

MODELLING DISEASE SPREAD - MIDTERM CHECKPOINT

(GROUP 10) SUBHODH KOTEKAL, DONGIL LEE, HONGYANG LI, BOYUAN YU

1. INTRODUCTION TO THE SIR MODEL

The Susceptible-Infected-Removed Model for Disease Spread, in which we will call it the SIR model in the following discussion, is a model that puts each individual in the population into exactly one of three categories: S , I and R . In the simplest setting of simulating a disease spread, the S category contains those individuals who are susceptible to the disease. Meanwhile, the I category stands for the infected individuals who can spread the disease, and the R category is the population of "removed" individuals who were previously infected, and are no longer able to spread the disease. In reality, one can think of the R category as people who have either recovered and immuned to the disease, or died. We also want to introduce two parameters b and k . The parameter b represents the number of people that an infected individual spreads disease to through interaction each day, and k is the fraction of infected population that is "removed" (become members in the R category). Together, we can formulate this model into a system of ordinary differential equations, where $s(t)$, $i(t)$ and $r(t)$ denote the size of the population of the corresponding category as a function of time variable t as the following:

$$s'(t) = -b \cdot s(t) \cdot i(t)$$

$$i'(t) = b \cdot s(t) \cdot i(t) - k \cdot i(t)$$

$$r'(t) = k \cdot i(t).$$

One can interpret the equations as follows: the first equation shows that the number of susceptible people who are newly infected (hence being moved from S category to I category) at each given time is positively correlated with $s(t)$, $i(t)$ and parameter b . Intuitively speaking, it means more people would be infected each day if there are more susceptible people, more infected people, or more interactions between populations. The third equation shows that the number of infected people being removed (hence being moved from I category to R category) at each given time is positively correlated with $i(t)$ and parameter k . It means that more people would be removed (or recovered) each day if there are more infected people or higher rate of "removing" (or recovery). The second equation can be seen as a result of the first and the third equation, as the sum of population of categories must be equal to the total population N at any given time.

2. STRUCTURE OF `SIR`

To conduct simulations of ODE version of the SIR model, we take an object oriented approach by creating the `SIR_ODE` class. A `SIR_ODE` object can be created by providing the parameters `b` and `k` as well as providing the initial data `i(0)` and the population size `N`. This object thus stores the parameters of a model of interest, and the object can be made to simulate the solution for a given interval of time. The parameters of the `SIR_ODE` can be easily changed. The purpose of this object is to neatly package all of the information of a particular instantiation of the SIR model. This `SIR_ODE` object is the primitive representation of the ODE version of the SIR model. Variants of the SIR model that will be used in the final project will be associated to corresponding objects that inherit from `SIR_ODE`.

To Build a discrete version of the SIR model, we also take an object oriented approach by first creating the `SIR_DISCRETE` class, which represents a person with three states: 'S', 'I', and 'R'. The person's state can be obtained and assigned. Then we initialize the simulation by providing the population size `N` and initial infected and susceptible people ratio: i_0 and s_0 . The simulation precedes by updating the states based on two parameters: `k` and `b`. Further variations for final project can be implemented by either inheriting from the existing class or changing the way to update states of people.

3. PRELIMINARY INVESTIGATIONS

To get a first sense of how the ODE version of the SIR model behaves, we started with a simple simulation. We considered the parameter configuration of $k = 0.05$, $b = 10$, $N = 10000$ and $i_0 = 0.0001$. In words, about 5% of those infected recover per day, there are about 10 interactions per day per individual that could spread the disease, the population size is 10000 individuals, and 0.01% of the population is infected at time 0. Using the `solve_ivp` function from `scipy`, we simulate the solution of $S(t)$, $I(t)$, $R(t)$ for $0 \leq t \leq 50$ and obtain Figure 1. With this configuration of parameters, the dynamics are striking in how fast the number of infected individuals rises. In particular, the number of susceptible people in the population drops to zero extremely quickly. The mirroring dynamics of the infected and recovered individuals agrees with intuition.

