{i_c} \right) s \eta^C_{(s-1,i_d,i_w+1,i_c)}(w,d) \\
& + \phi_W \beta_C s i_c \frac{1}{i_c} \eta^C_{(s-1,i_d,i_w+1,i_c)}(w-1,d) \\
& + \beta_C (1-\phi_W) i_w i_c \frac{i_c - 1}{i_c} \eta^C_{(s,i_d,i_w-1,i_c+1)}(w,d) \\
& + \beta_C (1-\phi_W) i_w i_c \frac{1}{i_c} \eta^C_{(s,i_d,i_w-1,i_c+1)}(w,d-1) \\
& + \beta_C (1-\phi_W-\phi_D) s i_c \frac{i_c - 1}{i_c} \eta^C_{(s-1,i_d,i_w,i_c+1)}(w,d) \\
& + \beta_C (1-\phi_W-\phi_D) s i_c \frac{1}{i_c} \eta^C_{(s-1,i_d,i_w,i_c+1)}(w-1,d-1) \\
& + \left(\beta_W i_w + \beta_C (1-\phi_D) i_c \frac{i_c - 1}{i_c} \right) i_d \eta^C_{(s,i_d-1,i_w,i_c+1)}(w,d) \\
& + \beta_C (1-\phi_D) i_c \frac{1}{i_c} i_d \eta^C_{(s,i_d-1,i_w,i_c+1)}(w-1,d) \\
& + \rho_W i_w \eta^C_{(s,i_d,i_w-1,i_c)}(w,d) \\
& + \rho_C i_c \frac{i_c - 1}{i_c} \eta^C_{(s,i_d,i_w,i_c-1)}(w,d) \\
& + \rho_C i_c \frac{1}{i_c} \delta_{w=0,d=0}.
\end{aligned}
\tag{5.20}
$$

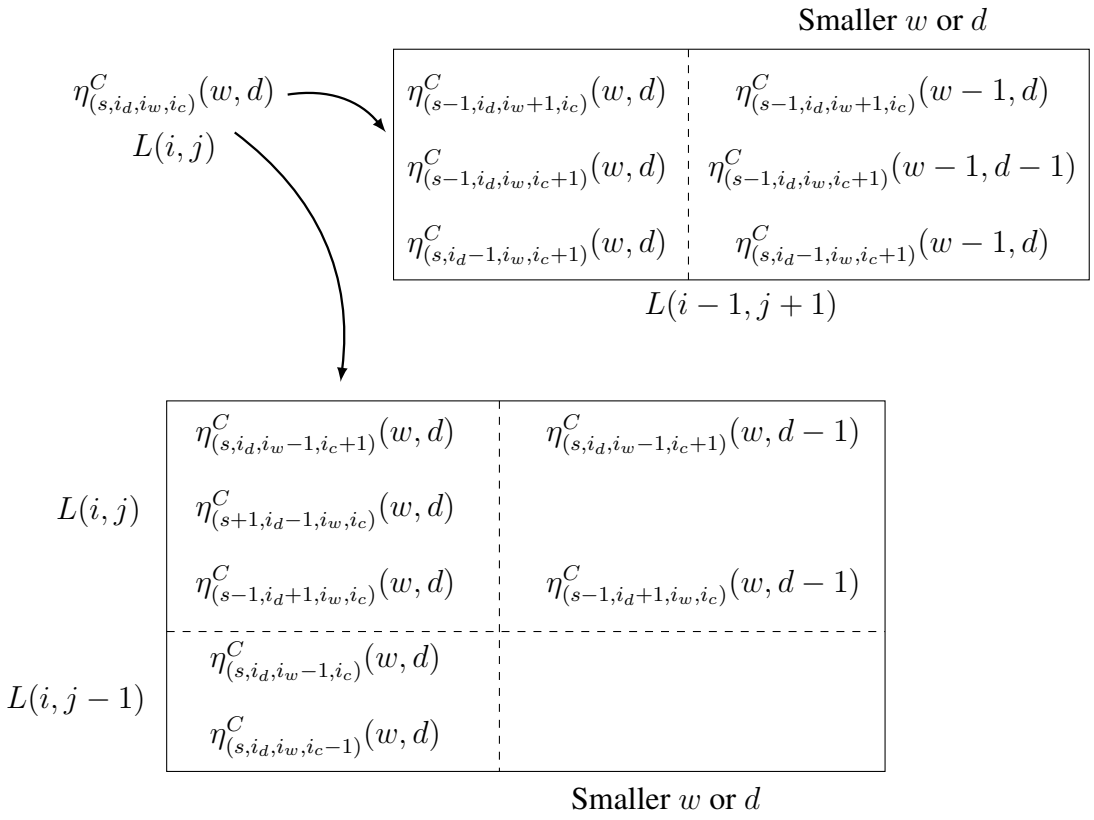

Figure 5.8: The dependencies of $\eta^C_{(s,i_d,i_w,i_c)}(w,d)$ from Equation (5.20).

(5.20) for increasing values of $w + d$, increasing values of i and increasing values of j, leading to Algorithm 5.11 from which we can obtain the joint probability distribution of $\left(R^C(W), R^C(D)\right)$. In order to fully implement Algorithm 5.11, we again define a column vector $\boldsymbol{\eta}^C_{i,j}(w,d)$ which gives the probabilities $\eta^C_{(s,i_d,i_w,i_c)}(w,d)$ for $s + i_d = i$ and $i_w + i_c = j$. From the dependencies shown in Equation (5.20) we can rewrite this equation as

$$\boldsymbol{\eta}^C_{i,j}(w,d) = \mathbf{A}_{i,j}\boldsymbol{\eta}^C_{i,j}(w,d) + \mathbf{b}^C_{i,j}(w,d), \tag{5.21}$$

where $\mathbf{A}_{i,j}$ is defined as in Equation (5.9), corresponding to jumps between states in sub-level $L(i,j)$, and vector $\mathbf{b}^C_{i,j}(w,d)$ contains the independent terms in Equation (5.20), that depend on levels $L(i-1,j+1)$ and $L(i,j-1)$.

Finally, one can analyse the probability distribution of R^W, in terms of the probabilities $\eta^W_{(s,i_d,i_w,i_c)}(w)$ as defined in Equation (5.19). We note that this counts the amount of WT-infections caused by a marked WT-infected individual until their removal, which includes any WT-infections caused by this marked individual even if they become co-infected, $I_W \to I_C \to$ Removal. So, naturally, we expect probabilities $\eta^W_{(s,i_d,i_w,i_c)}(w)$ to depend on probabilities $\eta^C(w,d)$. From now, we denote $\eta^C_{(s,i_d,i_w,i_c)}(w)$ as the marginal probabilities of $R^C(W)$:

$$\eta^C_{(s,i_d,i_w,i_c)}(w) = \sum_d \eta^C_{(s,i_d,i_w,i_c)}(w,d). \tag{5.22}$$

A first-step argument for this leads to Equation (5.23). If an event involves a WT-infected individual we say it directly involves our marked individual with probability $\frac{1}{i_w}$ and is does not directly involve our marked individual with probability $\frac{i_w-1}{i_w}$. The term involving $\eta^C_{(s,i_d,i_w-1,i_c+1)}(w)$ shows how $\eta^W_{(s,i_d,i_w,i_c)}(w)$ probabilities rely on $\eta^C_{(s,i_d,i_w-1,i_c+1)}(w)$ terms. We can, once again, depict the dependencies between the probabilities in Equation (5.23) graphically, shown in Figure 5.9.

$$\begin{aligned}\Delta_{(s,i_d,i_w,i_c)}\eta^W_{(s,i_d,i_w,i_c)}(w) &= \delta i_d \eta^W_{(s+1,i_d-1,i_w,i_c)}(w) + \phi_D\beta_C s i_c \eta^W_{(s-1,i_d+1,i_w,i_c)}(w)\\ &+ (\beta_W(i_w-1)+\phi_W\beta_C i_c)s\eta^W_{(s-1,i_d,i_w+1,i_c)}(w)\\ &+ \beta_W s\eta^W_{(s-1,i_d,i_w+1,i_c)}(w-1)\\ &+ \beta_C(1-\phi_W)i_c\left((i_w-1)\eta^W_{(s,i_d,i_w-1,i_c+1)}(w)\right.\\ &\left.+\eta^C_{(s,i_d,i_w-1,i_c+1)}(w)\right) + \beta_C(1-\phi_W-\phi_D)si_c\eta^W_{(s-1,i_d,i_w,i_c+1)}(w)\\ &+ (\beta_W(i_w-1)+\beta_C(1-\phi_D)i_c)i_d\eta^W_{(s,i_d-1,i_w,i_c+1)}(w)\\ &+ \beta_W i_d\eta^W_{(s,i_d-1,i_w,i_c+1)}(w-1)\\ &+ \rho_C i_c\eta^W_{(s,i_d,i_w,i_c-1)}(w) + \rho_W(i_w-1)\eta^W_{(s,i_d,i_w-1,i_c)}(w)\\ &+ \rho_W\delta_{w=0}.\end{aligned} \tag{5.23}$$

We can represent Equation (5.23) using column vectors and matrices in a similar manner to the one shown in Equation (5.21). Namely, we define a column vector $\boldsymbol{\eta}^W_{i,j}(w)$ which gives the probabilities $\eta^W_{(s,i_d,i_w,i_c)}(w)$ for $s+i_d=i$ and $i_w+i_c=j$. By making use of this column vector, we can rewrite Equation (5.23) as

$$\boldsymbol{\eta}^W_{i,j}(w) = \mathbf{A}_{i,j}\boldsymbol{\eta}^W_{i,j}(w) + \mathbf{b}^W_{i,j}(w), \tag{5.24}$$

where $\mathbf{A}_{i,j}$ is again defined as in Equation (5.9), corresponding to jumps between states in sub-level $L(i,j)$. Vector $\mathbf{b}^W_{i,j}(w)$ contains the independent terms in Equation (5.23), that depend on levels $L(i-1,j+1)$ and $L(i,j-1)$. We can then find solutions for this equation for a given state (s,i_d,i_w,i_c) using Algorithm 5.12.

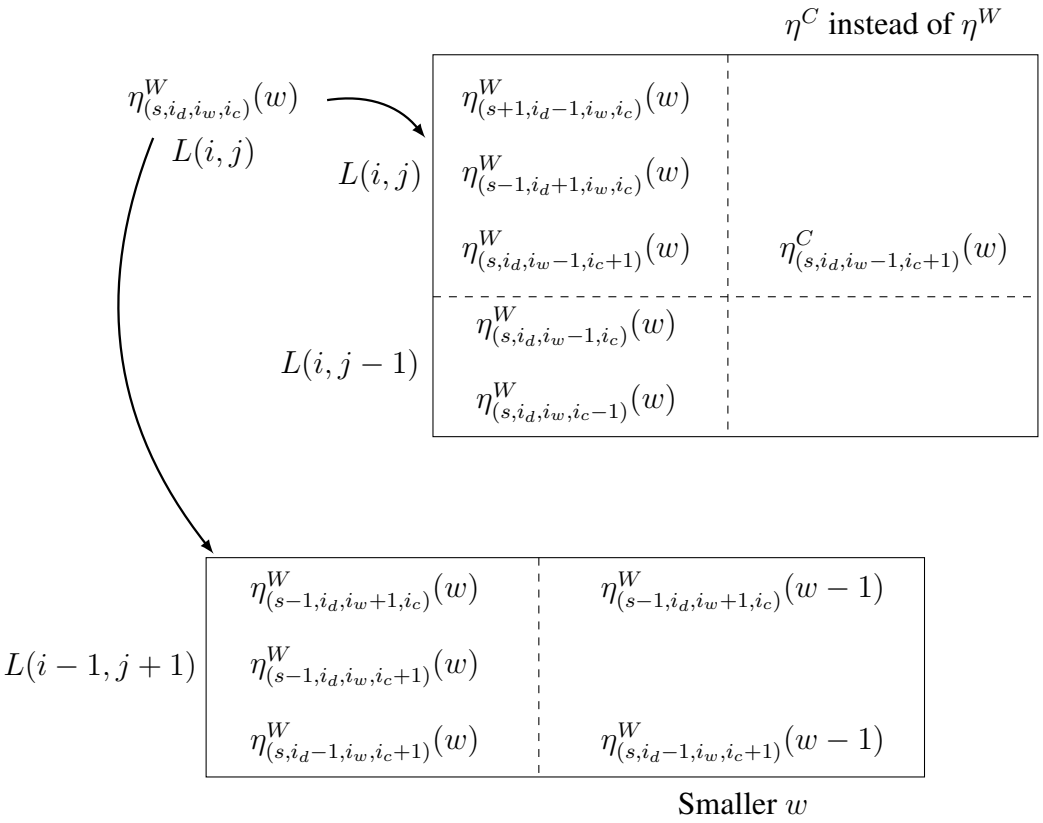

Figure 5.9: The dependencies of $\eta^W_{(s,i_d,i_w,i_c)}(w)$ from Equation (5.23).

Algorithm 5.11: An algorithm to calculate $\boldsymbol{\eta}^{C}_{i,j}(w,d)$, up to a desired value of $w+d=n_{\max}$.

For $i=1,\ldots,N$.
  For $j=1,\ldots,N-i$;
    For $s=0,\ldots,i$:
      $i_d=i-s$;
      For $i_w=0,\ldots,j$:
        $i_c=j-i_w$;
        If $i_c>0$:
          If $i_d>0$:
            $(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s+1,i_d-1,i_w,i_c)}=\frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}$;
          If $s>0$:
            $(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s-1,i_d+1,i_w,i_c)}=\frac{\phi_D\beta_C s(i_c-1)}{\Delta_{(s,i_d,i_w,i_c)}}$;
          If $i_w>0$:
            $(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s,i_d,i_w-1,i_c+1)}=\frac{(1-\phi_W)\beta_C i_w(i_c-1)}{\Delta_{(s,i_d,i_w,i_c)}}$;
For $n=0,\ldots,n_{\max}$:
  For $w=0,\ldots,\min(n,N)$;
    $d=n-w$;
    For $i=w,w+1,\ldots,N$:
      $j=0$;
      If $n=w=d=0$:
        For $s=0,1,\ldots,i$:
          $i_d=i-s$;
          $\left(\boldsymbol{\eta}^{C}_{i,0}(w,d)\right)_{\text{pos}(s,i_d,0,0)}=1$;
      For $j=1,\ldots,N-i$:
        If $i=0$:
          For $i_w=0,\ldots,j$:
            $i_c=j-i_w$;
            If $i_c=0$:
              $\left(\boldsymbol{\eta}^{C}_{0,j}(w,d)\right)_{\text{pos}(0,0,i_w,i_c)}=\delta_{w=0,d=0}$;
            Else:

$$
\begin{aligned}
\left(\boldsymbol{\eta}_{0,j}\right)^{C}(w,d)_{\text{pos}(0,0,i_w,i_c)} = \Big[&\beta_C(1-\phi_W)i_w\Big((i_c-1)\left(\boldsymbol{\eta}^{C}_{0,j}(w,d)\right)_{\text{pos}(0,0,i_w-1,i_c+1)}\\
&+\left(\boldsymbol{\eta}^{C}_{0,j}(w,d-1)\right)_{\text{pos}(0,0,i_w-1,i_c+1)}\Big)\\
&+\rho_W i_w(\boldsymbol{\eta}^{C}_{0,j-1}(w,d))_{\text{pos}(0,0,i_w-1,i_c)}\\
&+\rho_C\Big((i_c-1)\left(\boldsymbol{\eta}^{C}_{0,j-1}(w,d)\right)_{\text{pos}(0,0,i_w,i_c-1}+\delta_{w=0,d=0})\Big]\times\\
&\frac{1}{\Delta_{(0,0,i_w,i_c)}};
\end{aligned}
$$

Else:

For $s = 0, \ldots, i$:

$i_d = i - s$;

For $i_w = 0, \ldots, j$:

$i_c = j - i_w$;

If $i_c = 0$:

$\left(\mathbf{b}^C_{i,j}(w,d)\right)_{\text{pos}(s,i_d,i_w,i_c)} = \delta_{w=0,d=0}$;

Else:

$$
\begin{aligned}
\left(\mathbf{b}^C_{i,j}(w,d)\right)_{\text{pos}(s,i_d,i_w,i_c)} = \Big[& \phi_D \beta_C s \left(\boldsymbol{\eta}^C_{i,j}(w,d-1)\right)_{\text{pos}(s-1,i_d+1,i_w,i_c)} \\
& + (\beta_W i_w + \phi_W \beta_C (i_c - 1)) s \times \\
& \left(\boldsymbol{\eta}^C_{i-1,j+1}(w,d)\right)_{\text{pos}(s-1,i_d,i_w+1,i_c)} \\
& + \phi_W \beta_C s (\boldsymbol{\eta}^C_{i-1,j+1}(w-1,d))_{\text{pos}(s-1,i_d,i_w+1,i_c)} \\
& + \beta_C (1-\phi_W) i_w (\boldsymbol{\eta}^C_{i,j}(w,d-1)_{\text{pos}(s,i_d,i_w-1,i_c+1)} \\
& + \beta_C (1-\phi_D-\phi_W) s \left((i_c-1)\times\right. \\
& \left(\boldsymbol{\eta}^C_{i-1,j+1}(w,d)\right)_{\text{pos}(s-1,i_d,i_w,i_c+1)} \\
& \left. + \left(\boldsymbol{\eta}^C_{i-1,j+1}(w-1,d-1)\right)_{\text{pos}(s-1,i_d,i_w,i_c+1)}\right) \\
& + (\beta_W i_w + \beta_C (1-\phi_D)(i_c-1)) i_d \times \\
& \left(\boldsymbol{\eta}^C_{i-1,j+1}(w,d)\right)_{\text{pos}(s,i_d-1,i_w,i_c+1)} \\
& + \beta_C (1-\phi_D) i_d \left(\boldsymbol{\eta}^C_{i-1,j+1}(w-1,d)\right)_{\text{pos}(s,i_d-1,i_w,i_c+1)} \\
& + \rho_W i_w \left(\boldsymbol{\eta}^C_{i,j-1}(w,d)\right)_{\text{pos}(s,i_d,i_w-1,i_c)} \\
& + \rho_C \delta_{w=0,d=0} \\
& + \rho_C (i_c - 1) \left(\boldsymbol{\eta}^C_{i,j-1}(w,d)\right)_{\text{pos}(s,i_d,i_w,i_c-1)} \Big] \frac{1}{\Delta_{(s,i_d,i_w,i_c)}};
\end{aligned}
$$

Numerically solve Equation (5.21);

Algorithm 5.12: An algorithm to calculate $\boldsymbol{\eta}_{i,j}^{W}(w)$.

