Exam: Epidemic modelling

Candidate nr. 10037
Department of Physics, Norwegian University of Science and Technology, Trondheim
Norway
TFY4235 - Computational physics
(*Last updated on May 4, 2021*)

Abstract

Short abstract

Contents

PART
I

Introduction

PART
II

Code overview

PART III

Results and Discussion

1 Problem 2A: Deterministic SIR model

1.1 a)

The ODE-solver I chose to use is the one I made for exercise 2, as I am very familiar with this, and found it working reasonably good for these kind of purposes. The implementation is an object-oriented version of the solver methods shown in the lectures and should be easily understood by the documentation in `ode.py`.

I solve the deterministic SIR equations in terms of fractions of people. As seen in figure 1 the asymptotic expressions for $S(t)$ and $R(t)$ as given in the exam-sheet [Nor21]. The expressions for $S(\infty)$ and $R(\infty)$ are solved using the non-linear equation solver `fsolve` from `scipy`.

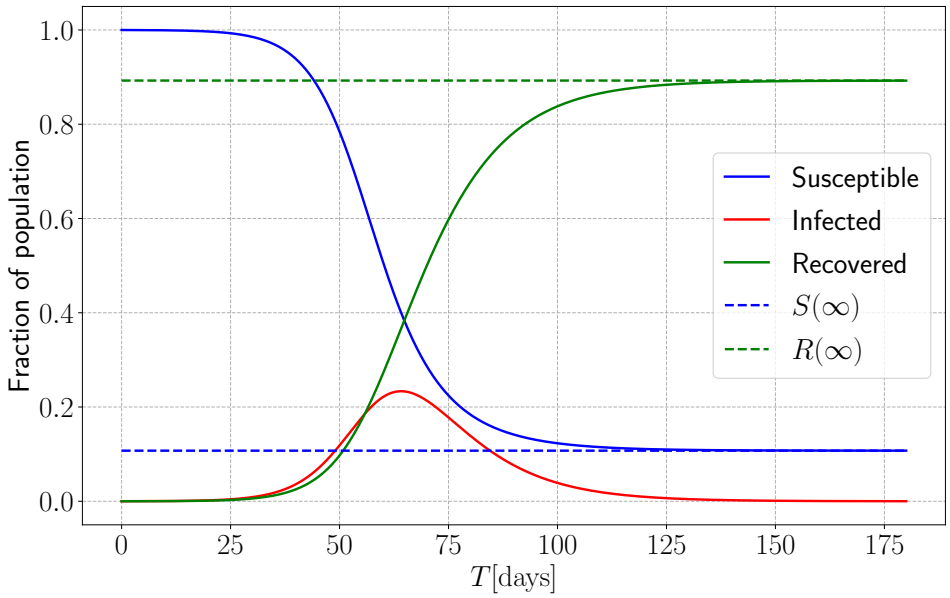

Figure 1: SIR equations with $\beta = 0.25\,\mathrm{day}^{-1}$, $\tau = 10\,\mathrm{day}$.

1.2 b)

For the early developments of the epidemic, the expression for I can be simplified to [Nor21]

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (\beta - 1/\tau)I. \tag{1}$$

The solution of this equation is given by

$$I(t) = I(0)\exp\left([\beta\tau - 1]\,\frac{t}{\tau}\right) = I(0)\exp\left([\mathcal{R}_0 - 1]\,\frac{t}{\tau}\right),$$

where we have introduced $\mathcal{R}_0 = \beta\tau$. As seen in figure 2, the fraction of infected people matches this expression quite closely during the first 40 days or so. This confirms the exponential growth in the beginning.

