Epidemics on Static and Adaptive Networks

Achterberg, M.A.

DOI
10.4233/uuid:2c93f1af-bf49-4353-b9b9-c6ed8d62d3c9

Publication date
2024

Document Version
Final published version

Citation (APA)
Achterberg, M. A. (2024). *Epidemics on Static and Adaptive Networks*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:2c93f1af-bf49-4353-b9b9-c6ed8d62d3c9

EPIDEMICS ON STATIC AND ADAPTIVE NETWORKS

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T. H. J. J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 9 januari 2024 om 15:00 uur

door

Massimo Alessandro ACHTERBERG

Ingenieur in de Toegepaste Wiskunde, Technische Universiteit Delft,
geboren te Rotterdam, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. P. F. A. Van Mieghem	Technische Universiteit Delft, promotor
Prof. dr. ir. R. E. Kooij	Technische Universiteit Delft, promotor
Onafhankelijke leden:	
Prof. dr. I. Z. Kiss	Northeastern University London, Verenigd Koninkrijk
Prof. dr. A. Pugliese	University of Trento, Italië
Prof. dr. O. Diekmann	Universiteit Utrecht
Prof. dr. F. H. J. Redig	Technische Universiteit Delft
Prof. dr. ir. G. Jongbloed	Technische Universiteit Delft, reservelid

Keywords: Mathematical epidemiology, Adaptive networks, Markov processes

Printed by: Ipskamp Printing, Enschede

Front & Back: Dennis van Duin

ISBN 978-94-6384-514-4

An electronic version of this dissertation is available at
https://repository.tudelft.nl.

CONTENTS

SUMMARY

The COVID-19 pandemic has had a disruptive impact on healthcare systems and everyday life of the majority of the people around the globe. Despite many years of research on network epidemiology, many key aspects of disease transmission and in particular the response of people to the spread of a disease, remain poorly understood. On the basis of epidemiological modelling lie the Susceptible-Infected-Susceptible (SIS) and Susceptible-Infected-Recovered (SIR) models. In this dissertation, we aim to improve the understanding of the spread of contagious diseases, with an emphasis on the interplay between disease spread and personal behaviour, applied to the SIS and SIR models.

The first part starts with the analysis of the eigenvalue spectrum of the infinitesimal generator of the Markovian SIS model with self-infections (Chapter 2). Based on the eigenvalue spectrum, which we believe encodes the majority of the dynamics, we derive an alternative definition of the epidemic threshold. We show that the epidemic threshold approximately coincides with the effective infection rate for which the third-largest eigenvalue is minimal. Contrary to the SIS process, where only an eigenvalue analysis is possible, the SIR process is completely solved on an arbitrary, heterogeneous network (Chapter 3). The benefit of the exact solution is demonstrated by analytically computing the time when the number of infections is maximal.

The second part concerns the interplay between the spread of a disease and the response of people to the disease spread. We develop the Generalised Adaptive SIS (G-ASIS) model to describe how individuals break and create links in the contact graph. The decisions for breaking or creating links are based on the viral state of the nodes attached to that link. For all 36 instances in the G-ASIS model, we analyse the relation between the epidemic threshold and the effective link-breaking rate (Chapter 4). We derive the first-order and second-order mean-field approximation of the G-ASIS model (Chapter 5) and illustrate that the second-order approximation is able to qualitatively approximate the Markovian model more accurately than the first-order approximation. The G-ASIS mean-field model is extended to arbitrary link-breaking and link-creation responses, which are not only related to the number of susceptible and infectious neighbours of a node, but may also depend on the presence of the virus in the whole population (Chapter 6). For all possible link-breaking and link-creation responses, epidemic waves cannot occur in the mean-field adaptive SIS process.

In the final part, we develop the Network-Inference-based Prediction Algorithm (NIPA) for forecasting the spread of contagious diseases on heterogeneous networks (Chapter 7). The contact graph is assumed to be unknown and is inferred by NIPA from the number of reported cases. NIPA is a hybrid method, combining epidemiological knowledge, machine-learning and networks. Network-based forecasting, and NIPA in particular, seems favourable for predicting epidemic outbreaks, which is demonstrated by showing that NIPA outperforms many other forecasting algorithms for estimating the spread of COVID-19.

SAMENVATTING

COVID-19 heeft een verwoestende impact gehad op zorgsystemen en het alledaagse leven van de meeste mensen op aarde. Ondanks jarenlang wetenschappelijk onderzoek naar epidemiologie blijven vele aspecten van ziekteverspreiding onvoldoende gekend, met name de reactie van mensen ten gevolge van de verspreiding van een besmettelijke ziekte. De Susceptible-Infected-Susceptible (vatbaar-besmettelijk-vatbaar) en Susceptible-Infected-Recovered (vatbaar-besmettelijk-genezen) modellen liggen aan de basis van vrijwel alle ziekteverspreidingsmodellen. In dit proefschrift pogen we ziekteverspreidingsmodellen beter te begrijpen, met de nadruk op de interactie tussen ziekteverspreiding en menselijk gedrag, toegepast op SIS en SIR modellen.

In het eerste gedeelte wordt het eigenwaardespectrum van de transitiematrix van het Markoviaanse SIS model met zelf-infecties onderzocht (Hoofdstuk 2). Op basis van het eigenwaardespectrum, die een groot gedeelte van de dynamica omvat, suggereren we een nieuwe definitie van de epidemische drempelwaarde. De epidemische drempelwaarde komt ongeveer overeen met de infectiesterkte waarvoor de twee-na-grootste eigenwaarde van de transitiematrix minimaal is. In tegenstelling tot het SIS proces waar alleen een eigenwaarde-analyse mogelijk is, wordt het SIR model op een heterogeen contactnetwerk exact opgelost (Hoofdstuk 3). De toegevoegde waarde van de exacte oplossing wordt gedemonstreerd door het maximale aantal tegelijke zieke mensen exact te bepalen.

Het tweede gedeelte beschrijft de samenhang tussen ziekteverspreiding en het gedrag van mensen op deze verspreiding. We ontwikkelen het Generalised Adaptive SIS (G-ASIS) model, welke beschrijft hoe individuele personen de contacten met hun naaste buren kunnen verbreken en herstellen. Het verbreken of herstellen van deze verbinding hangt af van de gezondheid van beide personen. Alle 36 mogelijke modelinstanties worden afgeleid en voor elk van deze modelinstanties onderzoeken we de relatie tussen de epidemische drempelwaarde en de effective verbrekingssterkte. We leiden de eerste-orde en tweede-orde gemiddelde-veldbenadering af van het G-ASIS model (Hoofdstuk 5) en illustreren dat de tweede-orde benadering het oorspronkelijke Markoviaanse model veel nauwkeuriger benadert. In Hoofdstuk 6 breiden we de gemiddelde-veld benadering uit naar algemene verbrekings- en herstellingsregels, die niet noodzakelijkerwijs afhangen van de gezondheidstoestand van de twee verbonden personen, maar ook kunnen afhangen van de aanwezigheid van de ziekte in de hele populatie. We laten zien dat onafhankelijk van de keuze van de verbrekings- en herstellingsregels, periodieke golven van infecties onmogelijk zijn in het gemiddelde-veld adaptieve SIS proces.

In het laatste gedeelte leiden we de Network-Inference-based Prediction Algorithm (NIPA) af voor het voorspellen van ziekteverspreiding in heterogene netwerken. De contactgraaf is in het algemeen onbekend en wordt daarom door NIPA geschat op basis van gerapporteerde ziektegevallen. NIPA is a hybride methode, die ziekteverspreiding, machine-learning en netwerktheorie combineert. Voorspellingen op basis van netwerk-

theorie, en NIPA in het bijzonder, lijken bijzonder nauwkeurig te zijn in het voorspellen van ziekteverspreiding, wat we demonstreren door te laten zien dat NIPA de verspreiding van COVID-19 nauwkeuriger kan voorspellen dan verschillende andere voorspellingsalgoritmes.

1

INTRODUCTION

1.1. CONTAGIOUS DISEASES

Contagious diseases have played a central role in shaping the world's history [1], since several devastating pandemics wiped out entire communities [2]. Thus, assessing and managing the spread of contagious diseases is an important, but difficult task. The severity of a disease is determined by its *transmissibility* and its *virulence*. Rather than focussing on the viral properties of the disease, we focus on the epidemiological, larger picture; *how many* people will be infected *on which time* and *in which areas*?

At least as important as the characteristics of the virus is the people's response to the disease – what countermeasures are being taken to slow down the virus spread and are those measures effective? Without the access to vaccinations or medicines, the only treatment of contagious diseases is to prevent getting infected, e.g. by staying outside of your infectious neighbours' proximity. In popular terms, people apply *social distancing* to reduce the risk of transmitting the disease.

Social distancing is one example of so-called non-pharmaceutical interventions (abbreviated as NPIs). NPIs come in various flavours. On a personal level, individuals can prevent themselves from getting infected by isolating themselves at home. Similarly, infected individuals can isolate to prevent *others* from getting infected. Those key epidemiological questions received a lot of attention from the scientific community during the COVID-19 pandemic. Many COVID-19 measures including no hand-shaking, washing hands, wearing face masks and curfews are nation-wide NPIs that essentially have the same goal: reducing the spread between infected and non-infected individuals.

1.2. BACKGROUND ON NETWORK EPIDEMIOLOGY

Epidemics are omnipresent in both human and animal populations. Epidemics often occur on a large scale, complicating experiments on the effectiveness of countermeasures. In particular, epidemic outbreaks cannot be modelled using *in-vitro*[1] experi-

[1] The Latin phrase *in-vitro* literally means *in glass* and is used to describe experiments that are conducted outside of their usual habitat, e.g. in a laboratory. As epidemics involve humans and the impact of epidemics

ments, because of its potentially devastating consequences. Instead, scientists rely on reports of the number of infected patients from health agencies and doctors. Using the number of reported cases, forecasts on the number of cases can be made under various scenarios using mathematical models.

The first scientist to provide a mathematical description of an epidemic outbreak was Daniel Bernoulli in 1766 [3]. Bernoulli analysed an outbreak of smallpox in London and developed a simple model to predict the number of future cases. A considerable advancement was made by Kermack and McKendrick in 1927. Kermack and McKendrick split up the population into different groups, called compartments, based on the viral state of the individuals [4]. Individuals can make a transition from one compartment to another compartment during a certain event, i.e. infection upon contact, indirect infection through air or via animals, recovery, death, etc.

Typical examples of compartments are the following: `Susceptible` (healthy, but may contract the disease upon contact with an infectious individual), `Exposed` (infected, not infectious yet but will become infectious in the near future), `Asymptomatic` (infected and infectious but shows no symptoms), `Infected` (infected and infectious and shows symptoms), `Quarantined` (infected cases who are locked away at home or in the hospital), `Recovered` (either temporary or permanently immune) and `Deceased` (passed away due to the disease).

We focus in this thesis on the two most elementary compartmental models: the Susceptible–Infected–Susceptible (SIS) and Susceptible–Infected–Recovered (SIR) models. In both models, susceptible individuals can be infected by their infectious neighbours and if successful, the susceptible person immediately becomes infected as well. Additionally, infected individuals can recover from the disease, either becoming susceptible again (SIS) or receiving permanent immunity (SIR).

Modelling the spread of epidemics can be performed on different aggregation levels. One may investigate the spread of a disease among single individuals, or instead look at groups of people, comprising of households, cities or countries. For both the individual and the group-based methods, we use the notation of a network to describe the spread of the disease in the population. The contact graph $G = (\mathcal{N}, \mathcal{L})$ contains the set $\mathcal{N}$ of N nodes, where each node represents one individual or one group. The set $\mathcal{L}$ of L links describes the connections between the nodes. The graph G is assumed to be simple[2], such that the graph G can be represented by an adjacency matrix A, whose elements a_{ij} indicate the existence ($a_{ij} = 1$) or non-existence ($a_{ij} = 0$) of a link between nodes i and j. We assume that the population remains constant over time, that is, there are no births and deaths in the population and there are no individuals emigrating from or immigrating to our system. Together with a description of the spread of an epidemic, e.g. the SIS or SIR process, the graph (structure) and infection process (service) describe an epidemiological complex network. The review by Pastor-Satorras *et al.* [6] covers many recent advancements in network epidemiology.

Throughout this thesis, we describe the spread of the disease by two processes: the infection process and the curing process. An infected node j may infect a susceptible node i with rate β_{ij} if node i and j are connected ($a_{ij} = 1$). Independently of the in-

is huge, such kind of experiments are simply impossible.

[2] A graph G is simple if it does not contain any self-loops or multiple links between two nodes [5].

fection process, infected node i can recover from the disease with rate δ_i. Any physical phenomenon spreads in continuous time, thus we assume in this thesis that the disease also evolves in continuous time (unless otherwise specified). The time distribution for both the infection and curing process can be any probability distribution, but in this thesis, we restrict ourselves to exponentially distributed transition times. Then the disease spread can be described as a continuous-time Markov process, which has many favourable properties.

1.3. ADAPTIVE NETWORKS

Most epidemic models assume that the contact graph G remains constant over time. In the presence of pharmaceutical solutions to eradicate a disease, of which vaccinations and medicines are prominent examples, we expect that the network barely changes over time. Unfortunately, in case of world-wide epidemics like COVID-19, pharmaceutical solutions are not immediately available and social distancing is the only available tool for disease prevention. Social distancing may happen based on individual decisions or governmental decisions. For example, if a person knows that a friend or family member is infected by the disease, the person will avoid contact with that particular friend or family member. In case of severe pandemics, like COVID-19, nation-wide lockdowns imposed by the government significantly reduce the number of links in the contact network. In either case, the contact network changes over time and the reason for changing the contact graph is based on the presence of the virus in the population. We call these networks *adaptive* or *coevolutionary* networks, because the structure of the contact network changes based on the spread of the contagious process. The overview in Figure 1.1 shows the types of commonly used networks in the literature. In this thesis, we primarily focus on static and adaptive networks.

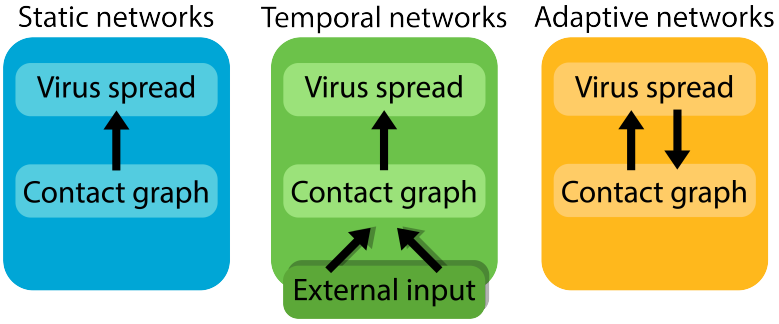


Figure 1.1: The different types of networks. Arrows represent the relation: "has influence on". Static networks have a fixed contact graph, which influences the spread of the disease. The contact graph of temporal networks changes over time due to external processes, like the movement of people. Adaptive networks allow individuals to break or create links with other individuals to prevent contracting the disease. Thus, the contact graph changes based on the spread of the virus in the population.

1.4. BEYOND EPIDEMICS

Throughout this thesis, we use terminology and notation from epidemiology to introduce and explain various concepts. However, the domain of spreading processes on networks is much broader than network epidemiology alone and also encompasses the spread of gossips, political preference, opinions [7], computer viruses [8], information in the human brain [9], raising awareness about a particular event or fact [10], fake news spreading [11], innovation spread [12], failure propagation [13] and internet packet routing [14]. Even though we primarily focus in this thesis on epidemiology, we emphasise that many obtained results for epidemic spreading can be generalised for other spreading phenomena that are also based on compartmental models.

1.5. MY CONTRIBUTIONS

Prior to discussing my contributions to the scientific literature of network epidemiology, let me start by mentioning that the field of network epidemiology is very well-researched and providing a comprehensive overview of network epidemiology is impossible due to the plethora of works from biologists, virologists, epidemiologists, data scientists, physicists, mathematicians and others. For example, a Google Scholar search for the keywords "network" and "epidemics" in March 2023 yields over 300,000 scientific papers. Instead of providing a generic literature overview, each chapter in this thesis will feature its own literature overview.

My contribution starts in Part I with elementary results on the well-known SI, SIS and SIR epidemic models. Chapter 2 discusses the Markovian SIS process with self-infections on the complete graph. The main focus is on the understanding of the eigenvalues (spectrum) of the infinitesimal generator of the underlying Markov chain. We show that key epidemiological behaviour of the SIS process, such as the epidemic threshold, is encoded in the eigenvalues. We proceed by investigating SI and SIR epidemics on heterogeneous graphs in Chapter 3. For the first time, we present an exact solution for epidemic processes without reinfections on heterogeneous networks. We demonstrate our exact solution by analytically solving the SI and SIR processes on heterogeneous networks.

Part II focusses on an important open problem in network epidemiology: what is the effect of personal decisions – to break or restore connections with other individuals to prevent them or yourself from contracting the disease – on the spread of the disease? Chapter 4 introduces the Generalised Adaptive SIS (G-ASIS) model to describe the spread of a disease in an adaptive network. We determine all possible rules to break or create connections in the graph. We analytically derive the average fraction of infected nodes in the metastable state and also derive bounds on the epidemic threshold. Chapter 5 builds on Chapter 4 and investigates the mean-field approximation of the G-ASIS model. Two different types of mean-field models are introduced and compared. We find that the second-order mean-field approximation is a much better approximation for the Markovian G-ASIS process than the first-order mean-field approximation. Chapter 6 generalises the mean-field model from Chapter 5 for arbitrary link-breaking and link-creation rules, which are not just limited to interactions between neighbours, but may depend on the presence of the virus in the whole population. For all link-adaption rules, we derive the epidemic threshold and show that limit cycles cannot exist.

In Part III, we shift our focus to a practical setting; how to forecast the spread of the COVID-19 pandemic? Chapter 7 introduces the Network-Inference-based Prediction Algorithm (NIPA), which is designed to forecast the future number of infected cases. NIPA is comprised of three steps. As a first step, the daily number of reported cases is preprocessed into a time series for an SIR model. By regarding the spread of COVID-19 as an epidemic spreading process on a network, the task of NIPA is to determine the infection link weights. NIPA assumes that the link weights are unknown, but are instead inferred from observations of the epidemic outbreak. Finally, using the inferred disease parameters, a forecast is produced based on an SIR model. We benchmark NIPA against other well-known epidemic models and introduce several variants of NIPA, e.g. including partial prior knowledge on the contact network. We show that network-based forecasting (and NIPA in particular) seems to be beneficial for predicting the spread of COVID-19.

We finalise this thesis with a conclusion, that summarises the majority of the obtained results. We further provide suggestions for future investigations in the field of network epidemiology.

I

EPIDEMICS ON STATIC NETWORKS

2

EIGENVALUE ANALYSIS OF ε-SIS DYNAMICS

We analyse continuous-time Markovian ε-SIS epidemics with self-infections on the complete graph. We focus on the complete graph because the majority of the graphs is analytically intractable, but many features of the ε-SIS process observed in the complete graph occur in other graphs. In this chapter, we illustrate that the time scales of the ε-SIS process are related to the eigenvalues of the tridiagonal infinitesimal generator of the SIS Markov chain. We provide a detailed analysis of all eigenvalues and illustrate that the eigenvalues show staircases, which are caused by the nearly degenerate (but strictly distinct) pairs of eigenvalues. We also illustrate that the ratio between the second-largest and third-largest eigenvalue is a good indicator of metastability in the ε-SIS process. Additionally, we show that the epidemic threshold of the Markovian ε-SIS process can be accurately approximated by the effective infection rate for which the third-largest eigenvalue of the transition matrix is the smallest. Finally, we derive the exact mean-field solution for the ε-SIS process on the complete graph with arbitrary initial conditions and show that the mean-field approximation does not correctly represent the metastable behaviour of Markovian ε-SIS epidemics.

This chapter is based on M. A. Achterberg, B. Prasse and P. Van Mieghem, *Analysis of continuous-time Markovian ε-SIS epidemics on networks*, Physical Review E 105, 054305, May 2022 [15].

2.1. INTRODUCTION

Over the last hundred years, many epidemic diseases have plagued humanity. Most epidemic outbreaks tend to emerge quickly, but take much longer to disappear [2]. One of the main reasons is re-infections. A common example of a disease with re-infections is influenza, which affects large portions of the population during the winter season. The influenza virus strain keeps mutating slightly, thereby bypassing the human immune system while maintaining most of its viral properties. Another example of recurring diseases are sexually transmitted diseases, such as chlamydia and gonorrhoea. Contrary to influenza, sexually transmitted diseases do not mutate, but people can simply be re-infected after recovering from the disease.

Initially, epidemic outbreaks spread exponentially fast, because most individuals in the population are susceptible to the new disease, as happened with the COVID-19 pandemic. In a closed and well-mixed population, the number of infected individuals stabilizes after a short time and continues to oscillate around the *prevalence*, which is defined as the average number of infected individuals. Then, the epidemic process is in the *metastable* or *quasi-stationary state*, because the number of infected individuals remains in the vicinity of the prevalence for a long period of time, whereafter the process eventually converges to its steady state. The steady state of most epidemic processes is the overall-healthy state, which corresponds to the situation where the disease has disappeared completely from the population.

We focus in this chapter on one of the simplest epidemic models, namely the SIS process. Besides the usual infection and curing processes, we consider a third, independent self-infection process with self-infection rate ε, which describes background or indirect infections. Infections may happen either through direct contact or indirectly, for example, after touching infected surfaces or inhaling air in a closed room previously contaminated by an infected individual. Thus, the Markovian ε-SIS model consists of three, independent Poisson processes: (i) the curing process with rate δ, (ii) infection process with rate β and (iii) self-infection process with rate ε. Given that the ε-SIS model consists of independent Poisson processes, we can describe the time-dependent behaviour of the ε-SIS model as a continuous-time Markov chain with 2^N states [16]. Continuous-time Markovian modeling implies that the infection and curing times are exponentially distributed. However, measurements in real epidemics seem to suggest that the infection time follows a bell-shaped distribution (such as Gamma, Weibull or lognormal distribution [17, 18]) which requires non-Markovian analysis, that is, unfortunately, considerably more complex (see e.g. [19]) and is currently insufficiently developed to compute time-dependent infection probabilities.

The continuous-time Markovian SIS process without self-infections on static networks has been investigated thoroughly [6]. Even on the complete graph, quantifying the average time spent in the metastable state appears challenging [20]. The major issue is that the metastable state is not stable, but collapses eventually, by a rare occurrence of successive curings, to the absorbing or overall-healthy state [21]. Several approaches have been proposed to quantify the metastable state in the Markovian SIS model. Jacquez and Simon [22] introduced an epidemic process, that prevents the original SIS process from entering the absorbing state at time t. Their reduced SIS model has a unique steady state, which can be related to the metastable state of the original SIS

model. Cator and Van Mieghem [23] constructed a similar modified process: If only a single node is infected, then the modified process forbids the curing of that node. Effectively, their modified SIS process is equivalent to the SIS process, with the exception that the transitions to the absorbing state have been removed. De Oliveira and Dickman [24] proposed to store the complete time lapse of the SIS model. Once the process converges to the absorbing state, the process jumps to a randomly selected sample from the history of the SIS process. Keeling and Ross [25] and Hill *et al.* [26] introduced a self-infection process ε on the complete graph. Besides the usual infection and curing processes, the nodes in the modified process can be infected by external sources, which are modelled as self-infections. Van Mieghem and Cator [16] generalised the ε-SIS model from the complete graph to general networks. Introducing a small amount of self-infections removes the absorbing state, establishes an irreducible Markov chain with a well-defined steady state different from the overall-healthy state and allows for the comparison between the ε-SIS model and the SIS model without self-infections. Finally, the 2^N-state Markov chain, described by 2^N linear differential equations, is often approximated by a mean-field approximation with N non-linear differential equations. The simplest mean-field approximation for the SIS model on networks is called the N-Intertwined Mean-Field Approximation (NIMFA) and assumes that the infection state of any two nodes is uncorrelated [27, 28, 29, 30]. A mean-field approximation is generally an adequate approximation for large, dense graphs with homogeneous transition rates and for infection rates sufficiently larger than the epidemic threshold. The epidemic threshold, denoted as τ_c, is defined as the largest infection rate for which the prevalence decays exponentially fast to zero [31]. The accuracy of NIMFA with respect to the Markovian SIS process is studied in [32] for various graph types. A key difference is the possibility of die-outs in the stochastic model, albeit with a very small probability, whereas the mean-field model either converges to the endemic equilibrium or to the all-healthy state, and excludes the possibility of sudden die-outs.

The continuous-time Markovian ε-SIS process on the complete graph can be described by a birth and death process (BDP). Birth and death processes can be solved by computing the probability generating function and solving the corresponding partial differential equation [33, 34]. Unfortunately, solving the partial differential equation seems infeasible for ε-SIS dynamics [35, Appendix A]. On the other hand, one may compute the Rayleigh-Ritz coefficient of the partial differential equations to derive bounds for the eigenvalues. There are also several approaches to compute the eigenvalues of BDPs exactly. A possibility is to consider the orthogonal polynomials that correspond to the BDPs, which are the Tricomi-Carlitz polynomials. The zeros of the Tricomi-Carlitz polynomials correspond to the eigenvalues of the BDP. Unfortunately, not many results are known for the zeros of the Tricomi-Carlitz polynomials [36]. Alternatively, one may solve for the eigenvalues directly, resulting in a continued fraction expansion [5], or one can also derive bounds on the eigenvalues, e.g. by the Cauchy interlacing theorem or using a Fokker-Planck approximation [37].

The eigenvalues of Markov chains and their relation to metastability have been studied in several works. Artajelo [38] studied the second-largest eigenvalue in general Markov chains in both continuous time and discrete time. Holme and Tupikina [39] computed the exact second-largest eigenvalue in SIS epidemics for all non-isomorphic graphs with

$3 \leq N \leq 8$ nodes. For an arbitrary graph size N, exact results of the ε-SIS model can be obtained only for a few graphs, such as the complete graph and the star graph [23]. For the complete graph with homogeneous transition rates and no self-infections, several analytical results have been obtained, such as the average time before extinction [37, 40, 31, 41] and the average time between the onset of the disease and the arrival at the metastable state [42, 43].

In this chapter, we study continuous-time Markovian ε-SIS epidemics on the complete graph from an eigenvalue perspective by computing *all* its eigenvalues. Although we realise that the complete graph is far from a realistic setting, we derive many qualitative properties of the ε-SIS process which, we believe, may also hold for other graphs. We describe the continuous-time Markov chain for the complete graph in Section 2.2. We additionally derive the general solution of the Markov chain. We introduce the concept of metastability in Section 2.3 and derive the exact mean-field solution in Section 2.4. We return to the Markovian ε-SIS process in Section 2.5 and provide a detailed analysis of all eigenvalues. We numerically identify the influence of the infection rate β, self-infection rate ε and network size N on the eigenvalues in Section 2.6. Finally, we present conclusions in Section 2.7.

2.2. THE ε-SIS PROCESS ON THE COMPLETE GRAPH

The Markov chain $\mathcal{M}$ of the ε-SIS process on the complete graph describes the number of infected individuals M. Since the population consists of N individuals, the number M of infected nodes ranges from zero to N. Thus, the Markov chain $\mathcal{M}$ has $N+1$ states with the transition rates [44, p. 474]

$$\begin{aligned} M &\mapsto M+1, \quad \text{at rate} \quad (\beta M+\varepsilon)(N-M) \\ M &\mapsto M-1, \quad \text{at rate} \quad \delta M \end{aligned}$$

where β denotes the infection rate, δ the curing rate and ε the self-infection rate in the complete graph K_N. The Markov chain $\mathcal{M}$ is drawn in Figure 2.1. The Markov chain is a birth and death process with birth rate $\Xi_k = (\beta k+\varepsilon)(N-k)$ that is quadratic in k and death rate $\mu_k = \delta k$ that is linear in k [1].

We compute the probability that k nodes are infected at time t in the Markov chain $\mathcal{M}$. Let $M(t)$ be the number of infected nodes at time t. By introducing $s_k(t) = \Pr[M(t) = k]$ as the probability that the number $M(t)$ of infected nodes at time t equals k, the following differential equations describe the exact dynamics of the Markov chain $\mathcal{M}$:

$$\begin{aligned} \frac{ds_0}{dt} &= \mu_1 s_1(t) - \Xi_0 s_0(t), \\ \frac{ds_k}{dt} &= -(\Xi_k+\mu_k)s_k(t) + \Xi_{k-1}s_{k-1}(t) + \mu_{k+1}s_{k+1}(t), \qquad k = 1,\ldots,N-1 \\ \frac{ds_N}{dt} &= \Xi_{N-1}s_{N-1}(t) - \mu_N s_N(t). \end{aligned}$$

[1] Contrary to the usual notation of the birth rate λ in birth and death processes and queueing theory, we denote the birth rate as Ξ to avoid confusion with the eigenvalue λ.

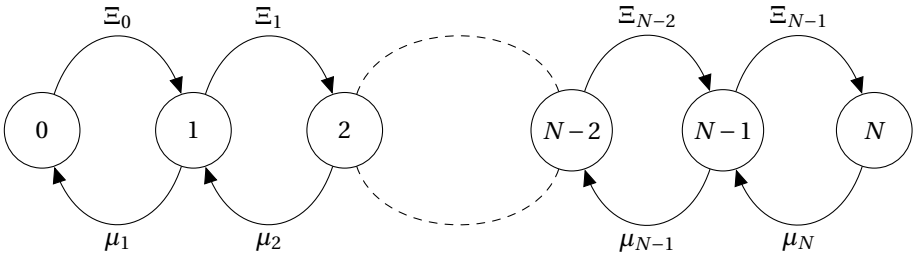


Figure 2.1: The Markov chain $\mathcal{M}$ of the ε-SIS process on the complete graph K_N with N nodes is a finite birth and death process with birth rate $\Xi_k = (\beta k + \varepsilon)(N-k)$ and death rate $\mu_k = \delta k$. The state $M = 0, ..., N$ of the Markov chain denotes the number of infected nodes in the graph. Furthermore, β is the infection rate, δ the curing rate and ε the self-infection rate.

If the curing rate $\delta > 0$, the ε-SIS process can be simplified by introducing the rescaled time $\tilde{t} = \delta t$. We additionally define the scaled birth rate $\tilde{\Xi}_k = \Xi_k/\delta = (\tau k + \varepsilon^*)(N-k)$ and the scaled death rate $\tilde{\mu}_k = \mu_k/\delta = k$, where the effective infection rate $\tau = \beta/\delta$ and effective self-infection rate $\varepsilon^* = \varepsilon/\delta$. We emphasise that the scaled time $\tilde{t}$, the effective infection rate τ and the effective self-infection rate ε^* are dimensionless variables. Introducing the $(N+1) \times 1$ vector $\mathbf{s} = (s_0, ..., s_N)^T$, the linear differential equations can be written in matrix notation:

$$\frac{d\,\mathbf{s}^T}{d\tilde{t}} = \mathbf{s}^T P, \tag{2.1}$$

where P is the $(N+1) \times (N+1)$ tridiagonal *transition matrix*,

$$P = \begin{pmatrix} -\tilde{\Xi}_0 & \tilde{\Xi}_0 & & & & \\ \tilde{\mu}_1 & -(\tilde{\Xi}_1 + \tilde{\mu}_1) & \tilde{\Xi}_1 & & & \\ & \tilde{\mu}_2 & -(\tilde{\Xi}_2 + \tilde{\mu}_2) & \ddots & & \\ & & & \ddots & & \\ & & & \ddots & \tilde{\Xi}_{N-2} & \\ & & & & \tilde{\mu}_{N-1} & -(\tilde{\Xi}_{N-1} + \tilde{\mu}_{N-1}) & \tilde{\Xi}_{N-1} \\ & & & & & \tilde{\mu}_N & -\tilde{\mu}_N \end{pmatrix}. \tag{2.2}$$

Together with the initial condition $\mathbf{s}_0 = \mathbf{s}(0)$, equation (2.1) describes the exact dynamics of the continuous-time Markovian ε-SIS process on the complete graph K_N, which can be solved to find

$$\mathbf{s}^T(\tilde{t}) = \mathbf{s}_0^T e^{P\tilde{t}}. \tag{2.3}$$

Using the probability vector of the number of infected nodes $\mathbf{s}(\tilde{t})$ at time $\tilde{t}$, one may compute the average fraction of infected cases $y(\tilde{t})$, commonly known as the *prevalence*,

as

$$y(\tilde{t}) = \frac{1}{N}\sum_{k=0}^{N} k \cdot s_k(\tilde{t}). \tag{2.4}$$

In the remainder of this chapter, we omit the tilde for the scaled time $\tilde{t}$ for readability.

We intend to show that metastability in the ε-SIS process is directly linked to the eigenvalues of the transition matrix P. We denote the eigenvalues $\lambda_1, \ldots, \lambda_{N+1}$, the right-eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_{N+1}$ and the left-eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_{N+1}$ of the $(N+1) \times (N+1)$ transition matrix P. The eigenvalues of the transition matrix P are real, because P is similar to a symmetric matrix $\tilde{P}$ and similarity preserves the eigenvalues. The transformed matrix $\tilde{P}$ is computed in Appendix A.1. The eigenvalues of P and $\tilde{P}$ cannot be computed analytically for $N > 4$ because it involves finding the roots of a characteristic polynomial with degree N. Thus one resorts to numerical methods to obtain the eigenvalues. Once the eigenvalues are known, the corresponding eigenvectors can be computed analytically [44, Appendix A.6.3].

Since all eigenvalues are real-valued, we may rank them in decreasing order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{N+1}$. Given that the tridiagonal matrix $\tilde{P}$ is symmetric and all off-diagonal terms are non-zero[2], all eigenvalues of P are distinct [45, Lemma 7.7.1]. The same conclusion follows by computing the probability generating function of the ε-SIS process and concluding that the resulting differential equation is of Sturm-Liouville type [35], which is known to have simple eigenvalues. The transition matrix P of the ε-SIS Markov chain has an unique, largest eigenvalue $\lambda_1 = 0$, which corresponds to the steady state π. The remaining eigenvalues $\lambda_2, \ldots, \lambda_{N+1}$ are negative and distinct. The solution (2.3) can be written as

$$\mathbf{s}(t) = \pi + \sum_{k=2}^{N+1} c_k e^{\lambda_k t} \mathbf{w}_k, \tag{2.5}$$

where λ_k is the eigenvalue corresponding to the right-eigenvector $\mathbf{w}_k$ of the ε-SIS process and the constant $c_k = \mathbf{v}_k^T \mathbf{s}(0)$ projects the initial vector $\mathbf{s}(0)$ on the k-th left-eigenvector $\mathbf{v}_k$ [25, equation (2.5)]. The vector $\mathbf{s}(t)$, whose components $s_k(t)$ specify the probability that k nodes at time t are infected, is decomposed in Eq. (2.5) into $N+1$ eigenstates of which the corresponding eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_{N+1}$ span the $(N+1)$-dimensional vector space. Each eigenvector $\mathbf{w}_k$ in Eq. (2.5) is weighted by the coefficient $c_k e^{\lambda_k t}$. The eigenvalue λ_k resembles a rate and has unit 1/time. The contribution of eigenvector $\mathbf{w}_k$ to the solution $\mathbf{s}(t)$ decays exponentially over time with decay rate equal to the eigenvalue λ_k (the contribution decays because $\lambda_k < 0$). The eigenvalue λ_k is thus inversely proportional to the average time that the corresponding eigenvector $\mathbf{w}_k$ significantly contributes to the solution $\mathbf{s}(t)$.