For $i = 1, \ldots, N$.
For $j = 1, \ldots, N - i$;
For $s = 0, \ldots, i$:
$i_d = i - s$;
For $i_w = 1, \ldots, j$:
$i_c = j - i_w$;
If $i_d > 0$:
$(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s+1,i_d-1,i_w,i_c)} = \frac{\delta i_d}{\Delta_{(s,i_d,i_w,i_c)}}$;
If $s > 0$:
$(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s-1,i_d+1,i_w,i_c)} = \frac{\phi_D \beta_C s i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
$(\mathbf{A}_{i,j})_{\text{pos}(s,i_d,i_w,i_c),\text{pos}(s,i_d,i_w-1,i_c+1)} = \frac{(1-\phi_W)\beta_C(i_w-1)i_c}{\Delta_{(s,i_d,i_w,i_c)}}$;
For $w = 0, \ldots, N$:
For $i = w, w + 1, \ldots, N$:
$j = 0$;
If $w = 0$:
For $s = 0, 1, \ldots, i$:
$i_d = i - s$;
$\left(\boldsymbol{\eta}_{i,0}^{W}(w)\right)_{\text{pos}(s,i_d,0,0)} = 1$;
For $j = 1, \ldots, N - i$:
If $i = 0$:
For $i_w = 0, \ldots, j$:
$i_c = j - i_w$;
If $i_w = 0$:
$\left(\boldsymbol{\eta}_{0,j}^{W}(w)\right)_{\text{pos}(0,0,i_w,i_c)} = \delta_{w=0}$;
Else:

$$\begin{aligned}\left(\boldsymbol{\eta}_{0,j}^{W}(w)\right)_{\text{pos}(0,0,i_w,i_c)} = \Big[&\beta_C(1-\phi_W)i_c\Big((i_w-1)\left(\boldsymbol{\eta}_{0,j}^{W}(w)\right)_{\text{pos}(0,0,i_w-1,i_c+1)} \\ &+ \left(\boldsymbol{\eta}_{0,j}^{C}(w)\right)_{\text{pos}(s,i_d,i_w-1,i_c+1)}\Big) \\ &+ \rho_C i_c(\boldsymbol{\eta}_{0,j-1}^{W}(w))_{\text{pos}(0,0,i_w,i_c-1)} \\ &+ \rho_W\Big((i_w-1)\left(\boldsymbol{\eta}_{0,j-1}^{W}(w)\right)_{\text{pos}(0,0,i_w-1,i_c)} + \delta_{w=0}\Big)\Big] \times \\ &\frac{1}{\Delta_{(0,0,i_w,i_c)}};\end{aligned}$$

Else:

For $s = 0, \ldots, i$:

$i_d = i - s$;

For $i_w = 0, 1, \ldots, j$:

$i_c = j - i_w$;

If $i_w = 0$:

$\left(\mathbf{b}^W_{i,j}(w)\right)_{\text{pos}(s,i_d,i_w,i_c)} = \delta_{w=0}$;

Else:

$$\begin{aligned}\left(\mathbf{b}^W_{i,j}(w)\right)_{\text{pos}(s,i_d,i_w,i_c)} = \Big[&(\beta_W(i_w-1)+\phi_W\beta_C i_c)\, s\left(\boldsymbol{\eta}^W_{i-1,j+1}(w)\right)_{\text{pos}(s-1,i_d,i_w+1,i_c)}\\ &+\beta_W s\left(\boldsymbol{\eta}^W_{i-1,j+1}(w-1)\right)_{\text{pos}(s-1,i_d,i_w+1,i_c)}\\ &+\beta_C(1-\phi_W)i_c\left(\boldsymbol{\eta}^C_{i,j}(w)\right)_{\text{pos}(s,i_d,i_w-1,i_c+1)}\\ &+\beta_C(1-\phi_W-\phi_D)si_c\left(\boldsymbol{\eta}^W_{i-1,j+1}(w)\right)_{\text{pos}(s-1,i_d,i_w,i_c+1)}\\ &+(\beta_W(i_w-1)+\beta_C(1-\phi_D)i_c)\, i_d\times\\ &\left(\boldsymbol{\eta}^W_{i-1,j+1}(w)\right)_{\text{pos}(s,i_d-1,i_w,i_c+1)}\\ &+\beta_W i_d\left(\boldsymbol{\eta}^w_{i-1,j+1}(w-1)\right)_{\text{pos}(s,i_d-1,i_w,i_c+1)}\\ &+\rho_C i_c\left(\boldsymbol{\eta}^W_{i,j-1}(w)\right)_{\text{pos}(s,i_d,i_w,i_c-1)}\\ &+\rho_W(i_w-1)\left(\boldsymbol{\eta}^W_{i,j-1}(w)\right)_{\text{pos}(s,i_d,i_w-1,i_c)}+\rho_W\delta_{w=0}\Big]\times\\ &\frac{1}{\Delta_{(s,i_d,i_w,i_c)}};\end{aligned}$$

Numerically solve Equation (5.24);

5.3 Numerical results

In this section we are going to look at the stochastic descriptors discussed in Sections 5.2.1-5.2.3. For each descriptor we will investigate the effect of varying different parameters that control the behaviour of the DIP on the corresponding summary statistics. In addition, here we will use baseline values of each parameter within the model that will stay constant unless they are the parameter being varied in each particular figure. The baseline parameters (and the initial condition of the populations) are shown in Table 5.1.

Parameter	Interpretation	Baseline value	Unit
N	Population Size	40	#
δ	DIP decay rate	$\frac{1}{4}$	weeks^{-1}
β_W	WT-infection rate	$\frac{3}{N}$	weeks^{-1}
β_C	Co-infection rate	$\frac{\beta_W}{3}$	weeks^{-1}
ϕ_W	WT-infection probability from a $I_C - S$ interaction	0.25	Unitless
ϕ_D	DIP-infection probability from a $I_C - S$ interaction	0.25	Unitless
ε_W	Recovery probability of WT-infected individual	0.9	Unitless
$1 - \varepsilon_W$	Death probability of WT-infected individual	0.1	Unitless
$1 - \varepsilon_C$	Death probability of co-infected individual	$\frac{1-\varepsilon_W}{3}$	Unitless
ε_C	Recovery probability of co-infected individual	$1 - \frac{1-\varepsilon_W}{3}$	Unitless
ρ_W	Removal rate of a WT-infected individual	1	weeks^{-1}
ρ_C	Removal rate of a co-infected individual	1	weeks^{-1}
$I_D(0)$	The initial number of individuals in the DIP-infected population	$0.5N$	#
$S(0)$	The initial number of susceptible individuals in the total population	$N - I_D(0) - 1$	#

Table 5.1: The parameters in the model, their interpretation and the baseline values used in the results unless explicitly stated.

It has been shown that the recovery period for a person infected with a WT-strain of flu is approximately 6 days (Vaidya *et al.* (2015),De Serres *et al.* (2010)). This informed the choice of $\frac{1}{\rho_W}$ = 1 week in this model, which was chosen to give an average recovery period of a week. We consider $\rho_C = \rho_W$ since we make the assumption that the DIP affects the infectivity of a co-infected individual, or their probability of recovery or death rather than their removal timescales. We chose $\frac{1}{\delta} = 4$ weeks from private communication with our experimental collaborators at Dstl, to represent the fact that the lifetime of defective interfering particles in a susceptible individual would likely be in the order of days or weeks, rather than months or years. We chose $\phi_W = \phi_D$ as a default choice as it is not clear whether the DIP would be more or less likely to be passed on by a co-infected individual than the WT virus. The value of 0.25 for these parameters was chosen such

$\bar{R}, \bar{D}$	$I_D(0) = 0.25N$	$I_D(0) = 0.5N$	$I_D(0) = 0.75N$
$I_W(0) = 1$	18.8, 1.75	**15.2, 1.20**	11.8, 0.8
$I_C(0) = 1$	6.29, 0.490	**5.21, 0.346**	4.31, 0.247

Table 5.2: The mean number of recovered ($\bar{R}$), and dead ($\bar{D}$) individuals for the parametric choices and initial conditions as described in Figure 5.10.

that there is a 50% chance of a co-infected individual passing on both the WT and the DIP strains.

The value $\beta_W = \frac{3}{N}$ was chosen such that without DIP protection, the basic reproduction number would be 3. This was chosen to be sufficiently high so that under these conditions an outbreak would be large enough such that varying DIP-related parameters would show a tangible effect on the behaviour of the descriptors. The choice of $(1 - \varepsilon_W) = 0.1$ was chosen to be representative of a pathogen with a high mortality rate (for comparison, it was found by Donaldson *et al.* (2009) that the A/H1N1 strain of flu had a $(1 - \varepsilon_W) = 0.0098$ in the age range of 65+ during an outbreak in 2009). This allows us to investigate the effects of varying ε_C on the outbreak outcomes. β_C was chosen to give an approximate basic reproduction number of 1 for the co-infected individuals and the baseline value of ε_C was chosen such that the ratio of $\beta_W : \beta_C$ matched the ratio $(1 - \varepsilon_W) : (1 - \varepsilon_C)$; this assumes a symmetric impact on the infectivity rate and the recovery probability by the DIP in a co-infected individual.

5.3.1 Size of the outbreak

In this section we will compute the probability distribution of $(R(\tau), D(\tau))$, in terms of probabilities $\alpha_{(s,i_d,i_w,i_c)}(r, d)$ for a wide range of parameter values and initial conditions. This joint probability distribution will be plotted as lower-triangular heat-maps, as we have the condition that $r + d \leq N$.

Figure 5.10 and Table 5.2, show the probabilities, $\alpha_{(s,i_d,i_w,i_c)}(r, d)$, and the mean values, $\bar{R}$ and $\bar{D}$, of $R(\tau)$ and $D(\tau)$ for different initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$. The entries which represent the baseline parameter values in Table 5.1 are shown in bold in Table 5.2; we will use this representation throughout this section. This allows us to track the impact that the

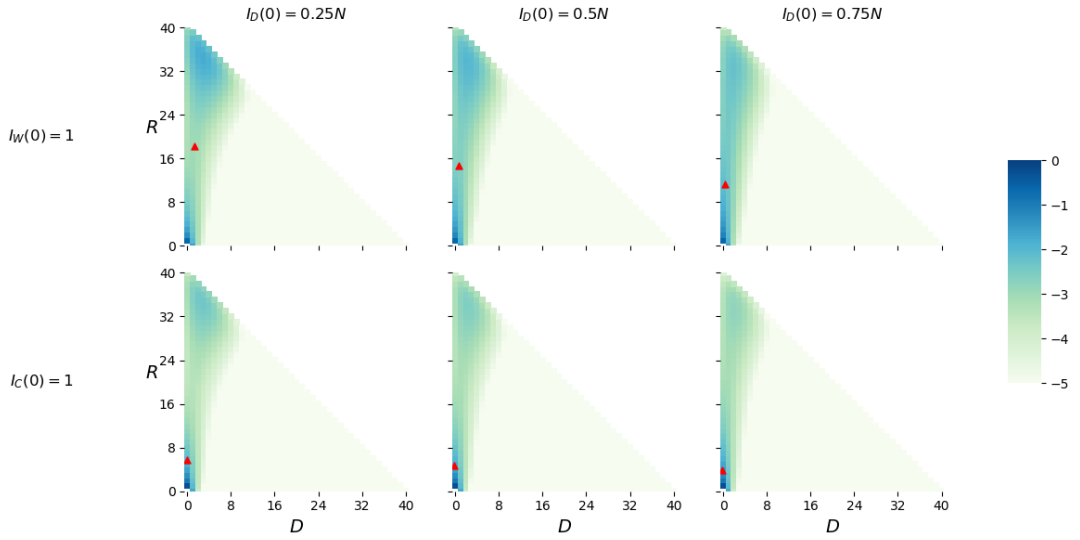

Figure 5.10: Heatmaps showing $\log_{10}(\alpha_{(s,i_d,i_w,i_c)}(r,d))$ for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (top row) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (bottom row). Red triangles represent the mean values of the random variables $(\bar{R}, \bar{D})$.

$\bar{R}\|r+d>1, \bar{D}\|r+d>1$	$I_D(0)=0.25N$	$I_D(0)=0.5N$	$I_D(0)=0.75N$
$I_W(0)=1$	$25.0, 2.31$	**20.1, 1.58**	$15.5, 1.03$
$I_C(0)=1$	$13.6, 1.11$	**11.0, 0.774**	$8.89, 0.539$

Table 5.3: The conditional mean number of recovered ($\bar{R}|r+d>1$), and dead ($\bar{D}|r+d>1$) individuals for the parametric choices and initial conditions as described in Figure 5.11.

initial number of DIP-infected individuals, as well as whether the first infection is a WT-infection or co-infection, has on the size of the outbreak. We can clearly see, and will see repeated for all parameter choices, that the type of first infection has a significant impact on the size of the outbreak. If the individual starting the outbreak is WT-infected the outbreak is likely to be substantially larger than if the individual that starts the outbreak is co-infected, which shows the impact that the DIP has as a mitigation at the population level. From Table 5.2 we can see that in all cases the mean number of infections drops from double to single figures between $I_W(0)=1$ and $I_C(0)=1$.

Now, looking at the effect of increasing the initial number of DIP-infected individuals, for both initial infection types, we see that increasing the number of DIP-infected individuals in the initial condition causes the mean size of the outbreak to decrease. This effect is more pronounced on the predicted number of dead individuals which show a more than 50% decrease from $I_D(0)=0.25N$ to $I_D(0)=0.75N$, whereas the number of recovered individuals also decreases but not by 50%. It is likely that this effect is caused by the higher number of DIP-infected individuals leading to a larger number of infections being co-infected individuals rather than WT-infected, which offers a higher recovery probability.

It is important to note that Table 5.2 implies that in order to protect the population, it is more effective to have the initial individual protected with the DIP, rather than protecting more individuals at the population level ($I_D(0)$). However, in practice, the probability of a co-infected individual starting the outbreak is likely dependent on how many individuals were originally protected by the DIP in the population, $I_D(0)$, when the outbreak started.

If the initially infected individual is removed before infecting any other individual, there is really no outbreak. Thus, it is of interest to analyse the size of the outbreak conditioned on at least one infection event occurring. In Figure 5.11 and Table 5.3, we show the

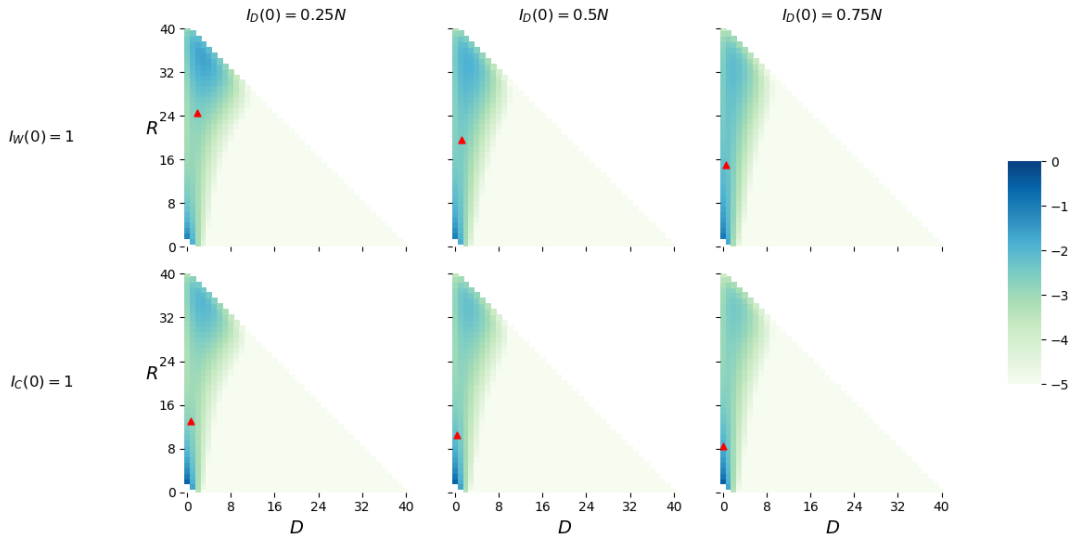

Figure 5.11: Heatmaps showing $\log_{10}(\alpha_{(s,i_d,i_w,i_c)}(r,d)|r+d>1)$ for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, and the individual starting the outbreak either being WT-infected $(I_W(0)=1)$ or co-infected $(I_C(0)=1)$, with the condition that at least one infection event occurs. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (top row) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (bottom row). Red triangles represent the conditional mean values $(\bar{R}|r+d>1, \bar{D}|r+d>1)$.