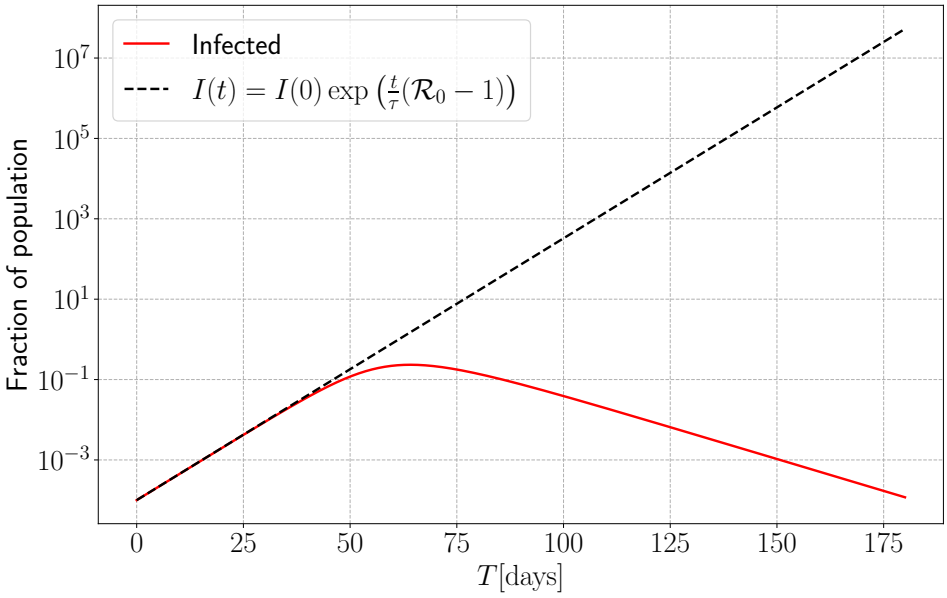

Figure 2: Infected people compared with the analytical approximation at the early stages.

1.3 c) Flattening the curve.

As seen from figure 1, the peak of the infected people is rather close to being at 0.2 when $\beta = 0.25$. Remark also that the peak occurs in the early stage of the epidemic. We use these observations to our favour when finding the *largest* β that will keep I/N[1] smaller than 0.2 during the pandemic. To do this, I choose the following procedure:

Choose a starting value for $\beta = \beta_0$;
Choose a tolerance `tol`;
Run the simulation of the SIR with the same parameters as in 2Aa, except with $\beta = \beta_0$.;
Calculate $\mathcal{I} := \max_{t\in[0,180]} I(t)$.;
while $err := |\mathcal{I} - 0.2| > tol$ **do**
  if $\mathcal{I} > 0.2$ **then**
    $\beta \leftarrow \beta \cdot 2^{\mathrm{err}}$
  else
    $\beta \leftarrow \beta \cdot 2^{-\mathrm{err}}$
  end
end
Run the simulation of the SIR with the same parameters as in 2Aa, except with the new value for β.;
Recalculate $\mathcal{I} := \max_{t\in[0,180]} I(t)$.;

Algorithm 1: Finding the largest beta keeping $\max I$ *less* than 0.2.

[1]Note that I use fractions of people instead of actual numbers of people in this section, so that is why I refer to I as if it is I/N in the procedure below.

This procedure is by no means any advanced piece of algorithm, but it ensures that beta is increased when the maximum of I is less than 0.2, and decreased when it is larger than 0.2. Further, it also ensures that the "nudging" of beta in each step is less when the error is less. To get an estimate for the *largest* β one can have to keep I/N *less* than 0.2, one should start off with an initial value β_0 for which the peak is less than 0.2, so that one reaches the limit from below. Using $\beta_0 = 0.2$, one finds with this the maximum value of $\beta = 0.28020370$, as given in table 1 together with the deviation $0.2 - \max_{t\in[0,180]} I(t)$ this value gives rise to.

1.4 d) Vaccinations

To find the minimum fraction of vaccinated people preventing an outbreak, $R(0)/N$, I use the following procedure:

Choose a starting value for $R(0) = R_0$;
Run the simulation of the SIR with the same parameters as in 2Aa, except with $R(0) = R_0$;
Calculate the `slope` of the fraction of infected people in a semi-log plot for the first 15[a] days of the simulation;
If the initial slope is negative it indicates that the outbreak dies out by itself;
while *slope*> 0 **do**
  $R(0) \leftarrow R \cdot 2^{\texttt{slope}}$;
  Run the simulation of the SIR with the same parameters as in 2Aa, except with the new value for $R(0)$.;
  Recalculate the `slope` in the semi-log axes.;
end

Algorithm 2: Finding the minimum fraction of initially vaccinated people for outbreaks to be impossible.

[a] As people are typically sick for 10 days ($\tau = 10$) I suppose this time scale is long enough.

Table 1: The maximum value of β giving a peak less than 0.2 of the infected fraction, and the minimum value of $R(0)$ (vaccinated) avoiding exponential growth.

Parameter	value	$0.2 - \max_{t\in[0,180]} I(t)$	Initial log-slope
β	0.28020370	$8.319 \cdot 10^{-7}$	—
$R(0)$	0.59987499	—	$-1.74 \cdot 10^{-15}$

2 Problem 2B: Stochastic SIR model

2.1 a)

As seen in the plot in figure 3, the 10 different realisations of the stochastic SIR model , shown in dashed lines in graded colours, seem to lie close to the deterministic solution.