In particular, the second-largest eigenvalue λ_2 (sometimes called the *convergence rate*, spectral gap, mixing rate or decay parameter) is inversely proportional to the average time required to converge towards the steady state [38, 46]. The convergence rate for continuous-time Markov chains and BDPs is thoroughly analysed in probability theory. For an overview of bounds of the convergence rate in Markov chains and BDPs, we refer to Van Doorn *et al.* [47] and Artalejo [38] and references therein. If the effective

[2] Here, we assume that the curing rate δ, the effective infection rate τ and the effective self-infection rate ε^* are non-zero.

self-infection rate $\varepsilon^* = 0$, then the average time of convergence $\mathrm{E}[T_{\text{extinction}}] = -1/\lambda_2$ to the steady state (or equivalently, the *extinction time*, survival time or absorption time) on the complete graph has the following exact relationship [40]:

$$\mathrm{E}[T_{\text{extinction}}] = \sum_{i=1}^{N} \sum_{j=0}^{i-1} \frac{(N-i+j)!}{i(N-i)!} \tau^j.$$

Nearly all works consider the SIS process without self-infections. Most proofs for the convergence rate (e.g. the proof in [40]) are based on the hitting time distribution of the absorbing state. By introducing the self-infection process, the absorbing state no longer exists and the proofs therefore do not hold for the ε-SIS process. Fortunately, the introduction of the self-infection process makes the Markov chain of the ε-SIS process irreducible, implying that the steady state exists and is also unique. The existence of the steady state greatly simplifies the analysis of the ε-SIS process for large times, because we can analyse the behaviour of the conceptually simpler steady state instead of the more complicated metastable state.

2.3. METASTABILITY IN THE ε-SIS PROCESS

Our primary motivation for researching the eigenvalues of the transition matrix P in equation (2.2) is the observation of plateau-behaviour in the ε-SIS process [35], which is illustrated in Figure 2.2. Each curve in Figure 2.2 is computed based on the exact solution (2.5), where initially a single node is infected. For small effective self-infection rates $\varepsilon^* \leq 10^{-5}$, Figure 2.2 depicts roughly three regimes for the time-varying prevalence: (I) initial phase, (II) metastable behaviour, and (III) convergence to the steady state. Phase (I) is characterised by the fast convergence to the metastable state. In the metastable phase (II), the prevalence y stays nearly constant for an extended period of time. Finally, phase (III) shows the exponential convergence to the steady state π.

Plateau-behaviour is generally caused by metastability of the dynamical process, where the infection and curing processes are temporarily in equilibrium (the physical explanation is presented in [21]). We consider the following definition of metastability for general dynamical processes that possess a steady state π.

Definition 2.1 *A dynamical process is **metastable** if the process stays in a certain state for an extended period of time before converging to the steady state π.*

We quantify metastability for the ε-SIS process using the eigenvalue ratio, which was first introduced by Jacquez and Simon [22, p. 85].

Definition 2.2 *The **eigenvalue ratio** ρ is defined as*

$$\rho = \frac{\lambda_3}{\lambda_2}. \tag{2.6}$$

The eigenvalue ratio ρ is an indicator for the existence of plateaus [22]. If the eigenvalue ratio ρ is large, then the second-largest eigenvalue λ_2 is much larger than the other eigenvalues $\lambda_3, \ldots, \lambda_{N+1}$. The influence of the right-eigenvectors $\mathbf{w}_2, \ldots, \mathbf{w}_{N+1}$ on the solution $\mathbf{s}(t)$ in (2.5), weighted by the exponentials $e^{\lambda_k t}$, will therefore converge much faster to zero for the eigenvalues $\lambda_3, \ldots, \lambda_{N+1}$ than for the second-largest eigenvalue λ_2.

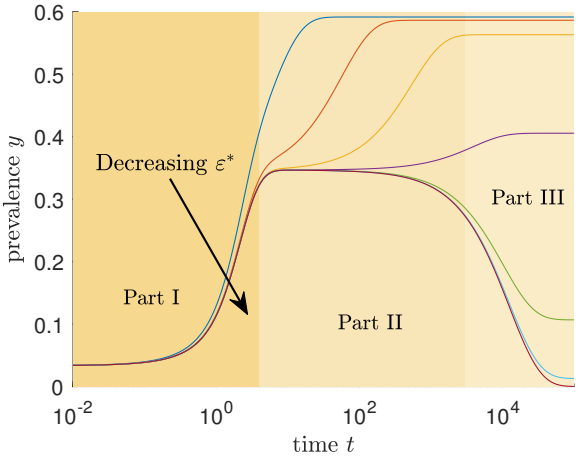


Figure 2.2: Plateau-behaviour in the Markovian ε-SIS process on the complete graph. Each curve is computed using the exact solution (2.5) and the parameters are $N = 30$ nodes, effective infection rate $\tau = 2.5\tau_c^{(1)} = 2.5/(N-1)$, the process starts with a single infected node and the effective self-infection rate $\varepsilon^* = \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, 0\}$. For $\varepsilon^* \leq 10^{-5}$, the background colours indicate the current phase of the ε-SIS process: (I) initial phase, (II) metastable behaviour and (III) convergence to the steady state.

The height of the plateaus in Figure 2.2 equals the prevalence y, which is defined in Eq. (2.4). Plateau-behaviour, as shown in Figure 2.2, is only clearly observed if the steady-state prevalence y_∞ is sufficiently different from the prevalence in the metastable state and if the effective self-infection rate ε^* is sufficiently small. Our definition of metastability in the ε-SIS process is:

Definition 2.3 *The ε-SIS process is **metastable** if the eigenvalue ratio $\rho \gg 1$ and the prevalence y_∞ in the steady state is sufficiently different from the prevalence in the metastable state.*

We analyse the case when the metastable prevalence y and the steady-state prevalence y_∞ are equal in Appendix A.2. We find the powerlaw relation $\varepsilon^* \sim \tau^{-N}$ if τ is sufficiently larger than the epidemic threshold τ_c. For further details, we refer to Appendix A.2.

If the ε-SIS process is metastable, the average time spent in the metastable state is roughly equivalent to the average time required to converge to the steady state, because the average time between the onset of the disease and the arrival at the metastable state is relatively short (see Figure 2.2 for an example).

2.4. MEAN-FIELD APPROXIMATION OF THE ε-SIS PROCESS

The majority of the research in network epidemiology is based on mean-field approximations of stochastic processes. However, the behaviour of the Markovian ε-SIS process is intrinsically different from its mean-field approximation, especially regarding the metastable state. Recently, Prasse *et al.* [48] solved the continuous-time mean-field SIS

process on the complete graph K_N with arbitrary initial conditions. We derive a similar result in Theorem 2.4 for the ε-SIS process with self-loops, whereby we added the infection rates τ_{ii} from each node i to itself. Adding self-loops only further upper-bounds the mean-field approximation from the Markovian solution and have not been added based on any physical intuition, but instead simplify the exact computation of the mean-field prevalence in Theorem 2.4. We refer to Appendix A.3 for the derivation of the mean-field equations, which approximate the exact, Markovian solution $\mathbf{s}(t)$ by the mean-field state vector $\mathbf{s}_{\mathrm{MF}}(t)$.

Theorem 2.4 *Consider the mean-field approximation of the ε-SIS process on the complete graph with homogeneous transition rates and self-loops, given by Eq. (A.6) in Appendix A.3, with **arbitrary initial conditions**. Then the mean-field state vector $\mathbf{s}_{\mathrm{MF}}(t)$ is equal to*

$$\mathbf{s}_{\mathrm{MF}}(t) = c_1(t)\mathbf{z}_1 + c_2(t)\mathbf{z}_2 \tag{2.7}$$

at every time t. Here, $\mathbf{z}_1$ and $\mathbf{z}_2$ are orthonormal agitation modes, which are given by

$$\mathbf{z}_1 = \frac{1}{\sqrt{N}}\mathbf{u}, \tag{2.8}$$

where $\mathbf{u} = (1,...,1)^T$ denotes the $N \times 1$ all-one vector, and

$$\mathbf{z}_2 = \left\| \left(I - \mathbf{z}_1\mathbf{z}_1^T\right)\mathbf{s}_{\mathrm{MF}}(0) \right\|_2^{-1} \left(I - \mathbf{z}_1\mathbf{z}_1^T\right)\mathbf{s}_{\mathrm{MF}}(0). \tag{2.9}$$

The functions $c_l(t) = \mathbf{z}_l^T\mathbf{s}_{\mathrm{MF}}(t) \in \mathbb{R}$, where $l = 1,2$, are the projection of the viral state vector $\mathbf{s}_{\mathrm{MF}}(t)$ on the agitation mode $\mathbf{z}_l$. The scalar function $c_1(t)$ equals

$$c_1(t) = \frac{1}{2\tau\sqrt{N}}\left(\tau N - 1 - \varepsilon^* + w_{\varepsilon^*}\tanh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)\right)$$

with the viral slope

$$w_{\varepsilon^*} = \sqrt{(1+\varepsilon^*-\tau N)^2 + 4\varepsilon^*\tau N}$$

and the constant

$$\Upsilon_{1,\varepsilon^*}(c_1(0)) = \operatorname{arctanh}\left(\frac{1}{w_{\varepsilon^*}}\left(2\tau\sqrt{N}\mathbf{z}_1^T\mathbf{s}_{\mathrm{MF}}(0) - \tau N + 1 + \varepsilon^*\right)\right).$$

The scalar function $c_2(t)$ equals

$$c_2(t) = \Upsilon_{2,\varepsilon^*}(c_2(0))\exp\left(-\frac{1+\varepsilon^*+\tau N}{2}t\right)\operatorname{sech}\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right),$$

with the constant

$$\Upsilon_{2,\varepsilon^*}(c_2(0)) = \mathbf{z}_2^T\mathbf{s}_{\mathrm{MF}}(0)\cosh\left(\Upsilon_{1,\varepsilon^*}(c_1(0))\right). \tag{2.10}$$

Proof. See Appendix A.3. □

We emphasise that Theorem 2.4 holds for arbitrary initial conditions. If the initial conditions are symmetric for all nodes, the solution $\mathbf{s}_{\text{MF}}$ equals the prevalence $y_{\text{MF}}(t)$ and simplifies to

$$y_{\text{MF}}(t) = \frac{\tau N - \varepsilon^* - 1}{2\tau N} + \frac{w_{\varepsilon^*}}{2\tau N} \tanh\left(\frac{w_{\varepsilon^*}}{2} t + \tanh^{-1}\left(\frac{\varepsilon^* + 1 - \tau N + 2\tau N y_0}{w_{\varepsilon^*}}\right)\right).$$

As derived in Appendix A.3, the mean-field prevalence upperbounds the Markovian prevalence $y(t)$. Theorem 2.4 states that the mean-field dynamics on the complete graph can be reduced from N equations to two agitation modes, where one is related to the initial condition and the other to the steady state. Hence, the metastable state (Phase (II) in Figure 2.2) does not exist under the mean-field approximation, because the existence of only two agitation modes does not allow for an intermediate, transient regime.

Figure 2.3 depicts the Markovian prevalence, mean-field prevalence and the exact mean-field prevalence with self-loops based on Theorem 2.4. Figure 2.3 illustrates that the mean-field approximation vastly overestimates the time-dependent fraction of infected nodes of the Markovian ε-SIS process, both with and without self-loops. If the effective infection rate τ is larger than the epidemic threshold τ_c, Figure 2.3a shows that, for $N = 30$, the discrepancy between the mean-field and Markovian prevalence is large everywhere, except at very small time scales or if ε^* is very large. Figure 2.3b is situated around the epidemic threshold τ_c, where the mean-field approximation is known to have the worst accuracy [32]. In the limit $N \to \infty$, the mean-field approximation error converges to zero, which we further illustrate in Section 2.6.3. Given that the metastable state is a key epidemiological quantity and that the metastable state does not exist under the mean-field approximation, we focus in the remainder of this chapter on the *Markovian* ε-SIS process.

Besides approximating the Markovian ε-SIS process by a mean-field approximation, various other approximation methods exist. We investigate an eigenmode approximation of the Markovian ε-SIS process on K_N in Appendix A.4. Unfortunately, for an accurate approximation, the number of eigenmodes scales proportional to the number of nodes N, rendering the approximation method infeasible for large networks.

2.5. SPECTRUM ANALYSIS OF THE ε-SIS PROCESS

Since the eigenvalues λ_k are key for the characteristic time scales of the ε-SIS process, we focus on the determination of the eigenvalues λ_k of the transition matrix P from Eq. (2.2). The eigenvalues are computed in several parameter limits in Theorem 2.5.

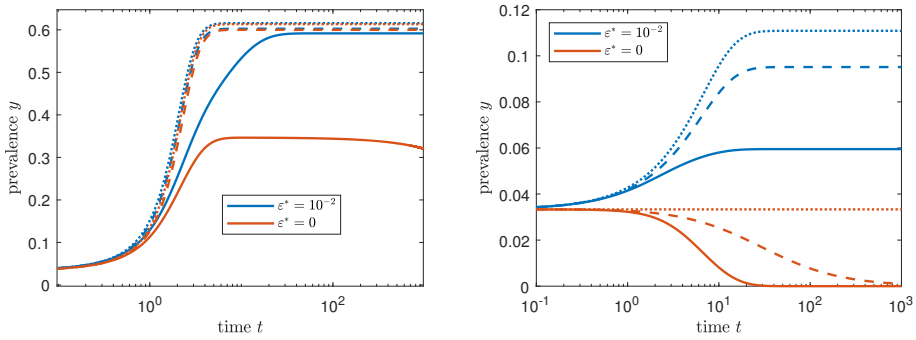


Figure 2.3: The exact Markovian solution (solid lines), the mean-field approximation without self-loops (dashed lines) and exact mean-field solution with self-loops from Theorem 2.4 (dotted lines) of the ε-SIS process on the complete graph with $N = 30$ nodes for various effective self-infection rates ε^* for (a) effective infection rate $\tau = 2.5/(N-1)$ and (b) effective infection rate $\tau = 1/(N-1)$.

Theorem 2.5 *The eigenvalues λ_k of the transition matrix P satisfy*[3]

$$\text{for } \varepsilon^* \to \infty \quad \lambda_k = -(k-1)\varepsilon^* - (k-1)(\tau(N+1-k)+1) + \mathcal{O}\left(\frac{1}{\varepsilon^*}\right),$$

$$\text{for } \tau \to \infty \quad \lambda_k = \begin{cases} -\frac{1}{4}\tau(N-1)(N+1) \pm \sqrt{\tau}(N+1)\sqrt{\frac{N-1}{2}} + \mathcal{O}(1), & \text{if } N \text{ odd, } k = \frac{N+2\pm1}{2} \\ -\tau\left(\frac{N}{2}-1\right)\left(\frac{N}{2}+1\right) + \frac{N^3}{12} + \frac{N^2}{12} - \frac{5N}{6} - \frac{1}{3} & \text{if } N \text{ even,} \\ \quad \pm\sqrt{1+\left(\frac{N}{2}\right)^3\left(\frac{N}{2}+1\right)^2\left(\frac{N}{2}-1\right)} + \mathcal{O}\left(\frac{1}{\sqrt{\tau}}\right), & k = \frac{N+2\pm2}{2}, \varepsilon^* = 1 \\ -\tau(k-1)(N-k+1) & \\ \quad + \left(-\varepsilon^*(N-k+1) + \frac{(k-1)(N-k+1)(N+1)}{(2k-N-3)(2k-N-1)}\right) + \mathcal{O}\left(\frac{1}{\sqrt{\tau}}\right), & \text{otherwise} \end{cases}$$

$$\text{for } \varepsilon^* < \frac{1}{N} \quad \lambda_{N+1} \gtrsim \begin{cases} -\left(\frac{1}{2x}+1+\frac{x}{2}\right)N, & \text{for } x \geq 1 \\ -2N, & \text{for } x < 1 \end{cases}$$

for $k = 1, 2, \ldots, N+1$ and where $x = \tau/\tau_c^{(1)}$ is the normalised effective infection rate.

Proof. See Appendix A.5. □

Theorem 2.5 states that the eigenvalue ratio ρ for large effective self-infection rates ε^* equals $\rho = \lambda_3/\lambda_2 \approx 2$. Thus metastability is not expected if the self-infection process dominates the infection and curing processes. For large, but finite effective infection rates τ and effective self-infection rates ε^*, the asymptotic expansions in Theorem 2.5 are only valid if the second term is strictly smaller than the first term, the third is smaller than the second, etc. For example, the second term in the expansion $\tau \to \infty$ must be strictly smaller than the first term, which only holds if $\tau > N+1$ (see Appendix A.5 for

[3] Contrary to our general consensus that the eigenvalues $0 = \lambda_1 > \lambda_2 > \ldots > \lambda_{N+1}$ are ordered, the eigenvalues λ_k in Theorem 2.5 are not necessarily ordered.

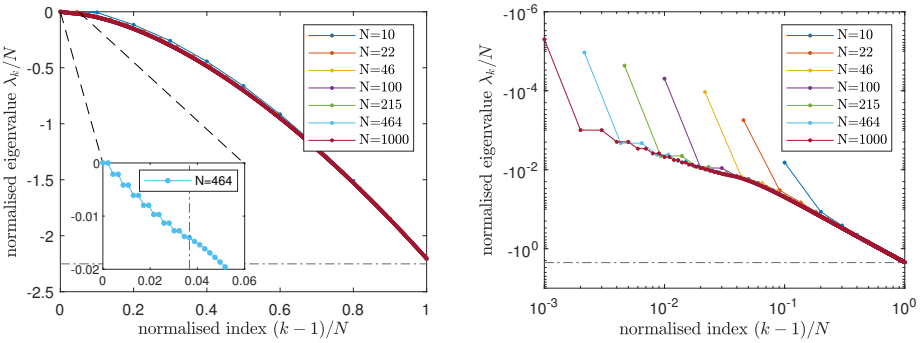


Figure 2.4: All eigenvalues λ_k of the transition matrix P on the complete graph with effective infection rate $\tau = 2/(N-1)$, effective self-infection rate $\varepsilon^* = 0.01/N$ on (a) linear-linear scale and (b) log-log scale for $k = 1,\ldots,N+1$. The eigenvalues have been normalised with respect to the number of nodes N. Normalised index 0 corresponds to the largest eigenvalue $\lambda_1 = 0$ and normalised index 1 corresponds to the smallest eigenvalue λ_{N+1}. The horizontal dash-dotted line indicates the lower bound for λ_{N+1} from Theorem 2.5. The largest eigenvalue $\lambda_1 = 0$ is not shown in subfigure (b) because of the logarithmic axis. The inset in (a) zooms in for $0 \le (k-1)/N \le 0.05$ and shows the critical index k_2 as a vertical dash-dotted line. In this example, the critical index $k_1 = 0$.

the derivation). Hence, the expansions in Theorem 2.5 provide valuable insights into the eigenvalues λ_k, even for finite values of the effective infection rate τ and effective self-infection rate ε^*.

We present here a full numerical analysis of the eigenvalues for finite parameter values. We compute the eigenvalues of the transition matrix P in Matlab using the command `eig`. Figure 2.4 illustrates the normalised eigenvalues λ_k/N versus the normalised index $(k-1)/N$ of the ε-SIS process for $k = 1,\ldots,N+1$. An interesting observation from Figure 2.4a is the negligible dependence of the network size N on the normalised eigenvalues λ_k/N. We further investigate the influence of the network size N by simulations in Section 2.6. Figure 2.4b shows that the second-largest eigenvalue λ_2 deviates significantly from the other eigenvalues $\lambda_3,\ldots,\lambda_{N+1}$. The difference between the second-largest eigenvalue λ_2 and the third-large eigenvalue λ_3 is the precise reason why we used the eigenvalue ratio ρ as the starting point for our Definition 2.3 of the metastable state in the ε-SIS process.

The inset of Figure 2.4a shows a "staircase" for the certain eigenvalues of the ε-SIS process. We propose to subdivide the eigenvalues into three regimes. We *define* the critical index k_1 as the start of the staircase and critical index k_2 as the end of the staircase. If the staircase does not exist, we take $k_1 = k_2 = 0$. The main plot of Figure 2.4a illustrates that the eigenvalues λ_k with $k > k_2$ roughly follow a quadratic relation between the normalised eigenvalues λ_k/N and the normalised index $(k-1)/N$. Theorem 2.5 states that the eigenvalues λ_k converge to $\tau(k-1)(N-k+1)$ if the effective infection rate τ tends to infinity, which explains the nearly quadratic form for $k > k_2$. On the other hand, the inset of Figure 2.4a shows staircase behaviour for $k_1 < k \le k_2$. The staircase is the result of two

nearly degenerate (but strictly distinct) eigenvalues of the ε-SIS process. If the effective infection rate τ tends to infinity, the eigenvalues converge to $\lambda_k \to \tau(k-1)(N-k+1)$, which are degenerate because $\lambda_k = \lambda_{N-k+2}$ for all $k = 1,\dots,N+1$. Even for large, finite effective infection rates τ, there is a small region $k_1 < k \leq k_2$ where staircase behaviour is observed, which is illustrated in Figure 2.4.

The staircase behaviour from Figure 2.4a is not always observed. Figure 2.5 illustrates the dependence of the critical indices k_1 and k_2 on the effective infection rate τ and the effective self-infection rate ε^*. Staircases are observed if the critical index $k_2 > 0$ in Figure 2.5, whereas the critical index k_1 indicates the start of the staircase and is only non-zero for a small region in the (τ, ε^*)-space. If the effective infection rate τ is below the epidemic threshold τ_c, both indices $k_1 = k_2 = 0$ and no staircases can be observed. In the limit $\varepsilon^* \to \infty$, Theorem 2.5 shows that staircases do not exist, because $\lambda_k \to -(k-1)\varepsilon^*$ for $k = 1,\dots,N+1$. In the intermediate regime, where $0 < \varepsilon^* < \infty$ and $\tau > \tau_c$, the critical index k_2 is often non-zero, indicating staircase behaviour. However, Figure 2.5 additionally shows small blue regions, where staircase behaviour is not observed. The blue regions are centred around $\varepsilon^* = \tau\left(n + \frac{1}{2}\right)$, where $n = 0, 1, 2, \dots$. Upon further inspection, the staircases seem best visible for $\varepsilon^* = \tau n$ where $n = 1, 2, \dots$.

For a given number of k infected nodes, the total effective infection rate $\tau k(N-k)$ and the total effective self-infection rate $\varepsilon^*(N-k)$. The effect of both infection processes is equally strong if its rates are equal, which implies that $\varepsilon^* = \tau k$. Thus, the staircases are best visible if the total rate of the infection process and self-infection process are equally large. We argue that one eigenvalue of the nearly-degenerate pair is due to the infection process and one eigenvalue corresponds to the self-infection process. Then, by varying either τ or ε^*, one of the eigenvalues of the pair must remain unchanged. Figure 2.6 shows the normalised eigenvalues for varying effective self-infection rates ε^* between τ and 2τ. Indeed, while the effective self-infection rate ε^* is varied, one eigenvalue of the nearly-degenerate pair changes whereas the other stays approximately constant. Thus, a plausible explanation for the staircase behaviour is a balance between self-infections and infections between nodes. While Figure 2.6 supports our explanation for the staircase behaviour, an analytic proof remains an open research question.

To summarise, we believe that the eigenvalues λ_k in the regime $2 < k \leq k_1$ are related to self-infection-dominated behaviour, the regime $k_1 < k \leq k_2$ describes when the influence of the infection and self-infection process is equally strong and the regime $k > k_2$ pertains to the infection-dominated behaviour. Equation (2.11) summarises our findings for the eigenvalues λ_k of the ε-SIS process:

$$\lambda_k \approx \begin{cases} 0 & \text{for } k = 1, \\ \text{convergence rate} & \text{for } k = 2, \\ \text{roughly linear} & \text{for } 2 < k \leq k_1, \\ \text{staircases} & \text{for } k_1 < k \leq k_2, \\ \text{roughly quadratic} & \text{for } k_2 < k \leq N+1. \end{cases} \tag{2.11}$$

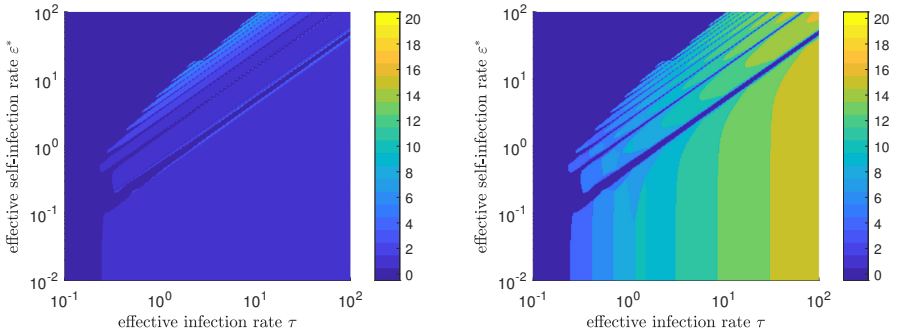


Figure 2.5: The critical indices k_1 (left) and k_2 (right) for different values of the effective infection rate τ and the effective self-infection rate ε^* on the complete graph K_N with $N = 20$ nodes.

2.6. NUMERICAL SIMULATIONS

We perform numerical simulations to examine how the entire set of eigenvalues changes with the effective self-infection rate ε^*, the effective infection rate τ and the network size N.

2.6.1. THE INFLUENCE OF THE EFFECTIVE SELF-INFECTION RATE ε^*

First, we examine the influence of the effective self-infection rate ε^* on the eigenvalues, by fixing the network size N and the effective infection rate τ. Figure 2.7 shows the absolute value of the eigenvalues of the transition matrix P for varying effective self-infection rates ε^*.

The ε-SIS process is governed by two infection processes: The infection process and the self-infection process. All events in the continuous-time Markovian ε-SIS process are independent, such that the rate to transition from a completely susceptible population to one infected node is solely governed by the self-infection process and may happen for every node independently, leading to a total rate $\varepsilon^* N$. Simultaneously, if a single node is infected, another node is infected with rate $\tau(N-1)$. If $\varepsilon^* N < \tau(N-1)$, the process is dominated by the infection process, whereas the process is dominated by self-infections if $\varepsilon^* N > \tau(N-1)$. For large network sizes N, the influence of the self-infection process and the infection process is equally large if $\tau \approx \varepsilon^*$. The vertical line in Figure 2.7 indicates the relation $\tau = \varepsilon^*$. The minimum of the eigenvalue ratio ρ and the relation $\tau = \varepsilon^*$ coincide in Figure 2.7c. At the intersection point $\tau = \varepsilon^*$, the minimal eigenvalue ratio ρ is approximately one. Then both eigenvalues are approximately equal and exhibit staircases in the eigenvalue spectrum, which was discussed in detail in Section 2.5.

The second-largest eigenvalue λ_2 changes significantly from $\varepsilon^* = 0$ to $\varepsilon^* = \tau$, but the remaining eigenvalues $\lambda_3, \ldots, \lambda_{N+1}$ stay nearly constant. To the right of the vertical line in Figure 2.7a, the remaining eigenvalues increase as well. In the limit of large effective self-infection rates ε^*, it holds that $\lambda_k \approx -(k-1)\varepsilon^*$ according to Theorem 2.5. The smallest eigenvalue $\lambda_{N+1} \approx -\varepsilon^* N$ and the second-largest eigenvalue $\lambda_2 \approx -\varepsilon^*$ are shown by

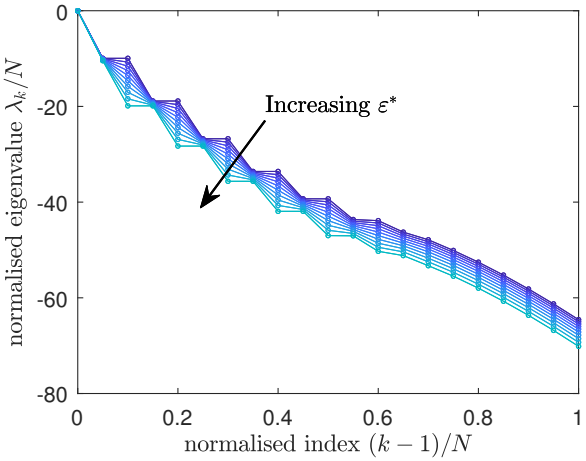


Figure 2.6: The normalised eigenvalues λ_k/N on the complete graph K_N with $N = 20$ nodes and $\tau = 10$ for different effective self-infection rates between $\varepsilon^* = 10$ and $\varepsilon^* = 20$. By varying the effective self-infection rate ε^*, one eigenvalue of the nearly-degenerate pair changes significantly whereas the other remains largely unchanged.

dashed lines in Figure 2.7a. Thus, the eigenvalues λ_k increase linearly with the effective self-infection rate ε^*, because self-infection process dominates the other processes and drives the ε-SIS process to the all-infected state.

If the effective self-infection rate $\varepsilon^* < \tau$, Figure 2.7 shows that the self-infection process barely influences the characteristic timescales of the ε-SIS process, which was also observed in Figure 2.2. A seemingly contradictory result was obtained by Van Mieghem [21], who showed that the steady-state prevalence y_∞ exhibits an explosive phase transition at certain small effective self-infection rates ε^*. The difference is that we consider a fixed effective infection rate τ and vary the effective self-infection rate ε^*, which contrasts [21], where the effective infection rate τ is varied for fixed self-infection rates $\varepsilon^* = 0$ and $\varepsilon^* > 0$. Performing a similar analysis as [21] on the eigenvalues, Figure 2.8 illustrates that the second-largest eigenvalue λ_2 is heavily influenced by the effective self-infection rate ε^*. For any finite effective self-infection rate $\varepsilon^* > 0$, there exists a phase transition for some effective infection rate τ, where the second-largest eigenvalue λ_2 converges to a constant for large effective infection rates τ. In the limit $\varepsilon^* \to 0$, no such transition is observed, which is in agreement with [21]. The other eigenvalues $\lambda_3, \ldots, \lambda_{N+1}$ remain largely unaffected by considering the limit $\varepsilon^* \to 0$.

2.6.2. THE INFLUENCE OF THE EFFECTIVE INFECTION RATE τ

Analogously to the effective self-infection rate ε^*, we analyse the influence of the effective infection rate τ on the eigenvalues of the transition matrix P in Figure 2.9. The vertical line in Figure 2.9 illustrates the mean-field epidemic threshold $\tau_c^{(1)} = \frac{1}{N-1}$, which is slightly smaller than the true epidemic threshold [23, 49]. Below the epidemic threshold τ_c, the eigenvalue ratio ρ in Figure 2.9c is small. Around the epidemic threshold τ_c,

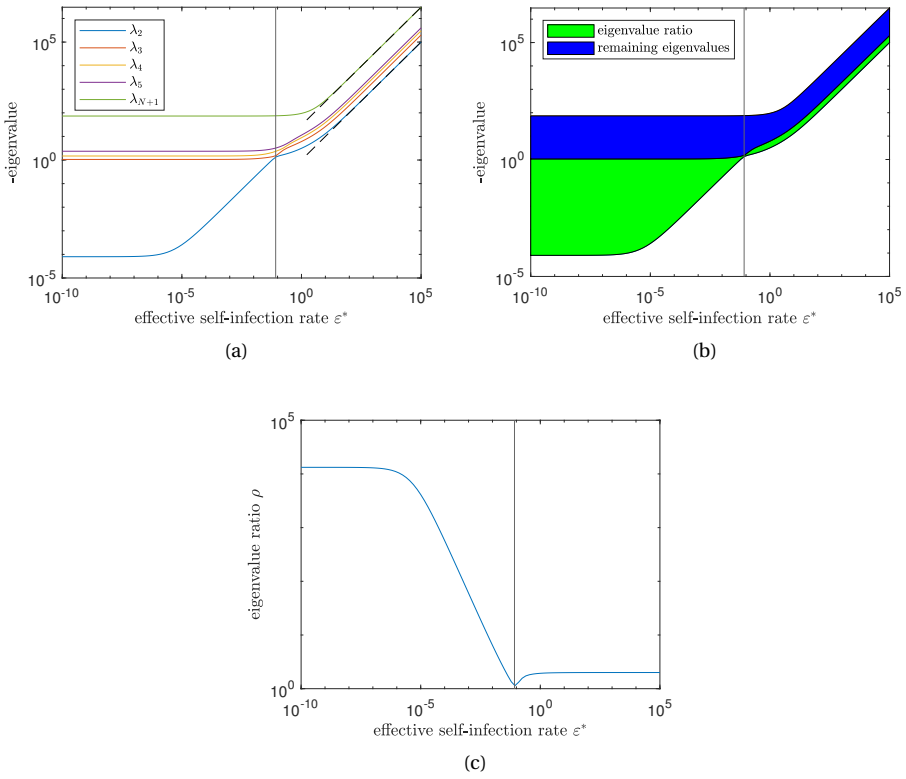


Figure 2.7: The eigenvalues of the transition matrix P for the complete graph with $N = 30$ nodes and effective infection rate $\tau = 2.5\tau_c^{(1)}$ for varying effective self-infection rates ε^* on a log-log scale. Subfigure (a) shows the second-to-fifth-largest eigenvalues $\lambda_2, \lambda_3, \lambda_4, \lambda_5$ and smallest eigenvalue λ_{N+1} as a function of the effective self-infection rate ε^*. The vertical line indicates the relation $\varepsilon^* = \tau$. Subfigure (b) is equivalent to (a), but here the eigenvalue ratio ρ is coloured green (light) and the area with the other eigenvalues is coloured blue (dark). Finally, (c) shows a plot of the eigenvalue ratio ρ versus the effective self-infection rate ε^*.

the eigenvalue ratio ρ increases rapidly as illustrated in Figure 2.9c.

For large effective infection rates τ, the second-largest eigenvalue λ_2 converges to $-\varepsilon^* N$ whereas the remaining eigenvalues $\lambda_3, \ldots, \lambda_{N+1}$ increase linearly with the effective infection rate τ, which is in line with Theorem 2.5. Hence, the eigenvalue ratio ρ tends to infinity if $\tau \to \infty$ and the system is considered metastable according to our Definition 2.3. For large effective infection rates τ and starting with a non-zero number of infected nodes, the remaining nodes will be infected extremely quickly. The spreading is only slowed down by the transition from 0 to 1 infected node. The metastable state here is the all-healthy state, which takes a considerable amount of time to leave, whereafter the process converges extremely fast to the all-infected state. The convergence rate from the metastable state to the steady state, which is $-\lambda_2$, exactly equals the rate to jump from 0 to 1 infected node, which is given by the scaled birth rate $\tilde{\Xi}_0 = \varepsilon^* N$.

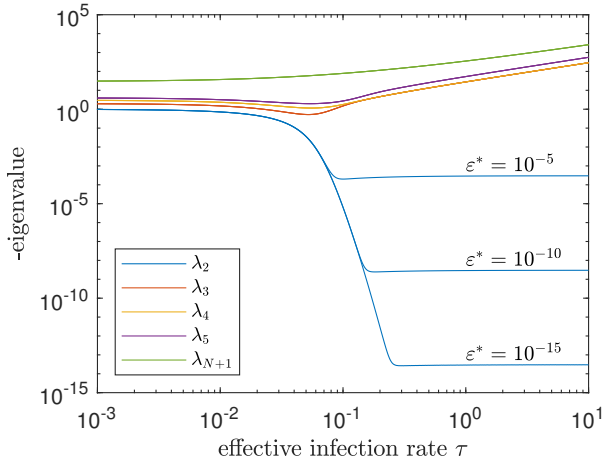


Figure 2.8: Illustration of the explosive phase transition for effective self-infection rates $\varepsilon^* > 0$ and no phase transition for $\varepsilon^* = 0$ for the second-largest eigenvalue λ_2. The remaining eigenvalues $\lambda_3, \ldots, \lambda_{N+1}$ are nearly indistinguishable for different ε^*. The network size equals $N = 30$.

