

Master's Thesis

Jeppe Finne Sørensen [gtc329@alumni.ku.dk](mailto:gtc329@alumni.ku.dk)

# Model for Cross Immunity in Malaria Endemics

Supervisor: Kim Sneppen; Niels Bohr Institute

10. of December 2018

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Malaria</b>	<b>4</b>
2.1	Symptoms and details . . . . .	5
2.2	Life cycle of malaria . . . . .	6
2.3	Immune response . . . . .	7
<b>3</b>	<b>Epidemic Models</b>	<b>10</b>
3.1	The SIR model and $R_0$ . . . . .	10
3.2	The SIS model . . . . .	12
3.3	SIR model with vital dynamics . . . . .	14
3.4	Final notes . . . . .	16
<b>4</b>	<b>Malaria Models</b>	<b>17</b>
4.1	Ross' model . . . . .	17
4.2	Final notes . . . . .	21
<b>5</b>	<b>The Malaria Model</b>	<b>22</b>
<b>6</b>	<b>Results</b>	<b>27</b>
6.1	The core components . . . . .	29
6.2	Replacement . . . . .	35
6.3	Strains . . . . .	41
6.4	Mutation . . . . .	47
6.5	Cross immunity . . . . .	53
<b>7</b>	<b>Discussion</b>	<b>61</b>
7.1	Other investigations . . . . .	61
7.1.1	Number of hosts . . . . .	61
7.1.2	Adjusting replacement . . . . .	62
7.1.3	Super-infection . . . . .	62
7.1.4	Initial conditions . . . . .	62
7.1.5	More antigens . . . . .	63
7.2	Data . . . . .	63
7.3	Real-Life applications . . . . .	63
7.4	Model problems and alternative model suggestion . . . . .	64
<b>8</b>	<b>Conclusion</b>	<b>66</b>

<b>9 Acknowledgements and References</b>	<b>68</b>
<b>10 Appendix</b>	<b>73</b>
10.1 Replacement steady state code . . . . .	73
10.2 Various time-series plots . . . . .	74
10.3 Number of data points for switching time . . . . .	78
10.4 Lower replacement plots . . . . .	78
10.5 Plots for systems with three antigens in each strain . . . . .	79

## **Abstract**

Malaria is a mosquito-borne infectious disease affecting humans and animals in mostly tropical and subtropical regions such as central Africa. The disease was responsible for an estimated 445,000 deaths in 2016. Despite costly eradication methods, the disease has been difficult to eliminate. Suppressing transmission through mosquito nets has been one of the most successful control strategies so far. Vaccines has not been successful because of the parasites antigenic diversity. Mathematical and computational models of malaria have provided insights of the disease and have helped selecting optimal control strategies. Some models are complex, taking into account multiple parameters including incubation, super-infections, age, or climate. In this thesis, an agent based stochastic malaria model is presented and examined. The model takes the antigenic diversity of malaria into account. This is achieved by including multiple strains that have antigens in common which results in cross immunity and competition between malaria strains. The model shows that malaria gains a large advantage from it's antigenic diversity. Greater antigenic diversity both increases the number of infected individuals and allowed for malaria to become endemic under suboptimal parameters. Moreover, at certain parameter values the system exhibit bi-stability by switching between dominating strains.

# Chapter 1

## Introduction

Malaria is a parasitic disease prevalent in almost all tropical and humid areas around the globe. Approximately 216 million cases were reported in 2016, most of them in Sub-Saharan Africa [1]. Symptoms include fever, vomiting, tiredness, and headache and may also cause complications such as seizures and coma [2]. In worst case scenarios malaria may result in death. In 2016 malaria had a death toll of an estimated 445,000 [1]. Malaria is so severe in some regions that it is a hindrance to economic growth [3].

Malaria is transmitted during the bites of infected female Anopheles mosquitoes. When mosquitoes bite infected individuals for a blood meal they have a chance to acquire malaria parasites from the blood stream. Infected mosquitoes can then transmit the parasites to healthy individuals [2]. There exist six types of malaria parasites that are able to infect humans which all goes under the name Plasmodium. Of the six, Plasmodium Falciparum and Plasmodium Vivax are the two most prevalent. Falciparum in Africa and Vivax in the rest of the world. Falciparum is the type that is most virulent and causes the most deaths world wide [1, 2].

Prevention methods for malaria includes medicine, mosquito elimination, and the prevention of bites [2, 4]. Most medications, like Mefloquine and Doxycline, decreases the chance of becoming infected. However, such medications are often not practical for natives, due to costs, impracticability, and side-effects from long-term use [5]. A cost-effective and practical prevention method is mosquito nets sprayed with insecticides [4, 6]. Another common prevention method is to reduce the number of Anopheles mosquitoes in local areas by for example spraying areas with insecticides and draining wet areas such as swamps [6]. However, it is difficult and costly to completely eliminate the mosquitoes in many areas. It would for example cost one-fifth of Tanzania's health budget to perform such a task, although it is feasible to do in some areas [7]. It has been possible to entirely eliminate malaria from some parts of the world, which was for example was done in Europe decades ago [8].

Medical treatment, although reducing complications and symptoms is not assured to remove malaria from the victim [2, 6]. The medicine also becomes less useful over time as malaria builds resistance against medication [2].

There does not exist vaccines for malaria that completely protects against it. Plasmodium has great antigenic variety by being able to create new strains with a different set of antigens over a short time period. As vaccines only targets a con-

strained set of antigens, this makes it difficult to build effective vaccines. [9, 10]. Thus, malaria remains one of the most difficult diseases to eliminate [1]. Research is required to gain new insights into new prevention methods and how to effectively use the ones we already have available.

Mathematical and computational models of Malaria can help to decide which control methods will be the most efficient for the money invested. They can also help to predict the severity of malaria. Malaria is an extremely complicated disease. Simplified models help in retrieving the most important parameters that determine the development of malaria, and can give new insights, which would be difficult to retrieve for complicated models [8, 11, 12].

Models often consist of a set of coupled, ordinary differential equations. They describe the rates of which a group of individuals transition to and from states, for example how quickly people become infected. The SIR epidemic model is a well-known one, which can predict how many people would become infected in an epidemic, if parameters can be approximated. Malaria models share much of the same foundation [13, 14].

Sir Ronald Ross was the first to develop a mathematical model of Malaria [11, 15], even before the *SIR* model of epidemics which was established in 1927 [14]. Ross may have drawn inspiration from Bernoulli's mathematical model for small pox [16, 17]. Ross' model is a differential model consisting of two coupled SI models - one for mosquitoes and one for humans. The model builds on the fact that the pathogen is transmitted by mosquitoes [11]. Moreover, he discovered and demonstrated that the parasite of malaria was mosquito-borne, which he received the Nobel Prize for<sup>1</sup> [18]. His mathematical model gave insights in how to effectively combat the malaria menace. Ross' model has since been used as a chassis for future malaria models [13, 15].

It took over 40 years before any extensions was made on Ross' pioneering work. This work was done by Macdonald and led to the Ross-MacDonald malaria models [12, 15]. He made various models with ideas such as super infection and inoculation [13, 15]. Further extensions have been made with considerations such as migration [19, 20], climate [21, 22], and age [23, 24] - a well-known one being the model constructed during the Garki Project [23, 25] [15].

In this thesis, a new malaria model is proposed. The model accounts for the great antigenic variation of malaria, but also the fact that the antigens is shared between strains, which results in cross immunity. Moreover, the model is not solely based on differential equation, but draws from agent based principles. It simplifies on earlier models by completely removing the mosquito vector and instead uses SIR like infection system. The model incorporates the idea of resistance and cross immunity in malaria as a significant part that drive the dynamics. The model is agent based, taking each individuals specific immunities into account. The antigenic variation and cross immunity will be modelled by treating each strain as independent entities, except that gaining resistance to one strain might also provide resistance to other strains.

Accounting for the the antigenic diversity and the human immunity response has been done in other models. Gu et al. [19], Mcqueen [26], Hoshen et al. [27] and

---

<sup>1</sup>However, the first to discover this first between Ross and Grassi is debated. See Cappana and Ernesto [18]

Gupta et al. [28] are two examples of mathematical models considering personal immunity and multiple antigens. The model of this thesis is different in the way shared antigens between strains is modelled [16, 13].

The first chapter of this thesis will explain details about malaria such as it's life cycle, common prevention methods, and the human immunity response. In chapter 3 common disease model will be detailed. They lay a basis for the analysis of malaria models. In chapter 4 an analysis of Ross' original model will be done. In chapter 5 The primary malaria model of this thesis will be explained. The results of investigating the model will then be done in chapter 6. The analysis will mostly be parametric investigations through simulations. But mathematical analysis will be done where applicable. The analysis will start with only the core components of the model, which is only infection and recovery. As the effects of each component is understood, more components will be added to the model. The chapter will end with showing the results of the complete model using the insights gained from the previous sections. Finally, a discussion of the model, improvements, and it's real-life application provide closure this thesis.

# Chapter 2

## Malaria

To make a model about malaria, one should have some understanding of it. In this chapter, a description of malaria is given. The disease in general, its life cycle, and the human immune response will be explained.

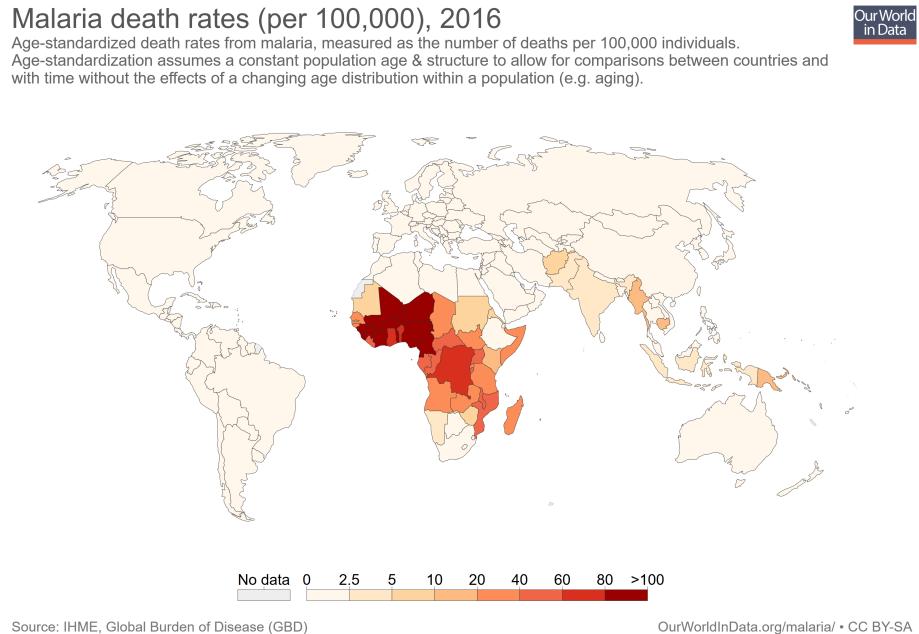
Malaria is one of the most threatening diseases in the world, killing 445,000 people in 2016. 216,000,000 cases of malaria were estimated in the same year [2].

Malaria is caused by parasites that live in the blood of humans and animals. It is transmitted through infected mosquitoes. By biting infected individuals, the mosquito becomes infected and can then transfer the parasite to a bitten creature through the mosquito's saliva [2].

Malaria mostly thrives where its vector, the Anopheles mosquito, also thrives. The mosquito enjoys humid and warm climates [1]. The conditions in the rain forest of central Africa, India and, south east Asia are perfect for the Anopheles mosquito. This is also the regions where malaria is most prevalent as shown in the world map of figure 2.1. The humid regions in northern South America and central America also provide great conditions for the mosquito, but effective control strategies have reduced the impact of malaria in these regions [29].

There exist hundreds of types of malaria parasites, and they are called Plasmodium. Only six types of Plasmodium infect humans [30]. The two most dominating ones are called falciparum and vivax. Falciparum is most prevalent in Africa and is the most lethal one. Vivax mostly operates outside of Africa [2].

In 1894 Ross went out to discover how malaria spread, which, at the time, was believed to be through miasma in the air. Ross discovered that the parasite lived in mosquitoes. He was able to demonstrate that mosquitoes were capable of transmitting the parasite between individuals [18]. This paved the way to properly fight the menace, by giving the idea of inhibiting the vector through for example mosquito nets or insect repellents, which remains some of the most effective control strategies today [4, 18].



**Figure 2.1:** World map of Malaria death rates measured as the number of death per 100,000 individuals in 2016 for each nation. Measured by WHO. Figure is from reference [31].

## 2.1 Symptoms and details

Death is the harshest consequence to strike infected individuals. However, only the most severe infections result in death, or if the infected individual is otherwise vulnerable. Most infections only give flu like symptoms. The first typical symptoms include fever, headache, and chills, which usually starts 10 to 15 days after an infectious bite. Infections may also include complexities such as vomiting, joint pain, seizures, and anaemia depending on the severity of the infection. It can be difficult, especially at the start of an infection, to properly diagnose malaria, as it can be mistaken with other pathogens such as influenza. Blood trials have to be made to be certain, which also makes it a costly affair. [2, 32].

Pregnant women and children are the people that suffer the most from malaria. Pregnant women have a higher chance to contacting cerebral malaria, and greatly increases infant mortality. Children has not yet had time to develop an immune system against malaria. This results in high mortality rates in those population groups, especially if they don't have genetic help like the sickle-cell trait [32, 33].

The infection time of malaria is highly variable. Single malaria infections can in rare cases last up to a decade. While symptoms do not last that long, the parasite may lie dormant in the blood-cells or in the liver. Over time symptoms may resurge. In high-infectious areas, people may have malaria their entire adult lives, but avoid severe symptoms because of their gradual immunity acquisition [34, 32].

Not only does malaria cause hardship or death for the individuals infected, it also hampers economic growth [3]. This in turn makes it harder to prevent malaria,

as the financial foundation is not there to educate local people about the dangers of malaria and to invest in more expensive control strategies [3].

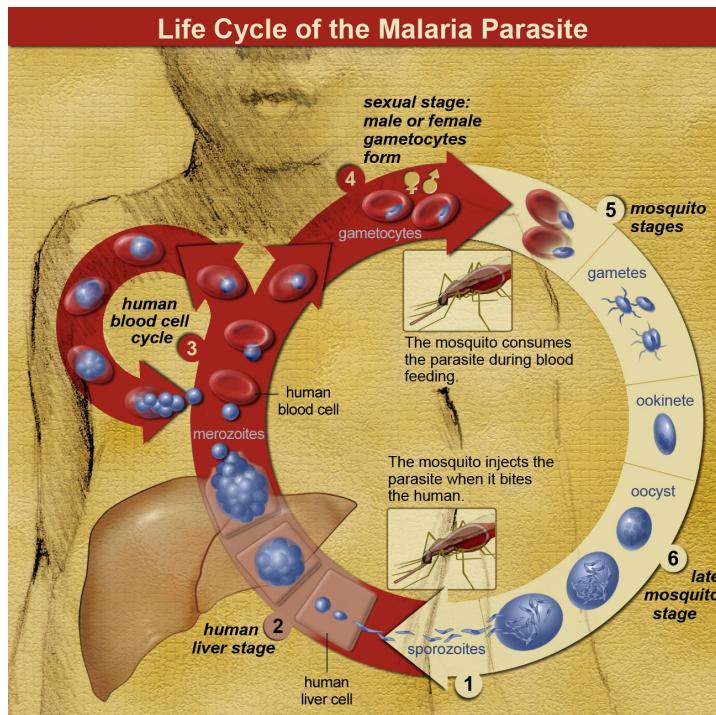
## 2.2 Life cycle of malaria

The malaria parasite, Plasmodium, undergoes a complicated life cycle with many different states - from the mosquito to the human body. This does not go into complete detail as that could be a thesis all of itself. The life cycle of Plasmodium parasites is extremely complex. Each part of the life cycle has its own function, shape, and structure [35].

A Plasmodium parasite can roughly be split into 6 different states - the transmission state, the liver state, human blood cell state, sexual state, and the early and late mosquito states [35]. The state cycle is nicely illustrated in figure 2.2.

The life of Plasmodium in humans starts in the liver as sporozoites after being transmitted from an infected mosquito (2). From there, the parasite develops into the blood stream as merozoites (3). A mosquito will absorb the parasites as gametocytes (4), where the parasite will produce new offspring (5). The offspring is called gametes and will develop into sporozoites (6), and begin the cycle anew when the mosquito transmits the parasite during another blood meal (1) [35, 36].

The main symptoms appear while the parasite is in the bloodstream, taking over individual red blood cells. In the liver, they usually do not harm the human; using it solely to reproduce and readies themselves for an attack on the red blood cells. The parasite reproduces sexually as gametocytes, but reproduces asexually otherwise [36].



**Figure 2.2:** The life cycle of malaria - from man to mosquito. Figure is from reference [37].

## 2.3 Immune response

When a Plasmodium parasite commandeers red blood cells, the immune system will become active and try to combat the pathogen. The immune system has a very difficult time to fight malaria, because malaria has great antigenic diversity which means that the body essentially has to fight a new pathogen every new infection. This also makes it practically impossible to become completely immune to malaria [9].

It takes time for the immune system to detect pathogens<sup>1</sup>. One big part of the adaptive immune system is a specific kind of protein called antibodies or immunoglobulin. Their function is to discover, control and stop pathogens, and to assist in a general immune response. To discover hostile pathogens, antibodies bind to the so-called antigens, which typically are on the surface of the pathogen [38]. Antigens are defined as substances which the antibodies can bind to [39]. A schematic of antibodies and antigens are shown in figure 2.3. When an antibody tries to bind to an antigen and fail, it can "re-assemble" to be able to bind to the pathogens antigens. After doing so that specific antibody are stored in memory and can be reproduced by the body if needed. When antibodies targets an antigen, the corresponding pathogen will be attacked, usually resulting in its death after some time. Antigens are expressed in a variety of ways depending on the pathogen. This immune response is one of the primary ways the immune systems responds to

<sup>1</sup>Pathogen is a micro-organism that can cause disease

a malaria infection and other pathogens like influenza [38, 10].

Vaccination uses this system, by introducing a light virulent strain of the pathogen to the system [38]. This is also true for malaria. Strains is a genetic variant of the pathogen, resulting in, for example, different suite of antigens or different virulence level. This is typically caused by mutation or by swapping genetic components [40]. The body will build antibodies for that pathogenic strain and keep them in memory, effectively making the vaccinated individual immune. If a pathogen species has a great strain diversity with many unique antigens, it becomes difficult to combat a disease with such vaccines. This is the case for malaria, but also for other problematic pathogens such as influenza [2, 9, 38, 10].

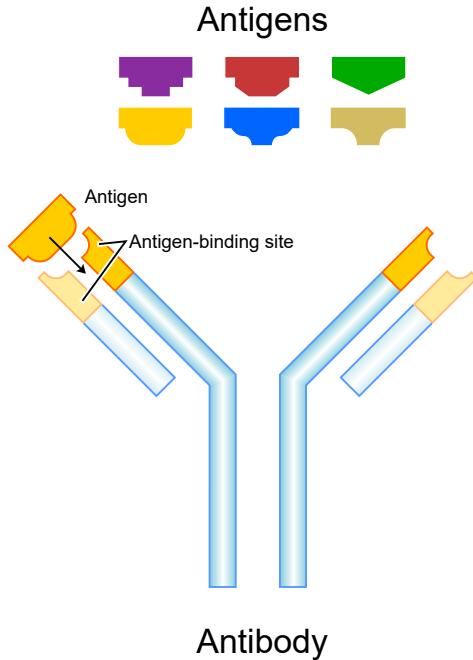
The number of different antigens in one species can be enormous. For Plasmodium, the antigenic proteins is expressed on the surface of infected blood cells, no unlike the picture of an infected blood cell shown in figure 2.4. These are also called the surface features of the strain [9]. A single Plasmodium strain has many antigens, but only expresses one on the surface at a time. The strain can spontaneously shift the antigen expressed on the surface, therefore requiring antibodies for all the antigens to become immune to a specific strain [41].

A big difficulty in the vaccination process of malaria, is the fact that it contains a big antigenic variety among the strains [42]. A cause for this is the group of antigenic proteins in Plasmodium called var genes. The var genes allow spontaneous switching to another surface feature causing the antibodies to no longer be able to bind to the surface protein. Each individual strain only knows a limited amount of antigens. However, the var genes are designed in such a way that new antigens can be created, by combining features of existing var genes, much like the shuffling of a deck of cards. This allows the types of Plasmodium that has these var genes to have a staggering amount of antigenic variety. It has been shown that up to 1/500 parasites undergoes a recombination per life cycle. This results in up to millions of new combinations produced each day in a single infected individual [10, 42, 41].

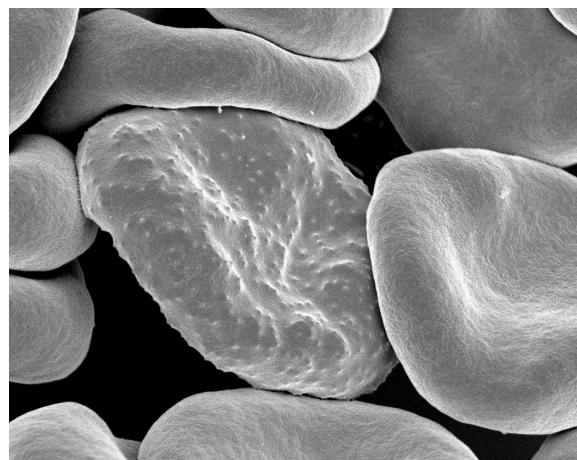
Some strains share some antigens between them. This allows one to use antibodies not only against the particular strain which the body is currently infected by, but potentially also future strain that shares some of the antigens with the current one. This is known as cross immunity [10, 42, 41]. The effects of cross immunity is one of the properties that will be examined through a malaria model in this thesis.

It is possible to be infected by multiple strains. This is called co-infection or super-infection [44]. The consequences of this is a discussed area and is not absolutely known. Whether it is an advantage or disadvantage for malaria is not with certainty known. It is postulated that the fact that the immune system has many more antigens to build antibodies for increases the virulence of malaria. The fighting for domination between strains in the body may circumvent such an advantage. Multiple different strain have been measured in the blood of humans, so super-infection is a known factor [45].

Knowing this, malaria remains a extremely complicated disease with many unknown factors [10]. The model presented in this thesis will can potentially give new insights into the behaviour of malaria.



**Figure 2.3:** Schematic of surface features of antigens and antibodies. The antibody shown as long "sticks" binds to the antigen of a pathogen. If the pathogen is seen as unwelcome the body will activate an immune response to kill the pathogen, which can be done in a variety of ways - sometimes by the antibody itself. If the antibody does not fit to the antigen, it can mutate and transform to find the correct binding proteins. Figure is from reference [43].



**Figure 2.4:** "Electron micrograph of red blood cells infected with *Plasmodium falciparum*, the parasite that causes malaria in humans. During its development, the parasite forms protrusions called 'knobs' on the surface of its host red blood cell which enable it to avoid destruction and cause inflammation. Using scanning electron microscopy, this image shows a knob-rich infected blood cell surrounded by knobless uninfected blood cells." Figure and citation is from reference [46].

# Chapter 3

# Epidemic Models

Epidemics and endemics can be studied using mathematical or computational models. They give insights by for example detailing transition mechanics or estimating the number of infected people at a given time. The fundamental epidemiological models will be presented here, namely SIR, SIS, and SIR with vital dynamics. They make a foundation for the malaria model presented in this thesis.

## 3.1 The SIR model and $R_0$

Probably the most well-known epidemic model is the SIR model. The SIR model is a differential compartment model that describes the rate for which individuals become infected and recover for epidemics. SIR stand for susceptible (denoted  $S$ ), infected (denoted  $I$ ) and recovered (denoted  $R$ ). All individuals in the population are in one of these compartments. Specifically, it is defined by a set of coupled ordinary differential equations that models the rate susceptible individuals become infected and infected individuals recover. The model is basically a number of individuals which changes state with rates based on a logical set of equations [47].

The SIR model is constructed around the idea that new infections happen when an infected individual meets a susceptible individual. This could for example be through sneezing or bodily contact. This means that the rate of new infections is proportional to the probability that infected individuals meet and have contact with susceptible individuals. Recovered individuals can not become infected again as they are considered immune, dead or something equivalent. A schematic is shown in figure 3.1 [47].



**Figure 3.1:** Diagram of the SIR model. Each node is a compartment. Black arrows denote transitions and red arrows denote influence on the transition rate from the origin compartment. Susceptible individuals become infected with a rate proportional to the infection rate  $\alpha$  and the number of infected individuals. Infected individuals recover with a rate of  $\beta$  and transitions to the recovered state.