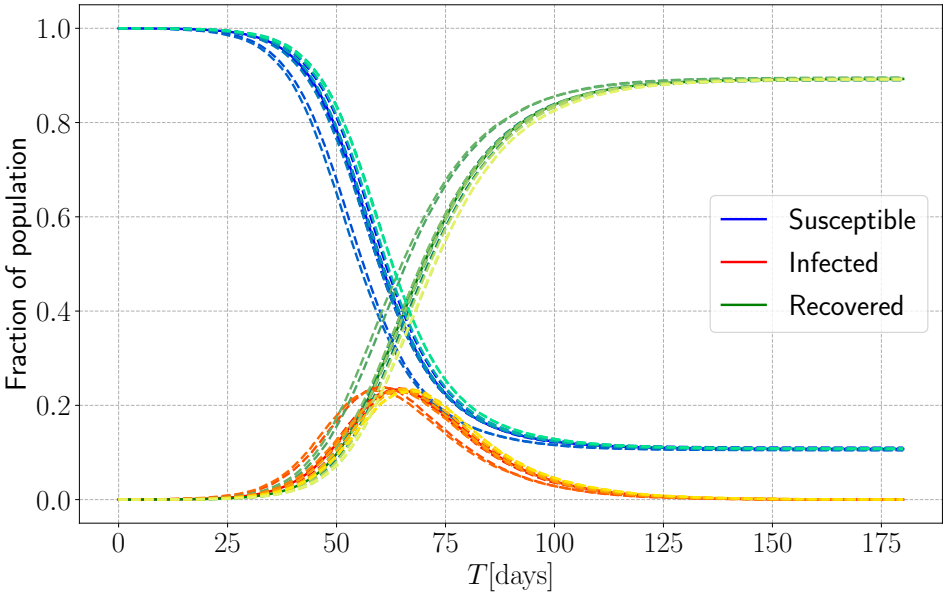

Figure 3: Solution of stochastic SIR equations with $\beta = 0.25\,\mathrm{day}^{-1}$, $\tau = 10\,\mathrm{day}$.

2.2 b)

As in the previous exercise, we plot the fraction of infected people together with the analytical model for the early stages. This is shown in figure 4. Here we clearly see that all the realisations are approximately linear in the semi-log plot in the first 40 days, or so, of the simulation as we observed in the previous exercise.

2.3 c) Probability of an outbreak

There will always be a certain probability for an outbreak disappearing by itself for the stochastic model. In the present subsection we estimate this probability for an initial number of infected people $= 1, 2, \ldots, 10$. This is done by the following procedure:

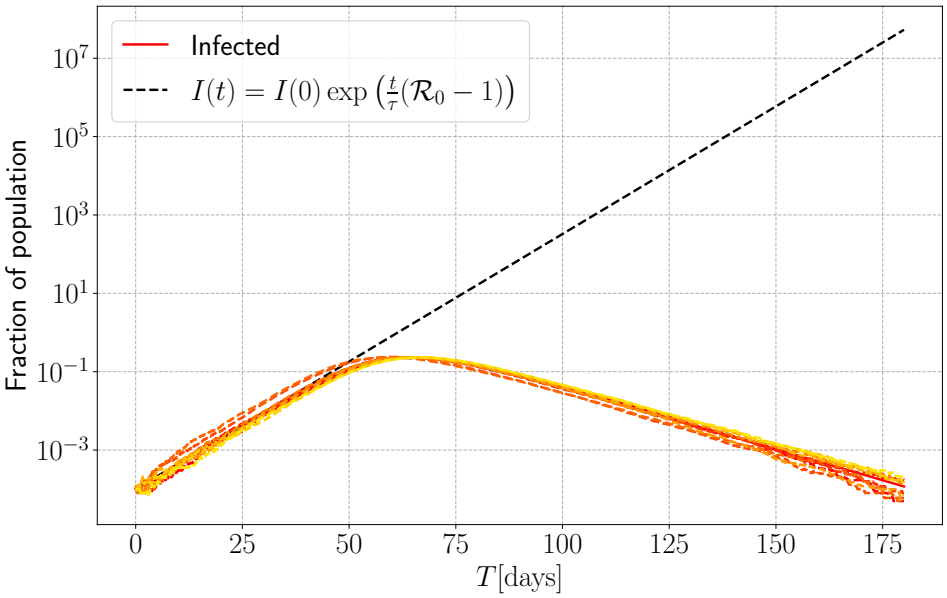

Figure 4: Infected people compared with the analytical approximation at the early stages. Stochastic and deterministic model.