We suspect that the reason for the such rapid increase increase is due to the choice of $b = 10$. To examine how social distancing might affect the dynamics, we ran the same simulation, except now we set $b = 1$. There is only one interaction per day per individual in this socially distanced society. The dynamics corresponding to this ODE system is given in Figure 2. These dynamics are quite different from those seen in Figure 1. We immediately notice that the time it takes for $S(t)$ to drop to zero is about ten times as long in Figure 2 than it is in Figure 1. Furthermore, we see that the highest number of infected people is much lower in Figure 2 than it is in Figure 1. As expected from the equations of the SIR model, the growth rate of $R(t)$ is the same in both figures. The parameter b dictates not only the rate at which susceptible individuals become infected, but also the

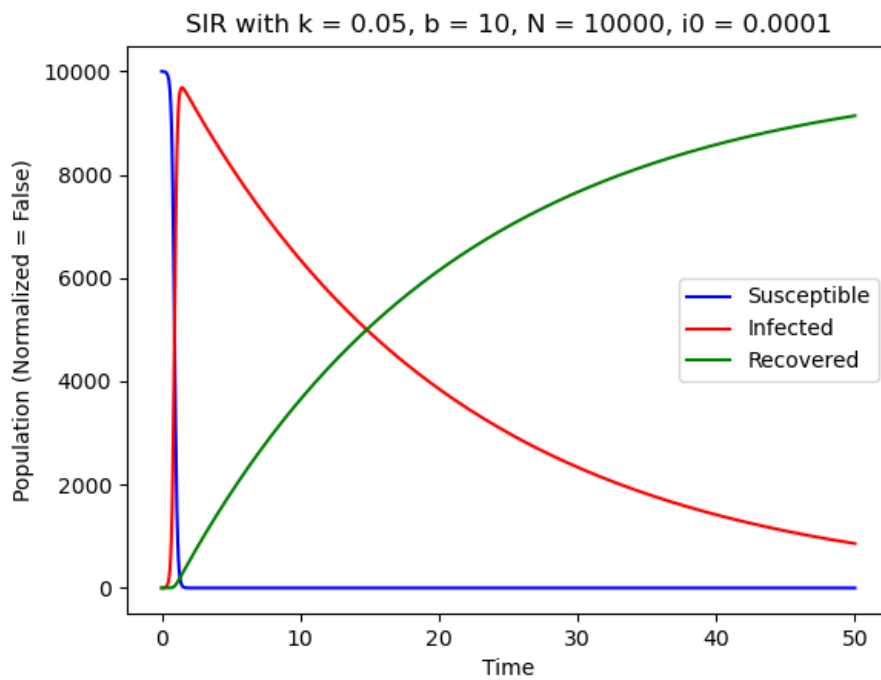


FIGURE 1. A simulation of a SIR model.

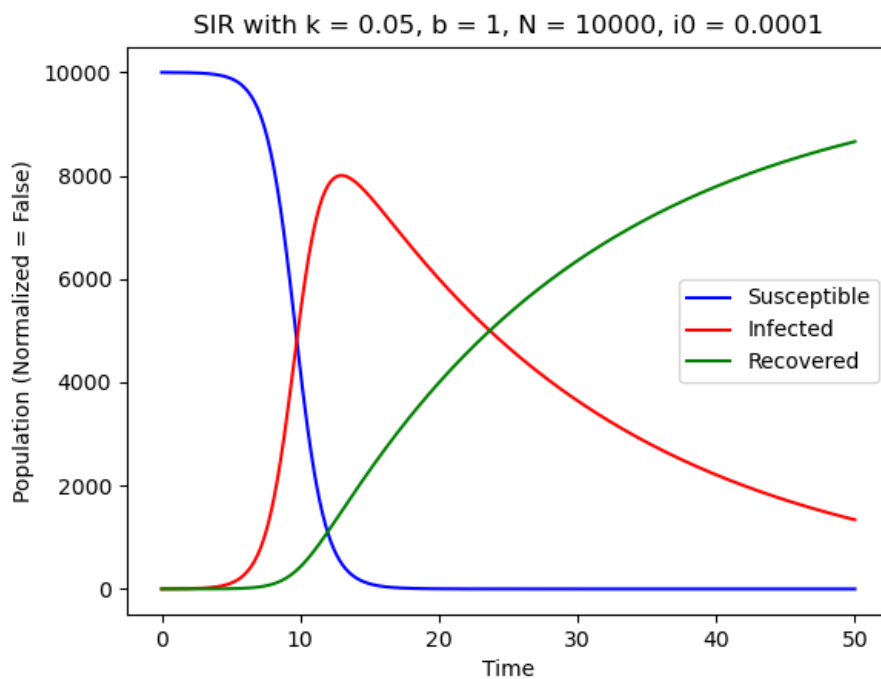


FIGURE 2. A simulation of an SIR model in a “socially distanced” society.