The system is expressed by three coupled differential equations [14, 47]

$$\frac{dS}{dt} = -\alpha I(t)S(t), \quad (3.1)$$

$$\frac{dI}{dt} = \alpha I(t)S(t) - \beta I(t), \quad (3.2)$$

$$\frac{dR}{dt} = \beta I(t). \quad (3.3)$$

where  $\alpha$  is the parameter governing the infection rate, and  $\beta$  is the parameter governing the rate of recovery.  $\alpha$  is equivalent to how often people meet multiplied by the probability that a transmission happens when infected and susceptible individuals have contact. The population size of the system is assumed to be constant, so the following constraints always hold:

$$S + I + R = N, \quad (3.4)$$

$$\frac{d}{dt}(S + I + R) = 0. \quad (3.5)$$

Often,  $N$  is set to equal 1 so each group is measured as proportions of the total population. This is assumed throughout this thesis and in the above equation system. This equation system will always result in the disease becoming extinct, since there will only be susceptible and recovered individuals left [47].

The model is simple, but can still give relevant information about epidemics in general. An important feature, denoted  $R_0$ , the basic reproduction rate, describes the severity of the disease. It is given by,

$$R_0 = \frac{\alpha}{\beta}. \quad (3.6)$$

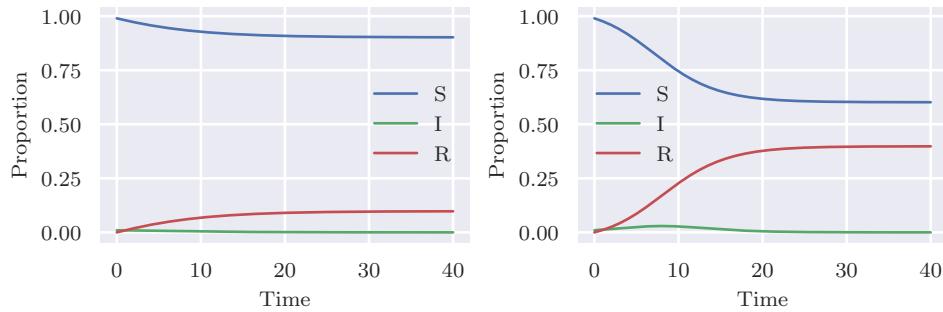
It is a general term used in epidemiology. It is defined as the number of secondary infections arising from a single infectious individual, in an otherwise uninfected population [47].

It is possible to calculate how many will be infected by an epidemic, given initial conditions as a function of  $R_0$ . If initial conditions are  $R(0) = 0$  and  $S(0) \approx 1$  the number of recovered individuals at  $R(\infty)$  will be the total number of infected individuals. These initial conditions are reasonable for new outbreak, as no one has become immune to it, and only a few in the system have become infected.  $S(\infty)$  is the number of individuals who have not been infected after the disease has died out. We divide eq. 3.1 with eq. 3.3

$$\frac{dS(t)}{dt} \frac{1}{\frac{dR(t)}{dt}} = \frac{-\alpha \cdot I(t) \cdot S(t)}{\beta \cdot I(t)} \quad (3.7)$$

The  $S(t)$  and  $R(t)$  terms are isolated

$$\frac{dS(t)}{dt} \frac{1}{S(t)} = -\frac{\alpha}{\beta} \cdot \frac{dR(t)}{dt}. \quad (3.8)$$



**Figure 3.2:** Two simulations of the SIR model. The left plot uses  $\alpha = 0.95$  and the right one uses  $\alpha = 1.25$ .  $\beta = 1.0$  in both cases. Initial values are  $S(0) = 0.99$  and  $I(0) = 0.01$ . They were simulated using a fourth order Runge-Kutta method [48] with  $dt = 0.01$

Integrating both sides, yield

$$\ln S(t) = -\frac{\alpha}{\beta} R(t). \quad (3.9)$$

We insert  $R_0$  and lift the logarithmic term

$$S(t) = \exp(-R_0 R(t)). \quad (3.10)$$

Finally, we set  $t \rightarrow \infty$  and use the fact that  $R(\infty) = 1 - S(\infty)$

$$S(\infty) = \exp(-R_0(1 - S(\infty))). \quad (3.11)$$

In extreme cases  $1 - S(\infty) \approx 1$ , and so we can make the approximation;

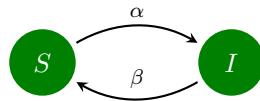
$$S(\infty) \approx \exp(-R_0). \quad (3.12)$$

This equation tells us how many will not have been infected based on the infectiousness of the disease. Simulations of the model are shown in figure 3.2.

Using the SIR model we can get an approximative prediction of how many will be infected during a new epidemic outbreak. This can be used to get an idea of how many people to vaccinate to prevent a serious outbreak. The model of course require one to determine  $\alpha$  and  $\beta$  to get any reasonable results. The parameters can be approximated from data by for example studying previous instances of epidemic incidents of the same kind, or by using the number of new infections at the outbreak of the disease [47, 49].

## 3.2 The SIS model

A variant of the SIR model is the SIS model [50]. As the model name suggests, the model is simplified to only contain two compartments - infected individuals and susceptible individuals. Unlike the SIR model, when an infected individual recovers, they do not gain immunity to the disease. Instead, infected individuals become susceptible again. This system is outlined in figure 3.3.

**Figure 3.3:** Diagram of the SI model.

The system is written by two differential equation,

$$\frac{dS}{dt} = \beta I - \alpha SI, \quad (3.13)$$

$$\frac{dI}{dt} = \alpha SI - \beta I. \quad (3.14)$$

The number of individuals are constant and represent a proportion of the population, so  $N = S + I = 1$ . The disease may not necessarily become extinct unlike the SIR model, as the infected individuals return to the susceptible compartment. There are two equilibrium states in this model. One where the disease becomes extinct and another where a constant proportion of the population is infected. This can be shown by solving the system's ODEs in steady state by setting the rates equal to 0.

$$0 = -\alpha SI + \beta I. \quad (3.15)$$

We use the fact that  $I = S - 1$  and therefore

$$0 = -\alpha I^2 + \beta I. \quad (3.16)$$

The second order equation is solved, and we get:

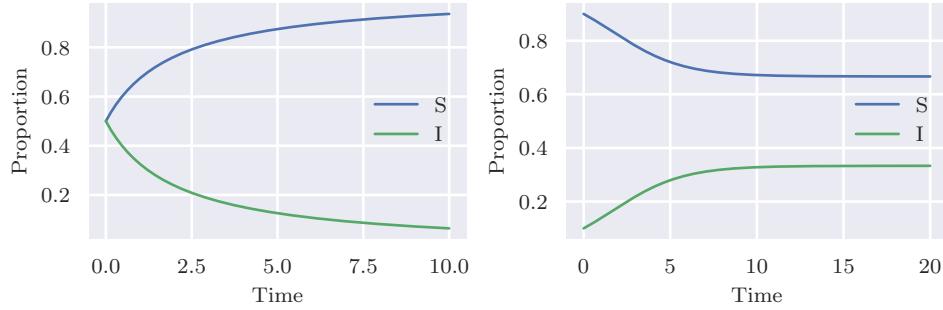
$$I(\infty) = 1 - \frac{\beta}{\alpha} \text{ and} \quad (3.17)$$

$$S(\infty) = \frac{\beta}{\alpha}. \quad (3.18)$$

This is the constant number of infected and susceptible there will be in the system, as long as  $\alpha/\beta > 1$ . If  $\alpha/\beta < 1$  the disease will become extinct, as individuals recover quicker than new individuals can become infected. There are two equilibrium states. One where  $I = 0$  and another where  $I = 1 - \beta/\alpha$ .

The value  $\alpha/\beta$  is also the basic reproductive number  $R_0$ , just as in the SIR model, as that is the number of newly infection a newly infected individual causes, given that all is susceptible. The model is very predictable if the parameters can be determined. Two simulations of the SIS model are shown in figure 3.4

The SIS model characterizes endemic diseases rather than epidemics that the SIR model characterizes. In the SIS model, the disease always survives as long as  $\alpha > \beta$ . This is also true for endemic diseases, such as Malaria. The first Malaria model - Ross' model, described in the next chapter relates very closely to the SIS model. Therefore, it does have more relevance to the particular disease model proposed in this thesis, than the SIR model.



**Figure 3.4:** Two simulations of the SIS model. The left plot uses  $\alpha = 0.95$  and the right one uses  $\alpha = 1.5$ .  $\beta = 1.0$  in both cases. Initial values are  $S(0) = 0.5$  and  $I(0) = 0.5$  for the left plot and  $S(0) = 0.9$  and  $I(0) = 0.1$  in the right plot. The right plot correspond with theoretical values of  $I = 1/3$  according to equation 3.17.

### 3.3 SIR model with vital dynamics

An extension to the SIR model is incorporating a term that makes individuals lose their immunity, this component is also called vital dynamics [50]. This can prevent the disease to become extinct as the susceptible compartment can be refilled.

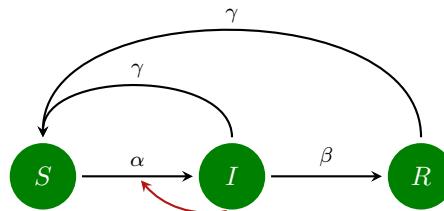
A positive term is added to the susceptible compartment. Individuals namely lose their immunity over time, or new people are put into the system through emigration or birth. The population is still kept constant for simplicity. A diagram of the system is shown in figure 3.5. The changes result in the following new equations [50]

$$\frac{dS}{dt} = -\alpha I(t)S(t) + \gamma(I(t) + R(t)), \quad (3.19)$$

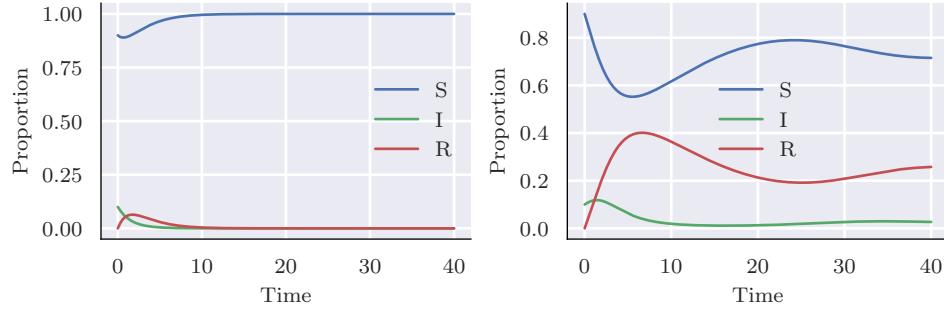
$$\frac{dI}{dt} = \alpha I(t)S(t) - \beta I(t) - \gamma I(t), \quad (3.20)$$

$$\frac{dR}{dt} = \beta I(t) - \gamma R(t), \quad (3.21)$$

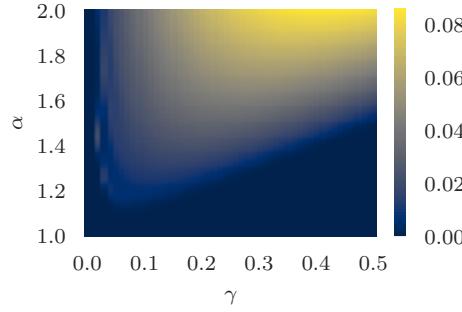
where  $\gamma$  is the parameter that adjusts the death rate. Two simulations of the model are shown in figure 3.6. In both cases, the system ends in a steady state, one where the disease dies and one where it does not. The additional term allow for the disease



**Figure 3.5:** Diagram of the SIR model with vital dynamics. The only difference from the SIR model is the fact that individuals can become replaced. This is represented by individuals in the  $I$  and  $R$  compartments transition into the  $S$  compartment with a rate proportional to  $R$ .



**Figure 3.6:** Two simulations of the SIR model with vital dynamics. The left plot uses  $\alpha = 0.95$  and the right one uses  $\alpha = 1.5$ .  $\beta = 1.0$  and  $\gamma = 0.1$  in both cases. Initial values are  $S(0) = 0.95$  and  $I(0) = 0.05$ .



**Figure 3.7:** 2D plot showing the proportion infected as a function of  $\alpha$  and  $\gamma$ . If the proportion is 0 then the disease die out.

to become endemic. The system show asymptotic stability when endemic, slowly moving towards the stable point as shown in figure 3.6 [50].

An interesting question this model poses, is at which parameters does the model become endemic? A parameter search can be done to answer this question. Simulations are done over a range of parameters, and through that we are able to identify the parameter range where the disease is endemic. This method will be used throughout the thesis. A parameter sweep is shown over  $\alpha$  and  $\gamma$  in figure 3.7 showing the proportion infected. A triangle appears where the proportion infected is above zero. Essentially, there exist an area where the disease is endemic and the system is asymptotically stable for all initial values (except  $I = 0$ ). Another insight is the fact that there exist a  $\gamma$  value that maximises the proportion infected. This specific  $\gamma$  value is not constant, but grows as a function of  $\alpha$ . Using steady state, the exact optimal  $\gamma$  can be calculated analytically - this will coincidentally be done later for the malaria model of this thesis. The triangle of survival can also be calculated.

Overall, the SIR with vital dynamics model simulates endemic diseases. But the replacement term complicate the model by introducing a new parameter which is not simple to predict the affects of.

### 3.4 Final notes

There exist many epidemic models that deal with different kind of diseases and take into account complications such as incubation, geography, or climate [47, 50]. We will now move on to describe malaria models specifically. They share many similarities with these basic epidemic models, but have some key differences. Three fundamental epidemic models have been discussed here, and as we shall see the model introduced in this thesis do share some similarities with them as well. Before the model of this thesis will be introduced, a short resume of malaria models will first be done.

# Chapter 4

## Malaria Models

The first malaria model was made by Sir Ronald Ross in 1914 [18]. Today, his model remains an important inspiration for malaria models. In this chapter, I will go through his malaria model and touch upon a few words on other relevant models. As we will see, the malaria model presented in this thesis has many difference from Ross' model.

### 4.1 Ross' model

Ross' model is based on the fact that mosquitoes transmitted the disease between humans. Therefore, both individuals and mosquitoes are incorporated. The model is a compartment based differential model. It is essentially equivalent to a double class SI model. In figure 4.1 a diagram of the model is shown. There are four total compartments in his model: Two for infected and susceptible humans denoted  $I_h$  and  $S_h$ , respectively and two for infected and susceptible mosquitoes denoted  $I_m$  and  $S_m$ . Infected mosquitoes transmit malaria to healthy individuals through bites and healthy mosquitoes become infected when biting an infected human. As infected individuals recover, they return to the susceptible compartment. The same goes for mosquitoes, although they are expected to die before they actually recover from malaria [11, 13, 51].

The system can be expressed by 4 differential equations [13, 51],

$$\frac{dS_h}{dt} = rI_h - abmI_mS_h, \quad (4.1)$$

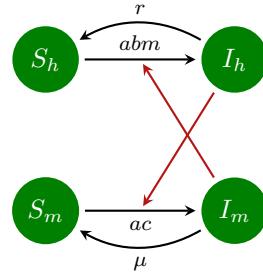
$$\frac{dS_m}{dt} = \mu I_m - acI_hS_m, \quad (4.2)$$

$$\frac{dI_h}{dt} = abmI_mS_h - rI_h, \quad (4.3)$$

$$\frac{dI_m}{dt} = acI_hS_m - \mu I_m. \quad (4.4)$$

A list of parameters and their definitions are shown in table 4.1.

To gain insights from the model, we find the reproduction number and perform steady state analysis. This model is stable with parameters for which malaria is



**Figure 4.1:** Diagram of Ross' model. Each node is a compartment for individuals or mosquitoes. Black lines represent transitions from one compartment to another, while red lines represent increased rate of transitioning proportional to the number of individuals in origin compartment.  $S_h$  denotes susceptible individuals,  $S_m$  denotes susceptible mosquitoes,  $I_h$  denotes infected individuals and  $I_m$  denotes infected mosquitoes. Individuals becomes bitten with rate  $a$ . Susceptible individuals becomes infected from a bite with probability  $b$ . Susceptible mosquitoes become infected from a bite against an infected individual with probability  $c$ . Individuals recover and return to the susceptible state with rate  $r$ . Mosquitoes die with rate  $\mu$ . The more mosquitoes there are  $m$ , the more bites an individual will receive a day. The more infected mosquitoes there are in the system, the larger the rate from the susceptible compartment to the infected compartment, as there are more infected mosquitoes to spread the disease.

endemic. To find the number of infected individuals, we will now determine the steady state solution.

The four equations can be condensed to only two equations since  $1 = S_h + I_h$  and  $1 = S_m + I_m$ . Using this property and setting the differential equations equal to 0 gives us the steady state equations

$$0 = abmI_m(1 - I_h) - rI_h, \quad (4.5)$$

$$0 = acI_h(1 - I_m) - \mu I_m. \quad (4.6)$$

Solving for  $I_h$  and  $I_m$  yields

$$I_h = \frac{\mu I_m}{ac(1 - I_m)}, \quad (4.7)$$

$$I_m = \frac{rI_h}{abm(1 - I_h)}. \quad (4.8)$$

Then, inserting  $I_h$  into the other equation

$$0 = r\mu - a^2bcm(1 - I_h) \left( 1 - \frac{rI_h}{abm(1 - I_h)} \right). \quad (4.9)$$

This reduces to a second order equation which has the solution

$$I_h = \frac{a^2bcm - r\mu}{a^2bcm + arc}. \quad (4.10)$$

This is the equation for the proportion of infected individuals at steady state. Increasing parameters  $a, b, c, m$  will make  $I_h$  converge towards 1, while increasing  $r$  and  $\mu$  will decrease the number of infected individuals. It is also possible for malaria

Symbol	Parameter	Standard Values
$a$	Man biting rate	0.01 to 0.5 day <sup>-1</sup>
$b$	Proportion of bites that produce infection in humans	0.02 to 0.5
$m$	Ratio of number of female mosquitoes to that of humans	0.5 to 40
$r$	Average recovery rate of humans	0.005 to 0.05 day <sup>-1</sup>
$c$	Proportion of bites by which one susceptible mosquito becomes infected	0.5
$\mu_2$	Per capita rate of mosquito mortality	0.05 to 0.5 day <sup>-1</sup>

**Table 4.1:** Parameters in Ross' model. Standard values and parameter descriptions are from Mandal et al. [13]

to die out if  $a^2bcm < r\mu$ , as that would decrease  $I_h$  to below 0. As  $a$  is squared in the equality, it is the most dominating parameter.

Two simulations of the system are shown in 4.2. In the left figure malaria becomes extinct since  $a^2bcm < r\mu$ . In the right figure,  $a$  is increased slightly from  $a = 0.05$  to  $a = 0.1$ . It reaches a steady state with 40% of the population being infected. The values retrieved in the simulations agree with the theoretical values 0 and 0.4. The theoretical values are retrieved by inserting the values given in the figure caption into equation 4.10).

Ross' model gives insight into how many individuals you can expect to be infected by malaria in an isolated region, as long as it is possible to approximate the parameters. Another way to approach the problem is to find the reproductive number  $R_0$  which is also tied to the steady state results.

The reproduction number can be calculated using the next generation method, which works well for models with multiple compartments [49, 52]. This method uses the dominant eigenvalue of the *next generation matrix*  $FV^{-1}$  to produce  $R_0$ .  $F$  and  $V$  are matrices of partial derivatives. The entries of  $FV^{-1}$  correspond to rates of new infections produced in the other compartment, multiplied by the time spend in that compartment [49, 52]. The two matrices are defined as

$$F = \frac{\partial F_i}{\partial x_j} \quad \text{and} \quad V = \frac{\partial V_i}{\partial x_j}, \quad (4.11)$$

where  $i$  and  $j$  go from 1 up to the number of compartments in the model.  $F_i$  is the rate of new appearances in compartment  $i$ , and  $V$  is the rate of transfer out of compartment  $i$ .  $x_i$  is the proportion of individuals in that compartment [49]. In Ross' model  $x_i$  represent the compartments of  $I_h$  and  $I_m$ .

Following equation 4.11 one finds that

$$F_{ij} = \begin{bmatrix} 0 & abm \\ ac & 0 \end{bmatrix} \quad \text{and} \quad V_{ij} = \begin{bmatrix} r & 0 \\ 0 & \mu \end{bmatrix}. \quad (4.12)$$

The upper right matrix element in  $F$  is the rate for which  $S_h$  transitions to  $I_h$ . The lower left matrix element is the rate for which  $S_m$  transitions to  $I_m$ . The upper left matrix element in  $V$  is the rate for which  $I_h$  recovers, and the lower right matrix element is the rate for which  $I_m$  dies.

The next generation matrix is then

$$FV^{-1} = \begin{bmatrix} 0 & abm \\ ac & 0 \end{bmatrix} \begin{bmatrix} 1/r & 0 \\ 0 & 1/\mu \end{bmatrix} = \begin{bmatrix} 0 & abm/\mu \\ 0 & ac/r \end{bmatrix}. \quad (4.13)$$

The dominant eigenvalue (spectral radius) is then:

$$\lambda = a \sqrt{\frac{bmc}{\mu r}}. \quad (4.14)$$

The dominant eigenvalue is related to  $R_0$  by describing the transitions into and out from the infected compartments.  $a$ ,  $b$ ,  $c$ , and  $m$  are parameters that increases the infectiousness of malaria, while  $\mu$  and  $r$  are parameters that decreases it.

There is essentially two ways of defining  $R_0$ . One defines the  $R_0$  by the number of infectives in the same class produced by a single infective of that class. The other definition gives the mean number of new infectives per infective in any class. The latter definition therefore depends on the number of classes in the system. [49]. The reproductive rate using the latter definition is

$$R_0 = \lambda^2 = \frac{a^2 bmc}{r\mu}. \quad (4.15)$$

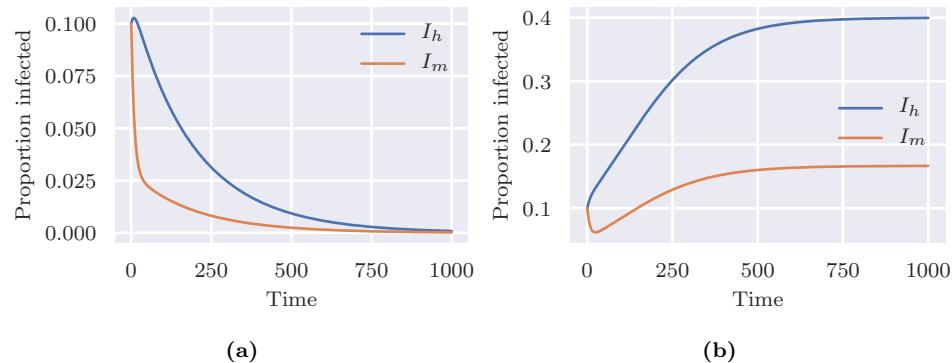
The reproduction number can identify when malaria is endemic and when it is not. The reproduction number corresponds with malaria becoming extinct for  $R_0 < 1$  and becoming endemic for  $R_0 > 1$ . Two simulations of Ross' model exemplify the use of  $R_0$  in figure 4.2.

$$(a) R_0 = \frac{0.05^2 \cdot 0.2 \cdot 0.5 \cdot 2.0}{0.1 \cdot 0.01} = 0.5, \quad (4.16)$$

$$(b) R_0 = \frac{0.1^2 \cdot 0.2 \cdot 0.5 \cdot 2.0}{0.1 \cdot 0.01} = 2.0. \quad (4.17)$$

The reproduction number of Ross' model can be used to identify how much a parameter can be adjusted for malaria to become extinct, and which parameters have the most weight. An important insight gained from eq 4.15, is that  $R_0$  increases as the biting rate  $a$  squared. This makes the biting rate the most impactful parameter in Ross' model [49]. For example, if the biting rate was known to be reduced by half, by for example introducing mosquito nets,  $R_0$  will be reduced by one fourth. The steady state value of  $I_h$  in equation 4.10 shows the direct change to when the parameters are adjusted. These insights gave Ross the idea that the most efficient control strategies would reduce the biting rate  $a$ . Mosquito nets was one such method, and is still one of the most used prevention methods today. [18, 13]. Ross' model is a great example of how mathematical models can give insight in diseases and help decide on efficient control strategies.

Ross' model is essentially a SIS with an additional class. The two models have many of the same properties. They are both stable systems thereby both simulating an endemic with a constant number of infected. The model both have parametric regions where stability is acquired.



**Figure 4.2:** Two simulations of Ross' model using  $b = 0.2$ ,  $m = 2.0$ ,  $r = 0.01$ ,  $c = 0.5$ , and  $\mu = 0.1$ .  $a = 0.05$  in figure (a) and  $a = 0.1$  for (b) (The model values used are from [47]). The proportion of infected individuals and infected mosquitoes are 0.1. In both cases the system ends in a steady state. In figure (a)  $R_0$  is smaller than 1 and therefore become extinct, while it becomes an endemic in (b) as  $R_0 = 1$ . The steady state values correspond with the theoretical values 0 and 0.4 respectively, according to equation 4.10.

## 4.2 Final notes

There have since been proposed many different malaria models, as Ross' model does not include many complications. It took over 40 years before a new malaria model was proposed. The first extension was done by MacDonald [15], who introduced incubation and super-infection. Newer models may account for factors such as geography, climate, or age [13]. Most of these models still maintain the differential compartment structure. Moreover, they typically also model the mosquito vector directly [13].