Choose the parameters of the model as those given in the exam sheet [Nor21], but with $T = 30$ days[a]. ;
Choose a batch-size B.;
for $I = 1, 2, \ldots, 10$ **do**
  Initialise an empty vector of length B: $\mathbf{X} = [0, \ldots, 0]$.;
  for $n = 1, \ldots, B$ **do**
    Run the simulation with initial number of infected people $= I$.;
    Calculate the `slope` in the semi-log axes for $I(t)$.;
    if *slope* $<= 0$ **then**
      $X_n = 0$
    else
      $X_n = 1$
    end
  end
  Estimate the probability of an outbreak for I initially infected by

$$p := P(\text{outbreak}|I) = \frac{1}{B}\sum_{n=1}^{B} X_n.$$

  Calculate the standard deviation of the estimate by

$$\sqrt{\text{Var}(\hat{p})} = \sqrt{\frac{p(1-p)}{B}}. \tag{2}$$

end

Algorithm 3: Calculating the probability of an outbreak as a function of the initial number of infected people, I.

[a] As the typical infection time is 10 days, I assume 50 days to be sufficient for detecting an outbreak in the stochastic model.

In performing this procedure, I use a batch size of 500 in each sweep. What we are estimating here is essentially a Bernoulli-distributed random variable, X: X can take the realisations 1 or 0 with probabilities p and $1-p$ respectively, and we assume them to be independent [Was04, p.26]. Then, as the variance of such a distribution is $p(1-p)$, the variance of the estimator for the probability, namely $\hat{p}$ i.e. the expectation value of X_n, is

$$\mathrm{Var}(\hat{p}) = \sum_{n=1}^{B} \mathrm{Var}\left(\frac{X_n}{B}\right) = \frac{1}{B^2}\sum_{n=1}^{B} \mathrm{Var}(X_n) = \frac{1}{B^2}\sum_{n=1}^{B} p(1-p) = \frac{p(1-p)}{B},$$

from which formula (2) follows.

The probability of an outbreak as a function of the initial number of infected people are shown in figure 5 together with the associated standard deviation. As seen from this plot, when there are more than 6 people initially infected, the probability is approximately 1 that an outbreak will happen. However, for e.g. 1 initially infected person, the probability is less than 0.6. This ultimately shows that the stochastic model has a more realistic feature to it than the deterministic one, in that these scenarios might occur.

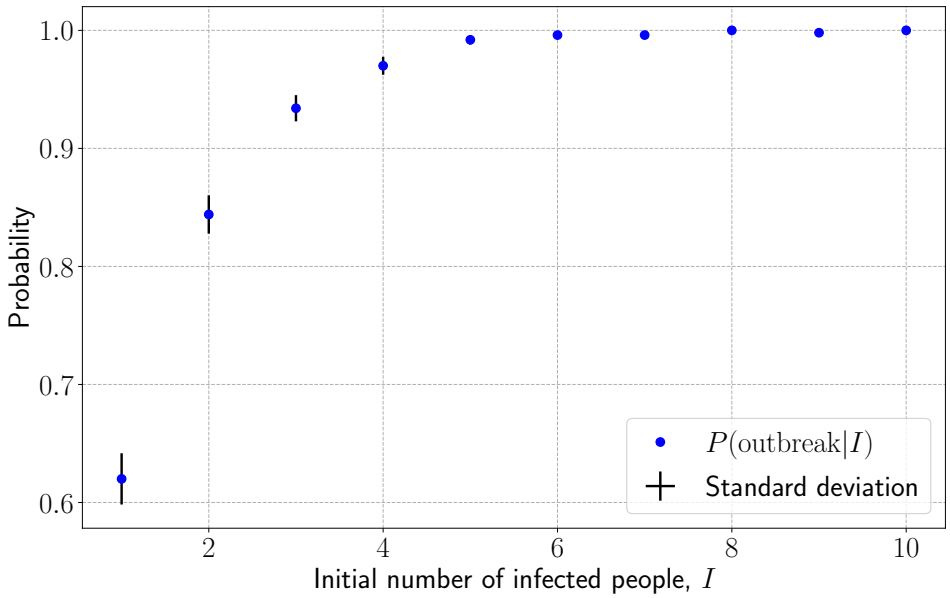

Figure 5: Probability of an outbreak as a function of initial number of infected people.

3 Problem 2C: Stochastic SEIIaR model

3.1 a)

The time-development of all the variables in the SEIIaR model is shown in figure 6, showing 10 realisations of the simulation. To compare it easily with the deterministic model, we contract E, I, I_a into one variable, so that they together represent all the infected people. The solutions of the deterministic model are shown together with the contracted versions of these simulations in figure 7.