peak of the total number of infected people at any given time. This relationship between b and the height of the red curve ($I(t)$) is perhaps what is behind the idea of “flattening the curve” via social distancing.

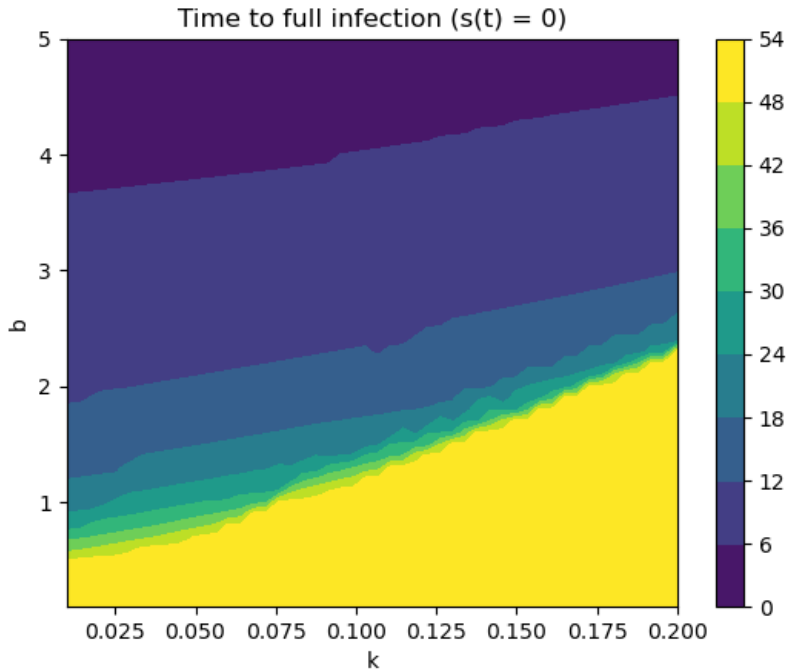


FIGURE 3. Phase diagram relating parameters and time to full infection.

We investigate further how the interplay between b and k can affect qualitative features of the dynamics. We first examine how the parameters affect the time it takes for all individuals to have been infected at some point (i.e. how do the parameters affect t^* where t^* is the earliest time that $S(t) = 0$). Running many simulations over different configurations of b and k , we obtain the phase transition diagram in Figure 3. The phase transition diagram agrees with basic intuition. Generally speaking, the time to full infection decreases as b increases. Further, the time to full infection decreases as k decreases. For low values of b and larger values of k , there is actually never a time that $S(t) = 0$. In other words, there are people in the population who never have been infected. Note that low values of b and larger values of k correspond to situations where people engage in social distancing for a disease that most people can recover quickly from. As such, such a parameter configuration is not very realistic. However, it is interesting to see that strict social distancing measures do work when recovery rates are small (i.e. low values of both b and k). In the lower left hand corner of the phase diagram, we do see that small b can lead to scenarios in which some people never get infected.

We also investigate how the choice of b and k affects the time to which over half the population is infected and not recovered ($i(t) > 0.5$). Such situations are severe ones in which the hospital system is overburdened with the tremendous number of active infections. While this event is related to the event in which $S(t) = 0$, investigating the time at which $i(t) > 0.5$ gives a more transparent indication to the severity of the outbreak. After all, there are many situations in which everyone may become infected but we always have $i(t) < 0.5$. This investigation thus sheds light on which parameter