## Chapter 5

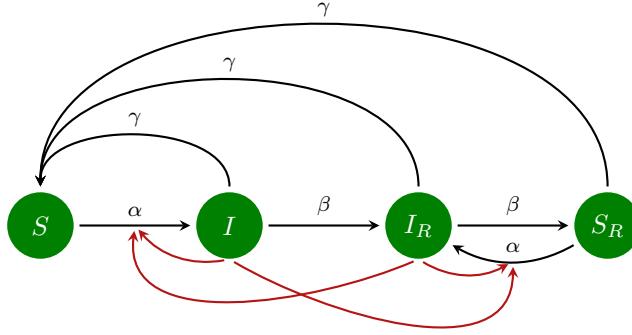
# The Malaria Model

In this chapter the malaria model of this thesis will be explained. The model will be explained such that the explanation starts with the core components. The greater complexities, which allows for cross immunity, will be explained thereafter. Lastly, notes on how this model is different and justifications for these differences will be presented.

This model is an agent based model with heavy inspiration from the SIR model. The system is not expressed as a set of differential equations. It is instead defined by a set of rules of interaction much like Conway's game of life [53], although this model is non-dimensional. The agents in the system are human individuals in an isolated region. These individuals are able to host the malaria parasite, and will henceforth be denoted as hosts. The hosts are assumed to be in their own independent system that interact little with the outside world. The number of hosts in the system is constant for simplicity reasons.

Hosts can be in a number of different states. These states are the susceptible state, the infected state, the resistant state, and the infected and resistant state. An overview of the possible states and the hosts transitions between them are given in figure 5.1. The susceptible state ( $S$ ) represents individuals that can easily be infected by malaria, and has no defence system against it. The infected state ( $I$ ) represents individuals that are infected by malaria. The resistant state ( $S_R$ ) represents individuals that have build defences by having developed antibodies against malaria. However, they can still become infected. The infected and resistant state ( $I_R$ ) represents individuals who are infected by malaria but have also developed an immune system against it.

Computationally, the system develops through single events that are chosen at each iteration. Events are defined by a change in the system, such as a susceptible host becoming infected and going from the susceptible state to the infected state. All events are in essence hosts transitioning from one state to another. There exist 4 possible events; infection, resistance, replacement, and mutation. At every iteration, one of these events are selected randomly, and the probability to be selected is proportional to their rate (Gillespie [54]). Each event's rate is calculated at each iteration, and is defined by an equation that changes as the system updates. The rates are written in table 5.1. Mathematically, the probability of an event to



**Figure 5.1:** A diagram of this thesis' malaria model, showing the transitions individuals' undergoes for a single strain with only a single antigen. Each node (green circle) represent a compartment an individual can be in. An individual can be susceptible ( $S$ ), resistant and susceptible ( $S_R$ ), infected ( $I$ ), or infected and resistant ( $I_R$ ). The black lines represent transition from one compartment to another. I.e. susceptible individuals becoming infected and transitioning from the susceptible compartment to the infected one. The red lines represents increase in the rate of a transition at the transition the red line is pointing towards, proportional to the number of individuals in the compartment of the red line's sources. For example, the more infected individuals there are, the more likely a single susceptible individual is to become infected. Each letter above the lines is parameters that adjusts the probabilities of the transition events, as shown in table 5.2.

be selected is

$$p(E_i) = \frac{R(E_i)}{\sum_{i=1}^4 R(E_i)}, \quad (5.1)$$

where  $E_i$  is a vector consisting of four elements, with each element being a single event, and  $p(E_i)$  are the probabilities and  $R(E_i)$  are the rates of each event.

The infection event causes a randomly selected susceptible or resistant host to become infected, with a random strain chosen in proportion with the number of infected host each strain infects. The rate of infection is governed by the parameter  $\alpha$ . It represents a combination of the biting rate of mosquitoes, the probability of getting infected on a bite, and the density of mosquitoes<sup>1</sup>. The rate equation is taken from the SIR model, which means that there are no mosquito vector. This means that the rate of this event is proportional to  $(I + I_R)(S + R)$ . But, because of super-infection, the rate is closer to  $I + I_R$ . This is because that if the selected host is already infected by the randomly selected strain, a new event will be drawn instead. Also note that you can still become infected even if you have resistance to the particular strain.

The resistance event allows host to gain resistance to malaria through the building of antibodies. The probability of this increases proportionally the the total number of infected individuals. When the event happens an infected host is randomly selected. Two things can then happen. If the host does not have all the required antibodies it gains a relevant antibody. If the host already has all the necessary antibodies, it recovers. When an antibody is gained, that host remem-

<sup>1</sup>somewhat equivalent to  $a$ ,  $b$ ,  $c$ , and  $m$  in Ross' model and  $\alpha$  in the SIR model

Event	Description	Rate
<i>Infection</i>	A random host becomes infected with a strain from a random infected host	$\approx \alpha(I + I_R)$
<i>Resistance</i>	A random infected host gains an antibody for a relevant antigen. If it already has antibodies for the specific strain, it removes the infection and becomes resistant ( $S_R$ )	$\beta(I + I_R)$
<i>Replacement</i>	A random host loses its memory and infection	$\gamma$
<i>Mutation</i>	A random strains changes to another different random strain	$\mu(I + I_R)$

**Table 5.1:** All possible events for the malaria model. The state notations represents the proportion of the hosts in that state. The parameters governing these rates are given in table 5.2. All state values are in proportion to the total population. For infection, hosts cannot get infected by a strain it already is infected by which explains the  $\approx$  sign. For example, if there is only one strain in the system, the rate essentially become  $\alpha(I + I_R)(S + S_R)$ . See the text for more information.

Symbol	Parameter	Values used
$N$	Number of hosts	10000
$\alpha$	Infection rate	0.25 to 1.05
$\beta$	Resistance gain rate	1.0
$\gamma$	Replacement rate	See section 6.2
$\mu$	Mutation rate	$10^{-4}$
$S$	Number of strains in the system	1 to 6
$L$	Antigens in each strain	1 to 3

**Table 5.2:** Parameters for the malaria model. The number of hosts is considered constant but is written here for reference.

bers the antibody until it becomes replaced through the replacement event. A consequence of this system is the fact that hosts with no antibodies will have to be selected multiple times through this event to completely recover. This means, that resistant hosts recover at double the rate compared to susceptible hosts, if one antibody is required to become fully resistant to a malaria strain. The rate of this event is proportional with the parameter  $\beta$ . The parameter represents how quickly individuals build antibodies and recover from a malaria infection.

The replacement event kills a random host and introduces a new one to the system. This is done by removing all infections and all antibodies of a randomly selected host. The rate increases proportionally with the parameter  $\gamma$ .

The mutation event changes one randomly selected strain and changes it to another random strain. The rate of this event is governed by  $\mu$  and is proportional to the number of infected hosts. The event represents mutation in malaria or strains getting reintroduced by mosquitoes from other areas.

All the parameters are shown in table 5.2 as an overview.

Now that the core components has been explained, we begin to define exactly how strains operate in this model. Introducing multiple strain make the model more complex, compared to just having a single strain. One of the goals is to have

a model, which is capable of simulating cross immunity. That is gaining resistance to one strain also gains partial resistance against other strains. Multiple strains is therefore a vital part of the model.

To be able to identify individual strains each strain is given an unique number. When an infection event happens, a random host is chosen to become infected. The strain which will be transmitted is also randomly chosen with a probability proportional to the number of hosts each strain infects. It is not possible for a host to be infected by the same strain multiple times. If this happens, a new randomly chosen event occurs, as explained earlier. It is possible for hosts to be infected by multiple different strain, this is called super-infection or co-infection<sup>2</sup>.

When a host is chosen for the resistance event, and it doesn't have any antibodies against a strain it is infected by, it gets the number that identifies the strain as a new antibody. The host will then be considered resistant to that strain. If the chosen host already is resistant, it recovers, and the specific strain is removed from the host. The host keeps the antibody indefinitely or until replaced. When hosts are infected by multiple strains, a random strain is chosen to gain resistance against. super-infection therefore prolongs the infectious state.

To introduce the possibility of cross immunity, the strains must have some shared properties. This is done by defining each strain by a set of numbers instead of a single one. Each number will then represent one antigen. The number of antigens that define each strain is called the number of antigens in each strain. A specific set of antigens then identifies a specific stain, and exactly that set of antigens is not present in any other strain. However, numbers can be the same between some strains, so gaining antibodies against one strain can therefore also give antibodies against a related strain. When these numbers or antigens are shared between strains, cross immunity can occur. Hosts has to have antibodies against all the antigens in a single strain to be able to recover from it. Hosts only gain a single antibody from a single resistance event. This means, that if a strain is defined by two antigens, completely susceptible hosts will need to be chosen thrice to completely recover from an infection.

For example, consider a system where each strain is defined by exactly two antigens, which is to say that there are two antigens in each strain or  $L = 2$  as shown in table 5.2. If there are two strains in the system, and we would like cross immunity to be able to occur, a possible system of strains could be the following:

$$\text{strain 1} = [1, 2]; \quad \text{strain 2} = [2, 3], \tag{5.2}$$

where each number represents an antigen. As seen, antigen number 2 is shared between them while the strains still have an unique set of antigens. This would not be possible if each strain only had one antigen. Then, when hosts gain antibodies for the antigens in one strain, they also gain antibodies for one of the antigens in the other strain. This way, hosts already recovered from one strain, only needs to spend 2/3 of time recovering from the other strain. The number of antigens and the number of strains can be any value - higher values increases the complexity, as it increases the number of possible combinations of strains and antigens.

This is how strains and the immune system are represented in this model. This system is meant to be an approximation which can model the diversity of malaria

---

<sup>2</sup>Other malaria models also have this feature, see Macdonald [55] for example, and is known to occur in real life as detailed in section 2

strains and antigens, while giving insights into what changes it makes to the system if some of the antigens are mutually shared between strains. Some of the complications of this system will be discussed in the following chapters.

The model proposed here takes a different approach than other malaria models, such as the ones touched upon in chapter 4. Three main points make it distinct.

Firstly, it discontinues the idea of modelling the mosquito vector. As we saw in the section about Ross' model, it was found that the model shares properties with the SIS model described in section 3.2. The model here is also more complex than Ross' original model, because of the introduction of multiple strains. Removing a whole class simplifies the model without costing too much in the dynamical aspect.

Secondly, the model is a stochastic agent based model constructed around a Gillespie algorithm [54]. Most malaria models are constructed as differential compartment models [13, 16]. To be able to simulate complex malaria combinations, it was found that the required equation system would become too complex. An agent based approach gives more flexibility. However, it comes with the downside that it is more difficult to analyse the model mathematically.

Thirdly, but most importantly, the model provides a way to simulate unique strains with shared antigens, which is one the main points of investigation. Because of the great antigenic diversity in malaria, as explained in chapter 2, it can become very complex to design a model that can represents such diversity. This is also the fundamental reason for the two above points. Taking within-host dynamics into account has been done before by for example Gupta et. al. [28], MCqueen et. al. [26], and Hoshen et. al. [27]. The model of this thesis has some important distinctions from those approaches, which are summed up by the two above points, and by the fact that this model can simulate a great variety of malaria complexity while having relatively simple rules [13, 56].

Now that the basis of the model has been explained the results from analysing and simulating the model will be presented.

# Chapter 6

## Results

The main purpose of this thesis is to investigate, analyse, and discuss the malaria model described in the previous chapter. The primary interest of the model is finding the ramifications of having multiple strains of malaria which share antigens. In this chapter, I present the results and insights obtained from the model. Analysis, explanations, and discussions will also be presented while going throughout the chapter.

The model will be investigated over a range of parameters. To identify the impact of each parameter, a very simplified version of the model is first explored. All complex components are taken out of the system, so only the bare foundation is left. After the simplified model is explored, new model components are added one at a time. Each component addition will be compared to previous instances. That way, it is more simple to test how each component impacts the system, more simple to analyse, and avoids a complete search over the whole parameter space. In general, it makes it more straightforward to get an overview of the model.

The model will mostly be investigated stochastically. In some instances it is possible to make deterministic simulations, that allows for easier analysis. They will also be compared to each other, to get an idea of what differences stochasticity makes. Moreover, mathematical analysis will be applied when possible which will include steady state analysis.

A term used to gauge the competitiveness under certain parameters will be used is  $R_0$ , which was also explained in chapter 3. Throughout this chapter, we define a base  $R_0$  as

$$\text{base } R_0 = \frac{\alpha}{\beta} \cdot \text{Number of resistance events to recover} \quad (6.1)$$

The number of resistance events to recover is typically 2 or 3 and depends on the number antigens representing each strain. This  $R_0$  gives an overall idea of the infectiousness with a particular set of parameters.

Another measure that will be used throughout the chapter is the unit of time called generation (gen). Since the simulations are done using arbitrary time units, a different measure of time is the number of iterations done in a simulation. The iterations can get extremely large, and does not scale with the number of hosts in

the system. Therefore, the generation unit is used, which is defined as the number of iterations divided by the number of hosts in the system.

The initial values used throughout this chapter will be the same unless otherwise noted. These are  $N = 10.000$  with  $S = 9500$  and  $I = 500$ . Moreover, the number of infected and other compartments will always be written as a proportion of the total population.

The first part of this chapter will deal with the model stripped of all its components, except the most fundamental ones, basically turning the model into something akin to a SIS model. It will only include the infection and resistance events, as those are the two components necessary to have a disease model. Replacement is the next component to be added which will allow resistant hosts to lose their resistance. The first complex part introduces many strains with a single antigen into the system. The model will be investigated over a range of different number of strains. There after the mutation component will be examined. Lastly, the number of antigens each strain can contain will be adjusted, finally opening the potential for cross immunity. This is the part which is must difficult to analyse, as there is many ways to adjust the combination of strains with many antigens. This will be the last section of this chapter.

## 6.1 The core components

In this section, the results and analysis of the model with only the core component are presented. This means that all non essential parts of the model are removed. We do this to get a better understanding of the basic parts of the model, without any noise from other parts. The core components for having a working model is the infection and resistance components. With those, we have a system where individuals is able to become infected and able to recover. A diagram of this system is visualised in figure 6.1.

With this simplified model there exist four compartments, which is to say that it there exist four states a host can be in. The susceptible state  $S$  where the host can become infected, the infectious state  $I$  where malaria may spread from, the infectious and resistant state  $I_R$ , where malaria can still spread from, but a host recover faster, and the resistant state  $S_R$  where a host can become infected again, but still remains resistant, so it recover quicker compared to a non resistant host.

The model is simplified by setting most of the parameters described in table 5.2 to the lowest possible values. This means that  $\gamma = 0$ ,  $\mu = 0$ ,  $A = 1$ , and  $S = 1$ . It has the consequence that it is not possible for the replacement event or the mutation event to occur. Moreover, there exist only one strain of malaria in the system.

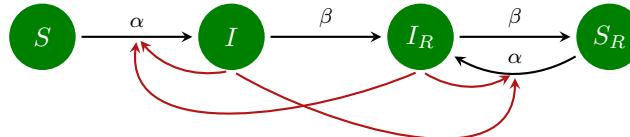
Before we examine the stochastic simulations of the model using the methods described in chapter 5, the model is first analysed deterministically and analytically. In this simple form the model is very close to a combination of the SIR and SIS models. It is also possible to write this system in differential form with 4 compartments, which allows for mathematical analysis as done in earlier chapters. The model with core component can be defined with the following ODEs:

$$\frac{dS}{dt} = -\alpha(I(t) + I_R(t)) \cdot S(t), \quad (6.2)$$

$$\frac{dI}{dt} = \alpha(I(t) + I_R(t)) \cdot S(t) - \beta I(t), \quad (6.3)$$

$$\frac{dI_R}{dt} = \alpha(I(t) + I_R(t)) \cdot S_R(t) + \beta I(t) - \beta I_R(t), \quad (6.4)$$

$$\frac{dS_R}{dt} = \alpha(I(t) + I_R(t)) \cdot S_R(t) - \beta I_R(t). \quad (6.5)$$



**Figure 6.1:** A diagram of the simplified model containing only one malaria strain, and no mutation and replacement components. A susceptible host ( $S$ ) will transition to the infected state ( $I$ ), and from that to the infected and resistant state ( $I_R$ ). It ends at the susceptible and resistant state ( $S_R$ ). The host will thereafter transition between the infected and resistant state, as it is not possible for a host to lose its resistance. The more infected hosts there are, the more probable it is for a susceptible host to become infected.

The coupled ordinary differential equations describes the development of each of the four compartments shown in figure 6.1. The parameters are the same as the ones shown in table 5.2. It is possible to convert the system to a set of differential equations as rates is proportional with probabilities [54]. The constraint  $S + I + I_R + S_R = N = 1$  applies. Also, recall the fact from chapter 5 that an infected host cannot become super infected with the same strain, which explains the rate of infection is proportional to  $(I + I_R)(S + S_R)$  and not  $(I + I_R)$ .

Simplifications can be done to the equation system. The only difference between the susceptible ( $S$ ) and the resistant ( $S_R$ ) hosts, is the fact that the resistant hosts recover at double the rate compared to non-resistant hosts. Infected hosts who are not resistant, first have to enter the *infected* compartment, and then the  $I_R$  compartment to be able to recover from malaria, while hosts in the  $S_R$  compartment transition directly to  $I_R$  compared when becoming infected. Basically, resistant hosts skip a whole state to recover. Since there is no way for a resistant host to become non-resistant, all hosts will in the end have become resistant. In essence, there will be no individuals in the  $I$  state just like the SIR model. If the basic reproduction number is above 0 in steady state all individuals will have been removed from the susceptible state. On long time scales the system can therefore be reduced to,

$$\frac{dI_R}{dt} = \alpha I_R S_R - \beta I_R, \quad (6.6)$$

$$\frac{dS_R}{dt} = \beta I_R - \alpha I_R S_R. \quad (6.7)$$

This simple two equation system ends with either malaria becoming extinct or endemic with a fixed proportion of the population being infected. Essentially, it settles in an equation system equivalent to the SIS model. For  $\alpha < \beta$  the disease will become extinct. For  $\alpha > \beta$  malaria will become endemic with a proportion of  $1 - \beta/\alpha$  being infected. This is shown by calculating the number the number of hosts in each compartment in steady state. Setting the equation system equal to 0 yields,

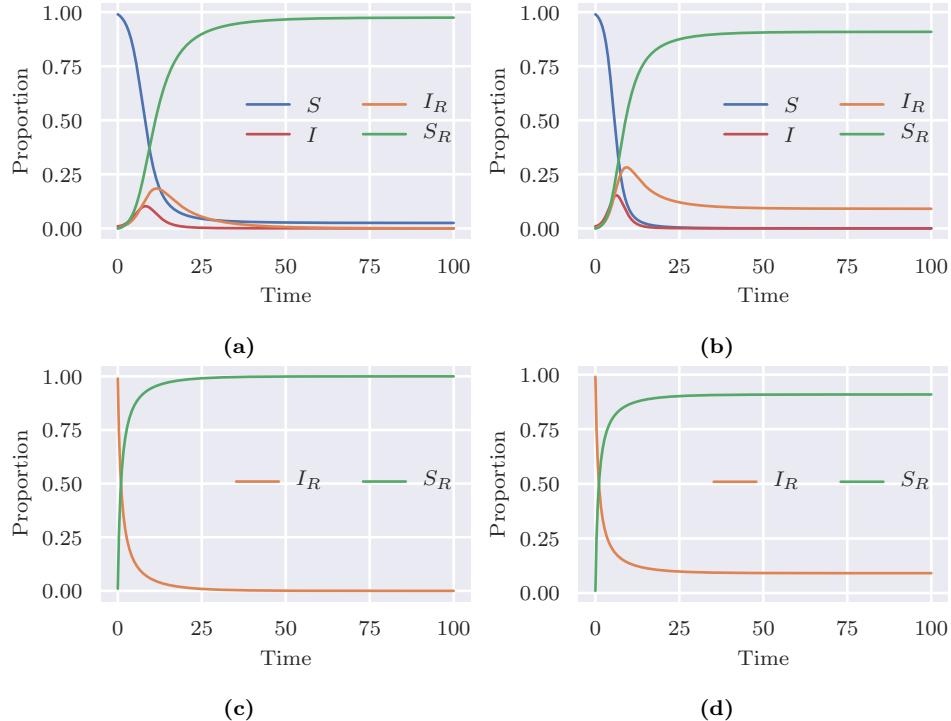
$$0 = \alpha I_R S_R - \beta I_R, \quad (6.8)$$

$$\rightarrow S_R(\infty) = S + S_R = \frac{\beta}{\alpha}, \quad (6.9)$$

$$\rightarrow I_R(\infty) = I + I_R = 1 - \frac{\beta}{\alpha}. \quad (6.10)$$

These steady state value gives the value of the number of infected individuals over long time scales.

Deterministic simulations show the same results as found above. Figures showing examples of simulations are shown in figure 6.2. The simulations are done with  $\alpha = 0.9$  and  $\alpha = 1.1$  while  $\beta$  is stuck at 1.0. This is done for both equations 6.2-6.5 and 6.6-6.7. In both cases the total number of infected at steady state is  $I + I_R = 0$  and  $I + I_R = 0.0909$ , for  $\alpha = 0.9$  and  $\alpha = 1.1$ , respectively. These values correspond to theoretical values according to equation 6.10. The simulations show exactly what has been hypothesised, namely that in the steady state everybody is resistant, and that the two equation systems are equivalent on long time



**Figure 6.2:** Deterministic simulations of equations 6.2-6.5 in figures (a) and (b) and 6.6-6.7 in figures (c) and (d).  $\alpha = 0.9$  for figures (a) and (c).  $\alpha = 1.1$  for figures (b) and (d).  $\beta = 1.0$  in all cases. The simulations are done using a fourth order Runge-Kutta method using  $\Delta t = 0.01$  over  $10^6$  steps. The x-axis correspond to one unit of time. The two kind of simulations are equal to each other in steady state, namely  $I_R = 0$  for (a) and (c) (In (a) the system is not in a completely steady state and  $I_R = 1 - 1.0/1.1 = 0.0909$  for (b) and (d)). This correspond with theoretical values according to equation 6.10. Specifically, the plots provide examples that the model in steady state is equivalent to a SIS model in steady state.

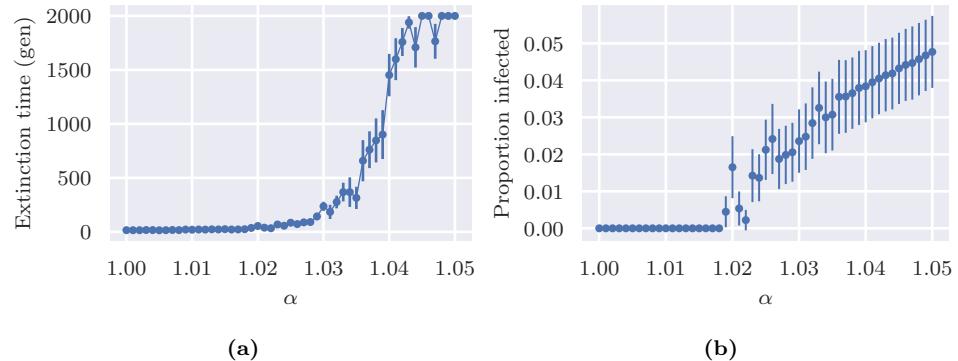
scales. The only difference is in figure ?? there still a small portion of susceptible individuals. This also happens in the SIR model, and comes from the fact that the disease becomes extinct before every individuals has been infected. This also means, that for guaranteeing everyone being at least through one infection, malaria must become endemic which happens if

$$\frac{\alpha}{\beta} > 1.0, \quad (6.11)$$

as in steady state a single infected individual will at least infect more than one other individual given that all other individuals are susceptible. Essentially, this model starts as an SIR model that over time converts into an SIS model.

The simulations show a great resemblance to the SIS model discussed in chapter 3. The SIS model achieves exactly the same steady state values.

Stochastic simulations give almost the same results as the theoretical results, but there are some key differences. This is evident by studying figures 6.3 where



**Figure 6.3:** Stochastic simulation of the simplified system expressed in equations 6.2-6.5 with  $N = 10000$  and  $\beta = 1.0$ . (a): Extinction time as function of infection rate. Extinction time is defined as the iteration where there is exactly 0 infected hosts. From  $\alpha = 1.03$  extinction time increases dramatically. (b): The mean proportion of infected hosts calculated from the steady state over 10 simulations. The theoretical values given in eq 6.10 for values  $\alpha > 1.036$ . The error bars represent the error on the mean for extinction time and square root variance of the distribution of a single simulation in the proportion infected. Both are estimated over 10 independent simulations.