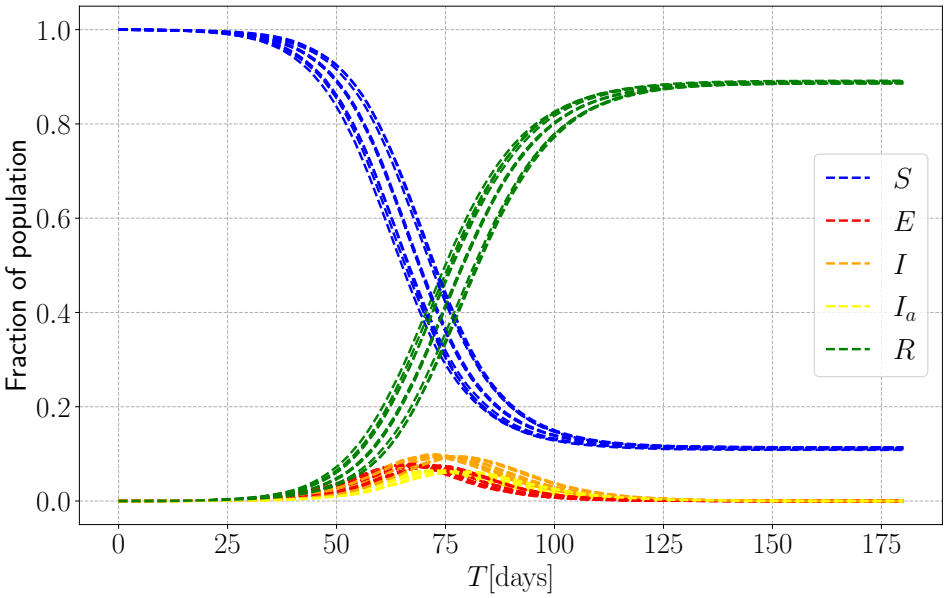

Figure 6: Solution of the stochastic SEIIaR-equations.

One prominent feature to notice here, is that the process overall seems to be *slower*, in the sense that it takes longer for people to get infected and eventually recover, but the peaks and asymptotic behaviours seem to align quite closely. This is probably due to the fact that this model includes a period in which people are infected but not yet able to infect anybody else, an incubation period of typical length $\tau_E = 3$ days. This will naturally delay the process, but it should not affect the peak nor the asymptotic behaviour. Ultimately, this shows that the SEIIaR model adds another layer of realism to our model in the sense that people do not get sick right away.

To test that the implementation is correct — by comparing with the deterministic and stochastic SIR model — we adjust the parameters of the SEIIaR to $\beta = 0.25$, $r_s = 1$, $r_a = 1$, $\tau_E = 0$, $\tau_I = 10$ and start with 10 out of 100 000 people initially infected. The results of doing 1000 such simulations and performing the average of the stochastic models is shown in figure 8. This shows that the two stochastic models are more or less identical, as expected.

3.2 b) Probability of outbreak dependence on r_s

The variable r_s describes how infectious a person is when he is in the infection state. Reducing this constant below 1 can therefore correspond to emulating the degree of self-

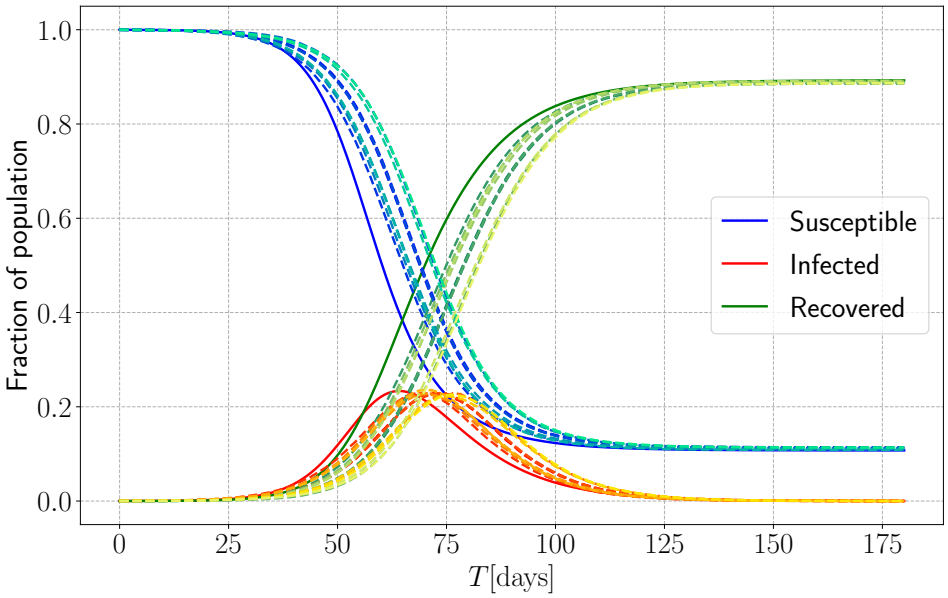

Figure 7: Comparison of the solution of the Stochastic SEIIaR-equations with the deterministic SIR-model. The number of infected people I in the stochastic model is $E+I+I_a$.