The epidemic threshold τ_c in the Markovian ε-SIS process exhibits a second-order phase transition[4]. Van Mieghem and Cator [16, p. 9] derived the following relation for the epidemic threshold τ_c for ε-SIS dynamics on the complete graph for large network sizes $N \gg 1$ and for small effective self-infection rates $\varepsilon^* < \frac{1}{N}$:

$$\tau_{c,\infty} = \frac{1}{N+2-2\sqrt{N+1}}. \tag{2.12}$$

We expand both approximations for the epidemic threshold τ_c for large network sizes $N \gg 1$:

$$\begin{aligned} \tau_c^{(1)} &= \frac{1}{N-1} &&= \frac{1}{N}\left[1 + \frac{1}{N} + \mathcal{O}\left(\frac{1}{N^2}\right)\right], \\ \tau_{c,\infty} &= \frac{1}{N+2-2\sqrt{N+1}} &&= \frac{1}{N}\left[1 + \frac{2}{\sqrt{N}} + \frac{2}{N} - \frac{1}{N\sqrt{N}} + \mathcal{O}\left(\frac{1}{N^2}\right)\right]. \end{aligned}$$

Figure 2.9 illustrates that the eigenvalues λ_2, λ_3 and λ_4 reach a minimum at a certain infection rate τ. We verify our hypothesis that the minimum of λ_2, λ_3 or λ_4 coincides with the true epidemic threshold by plotting the mean-field threshold $\tau_c^{(1)}$, Van Mieghem and Cator's threshold $\tau_{c,\infty}$ and the numerically obtained effective infection rates τ for which the eigenvalues λ_2, λ_3 and λ_4 attain a minimum in Figure 2.10. Additionally, we compute the steady-state prevalence y_∞ for the ε-SIS process, and take the derivative of the steady-state prevalence y_∞ with respect to the effective infection rate τ.

[4] A first-order, abrupt phase transition at the epidemic threshold τ_c exhibits a discontinuity in the first derivative of the steady-state prevalence y_∞. A second-order, continuous phase transition is an interval $[a, b]$ in which the behaviour of the ε-SIS process gradually changes from nearly exponential die-outs to long-lived epidemic outbreaks.

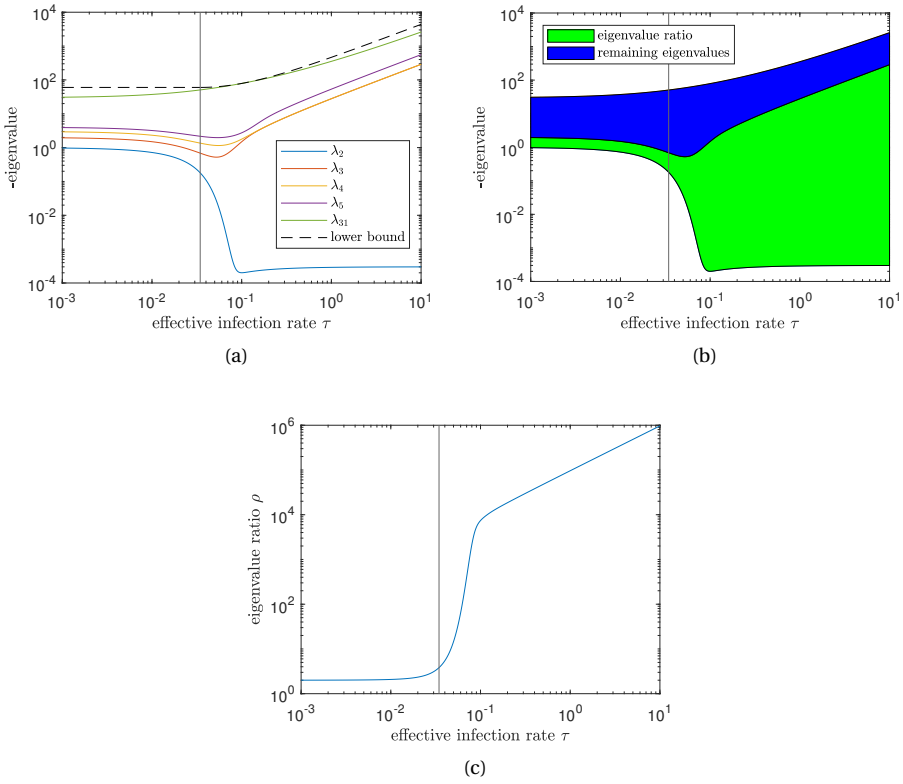


Figure 2.9: The eigenvalues of the transition matrix P for the complete graph with $N = 30$ nodes and effective self-infection rate $\varepsilon^* = 10^{-5}$ for varying effective infection rates τ. Subfigure (a) shows the second-to-fifth-largest eigenvalues $\lambda_2, \lambda_3, \lambda_4, \lambda_5$ and smallest eigenvalue λ_{N+1} as a function of the effective infection rate τ. The vertical line indicates the mean-field epidemic threshold $\tau_c^{(1)}$, which is slightly smaller than the true epidemic threshold τ_c. Subfigure (b) is equivalent to (a), but here the eigenvalue ratio ρ is coloured green (light) and the area with the other eigenvalues is coloured blue (dark). Finally, (c) shows a plot of the eigenvalue ratio ρ versus the effective infection rate τ. Finally, (a) illustrates the lower bound for the smallest eigenvalue λ_{N+1} from Theorem 2.5 by a dashed line.

Then the epidemic threshold follows as the effective infection rate τ for which $dy_\infty/d\tau$ is maximal[5]. The maximum of $dy_\infty/d\tau$ indicates for which effective infection rate τ, the steady-state prevalence grows the fastest, which is presumably a good indicator of the epidemic threshold. Figure 2.10 depicts that the effective infection rate τ where the minimum λ_3 is attained, is very close to the epidemic threshold based on the maximum of $dy_\infty/d\tau$. Interestingly, Van Mieghem and Cator's epidemic threshold $\tau_{c,\infty}$ is always larger than the mean-field threshold $\tau_c^{(1)}$, but always lower than the other estimates.

Figure 2.11 shows the steady-state prevalence y_∞ and the eigenvalue ratio ρ for varying infection rates τ. Above the epidemic threshold τ_c, the eigenvalue ratio ρ increases

[5]We further investigate the maximum of $dy_\infty/d\tau$ as a function of the network size N in Section 2.6.3.

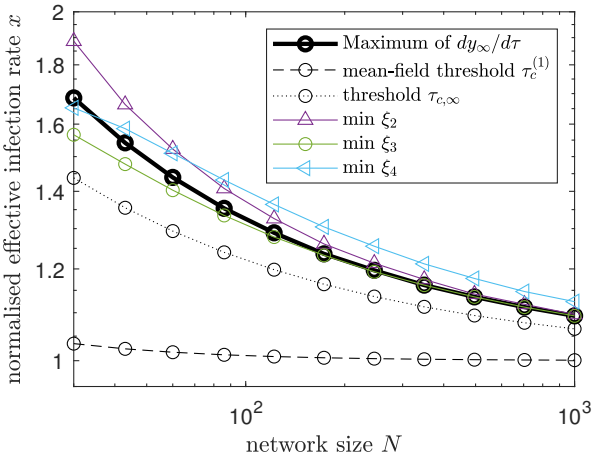


Figure 2.10: Several formulas for the suggested epidemic threshold τ_c for various network sizes N.

significantly and the time-dependent ε-SIS process will show metastable behaviour. Below the epidemic threshold τ_c, metastability is never observed because the eigenvalue ratio ρ is small. We conclude here that the epidemic threshold τ_c not only describes for which infection rates τ the epidemic persists or dies out, but is additionally a good descriptor whether the ε-SIS process exhibits metastable behaviour.

For non-complete graphs, we expect that a similar conclusion will hold, but the 2^N-sized Markov chain for general graphs prevents us from rigorously demonstrating this claim. For connected Erdős-Rényi graphs with $N \leq 12$ nodes and a randomly chosen link-connectivity p, simulations indicate that the eigenvalue ratio ρ and the steady-state prevalence y_∞ show a similar plot as for the complete graph in Figure 2.11.

2.6.3. THE INFLUENCE OF THE NETWORK SIZE N

Finally, we investigate the effect of the network size N on the ε-SIS process. In the limit $N \to \infty$, the Markovian ε-SIS process exactly converges to the mean-field approximation [50]. Figure 2.12 shows the steady-state prevalence y_∞ for various network sizes and the mean-field steady-state prevalence $y_\infty^{(1)} = 1 - 1/x$. By increasing the network size N, the steady-state prevalence y_∞ converges to the mean-field approximation. One of the methods in Section 2.6.2 to estimate the epidemic threshold τ_c is based on the computation of the derivative $dy_\infty/d\tau$, which in plotted in Figure 2.12b. We estimate the epidemic threshold τ_c based on the peak of $dy_\infty/d\tau$, which converges to the mean-field threshold $x = 1$ if the network size N increases to infinity. All curves have an $1/x^2$ tail, which agrees with the mean-field steady-state prevalence $dy_\infty^{(1)}/d\tau = 1/x^2$.

We further investigate the eigenvalues of the ε-SIS process in Figure 2.13. In Figure 2.13, we scale the effective infection rate τ/N and the effective self-infection rate ε^*/N. If the effective infection rate $\tau = 1.5\tau_c^{(1)}$ is larger than the epidemic threshold τ_c, as shown in Figure 2.13a, then by fixing k, all eigenvalues λ_k converge to a constant value in the limit $N \to \infty$. For example, the second-largest eigenvalue λ_2 stays nearly constant for

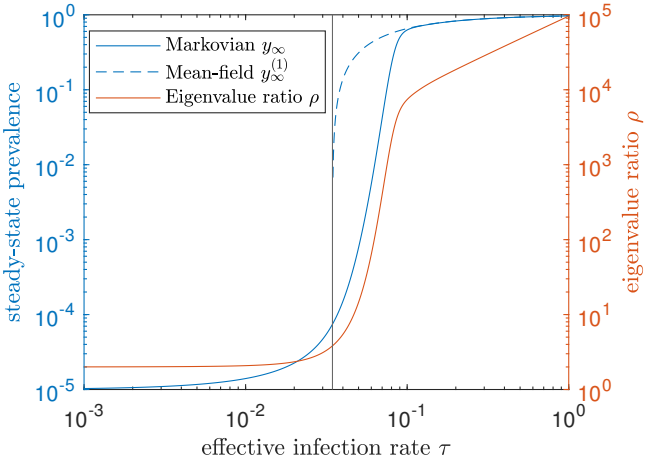


Figure 2.11: The steady-state prevalence y_∞, the mean-field steady-state prevalence $y_\infty^{(1)}$ and the eigenvalue ratio ρ with $N = 30$ nodes and effective self-infection rate $\varepsilon^* = 10^{-5}$ for varying effective infection rate τ. The vertical line indicates the mean-field epidemic threshold $\tau_c^{(1)} = \frac{1}{N-1}$, which is slightly smaller than the true epidemic threshold τ_c.

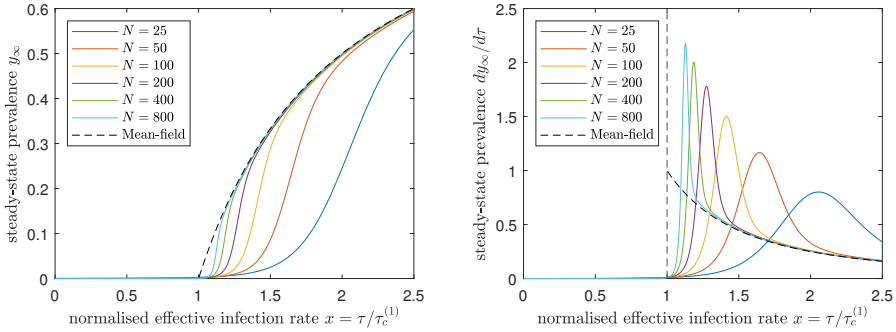


Figure 2.12: The normalised effective infection rate $x = \tau/\tau_c^{(1)}$ versus the steady-state prevalence y_∞ for varying network sizes N and the mean-field steady-state prevalence $y_\infty^{(1)} = 1 - \frac{1}{x}$ and $\varepsilon^* = 0.01/N$.

$N \gtrsim 100$. Closer to the epidemic threshold, namely for $\tau = 1.2\tau_c^{(1)}$ in Figure 2.13b, convergence occurs near $N \gtrsim 600$.

The mean-field approximation of the ε-SIS process is often used in network theory to reduce the computational complexity of the 2^N-sized Markov chain. In the limit of the network size N approaching infinity, the mean-field approximation becomes the exact solution, at least for the complete graph [50]. In practice, however, finite-sized networks are also approximated by mean-field methods, introducing an approximation error. It is known that mean-field methods perform poorly around the epidemic threshold,

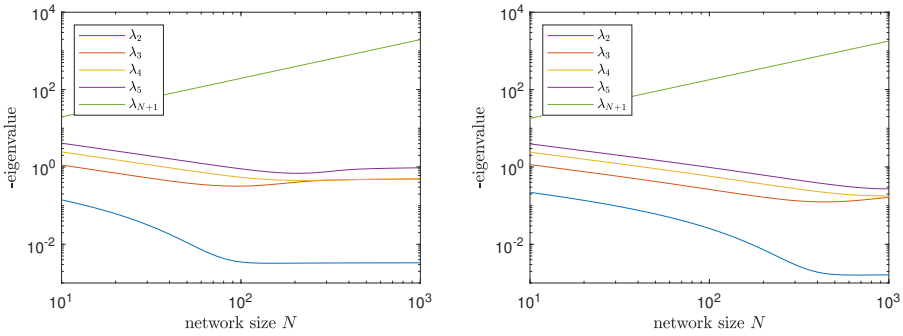


Figure 2.13: The eigenvalues of the transition matrix P for the complete graph with effective self-infection rate $\varepsilon^* = 0.01/N$. Both subfigures show the second-to-fifth-largest eigenvalues $\lambda_2, \lambda_3, \lambda_4, \lambda_5$ and smallest eigenvalue λ_{N+1} as a function of the network size N. Subfigure (a) corresponds to effective infection rate $\tau = 1.5\tau_c^{(1)} = 1.5/(N-1)$ and (b) to $\tau = 1.2\tau_c^{(1)} = 1.2/(N-1)$.

because the assumed independence of the stochastic variables does not hold. Around the epidemic threshold, the number of nodes N in the complete graph required to obtain a reasonable accuracy with the mean-field method, tends to increase closer to the epidemic threshold. We illustrate the aforementioned statement by focussing on the second-largest eigenvalue λ_2 of the transition matrix P. We define the *critical network size* N_c as the smallest network size N for which the second-largest eigenvalue λ_2 has a relative error of less than 10^{-6} compared to λ_2 in the thermodynamic limit (i.e. when $N \to \infty$). Figure 2.14 illustrates that the critical network size N_c increases if the epidemic threshold is approached from above.

2.7. CONCLUSION

In this chapter, we analysed the continuous-time Markovian ε-SIS process on the complete graph with N nodes. The transition matrix corresponding to the underlying Markov chain has $N+1$ distinct eigenvalues, of which the largest eigenvalue is zero and corresponds to the steady state. The remaining eigenvalues are all negative and distinct. Metastable behaviour can be observed in the ε-SIS process if the ratio between the second-largest and third-largest eigenvalue of the transition matrix is sufficiently large. The remaining eigenvalues are nearly degenerate for large effective infection rates, which results in staircases in the eigenvalue plot. The staircases are best visible if $\varepsilon^* = n\tau$, where $n = 1, 2, \ldots$. The epidemic threshold can be accurately estimated using the effective infection rate for which the third-largest eigenvalue of the transition matrix is the smallest. We additionally showed that the epidemic threshold does not only distinguish between small and large epidemic outbreaks, but additionally describes when the ε-SIS process may exhibit metastable behaviour.

Even though we confined ourselves in this chapter to the complete graph, we believe that some of our results are physical characteristics of ε-SIS dynamics and may hold for

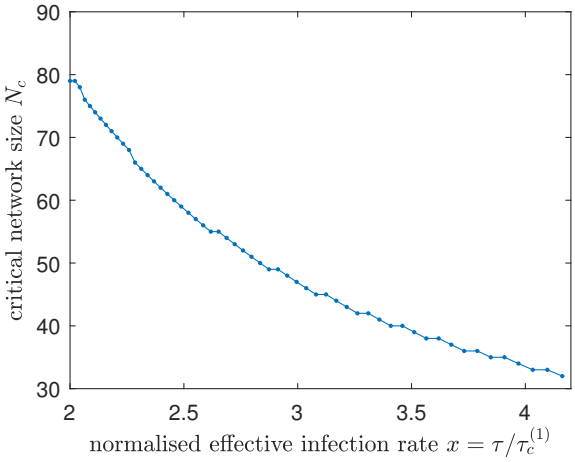


Figure 2.14: The critical network size N_c as a function of the normalised effective infection rate $x = \tau/\tau_c^{(1)}$. The analysis is conducted on the complete graph with effective self-infection rate $\varepsilon^* = 0.01/N$. The critical network size N_c, which indicates when the ε-SIS process can be safely approximated using a mean-field method, increases if the normalised effective infection rate x approaches the epidemic threshold $x_c = 1$.

any graph. Monte Carlo simulations show metastable behaviour for the ε-SIS process on general graphs [35], in non-Markovian SIS epidemics [51] and also in SIS processes on time-varying or adaptive networks [52]. Therefore, our results on the complete graph may describe the general physical behaviour of metastability in ε-SIS dynamics on networks.

3

ANALYTIC SOLUTIONS OF COMPARTMENTAL EPIDEMICS WITHOUT REINFECTIONS

Many epidemic processes on networks can be modelled by compartmental models. The corresponding compartmental graph describes how the viral state of the nodes changes from one compartment to another. If the compartmental model does not allow for re-infections, i.e. the compartmental graph does not contain loops, we provide an analytic closed-form solution of the continuous-time Markovian compartmental model on heterogeneous networks. The exact eigenvalues of the linear Markovian process are related to cut sets in the graph between nodes with states in different compartments. We illustrate our finding by analytically solving the continuous-time Markovian SI and SIR processes on heterogeneous networks. We show that analytic extensions to non-Markovian dynamics, temporal networks and simplicial contagion are possible.

3.1. INTRODUCTION

Most research in network epidemiology focuses on so-called compartmental models, in which the population is subdivided into several groups based on the current viral state of the individual [6]. General epidemic models are comprised of c compartments. We assume that interactions only occur between neighbouring nodes and that after such an interaction, at most one of those nodes changes its viral state. Then the transitions between compartments can be classified into two types: (i) link-based transitions, such as the infection of a particular node by another node, which requires a link between the nodes and (ii) node-based transitions, such as nodes recovering from the disease [53]. Many results have been obtained in the well-mixed assumption, i.e. all individuals can directly contact all other individuals [54] and the contact graph is a complete graph. In reality, each individual has different characteristics and cannot contact any other member of the population, leading to heterogeneous behaviour during an epidemic [6]. Therefore, we focus in this chapter on epidemic processes on heterogeneous networks.

Properties of such network-based models are usually computed using mean-field approximations [8] or Monte-Carlo simulations [55]. The downside is that Monte-Carlo simulations are known to converge slowly, which is problematic if the probability of a certain event is small. On the other hand, the mean-field approximation performs well for homogeneous, well-mixed graphs but is less precise for heterogeneous networks [32]. Instead of resorting to either Monte-Carlo simulations or mean-field approximations, we focus in this chapter on determining the *exact* time-dependent solution for general Markovian compartmental models.

For the computation the exact time-dependent solution, we require an efficient enumeration of all states of the underlying Markov chain. Since each of the N nodes can attain c viral states, the number of states equals c^N. The set of all possible states is denoted by $\mathbf{S}$. Providing explicit constructions for the states[1] and the corresponding infinitesimal generators on heterogeneous networks for general Markovian compartmental models is a tedious task. Several researchers have attempted to construct infinitesimal generators for specific epidemic models on networks. For example, Van Mieghem *et al.* [56] provided a construction based on binary numberings for SIS epidemics and for SIS epidemics with self-infections [16]. Simon *et al.* [27] suggested a tridiagonal block structure for the infinitesimal generator of the SIS process. Each block contains configurations with the same number of infected nodes k. Configurations with k infected nodes can only make transitions to configurations with $k+1$ and $k-1$ infected nodes, such that the infinitesimal generator Q has a block-tridiagonal structure. Economou *et al.* [57] developed a similar tridiagonal block structure, but used a different ordering within each block. For SIR epidemics, López-García [58] groups configurations based on the number of recovered nodes. Within each block with the same number of recovered nodes, configurations are grouped based on the number of infected nodes. Within each block with the same number of recovered *and* infected nodes, a lexicographical ordering is applied. The Generalised Epidemic Mean-Field (GEMF) model by Sahneh *et al.* [53] describes disease transmission in general compartmental models, where the state

[1] Throughout this chapter, we will use *state* and *configuration* when referring to one element of the state space $\mathbf{S}$, interchangeably.

space is constructed using a tensor product formulation. Merbis [59] proceeded with a similar tensor product construction to derive the time-dependent equations for general epidemic models. In this tensor formulation, Merbis and Lodato [60] derived an exact solution for the SI process on unweighted graphs. Another representation in terms of a tensor-product formulation for SIR epidemics is provided by Dolgov and Savostyanov [61].

General compartmental epidemic models, as formalised in the GEMF framework [53], can be solved by the eigendecomposition of the corresponding infinitesimal generator. If the compartmental graph is a tree, i.e. the compartmental graph does not contain any loops, then the corresponding Markov graph is also a tree, whose infinitesimal generator Q only has upper-triangular non-zero elements. Consequently, the eigenvalues and eigenvectors can be determined efficiently. If loops or cycles occur, as in the ε-SIS process from Chapter 2, then the infinitesimal generator Q is no longer upper-triangular and our solution method is not applicable, because the eigenvalues and eigenvectors can only be determined numerically, which is infeasible for large graphs.

In this chapter, we present for the first time an analytic closed-form solution of continuous-time Markovian epidemic processes without reinfections on heterogeneous networks. We demonstrate our method by considering SI epidemics in Section 3.2. We present an in-depth explanation of the eigenvalues of the infinitesimal generator in Section 3.2.1 and derive the full solution in Section 3.2.2. We discuss the benefits and difficulties of the exact solution in Section 3.2.3 and compare our exact results with simulations in Section 3.2.4. We then generalise our exact solution on non-Markovian dynamics, temporal networks, simplicial contagion and self-infections in Section 3.2.5. We also provide an extensive analysis on SIR epidemics in Section 3.3. We provide exact eigenvalues in Section 3.3.1, SIR eigenvectors in Section 3.3.2 and the full solution in Section 3.3.3. We demonstrate the power of our exact method by analytically computing the time of the epidemic peak in Section 3.3.4. Finally, we conclude in Section 3.4.

3.2. THE SI PROCESS

The Susceptible-Infected (SI) process describes the spread of information, diseases, innovations or neural activity over a network. Starting from a single infected node, the infection spreads to all its neighbours, which again infects their neighbours, until the whole network is infected [62]. The Markovian SI process is actually a Markovian discovery process, or, equivalently, a stochastic shortest path process [44, Chapter 16]. The SI process describes spreading phenomena without damping or curing (for example, HIV viruses in humans or the spread of innovations) or the spread evolves extremely fast, effectively flooding the population and prohibiting curing or recovering (for example, epileptic seizures in the human brain [63] and the spread of information on social media).

The SI process subdivides the nodes into two compartments: the set $\mathcal{I} \subseteq \mathcal{N}$ of infected (I) nodes are infected, aware or activated and the complementary set $\mathcal{S} = \mathcal{N} \setminus \mathcal{I}$ of susceptible (S) nodes are healthy, unaware or idle. We further assume that an infection over the direct link from node i towards a susceptible node j is a Poisson process with rate β_{ij} and is independent of all other links. By this assumption, the SI process is a continuous-time Markov process containing 2^N states, because each of the N

nodes in the graph is either susceptible or infected. The $2^N \times 1$ state vector is $\mathbf{s}(t) = (s_0(t), s_1(t), \ldots, s_{2^N-1}(t))^T$, whose element $s_i(t)$ describes the probability that the SI process is in state i at time t. Since the process must certainly be, for any time t, in one of the possible states, $\sum_{i=0}^{2^N-1} s_i(t) = 1$ for all times t. The transition rates between the states are denoted by the $2^N \times 2^N$ infinitesimal generator Q.

We describe the SI process using the labeling in [56]. The idea of the construction in [16, 56] is that the state i represents the viral state of all N nodes in the graph. We define the binary variables $x_k(i)$ indicating whether node k is infected in configuration i. Then the *viral state vector* $\mathbf{x} = (x_N, x_{N-1}, \ldots, x_2, x_1)^T$ describes the viral state of all nodes. By regarding the viral state vector $\mathbf{x}$ as a binary number, the configuration number i is computed as

$$i = \sum_{k=1}^{N} x_k(i) 2^{k-1},$$

where $x_k = 1$ if node k is infected and $x_k = 0$ if node k is susceptible. For the complete graph with $N = 4$ nodes, the possible transitions for the SIS process are shown in Figure 3.1. For example, the all-healthy state $i = 0$ has representation (0000), indicating that all nodes in state 0 are healthy. Similarly, for state 5 the binary representation is (0101), indicating that node 2 and 4 are healthy and node 1 and 3 are infected.

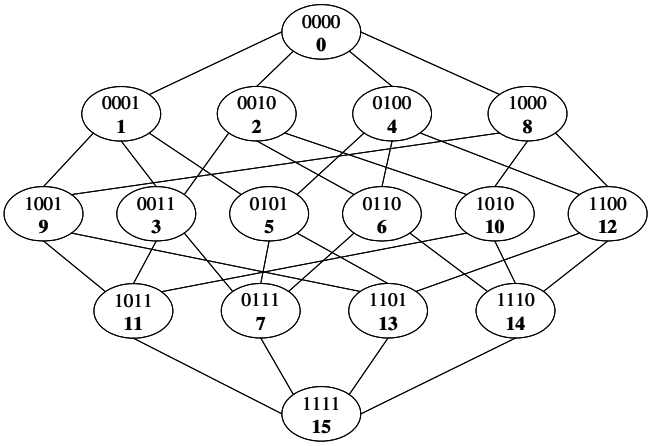


Figure 3.1: The Markov graph or Hasse diagram for the SIS process with $N = 4$ nodes from [56].

A fundamental property of Markov processes is that events in continuous time occur sequentially and simultaneous events do not happen almost surely. This means that transitions between states in the Markov chain can only occur if the binary representation of those states differs exactly one bit, which is illustrated in Figure 3.1. However, not all links in Figure 3.1 correspond to actual transitions in the SI process. Contrary to the SIS process that allows for the curing of nodes, the SI process only considers the infection process. Hence, the upward-pointing transitions in the Markov graph in Figure 3.1 cannot occur, because upward transitions correspond to the curing process. Furthermore, if the SI process is in the all-healthy state, then the spread never activates, because

there is no initial infection. Hence, the all-healthy state $i = 0$ cannot be entered nor left, thus state $i = 0$ has no incoming nor outgoing arrows in the SI process. Finally, we remark that the underlying graph G describes the existence of links between pairs of nodes. Naturally, the existence and weight of the links in the Markov graph are influenced by the existence of links in the original graph G.

Using the binary notation of the state space $\mathbf{S}$, we describe the transitions between the states using the $2^N \times 2^N$-dimensional infinitesimal generator Q. To simplify the notation, we introduce the $N \times N$, possibly asymmetric, weighted and non-negative adjacency matrix B with elements $\tilde{\beta}_{kl} = a_{kl}\beta_{kl}$. The elements q_{ij} of the infinitesimal generator Q for the SI process are [44, Chapter 17]

$$q_{ij} = \sum_{k=1}^{N} \tilde{\beta}_{mk} x_k(i), \qquad \text{if } j = i + 2^{m-1} \tag{3.1a}$$

$$\text{with } m = 1, 2, \ldots, N \text{ and } x_m(i) = 0.$$

$$q_{ii} = -\sum_{\substack{j=0 \\ j \neq i}}^{2^N - 1} q_{ij}, \tag{3.1b}$$

for $i, j = 1, 2, \ldots, 2^N$. The non-zero elements of the infinitesimal generator Q are visualised in Figure 3.2. The infinitesimal generator Q is an upper triangular matrix, because in absence of curing events, there is no transition from state i to state $j < i$ because the number of infected nodes is larger for state i than for state j. The infinitesimal generator Q for a graph with N nodes can also be constructed recursively in terms of the infinitesimal generator of a graph with $N - 1$ nodes and then adding one node [16].

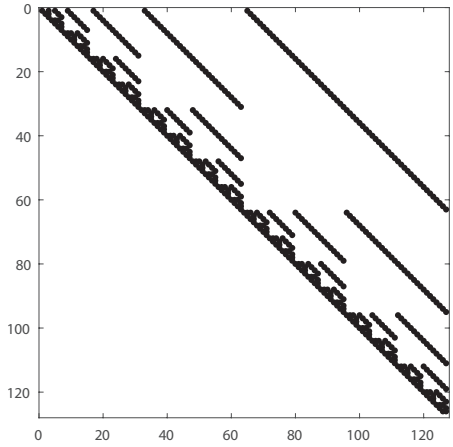


Figure 3.2: The structure of the infinitesimal generator Q for the SI process on a complete graph with $N = 7$ nodes.

The governing equation for the continuous-time Markov chain is

$$\frac{d\mathbf{s}^T(t)}{dt} = \mathbf{s}^T(t)Q, \tag{3.2}$$

whose solution is

$$\mathbf{s}^T(t) = \mathbf{s}^T(0)e^{Qt}.$$

The solution can be further detailed as [56, p. 4]

$$\mathbf{s}(t) = \pi + \sum_{i=1}^{2^N-1} e^{\lambda_i t} \sum_{m=0}^{n_i-1} \mathbf{r}_{i,m} \frac{t^m}{m!}, \tag{3.3}$$

where n_i denotes the multiplicity of eigenvalue λ_i, the vector $\mathbf{r}_{i,m}$ is related to the right- and left-eigenvectors of Q, the initial condition $\mathbf{s}(0)$ and the steady-state vector π, which equals $\pi = (0,0,\ldots,0,1)^T$. Unfortunately, the infinitesimal generator Q is asymmetric and is not even normal[2], preventing further simplifications in Eq. (3.3).

The SI process is completely described by the set (3.2) of 2^N linear differential equations. Frequently, an exponentially large state space is the fingerprint of a non-linear process. In particular, SIS and SIR epidemics exhibit a phase transition around the epidemic threshold. Below the threshold, the epidemic dies out exponentially fast, whereas above threshold, a non-zero fraction of the population remains infected over a long time. On the contrary, the SI process does not exhibit a phase transition, because the SI process always converges – given that the process does not start in the all-healthy state – to the all-infected state, irrespective how small, but non-zero, the infection rates $\beta_{ij} > 0$ are. For simple graphs with homogeneous transition coefficients[3], like the complete graph or the star graph, an exact analysis is possible [23] by exploiting symmetry of the graph [27]. Otherwise, exactly solving equation (3.2) for the SI process is infeasible for graphs with $N > 15$ nodes and one resorts to numerical simulations to solve (3.2). One can simulate the SI process by Monte Carlo simulations in discrete time [64] or in continuous time using a Gillespie algorithm [55].

The general solution (3.3) contains the eigenvalues and eigenvectors of the infinitesimal generator Q. The eigenvalues λ_i of the infinitesimal generator Q are of primary importance in the solution (3.3), because the eigenvalues are inversely proportional to the relaxation time of the corresponding eigenmode. The eigenvalue λ_i is therefore related to the average time for the SI process to make a transition from configuration i to another configuration j. The next section further elaborates on the eigenvalues of the SI process.

3.2.1. EIGENVALUES OF THE INFINITESIMAL GENERATOR Q

The 2^N eigenvalues of the $2^N \times 2^N$ infinitesimal generator Q of the SI Markov chain have negative real part and at least one eigenvalue is zero [44]. Since the infinitesimal generator of the SI process is upper triangular as illustrated in Figure 3.2, the eigenvalues of the SI process appear on the diagonal of the infinitesimal generator Q, hence, $\lambda_i = q_{ii} = -\sum_{j=i+1}^{2^N} q_{ij}$ for all $i = 0,1,\ldots,2^N-1$. Prior to describing the physical significance of the eigenvalue λ_i, we introduce some notation.

Each configuration i describes which nodes in the network are infected and which nodes are susceptible. For a given configuration i, we partition the set of nodes $\mathcal{N}$ into

[2] An $N \times N$ matrix A is *normal* if it commutes with its conjugate transpose: $AA^* = A^*A$.
[3] Homogeneous infection rates are $\beta_{kl} = \beta$ for all nodes k,l, such that $B = \beta A$.

two groups. The group $\mathcal{I}_i \subseteq \mathcal{N}$ defines the set of infected nodes in configuration i. Similarly, $\mathcal{S}_i$ contains the susceptible nodes in configuration i. Each node is either infected or susceptible, so $\mathcal{N} = \mathcal{I}_i \cup \mathcal{S}_i$ for all configurations i. We define the *cut set* as the set of links with one node in $\mathcal{S}_i$ and one node in $\mathcal{I}_i$.

The eigenvalue λ_i is equal to the sum over all possible transitions from configuration i to any other configuration j, where configuration j differs exactly one bit (one node) from configuration i. Starting in configuration i, visualised in Figure 3.3, the possible configurations j are those where one susceptible node will be infected by one of the already infected nodes. Hence, the eigenvalue λ_i equals (minus) the sum over all weighted links in the S-I cut set;

$$\lambda_i = -\sum_{k \in \mathcal{I}_i} \sum_{l \in \mathcal{S}_i} \tilde{\beta}_{kl}, \tag{3.4}$$

or, following the notation of [44];

$$\lambda_i = -\sum_{k=1}^{N} \sum_{l=1}^{N} x_k(i)(1 - x_l(i))\tilde{\beta}_{kl}. \tag{3.5}$$

If the underlying graph G has homogeneous link weights and $\tilde{\beta}_{kl} = \beta a_{kl}$, where a_{kl} specifies whether a link exists between node k and l, then λ_i / β *equals (minus) the number of links in the S-I cut set.*

Equation (3.4) provides the exact eigenvalue λ_i for every configuration i. The time-dependent solution (3.3) shows the implication of the eigenvalues λ_i. If the eigenvalue is large (in modulus), the corresponding state will converge exponentially fast to zero. For small (in modulus) eigenvalues, the convergence is much slower. The SI process can be regarded as a discovery process, that starts with a set of discovered nodes and discovers adjacent nodes. If $t \to \infty$ on a connected graph, all nodes will be discovered. Large (in modulus) eigenvalues have a small contribution to the total discovery time, but small (in modulus) eigenvalues will have a significantly larger contribution to the total discovery time in the SI process. In particular, a small (in modulus) eigenvalue λ_i of a configuration i corresponds to a small weight of the S-I cut set (see Figure 3.3) and it takes a long time to transfer the disease between these two groups of nodes. The minimal eigenvalue (equivalent to find the minimum weighted cut in the graph) can be obtained efficiently using the Stoer-Wagner algorithm [65].

Finally, equation (3.4) illustrates that eigenvalue $\lambda = 0$ occurs twice; the all-healthy state (where $\mathcal{I}_i = \emptyset$) and the all-infected state (where $\mathcal{S}_i = \emptyset$) have an empty cut set and the corresponding eigenvalue is zero. Hence, the all-healthy state and the all-infected are both steady states of the SI process. The all-healthy state is unstable, because adding a single infected node will lead to more infected nodes, whereas the all-infected state is stable.