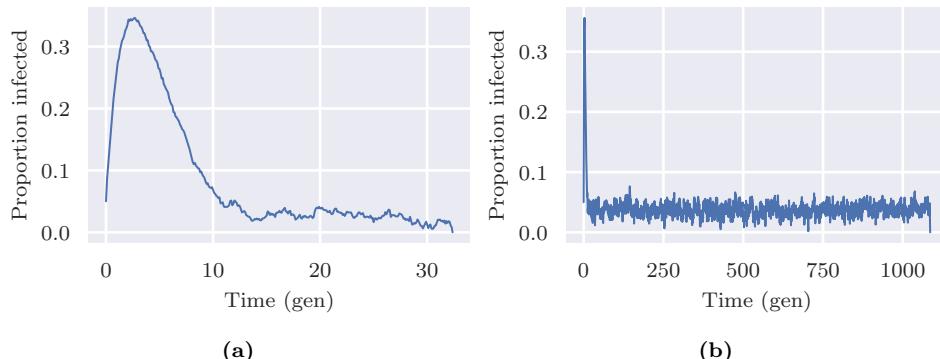
the time of extinction and proportion infected are shown. The simulations have parameters  $\alpha$  between 1.0 and 1.05 while  $\beta$  is constant at 1. The model is run with a number of hosts equal to  $N = 10.000$ .

It is expected that malaria becomes endemic on all simulated points for  $\alpha > \beta$  which happened in the deterministic case. With a simple test that counts any simulations where malaria survives for at least 500 generations, malaria is only consistently endemic when  $\alpha > 1.04$ . Even though the average over 10 repeated cases is above 500 generation, malaria does not consistently survive for long durations in each case which is also expressed by the error bars. basically, malaria has a non-insignificant probability to randomly die out for  $\alpha < 1.04$ . This is a difference of  $\alpha_{sto} - \alpha_{det} = 0.04$  compared to the deterministic case. The proportion of infected remains approximately the same, especially for larger  $\alpha$  as shown in figure 6.3b. For  $\alpha > 1.04$  it is clear that the system settles in a consistent endemic, as the proportion infected becomes nearly constant between separate simulations. The error-bars still have a significant value because of the random walk around the mid point as demonstrated in figure 6.4b.

In conclusion, the introduction of stochasticity decrease the probability of becoming endemic, but does not change the mean number of infected if malaria becomes endemic and stable. Only in a small range around the critical point does stochasticity make a significant difference on the endemic probability.

What is the source of this discrepancy between the deterministic and stochastic systems? Why does stochasticity reduce the survivability of Malaria? The answer is in the nature of stochasticity itself. This is demonstrated in figure 6.4. Here, figures of the proportion infected over the number of iterations is shown for  $\alpha = 1.02$  and  $\alpha = 1.04$ .

Malaria spreads quickly to a large portion of the population. As hosts become resistant, the steady state expressed in equations 6.6-6.7 is reached. Here, the left



**Figure 6.4:** Stochastic simulations of the system expressed in equations 6.2-6.5 with  $N = 10.000$  and  $\beta = 1.0$ . The figures show development of the number of infected as a function over time with  $\alpha = 1.02$  in (a) and  $\alpha = 1.04$  in (b). In (a) the number of infected stabilises at around the tenth generation, but dies off after a short amount of time because of stochastic processes. In (b) the number of infected stabilises at around  $t = 10\text{gen}$  and randomly dies at  $t = 1100\text{gen}$ .

most plot does reach an endemic state, but quickly becomes extinct, even though it would have been permanently endemic in the deterministic state. In the right figure, the system becomes stable, as it has a constant proportion infected for many generation. However, at some point, the number of infected reaches 0 even though malaria was in a stable state. The reason for this, is because of a chain of random events - specifically resistance events. The closer the number of infected is to 0 the more probable it is for malaria to randomly become extinct as the behaviour is much like a biased random walk.

Increasing the number of hosts in the system does improve the survivability of malaria, as it decreases the size of the proportional fluctuations. The stochastic system approaches the deterministic system as  $N \rightarrow \infty$ . However, this also increases computational time.

The concept of stochasticity in a malaria model is realistic. Malaria and diseases in general is namely affected by stochastic processes. Introducing malaria into a system with few individuals makes it much less probable for malaria to become endemic compared to a system with a large population. For example, it is much more likely that a mosquito never bites and transmits malaria to healthy individuals, if there only is 10 compared to say a 100 individuals before each individual has recovered. Basically, 100 events of one type happening in a row is much more unlikely than 10 events happening in a row. For this reason, the fact that the stochastic model exhibits this behaviour is a feature not a bug.

## Summary

It is shown that the model with only the core components is essentially identical to an SIS disease model over long time scales. Before that, the model can be compared to an SIR model. It was possible to translate the rules of the model outlined in chapter 5 to set of ordinary differential equations. This allowed us to reveal the

effect of the core components and analytically calculate when the system is in steady state and how many infected there will be as a function of model parameters. It was found that malaria becomes endemic when  $\alpha > \beta$  and the proportion of healthy hosts is  $S_R = \beta/\alpha$  at steady state.

Stochastic simulations almost give the same result as the deterministic results. The number of infected is the same in the two results as long as malaria is endemic. However, malaria can suddenly become extinct in the stochastic case through pure randomness, especially when the number of infected in the endemic state is low. This results in a different critical point going from extinction to survival. The critical point, which defines when malaria becomes consistently endemic, was calculated to be around  $\alpha/\beta > 1.04$  for  $N = 10000$ .

Overall, simplifications to the model behaves as expected and show predictable dynamics. It lays an important foundation for analysis when adding the more complex components.

## 6.2 Replacement

In this section, we build upon the groundwork done in the last section. A problem with the previous model using only the components of infection and resistance, was that the hosts would be permanently resistant. Basically, the purpose of immunity only mattered at the initialization of the system. The model needs an additional component that allows hosts to lose their resistance. This can be represented by arrival of new hosts through birth or immigration, or people simply losing their resistance over time, as discussed in section 5.

Replacement is the additional component that will be examined in this section. Functionally, the component works by removing a host from the system and inserting a new one. This means that a host which is replaced loses all resistances and any infections it might have. This can happen to a host in any compartment, but results in no changes for susceptible hosts. A diagram of the system with this addition is shown in figure 6.5.

The model can still be represented by a set of differential equations. The replacement component is represented by adding a single term in each of the compartments in equations 6.2-6.5. The term removes hosts from the  $I$ ,  $I_R$ , and  $S_R$  compartments and adds them to the  $S$  compartment, equal to the parameter  $\gamma$  multiplied by the number in the specific compartment. It is expressed as,

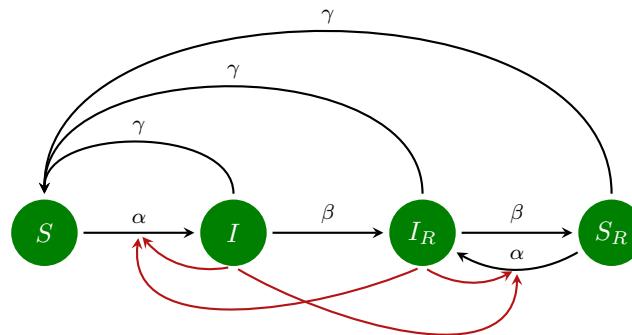
$$\frac{dS}{dt} = -\alpha(I(t) + I_R(t))S(t) + \gamma(1 - S(t)), \quad (6.12)$$

$$\frac{dI}{dt} = \alpha(I(t) + I_R(t))S(t) - \beta I(t) - \gamma I(t), \quad (6.13)$$

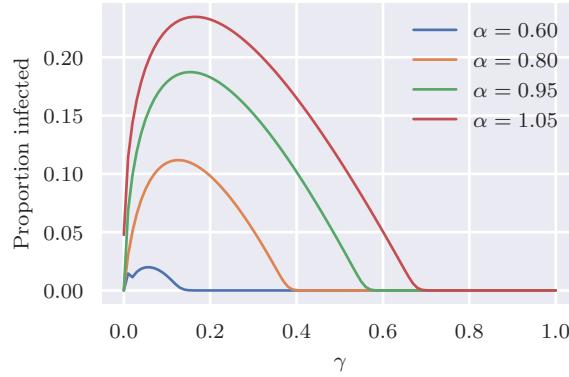
$$\frac{dI_R}{dt} = \alpha(I(t) + I_R(t))S_R(t) - \beta I_R(t) - \gamma I_R(t), \quad (6.14)$$

$$\frac{dS_R}{dt} = -\alpha(I(t) + I_R(t))S_R(t) + \beta I_R(t) - \gamma S_R(t). \quad (6.15)$$

The replacement component is expressed as the terms with the constant  $\gamma$ . The term  $(1 - S(t))$  in 6.12 is exactly that because the constraint  $1 = N = S(t) + I(t) +$



**Figure 6.5:** Transition diagram with the added replacement event. The diagram is equivalent to figure 5.1.



**Figure 6.6:** Plot of the proportion infected at steady state as function of  $\gamma$  for the four alpha values shown in the legend. The plots are visualisations of the solution of equation system 6.16-6.19.

$I_R(t) + S_R(t)$  holds.

The analysis starts by finding a steady state solution. We follow the same procedure as in the last section. The whole system is written and set equal to 0:

$$0 = -\alpha(I(t) + I_R(t))S(t) + \gamma(1 - S(t)), \quad (6.16)$$

$$0 = \alpha(I(t) + I_R(t))S(t) - \beta I(t) - \gamma I(t), \quad (6.17)$$

$$0 = \alpha(I(t) + I_R(t))S_R(t) - \beta I_R(t) - \gamma I_R(t), \quad (6.18)$$

$$0 = -\alpha(I(t) + I_R(t))S_R(t) + \beta I_R(t) - \gamma S_R(t). \quad (6.19)$$

In the last section, it was easy to find the a steady state solution because the system could be reduced to only two equations. Here, however, hosts do not remain resistant and it is therefore not possible to make such a reduction. Despite this, the system is still solvable. The derivation of solution is too involved to show directly, so the solution was computed using Matlab. The code for the solution is shown in the appendix. The steady state solution was essentially to solve a fourth order equation using appropriate constraints.

$$S(\infty) = \frac{\beta^2 + 2\beta\gamma + \gamma^2}{2\alpha\beta + \alpha\gamma} \quad (6.20)$$

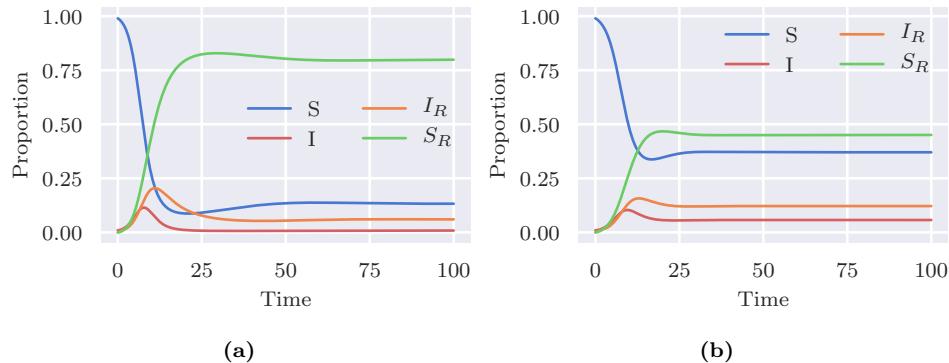
$$I(\infty) = \frac{\gamma(\beta^2 - 2\beta\gamma + 2\alpha\beta - \gamma^2 + \alpha\gamma)}{2\alpha\beta^2 + 3\alpha\beta\gamma + \alpha\gamma^2} \quad (6.21)$$

$$I_R(\infty) = \frac{\gamma(-\beta^3 - 2\beta^2\gamma + 2\alpha\beta^2 - \beta\gamma^2 + \alpha\beta\gamma)}{2\alpha\beta^3 + 5\alpha\beta^2\gamma + 4\alpha\beta\gamma^2 + \alpha\gamma^3} \quad (6.22)$$

$$S_R(\infty) = \frac{-\beta^4 - 2\beta^3\gamma + 2\alpha\beta^3 - \beta^2\gamma^2 + \alpha\beta^2\gamma}{2\alpha\beta^3 + 5\alpha\beta^2\gamma + 4\alpha\beta\gamma^2 + \alpha\gamma^3} \quad (6.23)$$

The other solution is when every compartment is 0 except for  $S$  which is then 1. This happens when  $\gamma$  is positive and malaria does not become endemic.

Plots of the solution for four different values of  $\alpha$  are presented in figure 6.6. They reveal that the replacement component increases the proportion of infected



**Figure 6.7:** Deterministic simulations of equations 6.12-6.15. (a) and (b) have parameters  $\gamma = 0.01$  and  $\gamma = 0.1$ , respectively with  $\alpha = 0.95$ . The total proportion of infected in steady state is  $I + I_R = 0.0685$  in (a) and  $I + I_R = 0.1792$  in (b). The larger replacement value is a significant advantage for malaria. This corresponds with theoretical values according to equations 6.20-6.20. A much larger proportion of the population is infected just from increasing  $\gamma = 0.01$  to  $\gamma = 0.1$ .

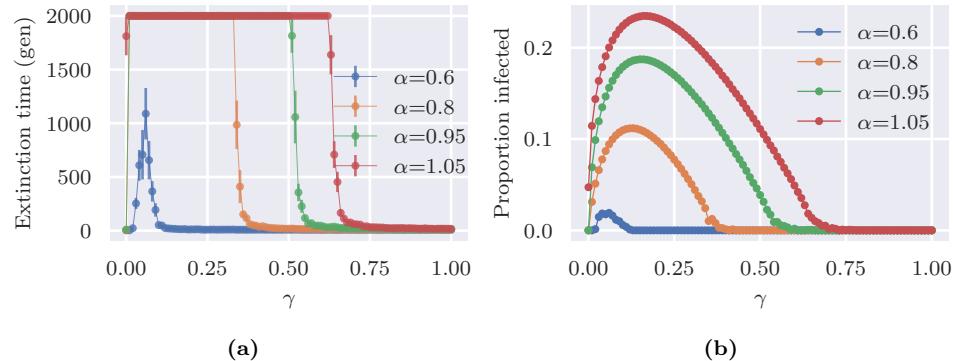
hosts and enhances malaria's survivability. For values as low as  $\alpha = 0.51$ , malaria is capable of surviving for a range  $\gamma$  values. This is a huge survivability boost compared to the system with no replacement. However, this is not true for large  $\gamma$  values, where the component can become detrimental to malaria. As an example, for  $\alpha = 1.05$  malaria is not able to survive for  $\gamma > 0.7$ .

This fact becomes more evident when studying figures 6.7, where the time-series are shown for two specific cases of  $\alpha$  and  $\gamma$ . A big increase is shown in the proportion infected by increasing  $\gamma = 0.01$  to  $\gamma = 0.1$  resulting in an increase of  $\Delta I + \Delta I_R = 0.1107$ . The fact that a much smaller proportion of the population is not resistant makes the mean time to recover significantly larger. If  $\gamma$  would be increased further, it would become a detriment for malaria as infected hosts also gets replaced. Therefore, large values of  $\gamma$  can make malaria become extinct, even if it would not otherwise.

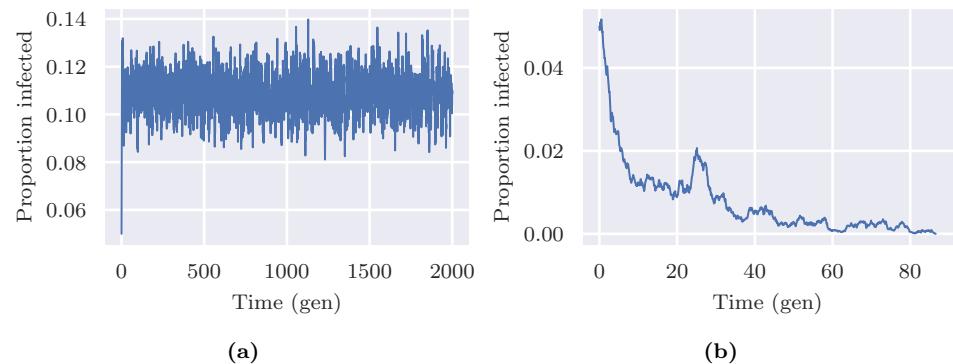
The system shows many of the same dynamics as an SIR model with vital dynamics discussed in section 3.3. Both show asymptotic stability and greatly depends on the rate of replacement. This gives an indication that the model is not unrealistic.

Stochastic simulations of this model show the same approximate behaviour as the deterministic simulations. This is seen in figure 6.8 which shows stochastic simulations over the same parameter space as in figure 6.6. The difference is the same as in the simplified model - stochasticity makes malaria becomes extinct at certain values where malaria would otherwise become endemic. For example, with  $\alpha = 0.6$  and  $\gamma \approx 0.1$  malaria becomes endemic in the deterministic case, but not in the stochastic case. The explanation for this is the same as the one given in the last section. Again, these differences become smaller with a larger number of hosts.

Two time-series examples are shown in figures 6.9. One is an example of Malaria becoming endemic, where it would not without the replacement component. In the other malaria does not become endemic, even though it would become endemic without replacement. They exemplify the fact that replacement is an advantage



**Figure 6.8:** Stochastic simulations of equation 6.12-6.15 showing extinction time and mean infected as a function of  $\gamma$  for different values of  $\alpha$ . The probability that malaria survives increases with the introduction of replacement and decreases around  $\gamma = 0.55$ , which is approximately the theoretical values according to equation 6.20-6.20. Error bars are omitted from the right plot for visual clarity.



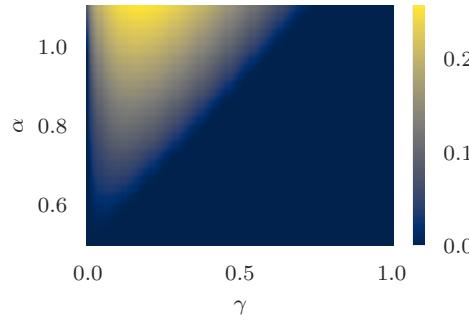
**Figure 6.9:** Stochastic simulations of equation 6.12-6.15 showing an evolution of the number of infected for two different cases. (a) has  $\alpha = 0.8$  and  $\gamma = 0.17$  (optimal is 0.16961) and (b) has  $\alpha = 1.05$  and  $\gamma = 0.75$ , exemplifying two values of  $\gamma$  that is advantageous and disadvantageous, respectively.

for malaria within lower values of  $\gamma$ , but is a detriment for large values of  $\gamma$ .

A thing to note from the previous figures is that there seem to exist an optimal replacement rate. That is to say there exists one and only one replacement rate that maximises the number of infected for any given infection rate. This is also evident from studying figure 6.10, which shows the mean number of infected as a contour plot over  $\alpha$  and  $\gamma$ . The maximum number of mean infected increases as alpha increases, and so does the optimal  $\gamma$  which maximises this. The same thing is seen in the SIR with vital dynamics model described in section 3.3.

The optimal replacement can be calculated analytically which is in an optimization problem, by taking the derivative of  $I + I_R$  and setting it equal to 0. Solving that equation gives multiple solutions<sup>1</sup>, but only one real solution. The real solution

<sup>1</sup>The solution was computed using MATLAB. The code doing so is shown in the appendix



**Figure 6.10:** 2D deterministic plots with  $\alpha$  and  $\gamma$  as variables showing the mean number of infected under a deterministic simulation.

is

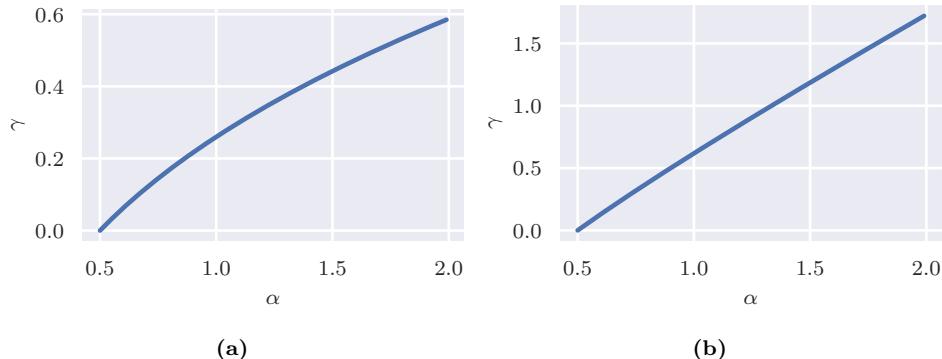
$$\gamma_{\text{optimal}} = (2\alpha)^{1/3} - 1. \quad (6.24)$$

The values gained from this function corresponds with simulations. The optimal replacement rate increases with  $\alpha$  as a power to the one third, so it only increases slowly with  $\alpha$ . A visualisation of the function is shown in figure 6.6. It is evident that the effects of replacement are significant. For example, at  $\alpha = 1.05$  the proportion infected more than doubles, when going from  $\gamma = 0$  to  $\gamma_{\text{optimal}} = 0.281$ , and it makes malaria endemic as low as  $\alpha > 0.5$

Studying figure 6.10 reveals a survival triangle, which means a region with the shape of a triangle where malaria becomes endemic.  $\gamma$  can become so large that malaria can not become endemic, even though it might do without the replacement component. The limits for this triangle can be found analytically by calculating for which  $\alpha$  there are no infected. The code that calculates it is shown in the appendix. It gives.

$$\gamma_{\text{extinction}} = \frac{\alpha}{2} + \frac{\sqrt{\alpha(\alpha + 4)}}{2} - 1. \quad (6.25)$$

This is almost a linear increase as shown in figure 6.11. This corresponds with simulations where the proportion infected goes to 0 in figure 6.10.



**Figure 6.11:** Plots of the optimal replacement rate  $\gamma$  (a) and the upper range of replacement rate which causes Malaria extinction (b). They are based on equation 6.24 and 6.25, respectively. At  $\alpha < 0.5$  Malaria is not able to survive.

### Summary

The component of replacement is an important addition for malaria to become endemic, not unlike how Vital Dynamics enables endemicity in the SIR model, as it makes hosts to lose resistance and become susceptible again. And analytical solution was found of this system in steady state.

In this section, it has been shown that for low  $\gamma$  values replacement advantageous for malaria. For gamma values greater than the ones derived in equation 6.25 malaria becomes extinct. Replacement can therefore be a detriment to malaria for large values. This show that Plasmodium has no advantage gained from killing it's victims.

A mathematical term was found for the most optimal value of  $\gamma$  shown in equation 6.24 . For the optimal  $\gamma$ , it is possible for Malaria to become endemic for  $\alpha > 0.5$ , which is significantly larger than without replacement.

### 6.3 Strains

In this section we will examine the system when additional strains are introduced<sup>2</sup>. The strains behave as described in chapter 5. The number of strains will be randomly distributed at initialization between the infected hosts. Also recall that super-infection is present, which means that a host can be infected by multiple strains. The approach of analysis will be the same way as the previous sections. First, we'll see what information can be gained from an analytical approach. Afterwards, stochastic simulations will be examined. The number of strains will be investigated for a range from 1 to 25.

To start, we will investigate the model with only two strains in the system, for simplicity reasons. With two strains a host can be susceptible, infected, infected and resistant, and resistant for each of the two strains. In other words, a host can be in any combination of the four compartments for each strain. For example, a host can be infected by the first strain, but be resistant to the other one, which would put the host in the  $II_R$  compartment. This results in a total of 16 compartments for two strains. The system with two strains is visualised in figure 6.12.

The fact that the system has so many compartments complicates the model substantially. It is, however, still possible to write the model in differential form. However, it is extremely unwieldy because of the number of equations and the number of terms in each equation. This can be exemplified by writing the susceptibility compartment to both strains,

$$\begin{aligned} \frac{dSS}{dt} = & -\alpha SS(t)[IS(t) + SI(t) + II(t) \\ & + I_RS(t) + SI_R(t) + IRI_R(t) + IS_R(t) + SRI(t) \\ & + I_RS(t) + SI_R(t) + II_R(t) + IRI(t)] + \gamma(1 - SS(t)), \end{aligned} \quad (6.26)$$

where each double lettered symbol represents a compartment, one letter for each strain. A host that is susceptible to both strains, can become infected by any host that contains either, and they all do so with the same factor. The  $\alpha$  term is therefore the rate of transition away from that compartment (that is becoming infected). It contains a single factor for the size for each of the 12 compartments that contains at least an infection from any strain, which is written as the sum of all compartments that contains an  $I$  or an  $I_R$ . The  $\gamma$  term is the replacement term, as all host in all states return to the  $SS$  compartment when replaced. The differential equations of the 15 other equations has been omitted for brevity, but follows the same formula.