isolation when people are symptomatic. We investigate the probability of an outbreak as a function of this self-isolation-rate r_s by the following procedure (which is essentially the same as that shown in $2Bb$ expect we are now finding the probability as a function of r_s):

The results of this calculation using a batch size of $B = 500$ and 100 values of r_s are shown in figure 9. The behaviour is as expected: when people are hardly infectious when symptomatic the probability of an outbreak is close to 0. This is also probably also a result of the fact that $r_a = 0.1$, i.e. it is not particularly likely to infect anyone when you are asymptomatic. If r_a and f_a was higher, one would expect the probability distribution to stagnate at some finite value when $r_s \to 0$, or at least go to 0 slower. This is demonstrated in the same figure, where we set $r_a = 1$ and perform the same test as above. This again is an intuitive confirmation that the model behaves as expected, and therefore is correctly implemented.

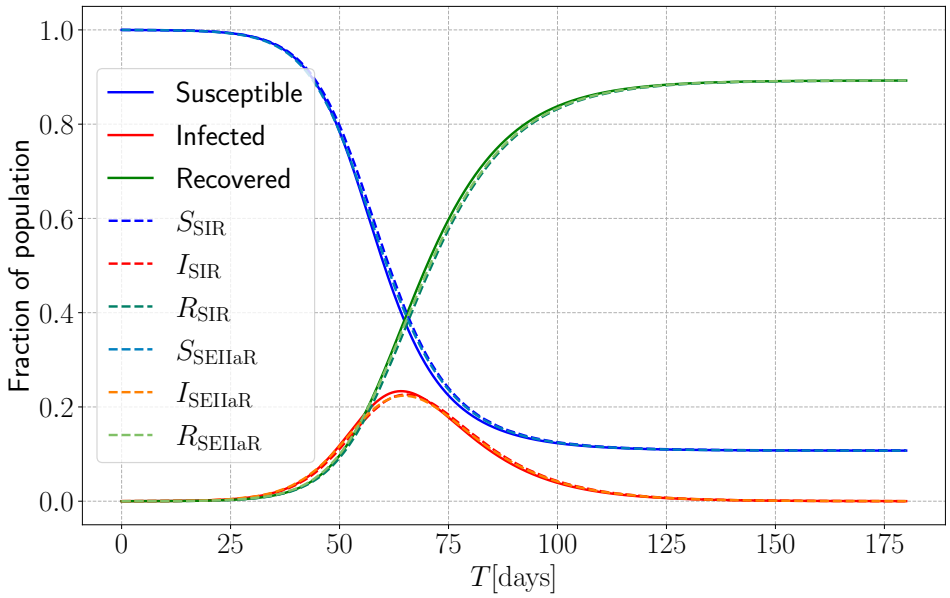

Figure 8: Solution of the stochastic SEIIaR-equations compared with the stochastic and deterministic SIR-equations for the case of identical parameters.

