UNIVERSITEIT VAN AMSTERDAM

Introduction to Computational Science

Exploration of stochastic SIR dynamics and meta population models

---

by

Caspar William Bruenech (*13104055*)
Bart de Mooij (*11217278*)

19 October 2020

Contents

1 Introduction

Currently, there are about 1400 known species of human pathogens.[5] Some of these can cause diseases that lead to serious health issues or even epidemics, such as COVID-19. COVID-19 has a major impact on societies all over the world, causing governments to implement restrictions as well as (partial) lockdowns. Therefore, infectious diseases not only form a health related threat but also an economical one. Governments are eager to obtain more insight into the spread of the disease to enable effective decision making on how to combat it.

Infectious diseases are caused by pathogenic microorganisms, such as bacteria and viruses. Viruses cannot reproduce nor survive without a host and are much smaller than bacteria. The main purpose of a virus is to deliver its genome into a host cell, resulting in an expression of the virus' genome.[1] This can potentially lead to harmful diseases like Influenza, Ebola and HIV. Infectious individuals can spread the virus among a population by transferring it to other hosts. The transmission of these infectious diseases differs upon disease and can be either direct or indirect.[4] For direct transmission, an individual gets infected via close contact with an infectious individual, droplet spread or a secondary host. In contrast, indirect transmission is passed via the environment, which can, for example, occur upon contact with a surface on which pathogens are present.

Infectious diseases can be modelled by using SIR models or more sophisticated variations of them. The name, SIR, of these models refers to the infection status where a distinction has been made between the number of susceptible (S), infected (I) and recovered or removed (R) individuals. Susceptibles are naive individuals who are susceptible to the disease, infected people are infected and able to transmit the pathogen to others. The recovered or removed compartment consists of recovered individuals and might not be contributing to further transmission of the disease, depending on the model in question. In the model, the number or fraction of individuals in the S, I and R compartments are obtained over time, displaying whether an epidemic might occur or not. In addition, the SIR model is also used to obtain insights into disease spread as well as establishing relative importance between different processes and actions. This can help governments with understanding and subsequent decision making. However, predicting future trends is too difficult due to the highly simplified representations of very complex systems.

Deterministic SIR models can be used to obtain the output of certain systems. Deterministic models use a set of equations to describe the output. When given the same parameter values and initial conditions, these calculations will always follow the same trajectory giving equal results. However, complex real world systems will rarely give the same outcomes due to chance coming into play. Therefore,

stochasticity should be introduced into computational models, which influences the SIR equations. This can, for example, take variations between individuals or fluctuating environmental factors into account. Stochastic effects cause the outcome of a simulation using the same initial conditions to be different from each other. Generally, these effects become more important whenever the number of infectious individuals is relatively small. This can be caused by a small population size, successfully applied control measurements or when an infectious disease has just invaded a population.

The five key features in stochastic SIR type models are variability between simulations, variances and co-variances, increased transients, stochastic resonance and extinctions. Simulations vary from each other because it is generally impossible to predetermine the precise prevalence of a disease. In addition, there will be a variance in the prevalence of a disease due to stochastic processes and mean population level can deviate from the deterministic equilibrium. Stochastic effects result in increased transients, which are caused by stochastic perturbations away from the endemic equilibrium and consequently countered by restorative forces of the endemic attractor. Stochastic perturbations can also excite oscillations close to the natural frequency of deterministic SIR equations. Lastly, stochasticity will result in extinctions due to chance fluctuations even when there is a basic reproductive rate of above one.

Imports can reintroduce diseases in a population, where no one is infected anymore. Subsequently, this infectious individual can transfer the disease to others within the population, causing the number of infecteds to rise again. The imports of diseases are essential to reintroduce a disease after extinction and are, therefore, required to obtain long term persistence.

Stochastic SIR dynamics can be explored, using Gillespie's Direct Method. This is an event-driven approach that incorporates the random nature of events. Gillespie's Direct Method uses the rates of all possible events to obtain the probability for each one of them to occur. Subsequently, it randomly selects one of the events based on these probabilities and repeats this process through time.

A metapopulation consists of groups of populations, which are spatially separated but consists of the same species. It allows to capture the spatial components of the system. Meta-populations are used because populations are usually not evenly spread, individuals move around between different populations and transmission is a local process. These calculations allow to determine the rate of spatial spread as well as help with understand how to optimize control measures. As a result, meta-population models are most applicable to human diseases.

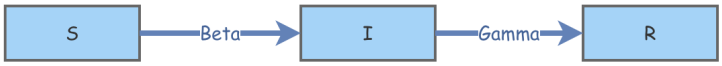

Figure 1: The compartments including the flow between them of a standard SIR model.

This research aims to explore the stochastic dynamics of the SIR model using their five hallmarks as well as investigating the effect of coupling between sub-populations in a meta-population model.

This report will explain the methodology of the research. After that, the results of the stochastic SIR model including its five key features and the meta-population models will be given. Subsequently, there will be a discussion after which a conclusion will be made based on the discussion and obtained results.

2 Methodology

Standard SIR model. The standard SIR model assumes that individuals obtain lifelong immunity once they moved to the R compartment. That means that once an individual enters the R compartment, it will stay in the R compartment forever. In addition, demography is not taken into account, resulting in a constant population size. The movement between compartments is characterized by certain parameters. The force of infection, β, describes the rate at which susceptibles are getting infected and, thus, move from the S to the I compartment of the model. This infection rate depends on the underlying population contact structure as well as the probability of transmission given contact. Subsequently, the rate of recovery, γ, describes the movement from the I to R compartment and is given by

$$\gamma = \frac{1}{infectious\,period} \tag{1}$$

The flow between the compartments including the corresponding parameters are shown in figure 1.