By combinatorics, the number of compartments as a function of the number of strains are

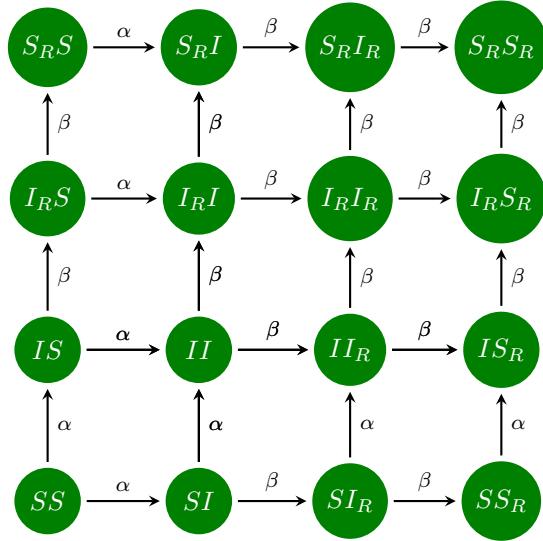
$$\text{Compartments} = 4^S, \quad (6.27)$$

where  $S$  is the number of unique strains in the system. That is an enormous increase of compartments with just a little increase in the number of strains.

The sheer number of compartments in the system, and the number of terms in the equations make analytical investigation unproductive. Therefore, we will rely on simulations to examine the system. As we have already examined the difference between stochastic and deterministic simulations, the simulations will only be done

---

<sup>2</sup>The number of antigens in a single strain remains 1



**Figure 6.12:** Diagram of the model with 2 strains. Each letter in a compartment denotes what state a host is in for each strain. For example, the  $II_R$  compartment denotes hosts that are infected by the first strain, and infected and resistant to the other strain. Replacement has been omitted for visual clarity. All compartments become replaced with a rate  $\gamma$  which results in returning to the  $SS$  compartment. Moreover, the red lines visualising in previous diagrams is not shown, again for visual clarity. Each host infected by a strain would increase the probability of being infected by that strain as outlined in chapter 5. If a third would be inserted into the diagram, it would fill out the third dimension with a total of  $4^3 = 64$  compartments.

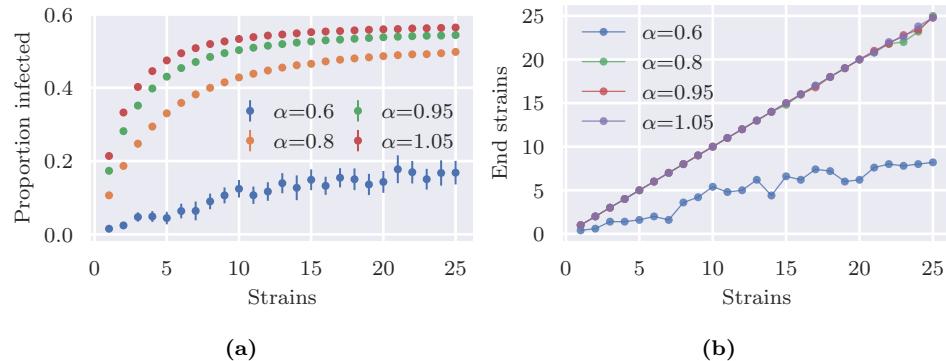
agent-based and stochastically as described in chapter 5.

To show the effects of additional strains the model is simulated with a varying number of strains. The optimal replacement rate  $\gamma_{\text{optimal}}$  written in equation 6.24, will be used throughout unless otherwise mentioned. By using the optimal replacement speed, the system will behave properly and we remove a parameter at the same time. This will also simplify the investigation. Moreover using  $\gamma_{\text{optimal}}$  also equalizes the system when varying  $\alpha$ , making comparison easier.

A plot of stochastic simulations showing the proportion infected for  $\alpha = 0.6$ ,  $\alpha = 0.8$ ,  $\alpha = 0.95$ ,  $\alpha = 1.05$  with  $\gamma_{\text{optimal}}$  as a function of number of strains in the system is shown in 6.13a. It shows a general increase in the proportion infected as the number of strains increases.

Focusing on the simulations where  $\alpha \geq 0.8$ , the increase in the proportion infected as a function of strains is consistent. This is further indicated from the almost non-existent error-bars. There are two reasons why multiple strains increase the proportion infected.

Firstly, the more strains there are in the system the less chance there is for a given host to be resistant to that host. So, the more strains there are in the system, hosts that become infected has less chance to have the corresponding antibodies. This becomes especially apparent by looking at the average number of resistances



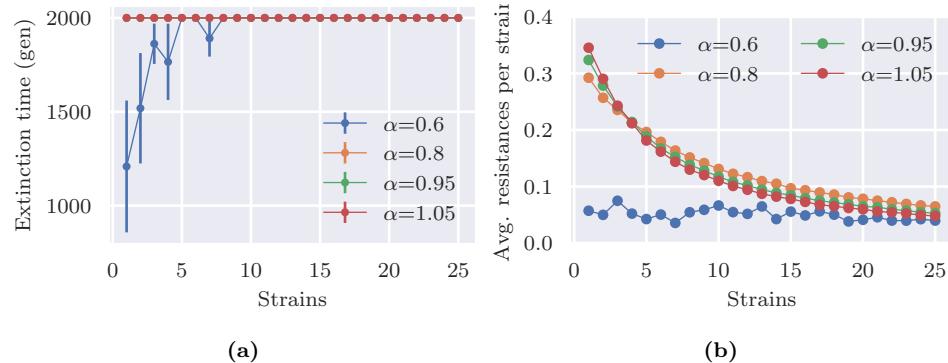
**Figure 6.13:** Proportion infected (a) and number of surviving strains after 2000 generations (b) as a function of starting strains for four different values of  $\alpha$ . The replacement rate is the optimal replacement expressed in equation 6.24. Note that in the right plot, the points of  $\alpha = 0.8$ ,  $\alpha = 0.95$ , and  $\alpha = 1.05$  are on top of each other. The values are calculated from the mean of 10 independent simulations.

over the whole population shown in figure 6.14b, which show are tending decrease of the average number of resistances in each strain. The average number of resistances in each strain is directly related to the probability of a given host has a given antibody when infected.

The second reason is the fact that a single host can become infected by multiple strains, e.g. super-infection. This translates to a greater proportion of the population that can be infected by a strain when an infection event happens, practically increasing the probability of an infection event. For example, a host in the state of *SI* is able to infect a host with the *IS* state. Moreover, if a host is infected by multiple strains, it has to get rid of them independently. It takes a host with no relevant antibodies and infected by two strains double the amount of time to build resistance and recover compared to being only infected by one strain. This can be seen in figure 6.12 following the *II* compartment to the *S<sub>R</sub>S<sub>R</sub>* compartment. Hosts in that compartment has to transition fours times before it recovers<sup>3</sup>. The mean time to recover for hosts with a single infection is  $1/(2\beta)$  but  $1/(4\beta)$  for hosts with two infections. This, combined with the fact that a larger proportion that can be infected by each strain, as explained above, makes multiple strains very effective for increasing the amount infected.

For  $\alpha = 0.6$  the increase in the proportion infected is not smooth, as evidenced by the big error bars shown in figure 6.13a. A general increasing tendency exists in the proportion infected, but it is less consistent compared to the higher infection rates. The reason for this is because it is possible for individual strains to become extinct. This is shown in figure 6.13b where the number of surviving strains in the system after 2000 generations is plotted as a function of the amount of starting strains. The amount of strains at the end of the simulation is always lower than the number of starting strains for  $\alpha = 0.6$ . The higher the amount of strains at the end of simulations the larger is the proportion of infected as well. The strains evidently become extinct over time, but what is the reason for this?

<sup>3</sup>For example *II*  $\rightarrow$  *I<sub>R</sub>I*  $\rightarrow$  *I<sub>R</sub>I<sub>R</sub>*  $\rightarrow$  *I<sub>R</sub>S<sub>R</sub>*  $\rightarrow$  *S<sub>R</sub>S<sub>R</sub>*



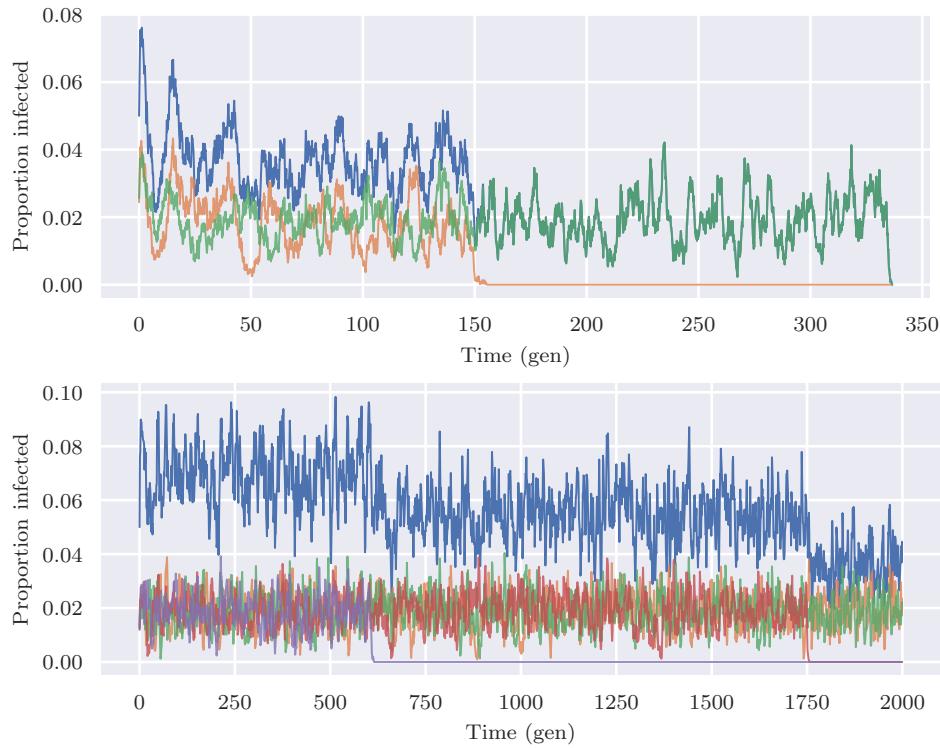
**Figure 6.14:** Figures showing extinction time and resistances in each strain from left to right. They use  $\alpha = 0.6$  and  $\gamma_{optimal}$ . The average number of resistances in each strain is calculated from the sum of all antibodies in all hosts divided by the number of hosts and strains therefore taking a value between 0 and 1.

The reason is the same reason why malaria become extinct in some regions of alpha in the previous sections. The strains become extinct at random. The random fluctuations are explicitly shown in figure 6.15. In both cases shown, the system goes into a steady state, where malaria is seemingly endemic and stable. But, only for a time. As time progresses, there is a chance that one strain becomes extinct. This results in there being one less strain in the system. This may happen any number of times, such as the two times it happens in the lower figure of figures 6.15. The consequences of this, is that inserting more strains into the system has greater effect in systems with high base  $R_0$ . For low  $\alpha$  values, malaria takes longer time to become extinct, but will become extinct eventually as each strain dies. Also notice that each time a strain dies, the total number of infected decreases with it identified at  $t \approx 600$  and  $t = 1750$  in the lower plot of figure 6.15.

A peculiarity of increasing the amount of strains, is the fact that it does not significantly increase the probability of Malaria becoming endemic. This fact becomes evident from studying figure 6.16a. There, the extinction time is plotted as a grid over infection rate and the number of strains. A very clear boundary around  $\alpha = 0.6$  is shown. Most simulations above  $\alpha > 0.6$  survives for many generation while simulations below  $\alpha < 0.6$  does not. This agrees with equation 6.25. This is independent of the number of strains in the system. The probability of Malaria becoming endemic is the same if  $S = 1$  and  $S = 10$ . Only where Malaria would already become endemic, does additional strains have an effect, except for maybe a longer extinction time.

What is the cause of this? The explanation can be found in figures 6.15. Because the infection rate is low, the proportion infected is low enough for stochastic properties to have effect. Each strain is essentially not competitive enough to remain endemic, and they do not support each other in remaining alive and can be considered their own individuals pathogens.

In the average resistance plot of figure 6.14b, an odd behaviour emerge. When the number of strain is below three there are more resistant hosts for larger values of  $\alpha$ . As the number of strains increase the behaviour changes and it is the system

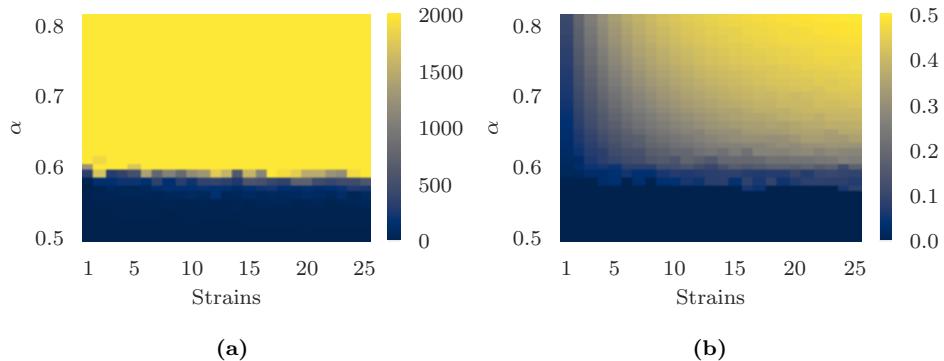


**Figure 6.15:** Two time-series plots using  $\alpha = 0.6$  and  $\gamma_{optimal}$  with a different number of strains. The upper one has two strains and the lower one has four strains. The blue solid line is the total number of infected, while the other lines is the number of infected for each individual strain. In the upper figure one strain become extinct at generation  $t \approx 150$  and at  $t \approx 340$ , resulting in malaria's extinction. In the lower figure, one strain becomes extinct at around generation  $t \approx 600$  and another one at  $t \approx 1750$ , allowing malaria to survive for a longer time.

with the largest  $\alpha$  that gets the least number of resistances in each strain<sup>4</sup>. Between  $\alpha = 1.05$  and  $\alpha = 0.8$  the relative difference maximises at  $S = 25$  by a factor of 1.35. This is odd behaviour? Why is there this sudden change?

The reason why there is lower average resistance in each strain for  $\alpha = 1.05$  than  $\alpha = 0.8$  for  $S > 4$  is because of the fact that the hosts cannot keep up with new infections. Because of super-infection, new infection keep accumulating before the hosts can build antibodies and recover from the other strains they are already infected by. When many strains are in the system, the infection rate is equal to  $\alpha I$ , as explained in chapter 5. Higher  $\alpha$  values then lower the chance for resistance events despite the resistance rate grows proportionally with  $I + I_R$ . It essentially comes down to the probability of recovery events compared to the sum of the other events. This is the reason for this odd behaviour.

<sup>4</sup>Except, for  $\alpha = 0.6$ , obviously, since the dynamics at that rate is distinct from the others



**Figure 6.16:** Contour plots showing the extinction time (a) and proportion infected (b) as a function of  $\alpha$  and number of starting strains. The units are generations and proportion infected, respectively. Malaria always become endemic for all simulations where  $\alpha > 0.6$ .

### Summary

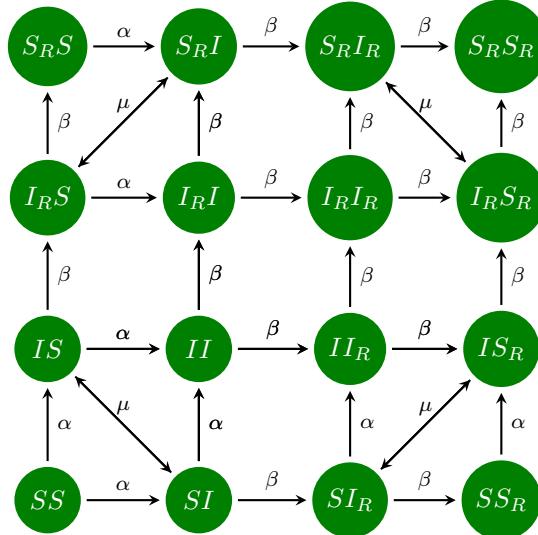
In this section, additional strain was introduced. It was shown that including strains increased the proportion infected. The strains were practically independent diseases, and hosts therefore had to build many antibodies to be completely resistant. However, it did not increase the probability for malaria to become endemic in regions of  $\alpha$  where Malaria would otherwise not become endemic. It was found, that each individual strain died out separately, which caused for a higher extinction time. The disease would always become extinct eventually, however. The reason for this was because of stochasticity. Because of the permanent extinction of strains the mutation component will be introduced.

## 6.4 Mutation

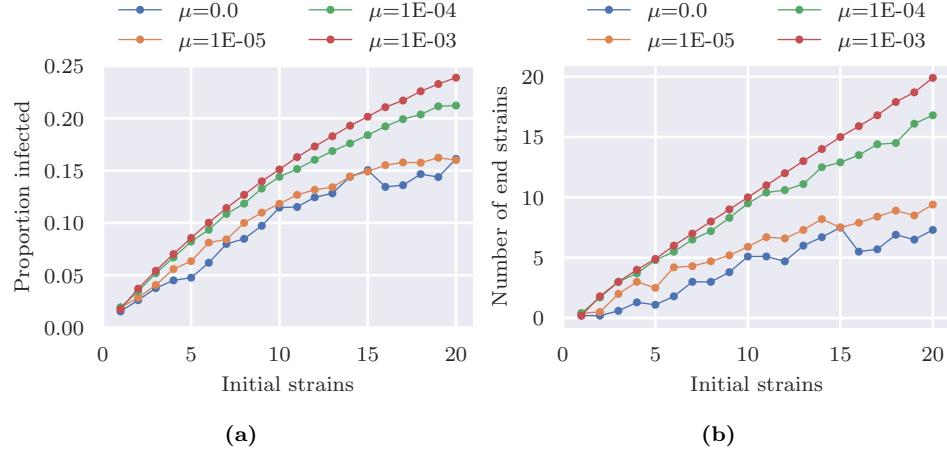
In the last section, we saw that separate strains could permanently become extinct, especially for  $\alpha$  values below 0.8. Because of this, it is more difficult to compare systems with low  $R_0$  for different amount of strains. Moreover, it is a necessary component to correctly identify cross immunity, which will be discussed in the next section.

Therefore, there has to exist some kind of safety valve, that makes sure that a single strain is not permanently extinct. To resolve the problem, the component of mutation is introduced. The component is an event type that changes a randomly chosen strain to another randomly chosen strain. The rate of this event is equal to the parameter  $\mu$  multiplied by the total number of infected, as shown in table 5.1. A diagram of compartment transitions is shown in figure 6.17. The event represents both infections from outside sources and active mutation in malaria parasites. With this simple component, a strain will never become fully extinct, as it will always have a chance to get back into the system. For example, if a strain infecting a host may mutate into another strain event if the other strain currently infects 0 individuals

To investigate the impact of introducing a mutation component, we'll use the same parameters as in the last section, but investigate them with a small variety of mutation rates. Then the different scenarios can be compared.



**Figure 6.17:** Diagram of the model with two strains including the component of mutation. The diagram is almost identical to figure 6.12 except the eight new possible transitions. The new transitions comes from the mutation component and contain  $\mu$ . They allow any infected hosts strain to be converted to any other strain in the system, even if the strain has become extinct. The primary purpose of mutation is to avoid single strains becoming extinct. Note that the mutation transition arrows all point to and from the nodes in question. The replacement lines and the support lines have been omitted just as in figure 6.12



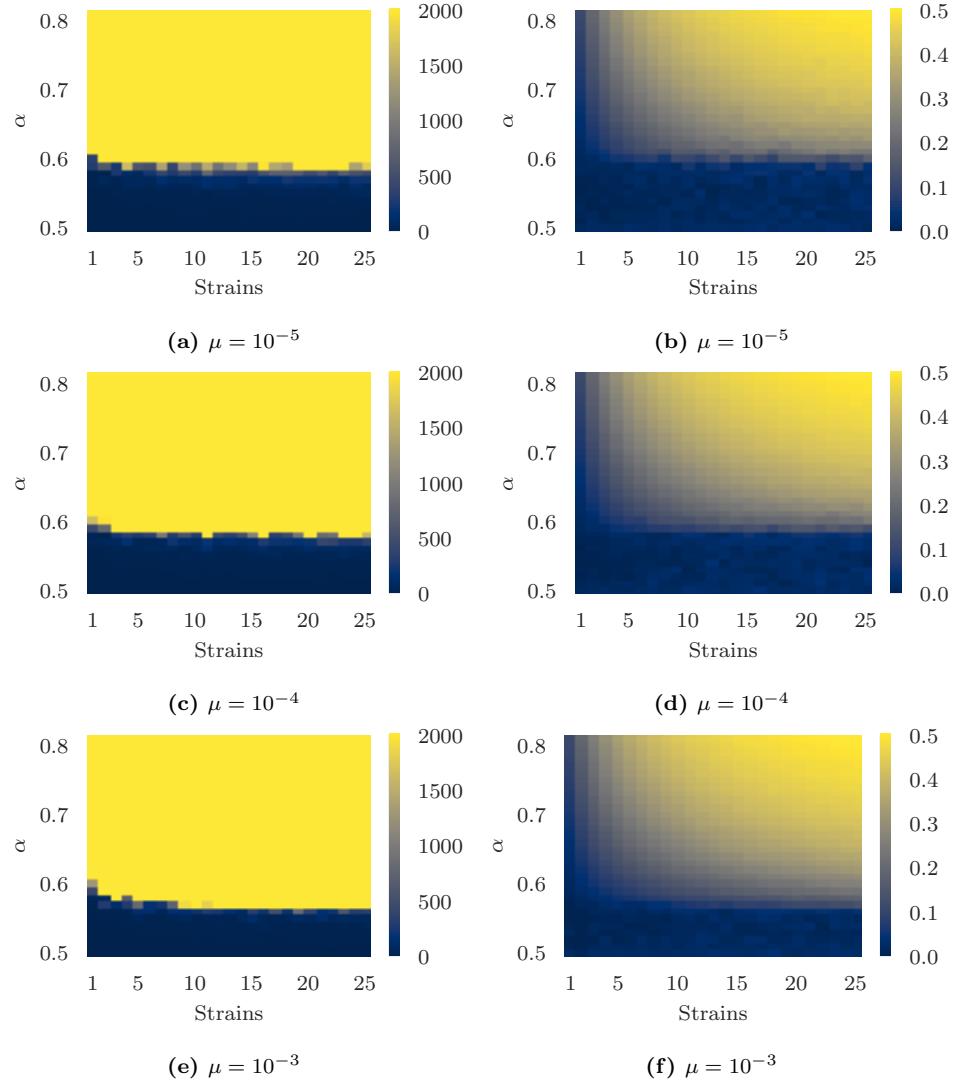
**Figure 6.18:** Plot of the proportion infected (a) and the number of strains at end of simulation (b) over various mutation rates shown in the legend.  $\alpha = 0.6$  and the optimal  $\gamma$  is used in all cases. The average is taken over 10 simulations. A general increase in the proportion infected for greater  $\mu$  is clearly visible. The number of strains at end of simulation also increases at greater mutation rates, and is even able to be equal to the number of starting strains consistently for  $\mu = 10^{-3}$ .

Without the mutation component each unique strain was able to avoid extinction if the infection rate was high enough, which was included values of  $\alpha > 0.65$ . Consistent survival was definitely seen for  $\alpha > 0.8$ . For this reason we will look at  $\alpha$  values lower than 0.8 and mainly  $\alpha = 0.6$ . One main expectation is the fact that the introduction of the mutation component would increase the range of  $\alpha$  values where malaria becomes endemic (i.e. between  $\alpha = 0.5$  and  $\alpha = 0.6$ ). It is also likely that a greater proportion of the population will be infected, but only for  $\alpha$  values where every strain did not consistently survive. It will be tested if the mutation component will help malaria become endemic, if it increases the proportion infected, and if it makes sure that strains does not completely become extinct on large time scales.

The first thing to consider is how various mutation rates changes the system for a specific infection rate. For now, we only consider the system with  $\alpha = 0.6$ , as that is a value where malaria is almost consistently endemic which we saw in the previous section.

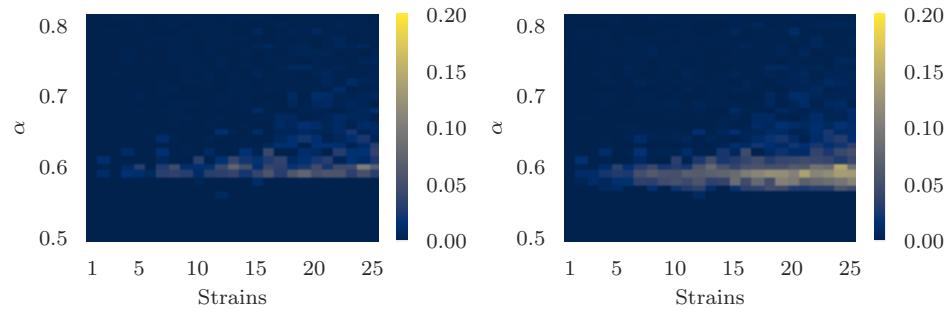
In figure 6.18 plots of the mean infection and number of strains at the end of simulation are shown for 4 different values of mutation rates  $\mu$  namely  $\mu = 0.0$ ,  $\mu = 10^{-5}$ ,  $\mu = 10^{-4}$ , and  $\mu = 10^{-3}$ . We see that the mean number of infected increases with higher parameters. As expected, the number of strains at end of the simulation also increases. For  $\mu = 10^{-3}$  the number of strains at end of simulation is almost always the same as the starting number of strains. This means, that the mutation component does what is expected, namely keeping single strain artificially alive.