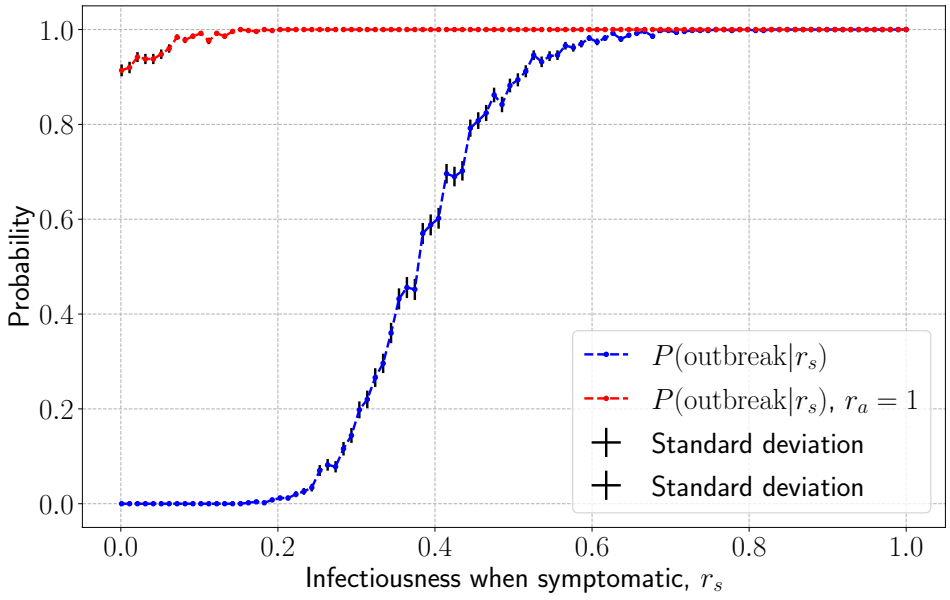

Figure 9: Probability of an outbreak as a function of r_s.

Choose the parameters of the model as those given in the exam sheet [Nor21], but with $T = 30\,\text{days}$[2]. ;
Choose a batch-size B.;
Choose a number of values n of r_s to try;
Select n values of r_s, equally spaced between 0.001 and 1 :

$$\mathbf{R} = [R_1, \dots R_n]$$

for $i = 1, 2, \dots, n$ **do**
  Initialise an empty vector of length B: $\mathbf{X} = [0, \dots, 0]$.;
  for $n = 1, \dots, B$ **do**
    Run the simulation with $r_s = R_i$.;
    Calculate the `slope` in the semi-log axes for $I(t)$.;
    if *slope* <= 0 **then**
      $X_n = 0$
    else
      $X_n = 1$
    end
  end
  Estimate the probability of an outbreak for I initially infected by

$$p \coloneqq P(\text{outbreak}|r_s) = \frac{1}{B}\sum_{n=1}^{B} X_n.$$

  Calculate the standard deviation of the estimate by

$$\sqrt{\text{Var}(\hat{p})} = \sqrt{\frac{p(1-p)}{B}}.$$

end

Algorithm 4: Calculating the probability of an outbreak as a function of r_s.

4 Problem 2D: Stochastic SEIIaR Commuter model

4.1 a) Commuter model for a two-town system

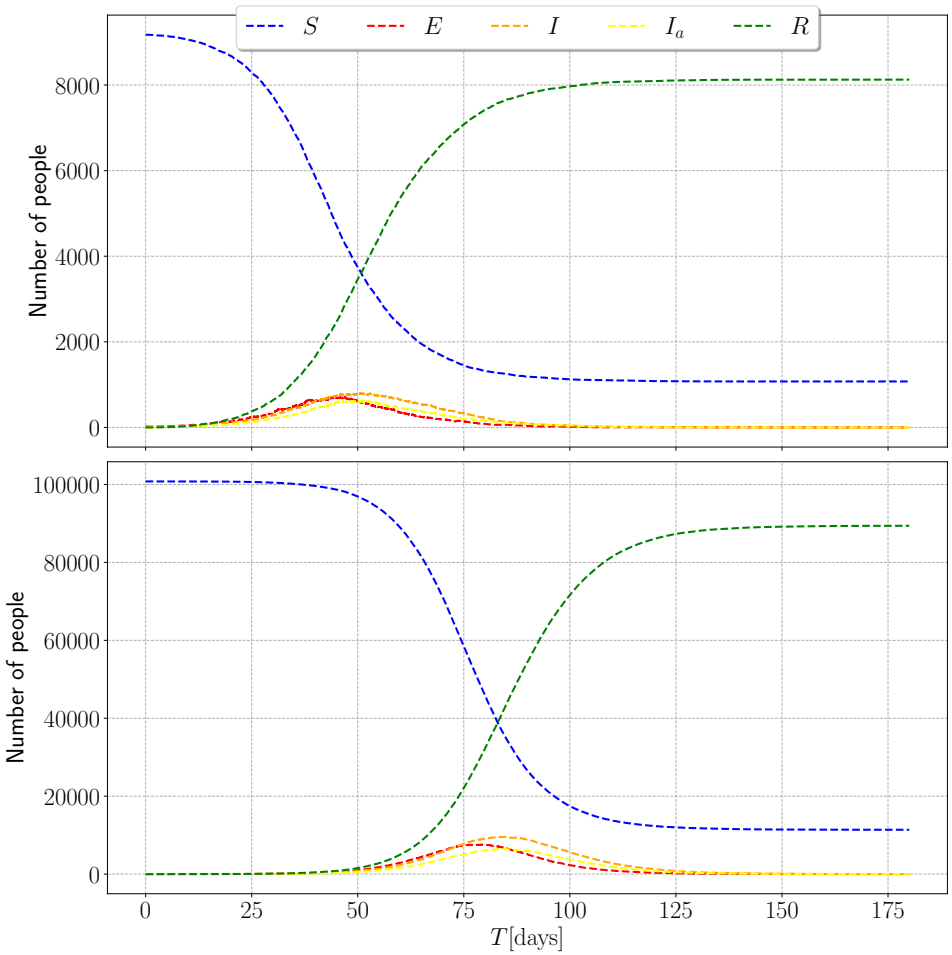

Figure 10: Solutions of Stochastic SEIIaR commuter model for the 2-city case.