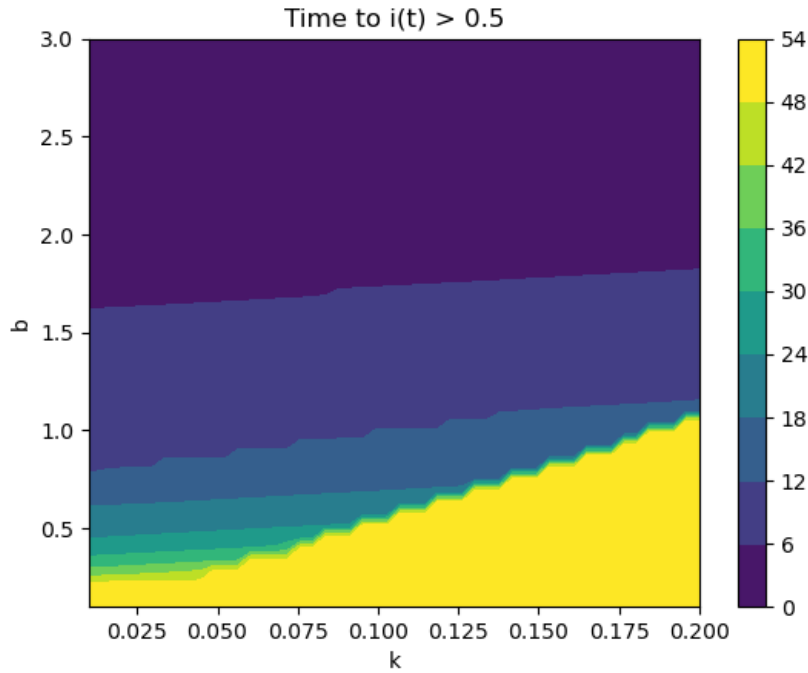


FIGURE 4. Phase diagram relating parameters and time to $i(t) > 0.5$.

configurations yield severe outcomes. Running the simulation over many configurations of the parameters, we obtain the phase diagram in Figure 4.

While the phase diagram in Figure 4 is similar to the diagram in Figure 3, it must be noted that the yellow region is shifted down (note the scale for b is different in the two plots). The diagram agrees with basic intuition. When b is small and k is large, there is low spread of disease from infected members to susceptible members, and many infected members are quickly recovering. The story is the same as in 3 for the other phase diagram regions as well. As we mentioned before, it's interesting to see that the yellow region is shifted down compared to Figure 3. This result is surprising in that we had originally thought that there may be situations in which $s(t)$ hits zero yet $i(t) < 0.5$ always. Instead, it seems that the reverse is much more likely, in that $i(t) > 0.5$ is achieved yet $s(t) > 0$ always. This is seen in the regions that are yellow in Figure 3 and non-yellow in 4 (often b is moderate).

We now switch to discrete SIR model and run simulations to see if discrete and continuous models agree on the dynamics of disease spread. We start by using the same set of parameters that resulted in Figure 1 for our first discrete simulation. From Figure 5, we can see that the discrete model is in agreement with the ODE version in terms of its overall projection on how the state of the individuals will change over time. In particular, the number of infected reaches its peak at around the same time and susceptible people quickly drops to zero as shown in ODE model. The "mirroring dynamics" of the infected and the recovered observed in Figure 1 remains largely the same as well. A small difference is that the curves for discrete model are not as smooth as the ODE version since both the Time and number of people in different states change in discrete fashion.