$\bar{R}, \bar{D}$	$\frac{1}{\delta} = 2$ weeks	$\frac{1}{\delta} = 4$ weeks	$\frac{1}{\delta} = 8$ weeks
$I_W(0) = 1$	16.7, 1.47	**15.2, 1.20**	14.1, 1.02
$I_C(0) = 1$	5.81, 0.441	**5.21, 0.346**	4.75, 0.279

Table 5.4: The mean number of recovered ($\bar{R}$), and dead ($\bar{D}$) individuals for the parametric choices and initial conditions as described in Figure 5.12.

probabilities, $\alpha_{(s,i_d,i_w,i_c)}(r,d)$, and the mean values, $\bar{R}$ and $\bar{D}$, of $R(\tau)$ and $D(\tau)$ for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$, conditioned on the first event within the system not being the initial infected individual recovering or dying. We will denote these conditional probabilities by $\alpha_{(s,i_d,i_w,i_c)}(r,d)|r+d>1$, and conditional mean values by $\bar{R} \mid r+d>1$, $\bar{D} \mid r+d>1$.

As expected, varying $I_D(0)$ and the type of initial infected individual shows similar trends as those seen in Figure 5.10 and Table 5.2. However, the effect on the conditional means is more pronounced. Now, for the case with the largest population protection $(I_D(0) = 0.75N)$, for $I_W(0) = 1$, the conditional average size of the outbreak is, $(\bar{R} \mid r+d>1) + (\bar{D} \mid r+d>1) = 16.53$, whereas without the guarantee of an infection event occurring the average size of the outbreak is $\bar{R} + \bar{D} = 12.6$. Again, a more pronounced effect on the conditional mean number of dead individuals is shown by increasing the initial value of $I_D(0)$.

Figure 5.12 and Table 5.4, show the probabilities, $\alpha_{(s,i_d,i_w,i_c)}(r,d)$, and the mean values, $\bar{R}$ and $\bar{D}$, of $R(\tau)$ and $D(\tau)$ for different rates at which the DIP protection in DIP-infected individuals decays, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$.

The rate of DIP decay has an effect on the mean size of the outbreak, although not as large an effect on the size of the outbreak as changing the initial value of $I_D(0)$, meaning that ensuring as many individuals are protected as possible is likely to be more effective than increasing the length of protection for a smaller proportion. Increasing the length of protection that the DIP gives causes the total number of infections to decrease and a higher proportion of these infections to result in recovery. This will be due to the fact that if the DIP decays slower, a larger proportion of infection events will lead to co-infected

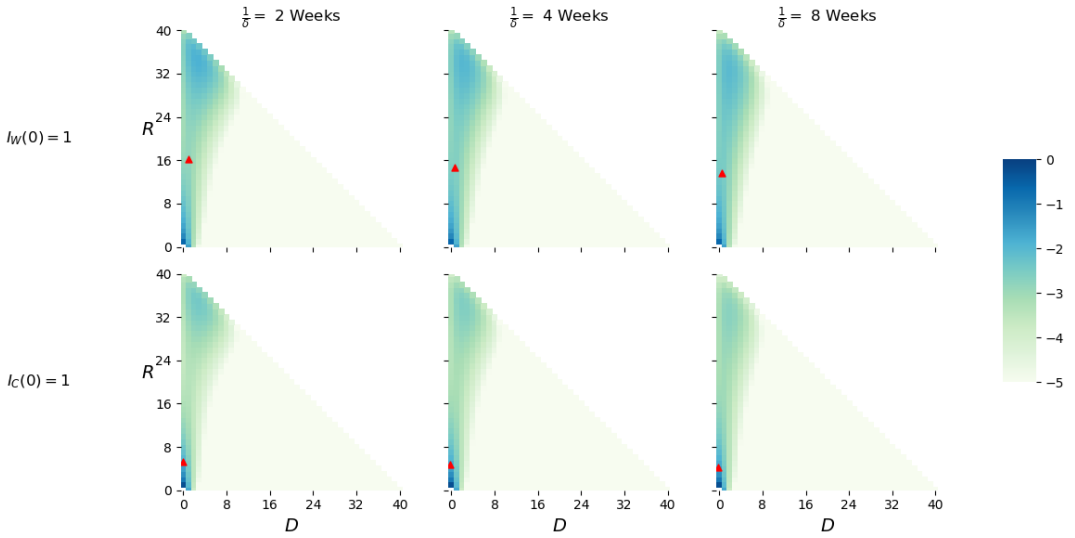

Figure 5.12: Heatmaps showing $\log_{10}(\alpha_{(s,i_d,i_w,i_c)}(r,d))$ for different rates at which the DIP protection in DIP-infected individuals decays, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (top row) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (bottom row). Red triangles represent the mean values of the random variables $(\bar{R}, \bar{D})$.

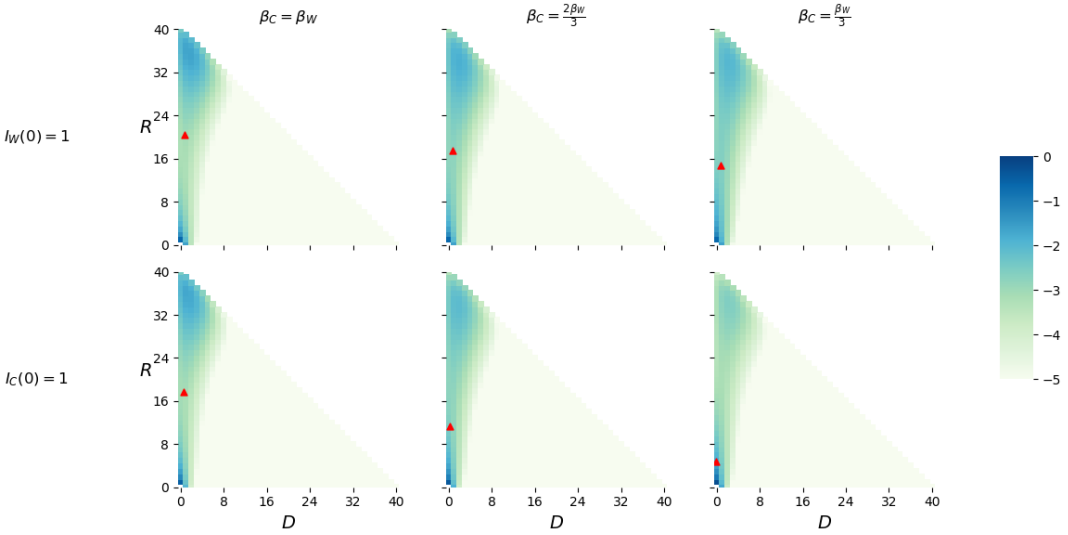

Figure 5.13: Heatmaps showing $\log_{10}(\alpha_{(s,i_d,i_w,i_c)}(r,d))$ for different co-infection rates, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (top row) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (bottom row). Red triangles represent the mean values of the random variables $(\bar{R}, \bar{D})$.

$\bar{R}, \bar{D}$	$\beta_C = \beta_W$	$\beta_C = \frac{2\beta_W}{3}$	$\beta_C = \frac{\beta_W}{3}$
$I_W(0) = 1$	20.9, 1.25	18.07, 1.25	**15.2, 1.20**
$I_C(0) = 1$	18.3, 0.954	11.8, 0.709	**5.21, 0.346**

Table 5.5: The mean number of recovered $(\bar{R})$, and dead $(\bar{D})$ individuals for the parametric choices and initial conditions as described in Figure 5.13.

individuals rather than WT-infected individuals which will give a higher probability of recovery.

Figure 5.13 and Table 5.5, show the probabilities, $\alpha_{(s,i_d,i_w,i_c)}(r,d)$, and the mean values, $\bar{R}$ and $\bar{D}$, of $R(\tau)$ and $D(\tau)$ for different co-infection rates, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$. Varying the co-infection rate in this way reflects the level of infectivity reduction the DIP gives in a co-infected individual; for example, when $\beta_C = \beta_W$, this corresponds to the DIP giving no reduction in the infectiousness of an individual that is co-infected. Conversely, when $\beta_C = \frac{\beta_W}{3}$ this is the case that the DIP makes a co-infected

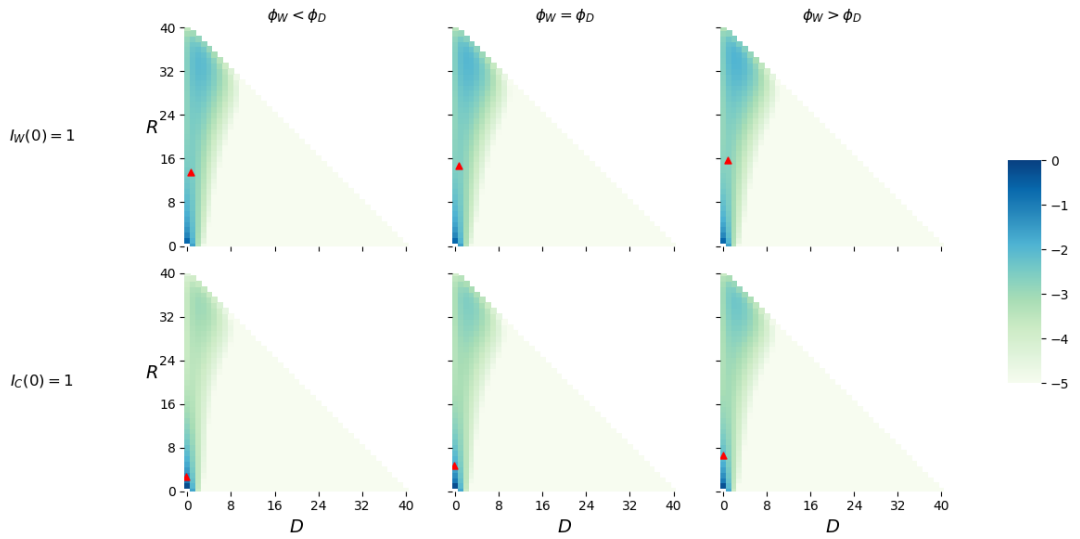

Figure 5.14: Heatmaps showing $\log_{10}(\alpha_{(s,i_d,i_w,i_c)}(r,d))$ for different probabilities of the outcome of a $I_C - S$ infection interaction, $\phi_W < \phi_D, \phi_W = \phi_D$ or $\phi_W > \phi_D$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ or co-infected $(I_C(0) = 1)$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (top row) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (bottom row). Red triangles represent the mean values of the random variables $(\bar{R}, \bar{D})$.

individual 3 times less infectious than a WT-infected individual.

This reduction in infectivity (when $\beta_C < \beta_W$) clearly has a large impact on the effectiveness of the protection that DIP provides to a population. Reducing the infectivity of a co-infected individual directly correlates to a reduction in the number of individuals affected by an outbreak. It is interesting to note, however, that decreasing the infectivity of a co-infected individual actually causes the proportion of individuals affected by the outbreak that die to increase (although the expected absolute number of deaths is still lower). This is something that any future endeavours to create Therapeutic Interfering Particles (TIPs) will need to be mindful of, that a balance between protection and infectivity is needed as if not enough people become infected with the protective DIP, it will be less effective at the population level.

Figure 5.14 and Table 5.6 show the effect of varying the probability that a co-infected individual passes on either the DIP-strain or the WT-strain. For this figure and throughout

$\bar{R}, \bar{D}$	$\phi_W < \phi_D$	$\phi_W = \phi_D$	$\phi_W > \phi_D$
$I_W(0) = 1$	14.0, 1.09	**15.2, 1.20**	16.3, 1.31
$I_C(0) = 1$	3.17, 0.175	**5.21, 0.346**	7.16, 0.513

Table 5.6: The mean number of recovered ($\bar{R}$), and dead ($\bar{D}$) individuals for the parametric choices and initial conditions as described in Figure 5.14.

	$\phi_W < \phi_D$	$\phi_W = \phi_D$	$\phi_W > \phi_D$
ϕ_W, ϕ_D	0.1, 0.4	0.25, 0.25	0.4, 0.1

Table 5.7: The precise values used when investigating the parameter regimes $\phi_W < \phi_D, \phi_W = \phi_D, \phi_W > \phi_D$.

this section when investigating these parameter regimes we use the parameter values for ϕ_W, ϕ_D shown in Table 5.7, which have been chosen so that $(1 - \phi_W - \phi_D) = 0.5$ in all scenarios. We can see that the observed effect is consistent across both initial infection types; increasing the probability that the WT-strain is passed (ϕ_W) rather than the DIP-strain (ϕ_D) causes the size of the outbreak to increase. This is expected as a WT-infected individual is more infectious than a DIP-infected individual under the baseline parameter values shown in Table 5.1. Increasing the probability that the WT-strain is passed to an individual also means that the proportion of affected individuals within an outbreak that die increases, as WT-infected individuals have a lower chance of recovery than co-infected individuals.

5.3.2 Number of co-infected individuals

In this section the impact of varying pairs of parameters within the model on the stochastic descriptor C, the cumulative number of co-infected individuals during the outbreak, described in Section 5.2.2, will be investigated.

Figure 5.15 and Table 5.8 show the effect of varying the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$ and the rate at which the DIP's protective

$\bar{C}\|I_W(0) = 1, \bar{C}\|I_C(0) = 1$	$I_D(0) = 0.25N$	$I_D(0) = 0.5N$	$I_D(0) = 0.75N$
$\frac{1}{\delta} = 2$ weeks	3.43, 2.47	5.21, 2.76	5.86, 2.86
$\frac{1}{\delta} = 4$ weeks	4.64, 2.82	**6.62, 3.14**	6.87, 3.14
$\frac{1}{\delta} = 8$ weeks	5.47, 3.08	7.40, 3.35	7.18, 3.21

Table 5.8: The mean number of co-infected individuals, $\bar{C}$, for the parametric choices and initial conditions as described in Figure 5.15.

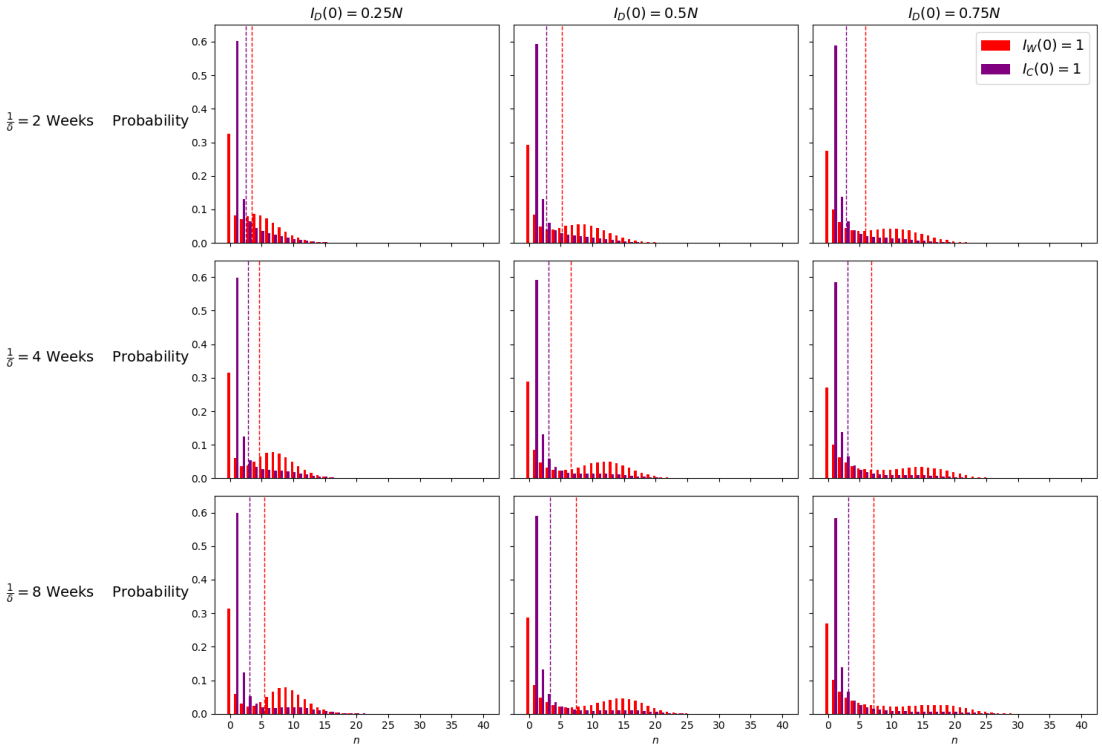

Figure 5.15: Histograms showing the probability distribution of C, $\{\xi(n), n = 0, \ldots, 40\}$, for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, a range of rates at which the DIP protection decays from a DIP-infected individual, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ (red) or co-infected $(I_C(0) = 1)$ (purple). Dashed lines indicate the mean values, $\bar{C}$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (red) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (purple).