The number of individuals present in the S, I and R compartments change over time and are given by the following first order differential equations

$$\frac{dS}{dt} = -\beta SI \tag{2}$$

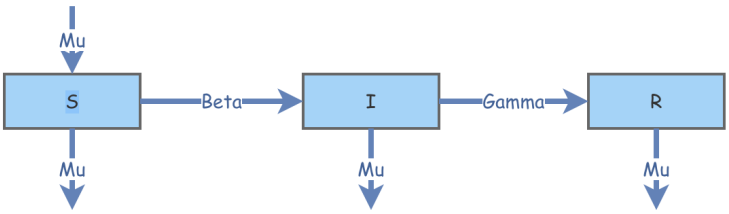

Figure 2: The compartments including the flow between them of a SIR model with demography.

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

where S, I and R are the number or fraction of individuals present in the corresponding compartment. The parameter determining if the infectious disease will become an epidemic is the basic reproductive rate R_0 which is given by

$$R_0 = \frac{\beta}{\gamma} \tag{5}$$

SIR model with demography. The SIR model with demography takes births and deaths into account. Births lead to fresh, new susceptibles, which is an important factor for an endemic state. On the other hand, all compartments suffer a natural mortality of μ individuals per time period. In the simple form, the birth and death rate are assumed to be equal, resulting in a constant total population size. However, more sophisticated models include different birth and death rates and, therefore, a varying population size. The flow of this model is shown in figure 2.

The first order differential equations of the SIR model with demography are as follows

$$\frac{dS}{dt} = \mu - \beta SI - \mu S \tag{6}$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I \tag{7}$$

$$\frac{dR}{dt} = \gamma I - \mu R \tag{8}$$

where the positive μ in equation 6 refers to the birth rate and all the negative μ terms to the natural death rate in the corresponding compartment. The introduction of demography also results in an adjusted basic reproductive rate which is given by

$$R_0 = \frac{\beta}{\gamma + \mu} \tag{9}$$

SIR model with imports. Infectious diseases can be re-introduced into a population by, for example, an infectious outsider who visits the population for a limited time. Even though there are multiple mechanisms for importing the disease, they are commonly modeled in two ways. However, the models in this report assumed that for every import, a susceptible was exported. Therefore, once an import occurs the susceptible pool decreases by one, while the infected pool increases by one. This also accounts for the case when an external source causes an infection, for example, when a susceptible visits another population and gets infected there before returning to its own population. The rate at which these imports occur for human pathogens is given by

$$import\ rate = \eta\sqrt{N} \tag{10}$$

Stochastic SIR. In this experiment we will be studying a stochastic implementation of the SIR model by introducing randomness into the dynamics. There are various ways of implementing this, including the introduction of noise into the transmission term, which is easy to implement and allows the study of more realistic dynamics such as extinction of the disease. However, this approach does not take into account the dynamics caused by variations in the individuals of the populations, thus a lot of the details of the dynamics when looking at smaller populations is absent. To circumvent this, one can perform a so-called event-driven approach to simulating the model, wherein the time is advanced by performing a series of events that affect the individuals of the populations. As we are now working with the absolute number of individuals, we denote the number of individuals in the susceptible, infected, and recovered compartments as X, Y, and Z, respectively. Demography is included into this model, which results in six possible events each with its own rate of occurrence, possibly happening at each time step:

- Birth, with rate μN, resulting in the addition of a susceptible individual, i.e. $X \to X + 1$.
- Transmission of infection at a rate $\beta\frac{XY}{N}$, causing an individual to move from the X compartment to the Y compartment, i.e. $X \to X - 1$ and $Y \to Y + 1$.

- Natural death in each of the three compartments (individual events), with rates μX, μY, μZ, resulting in $P \to P - 1$ for $P \in X, Y, Z$.
- Recovery, an infected individual recovers from the disease, causing the change $Y \to Y - 1$, $Z \to Z + 1$.

Gillespie's Direct Algorithm was used to simulate this stochastic model[2], which can be written as

Algorithm 1: Gillespie's Direct Algorithm

Label all events $E_1, ..., E_n$;
while *time* $\leq$ *maxtime* **do**
  Set rates $R_1, ..., R_n$;
  Set sum of rates $R = \sum_{m=1}^{n} R_m$;
  Set time until next event $\delta t = -\frac{\log RAND_1}{R}$;
  Set $P = RAND_2 \times R$;
  Set cumulative sum of all events R_{sum};
  Set event $p = \min R_{sum}$;
  Set time step $time = time + \delta \times time$;
end

Where $RAND_1$ and $RAND_2$ are two random, uniformly distributed numbers in $(0, 1)$. By utilizing this algorithm, the spread of disease in a smaller population can be simulated, (hopefully) producing more accurate results than the deterministic model. The goal is to design a set of experiments which clearly show five aspects of the stochastic dynamics:

1. Variability: due to the randomness in the model, and unlike the deterministic model, it is not expected to produce the same results when running multiple realizations, using the same initial conditions and parameter values. This variability will visualized in the model by running a series of realizations and plotting the end values of the X, Y, and Z compartments.
2. Stochastic resonance: A deterministic SIR model with demography will often approach the equilibrium values in a series of decaying oscillations in the number of infected, i.e. a series of epidemics with weaker outbreaks. As a result of how the stochastic model oscillates around the equilibrium, these oscillations can resonate with the damped oscillations of the underlying deterministic dynamics, causing epidemics with larger amplitudes and longer periods. This is visualize by comparing a number of stochastic realizations to a deterministic model with the same parameters. In order to observe resonance it is required to make sure the parameters were chosen in such a way

that the deterministic model has clear oscillatory behaviour. This should result in a higher amplitude and an epidemic with a larger duration for the stochastic model compared to the deterministic model.