Given a configuration i, then the number of infected nodes k in configuration i can be calculated as

$$k = \sum_{l=1}^{N} x_l(i). \tag{3.6}$$

For several underlying graphs, the eigenvalues can be computed analytically.

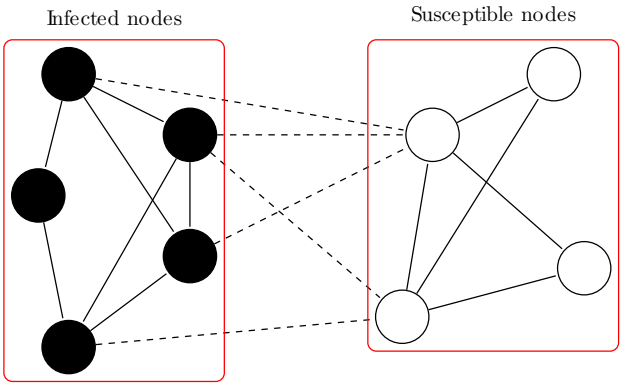


Figure 3.3: The cut set in SI epidemics on a graph with 9 nodes. The eigenvalue λ_i that belongs to configuration i is (minus) the sum over all links in the cut set. The links in the cut set are illustrated by dashed lines.

For the **complete graph** with homogeneous infection rates, the eigenvalues are

$$\lambda_i = -\beta k(N-k), \qquad \text{with multiplicity } \binom{N}{k}$$

where $0 \le k \le N$ is the number of infected nodes in configuration i. The SI process on the complete graph with homogeneous infection rates and starting with a fixed number k of infected nodes is presumably the easiest SI process, because in this case, the SI process can be transformed to a birth and death process, whose time-dependent solution can be calculated exactly [44, Ch. 16]. Starting from k infected nodes and ending with l infected nodes, the total spreading time is distributed as the sum of independent exponential random variables with parameters λ_m, where $k \le m \le l-1$ represents the number of infected nodes [62, 66].

For the **star graph** with homogeneous infection rates, the eigenvalues are

$$\lambda_i = \begin{cases} -\beta(N-1-k), & \text{if hub infected, multiplicity } \binom{N-1}{k} \\ -\beta k, & \text{if hub healthy, multiplicity } \binom{N-1}{k} \end{cases}$$

where $0 \le k \le N-1$ is the number of infected leaf nodes in configuration i. In this case, the eigenvalues can be simplified to the form $\lambda = -\beta k$ with multiplicity $2\binom{N-1}{k}$ for $0 \le k \le N-1$.

For the **cycle graph** with homogeneous infection rates, the eigenvalues are

$$\lambda_i = -2\beta p(k,i),$$

where $p(k,i)$ depends on the number of susceptible neighbours of the k infected nodes in configuration i. Contrary to the complete graph and the star graph, the position of the infected nodes is crucial in the cycle graph. Starting in configuration i with k infected nodes, the k infected nodes can be grouped into $p \le k$ connected, infectious components. The total number of susceptible neighbours is $2p$, because each infected component connects to two susceptible neighbours. After infecting one of the susceptible

neighbours, the number of connected, infected components either stays constant or reduces by one, because two components will merge if the in-between susceptible node has been infected. In that case, the number of components reduces to $p-1$. The number of connected, infected components can therefore only decrease over time. Since the eigenvalue λ_i only decreases (in modulus) over time, the discovery speed of the SI process also slows down over time.

In the **Erdős-Rényi graph** [67] with homogeneous infection rates with weight $\beta = 1$, links exist independent of other links with probability p. The eigenvalue probability distribution $\Pr[\lambda = -l]$ of a random configuration i follows from the law of total probability;

$$\Pr[\lambda = -l] = \sum_{k=0}^{N} \Pr[\lambda = -l \mid \text{state } i \text{ has } k \text{ infected nodes}] \times \Pr[\text{state } i \text{ has } k \text{ infected nodes}].$$

Suppose that configuration i consists of k infected nodes. Then the eigenvalue $\lambda_i = -l$ is (minus) the number of links in the cut set of configuration i. The cut set contains maximally $k(N-k)$ links, where each link exists independently of the other links with probability p. Hence, the probability distribution is a Binomial distribution with $k(N-k)$ possible links and probability p;

$$\Pr[\lambda = -l \mid \text{state } i \text{ has } k \text{ infected nodes}] = \binom{k(N-k)}{l} p^l (1-p)^{k(N-k)-l}.$$

The probability for configuration i to have k infected nodes is equal to the probability for any graph with N nodes to have k infected nodes, which equals $2^{-N}\binom{N}{k}$. Combining all, we find

$$\Pr[\lambda = -l] = \left(\frac{p}{1-p}\right)^l 2^{-N} \sum_{k=0}^{N} \binom{k(N-k)}{l} \binom{N}{k} (1-p)^{k(N-k)}, \tag{3.7}$$

with the convention that $\binom{k(N-k)}{l}$ is zero if $l < 0$ or $l > k(N-k)$. Hence, the eigenvalue λ is bounded between $\lambda = 0$ and $\lambda = -k(N-k)$. We plot the exact solution (3.7) with numerical simulations in Figure 3.4, which shows excellent agreement.

3.2.2. THE EXACT SI SOLUTION

A more explicit solution (3.3) of the SI process requires the computation of the eigenvectors of the infinitesimal generator Q. We denote by $\mathbf{v}_0, \mathbf{v}_1, \ldots \mathbf{v}_{2^N-1}$ and $\mathbf{w}_0, \mathbf{w}_1, \ldots \mathbf{w}_{2^N-1}$ the corresponding right- and left-eigenvectors of Q, respectively. The left-eigenvector $\mathbf{w}_{2^N-1}$, corresponding to the all-infected state $i = 2^N - 1$ with eigenvalue $\lambda_{2^N-1} = 0$, is equal to the steady-state vector π:

$$\mathbf{w}_{2^N-1} = \pi = (0, 0, \ldots, 0, 1)^T. \tag{3.8}$$

The corresponding right-eigenvector is the all-ones vector $\mathbf{v}_{2^N-1} = (1, 1, \ldots, 1)^T$. Since there are no transitions possible to and from the all-healthy state $i = 0$ in the SI process, *we remove the all-healthy state from the Markov chain*, which reduces the Markov chain to $2^N - 1$ states.

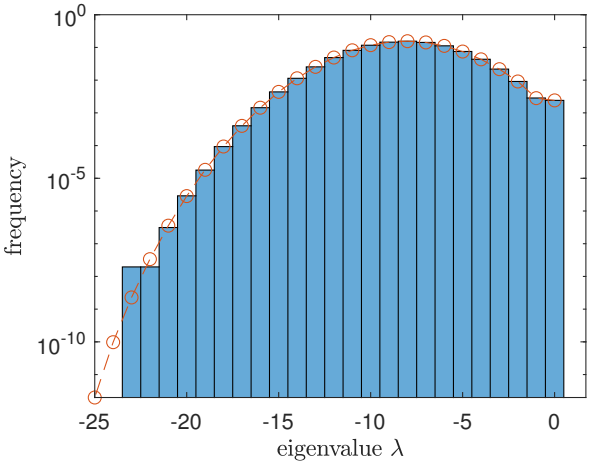


Figure 3.4: The eigenvalue distribution of the infinitesimal generator Q of the SI process in an Erdős-Rényi graph with $N = 10$ nodes and link-existence probability $p = 0.36$ averaged over 100,000 simulations. The exact solution (3.7) is shown as a red dashed line. The simulation results do not follow the exact solution for $\lambda \leq -20$ because the number of simulations is too small compared to the probability of occurrence of these eigenvalues.

The SI process for homogeneous transitions rates was recently solved by Merbis and Lodato [60], thus we focus on heterogeneous infection rates β_{ij}. In many practical applications, the weighted infection rates β_{ij} are real numbers and we can safely assume the following:

Assumption 3.1 *The eigenvalues of the infinitesimal generator Q are distinct.*[4,5,6]

Following Assumption 3.1, the infinitesimal generator Q is diagonalisable and the left- and right eigenvectors from different eigenvalues are orthogonal: $\mathbf{w}_i^T \mathbf{v}_j = \delta_{ij}$ for all $1 \leq i, j \leq 2^N - 1$, where δ indicates the Kronecker delta function [5]. Then the solution $\mathbf{s}(t)$ in Eq. (3.3) simplifies to

$$\mathbf{s}(t) = \pi + \sum_{i=1}^{2^N-2} c_i e^{\lambda_i t} \mathbf{w}_i \tag{3.9}$$

where the constant $c_i = \mathbf{s}(0)^T \mathbf{v}_i$. We remark that (3.9) is equivalent to (2.5) from Chapter 2 for ε-SIS epidemics on the complete graph. Eq. (3.9) does not sum over the all-healthy state $i = 0$, because the all-healthy state cannot be reached. Moreover, any infinitesimal

[4] The eigenvalues may also be degenerate, as long as the algebraic multiplicity equals the geometric multiplicity for all eigenvalues. Then the matrix is diagonalisable and Eq. (3.9) holds. It is, however, unclear when the algebraic and geometric multiplicities are equal for general graphs.

[5] Suppose that the weighted infection rates $\tilde{\beta}_{ij}$ are real and independently distributed, the probability that two eigenvalues are equal is almost surely zero on a finite graph.

[6] All graphs with homogeneous transition rates do not satisfy Assumption 3.1. We know that in homogeneous graphs, the eigenvalue equals the number of links in the cut set. Consider one infected node and all other nodes susceptible, then the size of the cut set exactly equals the infected node's degree. Since every connected graph has at least two nodes with the same degree [5], at least one eigenvalue is degenerate.

generator Q has row sum zero and thus a zero eigenvalue with corresponding all-one right-eigenvector $\mathbf{v}_{2^N-1} = \mathbf{u}$. The corresponding left-eigenvector $\mathbf{w}_{2^N-1}$ is the steady-state vector π, which we placed in front. All left-eigenvectors $\mathbf{w}$ (except $\mathbf{w}_{2^N-1}$) are orthogonal to the right-eigenvector $\mathbf{v}_{2^N-1} = \mathbf{u}$, which implies that each left-eigenvector sums element-wise to zero [5, art. 140–142].

The set of nodes that are initially infected at time $t = 0$ remain infected for all times $t > 0$, because nodes cannot cure. Consider a configuration i in the Markov graph. If one of the initially infected nodes is susceptible in configuration i, then configuration i cannot be reached from the initial state and configuration i can be removed from the Markov graph.

Using Assumption 3.1 and exploiting the upper-triangular structure of the infinitesimal generator Q, we can explicitly compute the eigenvectors. We say that state i can *reach* state j (denoted as $i \sim j$) if there is a directed path in the Markov graph from state i to state j.

Let us first consider the right-eigenvectors $\mathbf{v}_i$. For configurations i that correspond to one infected node, its right-eigenvector is the basis vector $\mathbf{e}_i$. For configurations i with two infected nodes, the right-eigenvector will have non-zero elements at the positions that correspond to all states that can reach state i (including state i itself). Let us consider an example of a graph with $N = 3$ nodes, provided in Figure 3.5.

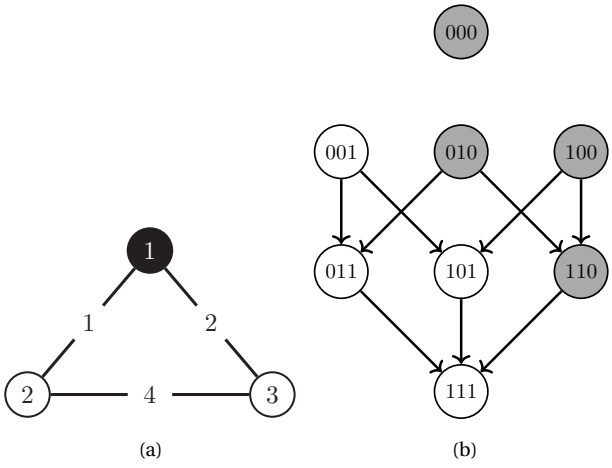


Figure 3.5: (a) Example of a weighted graph with $N = 3$ nodes and (b) the corresponding Markov graph. The grey nodes can be removed from the Markov graph, because the SI process starts with node 1 infected (indicated by the black node) and all grey nodes indicate configurations where node 1 is susceptible.

Example 3.2 *The example graph in Figure 3.5a consists of 3 nodes. Node 1 is initially infected whereas node 2 and 3 are susceptible. The Markov graph in Figure 3.5b is comprised of only four states: (00), (01), (10) and (11), where $(x_3 x_2)$ represent the viral state of nodes 3 and 2, respectively. We have omitted the viral state of node 1, as node 1 is always infected.*

The infinitesimal generator Q equals

$$Q = \begin{pmatrix} q_{00} & q_{01} & q_{02} & 0 \\ 0 & q_{11} & 0 & q_{13} \\ 0 & 0 & q_{22} & q_{23} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where $q_{00} = -3, q_{01} = 1, q_{02} = 2, q_{11} = -6, q_{13} = 6, q_{22} = -5, q_{23} = 5$. For configuration 0=(00) with eigenvalue $\lambda_0 = q_{00} = -3$, the right-eigenvector $\mathbf{v}_0$ equals the basis vector $\mathbf{e}_0 = (1,0,0,0)^T$. For configuration 1 = (01), the eigenvector can be computed using the definition;

$$(Q - q_{11}I)\mathbf{v}_1 = 0 \iff \left(\begin{array}{cccc|c} q_{00}-q_{11} & q_{01} & q_{02} & 0 & 0 \\ 0 & 0 & 0 & q_{13} & 0 \\ 0 & 0 & q_{22}-q_{11} & q_{23} & 0 \\ 0 & 0 & 0 & -q_{11} & 0 \end{array}\right).$$

The last three rows reveal that $(\mathbf{v}_1)_2 = (\mathbf{v}_1)_3 = 0$. Now we choose $(\mathbf{v}_1)_1 = 1$. Then the solution follows as $(\mathbf{v}_1)_0 = \frac{q_{01}}{q_{11}-q_{00}} = -\frac{1}{3}$. We conclude that $\mathbf{v}_1 = (-1/3, 1, 0, 0)^T$ and the only non-zero elements in $\mathbf{v}_1$ correspond to configurations that can reach configuration 1. Analogously, we can find $\mathbf{v}_2 = (-1,0,1,0)^T$. For $\mathbf{v}_3$, we obtain

$$(Q - q_{33}I)\mathbf{v}_3 = 0 \iff \left(\begin{array}{cccc|c} q_{00} & q_{01} & q_{02} & 0 & 0 \\ 0 & q_{11} & 0 & q_{13} & 0 \\ 0 & 0 & q_{22} & q_{23} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

We choose $(\mathbf{v}_3)_3 = 1$. Then the elements $(\mathbf{v}_3)_2 = -\frac{q_{23}}{q_{22}}$ and $(\mathbf{v}_3)_1 = -\frac{q_{13}}{q_{11}}$. The element $(\mathbf{v}_3)_0$ can be computed in the same manner and equals $(\mathbf{v}_3)_0 = -\frac{q_{01}(\mathbf{v}_1)_1 + q_{02}(\mathbf{v}_1)_2}{q_{00}}$. Here, the iterative nature of the construction is clear; $(\mathbf{v}_3)_0$ depends on $(\mathbf{v}_3)_1$ and $(\mathbf{v}_3)_2$. □

In general, the right-eigenvector $\mathbf{v}_i$ of a certain configuration i can be constructed iteratively;

$$\begin{aligned}
&(\mathbf{v}_i)_j = 0, && \text{if state } j \text{ cannot reach state } i \\
&(\mathbf{v}_i)_i = 1, && \text{by construction} \\
&(\mathbf{v}_i)_h = \frac{q_{ih}}{\lambda_i - \lambda_h} = \frac{\sum_{l=1}^{N} \tilde{\beta}_{ml} x_l(h)}{\lambda_i - \lambda_h}, && \text{if } i - h = 2^m, m = 0, 1, \ldots, N-1 \\
&(\mathbf{v}_i)_g = \sum_{h=0}^{2^N-1} \frac{q_{hg}}{\lambda_i - \lambda_g} (\mathbf{v}_i)_h, && \text{if } h - g = 2^n, n = 0, 1, \ldots, N-1 \\
&\quad \vdots
\end{aligned}$$

Configurations h and i differ only at position m, which corresponds to node $m+1$ being infected in configuration i and susceptible in configuration h. Similarly, configuration g and h differ only at position n corresponding to node $n+1$, etc. The iterative formulation consists of at most $N-1$ steps, because at each iteration, one node turns from susceptible into infected, there are N nodes in the graph and we start with at least one infected node. We emphasise that the eigenvector construction is infeasible for degenerate eigenvalues.

The left-eigenvectors $\mathbf{w}_i$ can be derived similarly, which leads to

$$
\begin{aligned}
(\mathbf{w}_i)_h &= 0, && \text{if state } h \text{ cannot be reached, starting from state } i \\
(\mathbf{w}_i)_i &= 1, && \text{by construction} \\
(\mathbf{w}_i)_j &= \frac{q_{ij}}{\lambda_i - \lambda_j} = \frac{\sum_{l=1}^{N} \tilde{\beta}_{ml} x_l(j)}{\lambda_i - \lambda_j}, && \text{if } j - i = 2^m, m = 0, 1, \ldots, N-1 \\
(\mathbf{w}_i)_k &= \sum_{j=0}^{2^N-1} \frac{q_{jk}}{\lambda_i - \lambda_k} (\mathbf{w}_i)_j, && \text{if } k - j = 2^n, n = 0, 1, \ldots, N-1 \\
&\vdots
\end{aligned}
$$

By construction, the eigenvectors are orthonormal $\mathbf{w}_i^T \mathbf{v}_i = 1$ for all configurations i, because the eigenvectors have only one non-zero value in common, which is the 1 at position i. Using the right-eigenvectors $\mathbf{v}_i$ and the left-eigenvectors $\mathbf{w}_i$, the solution (3.9) can be computed explicitly.

3.2.3. COMPUTATIONAL FEASIBILITY

Irrespective of the application domain, the quantity of interest of the SI process should be computable in a reasonable time. The $2^N \times 2^N$ infinitesimal generator Q cannot be computed nor stored if the number of nodes $N \geq 25$. Fortunately, the exact solution $\mathbf{s}(t)$ under Assumption 3.1 can be computed without explicitly constructing the infinitesimal generator Q.

Only the eigenvalues and eigenvectors of the infinitesimal generator Q are required in the solution (3.9). For a given configuration i, the eigenvalue λ_i can be computed based on equation (3.5) in $\mathcal{O}(N^2)$ operations. To compute all 2^N eigenvalues, we need $\mathcal{O}(N^2 2^N)$ operations. The iterative procedure to compute a single left- and right-eigenvector takes $\mathcal{O}(N2^N)$ operations if the eigenvalues are known. Since there are 2^N eigenvectors, the total number of required operations to compute all eigenvectors is $\mathcal{O}(N4^N)$. Hence, for networks with $N > 20$ nodes, computing all eigenvalues and eigenvectors remains infeasible.

Even though the computation of all eigenvalues and eigenvectors is infeasible for large networks, for certain quantities of interest, not all eigenvalues or eigenvectors are required. Here, we consider two exemplary scenarios. First, we consider the probability $\Pr[X_2(t) = 1]$ that node 2 is infected at time t. To compute $\Pr[X_2(t) = 1]$, we sum over all states in $\mathbf{S}$ in which node 2 is infected. We introduce the $2^N \times 1$ vector $\mathbf{m} = (0, 1, 1, 0, 0, 1, \ldots, 1)^T$, which is 1 at position j if the binary representation of configuration j has a 1 at position 2, and zero otherwise. Then $\Pr[X_2(t) = 1] = \mathbf{m}^T s(t)$ and

using (3.9), we find

$$\Pr[X_2(t)=1] = \mathbf{m}^T\pi + \sum_{i=1}^{2^N-2} c_i e^{\lambda_i t}\mathbf{m}^T\mathbf{w}_i. \tag{3.10}$$

The inner product $\mathbf{m}^T\pi = 1$, because node 2 will be infected after infinitely long time. We know that each left-eigenvector $\mathbf{w}_i$ sums element-wise to zero. Then for all configurations i for which the non-zero elements of the left-eigenvector $\mathbf{w}_i$ overlap with the non-zero elements of $\mathbf{m}$, the inner product is zero. Only the states i that can be reached from state 2 will have a zero inner product, which are 2^{N-1} states. Since c_i is non-zero for 2^{N-1} states, we conclude that the sum in (3.10) can be simplified from 2^N-2 states to 2^{N-2} states. Unfortunately, summing over 2^{N-2} states is still exponentially large and solving for large networks remains impossible.

As a second example, we consider the probability that all nodes are infected at time t. We multiply the solution $\mathbf{s}(t)$ with the vector $\mathbf{m} = (0,0,\ldots,0,1)^T$, which is the all-zeros vector, except the last element is one. Regarding the product $m^T\mathbf{w}_i$ in (3.10), we conclude that only the last element $(\mathbf{w}_i)_{2^N-1}$ is of importance. Unfortunately, computing the last element of the left-eigenvector $\mathbf{w}_i$ is as difficult as computing the whole vector, because of the iterative construction of the eigenvector $\mathbf{w}_i$.

The provided examples consider two limit cases, namely the infection probability of a single node and all nodes, respectively. For both examples, exponentially many eigenvalues and eigenvectors are required to build up the solution (3.9). The examples illustrate that calculating the exact, time-varying solution of the infection probability in SI epidemics on large, arbitrary graphs, is generally infeasible.

Fortunately, as we concluded earlier, the large (in modulus) eigenvalues have a small effect on the dynamics at large times. Presumably, the exact solution (3.9) can be accurately approximated using a truncation method. The number of modes is truncated at index m, such that the approximated solution $\tilde{\mathbf{s}}$ equals

$$\tilde{\mathbf{s}}(t) \approx \pi + \sum_{i=1}^{m} c_i e^{\lambda_i t}\mathbf{w}_i. \tag{3.11}$$

The approximation of Eq. (3.9) by Eq. (3.11) introduces the error

$$e(t) = \|\mathbf{s}(t) - \tilde{s}(t)\| = \left\| \sum_{i=m+1}^{2^{N-1}} c_i e^{\lambda_i t}\mathbf{w}_i \right\| \leq \sum_{i=k+1}^{2^{N-1}} e^{\lambda_i t}\|c_i\mathbf{w}_i\|,$$

where $\|\cdot\|$ is a vector norm and we used the triangle inequality $\|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$. Using the fact that the eigenvectors are normalised, we find

$$e(t) \leq \sum_{i=m+1}^{2^{N-1}} e^{\lambda_i t}\|c_i\mathbf{w}_i\| < e^{\lambda_{m+1}t} \sum_{i=m+1}^{2^{N-1}} |c_i|,$$

such that the error scales as $\mathcal{O}(e^{\lambda_{m+1}t})$. Similarly to ε-SIS dynamics (see Appendix A.4), the truncation method performs poor for small times, see Figure 3.6. Only a fraction of the eigenmodes does not accurately describe the dynamics. The main reason is as follows. Suppose the SI epidemic starts at a single node that has only one link with a

very small weight. Under the truncation approximation, the eigenmode corresponding to the infection from the seed node to the second node will be disregarded, causing a discrepancy between the true solution and the truncation approximation.

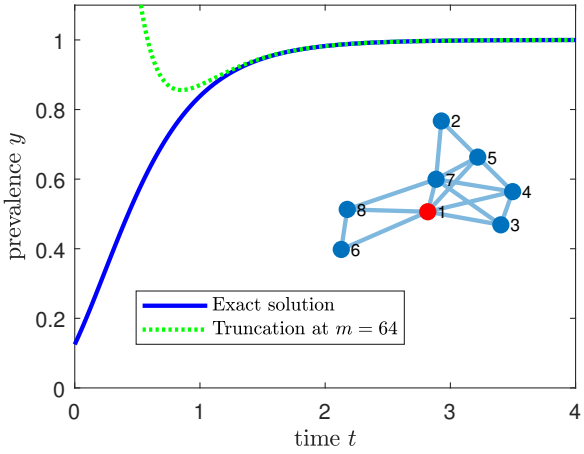


Figure 3.6: The exact solution (solid line) and the truncated solution (dotted line) with $m = N^2$ out of 2^N states of the Markovian SI process on an Erdős-Rényi graph with $N = 8$ nodes. The infection rates β_{ij} are chosen uniformly at random between $\beta_{\min} = 0.9$ and $\beta_{\max} = 1.1$. Initially, node 1 is infected.

3.2.4. NUMERICAL SIMULATIONS

We illustrate the accuracy of our exact solution method in Figure 3.7 for two graphs: an Erdős-Rényi graph with $N = 10$ nodes and link-connectivity $p = 0.33$ and a cycle graph with $N = 12$ nodes. The simulations start with node 1 infected. The results are averaged over 10,000 Monte Carlo simulations. We compare our exact solution and the Monte-Carlo simulations with the N-Intertwined Mean-Field Approximation (NIMFA), which assumes every pair of random variables is uncorrelated [29]. The exact solution coincides with the Monte Carlo simulations, whereas the mean-field approximation vastly overestimates the time-dependent prevalence [68].

3.2.5. EXTENSIONS OF MARKOVIAN SI EPIDEMICS

For Markovian SI processes on heterogeneous networks, we have derived the explicit solution (3.9). Here, we extend our results on the SI process to non-Markovian epidemics, temporal networks, simplicial contagion and the inclusion of self-infections. Additionally, we briefly discuss the difficulties of computing the exact solution for general compartmental models.

NON-MARKOVIAN DYNAMICS

Most epidemic models assume a memoryless Markov process, i.e. the probability to infect a neighbour is exponentially distributed over time. Non-Markovian effects in Markov

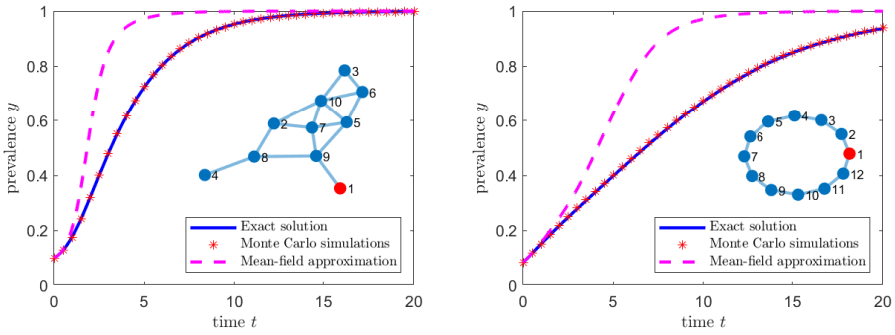


Figure 3.7: The exact solution (solid line), Monte Carlo simulations (asterisks) and the mean-field approximation (dashed line) of the Markovian SI process on (a) an Erdős-Rényi graph with $N = 10$ nodes and $p = 0.33$ and (b) a cycle graph with $N = 12$ nodes. The infection rates β_{ij} are chosen uniformly at random between $\beta_{\min} = 0.1$ and $\beta_{\max} = 1.1$. Initially, node 1 is infected. The mean-field approximation deviates in both cases significantly from the exact solution.

chains can be taken into account using fractional calculus, as was recently derived by Van Mieghem [69]. The governing equations (3.2) of the Markov chain change to

$$D^{\alpha}\mathbf{s}_{\alpha}^{T}(t) = \mathbf{s}_{\alpha}^{T}(t)Q^{\alpha}, \tag{3.12}$$

where D^{α} is the Caputo fractional derivative and $0 < \alpha < 1$. As derived by Van Mieghem [69], the solution (3.9) becomes

$$\mathbf{s}(t) = \pi + \sum_{i=1}^{2^N-1} c_i E_{\alpha,1}\left((\lambda_i t)^{\alpha}\right)\mathbf{w}_i \tag{3.13}$$

where $E_{\alpha,1}(z)$ is the Mittag-Leffler function [70] of the complex number z. Up to our knowledge, Eq. (3.13) is the first exact solution of an epidemic, non-Markovian process on an heterogeneous network. The exact solution (3.13) is plotted in Figure 3.8 for $0.5 \leq \alpha \leq 1$. Decreasing α increases the prevalence at short times but slows down the disease spread for large times. Compared to the Markovian case $\alpha = 1$, decreasing α increases the variance of the infection time distribution, causing a large probability of both very slow and very fast infections. At short times, this results in a larger infection probability, but it also takes much longer for all nodes to become infected.

TEMPORAL NETWORKS

So far, we have assumed that the network structure is fixed. In reality, networks are varying over time due to movements of individuals. Given a temporal graph G that changes finitely many times in the interval $[0, T]$, each graph G_n in an interval $[t_n, t_{n+1}]$ is represented by the adjacency matrix $A^{(n)}$. We assume that the graph remains connected at all times. Then we may compute the eigenvectors and eigenvalues for each of those

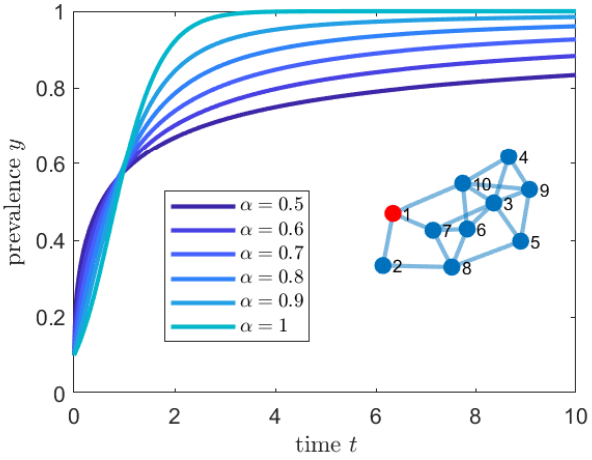


Figure 3.8: The exact solution of the non-Markovian SI process with Caputo fractional derivative α on an Erdős-Rényi graph with $N = 10$ nodes. The infection rates β_{ij} are chosen uniformly at random between $\beta_{\min} = 0.9$ and $\beta_{\max} = 1.1$. Initially, node 1 is infected. The case $\alpha = 1$ coincides with the Markovian case.

adjacency matrices $A^{(n)}$ and the solution becomes

$$\mathbf{s}(t) = \pi + \sum_{i=1}^{2^N-1} c_i^{(n)} e^{\lambda_i^{(n)} t} \mathbf{w}_i^{(n)}, \qquad t_n \le t \le t_{n+1} \tag{3.14}$$

where $c_i^{(n)} = \mathbf{s}(t_n)\mathbf{v}_i^{(n)}$. Within each interval $t_n \le t \le t_{n+1}$, the dynamics are equivalent to the case of the fixed network. Going to the next time interval requires different eigenvalues and eigenvectors, and additionally, the initial condition of the new interval must equal the final state of the previous interval. Although the eigenvectors and eigenvalues may be different for each interval, the steady-state vector π is always the all-ones vector, representing the state in which all nodes are infected, because the graph remains connected for all times.

SIMPLICIAL CONTAGION

In the study of opinion dynamics on networks, it was observed that multiple neighbours of a node may be required to persuade a node to adopt a particular strategy or opinion. Such higher-order spreading phenomena on networks are known as simplicial contagion [71], which was first applied to the SIS model by Iacopini *et al.* [72]. The conceptual idea is that, besides node-node interactions, connected infected triangles may increase the probability of converting a particular node to the other opinion, even larger than the sum of the individual infection rates. Adding simplicial contagion to the SI process does not alter the *structure* of the infinitesimal generator Q, because simplicial infections require the existence of triangles in the original network structure, whose individual links already had contributions to the infinitesimal generator. Instead, the *link weights* corresponding to triangles are increased for certain configurations in the infinitesimal generator Q. The eigenvalues are no longer the weighted sum over all links in the $S-I$ cut set,

but additionally include the weighted sum over all 2-infected-1-susceptible-node triangles in the graph. The eigenvectors can be computed in the usual way, such that the time-dependent solution (3.9) can be recovered.

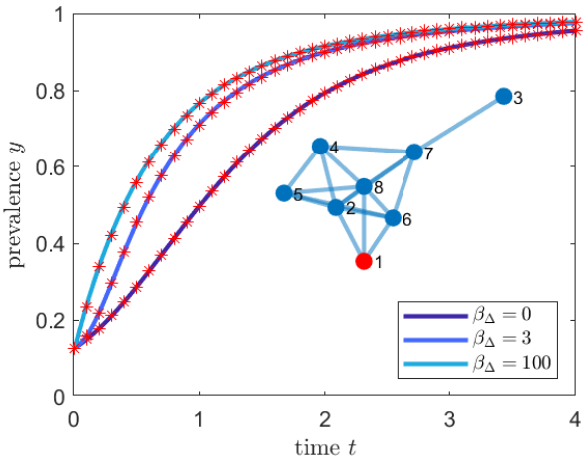


Figure 3.9: The exact solution of the SI process with simplicial contagion on an Erdős-Rényi graph with $N = 8$ nodes. The infection rates β_{ij} are chosen uniformly at random between $\beta_{\min} = 0.2$ and $\beta_{\max} = 1.0$ and the triangle interaction rate β_Δ is taken equal to 0,3 and 100. Initially, node 1 is infected. The Monte Carlo simulations are shown as red asterisks.

Figure 3.9 illustrates the impact of simplicial contagion on the classical SI process. Adding the homogeneous simplicial infection rate β_Δ increases the prevalence y, but only up to a certain point, because some parts of the graph do not contain any triangles and therefore nodes cannot be infected by higher-order simplexes.

SELF-INFECTIONS

Contrary to the standard SI process, where the spread or discovery of items is solely related to the underlying graph, some spreading phenomena can be triggered by external processes, which are unrelated to the spread on the graph. For example, individuals or companies may adopt a certain product or technology without any interference with others. The situation in which a node becomes infected spontaneously without inference with other nodes, is described as a self-infection with rate ε. The Bass model describes the spread of direct infections and self-infections of new products used in companies [12] and is actually equivalent to the SI process with self-infections [16].

The infinitesimal generator Q of the ε-SI process equals

$$q_{ij} = \varepsilon_m + \sum_{k=1}^{N} \tilde{\beta}_{mk} x_k(i), \qquad \text{if } j = i + 2^{m-1} \tag{3.15}$$

$$\text{with } m = 1, 2, \ldots, N \text{ and } x_m(i) = 0.$$

$$q_{ii} = -\sum_{\substack{k=0 \\ k \neq i}}^{2^N - 1} q_{ki}. \tag{3.16}$$

Thus, adding self-infections can both increase the link weight and add more links in the Markov graph. The resulting infinitesimal generator Q will remain upper-triangular, because the total number of infected nodes can only increase. In this case, the eigenvalues λ_i describe the weighted sum of the links in the S-I cut set plus the total (possibly weighted) sum of the self-infection rate of all susceptible nodes in configuration i.

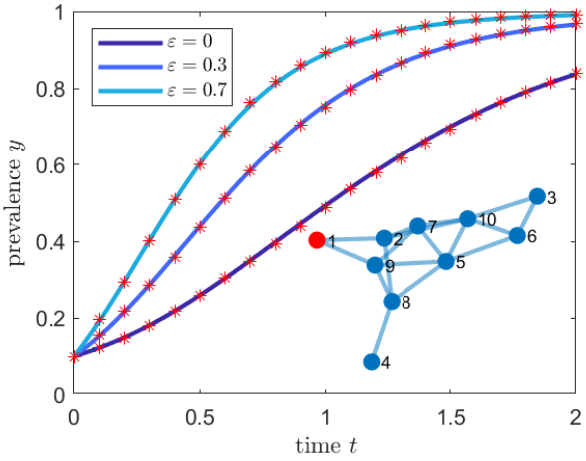


Figure 3.10: The exact solution of the Markovian SI process with self-infections on an Erdős-Rényi graph with $N = 10$ nodes. The infection rates β_{ij} are chosen uniformly at random between $\beta_{\min} = 0.9$ and $\beta_{\max} = 1.1$. The self-infection rate ε is taken equal to 0, 0.3 and 0.7. Initially, node 1 is infected.