$\bar{C}\|I_W(0)=1, \bar{C}\|I_D=1$	$I_D(0)=0.25N$	$I_D(0)=0.5N$	$I_D(0)=0.75N$
$\beta_C=\beta_W$	10.2, 12.5	14.4, 14.5	16.5, 15.7
$\beta_C=\frac{2\beta_W}{3}$	7.09, 6.86	10.3, 8.07	11.5, 8.60
$\beta_C=\frac{\beta_W}{3}$	4.64, 2.82	**6.62, 3.14**	6.87, 3.14

Table 5.9: The mean number of co-infected individuals, $\bar{C}$, for the parametric choices and initial conditions as described in Figure 5.16.

effect on an individual decays, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, on the probability distribution of the cumulative number, C, of co-infected individuals, $\{\xi(n), n = 0, \ldots, 40\}$, and its mean $\bar{C}$, for the cases in which the initial infection is a WT-infection, $I_W(0) = 1$, or a co-infection, $I_C(0) = 1$. It is interesting to note that varying $\frac{1}{\delta}$ has a small impact on $\bar{C}$ for the initial condition $I_C(0) = 1$, this is likely due to the fact that with these conditions an outbreak has a high probability of being relatively small and as such, not many DIP-affected individuals have the opportunity to decay to susceptible individuals during the outbreak. However, varying $\frac{1}{\delta}$ has a larger effect on the mean value of $\bar{C}$ for the initial condition $I_W(0) = 1$. This is related to the bi-modal shape of the red histogram for C, which implies that the outbreak is more likely to be larger for this initial condition and therefore the effect of DIP-infected individuals decaying to susceptible individuals will have a larger impact on the cumulative number of co-infected individuals. As $\frac{1}{\delta}$ increases, each of the red histograms are shifted to the right whilst retaining their bi-modal shape. However, as the initial number of individuals in the DIP-infected population increases the curve shifts to the right but instead of retaining its shape, it flattens and widens. This is likely related to the fact that in order to observe more co-infections, more DIP-infected individuals are required in the population.

Figure 5.16 and Table 5.9 show the effect of varying the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, and the infection rate of a co-infected individual, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, on the probability distribution of the cumulative number, C, of co-infected individuals, $\{\xi(n), n = 0, \ldots, 40\}$, and its mean, $\bar{C}$, for the cases where the initial infection is a WT-infection, $I_W(0) = 1$, or a co-infection, $I_C(0) = 1$. It can be seen that varying β_C has a significant impact on the distribution of the cumulative number of co-infected individuals as we would expect. When β_C is relatively large (equal to β_W), the outbreak is very likely to be sustained and more co-infections are likely to occur. Whereas when β_C is relatively small (equal to $\frac{\beta_W}{3}$) the outbreak is much more likely to be shorter and less co-infections will occur. This is reflected in the mean

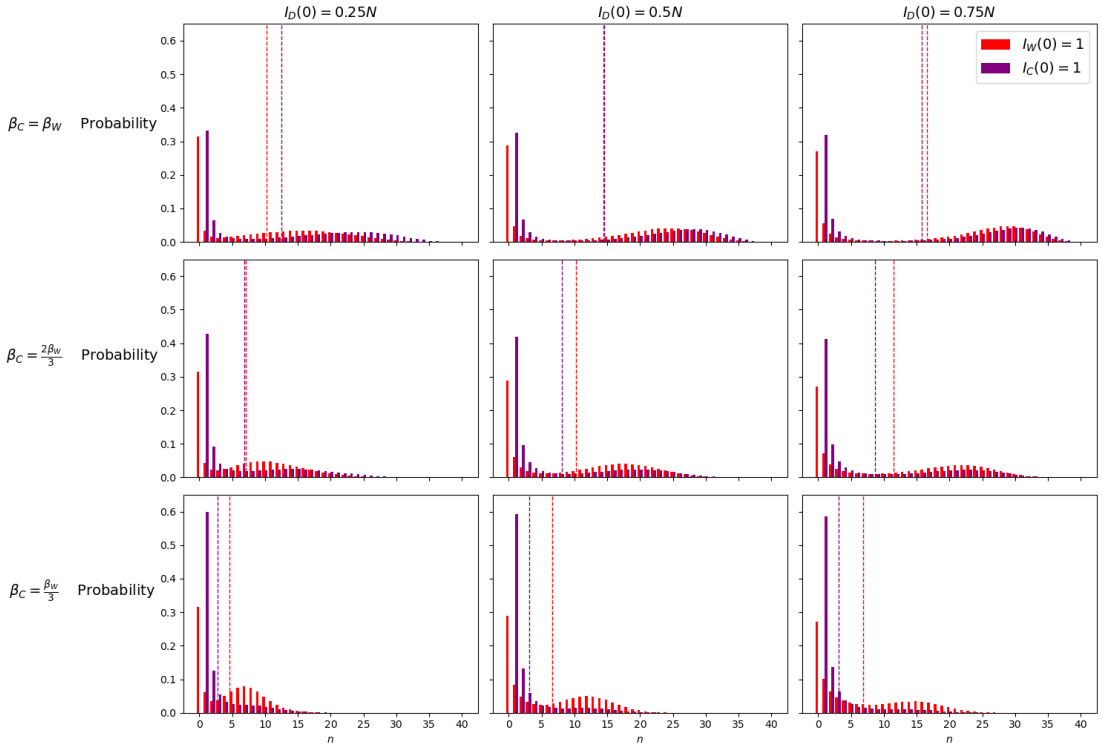

Figure 5.16: Histograms showing the probability distribution of $C, \{\xi(n), n = 0, \ldots, 40\}$, for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, a range of co-infection rates, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ (red) or co-infected $(I_C(0) = 1)$ (purple). Dashed lines indicate the mean values, $\bar{C}$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (red) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (purple).

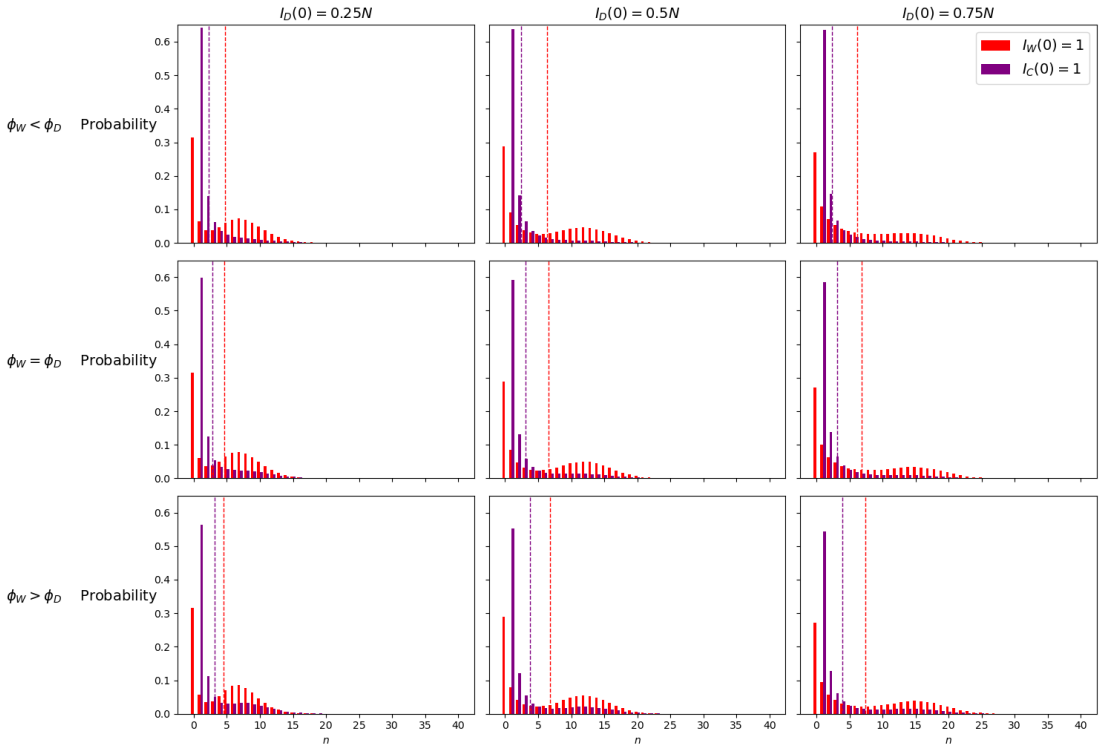

Figure 5.17: Histograms showing the probability distribution of $C, \{\xi(n), n = 0, \ldots, 40\}$, for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, three different cases of probabilities of the outcome of a $I_C - S$ infection event, $\phi_W < \phi_D, \phi_W = \phi_D$, or $\phi_W > \phi_D$ with specific probabilities given in Table 5.7, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ (red) or co-infected $(I_C(0) = 1)$ (purple). Dashed lines indicate the mean values, $\bar{C}$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (red) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (purple).

values, $\bar{C}$, for both initial conditions. When the initial number of DIP-infected individuals is increased, the mean expected number of co-infected individuals, $\bar{C}$, also increases. This, while seemingly counter-intuitive that more infections would be observed with more protection at the population level, makes empirical sense as a DIP-infected person only needs exposure to the WT to become co-infected and therefore a higher proportion of all infections will be co-infection events.

Figure 5.17 and Table 5.10 show the effect that varying the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.5N, 0.75N\}$, and the probability that a co-

$\bar{C}\|I_W(0)=1, \bar{C}\|I_C(0)=1$	$I_D(0)=0.25N$	$I_D(0)=0.5N$	$I_D(0)=0.75N$
$\phi_W < \phi_C$	7.42, 2.30	6.34, 2.39	6.21, 2.35
$\phi_W = \phi_C$	4.64, 2.82	**6.62, 3.14**	6.87, 3.14
$\phi_W > \phi_C$	4.52, 3.20	6.80, 3.81	7.45, 3.95

Table 5.10: The mean number of co-infected individuals, $\bar{C}$, for the parametric choices and initial conditions as described in Figure 5.17.

$\bar{C}\|I_W(0)=1, \bar{C}\|I_C(0)=1$	$\frac{1}{\delta}=2$ weeks	$\frac{1}{\delta}=4$ weeks	$\frac{1}{\delta}=8$ weeks
$\beta_C = \beta_W$	12.3, 13.0	14.4, 14.5	15.6, 15.4
$\beta_C = \frac{2\beta_W}{3}$	8.36, 6.97	10.3, 8.07	11.3, 8.72
$\beta_C = \frac{\beta_W}{3}$	5.21, 2.76	**6.62, 3.14**	7.40, 3.35

Table 5.11: The mean number of co-infected individuals, $\bar{C}$, for the parametric choices and initial conditions as described in Figure 5.18.

infected individual will pass on solely the WT-strain, ϕ_W, or DIP-strain, ϕ_D, with parameter regions $\phi_W < \phi_D, \phi_W = \phi_D$, or $\phi_W > \phi_D$ (the specific values used for ϕ_W and ϕ_D are given in Table 5.7), during an $I_C - S$ infection interaction, has on the probability distribution of the cumulative number, C, of co-infected individuals, $\{\xi(n), n = 0, \ldots, 40\}$, and its mean, $\bar{C}$, for the cases where the initial infection is a WT-infection, $I_W(0) = 1$, or a co-infection, $I_C(0) = 1$. Interesting results can be seen as ϕ_W increases relative to ϕ_D; when $I_D(0) = 0.25N$ we observe that ϕ_W increasing causes the mean number of co-infected individuals to decrease for the initial condition $I_W(0) = 1$. This is likely due to the fact that for these parameter values and initial conditions, the majority of individuals within the outbreak will be infected with the WT and not exposed to the DIP. However, as the initial number of DIP-infected individuals increases, increasing the value of ϕ_W relative to ϕ_D causes the mean number of co-infected individuals to increase for both initial conditions. This can be explained by the fact that ϕ_W increasing means that the WT strain is more likely to be passed on, and as the population has a large number of DIP-infected individuals, it becomes likely that they will become co-infected. It can also be observed that the number of initial DIP-infected individuals has an effect on the distributions; in each case increasing this quantity both shifts the distribution to the right and flattens it slightly, meaning that there is more likely to be a larger number of co-infected individuals over a wider range.

Figure 5.18 and Table 5.11 show the effect that varying the infection rate for a co-infected individual, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, and the rate at which a DIP-infected individual has

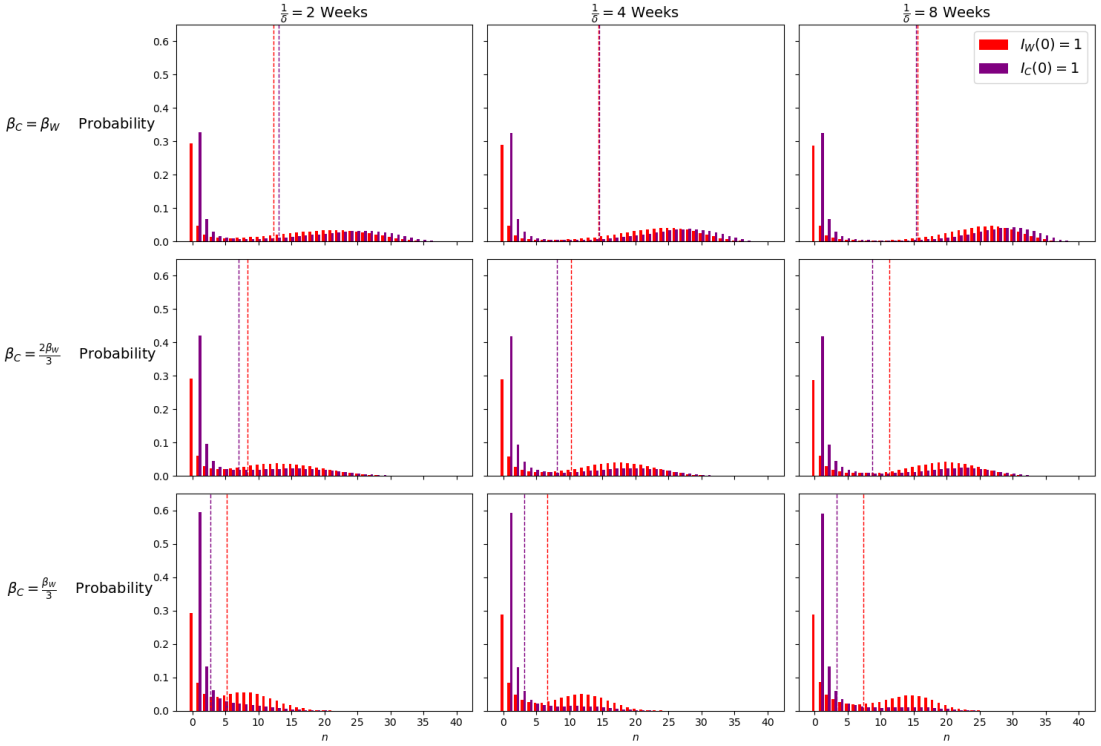

Figure 5.18: Histograms showing the probability distribution of $C, \{\xi(n), n = 0, \ldots, 40\}$, for different rates at which the DIP protection decays in a DIP-infected individual, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, a range of co-infection rates, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ (red) or co-infected $(I_C(0) = 1)$ (purple). Dashed lines indicate the mean values, $\bar{C}$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (red) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (purple).

$\bar{C}\|I_W(0)=1, \bar{C}\|I_C(0)=1$	$\phi_W < \phi_D$	$\phi_W = \phi_D$	$\phi_W > \phi_D$
$\beta_C = \beta_W$	13.8, 12.1	14.4, 14.5	14.1, 15.5
$\beta_C = \frac{2\beta_W}{3}$	9.66. 6.11	10.3, 8.07	10.4, 9.31
$\beta_C = \frac{\beta_W}{3}$	6.34, 2.39	**6.62, 3.14**	6.80, 3.82

Table 5.12: The mean number of co-infected individuals, $\bar{C}$, for the parametric choices and initial conditions as described in Figure 5.19.

their protection decay, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, has on the probability distribution of the cumulative number, C, of co-infected individuals, $\{\xi(n), n = 0, \ldots, 40\}$, and its mean, $\bar{C}$, for the cases where the initial infection is a WT-infection, $I_W(0) = 1$, or a co-infection, $I_C(0) = 1$. It is interesting to note the seemingly counter-intuitive result that as $\frac{1}{\delta}$ increases, and therefore individuals are more likely to have some protection from the disease for longer, the mean cumulative number of co-infected individuals actually increases. However, this remains a valid result since our focus here is not on the number of individuals infected in total, but on the number of co-infected individuals. Therefore, if individuals have a higher chance to be in I_D (as the rate of DIP decay is lower), then there is a higher probability of these individuals becoming co-infected as they only need exposure to the WT strain. The result that as β_C decreases the average number of co-infected individuals decreases can be expected as the overall chance of an individual being infected by a co-infected individual decreases. It is also notable and intuitive that this effect is more pronounced in the case when $I_C(0) = 1$. The histograms themselves represent this behaviour too, as $\frac{1}{\delta}$ increases, each of the histograms are shifted to the right whilst retaining a similar shape. Similarly, the histograms retain a similar profile as β_C decreases, whilst shifting to the left.