3. Increased transients: While the stochastic model does not directly follow the deterministic path of the ODE model, the changes in the population (i.e. the rates) are the same. Thus, there is an underlying deterministic behaviour in the model, with the addition of random behaviour around the equilibrium. This means that when the model is far from the equilibrium, it will quickly move towards it, and thus behave very similar to the deterministic model. To visualize this, the initial conditions are set to be far from the endemic equilibrium and a number of stochastic realization were run before comparing it to the deterministic model. We expect to see a smaller variance in stochastic models when the value is far from the equilibrium, thus following the deterministic model more closely.

4. Negative covariance and variance: Similar to the increased transient behaviour of the stochastic model, the underlying deterministic dynamics affect the model such that there is a general negative covariance between the number of infected and number of susceptibles. Additionally, the random behaviour of the model will cause perturbations in the populations and, consequently, there will be a continuous variation in the number of infected and susceptibles. This variance combined with the negative covariance will cause the stochastic model to deviate from the equilibrium values. To compute this, an experiment with parameters, such that the equilibrium values of the number of infected is not close to zero, was designed. Then, a series of realizations were computed as well as the mean covariance and variance over all runs. This will be performed for a varying number of total individuals in the system, and it is expected to observe a negative covariance between the number of susceptibles and infecteds for all experiments.

5. Pathogen extinctions: Diseases will eventually always go extinct in the case when no demography is included within the deterministic, continuous SIR model. However, introducing demography usually results into an endemic state rather than disease eradication due to births continuously refreshing the susceptible pool. In stochastic, individual based models, the disease will also be eradicated eventually as this will cause the transmission rate $\beta XY/N$ to always be zero, even when new susceptibles are born. In order to achieve an endemic-like state, imports of infected individuals from outside the population need to be included. This extinction characteristic is studied by looking at the average number of extinctions per year and the so-called "first-passage time", or the average time until the first extinction occurs,

both as a function of the total population size. When computing the average number of extinctions per year, imports are included, i.e. the possibility of new infecteds arriving from the "outside". For the first passage time experiment, the import parameter η will be set to 0 as only the first extinction is of interest.

These characteristics were estimated when the dynamics had reached the equilibrium state. Therefore, a so-called "burn-in" phase was included in the implementation, where an initial amount of time was simulated without saving any values. After the burn-in phase, the longer simulation was begun by using the end values of the burn-in phase as the starting conditions.

Metapopulations. In the stochastic SIR model we studied, it was assumed that the individuals were in a closed population. The possibility of importing infected individuals from the outside was included, where an infinite, unchanging pool of infected individuals was assumed. This is not effectively the same as having more populations, since how they affect each other was not taken into account. Metapopulations, while being one of the most simple spatial models, can produce realistic time series data of how a disease spreads in smaller human communities. By dividing the entire population into subpopulations, each with their own dynamics, a limited interaction between the populations is included. This allows the infection to spread both internally in each subpopulation as well as externally between different subpopulations. In these calculations, demography is ignored and all the subpopulations are assumed to be of the same size (i.e. contain the same total number of individuals). For this experiment, Gillespie's Direct Algorithm was used to simulate the spread of the infection. Since demography is not included, there are three events per subpopulation possible at each time step:

- Transmission of infection within the subpopulation with rate $\beta X_i Y_i / N_i$.
- Transmission from subpopulation j to subpopulation i with rate $\lambda_i X_i$.
- Recovery of an individual in subpopulation i with rate $\gamma_i Y_i$.

Whenever event (2) happens, an individual from subpopulation j simply infects one person in subpopulation i in a short time span before moving back to his own population and, thus, causing an increase of 1 in the infected compartment and a decrease of 1 in the susceptible compartment both of population i. The rate of infection from population j to population i depends on the *force of infection* λ_i, which is computed as [4]

$$\lambda_i = \frac{\beta_i}{N_i} \sum_j \rho_{ij} X_j \tag{11}$$

where ρ is a matrix quantifying strength of interaction between the populations, such that ρ_{ij} gives the strength of interaction to subpopulation i from subpopulation j. To study how the addition of metapopulations affect the dynamics of the SIR model, we will perform two experiments; one with two subpopulations, where $\rho_{i,i} = 1$, and $\rho_{i,j} < 1$, and one with more than two subpopulations. Multiple realizations with varying ρ_{ij}-values were performed to observe how this parameter affects the dynamics. For the 2 subpopulations experiment, it was required to decide what parameter to vary, as there is a vast number of potential parameter combinations when increasing the number of subpopulations by only a small amount. For both experiments, only one subpopulation was set to have one initial infected. Therefore, it may be most interesting to see how the pathogen spreads when the force of infection *from* the initially infected subpopulation *to* the other ones is varied. This is equivalent to varying the $n-1$th column in the ρ-matrix, which will be denoted as $\rho_{i,n-1}$.

3 Implementation & Results

3.1 Variability

Gillespie's Direct Algorithm was implemented into the SIR model to obtain information about the variability of the system. In these calculations the initial population size was set to 100 people, starting with initially only 1 infected. The parameters μ, γ, β and η were set to $\frac{1}{75}$, 1.0, 3.0 and 0.02, respectively. This simulation was ran 100 times and the obtained end values of the number of susceptibles, infecteds and recovered individuals are shown in figure 3.