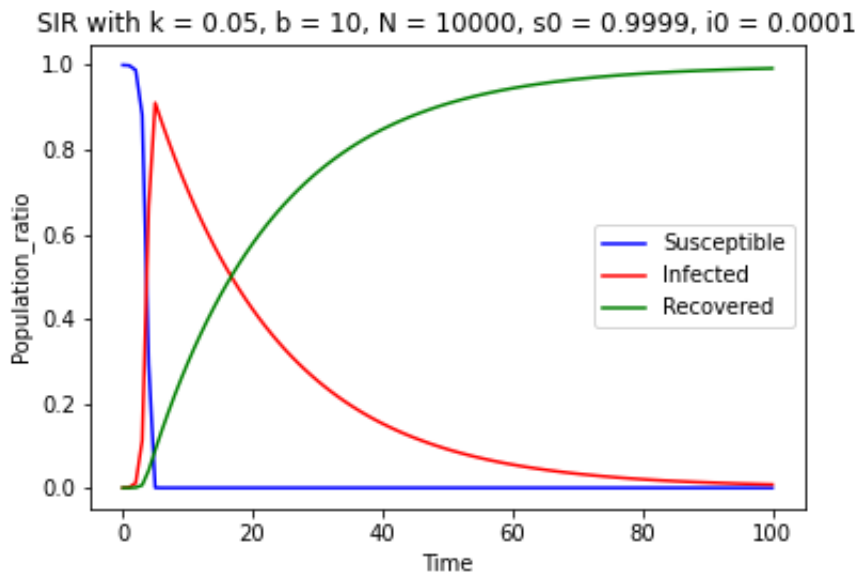


FIGURE 5. A simulation of discrete SIR model.

Next, we investigate whether the discrete model agrees or disagrees with the continuous model on how the number of interactions each person makes per day on average influences the dynamics of disease spread. In particular, we are interested in seeing if the discrete model will produce a different projection on the efficacy of social distancing on slowing down the spread of infectious disease. Running the same simulation with the parameter b now set to 1 instead of 10, as in ODE social distancing simulation, we obtain Figure 6. Except for quantitative details such as exactly when the maximum number of infected people are reached and how quickly people get infected or recovered, Figure 2 and Figure 6 are qualitatively very similar; it tells us that social distancing is an effective measure in fighting the infectious disease by 'flattening the curve' (as represented by the slope with which the red curve reaches its peak).

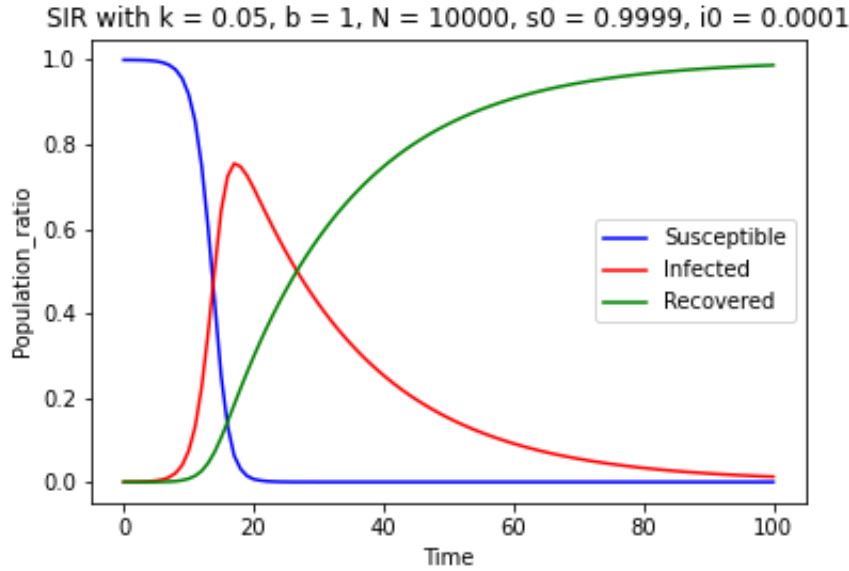


FIGURE 6. A simulation of discrete SIR model in a “socially distanced” society.

We can also use discrete model to investigate how b and k can affect the dynamics of the system. First, we look at the time for all people being infected ($s(t) = 0$). From the phase diagram in Figure 7, we can draw the similar conclusion as the ODE version that large b and small k lead to the decrease of time for full infection. For small enough b and large enough k , there will always be people that are not infected. The major difference between discrete and ODE results is the phase boundary. For discrete model, the boundary is more step-like and have more sudden changes along b and k axis. This is again due to the discrete change of data in discrete model.