4.2 b) Description of implementation & tests

To check that the system behaves as expected, we simulate the same scenario, only with another matrix which now represents *no* flow of workers to the other areas:

$$\mathbf{M} = \begin{bmatrix} 10000 & 0 \\ 0 & 100000 \end{bmatrix}. \tag{3}$$

The time-evolution of the different variables are shown in figure 11, in which it is apparent that the exposed people in the small city never infect those in the large city, as we expect. The expected behaviour is also observed when the exposed start out in area 2, but a plot of this is not included, for brevity.

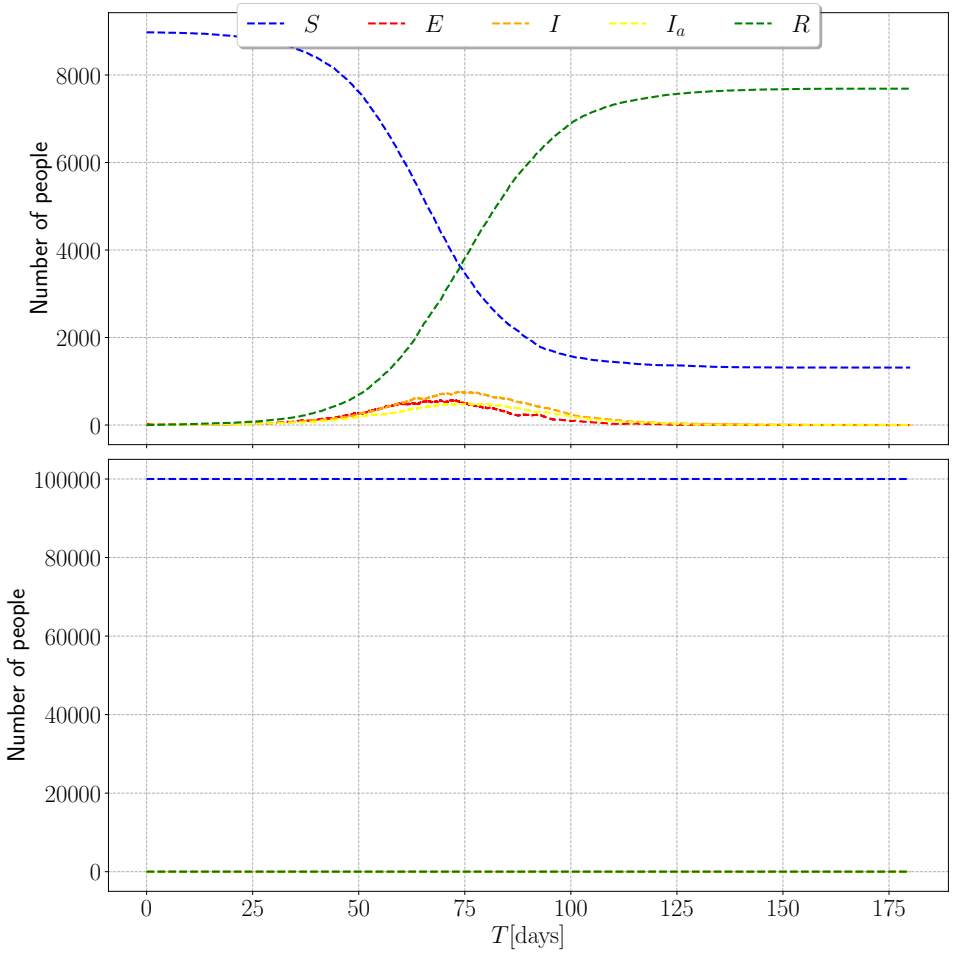

Figure 11: Solutions of Stochastic SEIIaR commuter model for the 2-city case, with the matrix **M** in equation (3).

5 Problem 2D: Larger tochastic SEIIaR Commuter model

References

[Nor21] Tor Nordam. Exam 2021 - Computational physics TFY4235, 2021.

[Was04] Larry Wasserman. *All of Statistics: A Concise Course in Statistical Inference*. Springer Texts in Statistics. Springer, New York, 2004.