3.2 Variance and Covariance

The second hallmark of stochastic SIR models is about the variance and the covariance. In these simulations, we used the parameter values $\mu = \frac{1}{75}$, $\gamma = 1.0$, $\beta = 3.0$, and $\eta = 0.02$. Figure (4) shows a log-log plot of the variance in the number of infected versus the mean of the number of infected. All values are the mean of 20 realizations.

The covariance between the number of susceptibles and infecteds of the stochastic SIR models were calculated. The co-variances were obtained at population sizes

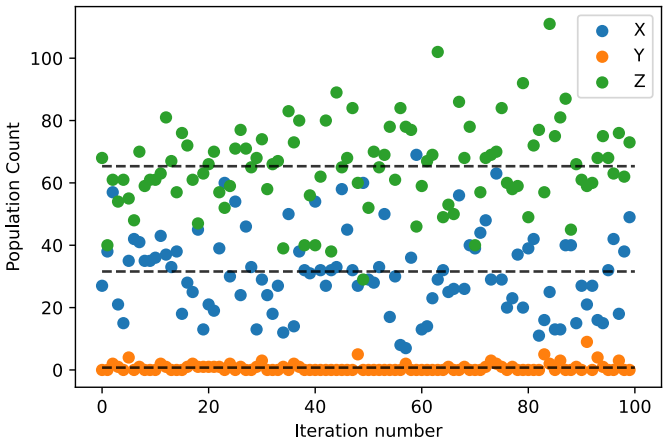

Figure 3: The variability of the stochastic SIR dynamics, showing the end values of X, Y and Z for a hundred simulations, where the dashed lines show the means of them.

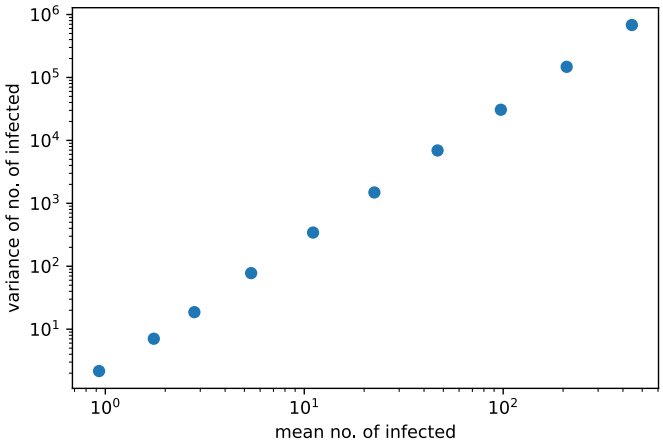

Figure 4: Variance in the number of infected people of the stochastic SIR model, showed by plotting the logarithm of the mean versus the logarithm of the variance on the x- and y-axis, respectively. Each data point is the average over 20 realizations.

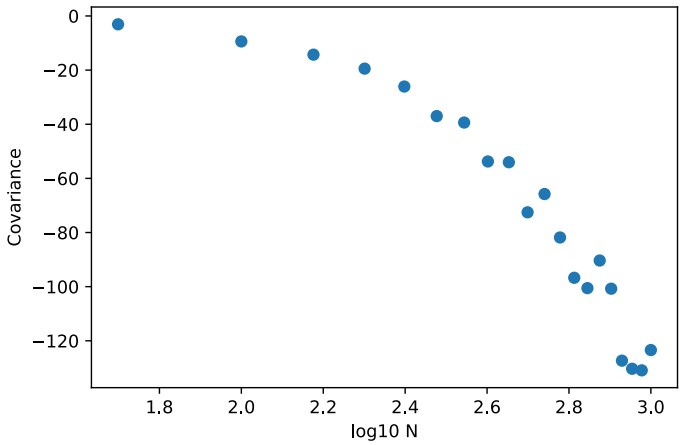

Figure 5: The co-variance between the susceptibles and infecteds of the stochastic SIR dynamics as a function of the total population size.

varying from 50 to a 1000 individuals, but always started with only one infected. In these simulations, μ was set to $\frac{1}{75}$, γ to 0.7, β to 3.0 and η to 0.02. Subsequently, this was plotted in figure 5.

3.3 Increased Transient

The increased transients were calculated using the SIR model with demography, excluding imports. The total population was set to 150 individuals, while the initial number of infecteds was 100. The parameters μ is $\frac{1}{75}$, γ 0.1 and β 3.0. This was simulated once deterministically and 300 times stochastically. The resulting graph of the increased transients of the number of infecteds is shown in figure 6, where the red line is the deterministic solution and all the transparent lines correspond to the different stochastic solutions.

3.4 Stochastic Resonance

The stochastic resonance of the SIR dynamics was simulated for the number of infecteds and, subsequently, compared to the deterministic SIR equations. These simulations also ignored the imports and, therefore, η was zero. The initial population size was 150 with 10 initial infecteds and 50 initial recovered, while the parameters μ, γ and β were set to $\frac{1}{8}$, 0.1 and 1.0, respectively. This was simulated 10 times, resulting in figure 7. In this graph, the red line is the deterministic

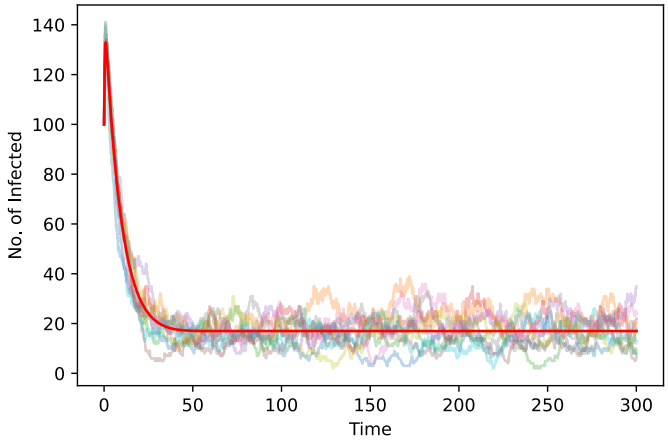

Figure 6: The increased transients of the stochastic SIR dynamics in the number of infecteds compared to the deterministic SIR equations.

solution, the transparent lines the different stochastic solutions and the blue line the mean of these stochastic simulations.