Adding the self-infection process to the SI process increases the prevalence significantly, as is shown in Figure 3.10. In the limit where the self-infection process is much larger than the infection process, each node adopts the innovation independent of all other nodes with an exponential distribution and the total prevalence is just the sum of N independent exponential random variables.

3.3. THE SIR PROCESS

Contrary to the SI process, whose applicability is limited, network-based SIR epidemics has a very rich and long history in modelling network epidemiology [6]. The SIR process is a transient process, because after being infected, nodes will recover and can never be re-infected.

We solve the SIR process on a heterogeneous network by proposing a labelling of the state space based on trinary numerals. We introduce x_k as the viral state of node k:

$$\begin{aligned} x_k &= 0, && \text{if node } k \text{ is susceptible} \\ x_k &= 1, && \text{if node } k \text{ is infected} \\ x_k &= 2. && \text{if node } k \text{ is recovered} \end{aligned}$$

The viral state vector $\mathbf{x} = (x_N, x_{N-1}, \ldots, x_1)^T$ describes the viral state of all nodes. Given a particular vector $\mathbf{x}$, we may compute the configuration number using the trinary numerals;

$$i = \sum_{k=1}^{N} x_k(i) \cdot 3^{k-1}.$$

The trinary numbering ensures that each viral state vector $\mathbf{x}$ corresponds to a different configuration number i, which ranges between $i = 0$ and $i = 3^N - 1$. We further introduce the variable $z_k(i)$, which equals one if node k is infected in configuration i and is zero otherwise. Then the infinitesimal generator Q of the SIR process equals

$$\begin{aligned} q_{ij} &= \sum_{k=1}^{N} \beta_{mk} z_k(i), && \text{if } j = i + 3^{m-1} \\ &&& \text{with } m = 1, 2, \ldots, N \text{ and } x_m(i) = 0 \\ q_{ij} &= \delta_m z_m(i), && \text{if } j = i + 3^{m-1} \\ &&& \text{with } m = 1, 2, \ldots, N \text{ and } x_m(i) = 1 \\ q_{ii} &= -\sum_{\substack{j=0 \\ j \neq i}}^{3^N - 1} q_{ji}, \end{aligned}$$

where δ_m is the curing rate of node m. The Markov graph with 3^N states is the graph corresponding to the infinitesimal generator Q and specifies all possible transitions. The Markov graph for a complete graph with $N = 3$ nodes is shown in Figure 3.11. The absorbing states are indicated by shaded circles. We observe that the Markov graph is a directed tree, i.e. a bipartite graph. Moreover, we can define the *layer number* $l = I + 2R$, where I represents the number of infected nodes and R the number of recovered nodes in each configuration. Transitions can only take place from states in layer l to states in layer $l + 1$. Layers with odd number l do not contain any absorbing states, because an odd layer number l implies that at least one node must be infected. If a configuration contains one infected node, that node can always recover, ruling out the existence of absorbing states in that layer.

Figure 3.12 shows all non-zero elements of the infinitesimal generator Q with $N = 4$ nodes. Some rows of Q are empty – these rows correspond to absorbing states in the

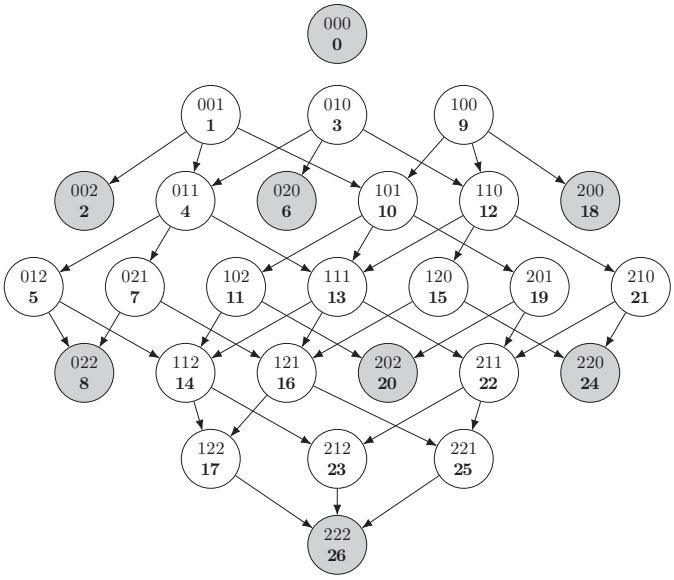


Figure 3.11: The Markov graph of the SIR process on the complete graph with $N = 3$ nodes. Each state $(x_3 x_2 x_1)$ encodes the configuration (viral state) of the nodes; susceptible (0), infected (1) and recovered (2). The configuration number i is shown in bold.

SIR process (the grey nodes in Figure 3.11). Although the infinitesimal generator Q has dimensions $3^N \times 3^N$, the total number of transitions is much lower. For a given configuration i, the number of possible transitions is given by N, because each node can change its viral state to at most one other state. Taking into account the possibility of not making any change, the number of non-zero elements in the infinitesimal generator Q is upper bounded by $(N+1)3^N$, which is much less than a dense matrix structure with 3^{2N} elements.

3.3.1. SIR EIGENVALUES

The eigenvalue λ_i of the infinitesimal generator Q that corresponds to configuration i equals

$$\lambda_i = -\sum_{k=1}^{N}\sum_{l=1}^{N} \beta_{kl} z_k(i)(1 - z_l(i)) - \sum_{k=1}^{N} \delta_k z_k(i). \tag{3.17}$$

The first component specifies the sum over the weighted links in the S-I cut set as in (3.5), whereas the second part contains the total weighted curing rate of the infected nodes in configuration i. Each configuration i that does not contain any infected nodes (i.e. $z_k = 0$ for all k) is an absorbing state and cannot be left. The total number of states without infected nodes, thus only containing susceptible and recovered nodes, equals 2^N. Hence, there are 2^N absorbing states. The eigenvalue of an absorbing state equals $\lambda = 0$ and eigenvalue $\lambda = 0$ has algebraic multiplicity 2^N. To determine the geometric multiplicity of eigenvalue $\lambda = 0$, i.e. solving $Q\mathbf{v} = \mathbf{0}$, we investigate the structure of the

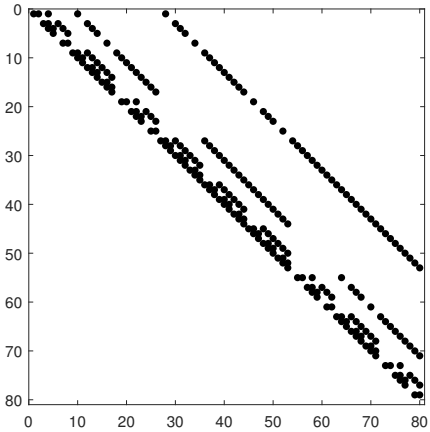


Figure 3.12: The structure of the infinitesimal generator Q for the SIR process on a complete graph with $N = 4$ nodes.

infinitesimal generator Q in Figure 3.12. There are 2^N rows with only zeros, allowing for 2^N free variables in the right-eigenvector $\mathbf{v}$. Hence, the algebraic multiplicity is equal to the geometric multiplicity for $\lambda = 0$ and moreover, the eigenvectors with eigenvalue zero can be chosen orthogonally.

It remains to verify the uniqueness of the non-zero eigenvalues. Similar as for SI epidemics, we require that all infection rates are different. For SIR epidemics, we additionally require that the underlying graph is complete. Consider the non-complete graph G from Figure 3.13 with a missing link between the orange and black node. Suppose the black node is infected and the orange node is either susceptible or recovered. All other nodes are considered susceptible (represented by the cloud). The eigenvalue λ in (3.17) equals the SI cut set plus the curing rate of the black node. The fact that the orange node is susceptible or recovered does not influence the eigenvalue λ, because the SI cut set and the sum over the curing rates are equivalent in both cases. Thus, any non-complete graph has one or more degenerate non-zero eigenvalues.

From simulations it appears that even in the case of missing links (and subsequently degenerate eigenvalues), the corresponding eigenspace is of full rank and orthogonal eigenvectors can be found. Explicit constructions for those eigenvectors are cumbersome and are omitted here. Instead, we can set infection rates $\beta_{kl} > 0$ arbitrarily small for every non-existing link. The resulting graph is the complete graph with heterogeneous infection rates, which can accurately approximate any realistic situation at the cost of a small error and avoids the computation of the complicated case of degenerate eigenvalues.

3.3.2. SIR EIGENVECTORS

In a similar fashion as for the SI process in Section 3.2.2, the right- and left-eigenvectors $\mathbf{v}_i$ and $\mathbf{w}_i$ of the SIR process can be constructed. We focus on the right-eigenvector $\mathbf{v}_i$,

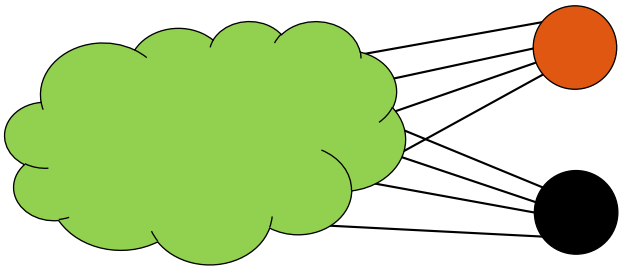


Figure 3.13: The graph G is represented by a green cloud containing susceptible nodes, an infected black node and the orange node. The link between the black and orange node is missing. Whether the orange node is susceptible or recovered does not influence the computation of the eigenvalue λ from Eq. (3.17), thus the eigenvalue λ is degenerate.

but the left-eigenvector $\mathbf{w}_i$ can be constructed similarly.

For a configuration i that corresponds to one infected node, the right-eigenvector $\mathbf{v}_i$ is the basis vector $\mathbf{e}_i$. For a configuration i with two infected nodes, the right-eigenvector $\mathbf{v}_i$ will have non-zero elements at the positions that correspond to all states that can reach state i (including state i itself). In fact, we follow the same procedure as for the SI process. The only exception occurs when $\lambda_i = 0$, which is a degenerate eigenvalue. Fortunately, for $\lambda = 0$ the algebraic and geometric multiplicity are equal, such that the eigenspace is of full rank and we may choose the eigenvectors of the zero eigenvalues as standard basis vectors: for configuration i with eigenvalue $\lambda_i = 0$, we choose $\mathbf{v}_i = \mathbf{e}_i$.

The right-eigenvector $\mathbf{v}_i$ of a certain configuration i can be constructed iteratively;

If $\lambda_\mathbf{i} = 0$; then $\mathbf{v}_i = \mathbf{e}_i$

If $\lambda_\mathbf{i} \neq 0$; then

$$(\mathbf{v}_i)_j = 0, \qquad \text{if state } j \text{ cannot reach state } i$$

$$(\mathbf{v}_i)_i = 1, \qquad \text{by construction}$$

$$(\mathbf{v}_i)_h = \frac{q_{ih}}{\lambda_i - \lambda_h} = \frac{\sum_{l=1}^{N} \tilde{\beta}_{ml} z_l(h) \cdot \mathbf{1}_{\{\mathbf{x}_m(h)=0\}} + \delta_m z_m(h)}{\lambda_i - \lambda_h}, \qquad \text{if } i - h = 3^m, m = 0, 1, \ldots, N-1$$

$$(\mathbf{v}_i)_g = \sum_{h=0}^{3^N-1} \frac{q_{hg}}{\lambda_i - \lambda_g} (\mathbf{v}_i)_h, \qquad \text{if } h - g = 3^n, n = 0, 1, \ldots, N-1$$

$\ldots$

Configurations h and i differ only at position m (corresponding to node $m+1$), where node $m+1$ is infected in configuration i and susceptible in configuration h or node $m+1$ is recovered in i and infected in h. Similarly, configuration g and h differ only at position n, corresponding to node n, etc.

3.3.3. SIR SOLUTION

After the derivation of the orthonormal eigenvectors for the zero and non-zero eigenvalues, the infinitesimal generator Q is diagonalisable and the time-dependent solution $\mathbf{s}(t)$ becomes

$$\mathbf{s}(t) = \sum_{i=0}^{3^N-1} c_i e^{\lambda_i t} \mathbf{w}_i, \tag{3.18}$$

where $c_i = \mathbf{s}(0)^T \mathbf{v}_i$ and $\mathbf{v}_i, \mathbf{w}_i$ are the right and left-eigenvector of Q, respectively. Defining the $3^N \times 1$ vector $\mathbf{m}$, whose elements m_i equal the number of infected nodes in configuration i, the time-varying prevalence $y(t)$ follows as

$$y(t) = \frac{1}{N} \sum_{i=0}^{3^N-1} c_i e^{\lambda_i t} \mathbf{m}^T \mathbf{w}_i,$$

which can be simplified to

$$y(t) = \sum_{i=0}^{2^N-1} \tilde{c}_i + \sum_{i=0}^{3^N-2^N-1} \tilde{c}_i e^{\lambda_i t}, \tag{3.19}$$

where $\tilde{c}_i = \frac{1}{N} c_i \mathbf{m}^T \mathbf{w}_i$. The first term in Eq. (3.19) contains all zero eigenvalues and the second term contains all non-zero (negative) eigenvalues. Surprisingly, the time-dependent prevalence $y(t)$ in (3.19) is just a function of the eigenvalues λ_i and the (complicated) variables $\tilde{c}_i$.

We can further simplify (3.19). The contribution of all non-zero eigenvalues in the second term of (3.19) converges exponentially fast to zero for $t \to \infty$, whereas the first term in (3.19) corresponding to the zero eigenvalues, is fixed. Since the prevalence $y \to 0$ for $t \to \infty$, it follows that the first term must be zero. Hence, the solution (3.19) reduces to

$$y(t) = \sum_{i=0}^{3^N-2^N-1} \tilde{c}_i e^{\lambda_i t}. \tag{3.20}$$

Figure 3.14 shows the exact solution (3.20) and Monte-Carlo simulations, which match perfectly. For SI epidemics, we considered non-Markovian dynamics using fractional derivatives, included time-varying contact networks, added higher-order simplicial contagion and added self-infections. The same extensions can straightforwardly be applied to SIR epidemics.

3.3.4. EPIDEMIC PEAK TIME IN SIR

A key property of SIR epidemics is the *epidemic peak time*, which is the time at which on average most nodes are infected. Knowing when the number of infections reaches the peak allows decision-makers to employ timely countermeasures. If the infection rate in the SIR process is above the epidemic threshold, then an outbreak occurs with large probability. After potentially a long time, the disease dies out, exhibiting a peak in the number of cases at the epidemic peak time. Due to the heterogeneous infection and curing rates, multiple local extrema may be observed [73].

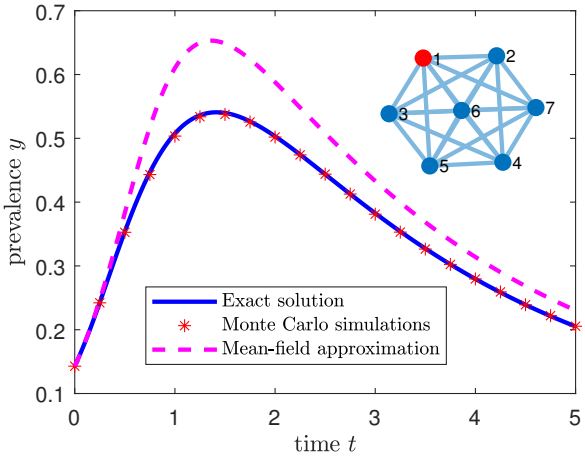


Figure 3.14: The exact solution of the Markovian SIR process on a complete graph with $N = 7$ nodes. The infection rates β_{ij} are chosen uniformly at random between $\beta_{\min} = 0.2$ and $\beta_{\max} = 1.0$ and the curing rates δ_i are chosen uniformly between $\delta_{\min} = 0.1$ and $\delta_{\max} = 0.6$. Initially, node 1 is infected.

The epidemic peak can be derived explicitly for the Markovian SIR process. The derivative of $y(t)$ in (3.20) equals

$$y'(t) = \sum_{i=0}^{3^N-2^N-1} \tilde{c}_i \lambda_i e^{\lambda_i t}. \tag{3.21}$$

The epidemic peak time t_{peak} obeys $y'(t_{\text{peak}}) = 0$. We determine the epidemic peak time t_{peak} using the Newton-Raphson and second-order Newton-Raphson method (see Appendix B.1 for the derivation). Methods like Newton-Raphson are ultimately suitable for the determination of the epidemic peak time using our solution, since the derivatives of $y(t)$ can be straightforwardly computed with high precision.

Starting with initial guess $\tilde{t}_0$, the Newton-Raphson method finds a root t^* of $y'(t)$ iteratively as

$$\tilde{t}_{k+1} = \tilde{t}_k - \frac{y'(\tilde{t}_k)}{y''(\tilde{t}_k)}, \tag{3.22}$$

and converges, provided that the initial point $\tilde{t}_0$ is sufficiently close, to the root t^*. Performing Newton-Raphson on Eq. (3.21) with starting point $\tilde{t}_0 = 0$ yields for $k \geq 0$:

$$\tilde{t}_{k+1} = \tilde{t}_k - \frac{\sum_{i=0}^{3^N-2^N-1} \tilde{c}_i \lambda_i e^{\lambda_i \tilde{t}_k}}{\sum_{i=0}^{3^N-2^N-1} \tilde{c}_i \lambda_i^2 e^{\lambda_i \tilde{t}_k}}.$$

The Newton-Raphson method converges quadratically to a local root, which is not necessarily equal to the global maximum or minimum. Similarly, the second-order Newton-

Raphson method is given by (see Appendix B.1)

$$\tilde{t}_{k+1} = \tilde{t}_k + \frac{-y''(\tilde{t}_k) \pm \sqrt{y''(\tilde{t}_k)^2 - 2y'(\tilde{t}_k)y'''(\tilde{t}_k)}}{y'''(\tilde{t}_k)}. \tag{3.23}$$

We verify the accuracy of both approaches by comparing the peak time estimate with the true epidemic peak time t_{peak} as follows: On the complete graph with $N = 7$ nodes, we repeat 100 times: Draw uniformly distributed $\beta_{ij} \in [0.05, 0.25]$ and $\delta_i \in [0.3, 0.5]$ and compute for each set of parameters the exact peak time using Eq. (3.21) and the bisection method. Figure 3.15a shows the time-varying prevalences $y(t)$ for each of the 100 trials. The epidemic peak is approximately located at $t_{\text{peak}} \approx 2$. Figure 3.15b shows the absolute difference between the real peak time t_{peak} and the peak time estimated by the Newton-Raphson method for $k = 0, \dots, 5$ iterations. For almost all trials, unfortunately, the error did not converge to zero. Contrary, the error of the second-order Newton-Raphson (2nd-NR) method in Figure 3.15c rapidly converges to zero for almost[7] all trials. Due to the cubic convergence of 2nd-NR, only three iterations suffice to accurately approximate the true epidemic peak time t_{peak}.

There are multiple reasons why the Newton-Raphson method does not converge to the true epidemic peak time t_{peak}. For example, for trials that satisfy $y'(0) > 0$ and $y''(0) > 0$ and starting at $\tilde{t}_0 = 0$, Eq. (3.22) tells us that $\tilde{t}_1 < 0$. Subsequent estimates $\tilde{t}_k$ for $k \geq 1$ will converge to even smaller (negative) values. For other trials, it holds that $y''(0) > 0$ but $y''(0)$ is very small. The Newton-Raphson estimate $\tilde{t}_1$ is then so large, that subsequent estimates $\tilde{t}_k$ will diverge to $+\infty$, as $t^* = \infty$ is also a valid solution of $y'(t^*) = 0$. Additionally, the prevalence $y(t)$ can be non-monotonic and exhibits multiple peaks. Figure 3.16 shows that such an effect can already appear in small graphs (although it is not very pronounced).

Figure 3.16 additionally demonstrates that the mean-field approximation performs very poorly. It is known that mean-field approximations perform poorly on small graphs [74], but allowing for heterogeneous transition rates instead of homogeneous transition rates leads to even worse mean-field estimates.

3.3.5. PROBABILITY OF k INFECTED NODES

Several properties of the SIR process can be derived from the time-dependent solution $\mathbf{s}(t)$. For example, the probability that k nodes are simultaneously infected at time t additionally characterises an epidemic outbreak and can be used to assess the maximum impact of the disease on the population. Given the solution $\mathbf{s}(t)$, the probability of k infected nodes can be computed by projecting on the vector $\mathbf{m}$, whose elements $m_i = 1$ if configuration i has k infected nodes. In practice, we count the number of ones in $\mathbf{x}_i$ to compute m_i. Figure 3.17 shows the probability of k infected nodes in a graph with 7 nodes. At time $t = 0$, the probability of 1 infected node is 1, whereafter it converges quickly to smaller values. In this example, the prevalence is rather high, implying that many nodes can be infected simultaneously. Figure 3.17 illustrates the probability that all nodes are infected simultaneously is 0.1 at its maximum value, which is very high.

[7] It might happen that divergence occurs if the third derivative $y'''(t)$ is almost zero, in which case a higher-order Newton-Raphson should be implemented.

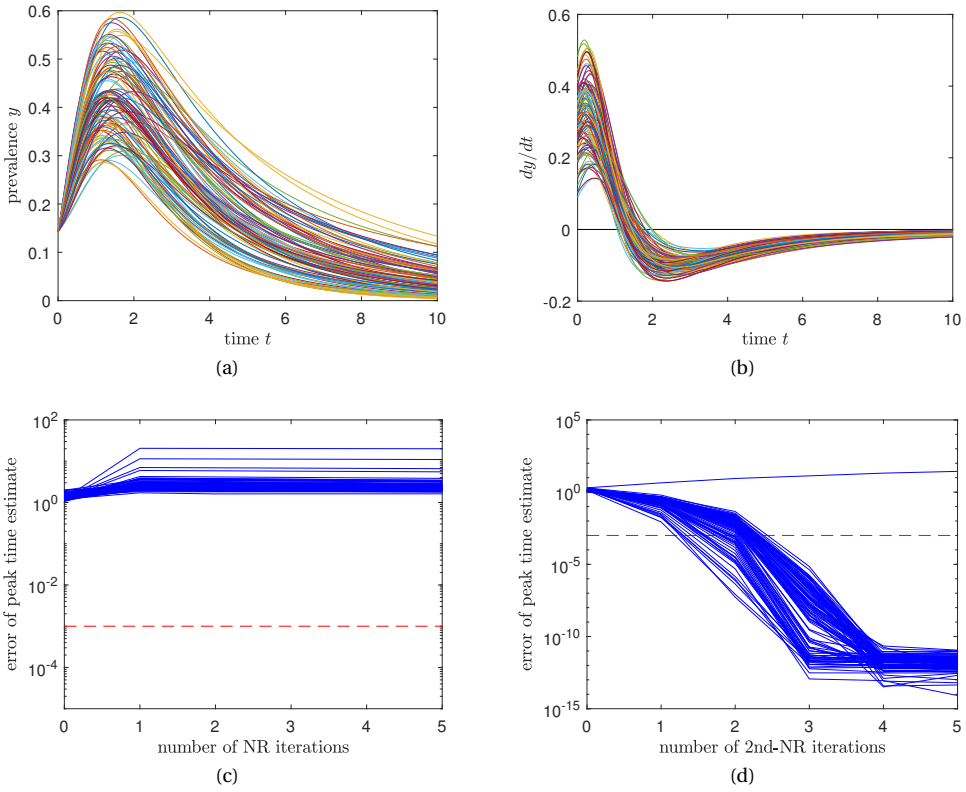


Figure 3.15: (a) The time-varying prevalence $y(t)$ in SIR epidemics on the complete graph with $N = 7$ nodes for 100 trials. Each trial randomly generated uniform random variables $\beta_{ij} \in [0.05, 0.85]$ and $\delta_i \in [0.1, 0.9]$. (b) The derivative of the prevalence y with respect to time t for each of the 100 trials. (c-d) The error of finding the true epidemic peak using the (c) Newton-Raphson method and the (d) second-order Newton-Raphson method with 0–5 iterations. Initially, $\tilde{t}_0 = 0$. The minimal error for finding the peak time is 10^{-10}, because the "exact" peak is also determined numerically using the bisection method. Newton-Raphson in (c) does not converge, whereas second-order Newton-Raphson in (d) almost always converges quickly.

3.3.6. PROBABILITY OF GROUP-LEVEL INFECTIONS

Many stochastic or deterministic models of epidemic on networks focus on the prevalence or the probability of k simultaneously infected nodes. The benefit of our exact method is the ability to exactly determine the probability that any group of k nodes is simultaneously infected:

$$y^{(k)}(t) = \frac{1}{\binom{N}{k}} \sum_{\mathcal{S} \subseteq \mathcal{N}, |\mathcal{S}| = k} \Pr[X_{i_1} = 1, X_{i_2} = 1, \ldots, X_{i_k} = 1].$$

For $k = 1$, we recover the prevalence $y(t)$. We know that $y^{(1)}(t) \geq y^{(2)}(t) \geq \ldots \geq y^{(N)}(t)$, because $\Pr[X = 1, Y = 1] \leq \Pr[X = 1]$ for any two random variables X and Y. Figure 3.18 shows the joint infection probability of k groups of nodes on a graph with $N = 7$ nodes. All curves roughly exhibit their peak at the same epidemic peak time, but the height of

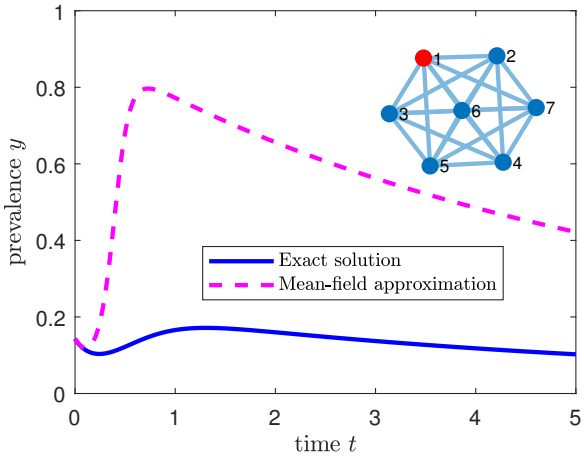


Figure 3.16: The exact solution (solid line) and mean-field approximation (dashed line) for SIR epidemics on the complete graph with $N = 7$ nodes. The heterogeneous mean-field approximation from Eq. (B.2) poorly matches the exact solution. Additionally in this example, the Markovian prevalence $y(t)$ is not monotonic.

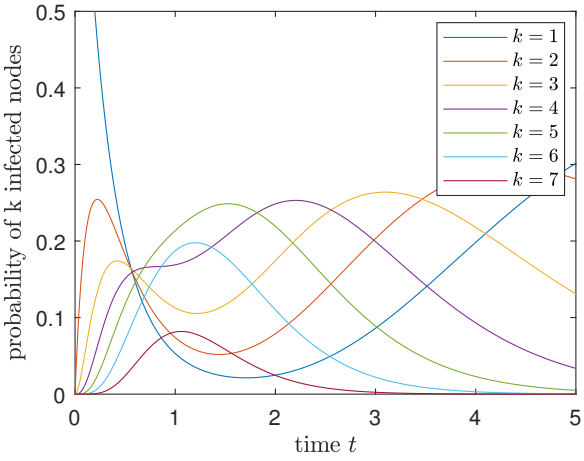


Figure 3.17: The probability of k infected nodes in SIR epidemics on a complete graph with $N = 7$ nodes and uniformly generated $\beta_{ij} \in [0.4, 0.9]$ and $\delta_i \in [0.1, 0.6]$. The probability of 1 infected node starts at 1, whereas all other curves start at 0.

the peak decreases with increasing k as expected.

The SIR epidemic model exhibits a phase transition around the epidemic threshold. Below the epidemic threshold, the disease dies out exponentially fast and above the threshold, the disease persists and infects a significant part of the population. It is conjectured that around the epidemic threshold, the probability of k groups being infected is roughly equal for several values of k. The epidemic threshold τ_c for Markovian SIR

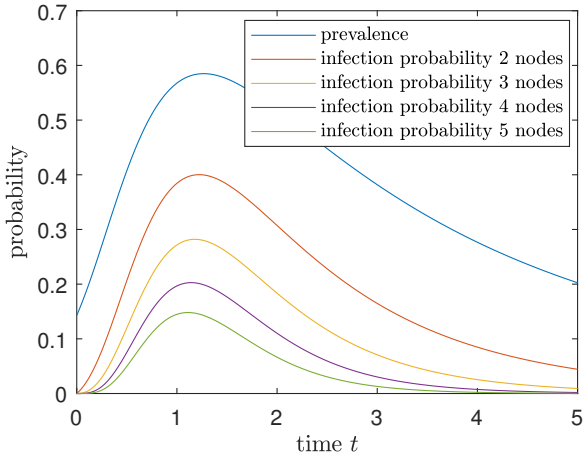
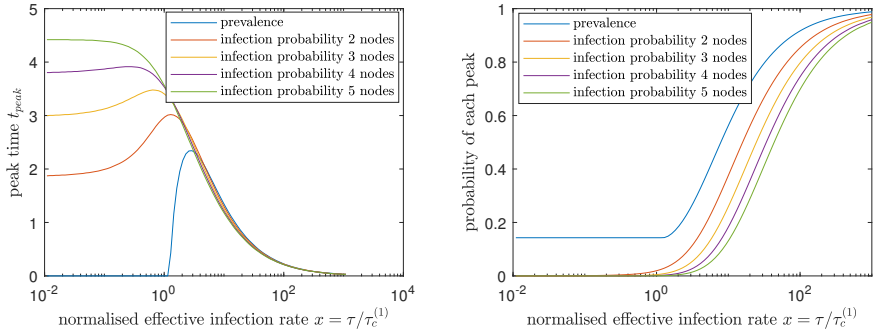


Figure 3.18: The joint infection probability of k groups of nodes in SIR epidemics on a complete graph with $N = 7$ nodes and uniformly generated $\beta_{ij} \in [0.4, 0.9]$ and $\delta_i \in [0.1, 0.6]$.

dynamics is not known, but lower-bounded by the mean-field threshold $\tau_c^{(1)}$, which is given in Eq. (B.3) in Appendix B.2. Figure 3.19 shows the peak time in (a) and height of the peak in (b). Around the epidemic threshold, which is approximated located at the normalised effective infection rate $x = \tau/\tau_c^{(1)} = 1$, we do not observe that all curves are approximately of the same order of magnitude. Most likely, a graph with $N = 7$ nodes is too small to exhibit a clear epidemic threshold.



(a) (b)

Figure 3.19: Finding the (a) time and (b) size of the epidemic peak for k groups of infected nodes in SIR epidemics on a complete graph with $N = 7$ nodes and uniformly generated $\beta_{ij} \in [0.4, 0.9]$ and $\delta_i \in [0.1, 0.6]$. Around the $x = 1$, which lower-bounds the epidemic threshold, all curves do not have the same magnitude.

3.3.7. THE LAPLACE TRANSFORM

In the limit of large networks, the solution (3.20) is composed of exponentially many terms and cannot be computed. Fortunately, we can rewrite (3.20) using Abel summation in the limit $N \to \infty$ to find (see Appendix B.3 for the derivation):

$$y(t) = t \int_0^\infty e^{-xt} g(x) dx, \qquad (t > 0) \tag{3.24}$$

and where

$$g(x) = \sum_{l=0}^{[\lambda^{-1}(x)]} \tilde{c}_l, \tag{3.25}$$

where $[x]$ indicates the integer part of x and $\lambda^{-1}(x)$ is the inverse eigenvalue function. The function $g(x)$ is the inverse Laplace transform of $\frac{y(t)}{t}$ and contains all information about the prevalence, but is expressed in the frequency x domain, rather than the time t domain. For the complete graph with $N = 7$ nodes, Figure 3.20a shows that the contribution of each frequency x exhibits a complicated structure of $g(x)$. Even for a graph with only 7 nodes, Figure 3.20b shows that the prevalence $y(t)$ from (3.24) approximately agrees with the exact solution (3.19). Equation (3.24) is derived under the assumption $N \to \infty$, which is accurate everywhere, except around $t = 0$. For a finite graph, the initial condition $y(0) = 1/N$ is a finite number, whereas for $N \to \infty$, the initial prevalence converges to zero, causing a discrepancy between the two solutions. We conclude that for most networks of a realistic size, Eq. (3.24) seems to be an accurate approximation of the solution (3.18).

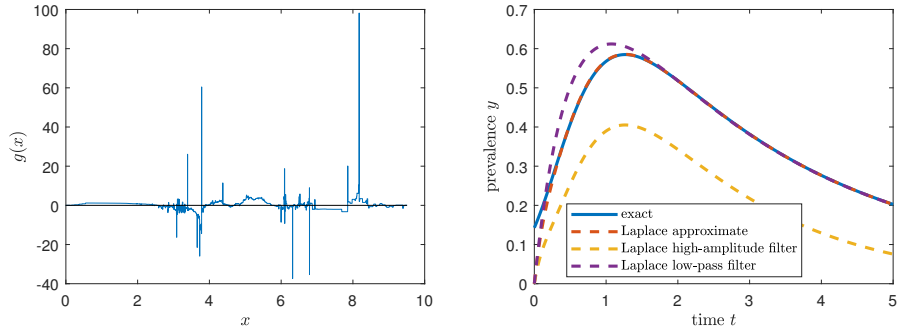


Figure 3.20: (a) The function $g(x)$ from Eq. (3.25) and (b) the exact prevalence (3.18) of the SIR process on a graph with $N = 7$ nodes and uniformly generated $\beta_{ij} \in [0.4, 0.9]$ and $\delta_i \in [0.1, 0.6]$. Additionally, (b) shows the prevalence computed by the Laplace transform (Eq. (3.24)), the high-amplitude filter with cut-off amplitude $\epsilon = 1$ and low-pass filter with frequency $x \in [0, 2]$. The low-pass filter uses approximately 330 out of 2059 non-zero eigenvalues.

Instead of the whole function $g(x)$, partial information on $g(x)$ may be sufficient to reconstruct the time-varying prevalence $y(t)$. We apply two techniques from signal

processing to $g(x)$. Method 1 performs a high-amplitude filter, keeping only frequencies with amplitude larger than some bound ϵ and all other contributions are set to zero;

$$g_1(x) = \begin{cases} g(x), & g(x) \geq \epsilon \\ 0. & \text{otherwise} \end{cases}$$

Second, we employ a low-pass filter, only keeping the contributions of frequencies in the interval $[0, f_0]$;

$$g_2(x) = \begin{cases} g(x), & x \leq f_0 \\ 0. & \text{otherwise} \end{cases}$$

The main reason behind $g_2(x)$ is that small (in modulus) frequencies correspond to small (in modulus) eigenvalues, which have the largest contributions in the Laplace transform (3.24). Figure 3.20b shows the result of applying both filters to $g(x)$. The low-pass filter is superior to the high-amplitude filter, because small values of x have a more significant contribution in the Laplace transform (3.24). Even for an amplitude cut-off $\epsilon = 1$ in Figure 3.20b, which is relatively small, the low-pass filter better approximates the true solution.