Figure 5.19 and Table 5.12 show the effect that varying the infection rate for a co-infected individual, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, and the probability that a co-infected individual will pass on solely the WT-strain, ϕ_W, or DIP-strain, ϕ_D, has on the probability distribution of the cumulative number, C, of co-infected individuals, $\{\xi(n), n = 0, \ldots, 40\}$, and its mean, $\bar{C}$, for the cases where the initial infection is a WT-infection, $I_W(0) = 1$, or a co-infection, $I_C(0) = 1$. One of the things to note is that similar behaviour to the results in Figure 5.18 can be seen here. That is, in the top row where $\beta_C = \beta_W$, the mean, $\bar{C}$, is larger for the initial condition $I_W(0) = 1$ in some parameter regions and larger for the initial condition $I_C(0) = 1$ in other parameter regions. It is interesting to note that as ϕ_W becomes larger than ϕ_D, the mean number of co-infected individuals increases,

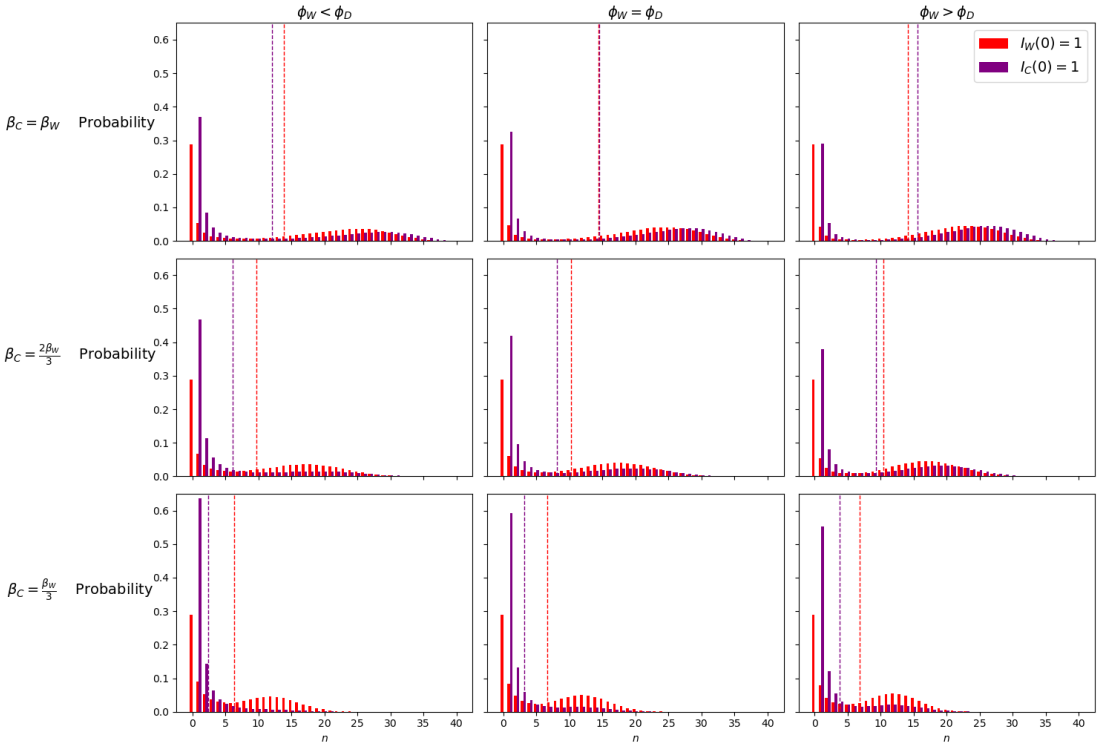

Figure 5.19: Histograms showing the probability distribution of $C, \{\xi(n), n = 0, \ldots, 40\}$, for different rates of infection for a co-infected individual, $\beta_C \in \{\beta_W, \frac{2\beta_W}{3}, \frac{\beta_W}{3}\}$, three different cases of probabilities of the outcome of a $I_C - S$ infection event, $\phi_W < \phi_D, \phi_W = \phi_D$ or $\phi_W > \phi_D$ (specific values given in Table 5.7), and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ (red) or co-infected $(I_C(0) = 1)$ (purple). Dashed lines indicate the mean values, $\bar{C}$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (red) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (purple).

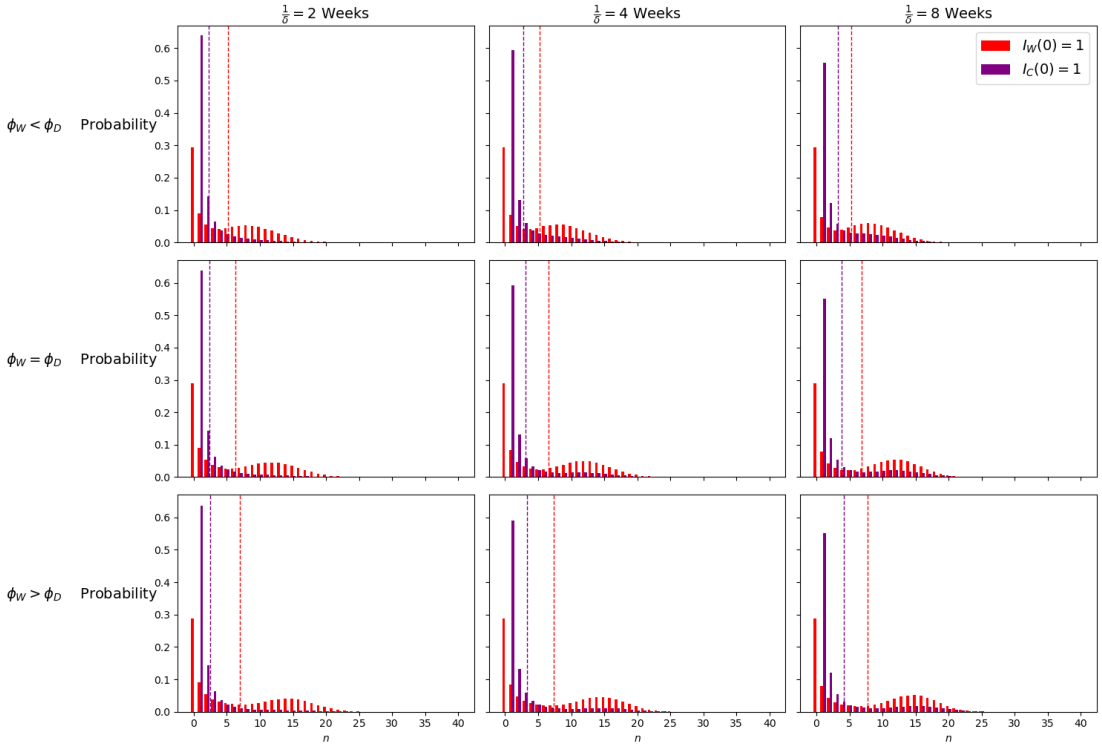

Figure 5.20: Histograms showing the probability distribution of $C, \{\xi(n), n = 0, \ldots, 40\}$, for different rates at which the DIP protection decays in a DIP-infected individual, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, three different cases of probabilities of the outcome of a $I_C - S$ infection event, $\phi_W < \phi_D, \phi_W = \phi_D$ or $\phi_W > \phi_D$ (specific values given in Table 5.7), and the individual starting the outbreak either being WT-infected $(I_W(0) = 1)$ (red) or co-infected $(I_C(0) = 1)$ (purple). Dashed lines indicate the mean values, $\bar{C}$. Baseline parameters as in Table 5.1. Initial states $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$ (red) or $(N - I_D(0) - 1, I_D(0), 0, 1)$ (purple).

even though a susceptible individual becomes less likely to become infected with the DIP. This is due to the higher probability of ϕ_W meaning that more individuals will be WT-infected and, on average, the outbreak itself will be larger. Again as β_C decreases, the mean number of co-infected individuals decreases and the histogram shifts to the left whilst retaining its bi-modal shape. As ϕ_W increases in relation to ϕ_D the histogram for the initial condition $I_C(0) = 1$ shows interesting behaviour. This histogram does not shift to the right, but probabilities in the approximate region, $n \in \{25, \ldots, 35\}$, become more concentrated around the second peak.

$\bar{C}\|I_W(0)=1, \bar{C}\|I_C(0)=1$	$\frac{1}{\delta}=2$ weeks	$\frac{1}{\delta}=4$ weeks	$\frac{1}{\delta}=8$ weeks
$\phi_W < \phi_C$	5.14, 2.25	5.21, 2.76	5.24, 3.20
$\phi_W = \phi_C$	6.34, 2.39	**6.62, 3.14**	6.80, 3.82
$\phi_W > \phi_C$	6.93, 2.45	7.40, 3.35	7.76, 4.20

Table 5.13: The mean number of co-infected individuals, $\bar{C}$, for the parametric choices and initial conditions as described in Figure 5.20.

Figure 5.20 and Table 5.13 show the effect that varying the rate at which a DIP-infected individual reverts to a susceptible individual, $\delta \in \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\}$ weeks^{-1}, and the probability that a co-infected individual will pass on solely the WT-strain, ϕ_W, or DIP-strain, ϕ_D, has on the probability distribution for the cumulative number, C, of co-infected individuals, $\{\xi(n), n = 0, \ldots, 40\}$, for the cases that the initial infection is a WT-infection, $I_W(0) = 1$, or a co-infection, $I_C(0) = 1$. The chance of no outbreak is the same for an initially WT-infected individual but the histogram is pushed to the right as ϕ_W becomes larger in proportion to ϕ_D, which is to be expected as WT is more infectious so there is more likely to be a sustained outbreak. For $I_C(0) = 1$, the probability of the outbreak causing no further co-infected individuals decreases as $\frac{1}{\delta}$ increases. This is likely due to the fact that DIP-infected individuals can only become co-infected in an infection event whereas a susceptible individual could become WT-infected or DIP-infected rather than co-infected. The mean number of co-infected individuals, $\bar{C}$, increases by a small amount as $\frac{1}{\delta}$ is increased, meaning that this decay rate affects the number of co-infected individuals. This seems counter-intuitive, that a longer decay rate results in more infections; however, this quantity does not represent the size of the outbreak, only the number of co-infected individuals. So in tandem with the results regarding the size of the outbreak, this increase just means a higher proportion of the infections are co-infections. As ϕ_W grows in relation to ϕ_D, the mean number of co-infected individuals grows and the histogram is pushed to the right whilst retaining its shape.

5.3.3 Reproduction number

In this section, the focus is on the effect that varying pairs of parameters within the model has on the reproduction numbers $R^C(W), R^C(D)$ and R^W, the stochastic descriptors described in Section 5.2.3.

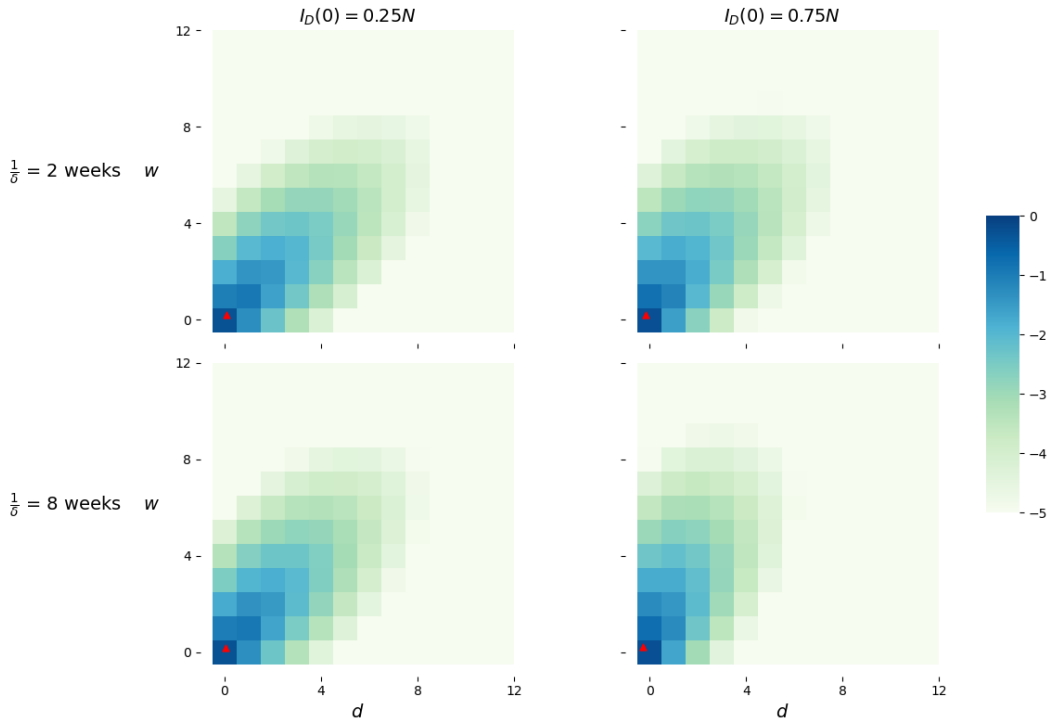

Figure 5.21: Heatmaps showing the impact on $\log_{10}(\eta^C_{(s,i_d,i_w,i_c)}(w,d))$ of different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the rate at which the protection afforded by the DIP decays in a DIP-infected individual, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1}. Baseline parameters as in Table 5.1. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 0, 1)$. Red triangles represent the mean values of the random variables, $\left(\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]\right)$.

5.3.3.1 A marked co-infected individual: $R^C(W)$ and $R^C(D)$

Here we consider the distribution of $\left(R^C(W), R^C(D)\right)$, namely the probability of a marked co-infected individual directly causing W wild-type infections and D DIP-infections before the co-infected individual either recovers or dies.

Within Figure 5.21 and Table 5.14, the impact that changing the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the rate at which the DIP protection decays within a DIP-infected individual, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1}, has on the probability distribution of the number of WT-infections, $R^C(W)$, and DIP-infections, $R^C(D)$, caused by a marked co-infected individual is studied.

$\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]$	$I_D(0) = 0.25N$	$I_D(0) = 0.75N$
$\frac{1}{\delta} = 2$ weeks	0.695, 0.581	0.704, 0.339
$\frac{1}{\delta} = 8$ weeks	0.698, 0.542	0.708, 0.221

Table 5.14: The mean number of WT-infections, $\mathbb{E}\left[R^C(W)\right]$, and DIP-infections, $\mathbb{E}\left[R^C(D)\right]$, caused by a marked co-infected individual for the parametric choices and initial conditions as described in Figure 5.21.

$\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]$	$I_D(0) = 0.25N$	$I_D(0) = 0.75N$
$\beta_C = \beta_W$	1.83, 1.46	1.85, 0.661
$\beta_C = \frac{\beta_W}{3}$	0.697, 0.557	0.706, 0.268

Table 5.15: The mean number of WT-infections, $\mathbb{E}\left[R^C(W)\right]$, and DIP-infections, $\mathbb{E}\left[R^C(D)\right]$, caused by a marked co-infected individual for the parametric choices and initial conditions as described in Figure 5.22.

The heatmaps show the distribution of $\left(R^C(W), R^C(D)\right)$; we can see that as the initial number of DIP-infected individuals increases the number of DIP-infections tends to have a slight reduction. This is also shown in the mean values where the $\mathbb{E}[R^C(D)]$ decreases when $I_D(0)$ increases. For $I_D(0) = 0.25N$, as $\frac{1}{\delta}$ increases; we see that $\mathbb{E}[R^C(D)]$ decreases by a very small amount. However, when $I_D(0) = 0.75N$, the same increase in the DIP decay period takes a third off the mean value of $R^C(D)$. This is likely due to the fact that DIP-infected individuals protection taking longer to decay into susceptible individuals is more significant on the profile of the outbreak when the susceptible population is already small. It is important to note that R^C remains largely unaffected by the changes to these parameters; this is largely intuitive as the wild-type strain can infect either a susceptible individual or DIP-infected individual.

In Figure 5.22 and Table 5.15, the impact that changing the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the infection rate for co-infected individuals, $\beta_C \in \{\beta_W, \frac{\beta_W}{3}\}$, has on the number of WT-infections, $R^C(W)$, and DIP-infections, $R^C(D)$, caused by a marked co-infected individual is studied.