3.5 Extinctions

To study extinctions, two experiments were run; one for computing the average number of extinctions per year as a function of the total population count, and one for computing the so-called *first passage time*, or average time until the first extinction. For the first experiment, the parameter values were γ of 0.7, μ of $\frac{1}{50}$, β of 10.0 and η of 0.02. These were used over the range $N \in 10, 10^5$, with 20 equally spaced intervals. For each value of N, 20 iterations were performed and the mean number of extinctions per year were computed. The resulting plot is shown in figure (8).

For the first-passage time experiment, we utilized the parameter values to a γ of 1.0, μ of $\frac{1}{50}$, β of 3.0 and η of 0.02 taken over $N \in 10^2, 10^4$, again as 20 equally spaced values. Similarly, each data point was computed using the mean of 20 realizations. A burn-in time of 50 and a simulation time of 150 were performed to ensure that the stochastic simulations were started at, or near, endemic equilibrium. The result is shown in figure 9.

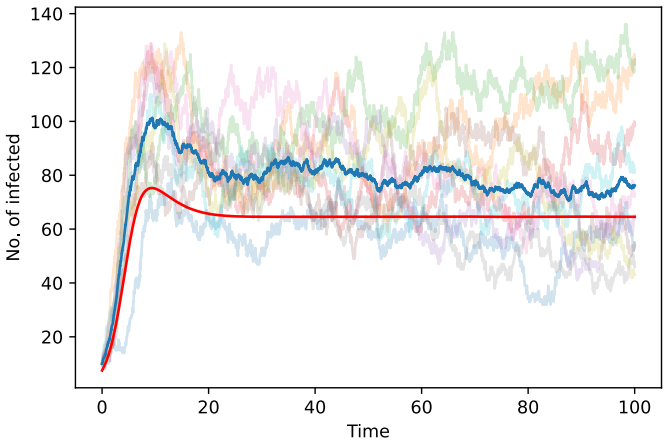

Figure 7: The stochastic resonance of the stochastic SIR dynamics in the number of infecteds compared to the deterministic SIR equations.

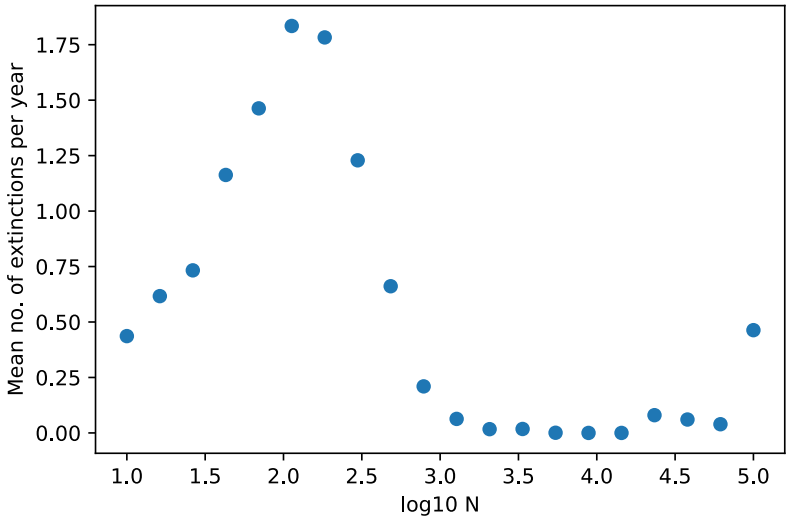

Figure 8: Mean number of extinctions per year as a function of total population count. Each mean value is computed using 20 realizations with the same parameter values.

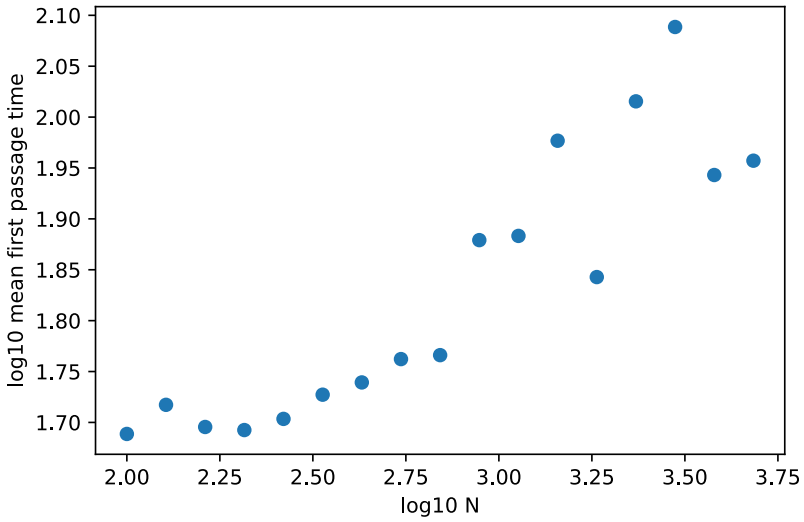

Figure 9: A log-log plot of the mean first passage time till extinction plotted against the total population size.