While the proportion infected increases for  $\alpha = 0.6$  by inserting the mutation component, the mutation also increases the regions of  $\alpha$  where malaria becomes endemic, as shown in figures 6.19 and 6.20. The greater the mutation rate is, the



**Figure 6.19:** Contour plots showing the extinction time and proportion infected on the left and right side, respectively. The colour-bars is measured in generations and proportion infected, respectively. This is done for 3 different mutation rates denoted on each figure caption. The figures of extinction time show an increasing tendency of malaria becoming endemic in the regions between  $\alpha = 0.55$  and  $\alpha = 0.6$  for greater mutation rate becomes. This is seen by the yellow region (maximum survival) encroaching into the blue region. The proportion infected does not increase for about  $\alpha > 0.65$ , where each strain would consistently survive.

greater the regions of infection rates where malaria becomes endemic. For 20 strains the constraining infection rate is about  $\alpha = 0.61$  with  $\mu = 0.0$  as shown in figure 6.16. This increases to  $\alpha = 0.6$  for  $\mu = 10^{-5}$ ,  $\alpha = 0.59$  for  $\mu = 10^{-4}$ , and  $\alpha = 0.57$  for  $\mu = 10^{-3}$ . It is not be a large difference, but small differences can have an impact on whether malaria becomes endemic or not. The fact that malaria can



**Figure 6.20:** Residuals of proportion infected between the system with no mutations and with mutations. The left one has  $\mu = 10^{-4}$  and the right one have  $\mu = 10^{-3}$ .

become endemic for  $\alpha = 0.58$  could be the deciding factor for malaria's survival.

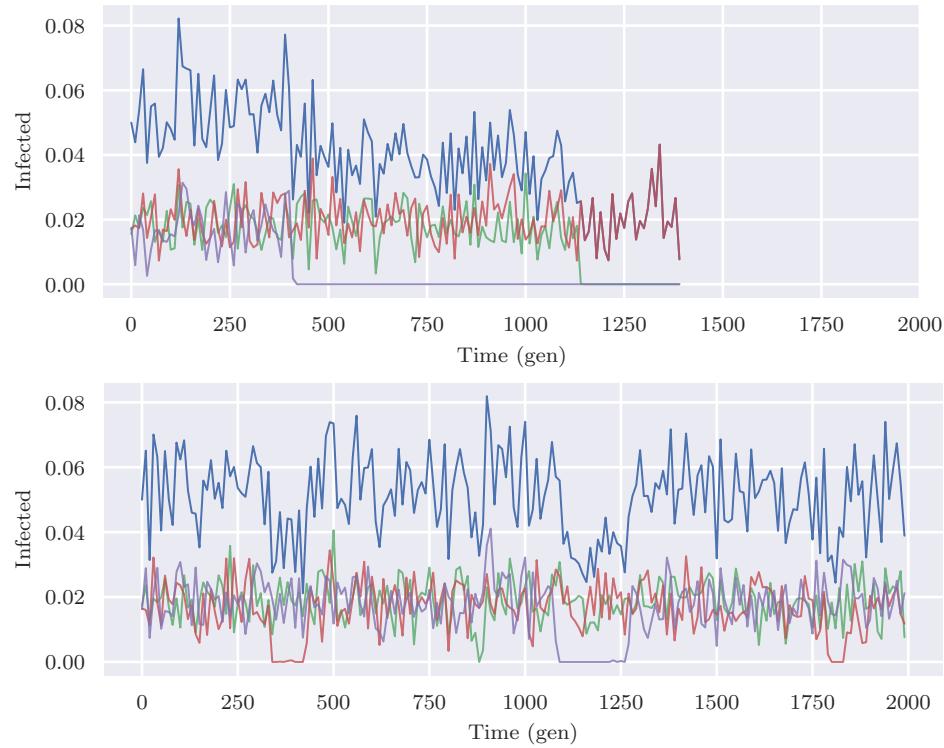
The reason for the increase was pointed out in the last section. For large infection rates the individual strains are able to avoid becoming extinct. The mutation component only makes a large differences in regions of  $\alpha$  where individual strains have a good probability of becoming extinct.

The proportion infected, shown in the right row of figures 6.19, only increase in values of  $\alpha$  where strains could become extinct, which means that  $\alpha < 0.65$ . For greater  $\alpha$ , the mutation component simply makes no difference to the system, because individual strains have very low probability of becoming extinct. This is shown by the constant number of strains in figure 6.13, and the residual plots of 6.20. The maximum found increase caused by the mutation rate was 0.18 at  $\alpha = 0.6$  and  $\mu = 10^{-3}$

How the mutation component changes the system dynamically is illustrated in the time-series plots of 6.21, showing the behaviour with and without the mutation component. There is a very clear difference between the 2 time-series plots. In the upper one without mutation component each individual strain dies out one at the time, which leads to the extinction of malaria. With a mutation rate of  $\mu = 10^{-4}$  the individual strains are after some generations able to re-emerge. Therefore malaria is able to remain endemic even over long periods of time.

Consider the small increase of the extinct strain at  $t \approx 400$  in the lower plot of figure 6.21. There, a mutation event happened causing the dead strain to re-emerge, but the strain quickly goes extinct again. This exemplifies that an extinct strain not only needs a mutation event to re-emerge, but also have to get a series of transmissions. Larger rates, would result in more changes for extinct to re-emerge. But it would also cause dominating strains to shift, which results in a lower chance for a strain to become extinct in the first place, and to have a more event distribution of the proportion of the population each strain infects. This also explains the more consistent increase of the number of end strains at larger mutation rates in figure 6.18.

To reduce the number of parameters going forward. a single mutation rate will be used. That poses the question of which  $\mu$  value to use. The value chosen, must be small enough so that individual strains cannot become permanently extinct, unless the malaria as a whole becomes extinct. Extinct strain must be able to



**Figure 6.21:** Plot showing extinction time and mean infected as a function of infection rate for  $\alpha = 0.6$ . The upper plot has no mutation component, while the lower one does with a value of  $\mu = 10^{-4}$ . Both time-series plots has exactly three strains to start with. The left one dies before reach 2000 generation - the strains die one for one. Strains also become extinct in the right one - three times at around 400, 1100, and 1800. But, because of the mutation rate the strains come back into the system after some time. Note that the time-series plots are shown with some granularity, and does not show every time-step.

re-emerge often enough, so they are able to calculate using a reasonable time frame. This excludes mutation rates smaller than  $10^{-5}$ .

However, the rate must not become too large. A large mutation rate would not allow differentiation between the separate strains, as explained earlier. This excludes values larger than  $10^{-3}$ , as that is a value where the number of end strains for  $\alpha = 0.6$  is almost at maximum as shown in figure 6.18.

The mutation rate chosen will be  $\mu = 10^{-4}$ . The choice of this value is a bit arbitrary. But, it was chosen because it fulfills the above requirements as shown in this section.

### Summary

It has been shown that the mutation component enables strains below  $\alpha < 0.65$  to survive. This results in marginally higher chance of becoming endemic for  $\alpha$  between 0.57 and 0.6. Moreover, the number of strains increases the proportion of people infected around these values. However, the mutation component has almost

no effect on the proportion infected for larger infection rates. The reason for this, is the fact that mutation ensures that a single strain does not completely become extinct, but individual strain only become extinct with good probability for lower  $\alpha$  values. If the mutation rate becomes very large the distinction between strains become less pronounced. At  $\mu = 10^{-3}$  a maximum increase in the proportion infected was found to be 0.18 at  $\alpha = 0.6$ . A mutation rate of  $\mu = 10^{-4}$  has been chosen to be used going forward to reduce the parameter space.

## 6.5 Cross immunity

In this section the final component will be added to the model and examined. This is the most complicated component as it greatly increases the parameter space. It is crucial not to overcomplicate the problem, and instead analyse the idea of cross immunity with brevity in mind. Throughout the section, we will mostly examine systems with two antigen in each strain.

For a host to recover from a malaria infection, a host has to have all antibodies for all antigens in a given strain. For example, if a malaria strain with two antigens infects a host with no antibodies, the host has to spend time to get antibodies for both of the antigens. Only thereafter is the host able to recover.

Cross immunity emerges when there exist mutual antigens between individual strains. When a host gains antibodies for antigens in a particular strain, which shares some of its antigens with another strain, the host already have antibodies in its memory to combat the other strain. This ultimately result in a lower mean time to recover. The impact of shared antigens among strains is exactly what will be examined.

First, we will recalculate  $\gamma_{optimal}$  and talk about the system from the information we have gained in the previous sections. Then, two almost identical systems will be compared and analysed, with the only difference being that in one system there are mutual antigen between strain and in the other there is not. This makes for easier evaluation of the difference cross immunity can do. This will be done both by examining the proportion infected over a range of  $\alpha$  values. Lastly, a few other simulations with a greater number of strains will be done.

To be able to distinguish the individual strains and antigen, each antigen is denoted by a number. Every strain contains multiple of these numbers, each being a antigen defining that particular strain. For example, consider a system with two strains each containing two antigens. One strain might contain antigens denoted as 1 and 2 and the other strain could have 2 and 3, just as the example given in chapter 5. Such a system would be written in the following way:

$$\text{strain}[1] = [1, 2], \quad \text{strain}[2] = [2, 3]. \quad (6.28)$$

The same notation is used for hosts and the antibodies they have. A host having antibodies for antigens 1 would be denoted as follows:

$$\text{Host}[i] \text{ antibodies} = [1] \quad (6.29)$$

If the host is infected by the first strain, and is selected in a resistance event, he would gain one the following antibody memory.

$$\text{Host}[i] \text{ antibodies} = [1, 2] \quad (6.30)$$

It would then require another two additional resistance event to completely get rid of the infection, but the host would still have the antibodies in memory.

This notation will be used throughout this section.

A consequence of the increased number of antigens in each strain, is that it takes longer time for a host to recover if the given host does not have any relevant antibodies. The new mean time for recovery is proportional to the number of

		Antigens in each strain					
		1	2	3	4	5	6
...		2/ $\beta$	3/ $\beta$	4/ $\beta$	5/ $\beta$	6/ $\beta$	7/ $\beta$
0 antibodies		1/ $\beta$	2/ $\beta$	3/ $\beta$	4/ $\beta$	5/ $\beta$	6/ $\beta$
1 antibodies		—	1/ $\beta$	2/ $\beta$	3/ $\beta$	4/ $\beta$	5/ $\beta$
2 antibodies		—	—	1/ $\beta$	2/ $\beta$	3/ $\beta$	4/ $\beta$
3 antibodies		—	—	—	1/ $\beta$	2/ $\beta$	3/ $\beta$
4 antibodies		—	—	—	—	1/ $\beta$	2/ $\beta$

**Table 6.1:** Table showing the average time to recover from a malaria strain as a function of number of antigens in each strain (column) and number of antibodies a given host has against antigens in that strain (row). The mean time follows the simple equation of  $1 + \text{antigens in strains} - \text{relevant antibodies}$ . Also recall that  $\beta = 1$  and is constant in all cases.

antigen in the strain, because it requires that many additional events to recover. An overview of the mean time with for various values of antigens and antibodies are shown in table 6.1.

This has a great affect on the survivability of malaria. The increased time for a host to recover, allows for malaria to become endemic at lower infection rates.  $R_0$  increases when the number of antigens is increased. The new calculation of the base  $R_0$  (as defined in the start of this chapter) is

$$R_0 = \frac{\alpha \cdot (L + 1)}{\beta} \quad (6.31)$$

where  $L$  is the number of antigens in each strain. In a system having strains with 2 antigens it is possible for malaria to become endemic as low as an infection rate of  $\alpha = 1/3$ . This is a change from  $\alpha = 1/2$  to  $\alpha = 1/3$ , compared to a system with only one antigen in each strain. However, this is only the potential increase, you also have to take into account that people will have more antibodies as a whole. This fact becomes more significant if the system contain shared antigens among the strains.

A consequence of the adjusted base  $R_0$ , is that the equation of the optimal replacement rate given in 6.24 is not true. That equation was computed for one antigen in each strain. For two antigens the optimal replacement rate is negative between  $\alpha = 1/3$  and  $\alpha = 1/2$ , which should not be the case, as  $R_0$  is above 0 for  $\alpha > 1/3$ .

It is possible to approximate a new optimal replacement rate. An adjusted replacement rate is used, shown in table 6.2, which takes the addition of antigens into account. It is also possible to generalise the formulae of 6.24 to any replacement rate. The calculations for this are done in MATLAB and is shown in the appendix. It gives the equation

$$\gamma_{\text{optimal}}(\alpha, \beta_{\text{adjusted}}) = (2 \cdot \alpha \cdot \beta_{\text{adjusted}}^2)^{1/3} - \beta_{\text{adjusted}}, \quad (6.32)$$

where  $\beta$  will be the new adjusted  $\beta$ . This equation will be used for replacement rates going forward.

A note about the adjusted  $\beta$ , is that the value does not actually change. The value will be kept constant at 1. The adjusted  $\beta$  is another way to put, that  $R_0$

Surface Features:	1	2	3	4	5	6
Adjusted $\beta$ :	$\frac{2}{2}$	$\frac{2}{3}$	$\frac{2}{4}$	$\frac{2}{5}$	$\frac{2}{6}$	$\frac{2}{7}$

**Table 6.2:** Table showing adjusted  $\beta$  values for equation 6.32 as a function of the number of antigens in a strain.

Name	$\beta$	$\gamma$	$\mu$	Strains
Cross:	1	Eq.6.32	$10^{-4}$	[1, 2], [2, 3], [3, 4], [4, 1]
No cross:	1	Eq.6.32	$10^{-4}$	[1, 2], [3, 4], [5, 6], [7, 8]
Cross big:	1	Eq.6.32	$10^{-4}$	[1, 2], [2, 3], [3, 4], [4, 5], [5, 6], [6, 1]
No cross big:	1	Eq.6.32	$10^{-4}$	[1, 2], [3, 4], [5, 6], [7, 8], [9, 10], [11, 12]
Odd:	1	Eq.6.32	$10^{-4}$	[1, 2], [2, 3], [3, 4], [4, 5], [5, 1]
Simple:	1	Eq.6.32	$10^{-4}$	[1, 2], [3, 4],

**Table 6.3:** Parameters for simulations done in this section. The strains column indicate the strains that exist with which antigens in those systems.

changes to keep earlier derivations consistent. The equation can be written as a function of the base  $R_0$  instead:

$$\gamma_{\text{optimal}}(\alpha, R_0, L) = \left( 2 \cdot \alpha \cdot \left( \frac{\alpha \cdot (L+1)}{R_0} \right)^2 \right)^{1/3} - \frac{\alpha \cdot (L+1)}{R_0}. \quad (6.33)$$

It is difficult to gauge the effects of cross immunity. There exist many different combinations of strains when increasing the number of antigens in each strain. It is infeasible to make a parameter sweep for every kind of combination of strains and antigens. If there exist 6 different antigens and each strain contain two antigens, there exist  $\sum_{i=0}^{A-1} i = 15$  unique strains, which can be generalised to

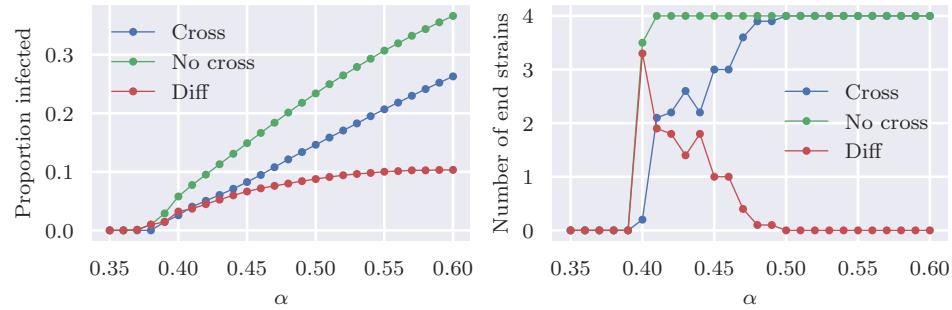
$$N = \sum_{i=0}^{A-1} i, \quad (6.34)$$

where  $A$  is the number of unique antigens. For three antigens in each strain then number becomes

$$N = \sum_{i=0}^{A-2} 3i + 1. \quad (6.35)$$

Because of these large numbers specific systems will have to chosen for examination.

To start, two similar systems will be tested. The two systems will have the same number strains and the same number of antigens in each strain. In one system no two strains will have the same antigens, but in the other all will share at least one antigen with another. This way, we will have a system where cross immunity is



**Figure 6.22:** Plots of the proportion infected and strains surviving in the system at end of simulation as a function of  $\alpha$  for two systems with four unique strains, but containing a different set of antigens. The system names are written in table 6.3. Diff is the absolute difference between the two systems.

present and another where it is not. The specifications of the systems is written in table 6.3. The names written in the table will be used throughout the section for reference. For now, the two relevant system are the ones with the names of *cross* and *no cross*. The *cross* system has four strains, where each strain shares one antigen with two other strains comparable to a one dimensional neighbourhood. In *no cross* no antigens are shared between strains.

The proportion infected and number of strains of the two system are shown in figure 6.22. There is a very clear difference between the two systems. For  $\alpha < 0.4$  malaria does not become endemic. For  $\alpha = 0.4$  it does in the *no cross* system but not in *cross*. So, it is slightly more probable for malaria to become endemic, if there is no cross immunity present.

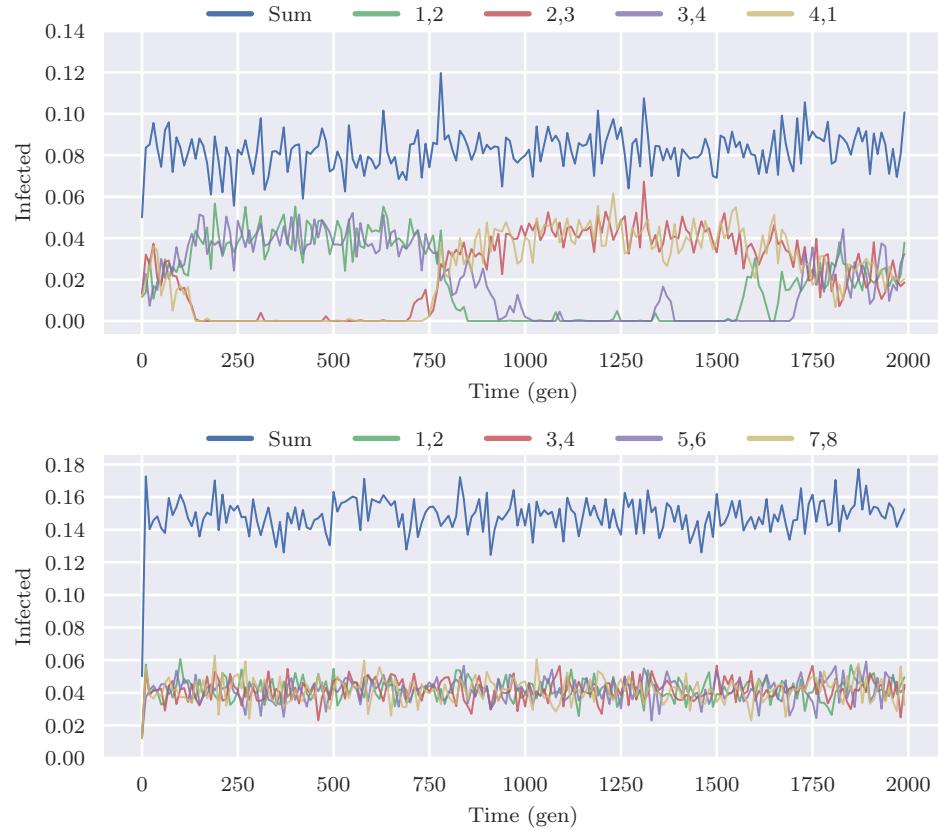
The proportion infected is greater for the *no cross* for  $\alpha > 0.4$  compared to *cross*. Essentially, cross immunity makes is a detriment for malaria. The difference between the two can be as large as 0.1.

This is further indicated by examining the number strains at end of simulation. There exist more strains for *no cross* between  $\alpha = 0.41$  and  $\alpha = 0.47$ . In this region the average number of strains at the end is lower than four strains unlike the system with no shared antigens.

Let us take further look into the course of these discrepancies. Two time-series plots for  $\alpha = 0.45$  have been plotted in figure 6.23 where the upper one shows the system of *cross* and the lower one shows the system of *no cross*.

*No cross* is identical to some the previous time-series plots, for example the lower one in figure 6.21. Every strain co-exists and malaria is in no risk of becoming extinct. They all share an equal amount of the population when correcting for random fluctuations. The system is practically identical to simulations in section 6.4 and are indistinguishable for large infection rates, when correcting for the additional antigens.

The *cross* time-series shows an interesting property. The system shows bistability. Two of the strains co-exist together while the other two strains are extinct. The strains co-exist in pairs, where the pairs are always the two that do not share any antigens. The dominating strains maximises the number of antigens be-

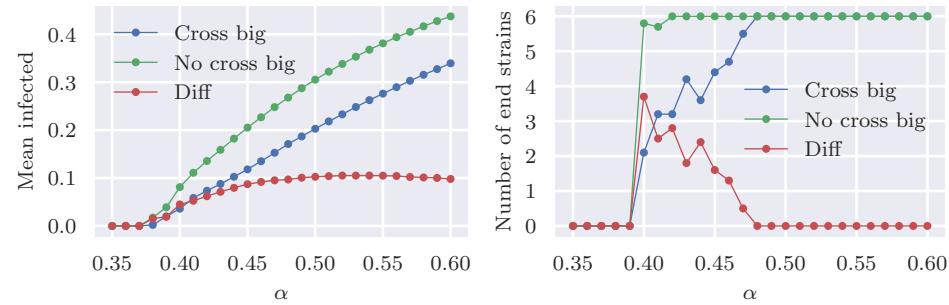


**Figure 6.23:** Time-series plots showing the proportion infected for each strain and the total number of infected. The infection rate is  $\alpha = 0.45$ . The upper plot is for the system with cross immunity and the lower one is the one without. The antigens of each strain is indicated in the legends. Note the difference on the y axis for the plots. Bi-stability is apparent in the upper plots, with switches between strain pairs happening at  $t = 800$  and  $t = 1750$ .

tween them. Strains [1, 2] and [3, 4] have four unique antigens between them while [1, 2], [2, 3] only have four. This maximises the mean time to recover, as there is more unique antigens in the system to build antibodies against. Thus, such pairing occurs.

This property emerges consistently as long as the infection rate is not large. When all strains can co-exist independently not pairing occurs. This happens for values of  $\alpha > 0.46$  as shown in 6.22 by the fact that the number of strains is almost at 4. The pairing can only happen if the infection rate is low enough, such that malaria is dependent on hosts having vulnerabilities in their immune-system for being endemic. Otherwise, there would be no reason why each strain could not co-exist. The pair dominating is decided at random.

Two systems with four strains have been examined. What happens when additional strains are put into the system? In figure 6.24 plots of a parameter sweep



**Figure 6.24:** Plots of the mean infected and strains left in the system at end of execution as a function of  $\alpha$  for two systems with six unique strains but containing different antigens. The systems are shown in the legends and are referred in table 6.3. Diff is the absolute difference between the two systems.

over alpha for *cross big* and *non cross big* are shown where each system has six strains. A time-series of those are shown in the appendix for  $\alpha = 0.45$ . This is also the case for *odd* and for a few other simulations with lower  $\alpha$  (details of parameters in table 6.3).

The same behaviour occurs as in *cross* and in *non-cross*. The proportion infected is larger across the board, and malaria is able to become endemic for  $\alpha > 0.38$ ; increasing the range where malaria becomes consistently endemic by at least 0.01.

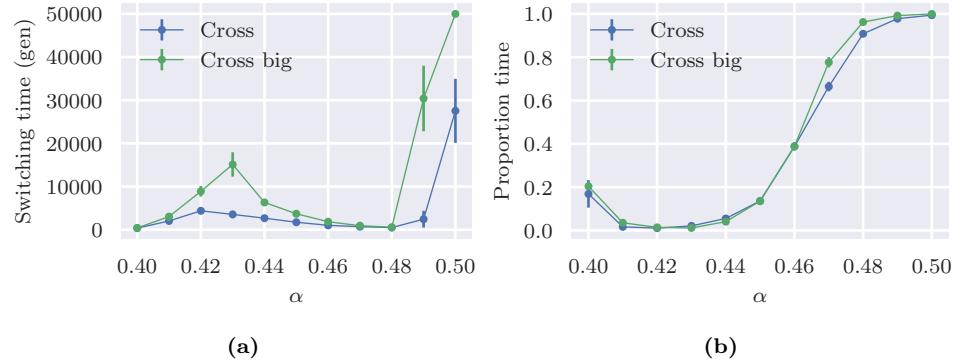
Bi-stability also happens in this case, but the switches occur less often. This is because it is less probable that three extinct strains re-emerge compared to two strain re-emerging. Time-series plots with these parameters can be found in the appendix, in addition to the setup named *odd* and time-series' for other  $\alpha$  values.