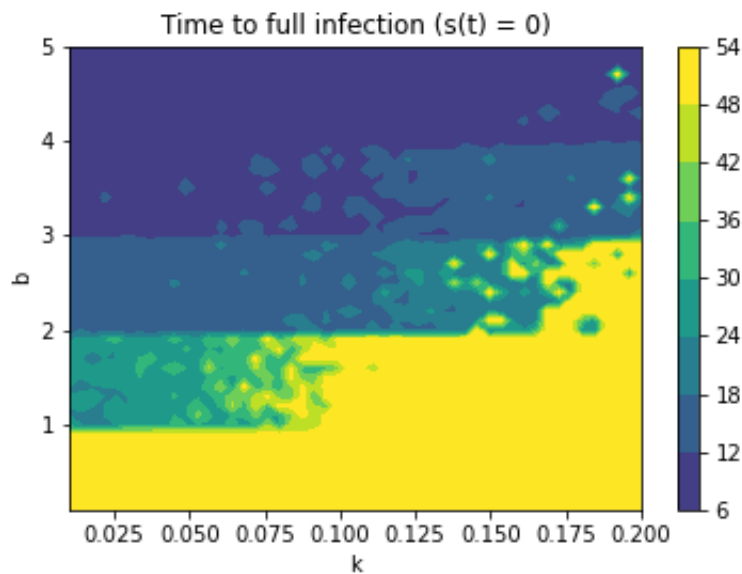


FIGURE 7. Phase diagram relating parameters and time to full infection for discrete model.

Finally, we can take a further look at the scenario where the infected people prevail in number ($i(t) > 0.5$). Similar to Figure 4, we can see that the phase boundary for discrete model, as shown in Figure 8, shifts towards smaller b and larger k values, which again indicates that there are some instances where $i(t) > 0.5$ is fulfilled while $s(t) > 0$.

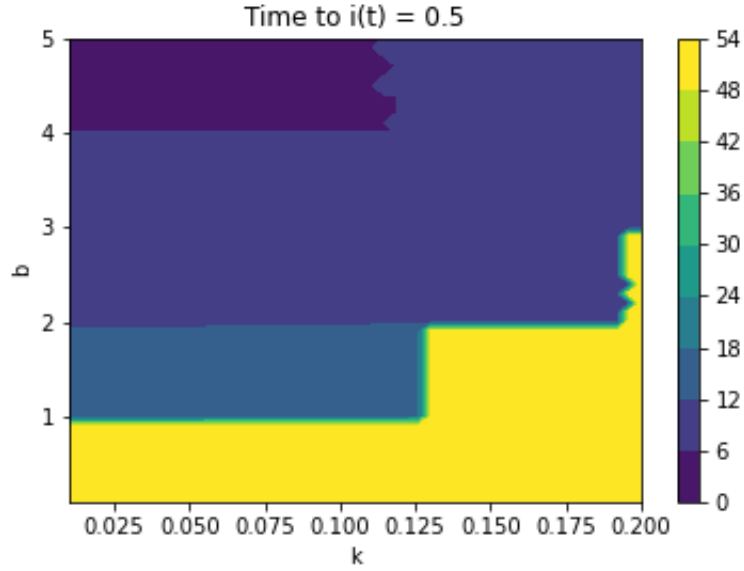


FIGURE 8. Phase diagram relating parameters and time to $i(t) > 0.5$ for discrete model.

4. PROPOSED VARIATIONS

4.1. Reinfection and spatial structure: a continuous SIRS model. Many diseases do not grant immunity to those who have recovered after being infected. In the case of viral infections, there are often different strains in the population at the same time. Those who have been infected once by one strain can be infected again with the other strain. One can then pose the interesting question: when is a disease completely eradicated from a population, and when does the disease recur?

To model reinfection, we can modify the SIR model to obtain the SIRS model in which individuals who are in the recovered group can transition to the susceptible group. We introduce a new parameter m to denote the rate at which recovered individuals become susceptible [1]. (Here, m stands for mutation with the interpretation being that recovered individuals become susceptible as the pathogens mutate.) Modifying the SIR model, we obtain the following system of coupled differential equations

$$\begin{aligned} s'(t) &= -b \cdot s(t) \cdot i(t) + m \cdot r(t) \\ i'(t) &= b \cdot s(t) \cdot i(t) - k \cdot i(t) \\ r'(t) &= k \cdot i(t) - m \cdot r(t). \end{aligned}$$

To cast our earlier questions in the context of this model, we can ask the following question. For which configurations of the parameters (b, k, m) do we get $i(t) = 0$ at some time point t (i.e. disease is eradicated)? For which configurations do we have $i(t) > 0$ for all time points t (i.e. disease is endemic)?