3.6 Meta-population Model

The first meta-population experiment involved using only two metapopulations, both with an equal number of individuals set to 100. The parameters γ_i and β_i were set to 1.0 and 3.0, respectively, for both populations. Population 1 contained initially 1 infected individual, while population 0 contained none. Four realizations were performed with varying values for $\rho_{i,j}$. For each realization, 100 simulations were performed and the mean number of infecteds per time step was computed. The result of this is shown in figure (10).

Subsequently, a similar experiment was performed with 6 metapopulations. The total population count of all subpopulations were set to be equal. In addition, the parameter values of β_i and γ_i were set to be the same as in the previous experiment with two subpopulations. In this case, population 5 had one initial infected, while all the others were initially disease-free. The force of infection matrix $\rho_{i,j}$ was generated with random values between 0 and 1 for $i, j \in 0..4$, while the last column was iterated over the values $(0.1, 0.9)$. A visualization of this is shown in figure (11). For each value, a 100 realizations were performed and, subsequently, the mean was computed. The resulting plot, showing all six populations for each value of $\rho_{,5}$ is displayed in figure (12).

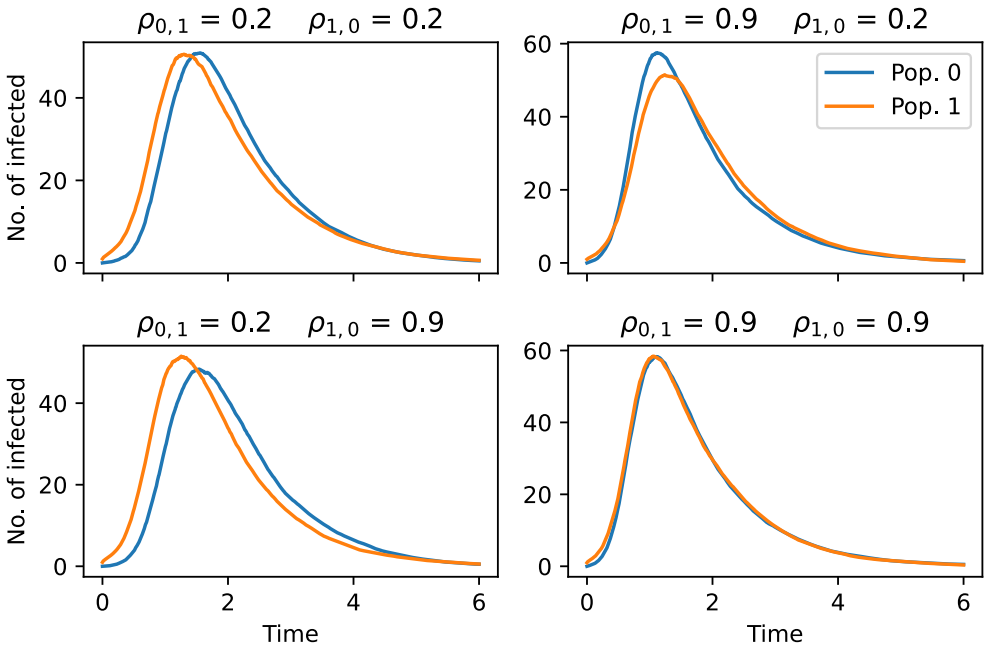

Figure 10: Plots of two metapopulations interacting with each other, using varying strengths of interactions between the subpopulations.

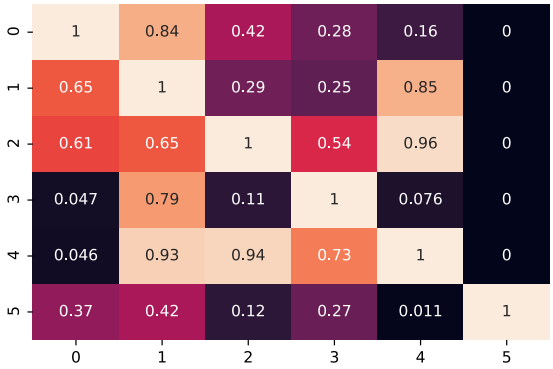

Figure 11: Annotated heatmap showing the values for the strength of interaction, $\rho_{i,j}$ between the six sub-populations.