We can see that increasing $I_D(0)$ has a significant impact on the expected number of DIP-infections, shifting the profile of the heatmap to the left and reducing the mean number of these type of infections by two thirds. The number of WT-infections also very slightly increases as this happens. This is likely as, under these conditions, it is more likely that the first event that happens is a $I_D \to I_C$ WT-infection and less likely that a $S \to I_W$ infection is the first event, compared to the case that the initial number

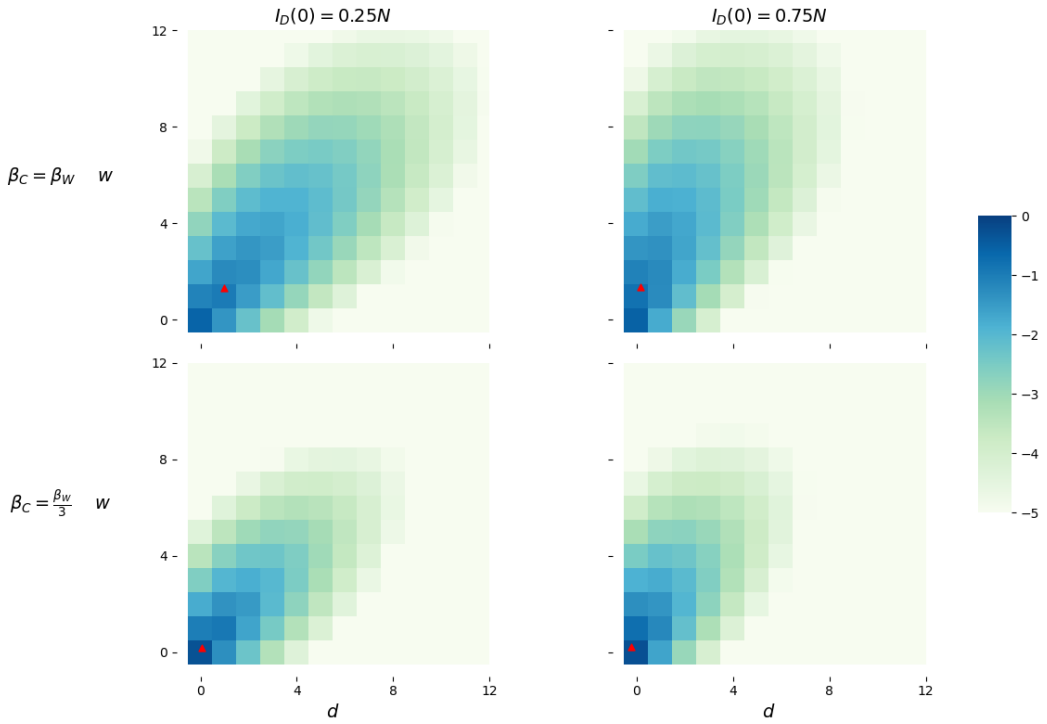

Figure 5.22: Heatmaps showing $\log_{10}(\eta^C_{(s,i_d,i_w,i_c)}(w,d))$ for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the infection rate of a co-infected individual, $\beta_C \in \{\beta_W, \frac{\beta_W}{3}\}$. Baseline parameters as in Table 5.1. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 0, 1)$. Red triangles represent the mean values of the random variables, $\left(\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]\right)$.

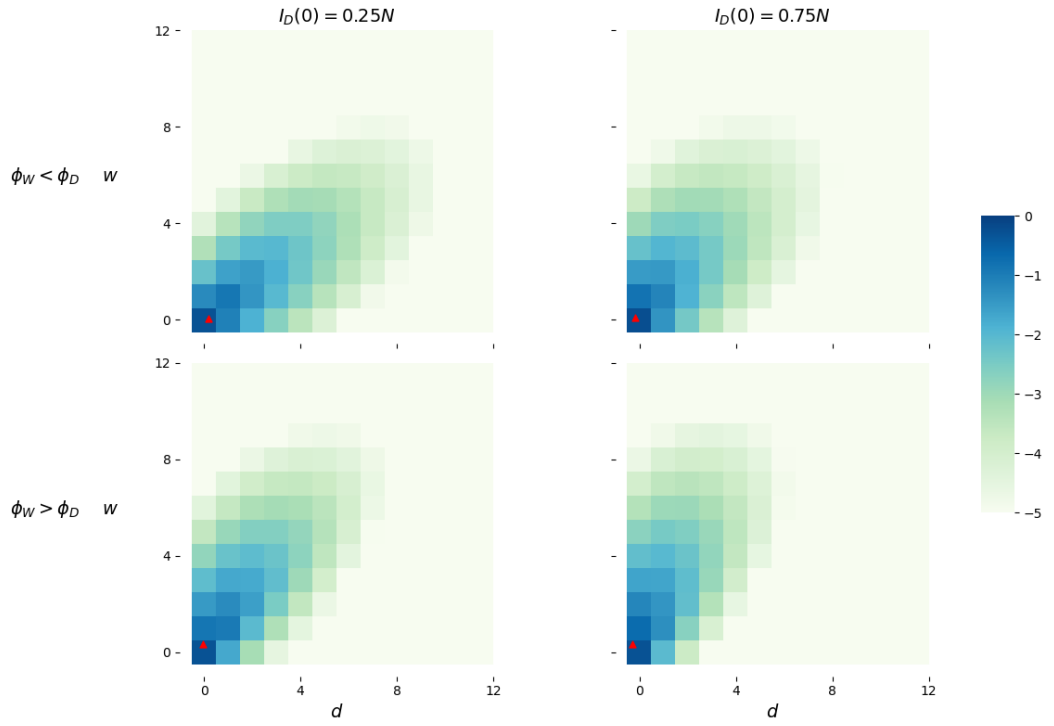

Figure 5.23: Heatmaps showing $\log_{10}(\eta^C_{(s,i_d,i_w,i_c)}(w,d))$ for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the probability that a co-infected individual passes on solely the WT-strain, ϕ_W, or the DIP, ϕ_D, in an $I_C - S$ infection interaction, $\phi_W < \phi_D$ or $\phi_W > \phi_D$ (specific values given in Table 5.7). Baseline parameters as in Table 5.1. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 0, 1)$. Red triangles represent the mean values of the random variables, $\left(\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]\right)$.

of DIP-infected individuals is low. Newly co-infected individuals are less infectious than newly WT-infected individuals and therefore the "competition" for our marked co-infected individual is lower and as such, the average number of WT-infections increases for our marked individual.

As expected, decreasing the rate of infection of a co-infected individual, β_C, has a pronounced effect on the number of both type of infections. This is reflected in the means for both types of infections; see Table 5.15.

In Figure 5.23 and Table 5.16, the impact that changing the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the probability of a co-infected individual

$\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]$	$I_D(0) = 0.25N$	$I_D(0) = 0.75N$
$\phi_W < \phi_C$	0.568, 0.674	0.572, 0.324
$\phi_W > \phi_C$	0.821, 0.443	0.837, 0.213

Table 5.16: The mean number of WT-infections, $\mathbb{E}\left[R^C(W)\right]$, and DIP-infections, $\mathbb{E}\left[R^C(D)\right]$, caused by a marked co-infected individual for the parametric choices and initial conditions as described in Figure 5.23.

$\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]$	$\frac{1}{\delta} = 2$ weeks	$\frac{1}{\delta} = 8$ weeks
$\beta_C = \beta_W$	1.84, 1.18	1.85, 0.985
$\beta_C = \frac{\beta_W}{3}$	0.700, 0.460	0.704, 0.382

Table 5.17: The mean number of WT-infections, $\mathbb{E}\left[R^C(W)\right]$, and DIP-infections, $\mathbb{E}\left[R^C(D)\right]$, caused by a marked co-infected individual for the parametric choices and initial conditions as described in Figure 5.24.

passing on solely the WT-strain, ϕ_W, or DIP-strain, ϕ_D, with parameters $\phi_W < \phi_D$ or $\phi_W > \phi_D$ (specific values given in Table 5.7) has on the probability distribution of the number of WT-infections, $R^C(W)$, and DIP-infections, $R^C(D)$, caused by a marked co-infected individual is studied.

We can see that increasing $I_D(0)$ has a significant impact on the expected number of DIP-infections, shifting the profile of the heatmap to the left and reducing the mean number of these type of infections by approximately half. Again, we see that the number of WT-infections also very slightly increases as this change occurs. As is to be expected, when a co-infected individual becomes more likely to pass on the WT-strain rather than the DIP-strain, this is reflected in the means for both types of infections.

In Figure 5.24 and Table 5.17, the impact that changing the rate of infection for co-infected individuals, $\beta_C \in \{\beta_W, \frac{\beta_W}{3}\}$, and the DIP-decay rate, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1}, has on the probability distribution of the number of WT-infections, $R^C(W)$, and DIP-infections, $R^C(D)$, caused by a marked co-infected individual is studied.

We can see a clear effect when the rate of infection of a co-infected individual, β_C, is decreased, lowering the expectation of both types of infection. As in Figure 5.21, increasing the length of protection from the DIP reduces the number of DIP-infections whilst slightly increasing the expected number of WT-infections.

Within Figure 5.25 and Table 5.18, the impact that changing the rate of infection for co-infected individuals, $\beta_C \in \{\beta_W, \frac{\beta_W}{3}\}$, and the probability that a co-infected individual

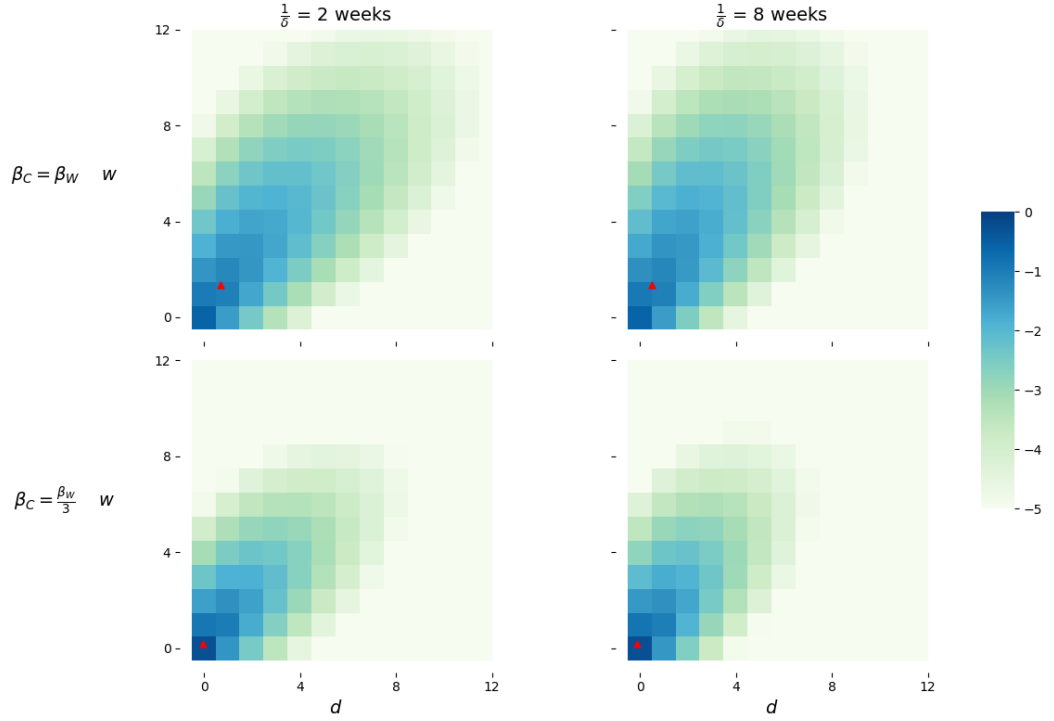

Figure 5.24: Heatmaps showing $\log_{10}(\eta^C(w,d))$ for differing rates at which the protection afforded by the DIP decays in a DIP-infected individual, $\frac{1}{\delta} \in \{2, 8\}$ weeks, and the rate of infection of a co-infected individual, $\beta_C \in \{\beta_W, \frac{\beta_W}{3}\}$. Baseline parameters as in Table 5.1. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (0.5N - 1, 0.5N, 0, 1)$. Red triangles represent the mean values of the random variables, $\left(\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]\right)$.

$\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]$	$\phi_W < \phi_D$	$\phi_W > \phi_D$
$\beta_C = \beta_W$	1.56, 1.30	2.10, 0.834
$\beta_C = \frac{\beta_W}{3}$	0.570, 0.499	0.830, 0.328

Table 5.18: The mean number of WT-infections, $\mathbb{E}\left[R^C(W)\right]$, and DIP-infections, $\mathbb{E}\left[R^C(D)\right]$, caused by a marked co-infected individual for the parametric choices and initial conditions as described in Figure 5.25.

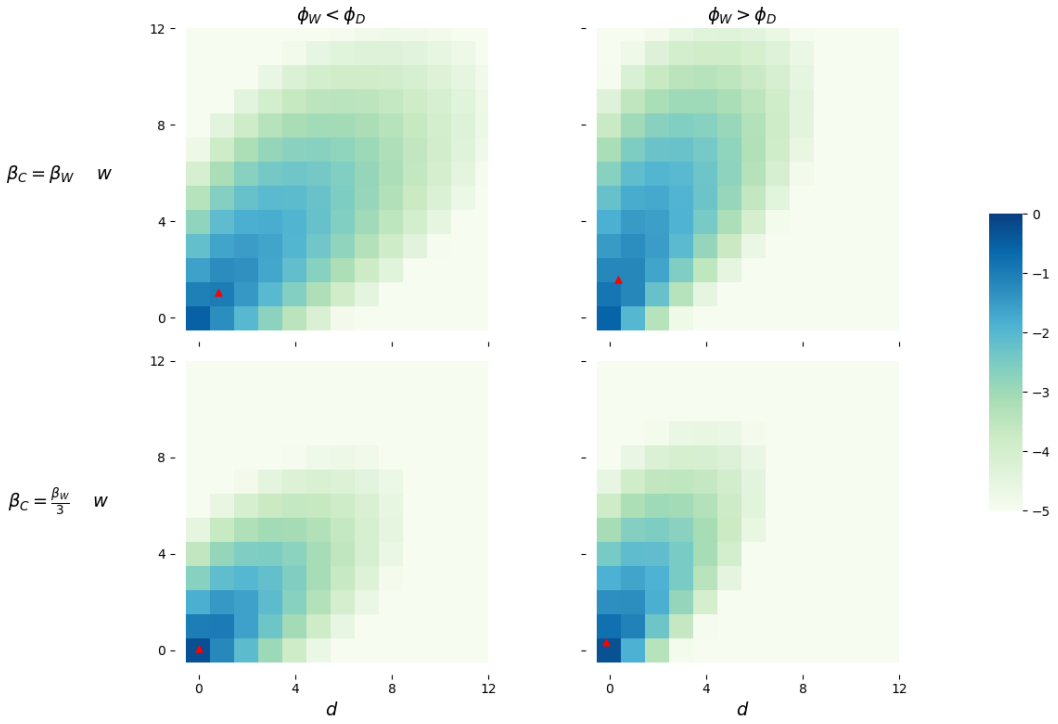

Figure 5.25: Heatmaps showing $\log_{10}(\eta^C_{(s,i_d,i_w,i_c)}(w,d))$ for differing rates of infection of a co-infected individual, $\beta_C \in \{\beta_W, \frac{\beta_W}{3}\}$ and the probability that a co-infected individual passes on solely the WT-strain, ϕ_W, or the DIP, ϕ_D, in an $I_C - S$ infection interaction, $\phi_W < \phi_D$ or $\phi_W > \phi_D$ (specific values given in Table 5.7). Baseline parameters as in Table 5.1. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (0.5N - 1, 0.5N, 0, 1)$. Red triangles represent the mean values of the random variables, $\left(\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]\right)$.