Interestingly, eigenmode truncation (as explained in Section 3.2.3) is only effective for large times, as shown in Figure 3.21. Eigenmode truncation also maintains the smallest (in modulus) eigenvalues, but apparently performs much worse compared to the low-pass filter of the Laplace transform. The primary reason for this discrepancy is that eigenmode truncation considers finite networks whereas the Laplace transform was derived under the assumption of an infinitely large network. Even for a network with $N = 7$ nodes, the Laplace transform (3.24) performs better using the same number of eigenmodes. The advantage of the latter is that the solution can no longer explode around $t = 0$ because the t-term in (3.24) guarantees $y(0) = 0$.

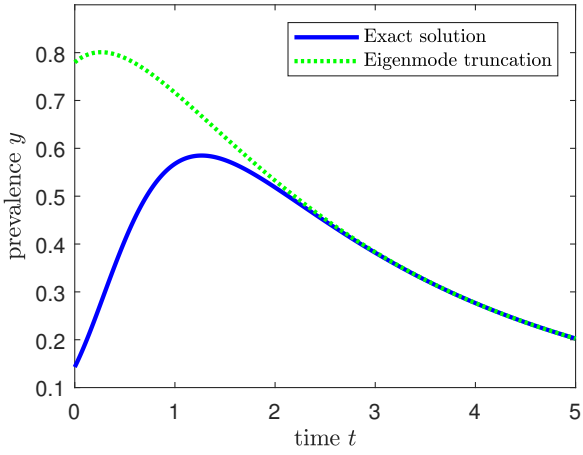


Figure 3.21: Eigenmode truncation in SIR epidemics on the complete graph with $N = 7$ nodes and uniformly generated $\beta_{ij} \in [0.4, 0.9]$ and $\delta_i \in [0.1, 0.6]$. Approximately 330 out of 2059 non-zero eigenvalues are used.

3.4. CONCLUSION

We investigated continuous-time Markovian compartmental models on networks. We focussed on transient epidemic compartmental models such as SI and SIR, where reinfections do not appear. Then the compartmental graph describing the transitions between compartments does not contain loops. We showed that such a continuous-time Markovian compartmental model admits an analytic solution using eigendecomposition of the infinitesimal generator. We demonstrated our method on the SI and SIR process. We derived the eigenvalues of the infinitesimal generators and showed that they are related to cut-sets in the graph. Assuming that all non-zero eigenvalues are distinct, we constructed the exact solution of the SI and SIR process on heterogeneous networks. Additionally, we showed that the transient compartmental model can be extended by adding self-infections, simplicial contagion, temporal networks and non-Markovian dynamics while maintaining an exact, analytic solution for the time-varying infection probabilities of all nodes.

The advantage of our exact approach is twofold. First, the analytic solution on small networks can serve as a calibration tool for (non)-Markovian compartmental simulators on networks and subsequently allows researchers to determine the number of simulation events to achieve a certain accuracy. Second, we showed that the underlying state space scales exponentially with the number of nodes. Much effort is devoted to simplify the state space, e.g. by aggregating states in the Markov graph. Our exact time-dependent solution can assess the quality of the proposed aggregation methods in an exact manner.

Finally, the advance in quantum computers and quantum algorithms may help to exactly determine the time-varying prevalence for much larger networks. There is an strong analogy between the viral state of a node in a network and a quantum state of a particle in a quantum device. Also, the entire state space of a quantum system is increasing exponentially in the number of qubits. Hence, we expect a computational breakthrough and an unraveling of the SIR epidemic phase transition with the further development of quantum computers.

II

EPIDEMICS ON ADAPTIVE NETWORKS

4

THE GENERALISED ADAPTIVE SIS PROCESS

In the classical SIS model, a disease or infection spreads over a given, mostly fixed graph. However, in many real complex networks, the topology of the underlying graph can change due to the influence of the dynamical process. In this chapter, we assume that the network adaptively changes its topology based on the presence of the virus in the network. An entire class of link-breaking and link-creation mechanisms, which we name Generalised Adaptive SIS (G-ASIS), is presented and analysed. For each instance of G-ASIS, the relation between the epidemic threshold and the effective link-breaking rate is determined to be either linear, constant or remains unknown. We confirm our theoretical results with numerical simulations.

This chapter is based on M. A. Achterberg, J. L. A. Dubbeldam, C. J. Stam and P. Van Mieghem, *Classification of link-breaking and link-creation updating rules in susceptible-infected-susceptible epidemics on adaptive networks*, Physical Review E 101, 052302, May 2020 [75].

4.1. INTRODUCTION

A major open problem in epidemiology is to understand the effect of people's personal responses to an epidemic outbreak. For example, an individual can decide to break contact with other individuals to prevent themselves or others from contracting the disease. In that case, the local contact network of the individual changes based on the spread of the disease. Such networks are called coevolutionary or adaptive networks [76, 77], because the contact network changes due to the spread of the disease in the population.

Many networks can be modelled as adaptive networks. For example, the brain connectome is a highly adaptive network [78]. Opinion networks, in which opinions are transferred between people, also adapt over time as people commonly prefer to contact people with similar opinions [79]. Also, time-evolving contact patterns during epidemics are highly adaptive [80].

One of the first adaptive epidemic models was introduced by Gross *et al.* [81], which describes the spread of an SIS epidemic with a link-rewiring process. Each infected node can infect its healthy neighbour with probability p. Independent of the infection process, infected individuals can cure with probability r. To model the adaptive behaviour of individuals, Gross *et al.* introduces a link rewiring process. Susceptible nodes may rewire their link with an infectious neighbour to a randomly chosen susceptible node with probability w. Gross's model was analysed extensively [82, 83, 84, 85] and several other rewiring schemes have been proposed [86, 87, 88, 89, 90]. Link-rewiring schemes have also been investigated in other epidemic models, such as SIR [91, 92], SIRS models [93] and on growing networks [94].

The seminal work of Gross *et al.* [81] allows for the rewiring of links in the network, but the total number of links in Gross's model is fixed. Actually, in real-world epidemics, the number of links in the contact network varies over time, due to natural fluctuations and disease countermeasures. Several studies have considered a time-varying number of links. Tunc *et al.* [7] proposed to break links and automatically restore them after a fixed time. Zhou *et al.* [95] investigated growing networks, in which links between susceptible and infected nodes can be broken. Sahneh *et al.* [96] considered the interplay between the disease spread and the spread of awareness on the disease in a multilayer network. Guo *et al.* [52] introduced the Adaptive SIS (ASIS) model , where links between susceptible and infected nodes are not rewired, but broken. Independently, a broken link between two susceptible nodes can be restored. Hence, the network evolves according to two processes: a link-breaking and a link-recreation mechanism. Aside from epidemiology, the methodology was successfully applied to model the spread of information propagation in the Adaptive Information Diffusion (AID) model [97].

In order to gain understanding of how the link dynamics affect the overall dynamics of adaptive networks, we propose the Generalised Adaptive SIS model (G-ASIS for short). The versatile G-ASIS model comprises the Adaptive SIS (ASIS) [52] and Adaptive Information Diffusion (AID) [97] model by allowing for all possible link-breaking and link-creation mechanisms. The G-ASIS model assumes that the links between nodes can be changed based on two processes. On the one hand, the links between two nodes can be broken with a certain probability. Another rule describes the possibility for links to be created between two disconnected nodes. The probability for the link-breaking and link-creation process is dependent on the current viral state of the two end-nodes of the

link. Hence, the underlying contact network *adapts* to the spread of the epidemic.

The G-ASIS model was developed independently by Kiss *et al.* [98]. The major difference is that they primarily focus on the mean-field approximation of the G-ASIS model, which is the focus of our Chapter 5. On the other hand, our results in this chapter concern the original Markovian G-ASIS model, which is significantly more complicated. Moreover, we additionally derive all possible link-updating rules and upper and lower bounds for the epidemic threshold.

The G-ASIS model assumes *local awareness* of the nodes on the disease (i.e. nodes decide to break or create links based on the viral state of the neighbours). On the contrary, nodes may hear about reports of the number of infected cases on the news, leading to awareness on the viral prevalence in the global population, which is called *global awareness* [10, 96]. The interplay between both types of awareness is an interesting and active research area [99]. We further discuss awareness in epidemics in Chapter 6.

This chapter is structured as follows. In Section 4.2, we derive and explain the G-ASIS model and discuss all possible updating rules of the network dynamics. In Section 4.3, we derive a lower bound for the epidemic threshold for each of the G-ASIS instances. An implicit relation for the epidemic threshold is also derived. Next, we present simulation results in Section 4.4 and finally, we summarize and discuss our findings in Section 4.5.

4.2. GENERALISED ADAPTIVE SIS MODEL

Throughout this chapter, we primarily use terminology and notation from epidemiology to introduce and explain various concepts, but the results also apply to general spreading phenomena, ranging from gossips, political preferences, opinions, information spread in the human brain, raising awareness about a particular event, innovation spread, cascading failures and other spreading processes.

4.2.1. MODEL DESCRIPTION

We consider the spread of a disease over a graph $G(\mathcal{N},\mathcal{L})$ where $\mathcal{N}$ is the set of N nodes and $\mathcal{L}$ is the set of L links. Every node i represents an individual which can be in two states: infected or healthy. The viral state of node i is denoted by the Bernoulli random variable $X_i(t)$, which equals $X_i(t)=1$ if node i is infected at time t and $X_i(t)=0$ if node i is healthy, but susceptible to the disease. An infected node i can infect a neighbouring susceptible node j via a Poisson process with rate β_{ij} if the two nodes are connected $a_{ij}(t)=1$ at time t. Independently, an infected node i can recover from the disease with Poisson rate δ_i. Thus, the state X_i of the node i changes as follows

$$\frac{d\,\mathrm{E}[X_i(t)]}{dt} = \mathrm{E}\Big[-\delta_i X_i(t) + (1-X_i(t))\sum_{j=1}^{N}\beta_{ij}X_j(t)a_{ij}(t)\Big]. \tag{4.1}$$

The right-hand side of (4.1) consists of two parts: an infected node i cures with rate δ_i and a susceptible node i can be infected by each of its neighbouring infected nodes j with rate β_{ij}.

Besides the evolution of the spreading process, the graph evolves over time as well. Here, we present the Generalised Adaptive SIS (G-ASIS) model, which assumes that the

link a_{ij} between node i and j changes based on two independent processes: (i) a link-creation process f_{cr} with Poisson rate ξ_{ij} and (ii) a link-breaking process f_{br} with Poisson rate ζ_{ij}. Both the link-breaking and the link-creation process depend on the viral state of the two nodes attached to the link. Similar to the viral state $X_i(t)$, the link a_{ij} between node i and j is also modelled as a Bernoulli random variable. We assume that the interaction between node i and j is symmetric, such that the adjacency matrix $A(t)$ with elements $a_{ij}(t)$ remains symmetric for all times t.

Not all links necessarily adhere to the link-breaking and link-creation mechanisms. Instead, some links may be (non)-existent permanently. We denote the set of permanently non-existing links in the network by $\mathcal{L}_0$ and the set of permanently existing links by $\mathcal{L}_1$. These links do not adhere to the link-creation and link-breaking mechanisms but instead are always non-existent and existent, respectively, for all times. The remaining set of links $\mathcal{L}_{\text{adaptive}} = \mathcal{L} \backslash \{\mathcal{L}_0 \cup \mathcal{L}_1\}$ evolve according to the link-creation and link-breaking mechanisms. We denote the number of links of each type by L_0, L_1 and L_{adaptive}, respectively.

Then the governing equation for the link a_{ij} is given by

$$\frac{d\,\mathrm{E}[a_{ij}(t)]}{dt} = \begin{cases} \mathrm{E}\Big[-\zeta_{ij} a_{ij}(t) f_{\text{br}}(X_i(t), X_j(t)) \\ \quad +\xi_{ij}(1-a_{ij}(t)) f_{\text{cr}}(X_i(t), X_j(t))\Big], & \text{if } (i,j) \in \mathcal{L}_{\text{adaptive}} \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

where f_{br} and f_{cr} are the link-breaking and link-creation mechanism, respectively. If the link-breaking rate $\zeta_{ij} = 0$ and the link-creation rate $\xi_{ij} = 0$ for all nodes i, j, then the process simplifies to the SIS process on a static network. On the other hand, if $f_{\text{br}} = f_{\text{cr}} = 1$, the network dynamics is decoupled from the disease dynamics and the network is a temporal network that evolves randomly. An extensive analysis of the decoupled, non-adaptive case is provided by Kiss *et al.* [98]. We emphasise that general temporal networks are not necessarily governed by independent link-breaking and link-creation mechanisms, but instead follow more complex patterns, including temporal correlations and cluster formation.

We proceed in the next section by deriving all possible link-breaking and link-creation mechanisms. We drop the explicit time-dependence of the random variables X_i and a_{ij} in the remainder of this chapter for clarity.

4.2.2. DERIVATION OF THE UPDATING RULES

The link-breaking mechanism f_{br} and link-creation mechanism f_{cr} in G-ASIS depend on the viral state $X_i(t)$ and $X_j(t)$ of node i and j, but not on $a_{ij}(t)$ nor explicitly on the time t. We next determine all possible updating rules for f_{br} and f_{cr}. For convenience, a rule is denoted by f and applies to f_{br} as well as to f_{cr}. Each rule f of a link between node i and j has Bernoulli random variables X_i and X_j as input. Each rule f is a linear or quadratic function of X_i and X_j that evaluates to zero or one, similar to a logical gate.

We classify the updating rules according to the number of possible inputs that give $f = 1$. Consider for example the rule $f = X_i X_j$, which is visualised in Figure 4.1. Then, $f = 1$ only for $X_i = X_j = 1$. Any other input for X_i and X_j yields $f = 0$. The number of permutations of this type can be computed as follows. There are four possible inputs

(combinations of X_i and X_j) and one positive outcome: $\binom{4}{1} = 4$. The complying rules are:

$$X_i X_j, \qquad (1-X_i)X_j,$$
$$(1-X_i)(1-X_j), \qquad X_i(1-X_j).$$

There are also rules for which two combinations of X_i and X_j yield $f = 1$. As an example,

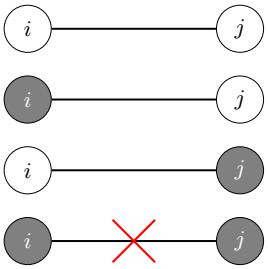


Figure 4.1: Schematic overview of two connected nodes. Grey nodes are infected nodes, white nodes are healthy nodes. The decision to break or create the link between node i and j depends on the viral states X_i and X_j. In this example, the link is broken only if $X_i = X_j = 1$, which corresponds to link-breaking rule $f_{\text{br}} = X_i X_j$.

consider the rule $f = (X_i - X_j)^2$. Then $f = 1$ if X_i is not equal to X_j. There are six rules of this type, because there are four inputs and two combinations; $\binom{4}{2} = 6$. These 6 rules are:

$$(X_i - X_j)^2, \qquad X_i, \qquad X_j,$$
$$1-(X_i - X_j)^2, \qquad (1-X_i), \qquad (1-X_j).$$

Thirdly, there are rules for which three combinations of X_i and X_j yield $f = 1$. For example, consider $f = 1 - X_i X_j$. The function's result is one if $X_i = 0$ or $X_j = 0$. The only situation to find $f = 0$ occurs when $X_i = X_j = 1$. All four rules of this type, namely $\binom{4}{3} = 4$, are:

$$1 - X_i X_j, \qquad 1-(1-X_i)X_j,$$
$$1-(1-X_i)(1-X_j), \qquad 1-X_i(1-X_j).$$

Two trivial rules have not yet been specified. The trivial rules $f = 1$ (which occurs in $\binom{4}{4} = 1$ case) and $f = 0$ (also in $\binom{4}{0} = 1$ case) are independent of the viral state of nodes X_i and X_j. Including the trivial rules, the total number of possible rules is $\sum_{k=0}^{4}\binom{4}{k} = 2^4 = 16$.

Each of the 16 possibilities for the function f can be rewritten, using the binomial property $\mathrm{E}[X_i^2] = \mathrm{E}[X_i]$, in the following parametrised form:

$$f(X_i, X_j) = a + bX_i + \tilde{b}X_j + cX_i X_j, \tag{4.3}$$

where the parameters $a, b, \tilde{b}, c \in \mathbb{Z}$. Since the assumed network is undirected, the function f must be symmetric in X_i and X_j, which implies that $\tilde{b} = b$ in (4.3). This removes eight asymmetric updating rules from the original derivation and simplifies (4.3) to

$$f(X_i, X_j) = a + b(X_i + X_j) + cX_i X_j, \tag{4.4}$$

where the parameters $a, b, c \in \mathbb{Z}$.

The trivial updating rules $f = 0$ and $f = 1$ are not particularly relevant. Choosing the updating rule $f = 0$ for either the link-breaking or link-creation mechanism removes the mechanism entirely from the governing equation (4.2). Hence, there is an exponentially fast convergence to the steady-state topology, without any dependence on the SIS process. The updating rule $f = 1$ is also a non-adaptive rule which is independent of the viral state of X_i and X_j. Hence, for our analysis of the epidemic threshold in the G-ASIS model, the non-adaptive rules are not incorporated.

After the removal of the non-adaptive and non-symmetric rules for the function f, only six updating rules remain. Therefore, the link-breaking mechanism f_{br} and link-creation mechanism f_{cr} each have six updating rules in the G-ASIS model. Since the link-breaking mechanism f_{br} and the link-creation mechanism f_{cr} can be chosen independently, and for each of them six updating rules are available, in total 36 Markov processes for topology updating are contained in G-ASIS. Each G-ASIS instance contains two mechanisms: a link-breaking mechanism f_{br} and a link-creation mechanism f_{cr}, which are given in general form by equation (4.4). An overview[1] of all updating rules is presented in Table 4.1.

Table 4.1: All updating rules for the link-breaking and the link-creation mechanism in the G-ASIS model. The rules for the link-breaking and link-creation mechanisms are structured. The inverse of any rule f is $1-f$. Also, taking the multiplication of two rules is equivalent to taking the intersection between the number of times a positive result for the rules is found.

rule f	a	b	c	gate
X_iX_j	0	0	1	AND
$1-X_iX_j$	1	0	-1	NAND
$(1-X_i)(1-X_j)$	1	-1	1	NOR
$1-(1-X_i)(1-X_j)$	0	1	-1	OR
$(X_i-X_j)^2$	0	1	-2	XOR
$1-(X_i-X_j)^2$	1	-1	2	XNOR

As an example, we consider the Adaptive SIS model, where the link between a susceptible node and an infected node is broken to prevent the spreading of the disease. Hence, the link-breaking mechanism f_{br} is equal to the updating rule $f_{\text{br}} = (X_i - X_j)^2$ and the corresponding parameters in (4.4) are $(a_{\text{br}}, b_{\text{br}}, c_{\text{br}}) = (0, 1, -2)$. When both end nodes of a link are susceptible, the link between the nodes is restored. The link-creation mechanism is therefore $f_{\text{cr}} = (1 - X_i)(1 - X_j)$ with parameters $(a_{\text{cr}}, b_{\text{cr}}, c_{\text{cr}}) = (1, -1, 1)$.

We assume that the infection, curing, link-breaking and link-creation processes are all independent Poisson processes, whose combined dynamics can be described by a continuous-time Markov chain. A schematic overview of the G-ASIS model is shown in Figure 4.2.

[1] The number of updating rules can also be derived by regarding power sets. For any updating rule f for the link a_{ij}, we write the symmetric nodal input as $\{X_i, X_j\}$. All possible symmetric input combinations for the rule f are $\mathcal{X} = \{\{0,0\},\{0,1\},\{1,1\}\}$. Each element in $\mathcal{X}$ can be zero or one, depending on whether the link can be changed. The total number of combinations is then given by the power set of $\mathcal{X}$, denoted as $2^{\mathcal{X}}$, which contains $2^3 = 8$ elements. Two elements from $2^{\mathcal{X}}$ correspond to the trivial rules $f = 0$ and $f = 1$. After removal of the trivial rules, we find six updating rules for each link-updating mechanism.

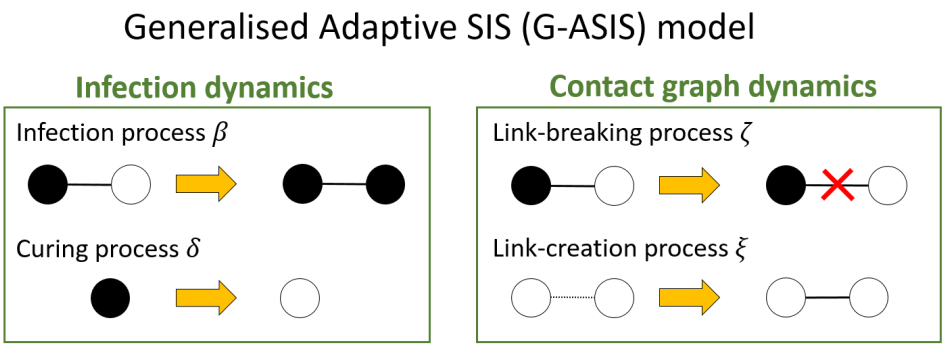


Figure 4.2: An overview of the processes in the G-ASIS model. We show a random example for the link-breaking and link-creation mechanisms, where the link is broken between susceptible and infected nodes and is created between two susceptible nodes. All possible updating rules for the link-breaking and link-creation mechanisms are specified in Table 4.1.

We confine ourselves in the remainder of this chapter to homogeneous curing, infection, link-breaking and link-creation rates. Furthermore, we assume that the set of permanently existing links $\mathcal{L}_1$ is empty and the set of adaptive links $\mathcal{L}_{\text{adaptive}}$ forms a connected graph G_{adaptive}. If the connectivity constraint of the adaptive links is not satisfied, the analysis must be repeated for each component individually. Despite the fact that we make these assumptions for analytic feasibility only, we expect that the conclusions and insights also hold for heterogeneous rates.

4.3. THEORETICAL RESULTS

One of the main concepts in epidemiology is the epidemic threshold τ_c. The epidemic threshold τ_c in a finite graph specifies a small interval for the effective infection rate $\tau = \beta/\delta$ in which the process quickly changes from the disease-free phase to the endemic phase [6]. The epidemic threshold τ_c can be defined as the largest value of the effective infection rate τ for which the prevalence y exponentially decays to zero over sufficiently large time [100]. Finding an analytical expression for the epidemic threshold is generally infeasible due to the complexity of the process. It is, however, possible to derive lower and upper bounds for the epidemic threshold.

4.3.1. LOWER BOUND ON THE EPIDEMIC THRESHOLD

Following [101], the epidemic threshold τ_c can be bounded from below. This methodology was also successfully applied to the static SIS model [44, Theorem 17.3.1]. We state one of our main results in Theorem 4.1.

Theorem 4.1 *The epidemic threshold τ_c for the G-ASIS model is bounded from below by*

$$\tau_c \geq \frac{1}{\lambda_1}\left(1 + \frac{\omega(\mathbf{1}_{\{a_{\text{br}}=0,b_{\text{br}}=1,c_{\text{br}}=-1\}}) - (\mathbf{1}_{\{a_{\text{cr}}=1,b_{\text{cr}}=0,c_{\text{cr}}=-1\}\cup\{a_{\text{cr}}=0,b_{\text{cr}}=1,c_{\text{cr}}=-2\}})}{(1-\mathbf{1}_{\{a_{\text{cr}}=1,b_{\text{cr}}=-1,c_{\text{cr}}=1\}}) + \delta/\xi}\right), \tag{4.5}$$

where λ_1 is the spectral radius (the largest eigenvalue) of the adjacency matrix of adaptive

graph G_{adaptive}, $\omega = \zeta/\xi$ is the effective link-breaking rate and $\mathbf{1}_x$ is the indicator function which is one if condition x is satisfied, and zero otherwise.

Proof. See Appendix C.1. □

For the SIS process on static graphs, it holds that $\tau_c \geq \frac{1}{\lambda_1}$ where λ_1 is the spectral radius of the adjacency matrix [51, 56]. Compared to the static case, Theorem 4.1 states that the lower bound for the epidemic threshold τ_c can both increase and decrease by introducing adaptive link-breaking and link-creation mechanisms. If the link-breaking coefficients satisfy $a_{\text{br}} = 0, b_{\text{br}} = 1, c_{\text{br}} = -1$ (corresponding to link-breaking rule $f_{\text{br}} = 1-(1-X_i)(1-X_j)$), then the epidemic threshold τ_c in (4.5) has a non-zero dependence on the effective link-breaking rate ω, regardless of the choice of the link-creation rule. Hence, for 6 out of the 36 instances of G-ASIS, the epidemic threshold τ_c increases at least linearly with the effective link-breaking rate ω. For the remaining 30 instances in G-ASIS, satisfying $f_{\text{br}} \neq 1-(1-X_i)(1-X_j)$, the lower bound in (4.5) is independent of the effective link-breaking rate ω and is similar to the lower bound of the classical SIS epidemic threshold. In epidemiology, a high epidemic threshold is preferable, because the disease only develops into an endemic for higher infection rates. Other areas of application, such as information spreading and human brain interactions, benefit from a low epidemic threshold as fast communication is advantageous for these phenomena.

4.3.2. UPPER BOUND ON THE EPIDEMIC THRESHOLD

We denote the fraction of infected nodes by $Z = \frac{1}{N}\sum_{i=1}^{N} X_i$. Above the epidemic threshold τ_c, the process remains for a long time in the *metastable state*, which was defined in Chapter 2. In the metastable state, stochastic variables are denoted with an asterisk (*). We denote by $y = \mathrm{E}[Z^*]$ the prevalence in the metastable state. Combining (4.1) and (4.2), an analytic, implicit quadratic relationship for the metastable prevalence y can be obtained, similarly as in [52] and [97].

Theorem 4.2 *If the adaptive graph G_{adaptive} is the complete graph, i.e. all links adhere to the link-breaking and link-creation mechanisms, the metastable prevalence y satisfies the quadratic equation*

$$y^2 + \left(\frac{2b_{\text{cr}}N\tau - (2b_{\text{cr}}+c_{\text{cr}})\tau + c_{\text{br}}\omega + c_{\text{cr}}}{c_{\text{cr}}N\tau}\right)y + \left(\frac{(N-1)a_{\text{cr}}}{c_{\text{cr}}N} - \frac{a_{\text{br}}\omega + a_{\text{cr}}}{c_{\text{cr}}N^2}\mathrm{E}\left[\sum_{i=1}^{N} d_i^*\right]\right.$$
$$\left. + \mathrm{Var}(Z^*) - \frac{(2b_{\text{br}}+c_{\text{br}})\omega + 2b_{\text{cr}} + c_{\text{cr}}}{c_{\text{cr}}N^2}\mathrm{E}\left[\sum_{i=1}^{N} d_i^* X_i^*\right]\right) = 0. \tag{4.6}$$

Proof. See Appendix C.2. □

The quadratic formula (4.6) for the prevalence y leads to an exact, implicit expression for the epidemic threshold τ_c:

Theorem 4.3 *If the adaptive graph G_{adaptive} is the complete graph, i.e. all links adhere to the link-breaking and link-creation mechanisms, the epidemic threshold τ_c in the G-ASIS*

model is implicitly given by

$$\tau_c = \frac{\frac{c_{\text{br}}}{c_{\text{cr}}}\omega + 1}{2\frac{b_{\text{cr}}}{c_{\text{cr}}}(1-N) + 1 - Nh(\omega,\xi,\tau_c)}, \tag{4.7}$$

where $h(\omega,\xi,\tau_c)$ is defined in (C.19) in Appendix C.3. Moreover, for 27 instances of G-ASIS, the epidemic threshold τ_c is upper-bounded by a linear function in ω and for 9 instances, the epidemic threshold is upper-bounded by a constant.

Proof. See Appendix C.3. □

Theorem 4.1 and 4.3 enable us to classify the instances of G-ASIS: some instances have a linear relation between the epidemic threshold τ_c and the effective link-breaking rate $\omega = \zeta/\xi$ and for the other instances, the epidemic threshold τ_c is independent of the effective link-breaking rate ω. Comparing the results of Theorem 4.1 and 4.3, two striking differences appear. First, the lower bound for the epidemic threshold τ_c in (4.5) is explicit, whereas (4.7) depends *implicitly* on the function $h(\omega,\xi,\tau_c)$. Second, the lower bound in (4.5) concludes that six instances have a linear relation between the epidemic threshold τ_c and the effective link-breaking rate ω, which contrasts the upper bound in Theorem 4.3 which has 27 linear-scaling instances. Subsequently, $27-6=21$ instances have an undetermined relation: their lower bound is independent of ω, whereas their upper bound scales linearly in ω. In Section 4.4, simulation results indicate that undetermined relations can exhibit both linear and constant behaviour. The relation between the epidemic threshold τ_c and the effective link-breaking rate ω can be summarised as follows:

6 instances:	linear in ω,
9 instances:	constant in ω,
21 instances:	undetermined.

4.4. NUMERICAL SIMULATIONS

The time in the governing equations (4.1) and (4.2) of the G-ASIS Markov process can be rescaled by the curing rate δ; hence, we always take $\delta = 1$. For the simulations in this section, the continuous-time G-ASIS Markov process is approximated by a sampled-time Markov chain [44] with a sufficiently small time step $\Delta t = 0.05$. All simulations assume a complete adaptive graph and start initially with a complete graph and all nodes infected. Each diagram is created by simulating a single simulation for 10^6 time units. The metastable prevalence y is obtained by averaging over all times in the interval $[10^5, 10^6]$. In all simulations, a small self-infection rate ε has been added, to ensure the G-ASIS process does not converge to the absorbing, all-healthy state.

4.4.1. PHASE TRANSITIONS

The relation between the prevalence y and the effective infection rate τ shows a phase transition from the all-healthy state to the endemic state around the epidemic threshold τ_c. Such phase transitions are shown for various instances of the G-ASIS model in Figure 4.3.

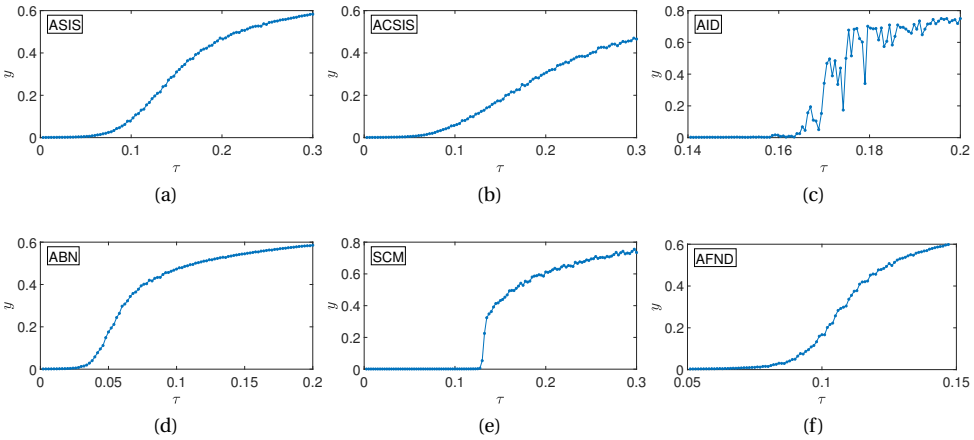


Figure 4.3: The relation between the effective infection rate τ and prevalence y for various instances of the G-ASIS model. We have taken $N = 40$, $\delta = 1$, $\varepsilon = 0.001$ and a complete initial network for all models. For (a,b,d,f), we have taken $\zeta = \xi = 1$, for (c) $\zeta = 0.5, \xi = 0.1$ and for (e) $\zeta = \xi = 0.1$.

The numerical results for the prevalence y in the ASIS model are presented in Figure 4.3a. Below the epidemic threshold $\tau_c \approx 0.05$, the prevalence y is zero. For effective infection rates $\tau > \tau_c$ the prevalence increases rapidly. The growth saturates as the effective infection rate τ increases and the prevalence y asymptotically increases to 1 as $\tau \to \infty$.

The Adaptive Contagious SIS (ACSIS) model is a variation on the ASIS model, where links are not only broken between susceptible and infected nodes, but also between two infected nodes. The reasoning behind ACSIS is that two people suffering from a disease are more likely to stay at home, effectively breaking links with each other. The epidemic threshold of ASIS (Figure 4.3a) and ACSIS (Figure 4.3b) are nearly equivalent, although the metastable prevalence y is generally lower in the ACSIS model. Due to the extra link-breaking rule in the ACSIS model, the disease is able to spread less quickly, causing the prevalence to decrease.

In contrast to the ASIS and ACSIS model, the Adaptive Information Diffusion (AID) model describes the spreading of information. In the AID model, nodes represent people and links their social interactions. The link between two susceptible nodes can be broken, because both nodes are not aware of the information and are no longer interested to upkeep their relationship. The link between susceptible and infected nodes can be created to enhance the propagation of information. Figure 4.3c shows that the metastable prevalence y heavily oscillates above the epidemic threshold τ_c. The reason is that the AID process is bistable, where both the all-healthy state and the endemic state are stable. Since Figure 4.3c is produced by a single iteration, a simulation may stay for a long time in either of the stable states, causing large differences across simulations with different infection rates τ. We further explain this phenomenon in Chapter 5.

The Adaptive Brain Network (ABN) model describes information transport in the human brain. Nodes represent different regions in the human brain and links specify the

connections between the brain regions. The nodes can be active (infected) or inactive (healthy) at any time T. From a control system point of view, the human brain incorporates two brain operational principles: (a) Hebbian learning, where two actively communicating nodes continuously try to improve their communication channel (i.e. increasing the weight of their link, known as their synaptic strength) and (b) homeostatic plasticity [102], which reduces the interaction strength between two connected nodes to prevent a positive coupling generated by Hebbian learning. The ABN model considers homeostatic plasticity to be the primary link-adaptation mechanism. Other G-ASIS instances may be used to describe Hebbian learning. Thus, new links can be created between two inactive nodes and existing links can be removed when both nodes are active. When one node is infected and one is healthy, the link between the nodes is preserved. The phase transition in Figure 4.3d is comparable to the ASIS model, although the epidemic threshold τ_c is smaller and the ascent of the prevalence y is steeper around the epidemic threshold.

Another instance of the G-ASIS model is the Scientific Collaboration Model (SCM), where nodes represent researchers showing interest (or not) in a particular research area. Links represent collaborations between researchers. Researchers can spread their interest to collaborating, connected researchers. Independently, researchers can lose, forget or do not pay attention to the research area. Besides the infection process, the network is evolves in the following way. Researchers can break their link if both are not interested in the research area. Since there is a potential collaboration between susceptible and infected researchers, their link persists. Finally, the link can be created between two researchers who are both interested, but are not yet collaborating. The phase diagram in Figure 4.3e is comparable to the ABN model in Figure 4.3d, however, the ascent of the prevalence y is very steep around the epidemic threshold τ_c.

The Adaptive Fake News Diffusion (AFND) model describes the spread of fake news in a social network. The nodes in the AFND model represent people who either believe or do not believe a fake news item. People are connected to other people over adaptive links. Infected nodes try to persuade healthy, neighbouring nodes to believe the fake news item. Simultaneously, infected nodes may 'cure' from the fake news. Links in the AFND model can be broken between susceptible and infected nodes based on social awareness against fake news. Simultaneously, two healthy nodes have no interest in keeping their relationship regarding the fake-news item, so their link can be broken. Additionally, links are assumed to be created between healthy and susceptible nodes since fake news items are mostly sensational and may trigger interest between the two people to connect with each other. Hence, the spreading of fake news causes links between susceptible and infected nodes to be created and broken simultaneously. The phase diagram of the AFND model, shown in Figure 4.3f, is similar to that of the ASIS and ACSIS model in Figure 4.3a and 4.3b respectively.