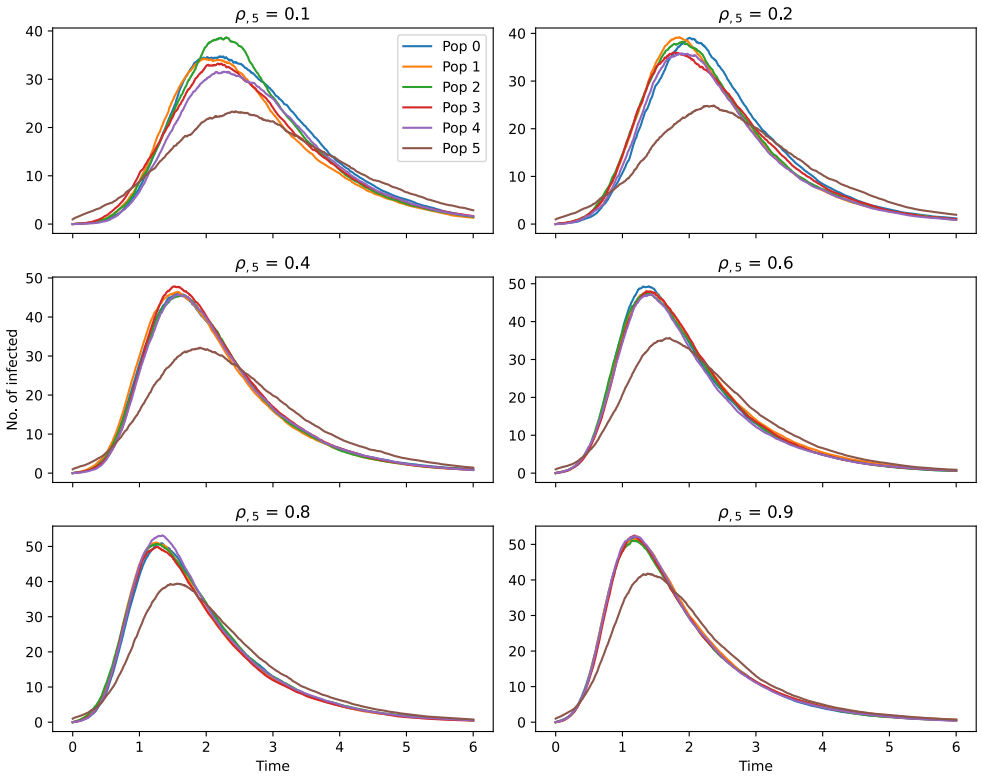

Figure 12: Plots of six metapopulations interacting with each other, using the interaction strength of subpopulation five to the other to change from 0.1 to 0.9 and the others to be randomly generated as shown in figure 11

4 Discussion

The variability of the stochastic SIR dynamics is shown in figure 3. This graph shows that for every iteration, different end values of the number of susceptibles, infected and recovered are obtained. This implies that there is a variability in the stochastic simulations, caused by the randomness in the events, which is in line with what was expected.

Figure (4) shows a log-log plot of the variance of the number of infected individuals versus the mean of the number of infected. The plot shows a clear linear behaviour between the variance and the mean, which implies a power-law relationship between the two quantities. This is indeed what was expected, as this has been previously observed to be true for a variety of stochastic epidemiological models [3]. This also tells us that even though a higher population count results in a model with fewer transients and, consequently, behaviour that is more similar to the deterministic model. Nevertheless, the absolute variance is higher due to the increased number of people and, thus, the larger the values are it can vary with due to, for example, a larger susceptible pool. The covariance between the susceptible and infected individuals is shown in figure 5. This graph shows that the covariance between them is always negative and becomes more negative with a higher population size. The reason for this is that a larger population size causes more susceptibles to be present, resulting in a faster and more efficient spread of diseases.

The increased transients of the stochastic SIR dynamics in the number of infecteds compared to the deterministic SIR equations is shown in figure 6. This graph shows that all the different stochastic simulations possess a different trajectory, but always moves back towards the deterministic solution once it deviates from it. It is observed that once the number of infecteds deviates further away from the deterministic equilibrium, it jumps back steeper towards this equilibrium. This is caused by the restorative forces that become stronger and more dominant further away from the equilibrium. This is clearly visible in the form of the smaller variance in the stochastic realizations when the number of infected is higher (i.e. further from the equilibrium), and is exactly what was expected.

The graph of the stochastic resonance given in figure 7 shows that the stochastic simulations deviate from each other. They are generally different than the deterministic solution. The blue line shows that the mean of the stochastic simulations and it is observed that this is steeper and higher than the deterministic one. The reason for this is that stochastic perturbations can excite oscillations close to the natural frequency of the deterministic dynamics. Therefore, the epidemic is enhanced as well in this graph and the endemic equilibrium is higher for

the stochastic SIR dynamics than for the deterministic one.

For studying the possibility of pathogen extinction due to stochasticity, the results in figures (8) and (9) were produced. The first figure shows the average number of extinctions per year as a function of total population size. The results of the mean number of extinctions per year shows that the number is small for smaller population sizes and increases exponentially before reaching a peak at around $N \approx 10^2$. Then, the number quickly degrades as the population size increases, before stabilizing close to 0 for $N \leq 10^3$. Intuitively, it is expected that, for smaller populations, the increased stochastic fluctuations would cause the number of extinctions to be higher as there is a higher likelihood of the number of infected jumping to zero. Consequently, it is expected to see that as the population grows larger, this number goes drastically down as the model behaves more and more like the deterministic model with smaller fluctuations. While this expected behaviour is indeed observed in figure (8), it is only apparent in the interval $N \in 10^2, 10^5$. For $N10^2$, the mean number of extinctions per year actually increases with the population count. One possible explanation for this is that the population is small enough, causing the fluctuations to be so large that the number of infected individuals jumps to larger values further away from the endemic equilibrium. Therefore, the number of infecteds are further away from 0, reducing the chance of extinction. As the population grows slightly larger, the fluctuations are still strong, but not strong enough to profoundly change the dynamics via a large random jump in the number of infected. Thus, at around $N = 10^2$, the fluctuations are strong enough to result in the number of infected often ending up in 0, but not strong enough to randomly jump to a large value, delaying possible extinctions.

For the first passage time results, (Fig. (9)), it is observed that the mean time until first extinction increases somewhat exponentially with the total population count with a large variance in the results for $N\,10^3$. The simulations with smaller population sizes, thus, the ones with more frequent and stronger fluctuations, are expected to have a smaller first passage time. Therefore, these results are in line with what was anticipated. However, it should be noted that when computing these results, there were iterations in which the model contained zero extinctions. Thus, some data points will be less accurate than others, as the mean first passage time will only have been computed over a few iterations. This is why there is a larger variance observed in the data for higher population counts. When the dynamics behave more akin to the deterministic model, there is a smaller likelihood of extinction and, therefore, only a few iterations resulting in an extinction time, which can be used for estimating the mean first passage time. Nevertheless, the results shows the expected behaviour; higher total population counts results in a more deterministic model, resulting in a smaller probability of extinction at each time step and increasing the mean time until the first extinction.