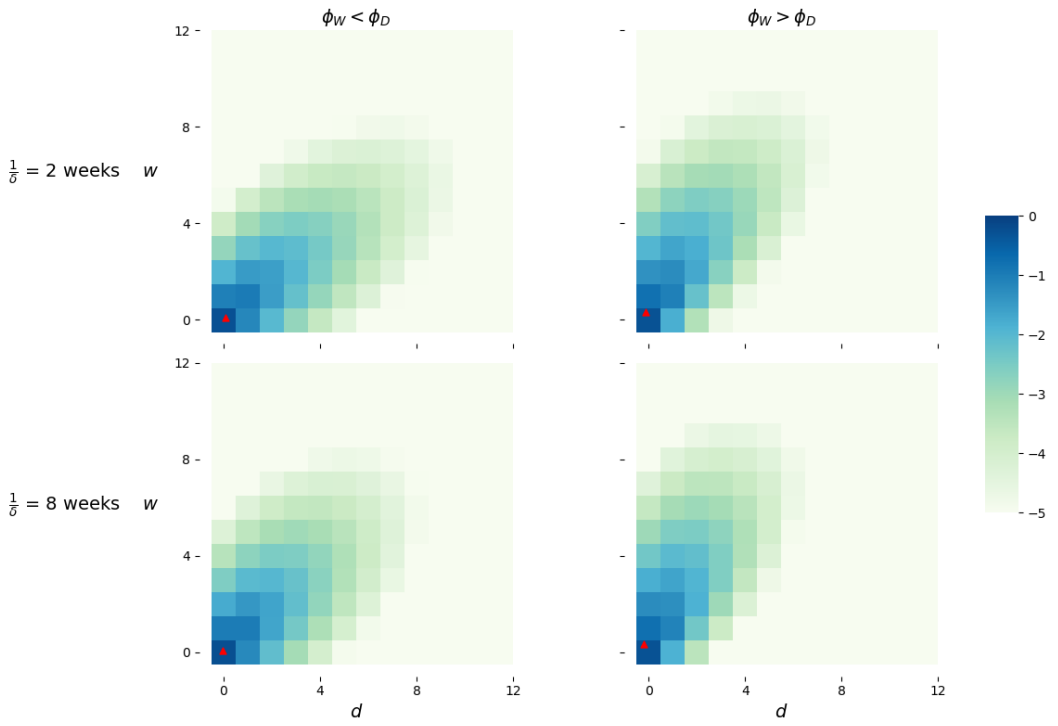

Figure 5.26: Heatmaps showing $\log_{10}(\eta^C_{(s,i_d,i_w,i_c)}(w,d))$ for differing rates at which the protection the DIP offers a DIP-infected individual decays, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1} and the probability that a co-infected individual passes on solely the WT-strain, ϕ_W, or the DIP, ϕ_D, in an $I_C - S$ infection interaction, $\phi_W < \phi_D$ or $\phi_W > \phi_D$. Baseline parameters as in Table 5.1. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (0.5N - 1, 0.5N, 0, 1)$. Red triangles represent the mean values of the random variables, $\left(\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]\right)$.

passes on solely the WT-strain, ϕ_W, or the DIP-strain, ϕ_D, with parameters chosen such that $\phi_W < \phi_D$ or $\phi_W > \phi_D$ (specific values given in Table 5.7), has on the probability distribution of the number of WT-infections, $R^C(W)$, and DIP-infections, $R^C(D)$, caused by a marked co-infected individual is studied.

Here we can clearly see the effects caused by varying the parameters. When the rate of infection of a co-infected individual, β_C, is decreased, the expectation of both types of infection lowers significantly and the heatmap profile shifts towards the lower end of the spectrum for both types of infection. We can also see that in the case that $\phi_W > \phi_D$, WT-infections become more prevalent whilst DIP-infections decrease.

In Figure 5.26 and Table 5.19, the impact that changing the rate at which a DIP-infected

$\mathbb{E}\left[R^C(W)\right], \mathbb{E}\left[R^C(D)\right]$	$\frac{1}{\delta} = 2$ weeks	$\frac{1}{\delta} = 8$ weeks
$\phi_W < \phi_C$	0.569, 0.557	0.571, 0.461
$\phi_W > \phi_C$	0.827, 0.365	0.832, 0.304

Table 5.19: The mean number of WT-infections, $\mathbb{E}\left[R^C(W)\right]$, and DIP-infections, $\mathbb{E}\left[R^C(D)\right]$, caused by a marked co-infected individual for the parametric choices and initial conditions as described in Figure 5.26.

$\mathbb{E}[R^W]$	$I_D(0) = 0.25N$	$I_D(0) = 0.75N$
$\frac{1}{\delta} = 2$ weeks	2.339	2.474
$\frac{1}{\delta} = 8$ weeks	2.360	2.523

Table 5.20: The mean number of infections caused by a marked WT-infected individual, $\mathbb{E}[R^W]$, for the parameter choices and initial conditions described in Figure 5.27.

individual reverts to a susceptible individual, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1}, and the probability that a co-infected individual passes on solely the WT-strain, ϕ_W, or the DIP-strain, ϕ_D, has on the probability distribution of the number of WT-infections, $R^C(W)$, and DIP-infections, $R^C(D)$, caused by a marked co-infected individual is studied. We see that increasing the length of protection the DIP gives causes a reduction in DIP-infection events whilst having a very small effect on the number of WT-infections. We also see that when ϕ_W is greater than ϕ_D, as opposed to less than, the number of WT-infections increases whilst the number of DIP-infections decreases.

5.3.3.2 A marked WT-infected individual: R^W

The number of infections caused by a marked WT-infected individual, R^W, is not significantly affected by changes to the DIP-related parameters; therefore we shall show the effect of varying one pair of DIP-related parameters for illustrative purposes only.

Figure 5.27 and Table 5.20 display the impact that changing the initial number of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$, and the rate at which a DIP-infected individual reverts to a susceptible individual, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1}, has on the probability distribution of the number of WT-infections caused by a marked WT-infected individual, R^W, and its mean, $\mathbb{E}[R^W]$.

As is to be expected, these DIP-related parameters do not have a significant impact on the number of WT-infections by a marked individual. However, note that whilst the impact is small, increasing the period over which the protection afforded by the

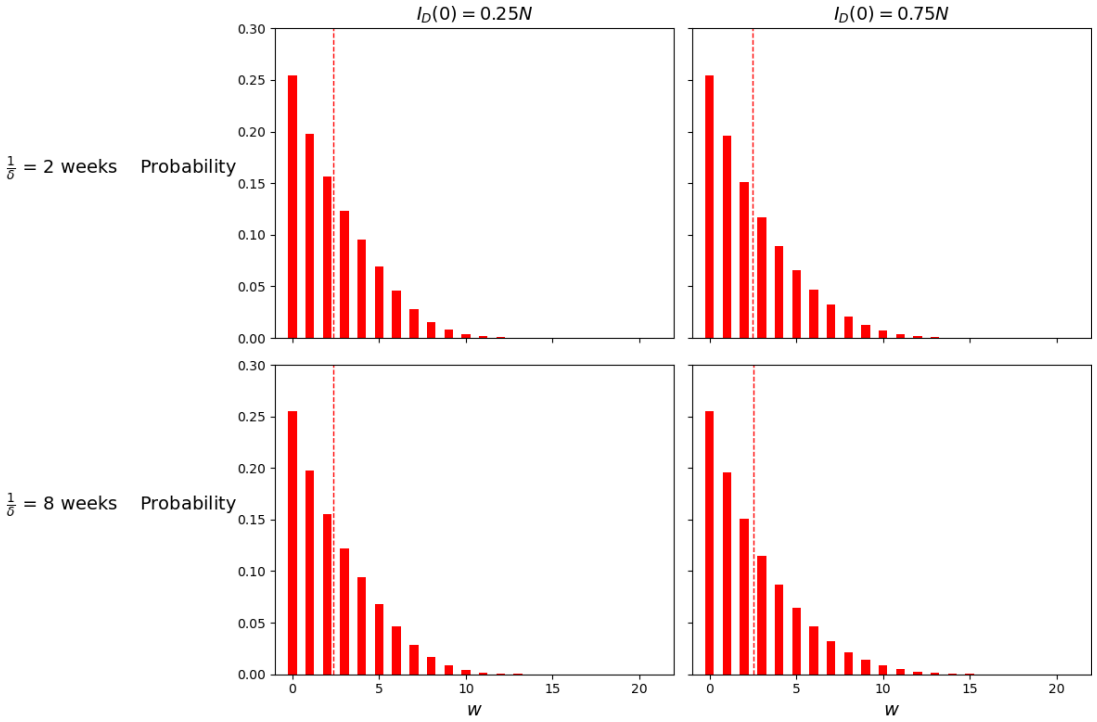

Figure 5.27: Histograms showing the probability distribution of R^W, $\{\eta^W_{(s,i_d,i_w,i_c)}, w = 0,\ldots,20\}$ for different initial numbers of DIP-infected individuals, $I_D(0) \in \{0.25N, 0.75N\}$ and the rate at which the protection afforded by the DIP decays in a DIP-infected individual, $\delta \in \{\frac{1}{2}, \frac{1}{8}\}$ weeks^{-1}. Initial state $(S(0), I_D(0), I_W(0), I_C(0)) = (N - I_D(0) - 1, I_D(0), 1, 0)$. Dashed lines indicate the mean value of the random variable, $\mathbb{E}[R^W]$.

DIP decays increases the expected number of infections by our marked WT-infected individual. This seems counter-intuitive, but can be explained by the fact that if there are more DIP-infected individuals, there is a higher chance that a WT-infection creates a co-infected individual than a WT-infected individual; this is beneficial to our WT-infected individuals chances of causing an infection as WT-individuals are more infectious than co-infected individuals. This logic also applies for the result showing that increasing the initial number of DIP-infected individuals also increases $\mathbb{E}[R^W]$.

5.4 Discussion

We have defined and analysed a compartmental epidemic model of viral infection in the presence of both a wild-type and defective interfering strain. This is an area of research that has not been explored to a great depth previously, but building on previously well known models of viral epidemic dynamics we aim to explore the effect of a DIP-strain of virus on the number of individuals infected and their outcomes for an outbreak in a closed population. Our interest in the impact the presence a DIP-strain of a pathogenic virus can have is motivated by research being done into using defective interfering particles of a given virus as a therapeutic agent (known as TIPs); this is being researched for viruses such as SARS-CoV-2 (Locke *et al.* (2024)).

Our focus within this chapter has been investigating stochastic descriptors for the model, such as the size of the outbreak or the cumulative number of co-infected individuals over the course of the outbreak. In Section 5.2, the methods used to develop algorithms capable of finding the probability distribution and expectation of these stochastic descriptors were presented, as well as the algorithms themselves. These algorithms allow complete probability distributions to be calculated for the summary statistics we consider. Furthermore, in Section 5.3, we investigated the effect of varying different DIP-related parameters on each of the summary statistics presented, with a view to understanding the extent to which the characteristics of the DIP strain can impact the outbreak.

Our work shows that the characteristics of the DIP can have a profound impact on an outbreak of a virus, with a higher proportion of the population being infected by the DIP (75%) at the onset of the outbreak causing a marked decrease in the expected size of the

outbreak compared to if the number of DIP-infected individuals was low (25%). This is mirrored in the exact reproduction number

The work within this chapter is largely theoretical at its current stage, however one potential extension would be to attempt to use published data to calibrate some of our parameters in a more precise way.

Chapter 6

Concluding remarks

In this thesis, mathematical models for *in vitro* and *in vivo* dynamics of the bacterium *Bacillus anthracis*, the causative agent of the disease anthrax, have been considered. We have also considered a model of viral epidemic dynamics for a pathogenic wild-type strain of a given virus and defective interfering strain of the same virus within a closed population.

In Chapter 3, a two-compartment stochastic model describing the behaviour of toxin-producing bacteria has been defined and analysed. By making use of first-step arguments we have studied several summary statistics, such as the lifespan of a bacterium or the number of toxins produced by the bacterium and its progeny. Our methods have been illustrated by focusing on the bacterium *B. anthracis* under the case of antibiotic treatment. A potential avenue for further investigation is to use *in vitro* data to better calibrate the parameters in the model for this scenario. Whilst we have only considered two-compartments within the scope of this thesis, a particular strength of our methodology is that it could be applied to any network of compartments, as long as the bacteria behave independently. For example, in the future this could be applied in the case of *B. anthracis* once it enters the blood stream and makes its way throughout a host.

We begin to consider quantifying the toxin production of *B. anthracis* in Chapter 3 and extend this work in Chapter 4. We have developed and proposed a delay differential equation model of *in vitro* dynamics of *B. anthracis* growth and PA production and degradation, with the motivation of quantifying PA production and degradation *in vitro*. As PA

is essential in binding other toxin proteins to cell surfaces, we believe that understanding its dynamics *in vitro* is an important step towards gaining a fuller understanding of *in vivo* toxin dynamics. A benefit of the approach we undertook is that the model can be extended and adapted as further data is produced; this is shown by the incorporation of separate decay terms of PA due to the experiment with protease inhibitors by Zai *et al.* (2016) and its application to the other considered datasets.

In the work on the delay differential equation (DDE) model, we have made use of a Bayesian approach to calibrating parameters and compared the results between four separate *in vitro* datasets. The homogeneity of some parameters across the datasets, such as the PA production rate, suggests that these parameter values could be used as prior beliefs whilst considering corresponding parameters in a model of *in vivo* anthrax infection.

We also have produced a stochastic analog of a previously published model of within-host anthrax infection, by Day *et al.* (2011), within Chapter 4 by making use of an approximation of the Gillespie algorithm, known as tau-leaping. This stochastic analog allowed us to compute the response probabilities of the model for various initial spore exposures, an important result that cannot be obtained from a deterministic model.

Within Chapter 5, a compartmental epidemic model of viral infection in the presence of both a wild-type and defective interfering strain is introduced and analysed. This model builds upon well known models of wild-type viral epidemics and expands their scope to include DIP-infections and co-infections. The work in this chapter is motivated by research into using defective interfering particles of a given virus as a therapeutic agent, where the characteristics of the DIPs will have a significant impact on their efficacy, such as in the case of SARS-CoV-2 (Locke *et al.* (2024)).

To analyse summary statistics of the compartmental epidemic model we use techniques involving first-step arguments, as was also shown in Chapter 3. Here, the model is too complex to find a closed analytic solution for the summary statistics directly, and we therefore use these first-step arguments as the basis of an algorithmic approach to obtain the probability distributions and expectations of our stochastic descriptors, such as the number of individuals infected throughout the course of an outbreak or the reproduction number for a marked co-infected individual. The developed algorithms are then used to

demonstrate the impact of the DIP-related parameters on each of the summary statistics. The work done in this chapter shows that the characteristics of the DIP strain have a significant impact on the size and mortality rate of an outbreak of a pathogenic virus.

Appendix A

Python codes

Some of the codes used within this thesis have been provided in a Github repository. Here I will briefly explain the content of each code and the figures within the thesis they were used to produce.

A.1 C3-Lifespan-of-a-bacterium

We calculate the analytic value of $\mathbb{E}[T_1]$, from Section 3.3.1, along a grid of given values of ν_{12} and ν_{21}, for three different values of μ_2. We then present these results in heatmaps with an associated colourbar, which produces Figure 3.3.

A.2 C4-Tau-Leap-Dose-Response

This code makes use of the tau-leaping algorithm (see Section 4.1.4.3) as a stochastic analog to the model described by Day *et al.* (2011). We carry out 1000 simulations for a number of initial doses of spores and calculate the response probability for these doses. The median of the simulations for each initial dose exposure is also calculated and compared to the solution of the deterministic model. These correspond to Figures 4.13, 4.14, 4.15 and 4.16.

A.3 C4-Dstl-ABC-SMC and C4-Dstl-Plotting

The first code (C4-Dstl-ABC-SMC) follows the ABC-SMC algorithm for 20 iterations with a sample size of 200 parameter sets for the model described in Section 4.2. These accepted parameter sets are saved and indexed for use in plotting in the next code (C4-Dstl-Plotting). The code for plotting loads the saved parameter sets from 10 separate runs of the ABC-SMC algorithm and uses these parameter sets to produce Figures 4.33, 4.34, 4.35 and 4.36.

A.4 C5-Size-Distribution

Here we algorithmically calculate $\boldsymbol{\alpha}_{i,j}(r, d)$, using Algorithm 5.3. We then use these solutions to show the impact of varying the initial number of DIP-infected individuals in the population, as well as presenting the differences between whether the individual that started the outbreak was WT-infected or co-infected. This is shown in Figures 5.10 and Figures 5.11.