The switching between dominating strain pairs happen because of strains ability to mutate. A stroke random event happen that leads to non-dominating strains to re-emerge, as was explained in section 6.4.

An interesting measure is how long it takes for a switch to occur. This would give an idea how stable the system is as a function of  $\alpha$ .

The switching time for *cross* and *crossbig* is shown in figure 6.25a. As seen, the switching time depends on alpha, but it is not a linear function. The switching time is extremely small for  $\alpha = 0.40$ . As  $\alpha$  approaches 0.42 the switching time increases. Afterwards, the switching time suddenly begins to decrease as it reaches its minimum at 0.47 with an average switching time of 619 generations. After this it steeply increases. The switching time is larger for *cross big* for all infection rates. The smallest deviation is at  $\alpha = 0.48$  which has about 1 standard deviation of separation with switching times of 499 and 570 and standard deviations of 44 and 58 for *cross* and *cross big*, respectively. The number of data points used for finding these values are given in the appendix in table 10.1

At some  $\alpha$  values the system is mostly in a state where all strains are surviving. That is to say a state where no pairs are dominating. The proportion time spent in this state is shown in figure 6.25b. It shows that most of the time above  $\alpha = 0.48$  is spent in a non-dominating state. This makes sense, as the high infection rates makes sure that everyone can co-exist. This explains the large switching times for  $\alpha > 0.48$ .



**Figure 6.25:** (a): Mean switching time between  $\alpha = 0.40$  and  $\alpha = 0.50$ . Malaria does not consistently become endemic at infection rates lower than 0.40 and switching time becomes very large at infection rates higher than 0.5, which is the reason why they have been omitted. The switching time was found from 5 repeated cases with a running time of  $2 \cdot 10^6$  generations. (b): Proportion of time spend in a state where all strains are endemic and relevant and no pairs are dominating. The error-bars show the error on the mean. The number of data points used to find the switching time is given in the appendix in figure 10.1.

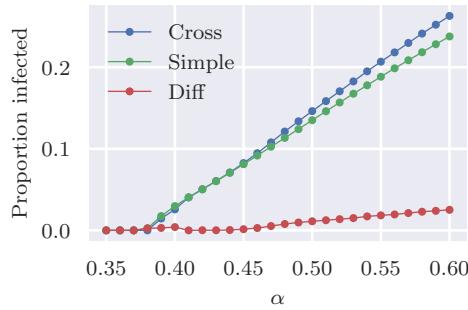
The lowest switching time appears at the largest  $\alpha$  values before all strain co-exist the greatest proportion of time. Larger  $\alpha$  values helps a non-dominating state to become present again, because it increases the probability of a non-dominant strain to gain many infections in a short time, which is requirement for re-emerging as explained in section 6.4. The time spent in a state where a pair is dominating is small, as it is four times as probable to be in a state where all strains exist. This explains why the lowest switching time is found at  $\alpha = 0.47$ . *Cross* and *cross big* approximately spends the same time where one pair is dominating except at  $\alpha = 0.4$ .

At  $\alpha = 0.4$  the switching time is very low, despite the argument presented above. The reason for this, is because malaria is not consistently endemic at this value. That results in instability through large relative fluctuations between strains. See the time-series for this value in figure 10.2 in the appendix as an example.

Compared to real life malaria this be-stable behaviour may not be entirely unrealistic. If  $R_0$  is very large all strains can live in co-existence despite their inner competition. If  $R_0$  is low but within survival, strains start to compete among themselves. Strains that has the most diverse antigens has the greatest chance to survive. However, as individuals become resistant to certain strains, other strains have the chance to become dominant.

The plots of figure 6.25 also shows that additional strains increases the switching time. This makes sense, as all the paired strains has to collectively encroach on the opposing territory.

In the *cross* system, only four unique antigens are in the system. In comparison, the *non cross* system has eight unique antigens. This has the consequence that hosts has many more possible antigens to build antibodies against. It is therefore not strange that the *non – cross* system has a higher proportion of the population infected. The *cross* system may be more similar to a system which have the same



**Figure 6.26:** Mean infected for cross and simple systems and their difference. The simulation parameters are shown in table 6.3.

number of unique antigens in the system. The *cross* system is compared to the *simple* system which is written in table 6.3.

Simulations do indeed indicate that the *cross* system is more similar to the *simple* system than the *no cross* system. This is shown in figure 6.26 where *cross* and *simple* are compared. The proportion infected between the systems are almost the same, with a slight divergence starting at  $\alpha = 0.46$ , where *cross* increases slightly more than *simple*. The reason for this difference is because of super-infection. In *cross* each host can have up to four infections at a time, while they can only have two in *simple*. This increases the proportion of the population that are able to be infected and therefore increases the chance for an infection event to happen.

From this, it can be gathered that one of the main parameters that guide the number of infected, is not only how strains are interconnected, but how many unique antigens there exist in the system. This shows that Malaria has a lot to gain by maximising its antigenic diversity.

### Summary

In this section, the number of antigens in each strain was increased. This allowed for shared antigens between strains and therefore cross immunity could occur. The increased number of antigens changed  $R_0$  and a new approximate optimal  $\gamma$  was calculated. Six different systems was simulated, analysed, and compared; some with shared antigens and some without. The systems with shared antigens always had a lower number of infected compared to similar systems and had a slightly shorter range of survival. It was shown, that the number of unique antigens in the system had a greater impact than the number of strains in the system. Moreover, the system with cross immunity had dominating pairs which switched over time in a small region of  $\alpha$ . The dominating pairs always maximised the number of unique antigens. The mean switching time was found for two systems, and was highly dependent on both the number of strains and on  $R_0$ . A main point to take from this investigation, is it is very advantageous for malaria to have a large antigenic diversity.

# Chapter 7

## Discussion

In this chapter, a discussion, perspective, and further examination of the model will be presented. An examination will be done on the consequences of the choices of certain parameters and what insights could be gained from other parameter sweeps. Arguments for the specific choices of research will be presented. We also discuss the perspective on the topic with particular emphasis on insights acquired from the model and their application to malaria dynamics around the globe. Finally, some of the complications of the models will be inspected, and a suggestion will be made of an alternative model that could solve some of these complications.

### 7.1 Other investigations

The model of this thesis is complex and have many parameters to tune. For this reason, there are many areas for potential research. In this section, we outline new scientific inquiries one could pursue with respect to the model.

#### 7.1.1 Number of hosts

In section 6.1 it was suggested that the number of hosts had an influence on the stochasticity of the system. Increasing the number of hosts would decrease the stochasticity of the system, and increase malarias probabilities of becoming endemic. This fact could be investigated by varying the number of hosts and keeping other parameters constant. It could be of interest to get an idea of how much the system changes from this - specifically the increase or decrease of the basic reproductive rate ranges in which malaria becomes endemic<sup>1</sup>. This will give more insight in how big the population number has for the survival of malaria.

Further more, it was postulated that decreasing the number of hosts would decrease switching time for systems which include cross immunity. This would also be of interest to investigate.

---

<sup>1</sup>Recall from section 6.2 that malaria was endemic for  $\alpha > 0.5$  in the deterministic case, while malaria was endemic for  $\alpha > 0.6$  in the stochastic case

### 7.1.2 Adjusting replacement

Replacement has throughout this thesis been a relatively unexplored parameter, as an approximate optimal  $\gamma$  value was calculated and used. This idea simplified the parameter space and normalised the systems when sweeping over  $\alpha$ . However, by doing this, the system behaved a specific undetermined way.

The optimal  $\gamma$  might not be entirely realistic. It is true that there are high birth rates and high mortality rate in areas where malaria is most severe [1, 49]. The chosen replacement rate might have been very large compared to the other parameters. For example, in a system with  $\alpha = 0.7$  and one antigen per strain the corresponding value is  $\gamma_{\text{optimal}} = 0.119$ , which means that the infection event does at most have a 6 times higher rate than the replaced event. In other words, for every sixth individual being infected with malaria an individual is removed and replaced with a new individual.

High replacement decrease the difference there is between systems with cross immunity and no cross immunity, as only a low portion of the population remember it's antibodies for a long time. Essentially, the replacement component may dominate the dynamics rather than the resistance component. This should be avoided, as it makes it more difficult to find the effects of cross immunity. Therefore, it is of interest to study the systems at other replacement rates than the ones tested.

The calculated optimal replacement rate assumes that there is only a single strain in the system. It does not take into account the introduction of additional strains combined which increases the true  $R_0$ . This assumption makes for fine approximation, but makes it so that we are not working with a true optimal gamma. This distorts the comparison between systems with different number of strains.

Comparisons of lower replacement rate is shown in the appendix in figure 10.6, as an example of the differences between lower and higher replacement rates.

### 7.1.3 Super-infection

Super-infection is something that has been measured in infected individuals [44, 45, 57] and has been included in other malaria models<sup>2</sup> [12, 55]. This is why we included super-infection in the model of this thesis. However, super-infection is another feature that complicates the model, and which functions we do not fully understand. It could for example reveal whether the two systems compared in figure 6.26 are actually the same as hypothesised.

### 7.1.4 Initial conditions

The choice of initial condition is something we have only lightly touched upon. The initial conditions have little effect on the behaviour of the system, and no effect if in steady state, as long as there are enough infected hosts to start the initial transmissions. The most significant exception is for systems with low  $R_0$ , as having more initial infected will increase the time till extinction. Initial conditions is therefore not something that have potential for new insights.

---

<sup>2</sup>Macdonald for example

### 7.1.5 More antigens

In this thesis, the model has only been examined for 1 and 2 antigens for each strain. There is good reason why it was not expanded to larger values. The model becomes increasingly complex as more antigens and strains are added to the system. The many ways to combine antigens between strains increases potentially as explained in section 6.5. Moreover, there are still insights to be gained by investigating the system for two antigens per strain.

Despite this, it would be of interest to examine the system for higher antigen per strain values. New dynamics and properties could emerge which simply does not occur for lower antigens per strain values. It also allows for new combinations of antigens between strains. Moreover, there might be many antigens for each individual strain in the malaria parasite [35, 42, 44] in real life. Higher values is possibly easier to apply for real life applications. For these reasons, it would be an interesting point to further research.

## 7.2 Data

A convincing model should be able to approximately simulate real life. Testing the model against real life data gives models legitimacy. It also makes it possible to adjust parameters to realistic values and potentially reduces the parameter space. The model presented in this thesis has not been tested against data, which is an inquiry of priority if the model should be further worked upon.

Unfortunately, good data for testing is not easy to get. Optimal data requires measurements on hosts to be done consequently on a lot of individuals in a relatively isolated region which is both time consuming, difficult, and expensive. It was not possible to get data for this project. Dietz [23] tested his malaria model with data from the African Savannah [13, 25].

## 7.3 Real-Life applications

One of the principal insights of the model is the advantage of antigenic diversity. For Plasmodium to gain great diversity, it has to have a strategy to generate new antigenic features. A change in a single antigen of a wide spread malaria strain will quickly become extinct as many people in the population will have most of required antibodies against the newly spawned strain. Many mutations must happen before a particular strain diversified itself enough to be competitive. Moreover, the new strain also has to become lucky by spreading to a set of new vulnerable hosts. This leads to very distinct malaria strains dominating the system. As the population builds antibodies for those strains, new strains with unknown antigens has the possibility to dominate the system. This synchronisation between strains was also found by Macqueen et. al. [26] which further indicates this sort of behaviour. Specifically, it was found that the antigenic diversity that determines the competitiveness of malaria, not the number of strains, as illustrated in section 6.5.

A feature of this model is its ability to simulate the effects of malaria vaccines. Malaria vaccines generally give antibodies against specific antigens. This essentially results in that a vaccine only helps against a fraction of malaria strains

[42]. Vaccinations in this model would be modelled by giving antibodies against a specific set of antigens in a proportion of the population. If model parameters can be estimated and malaria diversity identified, then this model could help estimate how many should be vaccinated to eradicate malaria or estimate the reduction of infectiousness in malaria after a vaccination campaign.

The greatest difficulty in parameter estimation is mapping the antigenic diversity of malaria and correctly simulating that in the model. If the model can approximate the diversity a difficulty still lies in the computational speed of the model, especially when considering that the population of 10,000 used in this thesis would have to be drastically increased to be truly representative.

## 7.4 Model problems and alternative model suggestion

The main purpose is to understand the dynamic interplay between antigenic diversity and cross immunity. This is modelled through strains sharing antigens between them, and by giving each host their separate antibody memory. There are some problems with this approach which are outlined below.

The model presented here can be considered complex, specifically because of how the antigenic diversity is modelled. There namely are a staggering amount of permutations, and that makes it difficult to chose how to exactly define each strain. For example, if one wanted to strains to share 2/3 of antigens, there must be at least three antigens per strain in the system.

It is bad design that increasing the antigens per strain also increases  $R_0$ , even if the antigenic diversity and other parameters remain constant. This can be corrected for, which was done when using the adjusted optimal  $\gamma$  in section 6.5, but solutions like that is very inelegant. Moreover, to create some systems with a specific set of shared antigens it can be necessary to add more antigens per strain, which in turn changes the system as it then requires another resistant event to recover. Put another way, changing the number of antigens per strain changes how the system works directly, even without changing anything else about the system. The resistance and strain system should not be dependent on each other, they should be two completely independent parameters.

In this model resistance to a particular strain reduces the time to recover. This makes sense and is in good accordance with how malaria resistance works<sup>3</sup> [10, 9]. It is problematic that the mean recovery time is a function of the number of antigens per strain. It makes it difficult to compare to other systems, as the number of antigens per strain should primarily be used to explain the diversity and shared features between strains. Moreover, the speed gained for recovery for having additional strain is only dependent on the number of antigen per strains following the values given in table 6.1.

I propose another approach to model antigenic diversity and cross immunity that solves some of the aforementioned problems. This proposed model essentially removes the way strains are defined and how resistance is build, and replaces them with a similar system. The model is still a agent based stochastic model, where

---

<sup>3</sup>With the limited knowledge there is on the subject.

hosts can be infected by zero or more malaria strains. Transmissions, replacement, and mutation also essentially function the same way.

In the new proposed model a strain is not defined by a set of numbers representing antigens. All strains are represented by one number identifying it no matter how many antigens a strain consist of. Infected hosts then build antibodies by iterative ticks using an Euler-method or equivalent [58]. At each time-step every infected host builds resistance against the strains they are infected by. A resistance to a particular strains is represented by a 0 and 1, where 1 is considered to have all relevant antibodies. Every host then has a chance to recover based on that number. The growth in resistance is random - some hosts may quickly build resistance while other do not. The growth of resistance is a parameter that can be adjusted and the probability distribution used to draw numbers can be any reasonable distribution.

Cross immunity is represented by a covariance matrix with rows and columns equal to the number of strains. Each element contain values ranging from 0 to 1, where the diagonal is always 1. Each element represent how many antigens are shared between strains. Whenever resistance is gained against one strain, resistance is also gained against other strains proportional to the values in the covariance matrix. This way, the mutual antigens between strains can easily be represented without changing anything else in the system.

The covariance matrix can be build in a number of ways. A simple matrix for example would be to have all neighbouring strains be equal to 0.5.

This alternative model provides an easy way to measure the proportion of shared antigens between strains. Mathematically, this can be done by taking the mean of the whole co-variance matrix disregarding the diagonal. The proposed model also allows to change how many and how much each strain share in antigens. It makes it very easy to create any kind of co-variance, without artificially creating a combination of strains that result in exactly the desired system.

Since there is no parameter adjusting the number of antigens per strain, base  $R_0$  does not depend on how many many antigens per strain there is in the system. The resistance component is flexible and can easily be adjusted. Even the function that governs the rate of resistance build-up can be adjusted, which was not true for the old model.

Lastly, the idea of a covariance matrix is a well-known idea and is well understood mathematically. This whole idea supposedly both makes the model easier to understand, more mathematical rigid, and more flexible.

# Chapter 8

## Conclusion

In this thesis, a new malaria model has been presented, simulated, and analysed. The model takes into account mutual diversity in antigenic features of malaria strains. The shared antigenic features results in cross immunity and changes the dynamics compared to a system without it. The model did so by defining a set of strains that shared antigens. Hosts had to build antibodies separately for each antigen to be able to gain faster recovery.

The model was studied by simplifying the model to its core components. After inspection new components were added to the model. From this, it was possible to determine an optimal replacement rate for malaria as a function of  $R_0$ . Analytical and simulated conclusions were in agreement.

The complete model was too complex to be written in a simple differential form, and limited information could be gained from analytical analysis. Analysis was therefore primarily done on stochastic simulations following the agent based rules of the model.

It was found that increasing the number of strains in the system increased the proportion of the population malaria infected, because of super-infection and because of the increased antigenic variety. It did not increase the range  $R_0$  where malaria became consistently endemic. For this, a mutation component was a necessity. Mutation allowed for extinct strain to re-emerge into the system. It was found that this changed the range of malaria becoming endemic from  $\alpha = 0.6$  to  $\alpha = 0.57$  for a mutation rate equal to  $\mu = 10^{-3}$ .

When the strains could contain multiple strains, it was shown that systems without cross immunity was more competitive than the system without. Cross immunity was always a disadvantage for malaria. It was shown that the most important factor for malaria's success was the maximisation of unique antigens in the system.

Systems with cross immunity did at some parameter values show bi-stability. Pairs of strains that minimised the shared antigens would dominate. At times the dominating pair would switch with the non-dominating pair. The switching time between dominating states was calculated for two systems. It was shown it was greatly dependent on the parameter of infection rate  $\alpha$ . Adding more strains to the system increased switching time.

It was suggested that this model could help with gauge the effects of vaccines in malaria. Both by giving insights in a system containing diverse malaria strains and

by being able to simulate vaccines targeting specific antigens. With adjustments, the model show promise to be able to gain insights of malaria.

The model have some complications that could be improved upon. Firstly, the model is complex. Secondly, increasing the number of antigens per strain made the system unpredictable in undesired ways. Lastly, the way replacement and transitioning from non-resistant to resistant states was implemented was very inflexible. An alternative model was proposed which could solve some of these problems and still contain the desired features. Despite this, the underlying mechanics of the model show prospect for further inquiry.

## **Chapter 9**

# **Acknowledgements and References**

Thanks to my super-visor Kim Sneppen for supporting me throughout the project. Also thanks to Aleksandar, Alexander, Anesten, Gustav, Henrik, Isabella, Marie, Mathias, Niels, Svend, Troels, and Zlatko for helping to correct parts of the thesis.

# Bibliography

- [1] *World Malaria Report 2017*. WHO, 2017. ISBN: 978-92-4-156552-3.
- [2] *Malaria*. <http://www.who.int/news-room/fact-sheets/detail/malaria>. June 2018.
- [3] Gollin Douglas and Zimmerman Christian. “Malaria: Disease Impacts and Long-Run Income differences”. In: *IZA DP* 2997 (2007).
- [4] “Malara Prevention Works - Let’s close the gap”. In: *WHO* (2017).
- [5] Fernando SD, Rodrigo C, Rajapakse S. “Chemoprophylaxis in malaria: Drugs, evidence of efficacy and costs”. In: *Asian Pacific Journal of Tropical Medicine* 4.4 (2011), pp. 330–336. DOI: [https://doi.org/10.1016/S1995-7645\(11\)60098-9](https://doi.org/10.1016/S1995-7645(11)60098-9).
- [6] Howitt P et al. “Technologies for global health”. In: *The Lancet* 380.9840 (2012), pp. 507–535. DOI: [10.1016/S0140-6736\(12\)61127-1](https://doi.org/10.1016/S0140-6736(12)61127-1).
- [7] Sabot O et al. “Costs and financial feasibility of malaria elimination”. In: *Lancet* 376.9752 (2010), pp. 1604–1615. DOI: [10.1016/S0140-6736\(10\)61355-4](https://doi.org/10.1016/S0140-6736(10)61355-4).
- [8] McKenzie F.E. “Why Model Malaria?” In: *Parasitology Today* 16.12 (2000), pp. 511–516. ISSN: 0169-4758. DOI: [https://doi.org/10.1016/S0169-4758\(00\)01789-0](https://doi.org/10.1016/S0169-4758(00)01789-0). URL: <http://www.sciencedirect.com/science/article/pii/S0169475800017890>.
- [9] Peter C Bull et al. “Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria”. In: *Nature medicine* 4.3 (1998), p. 358.
- [10] Denise L. Doolan, Carlota Dobaño, and J. Kevin Baird. “Acquired Immunity to Malaria”. In: *Clinical Microbiology Reviews* 22.1 (2009), pp. 13–36. ISSN: 0893-8512. DOI: [10.1128/CMR.00025-08](https://doi.org/10.1128/CMR.00025-08). eprint: <https://cmr.asm.org/content/22/1/13.full.pdf>. URL: <https://cmr.asm.org/content/22/1/13>.
- [11] Ross, Ronald Sir. *The Prevention of Malaria*. E. P. Dutton and Company, 1910.
- [12] Macdonald G. *The Epidemiology and Control of Malaria*. Oxford University Press, London, 1957.
- [13] Sandip Mandal, Ram Rup Sarkar and Somdatta Sinha. “Mathematical model of malaria - a review”. In: *Malaria Journal* 10:202 (2011), p. 19.

- [14] A. G. McKendrick W. O. Kermack. “A contribution to the theory of epidemics”. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 115.772 (1927), pp. 700–721. ISSN: 0950-1207. DOI: 10.1098/rspa.1927.0118. eprint: <http://rspa.royalsocietypublishing.org/content/115/772/700.full.pdf>. URL: <http://rspa.royalsocietypublishing.org/content/115/772/700>.
- [15] Smith et al. “Ross, Macdonald, and a Theory for the Dynamics and Control of Mosquito-Transmitted Pathogens”. In: *Chitnis CE, ed. PLoS Pathogens* 8(4).e1002588 (2012). DOI: 10.1371/journal.ppat.1002588.
- [16] Mandal Sandip, Ram Rup Sarkar, and Sinha Somdatta. “Mathematical models of malaria - a review”. In: *Malaria Journal* 10.202 (2011).
- [17] D. Bernoulli. “Essai d'une nouvelle analyse de la mortalité causée par la petite vérole (An attempt at a new analysis of the mortality caused by smallpox and the advantages of inoculation to prevent it)”. In: *Mém. Math. Phys. Acad. Roy. Sci., Paris* 1 (1766).
- [18] Ernesto Capanna. “Grassi versus Ross: who solved the riddle of malaria”. In: *International Microbiology* 9.1 (2006), pp. 69–74.
- [19] Gu, Weidong et al. “An individual-based model of Plasmodium falciparum malaria transmission on the coast of Kenya”. In: *Malaria Journal* 97 (2003), pp. 43–50.
- [20] Martens P and Hall L. “The population dynamics of malaria”. In: *Emerg Infect Dis* 6 (2000), pp. 103–109.
- [21] Paajimans KP, Read AF, and Thomas MB. “Understanding the link between malaria risk and climate”. In: *Proc Nat Acad Sci USA* 106 (2009), pp. 13844–13849.
- [22] Jetten TH, Martens WJM, and Takken W. “Model simulations to estimate malaria risk under climate change”. In: *J Med Entomol* 33.3 (1996), pp. 361–371.
- [23] Dietz K., Molineaux L., and Thomas A. “A malaria model tested in the African savannah”. In: *Bull. Org. mond. Santé, Bull. Wld Hlth Org.* 50 (1974), pp. 347–357.
- [24] Aron JL and May RM. “The population dynamics of malaria”. In: *London: Chapman and Hall* (1982), pp. 139–179.
- [25] Molineaux L. and Grammiccia G. *The Garki Project*. WHO, 1980. ISBN: 92-4-156061-4.
- [26] Philip G McQueen and F Ellis McKenzie. “Host control of malaria infections: constraints on immune and erythropoietic response kinetics”. In: *PLoS computational biology* 4.8 (2008), e1000149.
- [27] MB Hoshen et al. “Mathematical modelling of the within-host dynamics of Plasmodium falciparum”. In: *Parasitology* 121.3 (2000), pp. 227–235.
- [28] Gupta Sunetra and Galvani Alison. “An individual-based model of Plasmodium falciparum malaria transmission on the coast of Kenya”. In: *Phil. Trans. R. Soc. Lond. B* 354 (1999), pp. 711–719.
- [29] Gabriel Zorello Laporta et al. “Malaria vectors in South America: current and future scenarios”. In: *Parasites & vectors* 8.1 (2015), p. 426.