The metapopulation experiment containing two subpopulations was performed and the result is shown in figure 10. The initial infected person was placed in population 1. When the interaction strengths between both subpopulations were low (i.e. $\rho_{0,1}$ and $\rho_{1,0}$ are 0.2), a delay is observed. The reason for this delay is that population zero needs to get an infection via population one, which takes some time as the interaction between them is relatively weak. When the interaction strength of the population with initial infected to the other population was taken to be large and weak the other way around (i.e. $\rho_{0,1}$ is 0.9 and $\rho_{1,0}$ is 0.2). In this case, there is no delay observed because the coupling will quickly cause infection of someone at population 0. In the case when these two interaction strengths are reversed (i.e. $\rho_{0,1}$ is 0.2 and $\rho_{1,0}$ is 0.9), there is a delay again. This is caused by the weak coupling of the initially infected population to the disease-free population. In the last graph of this figure, both populations interact strongly with each other (i.e. $\rho_{0,1}$ and $\rho_{1,0}$ are 0.9). This shows no delays and the graphs almost completely overlap each other. These subpopulations interact so strongly with each other that they could almost be considered as one larger population. The interaction strength from the initially disease-free population to the initially infected population is, thus, not important in observing delays of infections in a subpopulation. In addition, from these graphs, it can be concluded that a weaker interaction strength of the initially infected population towards the initially disease-free population causes a delay in people getting infected.

Figure 12 shows the result from the six metapopulation model. The various subplots shows the mean values of the time series of the number of infected for each subpopulation, with varying values for the $\rho_{,5}$ column. This shows how the strength of interaction between the initially infected population and the other subpopulation affects the spread of the pathogen. It is observed that for smaller values of the parameter, the peak of the populations is smaller. For population 5, the initially infected population, this peak is around 23 for the smallest $\rho_{,5}$ values, and approximately 40 for the largest. The other populations show a similar behaviour, with a peak going from around 35 to 50, when the strength of interaction increases. The reason for this is due to the fact that with a stronger strength of interaction, the pathogen spreads more rapidly to the other subpopulations. This is due to the connection between the populations, resulting in more spread between the groups. With a small strength of interaction, only a few infections are caused by population 5. Therefore, the outbreak in each subpopulation is not amplified by the continuous spread from population 5. In this experiment, a substantially smaller peak in the number of infected in subpopulation 5 compared to the other subpopulation is also observed. The reason for this can be deducted by looking at the ρ-matrix (Fig (11). It is observed that row 5, which denotes the strength of interaction *from* all subpopulations $\rightarrow$ number 5, all have a relatively small value with none of them

being over 0.5. This means that there will not be as many infections caused in subpopulation 5 as a result of spread from a different subpopulation when compared to, for example, subpopulation 2, which has a substantially higher mean $\rho_{2,j}$-value. Additionally, it is observed that there is a delay in the outbreak in subpopulation 5, a delay which increases with the parameter value. This is also a consequence of the small value of $\sum_{j=0}^{5} \rho_{5,j}$, as the other subpopulations will have a larger effective R_0 due to the higher value of the strength of interaction from 5 to the others. Therefore, delaying the epidemic in subpopulation 5, even though it was the one that started with an initial infected individual.

5 Conclusion

The following can be concluded from the obtained results and the discussion of them:

- Stochastic simulations results in the variability of SIR dynamics.
- The number of susceptibles and infected are related via a power-law relationship.
- There is a negative co-variance between the number of susceptibles and number of infecteds and more negative for larger population sizes. Larger population sizes causes a faster and more efficient spread of diseases.
- Stochastic models can be conceptualized as resulting from random perturbations away from, and transient-like return towards, the deterministic attractor.
- Stochasticity can excite oscillations close to the natural frequency of the deterministic models, resulting in enhanced epidemics and stochastic resonance.
- Smaller populations exhibit a shorter first passage time. However, for really small populations, this is not observed possibly due to the number of infecteds initially grows far away from extinction.
- Larger population counts result in a more deterministic model containing a smaller probability of extinction.
- In metapopulations with two subpopulations, weaker interaction strength of the initially infected population towards the in initially disease-free population results in a delay of people getting infected in the initially disease-free

population. On the other hand, strong interactions between the two subpopulations causes them to behave similarly.

- With the addition of more subpopulations, in this case 6, it is observed that for a higher interaction strength between the initially infected subpopulation and the others resulted in a larger total number of infected individuals in addition to a more rapid arrival of the peak incident count. In addition, just like the two subpopulation case, the initially infected subpopulation can have a delayed outbreak if the strength of interaction to this population is small.

Further research can look into a larger variation of parameters used in these models and compare those effects with each other, especially for the interaction strengths in our metapopulation models. In addition, due to the limited computational power available for this research, next research could improve and verify our obtained results by using more simulations to obtain more accurate means as well as perform these simulations for larger population sizes.

References

[1] H. R. Gelderblom. *Medical Microbiology*. 4th edition. Galveston, Tex., 1996. ISBN: 0-9631172-1-1.

[2] D. T. Gillespie. "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions". In: *Journal of Computational Physics* 22 (4 1976). ISSN: 10902716. DOI: 10.1016/0021-9991(76)90041-3.

[3] M. J. Keeling and B. Grenfell. "Stochastic dynamics and a power law for measles variability". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 354 (1384 1999). DOI: 10.1098/rstb.1999.0429.

[4] M. J. Keeling and P. Rohani. *Modeling Infectious Diseases in Humans and Animals*. 1st edition. Princeton University Press, 2008. ISBN: 978-0-691-11617-4.

[5] "Microbiology by numbers". In: *Nat. Rev. Microbiol.* 9 (2011), p. 628. DOI: https://doi.org/10.1038/nrmicro2644.