4.4.2. RELATION BETWEEN THRESHOLD AND LINK-BREAKING RATE

Although the epidemic threshold was shown for various G-ASIS instances in Figure 4.3, the quantitative effect of the link-updating mechanisms on the spreading of the disease remains unclear. Thus, we investigate the relation between the epidemic threshold τ_c and the effective link-breaking rate ω for various G-ASIS instances in Figure 4.4. The

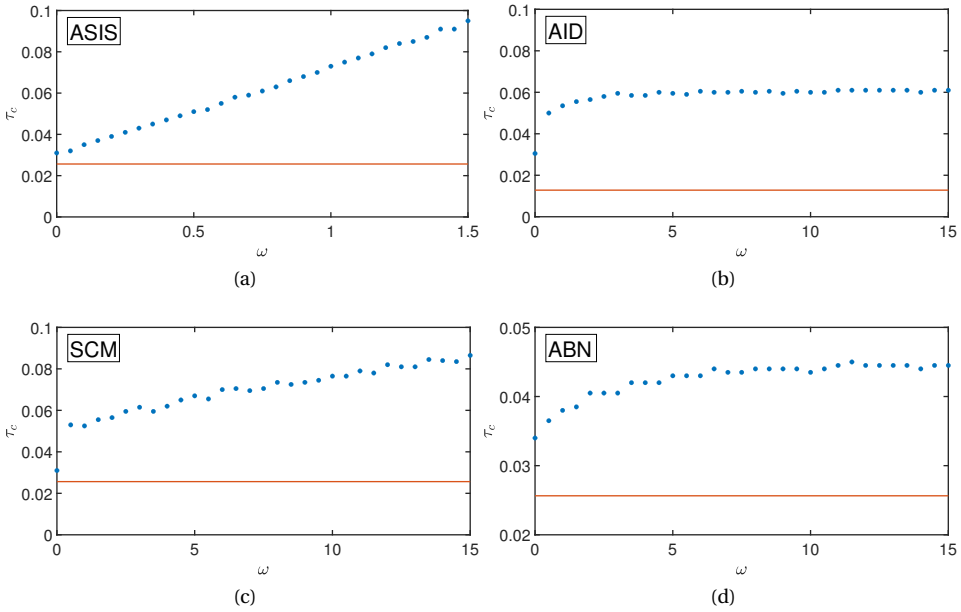
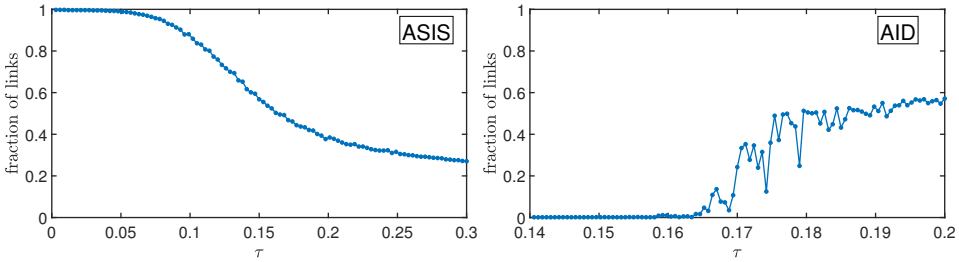


Figure 4.4: The epidemic threshold τ_c as a function of the effective link-breaking rate ω for four instances of the G-ASIS model. The data points are obtained from simulations and the solid line represents the lower bound from Theorem 4.1. The theory predicts a constant relationship (b) or is undetermined (a,c,d). We have taken $N = 40$, $\delta = \xi = 1$, $\varepsilon = 0.001$ and a complete initial network.

dots represent numerical simulations whereas the solid line represents the lower bound from Theorem 4.1. The result from Theorem 4.3 is not shown in Figure 4.4 because (4.7) is merely an implicit relation for the epidemic threshold τ_c. The AID model in Figure 4.4b shows nearly constant behaviour, which is in agreement with Theorem 4.3. For the ASIS, SCM and ABN models, shown in Figure 4.4a, 4.4c and 4.4d, respectively, the theory was not conclusive about the relation between the epidemic threshold τ_c and the effective link-breaking rate ω. Figure 4.4a shows a clear linear relationship and Figure 4.4d depicts a nearly constant relationship. In contrast, the relation in Figure 4.4c appears to be linear, but the relatively small slope indicates a weak relationship between the effective link-breaking rate ω and the epidemic threshold τ_c.

4.4.3. THE METASTABLE TOPOLOGY

In the G-ASIS model, the topology of the underlying graph is constantly changing over time. Nevertheless, the graph remains approximately constant in the metastable state. The characteristics of the metastable graph are of interest and contribute to the understanding of the interplay between disease spreading and topology updating. Any graph metric can be measured in the metastable state, but we focus here on the easiest metric: the number of links L. When the effective infection rate τ is smaller than the epidemic threshold τ_c, the metastable prevalence y is zero and the average number of links $\mathrm{E}[L]$



(a) (b)

Figure 4.5: The relation between the effective infection rate τ and the average fraction of links $\mathrm{E}[L]/(\frac{1}{2}N(N-1))$ for two instances of the G-ASIS model. We have taken $N = 40$, $\delta = 1$, $\varepsilon = 0.001$ and a complete initial network for all models. Subfigure (a) shows the ASIS model with $\zeta = \xi = 1$ and (b) illustrates the AID model with $\zeta = 0.5, \xi = 0.1$.

equals

$$\mathrm{E}[L] = \begin{cases} \frac{a_{\mathrm{cr}}}{a_{\mathrm{br}}\omega + a_{\mathrm{cr}}} \frac{1}{2} N(N-1), & \text{for } a_{\mathrm{cr}} \neq 0 \text{ or } a_{\mathrm{br}} \neq 0 \\ c, & \text{otherwise} \end{cases} \tag{4.8}$$

for any $c \in [0,1]$. Equation (4.8) follows from (4.6) by substituting an all-healthy population $y = 0$. However, the number of metastable links in the endemic state $\tau > \tau_c$ cannot be computed from neither (4.6) nor (4.7) in closed form. Thus, we resort to simulations to obtain the average number of links in the metastable state.

Figure 4.5 illustrates the fraction of links in the metastable state for two instances of G-ASIS. The ASIS model in Figure 4.5a starts with a completely connected graph at $\tau = 0$, because the prevalence is zero and the link-breaking mechanism between susceptible and infected nodes is rarely used. If the effective infection rate τ is larger than the epidemic threshold τ_c, the prevalence y is non-zero (see Figure 4.3a) and the link-breaking mechanism reduces the fraction of links. As the effective infection rate τ increases up to infinity, the prevalence y increases to 1 and the link-creation mechanism between two susceptible nodes is rarely activated. Hence, the fraction of links decreases to zero. For the AID model in Figure 4.5b, we observe opposite behaviour. If the effective infection rate τ is smaller than the epidemic threshold τ_c, the prevalence y is zero. In the AID model, links are broken between susceptible nodes and created between susceptible-infected pairs. If the prevalence is zero, there are no infected nodes and if a node gets infected, the node removes all connections to neighbouring nodes. Thus, the metastable graph does not contain any links. In the endemic state $\tau > \tau_c$, the prevalence y increases, which enables the creation of links in the network. As the effective infection rate τ approaches infinity, the prevalence increases to 1. Then the fraction of links also converges to 1 as $\tau \to \infty$, because the link-breaking rule between susceptible nodes in the AID model is rarely used, as there are hardly any susceptible nodes.

4.4.4. SUMMARY

Table 4.2 summarises the obtained results in this chapter. For the ACSIS model, the epidemic threshold τ_c is a linear function of the effective link-breaking rate ω and the

metastable state always exists. The AID model, which has a constant relation between the epidemic threshold τ_c and the effective link-breaking rate ω, possesses a bistable state above the epidemic threshold. Unfortunately, the relation between τ_c and ω in the ASIS, AFND, ABN and SCM models could not be determined. The simulations support the hypothesis that the lower bound is strict for the ABN model and the linear bound is correct for the ASIS, AFND and SCM model, indicating that both the upper and lower bound can be tight for different instances of the G-ASIS model.

Table 4.2: A selected set of instances from the G-ASIS model and their properties. The table assumes $\delta = 1$. The variable λ_1 denotes the spectral radius of the adjacency matrix of the adaptive graph G_{adaptive}.

Model name and appearance in literature	Updating rules (link-breaking) (link-creation)	Bistable state	Lower bound epidemic threshold τ_c	Upper bound epidemic threshold τ_c
ASIS model [7, 52, 97]	$(X_i - X_j)^2$ $(1-X_i)(1-X_j)$	No	$\frac{1}{\lambda_1}$	Linear
ACSIS model	$1-(1-X_i)(1-X_j)$ $(1-X_i)(1-X_j)$	No	$\frac{1}{\lambda_1}(1+\omega\xi)$	Linear
AID model [97]	$(1-X_i)(1-X_j)$ $(X_i - X_j)^2$	Yes	$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant
ABN model	$X_i X_j$ $(1-X_i)(1-X_j)$	No	$\frac{1}{\lambda_1}$	Linear
SCM model	$(1-X_i)(1-X_j)$ $X_i X_j$	No	$\frac{1}{\lambda_1}$	Linear
AFND model	$1-X_i X_j$ $(X_i - X_j)^2$	No	$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Linear

4.5. CONCLUSION

The Generalised Adaptive SIS (G-ASIS) model was introduced in this chapter to describe the spread of contagious processes on adaptive networks. The G-ASIS model consists of two adaptive mechanisms: links between nodes can be broken and created, based on the viral state of the nodes. We showed that for each mechanism, six updating rules are available. Hence, the G-ASIS model contains 36 adaptive processes. For all 36 instances, we derived a relation for the prevalence in the metastable state. We also showed that the relation between the epidemic threshold τ_c and the effective link-breaking rate ω is linear for 6 instances and constant for 9 instances, but could not be determined for the remaining 21 instances.

One possible method to further determine the relation between the epidemic threshold τ_c and the effective link-breaking rate ω, is using mean-field approximations. The next chapter explores mean-field approximations of the G-SIS model in further detail.

5

THE GENERALISED ADAPTIVE SIS PROCESS – MEAN-FIELD

The Generalised Adaptive SIS (G-ASIS) model from Chapter 4 extends the classical SIS contagious process to a time-evolving network. The network is assumed to change based on decisions of nodes to create or break connections with their susceptible or infected neighbours. Unfortunately, the complexity of the Markovian G-ASIS model is high and analytic results are scarce. To overcome the high complexity, we focus in this chapter on mean-field approximations. Our contribution is fourfold. First, we rigorously derive the first-order and second-order mean-field approximations from the continuous-time Markov chain. Second, we illustrate that the first-order mean-field approximation fails to approximate the epidemic threshold of the Markovian G-ASIS model accurately. Third, we show that the second-order mean-field approximation is a qualitative good approximation of the Markovian G-ASIS model. Finally, we show that the AID model exhibits a bistable metastable state, which contrasts most other G-ASIS instances that show a stable metastable state. Our theoretical results are supported by numerical simulations.

This chapter is based on M. A. Achterberg and P. Van Mieghem, *Moment closure approximations of susceptible-infected-susceptible epidemics on adaptive networks*, Physical Review E 106, 014308, Jul 2022 [103].

5.1. INTRODUCTION

The Generalised Adaptive SIS (G-ASIS) model was introduced in Chapter 4 to describe the spread of contagious phenomena on an adaptive network. Besides exact lower bounds on the epidemic threshold and implicit solutions for the metastable prevalence, no exact results could be derived for the Markovian G-ASIS model. Even though the G-ASIS model is a Markov chain, which is a linear process, exact solutions are scarce due to the exponentially large number of states in the Markov chain.

Several methods are known in the literature to approximate exponentially large state spaces. One important method are mean-field approximations, which assume that the states of (groups of) nodes are uncorrelated [104]. Mean-field models were widely popularised in network epidemiology by Pastor-Satorras and Vespignani in 2001, when they illustrated that in scale-free networks, the epidemic threshold converges to zero for infinitely large networks [8]. Ever since, many mean-field methods have been proposed for the static SIS model, including first-order mean-field [30, 56] and second-order mean-field approximations [105, 106]. The accuracy of the first-order mean-field approximation for the static SIS model has been investigated [32], but the determination of the accuracy of higher order mean-field approximations remains a challenging open problem [107].

Mean-field approximations in the G-ASIS model have already been explored by Kiss *et al.* [98], but no rigorous derivation from the Markovian model was given. Szabó *et al.* [108] provided a detailed bifurcation analysis for the second-order mean-field approximation of the G-ASIS model, but only for the case that links cannot be broken nor created between two infected nodes. Szabó *et al.* additionally proved that the number of steady states is maximally three and derived conditions when such number of steady states exists. A variation on the second-order mean-field approximation of the ASIS model has been studied by Szabó-Solticzky *et al.* [109].

In this chapter, we complement the aforementioned analysis by providing rigorous derivations by applying a first-order and second-order mean-field approximation to the Markovian G-ASIS model. Section 5.2 recalls the G-ASIS model. Section 5.3 discusses the first-order mean-field approximation for the G-ASIS model and shows that the first-order mean field fails to mimic the Markovian G-ASIS model. Contrary to the first-order mean field, the second-order mean-field approximation, derived in Section 5.4, is shown to be considerably more accurate. In Section 5.5, the first- and second-order mean-field approximations are compared with the Markovian G-ASIS model by numerical simulations. Finally, we present our conclusion and outlook in Section 5.6.

5.2. THE G-ASIS MODEL

Recall that the governing equations for the G-ASIS model are given by

$$\frac{d\,\mathrm{E}[X_i]}{dt} = \mathrm{E}\left[-\delta_i X_i + (1-X_i)\sum_{j=1}^{N}\beta_{ij}X_j a_{ij}\right], \tag{5.1a}$$

$$\frac{d\,\mathrm{E}[a_{ij}]}{dt} = \begin{cases} \mathrm{E}\Big[-\zeta_{ij}a_{ij}f_{\mathrm{br}}(X_i,X_j)+\xi_{ij}(1-a_{ij})f_{\mathrm{cr}}(X_i,X_j)\Big], & \text{if } (i,j)\in\mathcal{L}_{\mathrm{adaptive}} \\ 0, & \text{otherwise} \end{cases} \tag{5.1b}$$

where f_{br} and f_{cr} specify the link-breaking and link-creation mechanisms, respectively. All six non-trivial updating rules for the link-breaking f_{br} and link-creation f_{cr} mechanisms have been identified in Table 4.1 and can be written in the form

$$f(X_i,X_j) = a + b(X_i+X_j) + cX_iX_j, \tag{5.2}$$

where $a,b,c\in\mathbb{Z}$. The parameters a, b and c of the six non-trivial updating rules are listed in Table 4.1. Using Eq. (5.2), the governing equations (5.1) become

$$\frac{d\,\mathrm{E}[X_i]}{dt} = \mathrm{E}\left[-\delta_i X_i + (1-X_i)\sum_{j=1}^{N}\beta_{ij}X_j a_{ij}\right], \tag{5.3a}$$

$$\frac{d\,\mathrm{E}[a_{ij}]}{dt} = \begin{cases} \mathrm{E}\Big[-\zeta_{ij}a_{ij}(a_{\mathrm{br}}+b_{\mathrm{br}}(X_i+X_j)+c_{\mathrm{br}}X_iX_j) \\ \qquad +\xi_{ij}(1-a_{ij})(a_{\mathrm{cr}}+b_{\mathrm{cr}}(X_i+X_j)+c_{\mathrm{cr}}X_iX_j)\Big], & \text{if } (i,j)\in\mathcal{L}_{\mathrm{adaptive}} \\ 0. & \text{otherwise} \end{cases} \tag{5.3b}$$

Equations (5.3a) and (5.3b) describe the most general version of the G-ASIS model with heterogeneous infection, curing, link-breaking and link-creation rates. The fact that some links do not adhere to the link-breaking and link-creation dynamics (that either remain existent or non-existent for all times), is reflected by the last line of Eq. (5.3b).

5.3. FIRST-ORDER MEAN-FIELD APPROXIMATION

Even though the Markovian G-ASIS model (5.3) is a simple description of spreading processes on adaptive networks, its analysis is difficult. Each of the 36 instances of the G-ASIS model can be described by a Markov chain with $2^{N+L_{\mathrm{adaptive}}}$ states, which makes the computation for any connected graph with more than $N=20$ nodes infeasible. Only in some special cases, like the adaptive complete graph and the adaptive star graph [110], the huge state space can be exactly reduced using equitable partitions [111].

For all other graphs, the huge state space of the Markov chain can be approximated using mean-field approximations. Mean-field approximations constitute of one or more closure relations, which describe how the higher-order moments of the random variables in the process are approximated by lower-order moments of these random variables [112]. In contrast to the linear Markovian equations, the resulting mean-field equations are non-linear. Mean-field approximations induce an error, but also significantly reduce the dimensionality of the process [53].

A common mean-field approximation for static networks is the Heterogeneous Mean-Field (HMF) approximation [8], which is a first-order mean-field approximation that additionally considers a topological approximation by aggregating all nodes with the same degree in the same group. The HMF approximation is extended to adaptive networks by Marceau *et al.* [85] for a link-rewiring model and by Demirel *et al.* [94] for growing networks. The approximation by Marceau *et al.* not only considers the number of neighbours of each node, but also includes the number of *infected* neighbours in the mean-field approximation, thereby improving on the standard HMF approximation. We expect that a similar HMF approximation can be derived for the G-ASIS model using the framework of Devriendt and Van Mieghem [30].

However, the HMF approximation appears inferior [74] to the first-order mean-field approximation without any topological approximation, also known as the N-Intertwined Mean-Field Approximation (NIMFA) [29]. Thus, we focus on NIMFA from here onwards. NIMFA assumes that any pair of random variables X_i, X_j and a_{ij} is independent (hence, uncorrelated):

$$\begin{aligned}
\mathrm{E}[X_i X_j] &= \mathrm{E}[X_i]\mathrm{E}[X_j], \\
\mathrm{E}[X_i a_{ij}] &= \mathrm{E}[X_i]\mathrm{E}[a_{ij}], \\
\mathrm{E}[X_i X_j a_{ij}] &= \mathrm{E}[X_i]\mathrm{E}[X_j]\mathrm{E}[a_{ij}],
\end{aligned}$$

for all $i \neq j$. The first-order mean-field equations for G-ASIS are then given by

$$\frac{d\,\mathrm{E}[X_i]}{dt} = -\delta_i \mathrm{E}[X_i] + (1-\mathrm{E}[X_i]) \sum_{j=1}^{N} \beta_{ij} \mathrm{E}[X_j]\mathrm{E}[a_{ij}], \tag{5.4a}$$

$$\frac{d\,\mathrm{E}[a_{ij}]}{dt} = \begin{cases} -\zeta_{ij}\mathrm{E}[a_{ij}](a_{\mathrm{br}} + b_{\mathrm{br}}(\mathrm{E}[X_i]+\mathrm{E}[X_j]) + c_{\mathrm{br}}\mathrm{E}[X_i]\mathrm{E}[X_j]) \\ \quad +\xi_{ij}(1-\mathrm{E}[a_{ij}])(a_{\mathrm{cr}} + b_{\mathrm{cr}}(\mathrm{E}[X_i]+\mathrm{E}[X_j]) + c_{\mathrm{cr}}\mathrm{E}[X_i]\mathrm{E}[X_j]), & \text{if } (i,j) \in \mathcal{L}_{\text{adaptive}} \\ 0. & \text{otherwise} \end{cases} \tag{5.4b}$$

Although the number of equations in Eq. (5.4) is $N + L_{\text{adaptive}}$ and not N, we call (5.4) the adaptive N-Intertwined Mean-Field Approximation (aNIMFA), because of the analogy to the NIMFA equations for static networks [29]. Contrary to the NIMFA equations for the static SIS model, aNIMFA is not necessarily an upper bound for the Markovian dynamics [68].

The steady state of the NIMFA equations and the metastable state of the Markov process show similar behaviour for sufficiently large networks and for effective infection rates $\tau = \beta/\delta$ above the epidemic threshold [113]. For adaptive networks, however, we will show that the steady state of aNIMFA and the metastable state of the Markov process can deviate significantly. One of the reasons is as follows. The curing process in the SIS model only involves the state of the node itself. Any mean-field method will therefore exactly capture the curing process, because the assumed independence of random variables is irrelevant for the curing process. On the contrary, the joint probability of infection of n nodes in the network depends on the joint probabilities of infection of $n+1$ nodes. Any mean-field method, irrespective of its order (smaller than N), will approximate the infection process and induce an approximation error. On the other hand, the

link-breaking and link-creation processes involve the state of a link plus the states of the adjacent nodes. By using a first-order mean-field approximation, the link-creation and link-breaking processes will be approximated. In Section 5.4, we will construct a second-order mean-field approximation that only involves the approximation of the infection process and exactly captures the link-breaking and link-creation processes.

5.3.1. FIRST-ORDER MEAN FIELD ON THE COMPLETE GRAPH

The inaccuracy of the first-order mean-field approximation is exemplified by the easiest case, in which the infection, curing, link-breaking and link-creation rates are homogeneous parameters and the adaptive graph $\mathcal{L}_{\text{adaptive}}$ is the complete graph, such that $\mathcal{L}_0 = \mathcal{L}_1 = \emptyset$. Then, the aNIMFA equations become

$$\frac{d\,\mathrm{E}[X_i]}{dt} = -\delta \mathrm{E}[X_i] + \beta(1-\mathrm{E}[X_i]) \sum_{j=1, j\neq i}^{N} \mathrm{E}[X_j]\mathrm{E}[a_{ij}], \tag{5.5a}$$

$$\begin{aligned}\frac{d\,\mathrm{E}[a_{ij}]}{dt} = &-\zeta \mathrm{E}[a_{ij}](a_{\mathrm{br}} + b_{\mathrm{br}}(\mathrm{E}[X_i] + \mathrm{E}[X_j]) + c_{\mathrm{br}}\mathrm{E}[X_i]\mathrm{E}[X_j]) + \\ &\xi(1-\mathrm{E}[a_{ij}])(a_{\mathrm{cr}} + b_{\mathrm{cr}}(\mathrm{E}[X_i] + \mathrm{E}[X_j]) + c_{\mathrm{cr}}\mathrm{E}[X_i]\mathrm{E}[X_j]).\end{aligned} \tag{5.5b}$$

If the initial prevalence is the same for every node and the initial link-density is the same for every link, Eq. (5.5) can be simplified. Introducing the average fraction of infected nodes, also known as the prevalence, as $y = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}[X_i]$, the average link density $z = \frac{1}{N(N-1)} \sum_{i=1}^{N}\sum_{j=1}^{N} \mathrm{E}[a_{ij}]$, rescaling time by $\tilde{t} = t\delta$, defining $\tau = \beta/\delta$, $\tilde{\zeta} = \zeta/\delta$, $\tilde{\xi} = \xi/\delta$ and introducing the normalized effective infection rate $x = \tau(N-1)$, we obtain (dropping the tildes)

$$\frac{dy}{dt} = -y + x(1-y)yz, \tag{5.6a}$$

$$\frac{dz}{dt} = -\zeta z(a_{\mathrm{br}} + 2b_{\mathrm{br}}y + c_{\mathrm{br}}y^2) + \xi(1-z)(a_{\mathrm{cr}} + 2b_{\mathrm{cr}}y + c_{\mathrm{cr}}y^2). \tag{5.6b}$$

After substituting one of the 36 instances of the G-ASIS model, the two differential equations (5.6) with 3 parameters τ, ζ and ξ provide a first-order mean-field description of the G-ASIS model. The steady-state prevalence y_∞ and the steady-state link density z_∞ of Eq. (5.6) are given in Theorem 5.1.

Theorem 5.1 *The steady states (y_∞, z_∞) of equation (5.6) are the real-valued solutions of the cubic equation*

$$\begin{aligned}&c_{\mathrm{cr}}x\,y_\infty^3 + (2b_{\mathrm{cr}}x - c_{\mathrm{cr}}x + c_{\mathrm{cr}} + c_{\mathrm{br}}\omega)\,y_\infty^2 + \\ &\quad (2b_{\mathrm{cr}} + a_{\mathrm{cr}}x - 2b_{\mathrm{cr}}x + 2b_{\mathrm{br}}\omega)\,y_\infty + (a_{\mathrm{cr}} + a_{\mathrm{br}}\omega - a_{\mathrm{cr}}x) = 0\end{aligned} \tag{5.7}$$

where $\omega = \zeta/\xi$ is the effective link-breaking rate and z_∞ follows as

$$z_\infty = \frac{1}{x(1-y_\infty)}, \tag{5.8}$$

5

or the steady state equals the trivial (all-healthy) steady state

$$y_\infty = 0 \quad \text{and} \quad z_\infty = \begin{cases} \frac{a_{\text{cr}}}{a_{\text{br}}\omega + a_{\text{cr}}}, & \text{if } a_{\text{cr}} \neq 0 \text{ or } a_{\text{br}} \neq 0 \\ c, & \text{otherwise} \end{cases} \tag{5.9}$$

for any $c \in [0,1]$.

Proof. See Appendix D.1. □

Although for $a_{\text{br}} = a_{\text{cr}} = 0$, there are infinitely many steady states with prevalence $y_\infty = 0$ and link density z_∞, we will continue to call those states the *trivial steady state*, in line with classical SIS epidemics on static networks.

The existence of the trivial steady state $y_\infty = 0$ is illustrated by Theorem 5.1. In addition to the trivial steady state, we show in Theorem 5.2 that at least one non-trivial steady state exists for all instances in the G-ASIS model.

Theorem 5.2 *For each instance of the G-ASIS model, there is a non-empty* (x,ω)*-region where at least one non-trivial steady state exists.*

Proof. See Appendix D.2. □

Theorem 5.2 guarantees that the introduction of link-breaking and link-creation mechanisms to the standard SIS model is not able to destroy the endemic state completely. Moreover, the following relation for the mean-field epidemic threshold $\tau_c^{(1)}$ follows from the proof of Theorem 5.2.

Corollary 5.3 *For G-ASIS instances whose link-breaking* $f_{\text{br}}(y)$ *and link-creation* $f_{\text{cr}}(y)$ *mechanisms do not have coinciding zeros (see part (ii) in Proof 1 in Appendix D.2 for details), the first-order mean-field epidemic threshold equals*

$$\tau_c^{(1)} = \frac{1}{N-1} \frac{a_{\text{cr}} + a_{\text{br}}\omega}{a_{\text{cr}}}.$$

In the sequel, we consider some example instances of the G-ASIS model.

5.3.2. THE ASIS MODEL

The Adaptive SIS (ASIS) model describes the tendency of healthy people to prevent themselves from contracting the disease by avoiding contact with infected individuals. In the ASIS model, links can be broken between susceptible and infected nodes to prevent the disease from spreading and links can be created between susceptible nodes. Substituting the model parameters of ASIS (see Table 4.1) in equation (5.7) yields

$$x y_\infty^3 + (1 - 3x - 2\omega)\, y_\infty^2 + (-2 + 3x + 2\omega)\, y_\infty + (1 - x) = 0. \tag{5.10}$$

The solution $y_\infty = 1$ is not a valid steady state (according to Eq. (5.6a)) and can be removed. Dividing the polynomial in Eq. (5.10) by $y_\infty - 1$, reduces to the quadratic equation

$$x y_\infty^2 + (1 - 2x - 2\omega)\, y_\infty + (x - 1) = 0,$$

whose solutions are

$$y_\infty = 1 - \frac{1-2\omega}{2x} \pm \sqrt{\left(\frac{1-2\omega}{2x}\right)^2 + \frac{2\omega}{x}}. \tag{5.11}$$

The positive branch of (5.11) is infeasible, because y_∞ would be larger than one. For the steady-state solution y_∞ to exist, the expression under the square root in (5.11) must be non-negative and y_∞ must be bounded between zero and one. If one of these criteria is exactly satisfied, thus the expression under the square root is zero or y_∞ is either zero or one, the resulting condition exactly specifies when the solution y_∞ exists or not. In other words, the existence of y_∞ is described by a bifurcation parameter, also known as the epidemic threshold. Using the relation $x = \tau(N-1)$, the epidemic threshold for the ASIS model follows as

$$\tau_c^{(1),\mathrm{ASIS}} = \frac{1}{N-1},$$

and is independent of the effective link-breaking rate ω. To summarise, the solution is

$$y_\infty = \begin{cases} 1 - \frac{1-2\omega}{2\tau(N-1)} - \sqrt{\left(\frac{1-2\omega}{2\tau(N-1)}\right)^2 + \frac{2\omega}{\tau(N-1)}}, & \tau \geq \tau_c^{(1),\mathrm{ASIS}} = \frac{1}{N-1} \\ 0. & \text{always} \end{cases} \tag{5.12}$$

The steady-state solutions y_∞ are shown for various ω-values in Figure 5.1. Applying linear stability analysis (see Chapter 6 for details), we find that the all-healthy state $y_\infty = 0$ is stable for $\tau \leq \tau_c$ and is unstable otherwise. If the endemic state exists, it is always stable.

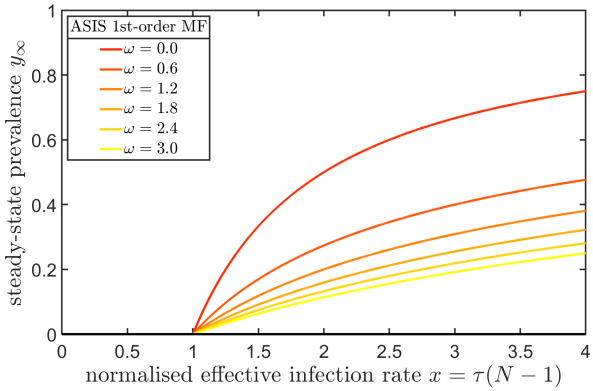


Figure 5.1: The transcritical bifurcation in the first-order mean-field ASIS model with $N = 40$ and $\xi = 0.5$. The epidemic threshold τ_c is fixed for varying effective link-breaking rates ω, which contrasts the Markovian ASIS model, where the epidemic threshold τ_c appears to scale linearly with the effective link-breaking rate ω.

5.3.3. THE AID MODEL

The Adaptive Information Diffusion (AID) model describes the spread of information among people. Links are created between susceptible (informationless) nodes and in-

fected (informative) nodes to enhance the spread of the news. Links are removed between susceptible nodes because both nodes are no longer interested to upkeep their relationship. Substituting the parameters from the AID model (see Table 4.1) into equation (5.7) yields

$$-2xy_\infty^3 + (4x-2+\omega)\,y_\infty^2 + (2-2x-2\omega)\,y_\infty + \omega = 0.$$

As before, the solution $y_\infty = 1$ is an invalid steady state. By removing $y_\infty = 1$, the cubic equation simplifies to the quadratic equation

$$2xy_\infty^2 + (-2x+2-\omega)\,y_\infty + \omega = 0,$$

whose solutions are

$$y_\infty = \frac{2x+\omega-2 \pm \sqrt{(2x+\omega-2)^2 - 8x\omega}}{4x}. \tag{5.13}$$

The epidemic threshold τ_c can be determined by checking when the steady-state solution y_∞ is bounded between zero and one and is real-valued. The epidemic threshold follows as

$$\tau_c^{(1),\mathrm{AID}} = \frac{\frac{1}{2}(\omega+2) + \sqrt{2\omega}}{N-1}. \tag{5.14}$$

The bifurcation diagram for the AID model is shown in Figure 5.2. The main difference between the ASIS and the AID model is that the epidemic threshold in the AID model increases for increasing ω, whereas the epidemic threshold remains constant for the ASIS model. The steady-state solution y_∞ is zero below the epidemic threshold and is non-zero at the epidemic threshold:

$$y_\infty\left(\tau_c^{(1),\mathrm{AID}}\right) = \frac{\omega+\sqrt{2\omega}}{\omega+2+\sqrt{8\omega}}.$$

To summarise, the solution is

$$y_\infty = \begin{cases} \frac{2\tau(N-1)+\omega-2 \pm \sqrt{(2\tau(N-1)+\omega-2)^2 - 8\tau(N-1)\omega}}{4\tau(N-1)}, & \tau \geq \tau_c^{(1),\mathrm{AID}} = \frac{\frac{1}{2}(\omega+2)+\sqrt{2\omega}}{N-1} \\ 0. & \text{always} \end{cases} \tag{5.15}$$

The bifurcation diagram in Figure 5.2 shows two non-trivial steady states, of which one the upper one is stable (solid line) and the other is unstable (dashed line). The stability of each branch was determined using linear stability analysis. The trivial steady state is always stable in the AID model. The existence of the two non-trivial steady states is illustrated in Figure 5.3 using different initial conditions.

An intriguing observation is that the prevalence y of the AID model below the epidemic threshold τ_c is zero, whereas the prevalence is non-zero while approaching the epidemic threshold from above. Similar behaviour was e.g. observed in rewiring models for SIR epidemics [114].

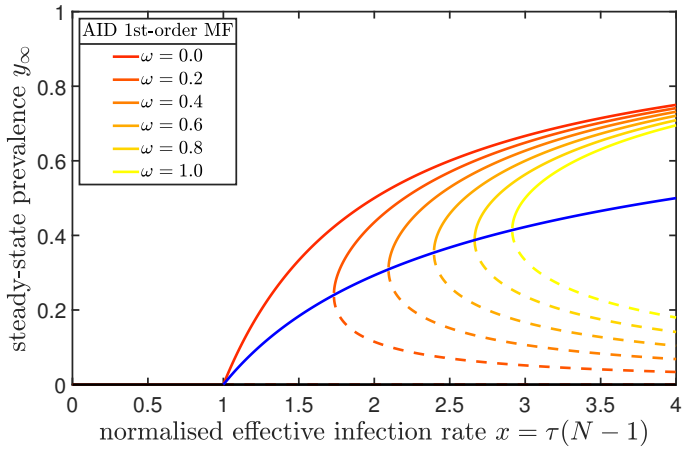


Figure 5.2: The saddle-node bifurcation in the first-order mean-field AID model with $N = 40$ and $\xi = 0.5$. The epidemic threshold τ_c moves over the blue line as the effective link-breaking rate ω increases. The Markovian AID model has a nearly constant epidemic threshold τ_c, but the mean-field AID threshold linearly depends on the effective link-breaking rate ω.

5

5.3.4. THE ABN MODEL

The ABN model describes the spread of information in the human brain. Each region is active (infected) or inactive (susceptible). The ABN model assumes that links can be created between susceptible (inactive) nodes and links are removed between infected (active) nodes. Substituting the parameters of the ABN model (see Table 4.1) yields

$$x y_\infty^3 + (1 - 3x + \omega)\, y_\infty^2 + (3x - 2)\, y_\infty + (1 - x) = 0. \tag{5.16}$$

Unfortunately, like many instances of the G-ASIS model, the cubic equation (5.16) cannot be further simplified. Using Corollary 5.3, the epidemic threshold τ_c follows as

$$\tau_c^{(1),\mathrm{ABN}} = \frac{1}{N-1}, \tag{5.17}$$

which agrees with the numerical results from Figure 5.4. The results are similar to the static SIS and ASIS model, which show the existence of a stable endemic steady state above the epidemic threshold. The stability of the all-healthy state changes at the epidemic threshold τ_c from stable to unstable, leading to a transcritical bifurcation, as visualised in Figure 5.4.

5.4. SECOND-ORDER MEAN-FIELD APPROXIMATION

The first-order mean-field approximation, discussed in Section 5.3, assumes that any pair of random variables is uncorrelated. In this section, we derive a higher-order mean-field approximation, which assumes that pairs, triplets, etc. of random variables are uncorrelated.