As stated, this model posits that any individual in the infected group can interact with any individual in the susceptible group. This is not a realistic assumption as people only interact with people who are nearby. In particular, interactions are spatially structured, and this spatial structure has implications for how fast the disease can spread [2]. This spatial structure will have real implications for simulating disease spread at large geographical scales (such as the spread within a country or a continent rather than just a city or a town). To account for this, we modify the SIRS model and define a two dimensional spatial SIRS model [3]. We obtain the following system of differential equations

$$\begin{aligned} \frac{\partial s(x, y, t)}{\partial t} &= -b \cdot s(x, y, t) \cdot \tilde{i}(x, y, t) + m \cdot r(x, y, t) \\ \frac{\partial i(x, y, t)}{\partial t} &= b \cdot s(x, y, t) \cdot \tilde{i}(x, y, t) - k \cdot i(x, y, t) \\ \frac{\partial r(x, y, t)}{\partial t} &= k \cdot i(x, y, t) - m \cdot r(x, y, t) \end{aligned}$$

where

$$\tilde{i}(x, y, t) = \int_{\mathbb{R}^2} \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(x-x')^2 + (y-y')^2}{2\tau^2}\right) dx' dy'.$$

In this model, the susceptible individuals at location (x, y) at time t interact with infected people at other locations at time t with weights exponentially decaying in the distance to those other locations. The parameter τ^2 determines the “interaction length”.

With this model, we can then investigate how travel restrictions affect disease spread. A travel restriction corresponds to a decrease in the value of τ^2 . For a choice of (b, k, m) we can ask whether there are choices of τ^2 that lead to the disease being eradicated. Moreover, we can investigate the relative importance of social distancing vs travel restrictions, i.e. what is the interplay between b and τ^2 ?

4.2. The effects of 'lock down' on disease spread: an agent-based SIRS model.

In face of the outbreak of Covid-19, social distancing interventions or more strict 'lock down' measurements are introduced to slowdown the spread of the disease. Although people in principle agree on the necessity of this measurement, in many western countries, many questions around it is still under debate even today. These questions include when to introduce 'lock down', how strict it should be, how long we should remain in 'lock down' or quarantine, and etc. To demonstrate the effects of 'lock down' on the dynamics of disease spread and shed light on the questions mentioned above, we propose to use a simple agent-based (discrete) SIR model to simulate the evolution of the disease.

To achieve this goal, there are some modifications needed to be added to the basic SIR model. In current model, each people can interact with b people to potentially transmit the virus, which is not realistic since people should only interact with those who are nearby. Thus, in order to introduce the spatial structure into the model, we transfer the discrete model onto a 2 dimensional grid with length L . You can image it as a country with area $L \times L$. And people can now only interact with each other when they are within a distance of D_0 . By tuning this new parameter D_0 , we can now mimic the effects of 'lock down' measurement [4].

For example, we can make D_0 time-dependent [4]:

$$D_0(t) = \begin{cases} D_0 & \text{for } t \leq T_0 \\ D_0 e^{-(t-T_0)/\lambda} & \text{for } t > T_0 \end{cases} \quad (1)$$

The T_0 represents the timing of 'lock down' while the λ represents how fast the 'lock down' take into effect. Assuming $\lambda > 0$, as $t \rightarrow +\infty$, $D_0 \rightarrow 0$. Thus, people are being confined into a small place and cannot affect each other. By changing T_0 and λ and monitoring how the three types of population evolve with time, we can then answer the questions relate to the effects of 'lock down'.

Furthermore, we can set a time T_1 when we release the 'lock down' measure and $D_0(T_1) = D_0$ to see how the pandemic will evolve if we end 'lock down' too early. Along with these parameters, we can also tune b to simulate the effectiveness of wearing masks or social distancing on slowing down the spread of disease. To present the results, some 2 dimensional phase diagrams can be useful. For agent-based simulation in grid, short animation illustrating how the infected population evolves (showing by color dots) can be very powerful.