References

ALAHMADI, A.A., FLEGG, J.A., COCHRANE, D.G., DROVANDI, C.C. & KEITH, J.M. (2020). A comparison of approximate versus exact techniques for Bayesian parameter inference in nonlinear ordinary differential equation models. *Royal Society open science*, **7**, 191315. 121

ALBRINK, W.S. (1961). Pathogenesis of inhalation anthrax. *Bacteriological reviews*, **25**, 268–273. 76

ALLEN, L.J. (2007). An introduction to mathematical biology. *Upper Saddle River, New Jersey*. 36

ALLEN, L.J. (2010). *An introduction to stochastic processes with applications to biology*. CRC Press. xiii, 7, 16, 20, 24, 36

ARTALEJO, J. & LOPEZ-HERRERO, M. (2013). On the exact measure of disease spread in stochastic epidemic models. *Bulletin of mathematical biology*, **75**, 1031–1050. 180

BANKS, D.J., BARNAJIAN, M., MALDONADO-AROCHO, F.J., SANCHEZ, A.M. & BRADLEY, K.A. (2005). Anthrax toxin receptor 2 mediates bacillus anthracis killing of macrophages following spore challenge. *Cellular microbiology*, **7**, 1173–1185. 4, 37

BRITTON, T. (2010). Stochastic epidemic models: a survey. *Mathematical biosciences*, **225**, 24–35. 36

BROCKWELL, P.J. (1986). The extinction time of a general birth and death process with catastrophes. *Journal of Applied Probability*, **23**, 851–858. 36

BURRELL, C.J., HOWARD, C.R. & MURPHY, F.A. (2017). Virion structure and composition. *Fenner and White's Medical Virology*, 27. 5

CAO, Y., GILLESPIE, D.T. & PETZOLD, L.R. (2006). Efficient step size selection for the tau-leaping simulation method. *The Journal of chemical physics*, **124**. 99, 102

CARRUTHERS, J., LÓPEZ-GARCÍA, M., GILLARD, J.J., LAWS, T.R., LYTHE, G. & MOLINA-PARÍS, C. (2018). A novel stochastic multi-scale model of francisella tularensis infection to predict risk of infection in a laboratory. *Frontiers in microbiology*, **9**, 1165. 36, 37

CARRUTHERS, J., LYTHE, G., LÓPEZ-GARCÍA, M., GILLARD, J., LAWS, T.R., LUKASZEWSKI, R. & MOLINA-PARÍS, C. (2020). Stochastic dynamics of francisella tularensis infection and replication. *PLoS computational biology*, **16**, e1007752. 36, 37

CASTRO, M., LÓPEZ-GARCÍA, M., LYTHE, G. & MOLINA-PARÍS, C. (2018). First passage events in biological systems with non-exponential inter-event times. *Scientific reports*, **8**, 1–16. 36

CHAMBERS, J., YARRARAPU, S.N.S. & MATHAI, J.K. (2018). Anthrax infection. 2

CHARLTON, S., HERBERT, M., MCGLASHAN, J., KING, A., JONES, P., WEST, K., ROBERTS, A., SILMAN, N., MARKS, T., HUDSON, M. *et al.* (2007). A study of the physiology of bacillus anthracis sterne during manufacture of the uk acellular anthrax vaccine. *Journal of applied microbiology*, **103**, 1453–1460. xi, xix, 111, 114, 115, 133, 134, 135, 136, 138, 139, 140, 146, 147

CHIANG, C., BONGIORNI, C. & PEREGO, M. (2011). Glucose-dependent activation of bacillus anthracis toxin gene expression and virulence requires the carbon catabolite protein ccpa. *Journal of bacteriology*, **193**, 52–62. 116

CHOI, P.J., CAI, L., FRIEDA, K. & XIE, X.S. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*, **322**, 442–446. 36

COTE, C.K., WELKOS, S.L. & BOZUE, J. (2011). Key aspects of the molecular and cellular basis of inhalational anthrax. *Microbes and infection*, **13**, 1146–1155. 2

DAY, J., FRIEDMAN, A. & SCHLESINGER, L.S. (2011). Modeling the host response to inhalation anthrax. *Journal of theoretical biology*, **276**, 199–208. xi, xv, xxvii, 1, 3, 56, 66, 68, 76, 78, 79, 80, 81, 82, 83, 87, 89, 90, 92, 109, 146, 147, 149, 226, 229

DE LA HIGUERA, L., LÓPEZ-GARCÍA, M., CASTRO, M., ABOURASHCHI, N., LYTHE, G. & MOLINA-PARÍS, C. (2019). Fate of a naive t cell: a stochastic journey. *Frontiers in immunology*, **10**, 194. 37

DE SERRES, G., ROULEAU, I., HAMELIN, M.E., QUACH, C., SKOWRONSKI, D., FLAMAND, L., BOULIANNE, N., LI, Y., CARBONNEAU, J., BOURGAULT, A.M. *et al.* (2010). Contagious period for pandemic (H1N1) 2009. *Emerging infectious diseases*, **16**, 783. 193

DONALDSON, L.J., RUTTER, P.D., ELLIS, B.M., GREAVES, F.E., MYTTON, O.T., PEBODY, R.G. & YARDLEY, I.E. (2009). Mortality from pandemic A/H1N1 2009 influenza in england: public health surveillance study. *Bmj*, **339**. 194

FRENSING, T. (2015). Defective interfering viruses and their impact on vaccines and viral vectors. *Biotechnology journal*, **10**, 681–689. 151

FRIEDLANDER, A.M. (1986). Macrophages are sensitive to anthrax lethal toxin through an acid-dependent process. *Journal of Biological Chemistry*, **261**, 7123–7126. 80

GILLESPIE, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, **81**, 2340–2361. 23, 36

GILLESPIE, D.T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of chemical physics*, **115**, 1716–1733. 96, 99

GOEL, A.K. (2015). Anthrax: A disease of biowarfare and public health importance. *World Journal of Clinical Cases: WJCC*, **3**, 20. 2

GÓMEZ-CORRAL, A. & LÓPEZ-GARCÍA, M. (2015). Lifetime and reproduction of a marked individual in a two-species competition process. *Applied Mathematics and Computation*, **264**, 223–245. 36

GÓMEZ-CORRAL, A. & LÓPEZ-GARCÍA, M. (2017). On sir epidemic models with generally distributed infectious periods: Number of secondary cases and probability of infection. *International Journal of Biomathematics*, **10**, 1750024. 36

HARDENBROOK, N.J., LIU, S., ZHOU, K., GHOSAL, K., ZHOU, Z.H. & KRANTZ, B.A. (2020). Atomic structures of anthrax toxin protective antigen channels bound to partially unfolded lethal and edema factors. *Nature communications*, **11**, 840. xiii, 5

HERZENBERG, L.A., PARKS, D., SAHAF, B., PEREZ, O., ROEDERER, M. & HERZENBERG, L.A. (2002). The history and future of the fluorescence activated cell sorter and flow cytometry: a view from stanford. *Clinical chemistry*, **48**, 1819–1827. 36

HODGES, D., KLEIN, F., MAHLANDT, B., JONES JR, W., HAINES, B., RHIAN, M. & WALKER, J. (1965). Role of the lymphatics in the pathogenesis of anthrax. *The Journal of infectious diseases*, 481–494. 3

HUANG, E., PILLAI, S.K., BOWER, W.A., HENDRICKS, K.A., GUARNIZO, J.T., HOYLE, J.D., GORMAN, S.E., BOYER, A.E., QUINN, C.P. & MEANEY-DELMAN, D. (2015). Antitoxin treatment of inhalation anthrax: a systematic review. *Health security*, **13**, 365–377. 77

INGLESBY, T.V., O'TOOLE, T., HENDERSON, D.A., BARTLETT, J.G., ASCHER, M.S., EITZEN, E., FRIEDLANDER, A.M., GERBERDING, J., HAUER, J., HUGHES, J. *et al.* (2002). Anthrax as a biological weapon, 2002: updated recommendations for management. *JAMA*, **287**, 2236–2252. 76, 77

JERNIGAN, D.B., RAGHUNATHAN, P.L., BELL, B.P., BRECHNER, R., BRESNITZ, E.A., BUTLER, J.C., CETRON, M., COHEN, M., DOYLE, T., FISCHER, M. *et al.* (2002). Investigation of bioterrorism-related anthrax, united states, 2001: epidemiologic findings. *Emerging infectious diseases*, **8**, 1019. 3

JOHNSON, J., NOWICKI, M.O., LEE, C.H., CHIOCCA, E.A., VIAPIANO, M.S., LAWLER, S.E. & LANNUTTI, J.J. (2009). Quantitative analysis of complex glioma cell migration on electrospun polycaprolactone using time-lapse microscopy. *Tissue Engineering Part C: Methods*, **15**, 531–540. 36

KANG, T.J., FENTON, M.J., WEINER, M.A., HIBBS, S., BASU, S., BAILLIE, L. & CROSS, A.S. (2005). Murine macrophages kill the vegetative form of bacillus anthracis. *Infection and immunity*, **73**, 7495–7501. 3

KIMMEL, M. & AXELROD, D.E. (2002). Branching processes in biology. 36

KOEHLER, T.M. (2009). Bacillus anthracis physiology and genetics. *Molecular aspects of medicine*, **30**, 386–396. 112

KOEHLER, T.M., DAI, Z. & KAUFMAN-YARBRAY, M. (1994). Regulation of the bacillus anthracis protective antigen gene: Co2 and a trans-acting element activate transcription from one of two promoters. *Journal of bacteriology*, **176**, 586–595. 112

KRUTZIK, P.O. & NOLAN, G.P. (2006). Fluorescent cell barcoding in flow cytometry allows high-throughput drug screening and signaling profiling. *Nature methods*, **3**, 361–368. 36

KULKARNI, V.G. (2016). *Modeling and analysis of stochastic systems*. CRC Press. 16

KYPRIANOU, A.E. & PALAU, S. (2018). Extinction properties of multi-type continuous-state branching processes. *Stochastic Processes and their Applications*, **128**, 3466–3489. 36

LEPPLA, S.H. (2000). Anthrax toxin. In *Bacterial protein toxins*, 445–472, Springer. 67

LIU, S., MOAYERI, M. & LEPPLA, S.H. (2014). Anthrax lethal and edema toxins in anthrax pathogenesis. *Trends in microbiology*, **22**, 317–325. 67

LOCKE, M., GREBENNIKOV, D., SAZONOV, I., LÓPEZ-GARCÍA, M., LOGUINOVA, M., MEYERHANS, A., BOCHAROV, G. & MOLINA-PARÍS, C. (2024). Exploring the therapeutic potential of defective interfering particles in reducing the replication of sars-cov-2. *Mathematics*, **12**, 1904. 152, 223, 226

LÓPEZ-GARCÍA, M. (2016). Stochastic descriptors in an sir epidemic model for heterogeneous individuals in small networks. *Mathematical biosciences*, **271**, 42–61. 36, 180

LÓPEZ-GARCÍA, M., NOWICKA, M., BENDTSEN, C., LYTHE, G., PONNAMBALAM, S. & MOLINA-PARÍS, C. (2018). Quantifying the phosphorylation timescales of receptor–ligand complexes: a markovian matrix-analytic approach. *Open biology*, **8**, 180126. 35

MARRIOTT, A.C. & DIMMOCK, N. (2010). Defective interfering viruses and their potential as antiviral agents. *Reviews in medical virology*, **20**, 51–62. 6, 151

MUKHOPADHYAY, T.K. (2008). *Rapid vaccine development using a micro-scale platform*. University of London, University College London (United Kingdom). 134

NAGY, C.F., LEACH, T.S., KING, A. & GUTTENDORF, R. (2018). Safety, pharmacokinetics, and immunogenicity of obiltoxaximab after intramuscular administration to healthy humans. *Clinical pharmacology in drug development*, **7**, 652–660. 149

PASSALACQUA, K.D. & BERGMAN, N.H. (2006). Bacillus anthracis: interactions with the host and establishment of inhalational anthrax. *Future microbiology*, **1**, 397–415. 83

PERRAULT, J. (1981). Origin and replication of defective interfering particles. *Initiation Signals in Viral Gene Expression*, 151–207. 151

PETOSA, C., COLLIER, R.J., KLIMPEL, K.R., LEPPLA, S.H. & LIDDINGTON, R.C. (1997). Crystal structure of the anthrax toxin protective antigen. *Nature*, **385**, 833–838. 68

PFLUGHOEFT, K.J., SWICK, M.C., ENGLER, D.A., YEO, H.J. & KOEHLER, T.M. (2014). Modulation of the bacillus anthracis secretome by the immune inhibitor a1 protease. *Journal of bacteriology*, **196**, 424–435. 117

PITT, M., LITTLE, S., IVINS, B., FELLOWS, P., BARTH, J., HEWETSON, J., GIBBS, P., DERTZBAUGH, M. & FRIEDLANDER, A. (2001). In vitro correlate of immunity in a rabbit model of inhalational anthrax. *Vaccine*, **19**, 4768–4773. 75

PLOTKIN, S. & GRABENSTEIN, J.D. (2008). Countering anthrax: vaccines and immunoglobulins. *Clinical Infectious Diseases*, **46**, 129–136. 76

POPOV, S.G., VILLASMIL, R., BERNARDI, J., GRENE, E., CARDWELL, J., WU, A., ALIBEK, D., BAILEY, C. & ALIBEK, K. (2002). Lethal toxin of bacillus anthracis causes apoptosis of macrophages. *Biochemical and biophysical research communications*, **293**, 349–355. 80

RETA, A. (2017). Defective interfering particles and their role in disease progression and persistence. *Archives of Microbiology & Immunology*, **1**, 65–72. 5, 6

REZELJ, V.V., LEVI, L.I. & VIGNUZZI, M. (2018). The defective component of viral populations. *Current opinion in virology*, **33**, 74–80. 6

SCHMITT, K. & ZACCHIA, N.A. (2012). Total decontamination cost of the anthrax letter attacks. *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, **10**, 98–107. 3

SETLOW, P. (2003). Spore germination. *Current opinion in microbiology*, **6**, 550–556. 117

SHAFAZAND, S., DOYLE, R., RUOSS, S., WEINACKER, A. & RAFFIN, T.A. (1999). Inhalational anthrax: epidemiology, diagnosis, and management. *Chest*, **116**, 1369–1376. 79

SHARMA, S., BHATNAGAR, R. & GAUR, D. (2020). Bacillus anthracis poly-γ-d-glutamate capsule inhibits opsonic phagocytosis by impeding complement activation. *Frontiers in immunology*, **11**, 462. 4

SHIL, P. (2016). Mathematical modeling of viral epidemics: A review. *Biomedical Research Journal*, **3**, 195–215. 152

SIRARD, J.C., MOCK, M. & FOUET, A. (1994). The three bacillus anthracis toxin genes are coordinately regulated by bicarbonate and temperature. *Journal of Bacteriology*, **176**, 5188–5192. 112

SOMPAYRAC, L.M. (2022). *How the immune system works*. John Wiley & Sons. 1

STRAUSS, J.H. & STRAUSS, E.G. (2008). Overview of viruses and virus infection. *Viruses and Human Disease*, 1. 5

SUBRAMANIAN, G.M., CRONIN, P.W., POLEY, G., WEINSTEIN, A., STOUGHTON, S.M., ZHONG, J., OU, Y., ZMUDA, J.F., OSBORN, B.L. & FREIMUTH, W.W. (2005). A phase 1 study of pamab, a fully human monoclonal antibody against bacillus anthracis protective antigen, in healthy volunteers. *Clinical infectious diseases*, **41**, 12–20. 149

SWEENEY, D.A., HICKS, C.W., CUI, X., LI, Y. & EICHACKER, P.Q. (2011). Anthrax infection. *American journal of respiratory and critical care medicine*, **184**, 1333–1341. xiii, 5

THAKUR, A., RESCIGNO, A. & SCHAFER, D. (1973). On the stochastic theory of compartments: Ii. multi-compartment systems. *Bulletin of mathematical biology*, **35**, 263–271. 36

THOMPSON, K.A.S. & YIN, J. (2010). Population dynamics of an rna virus and its defective interfering particles in passage cultures. *Virology journal*, **7**, 1–10. 152

TONI, T., WELCH, D., STRELKOWA, N., IPSEN, A. & STUMPF, M.P. (2009). Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, **6**, 187–202. 27, 29, 120, 147

TSOULARIS, A. & WALLACE, J. (2002). Analysis of logistic growth models. *Mathematical biosciences*, **179**, 21–55. 30

TURNBULL, P.C., FRAWLEY, D.A. & BULL, R.L. (2007). Heat activation/shock temperatures for bacillus anthracis spores and the issue of spore plate counts versus true numbers of spores. *Journal of Microbiological Methods*, **68**, 353–357. 112

TURNER, B.M. & VAN ZANDT, T. (2012). A tutorial on approximate bayesian computation. *Journal of Mathematical Psychology*, **56**, 69–85. 27, 28

VAIDYA, N., MORGAN, M., JONES, T., MILLER, L., LAPIN, S. & SCHWARTZ, E. (2015). Modelling the epidemic spread of an H1N1 influenza outbreak in a rural university town. *Epidemiology & Infection*, **143**, 1610–1620. 193

VASCONCELOS, D., BARNEWALL, R., BABIN, M., HUNT, R., ESTEP, J., NIELSEN, C., CARNES, R. & CARNEY, J. (2003). Pathology of inhalation anthrax in cynomolgus monkeys (macaca fascicularis). *Laboratory Investigation*, **83**, 1201–1209. 75

WESTERA, L., DRYLEWICZ, J., DEN BRABER, I., MUGWAGWA, T., VAN DER MAAS, I., KWAST, L., VOLMAN, T., VAN DE WEG-SCHRIJVER, E.H., BARTHA, I., SPIERENBURG, G. *et al.* (2013). Closing the gap between t-cell life span estimates from stable isotope-labeling studies in mice and humans. *Blood, The Journal of the American Society of Hematology*, **122**, 2205–2212. 36

WILLIAMS, B., LÓPEZ-GARCÍA, M., GILLARD, J.J., LAWS, T.R., LYTHE, G., CARRUTHERS, J., FINNIE, T. & MOLINA-PARÍS, C. (2021). A stochastic intracellular model of anthrax infection with spore germination heterogeneity. *Frontiers in Immunology*, **12**, 688257. 139

WILLIAMS, B.F. (2022). *Mechanistic intracellular and within-host models of bacterial and viral infections*. Ph.D. thesis, University of Leeds. 149

ZAI, X., ZHANG, J., LIU, J., LIU, J., LI, L., YIN, Y., FU, L., XU, J. & CHEN, W. (2016). Quantitative determination of lethal toxin proteins in culture supernatant of human live anthrax vaccine bacillus anthracis a16r. *Toxins*, **8**, 56. xi, xv, xviii, xix, 4, 37, 56, 67, 68, 69, 72, 111, 114, 115, 116, 117, 122, 123, 124, 125, 127, 128, 133, 134, 135, 138, 139, 140, 141, 146, 147, 149, 226