- [30] Adriana Calderaro et al. “Accurate identification of the six human Plasmodium spp. causing imported malaria, including Plasmodium ovale wallikeri and Plasmodium knowlesi”. In: *Malaria Journal* 12.1 (Sept. 2013), p. 321. ISSN: 1475-2875. DOI: [10.1186/1475-2875-12-321](https://doi.org/10.1186/1475-2875-12-321). URL: <https://doi.org/10.1186/1475-2875-12-321>.
- [31] Seattle, United States: Institute for Health Metrics and Evaluation (IHME) 2017. *Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2016 (GBD 2016) Results*. <http://ghdx.healthdata.org/gbd-results-tool>.
- [32] Alessandro Bartoloni and Lorenzo Zammarchi. “Clinical aspects of uncomplicated and severe malaria”. In: *Mediterranean journal of hematology and infectious diseases* 4.1 (2012).
- [33] Marcus J Rijken et al. “Malaria in pregnancy in the Asia-Pacific region”. In: *The Lancet infectious diseases* 12.1 (2012), pp. 75–88.
- [34] Elizabeth A Ashley and Nicholas J White. “The duration of Plasmodium falciparum infections”. In: *Malaria journal* 13.1 (2014), p. 500.
- [35] Laurence Florens et al. “A proteomic view of the Plasmodium falciparum life cycle”. In: *Nature* 419.6906 (2002), p. 520.
- [36] Alan F Cowman, Drew Berry, and Jake Baum. “The cellular and molecular basis for malaria parasite invasion of the human red blood cell”. In: *J Cell Biol* 198.6 (2012), pp. 961–971.
- [37] National Institutes of Health (NIH). *Life Cycle of the Malaria Parasite*. [https://upload.wikimedia.org/wikipedia/commons/9/9a/Life\\_Cycle\\_of\\_the\\_Malaria\\_Parasite.jpg](https://upload.wikimedia.org/wikipedia/commons/9/9a/Life_Cycle_of_the_Malaria_Parasite.jpg). May 2016.
- [38] Janeway C. Garland Publishing, 2001. ISBN: 978-0-8153-3642-6.
- [39] U. S. National Library of Medicine. *Antigen*. <https://medlineplus.gov/ency/article/002224.htm>. 2005.
- [40] L Dijkshoorn, BM Ursing, and JB Ursing. “Strain, clone and species: comments on three basic concepts of bacteriology”. In: *Journal of medical microbiology* 49.5 (2000), pp. 397–401.
- [41] Antoine Claessens et al. “Generation of antigenic diversity in Plasmodium falciparum by structured rearrangement of Var genes during mitosis”. In: *PLoS genetics* 10.12 (2014), e1004812.
- [42] Kirsten Flick and Qijun Chen. “var genes, PfEMP1 and the human host”. In: *Molecular and Biochemical Parasitology* 134.1 (2004), pp. 3–9. ISSN: 0166-6851. DOI: <https://doi.org/10.1016/j.molbiopara.2003.09.010>. URL: <http://www.sciencedirect.com/science/article/pii/S0166685103003207>.
- [43] Fvasconcellos and National Human Genome Research Institute. *Antibody*. <https://commons.wikimedia.org/wiki/File:Antibody.svg>. May 2007.
- [44] Andrew A. Lover and Richard J. Coker. “Do mixed infections matter? Assessing virulence of mixed-clone infections in experimental human and murine malaria”. In: *Infection, Genetics and Evolution* 36 (2015), pp. 82–91. ISSN: 1567-1348. DOI: <https://doi.org/10.1016/j.meegid.2015.08.041>. URL: <http://www.sciencedirect.com/science/article/pii/S1567134815003640>.

- [45] Hill, W. G. and H. A. Babiker. “Estimation of numbers of malaria clones in blood samples”. In: *Proc. R. Soc. Lond. B Biol. Sci.* 262.1365 (1995), pp. 249–257. DOI: 10.1098/rspb.1995.0203.
- [46] Rick Fairhurst, Jordan Zuspann, National Institute of Allergy, and Infectious Diseases, National Institutes of Health. *Red blood cells infected with malaria*. <https://www.flickr.com/photos/nihgov/25534997493/in/photolist-EUrx8t-CvR53a-B3Ad52-ydGygr-wZzPff-C5BN5H>. Mar. 2016.
- [47] Murray, J. D. *Mathematical Biology I: An Introduction, Third Edition* (p. 315-395). Springer, 2002. ISBN: 0-387-95223-3.
- [48] Michael T Heath. *Scientific computing*. McGraw-Hill New York, 2002. ISBN: 978-0072399103.
- [49] J.M Heffernan, R.J Smith and L.M Wahl. “Perspectives on the basic reproductive ratio”. In: *Journal of the Royal Society Interface* 2.4 (2005). DOI: 10.1098/rsif.2005.0042.
- [50] Herbert W. Hethcote. *Three Basic Epidemiological Models*. Ed. by Simon A. Levin, Thomas G. Hallam, and Louis J. Gross. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989. ISBN: 978-3-642-61317-3. DOI: 10.1007/978-3-642-61317-3\_5. URL: [https://doi.org/10.1007/978-3-642-61317-3\\_5](https://doi.org/10.1007/978-3-642-61317-3_5).
- [51] Ross, Ronald Sir. “Some a priori pathometric equations”. In: *The British Medical Journal* 1.2830 (1915), pp. 546–547.
- [52] P. van den Driessche and James Watmough. “Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission”. In: *Mathematical Biosciences* 180.1 (2002), pp. 29–48. ISSN: 0025-5564. DOI: [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6). URL: <http://www.sciencedirect.com/science/article/pii/S0025556402001086>.
- [53] Martin Gardner. “Mathematical games: The fantastic combinations of John Conway’s new solitaire game “life””. In: *Scientific American* 223.4 (1970), pp. 120–123.
- [54] Daniel T. Gillespie. “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions”. In: *Journal of Computational Physics* 22.4 (1976), pp. 403–434. ISSN: 0021-9991. DOI: [https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3). URL: <http://www.sciencedirect.com/science/article/pii/0021999176900413>.
- [55] G. Macdonald et al. “Epidemiological Basis of Malaria Control”. In: *Bull World Health Organ* 15 (1956), pp. 613–626.
- [56] Nicole Mideo, Troy Day, and Andrew F Read. “Modelling malaria pathogenesis”. In: *Cellular microbiology* 10.10 (2008), pp. 1947–1955.
- [57] J. C. Anderson et. al. Timothy. “Microsatellite Markers Reveal a Spectrum of Population Structures in the Malaria Parasite Plasmodium falciparum”. In: *Molecular Biology and Evolution* Volume 17.Issue 10 (2000), pp. 1467–1482. DOI: <https://doi.org/10.1093/oxfordjournals.molbev.a026247>.
- [58] P. E. Kloeden and R. A. Pearson. “The numerical solution of stochastic differential equations”. In: (1976).

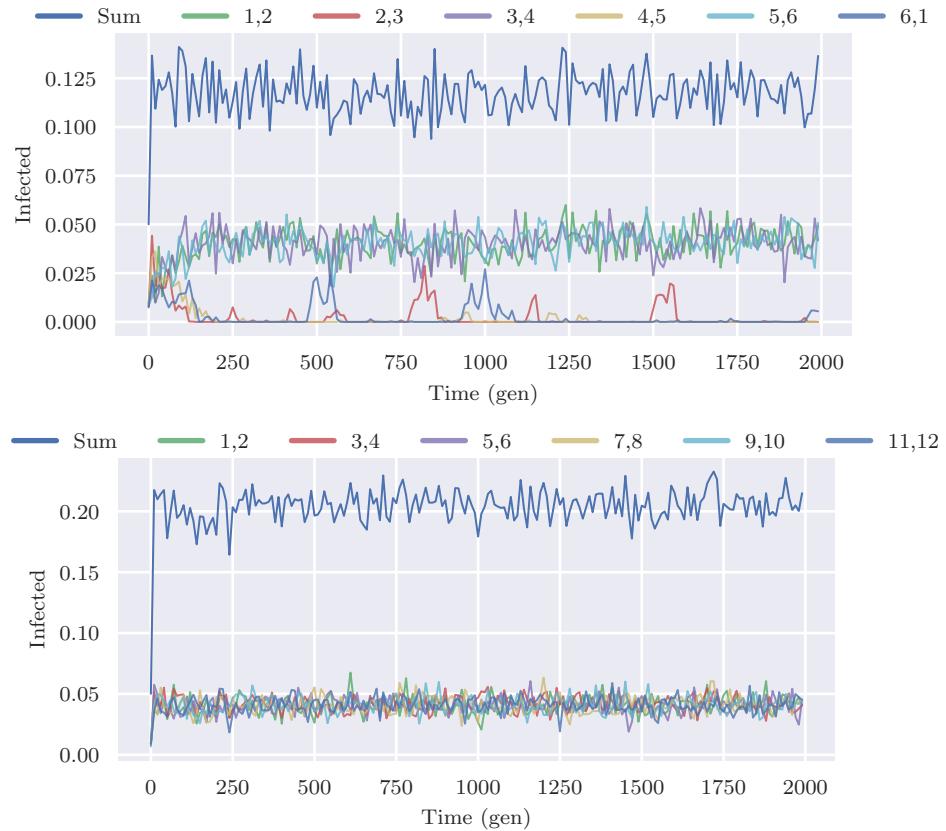
# Chapter 10

## Appendix

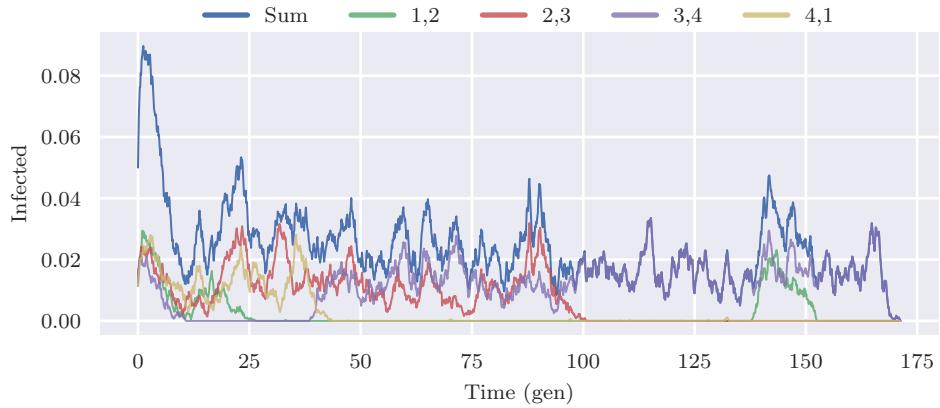
### 10.1 Replacement steady state code

```
1 clear; clc;
2
3 % Define variables. a is alpha, b is beta and g is gamma.
4 syms a b g S I I_R R
5
6 % Define equation system
7 eq1 = -a*S*(I+I_R)+g*(1-S) == 0;
8 eq2 = a*S*(I+I_R)-b*I-g*I == 0;
9 eq3 = b*I-b*I_R-g*I_R == 0;
10 eq4 = b*I_R-g*R == 0;
11 eq5 = S + I + I_R + R == 1;
12
13 % Write assumption: variables has to be positive and real.
14 assume(a, 'real')
15 assumeAlso(a > 0)
16 assume(g, 'real')
17 assumeAlso(g > 0)
18 assume(b, 'real')
19 assumeAlso(b > 0)
20
21 % Solve the system and print to command window
22 sol = solve([eq1, eq2, eq3, eq4, eq5], [S, I, I_R, R]);
23 Ssol = sol.S
24 Isol = sol.I
25 I_Rsol = sol.I_R
26 Rsol = sol.R
27
28 % Solve for optimal gamma and print
29 optimalForIDiff = diff(Isol + I_Rsol, g);
30 solveOptim = optimalForIDiff == 0;
31 OptimalGamma = solve(solveOptim, g, 'MaxDegree', 3)
32
33 % Solve for extinction gamme and print
34 ExtinctionEquation = Isol + I_Rsol == 0;
35 ExtinctionEquationSolution = solve(ExtinctionEquation, g)
```

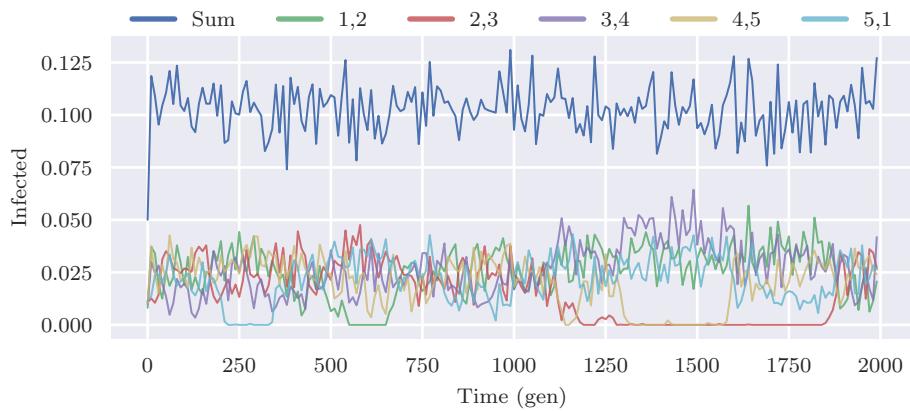
## 10.2 Various time-series plots



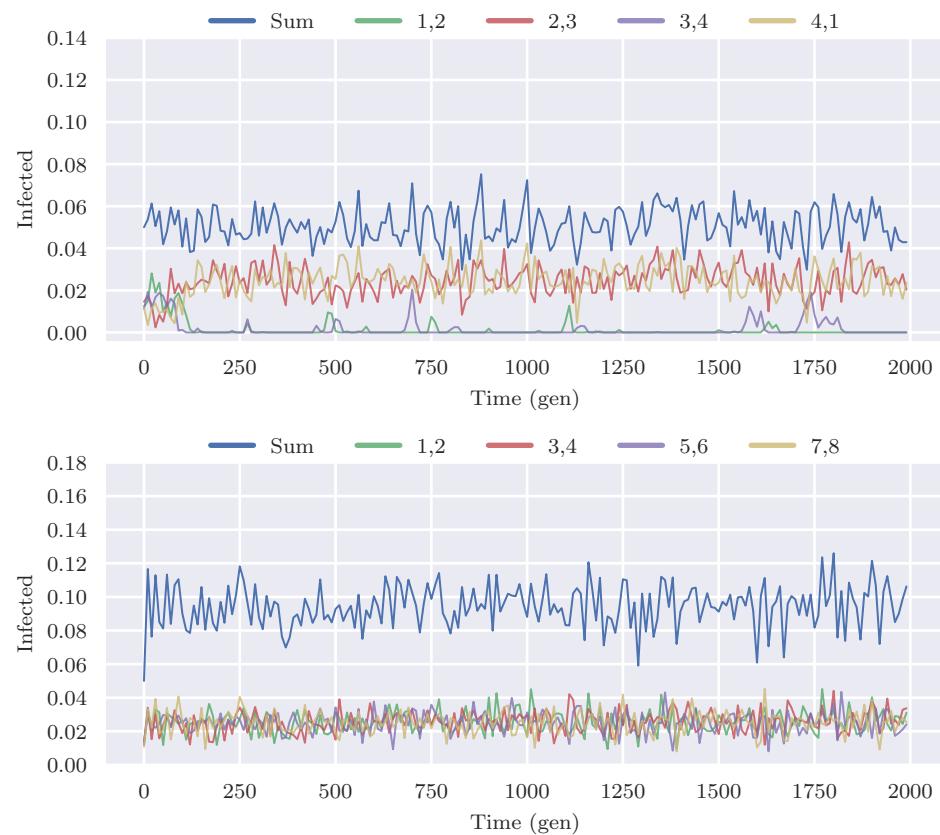
**Figure 10.1:** Time-series plots showing the mean number of infected for each strain and the total number of all infected for cross big and non cross big. The infection rate is  $\alpha = 0.45$ . The upper plot is for the system with cross immunity and the lower one does not. The surface features of each strain is indicated in the legends. Note the difference on the y axis for the plots. No switching between pairs happen in this instance



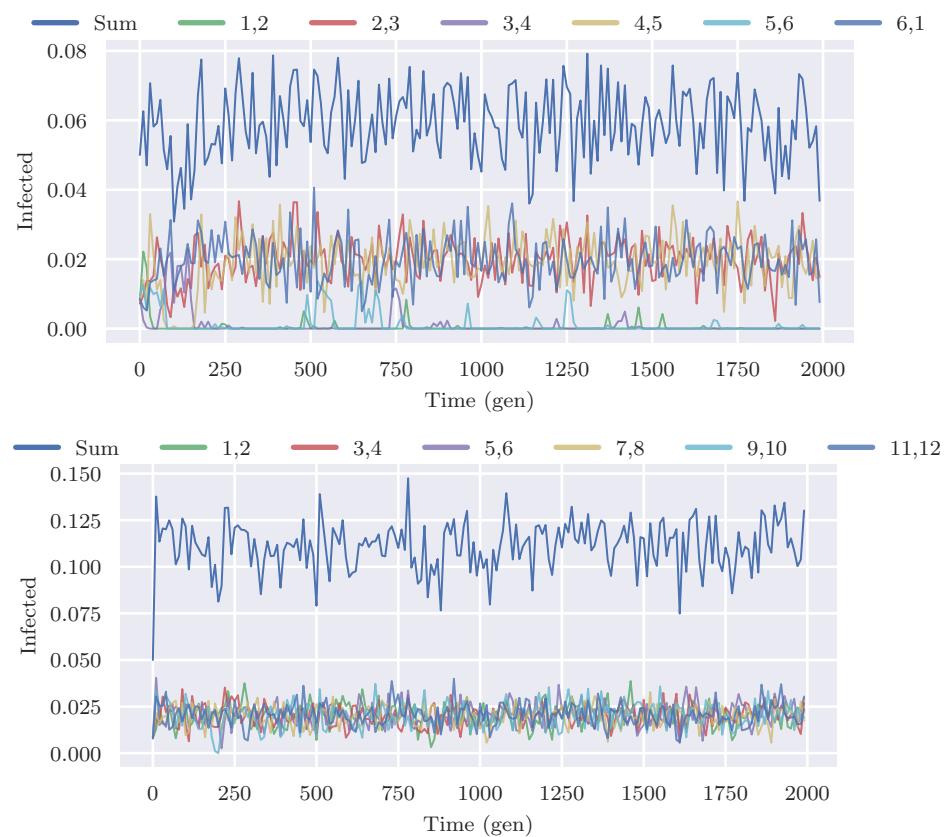
**Figure 10.2:**  $\alpha = 0.40$ . cross. The fluctuations for the individual strains makes the naive switching time extremely small. The time-series does not show the typical pairing like for larger  $\alpha$  values, because the strains are fragile.



**Figure 10.3:** Odd for  $\alpha = 0.45$ .



**Figure 10.4:** Comparing cross and no cross for  $\alpha = 0.42$ .



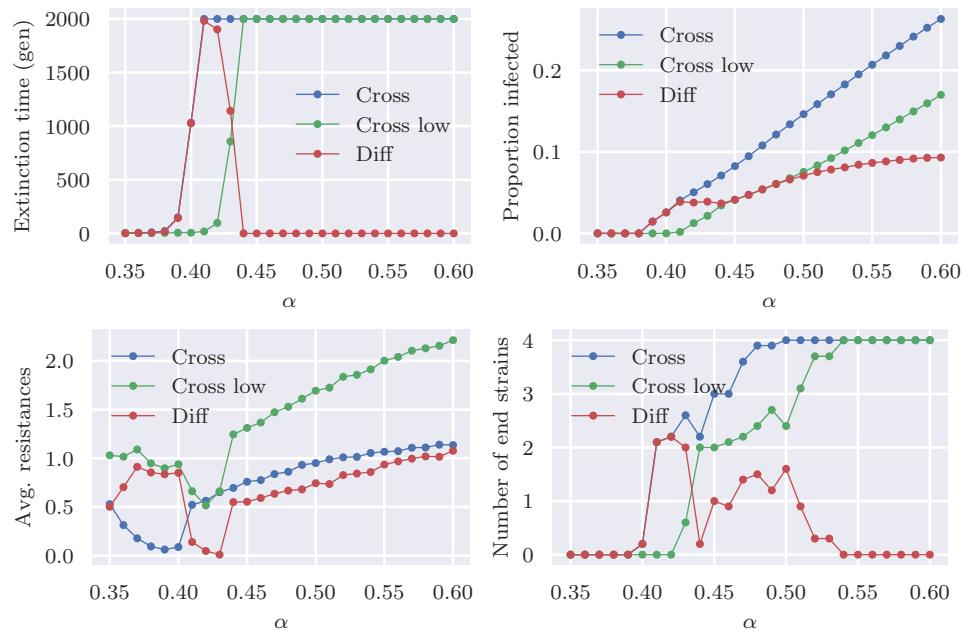
**Figure 10.5:** Big for  $\alpha = 0.42$

### 10.3 Number of data points for switching time

$\alpha$	0.40	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49	0.50
N cross	63	189	99	124	168	240	298	233	90	25	11
N cross big	297	139	49	26	65	111	160	114	31	10	10

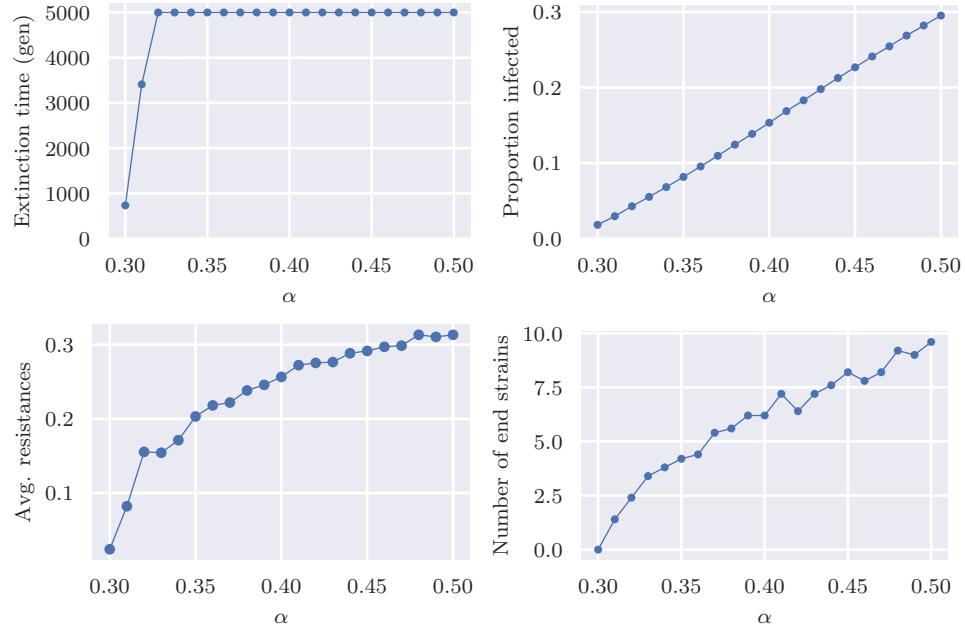
**Table 10.1:** Table showing the number of data points used for calculating the values in figure 6.25.

### 10.4 Lower replacement plots

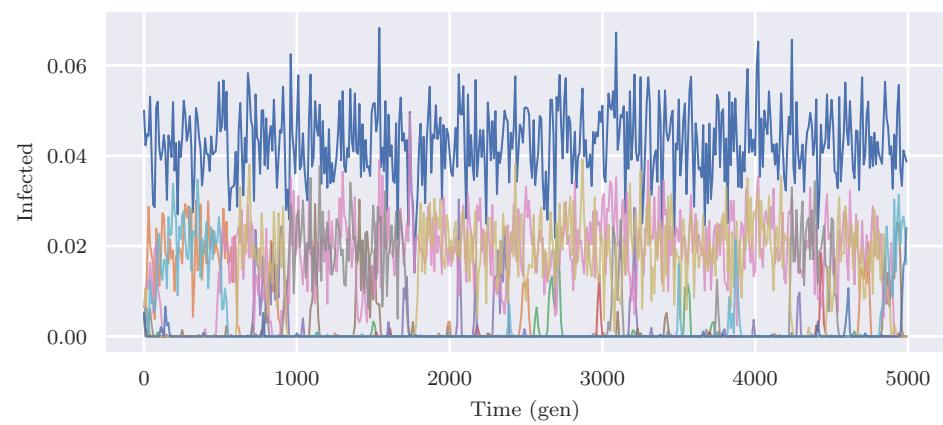


**Figure 10.6:** Figures showing cross of table 6.3, with optimal  $\gamma$  and with  $\gamma_{optimal}/4$ . The latter is named cross low.  $\mu = 10^{-4}$ . The plots show as hypothesised in section 7.1.2. The lower replacement rate is a detriment for all values of  $\alpha$  compared to the optimal  $\gamma$ .

## 10.5 Plots for systems with three antigens in each strain



**Figure 10.7:** Figures showing details of a system with three antigens in each strain. There is ten strains in the system. There are five unique antigens in the system, so each combination of antigens exist.  $\gamma$  is optimal and  $\mu = 10^{-4}$ .



**Figure 10.8:** Time series plots of the proportion infected and proportion infected for each strain. Parameters are  $\alpha = 0.32$ ,  $\gamma_{optimal}$ , and  $\mu = 10^{-4}$ . Only a few strains (2.5 on average) exists at a single time. The strains surviving almost always maximises the number of antigens.