4.3. SIR model that depends on current state. In reality, a single system of ordinary differential equation might not accurately capture the patterns in changes of $s(t)$, $i(t)$ and

$r(t)$ throughout the process. For example, when a significant size of the population are infected, the government intervention and media exposure to the virus may potentially alter people's behavior. Moreover, after medical supplies are allocated and feasible treatment has been found, infected people may recover quicker. In these situations, once $i(t)$ or $r(t)$ reaches a certain threshold, there is fundamental changes to the systems of differential equations. In particular, the parameters b and k would change in response to the real world incident, and the system of equation now becomes

$$\begin{aligned} s'(t) &= -b' \cdot s(t) \cdot i(t) \\ i'(t) &= b' \cdot s(t) \cdot i(t) - k' \cdot i(t) \\ r'(t) &= k' \cdot i(t) \end{aligned}$$

where

$$\begin{aligned} b' &< b \\ k' &> k \end{aligned}$$

in the discussion above. Using this model, we would be able to see the effects of public policies in the middle of simulation. We may want to answer, given a fixed initial condition, how does the time such alternation occur impact the peak of infection. Given the effectiveness of a policy or treatment (fix b' and k'), what would be the last day for such alternation to take place in order to cap the maximum infected population under certain threshold for various initial conditions? One way to answer these questions is to set an event such as $i(t) = 0.1$, which terminates the original ode simulation, and run a different simulation with the new parameters for the remaining time interval. Further analysis would be based on the concatenated solution of the two systems.

4.4. "Light at the end of the tunnel: SIRV model". With the exciting announcement from the biotech companies Pfizer and BioNTech that Covid-19 vaccines with 90 percent efficacy will soon be available to the public, one might ask how introducing 'vaccinated people' into the model will change the dynamics we saw in discrete disease spread model. Vaccines will not only slow down the spread of infectious disease in time but it will also affect how the disease will propagate in space. Change in the spatial dynamics of infectious disease might in turn have a profound impact on how quickly disease will spread in time.

In this variation, we will attempt to dissect the role of the vaccines in time and space domains in fighting infectious disease, to gain a more detailed understanding of (and hopefully more confidence on) how the vaccines might end Covid-19 pandemic. In order to achieve this we plan to simulate disease spread in a well-controlled, 2-dimensional toy world and keep track of a few metrics over time that will help us understand the temporal and spatial dynamics of the disease spread in the presence and absence of the vaccines. Without being exhaustive about details on assumptions and implementation strategies, we propose to introduce following key ideas:

1. 2D grid as a toy world: Each cell in an M by N 2-dimensional grid represents a person with an internal state. In addition to S,I,R we now have V representing 'vaccinated' into the model, making it a SIRV model.
2. The number of infected people as a metric for the efficacy of the vaccines in the time domain: Every person in the grid will be stationary throughout a simulation interacting only with its immediate neighbors (as might happen with maximum social distancing practice and travel restrictions). This will enable us to have a control over space domain and keep track of how the number of infected people evolve over time, both locally and globally as well as in the presence or absence of the vaccines.
3. 'Grid averaging' as a graphical metric for the efficacy of the vaccines in the space domain: Since it is difficult to summarize the spatial dynamics of infection with a number, we propose to use a graphical tool to help us understand how the vaccines may slow down the spread of disease in the space domain. This can be achieved by, for instance, by observing time-binned internal states of the grid (which we call 'grid averaging' that consists of splitting time into a few bins, averaging the internal states of each cell of the grid in each bin and displaying the averages). This is roughly equivalent to taking a snapshot of the grid state in discrete time.

We acknowledge the possibility that the inhomogeneity of the grid in terms of its internal state and the way the vaccines will be distributed (e.g. where to dispense and at what rate) may obscure the simulation results and make them hard to interpret. These variables will be controlled as much as possible, but they might be interesting variables to play with in the sense that they might help shed light on under what conditions the vaccines' efficacy in time and space domains may differ. For instance, interesting questions include if prioritized vaccination based on the internal states (susceptible first) will result in a faster recovery of the community as our intuition suggests.

REFERENCES

- [1] Fred Brauer. Compartmental Models in Epidemiology. 1945:19–79.
- [2] Romain Ducasse. Qualitative properties of spatial epidemiological models.
- [3] Norman T. J. Bailey. The simulation of stochastic epidemics in two dimensions. The Regents of the University of California.
- [4] Efthimios Kaxiras and Georgios Neofotistos. Multiple epidemic wave model of the covid-19 pandemic: Modeling study. Journal of Medical Internet Research, 22(7):e20912, 2020.