Confining ourselves to homogeneous infection, curing, link-breaking and link-creation rates and all links in the graph adhering to the link-breaking and link-creation mecha-

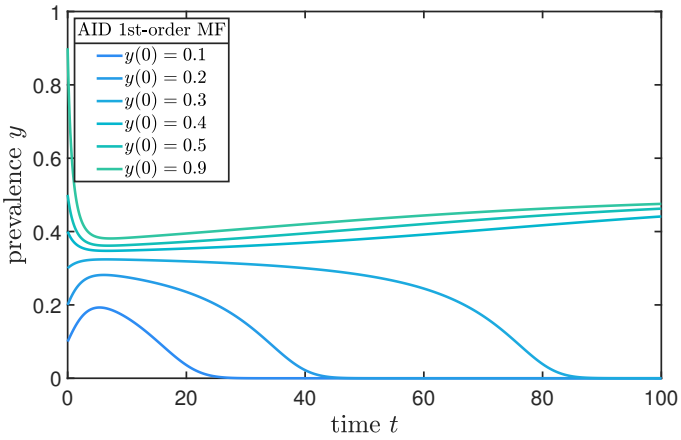


Figure 5.3: Numerical solutions of the first-order mean-field AID model using different initial conditions for the fraction of infected nodes, $z(0) = 0.5$, $\xi = \zeta = 0.1$ and $\tau = \frac{3}{N-1}$. As indicated by the bifurcation diagram in Figure 5.2, there are two stable steady states: $y_\infty = 0$ and $y_\infty \approx 0.5$. There is also an unstable steady state at $y_\infty \approx 0.35$.

nisms, the governing equations (5.3) simplify to

$$\frac{d\,\mathrm{E}[X_i]}{dt} = \mathrm{E}\left[-\delta X_i + \beta(1-X_i)\sum_{j=1}^{N} X_j a_{ij}\right], \tag{5.18a}$$

$$\begin{aligned}\frac{d\,\mathrm{E}[a_{ij}]}{dt} = \mathrm{E}\Big[&-\zeta a_{ij}(a_{\mathrm{br}} + b_{\mathrm{br}}(X_i + X_j) + c_{\mathrm{br}} X_i X_j) \\ &+ \xi(1-a_{ij})(a_{\mathrm{cr}} + b_{\mathrm{cr}}(X_i + X_j) + c_{\mathrm{cr}} X_i X_j)\Big].\end{aligned} \tag{5.18b}$$

Using the closure relation

$$\mathrm{E}[X_i X_j] \approx \mathrm{E}[X_i]\mathrm{E}[X_j], \tag{5.19}$$

the following governing equations can be derived (see Appendix D.3 for the derivation)

$$\begin{aligned}
\frac{dy}{dt} &= \tau\frac{N-1}{2} z_{\mathrm{SI}} - y, \\
\frac{dz_{\mathrm{SS}}}{dt} &= z_{\mathrm{SI}} - \tau(N-2) z_{\mathrm{SSI}} + \xi_{\mathrm{SS}}\left(\frac{N}{N-1}(1-y)^2 - \frac{1}{N-1}(1-y) - z_{\mathrm{SS}}\right) - \zeta_{\mathrm{SS}} z_{\mathrm{SS}}, \\
\frac{dz_{\mathrm{II}}}{dt} &= \tau z_{\mathrm{SI}} + \tau(N-2) z_{\mathrm{ISI}} - 2z_{\mathrm{II}} + \xi_{\mathrm{II}}\left(\frac{N}{N-1} y^2 - \frac{1}{N-1} y - z_{\mathrm{II}}\right) - \zeta_{\mathrm{II}} z_{\mathrm{II}}, \\
\frac{dz_{\mathrm{SI}}}{dt} &= -(1+\tau) z_{\mathrm{SI}} + \tau(N-2) z_{\mathrm{SSI}} - \tau(N-2) z_{\mathrm{ISI}} + 2z_{\mathrm{II}} + \xi_{\mathrm{SI}}\left(\frac{2N}{N-1} y(1-y) - z_{\mathrm{SI}}\right) - \zeta_{\mathrm{SI}} z_{\mathrm{SI}},
\end{aligned} \tag{5.20}$$

where y denotes the fraction of infected nodes, $z_{\mathrm{SS}}, z_{\mathrm{SI}}$ and z_{II} denote the fraction of links in the graph between susceptible-susceptible (S-S), susceptible-infected (S-I) and infected-infected (I-I) pairs of nodes, respectively. Finally, z_{SSI} and z_{ISI} denote the fraction

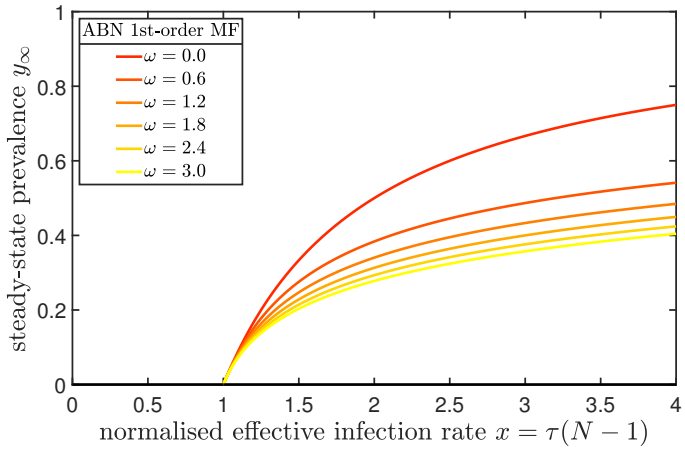


Figure 5.4: The transcritical bifurcation in the first-order mean-field ABN model with $N = 40$ and $\xi = 0.5$. The epidemic threshold τ_c remains constant as the effective link-breaking rate ω increases, which is in agreement with the Markovian ABN model.

of connected S-S-I and I-S-I triples in the graph, respectively. Any other triples are irrelevant, because S-S pairs and S-I pairs can only be infected by an external infected node, not by an external susceptible node. The external infected node must be connected to one of the susceptible nodes in the original node pair, leading to the triplet S-S-I or I-S-I.

Many possible closure relations exist, but in Eq. (5.20), the advantage of our closure relation (5.19) becomes clear: The unknown variables z_{SSI} and z_{ISI} are related to the effective infection rate τ only. Using other closure relations than Eq. (5.19) would lead to an additional approximation of the link-breaking and link-creation processes, which undoubtedly increases the approximation error.

As closure relations for z_{SSI} and z_{ISI}, we use the closure relations from the static SIS model [27, 81, 98, 105], which is derived as follows. We assume that the number of I-S-I triplets equals the number of links between susceptible and infected (S-I) nodes, multiplied by the average number of links between the susceptible node from the considered S-I pair and the remaining infected nodes in the network. The latter equals the number of S-I links divided by the number of susceptible nodes;

$$\frac{1}{2}N(N-1)(N-2)z_{ISI} \approx \frac{1}{2}N(N-1)z_{SI} \cdot \frac{\frac{1}{2}N(N-1)z_{SI}}{N(1-y)}.$$

The same holds for z_{SSI}, except that the infected node can connect to both susceptible nodes of the S-S node pair;

$$\frac{1}{2}N(N-1)(N-2)z_{SSI} \approx \frac{1}{2}N(N-1)z_{SS} \cdot \frac{\frac{1}{2}N(N-1)z_{SI}}{\frac{1}{2}N(1-y)},$$

which can be simplified to

$$z_{\mathrm{ISI}} \approx \frac{1}{2}\frac{N-1}{N-2}\frac{z_{\mathrm{SI}}^2}{1-y}, \qquad z_{\mathrm{SSI}} \approx \frac{N-1}{N-2}\frac{z_{\mathrm{SS}} z_{\mathrm{SI}}}{1-y}. \tag{5.21}$$

Using the closure relations (5.21), we obtain a second-order mean-field approximation of the G-ASIS model:

$$\begin{aligned}
\frac{dy}{dt} &= \tau\frac{N-1}{2} z_{\mathrm{SI}} - y,\\
\frac{dz_{\mathrm{SS}}}{dt} &= z_{\mathrm{SI}} - \tau(N-1)\frac{z_{\mathrm{SS}} z_{\mathrm{SI}}}{1-y} + \xi_{\mathrm{SS}}\left(\frac{N}{N-1}(1-y)^2 - \frac{1}{N-1}(1-y) - z_{\mathrm{SS}}\right) - \zeta_{\mathrm{SS}} z_{\mathrm{SS}},\\
\frac{dz_{\mathrm{II}}}{dt} &= \tau z_{\mathrm{SI}}\left(1 + \frac{N-1}{2}\frac{z_{\mathrm{SI}}}{1-y}\right) - 2z_{\mathrm{II}} + \xi_{\mathrm{II}}\left(\frac{N}{N-1}y^2 - \frac{1}{N-1}y - z_{\mathrm{II}}\right) - \zeta_{\mathrm{II}} z_{\mathrm{II}},\\
\frac{dz_{\mathrm{SI}}}{dt} &= -(1+\tau) z_{\mathrm{SI}} + \tau(N-1)\frac{z_{\mathrm{SI}}}{1-y}\left(z_{\mathrm{SS}} - \frac{1}{2} z_{\mathrm{SI}}\right) + 2z_{\mathrm{II}} + \xi_{\mathrm{SI}}\left(\frac{2N}{N-1}y(1-y) - z_{\mathrm{SI}}\right) - \zeta_{\mathrm{SI}} z_{\mathrm{SI}}.
\end{aligned} \tag{5.22}$$

We would like to intuitively justify equation (5.22), whereby we focus on the equation for the fraction of links between S-S pairs z_{SS}; the other equations follow analogously. The term $\xi_{\mathrm{SS}}\left(\frac{N}{N-1}(1-y)^2 - \frac{1}{N-1}(1-y) - z_{\mathrm{SS}}\right)$ was added for the following reason. The number of S-S links in the network increases with Poisson rate ξ_{SS} based on the number of non-existing links between pairs of susceptible nodes. Given the number of susceptible nodes $N(1-y)$, the maximum number of S-S links is $\frac{1}{2}(N(1-y))(N(1-y)-1)$. Knowing that the maximum number of links equals $\frac{1}{2}N(N-1)$, the maximum fraction of S-S links equals $\frac{N}{N-1}(1-y)^2 - \frac{1}{N-1}(1-y)$. However, we should subtract the fraction of currently active S-S links, which is given by z_{SS}. Additionally, by breaking S-S links with Poisson rate ζ_{SS}, the number of S-S links should decay exponentially to zero, so we subtract $\zeta_{\mathrm{SS}} z_{\mathrm{SS}}$. The other terms are related to the curing and infection process, which we will not explain here.

We derived the second-order mean-field model (5.22) from the original Markovian model (5.18) by the following three steps: (i) we considered homogeneous parameters and a complete adaptive graph, (ii) we assumed that the states of any pair of nodes X_i and X_j are independent (see Eq. (5.19)) and (iii) we approximated the infection process according to Eq. (5.21). The variables z_{ISI} and z_{SSI} describe three *connected* I-S-I and S-S-I triplets, consisting of three random variables for the nodal states and two random variables for the intermediate links. Then z_{ISI} and z_{SSI} are approximated in Eq. (5.21) by z_{SI} and y, which are composed of three and one random variable, respectively. Thus, approximation (iii) is a third-order approximation. At first sight, approximation (ii) in Eq. (5.19) seems to be a first-order closure relation, because we assume no correlation between any two random variables connected to each other by a link. However, the states X_i and X_j of the two nodes i and j do not directly influence each other, but can only propagate via the intermediate link a_{ij}, which itself is a stochastic variable. Thus, we argue that (ii) is actually a second-order closure relation. We conclude that our approximation (5.22) is a second-order *adaptive* mean-field approximation.

5.4.1. SECOND-ORDER MEAN-FIELD APPROXIMATIONS IN THE LITERATURE

Our second-order mean-field approximation (5.22) is not new, but was introduced by Kiss *et al.* [98] and further analysed in [108, 109]. Our contribution constitutes of a rigorous derivation of this second-order approximation, starting from the $2^{N+L_{\text{adaptive}}}$ Markov equations (5.18) towards the second-order mean-field approximation, which we presented in Appendix D.3 and in the previous subsection. We additionally analyse specific instances of the G-ASIS model in detail by comparing their first- and second-order mean-field approximations.

Our notation for the second-order mean-field equations (5.22) differs from other notations from the literature [81, 98, 109]. Most works use $[SS]$, $[II]$ and $[SI]$ to represent (twice) the average number of links between susceptible-susceptible, infected-infected and susceptible-infected nodes, respectively. First, we believe that definitions should be intuitive and should describe the actual number of links; not twice that value. Second, to bring the definitions of $z_{\text{SS}}, z_{\text{II}}$ and z_{SI} in line with the definition of the prevalence y, which is the average *fraction* of infected nodes, we have chosen to normalise all definitions in (D.6) by the maximum number of links, such that $z_{\text{SS}}, z_{\text{II}}$ and z_{SI} specify the fraction of links rather than the number of links.

Our second-order mean-field approximation (5.22) is equivalent to the formulation in Kiss *et al.* [98]. This follows by introducing $[I]$ and $[S]$ as the number of infected (susceptible) nodes and $[SS]$ and $[II]$ as twice the number of links between S-S and I-I node pairs, $[SI]$ as the number of S-I links, the link-breaking rate ω, link-creation rate α, curing rate τ and applying the following transformations to Eqs. (4.1) - (4.4) from Kiss *et al.* [98]

$$\begin{aligned}
[I] &:= Ny, \\
[S] &:= N(1-y), \\
[SS] &:= N(N-1)z_{\text{SS}}, \\
[SI] &:= \frac{1}{2}N(N-1)z_{\text{SI}}, \\
[II] &:= N(N-1)z_{\text{II}}, \\
\omega_{ab} &:= \zeta_{ab}, \\
\alpha_{ab} &:= \xi_{ab}, \\
\gamma &:= 1,
\end{aligned}$$

where a, b are any combination of S and I. Then we exactly recover the second-order mean-field approximation (5.22).

A small difference between [98] and our work is the chosen closure relation. Kiss *et al.* [98] and also Szabó *et al.* [108] consider

$$[ABC] \approx \frac{n-1}{n}\frac{[AB][BC]}{[B]},$$

where A, B and C are random variables and n is the average number of links per node. We consider the simple case where n is sufficiently large, such that $(n-1)/n \approx 1$.

A similar second-order adaptive model, with different nomenclature and terminology, was developed by Heesterbeek and Metz [115]. Their model considers the possibil-

ity of adding and removing of links between nodes with the same and different disease states. However, Heesterbeek and Metz [115] allow at most one connection to each node simultaneously, whereas we allow for each node to connect to all other nodes. Although the assumption of maximally one connection is rather restrictive, the resulting governing equations (S1) and (S2) of Heesterbeek and Metz [115] for the SIS process are similar to Eq. (5.22) under several transformations.

Leung and Diekmann [116] considered a similar situation, in which each node can establish a maximum number of connections, i.e. the maximum degree of each node is bounded. The connections can be established between nodes of similar and different disease states, however, the link-creation rate is assumed to be independent of the disease states of the nodes. We believe that Equation (3.1) in Leung and Diekmann [116] can be straightforwardly extended to allow for adaptive link-breaking and link-creation mechanisms. Still, the key difference is the boundedness of the maximum degree, which is not present in the G-ASIS model.

5.4.2. ANALYSIS OF THE SECOND-ORDER MEAN FIELD

We proceed our analysis of the second-order mean-field approximation by computing the steady states of Eq. (5.22).

Theorem 5.4 *The steady states of system (5.22) are the all-healthy state*

$$y_\infty = z_{\mathrm{II},\infty} = z_{\mathrm{SI},\infty} = 0, \qquad z_{\mathrm{SS},\infty} = \begin{cases} \frac{\xi_{\mathrm{SS}}}{\zeta_{\mathrm{SS}}+\xi_{\mathrm{SS}}}, & \text{if } \zeta_{\mathrm{SS}} \neq 0 \text{ or } \xi_{\mathrm{SS}} \neq 0 \\ c, & \text{otherwise} \end{cases} \tag{5.23}$$

for any $c \in [0,1]$*, or are the solution of the cubic equation*

$$\alpha_1 y_\infty^3 + \alpha_2 y_\infty^2 + \alpha_3 y_\infty + \alpha_4 = 0 \tag{5.24}$$

where the coefficients $\alpha_1, \alpha_2, \alpha_3$ *and* α_4 *depend on the effective infection rate* τ*, the link-breaking rate* ζ*, the link-creation rate* ξ *and the choice of the link-breaking and link-creation mechanisms and are given in Eq. (D.17) in Appendix D.4.*

Proof. See Appendix D.4. □

Although our model (5.22) is equivalent to Szabo *et al.* [108], we find a cubic equation (5.24) for the non-trivial steady states whereas [108] finds a fourth-order equation (see Eq. (11) in [108]). Actually, the trivial steady state $x = 0$ is a solution of their Eq. (11). After removing $x = 0$, their Eq. (11) simplifies to our Eq. (5.24).

Equation (5.23) was obtained earlier for the Markovian G-ASIS model below the epidemic threshold (recall Eq. (4.8)) and for the mean-field model in Section 5.3. Hence, the all-healthy state from all mean-field approximations is in compliance with the all-healthy state of the Markovian G-ASIS model.

Unfortunately, equation (5.24) cannot be further simplified without considering specific link-breaking and link-creation mechanisms. Thus, performing stability analysis or constructing steady-state solutions is hard for the general case. For the case $\zeta_{\mathrm{II}} = \xi_{\mathrm{II}} = 0$, Szabó *et al.* [108, Theorem 2] derived a formula for the epidemic threshold:

$$\tau_c = \frac{\zeta_{\mathrm{SI}} + \xi_{\mathrm{SI}} + 1}{(N-1)\frac{\xi_{\mathrm{SS}}}{\xi_{\mathrm{SS}}+\zeta_{\mathrm{SS}}} + N\xi_{\mathrm{SI}}}.$$

For specific instances of the G-ASIS model, we will provide an extensive analysis as follows. We use Maple to analytically compute the steady states of the cubic equation (5.22) for each of the 36 instances in the G-ASIS model. If the solutions remain very tedious, we determine the (three) steady states numerically.

5.4.3. THE ASIS MODEL

We revisit the ASIS model, in which links can be broken between infected-susceptible pairs and links can be created between susceptible nodes. Hence, the link-breaking rule is $f_{\text{br}} = (X_i - X_j)^2$ and the link-creation rule is $f_{\text{cr}} = (1 - X_i)(1 - X_j)$. Or, in the formulation of this section, $\zeta_{\text{SI}} = \zeta, \xi_{\text{SS}} = \xi$ and $\zeta_{\text{II}} = \zeta_{\text{SS}} = \xi_{\text{II}} = \xi_{\text{SI}} = 0$. Substituting the model parameters of the ASIS model into Eq. (5.24), we find

$$\begin{aligned}(N\tau)\, y_\infty^3 + (-3N\tau + \tau + \xi\omega + 1 - 2\omega)\, y_\infty^2 \\ + (3N\tau - 2\tau - 2\xi\omega - 2 - 2\omega)\, y_\infty + (-N\tau + \tau + \xi\omega + 1) = 0.\end{aligned} \tag{5.25}$$

Equation (5.25) has $y_\infty = 1$ as a solution, which is an invalid steady state according to Eq. (5.22). We can remove the solution $y_\infty = 1$ from Eq. (5.25) to find

$$(N\tau)\, y_\infty^2 + (-2N\tau + \tau + \xi\omega + 1 - 2\omega)\, y_\infty + (N\tau - \tau - \xi\omega - 1) = 0,$$

whose solution is

$$y_\infty = 1 - \frac{\tau + 1 + \xi\omega - 2\omega}{2\tau N} \pm \frac{1}{2\tau N}\sqrt{(\tau + 1 + \xi\omega - 2\omega)^2 + 8\tau N\omega}. \tag{5.26}$$

As before, the solution y_∞ must be real-valued and bounded between zero and one. Then only the negative branch of Eq. (5.26) appears a valid solution. Moreover, we can derive the epidemic threshold as

$$\tau_c^{(2),\text{ASIS}} = \frac{1 + \xi\omega}{N - 1}. \tag{5.27}$$

To sum up, we find

$$y_\infty = \begin{cases} 1 - \frac{\tau+1+\xi\omega-2\omega}{2\tau N} - \frac{1}{2\tau N}\sqrt{(\tau + 1 + \xi\omega - 2\omega)^2 + 8\tau N\omega}, & \text{for } \tau \geq \tau_c^{(2),\text{ASIS}} = \frac{1+\xi\omega}{N-1} \\ 0. & \text{always} \end{cases} \tag{5.28}$$

For various values of the effective link-breaking rate ω, Figure 5.5 shows the bifurcation diagram for the steady-state prevalence y_∞. The epidemic threshold τ_c in Eq. (5.27) scales linearly in the effective link-breaking rate $\omega = \zeta/\xi$, which is also illustrated in Figure 5.5. In contrast to the first-order mean-field threshold $\tau_c^{(1),\text{ASIS}} = \frac{1}{N-1}$, the second-order mean-field threshold $\tau_c^{(2),\text{ASIS}} = \frac{1+\xi\omega}{N-1}$ appears to show the correct linear scaling of the Markovian epidemic threshold. Hence, the second-order mean-field approximation is superior to the first-order mean-field approximation for estimating the epidemic threshold τ_c. We further elaborate on this observation in Section 5.5.

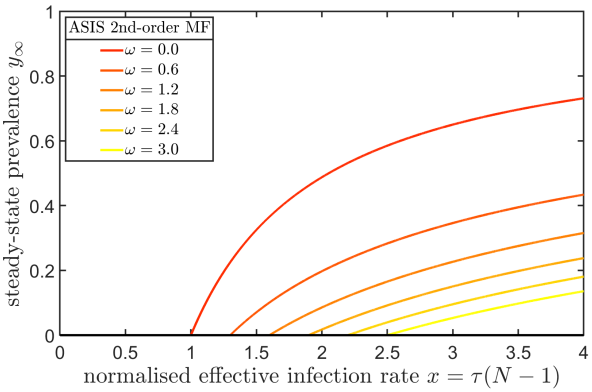


Figure 5.5: The transcritical bifurcation in the second-order mean-field ASIS model with $N = 40$ and $\xi = 0.5$. The epidemic threshold τ_c increases linearly with the effective link-breaking rate ω, which is in compliance with the Markov model.

5

5.4.4. THE AID MODEL

As a second example, we revisit the Adaptive Information Diffusion (AID) model. Similar to the ASIS model, we derive the steady states of Eq. (5.22) by inserting the parameters of the AID model. After the removal of the invalid steady state $y_\infty = 1$, we obtain the following solutions for the steady-state prevalence:

$$y_\infty = \begin{cases} 1 - \frac{2N\tau - \xi\omega + 2 - \omega \pm \sqrt{(2N\tau + \xi\omega - 2 + \omega)^2 - 8N\tau\omega}}{2N\tau(2-\xi\omega)}, & \text{for } 0 < \omega < \omega_c^{\text{AID}} \text{ and } \tau \geq \tau_c^{(2),\text{AID}}(\omega_c^{\text{AID}}) \\ 1 - \frac{2N\tau - \xi\omega + 2 - \omega - \sqrt{(2N\tau + \xi\omega - 2 + \omega)^2 - 8N\tau\omega}}{2N\tau(2-\xi\omega)}, & \text{for } \omega \geq \omega_c^{\text{AID}} \text{ and } \tau > \tau_c^{(2),\text{AID}} \\ 0, & \text{always} \end{cases} \tag{5.29}$$

where we used the effective link-breaking rate $\omega = \zeta/\xi$. We emphasise that $\zeta > 0$ and $\xi > 0$, otherwise the G-ASIS model would not be adaptive. We may compute the epidemic threshold τ_c explicitly in terms of the effective link-breaking rate ω

$$\tau_c^{(2),\text{AID}}(\omega) = \begin{cases} \dfrac{1 + \frac{\omega}{2} - \frac{\xi\omega}{2} + \sqrt{\omega\,(2-\xi\omega)}}{N}, & \text{for } 0 < \omega \leq \omega_c^{\text{AID}} \\ \dfrac{1 + \frac{1}{\xi}}{N}, & \text{for } \omega > \omega_c^{\text{AID}} \end{cases} \tag{5.30}$$

and the steady-state prevalence y_∞ at the epidemic threshold equals

$$y_\infty(\tau = \tau_c^{(2),\text{AID}}) = \begin{cases} 1 - \frac{4 - 2\xi\omega + 2\sqrt{\omega(2-\xi\omega)}}{(2 + \omega - \xi\omega + 2\sqrt{\omega(2-\xi\omega)})(2-\xi\omega)}, & \text{for } 0 < \omega \leq \omega_c^{\text{AID}} \\ 0. & \text{for } \omega > \omega_c^{\text{AID}} \end{cases} \tag{5.31}$$

The critical point ω_c follows by solving $\lim_{\omega \uparrow \omega_c} \tau_c^{(2),\text{AID}}(\omega) = \lim_{\omega \downarrow \omega_c} \tau_c^{(2),\text{AID}}(\omega)$ in (5.30), which leads to

$$\omega_c^{\text{AID}}(\xi) = \frac{2}{\xi(\xi+1)}. \tag{5.32}$$

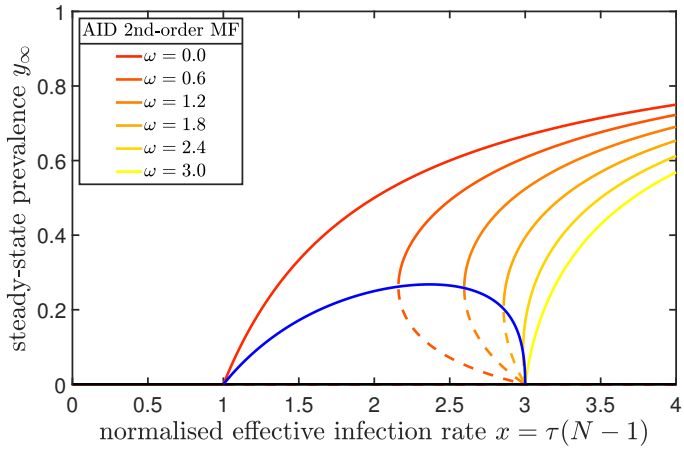


Figure 5.6: The bifurcation diagram for the second-order mean-field AID model with $N = 40$ and $\xi = 0.5$. The critical effective link-breaking rate equals $\omega_c^{\mathrm{AID}} = \frac{8}{3} \approx 2.67$. For $0 < \omega \leq \omega_c$, there is a saddle-node bifurcation at τ_c and a transcritical bifurcation at $x = 3$. For $\omega > \omega_c$, the bifurcation type is a transcritical. The blue line is a parametric curve from 0 to ω_c, where ω_c is given by equation (5.32).

5

Figure 5.6 depicts, as shown by the solution (5.29), three regions of solutions: (I) For $\omega < \omega_c^{\mathrm{AID}}$ and $\tau < \tau_c^{(2),\mathrm{AID}}$, the only steady state is the trivial, all-healthy state. Then (II), by applying a fixed $0 < \omega < \omega_c^{\mathrm{AID}}$, there are three steady states for $\tau_c^{(2),\mathrm{AID}} \leq \tau \leq \tau_c(\omega_c^{\mathrm{AID}})$, of which the upper and lower stable branches are shown in solid lines and the unstable middle branch by dotted lines. Furthermore, there is only one non-trivial stable state for large infection rates $\tau > \tau_c(\omega_c^{\mathrm{AID}})$. Finally (III) we consider the case $\omega > \omega_c$. For $\tau \leq \tau_c$ the only steady state is the all-healthy state. The location of the epidemic threshold is fixed (in terms of the effective link-breaking rate ω). Above the threshold $\tau > \tau_c$, the only stable steady state is the endemic state. Mostly importantly, the second-order mean-field AID model states that the epidemic threshold $\tau_c^{(2),\mathrm{AID}}$ converges to a constant value while the effective link-breaking rate $\omega \to \infty$. This contrasts the first-order mean-field approximation from Section 5.3, where the epidemic threshold τ_c increased up to infinity in the limit of the effective link-breaking rate ω to infinity. Further considerations are given in Section 5.5.

5.4.5. COMPARISON OF AID IN MEAN-FIELD AND MARKOV MODELS

The second-order mean-field approximation was primarily derived to gain a deeper understanding of the Markovian AID model. Trajanovski *et al.* [97] observed spurious oscillations for the Markovian AID model, indicating some kind of instability of the stochastic process. We argue here, albeit hand-waving in nature without providing any rigorous proof, that the metastable state in the Markovian AID model does not fail to exist, but there actually exist two metastable states simultaneously.

Our first reason to believe in the existence of two metastable states, is that the numerical evaluation of the exact, quadratic formula of the prevalence y, provided in Eq. (4)

from [97], reveals that not zero, but two non-trivial solutions y_1 and y_2 exist for the metastable prevalence y. We plot the prevalence y and the computed prevalences y_1 and y_2 in Figure 5.7, where y and y_1 overlap nearly perfectly in the metastable state. We emphasise that the computed prevalences y_1 and y_2 are only exact in the metastable state (when all time-derivatives are zero) and not in the transient regime. Our second reason is more technical and is provided in Appendix D.5.

The behaviour before arrival at the steady state is characterised by the bi-metastability phenomenon; the probability to leave the all-healthy state and the endemic state is both very low. Hence, we believe that the two Markovian prevalences y_1 and y_2 from Figure 5.7 are similar to the two non-trivial prevalences of the second-order mean-field approximation from Figure 5.6, where y_1 corresponds to the stable upper branch and y_2 to the unstable (dashed) lower branch of the second-order mean-field approximation.

Figure 5.7a shows that, starting the process near the endemic state, results in a fast convergence towards the endemic state. Additionally, the infection probability distribution in Figure 5.7b depicts a bell-shaped curve around $y \approx 0.85$. On the other hand, starting with half of the population infected and an empty graph, Figure 5.7c shows that the convergence towards the endemic state takes a longer time. At $t = 500$, the process from Figure 5.7c has not yet converged, because the infection probability distribution in Figure 5.7d is not yet equal to Figure 5.7b.

We provide an example of a much longer convergence time in Figure 5.8. Since our simulations involve only $N = 40$ nodes, we believe that for larger networks for some parameter values, the convergence time[1] might be longer than any feasible simulation time. The long convergence times in the AID model contrast the static SIS model and most other instances of the G-ASIS model, whose convergence times are generally much shorter. The consequence is that most AID epidemic outbreaks on large networks never arrive at the metastable state. Hence, the AID model cannot be fully understood from its steady-state distribution alone and research should focus on its time-dependent behaviour.

Figure 5.8a additionally shows $y(t) \pm \sigma(t)$, where $\sigma(t)$ is the standard deviation of the prevalence $y(t)$ at time t. The spread around the prevalence y is large, because the time-dependent infection probabilities $\Pr[y = j/N]$, plotted in Figure 5.8b, depicts a composition of two bell-shaped curves; one at $y = 0$ and another at $y \approx 0.7$.

The metastability of the all-healthy state is caused by the link-breaking and link-creation mechanism of the AID model, which break links between susceptible nodes and create links between susceptible-infected pairs. The all-healthy state corresponds to zero infected nodes and an empty graph. For an outbreak to occur, links must be created and the disease must spread simultaneously, whereas outbreaks in the SIS, ASIS and ABN model are initiated with a single infected node and a completely connected graph, which allows for an easier spread of the disease, because the links already exist in the graph.

We finalise our analysis of the second-order mean-field AID model by showing the

[1] To analyse the exact average convergence time, one possible method is to analyse the eigenvalues of the underlying Markov chain, as was performed for ε-SIS dynamics on static networks in Chapter 2. Unfortunately, an eigenvalue analysis for the Markovian G-ASIS model is infeasible due to the exponentially large state space, not even for the complete adaptive graph [110].

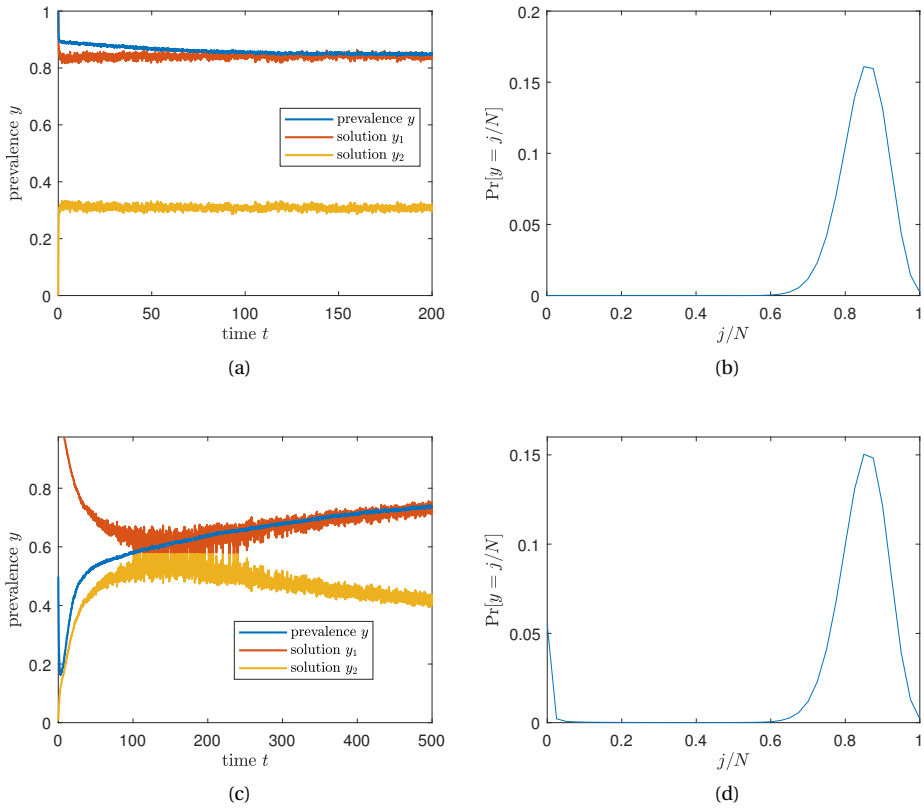


Figure 5.7: Illustration of the two metastable states in the AID model with $N = 40$, $\tau = 0.25, \zeta = 0.5$ and $\xi = 0.1$ for (a,c) the time-varying prevalence $y(t)$ and (b,d) the prevalence distribution $\Pr[y(t) = j/N]$ at time $t = t_{\text{end}}$ where j is the number of infected nodes. The upper plots (a,b) start with a complete graph and all nodes infected whereas the bottom row (c,d) initiates with an empty graph and half of the population infected. The results are averaged over 1,000 simulations and the computed prevalences y_1 and y_2 are based on the quadratic equation (4) from [97], but are only exact in the endemic state. We believe that y_1 is stable and y_2 is an unstable solution, analogous to the second-order mean-field solution from Figure 5.6. We refer to Section 5.5 for a description of our simulation method.

phase diagram in Figure 5.9. In region (I), the only steady state is the disease-free state. Region (II) is the bistable region, where one endemic state and the all-healthy state are stable and another unstable endemic state exists. In region (III), there is a unique, stable steady state and the all-healthy state is unstable.

5.4.6. THE ABN MODEL

The steady states of the Adaptive Brain Network (ABN) model satisfy the cubic equation

$$(N\tau\xi\omega + 2N\tau)\, y_\infty^3 + (-3N\tau\xi\omega - 6N\tau + 2\tau\xi\omega + 2\tau - 2\tau\omega + 2 + 2\omega)\, y_\infty^2 +$$
$$(3N\tau\xi\omega + 6N\tau - 4\tau\xi\omega - 4\tau + 2\tau\omega - \xi\omega - 4)\, y_\infty + (-N\tau\xi\omega - 2N\tau + 2\tau\xi\omega + 2\tau + \xi\omega + 2) = 0.$$

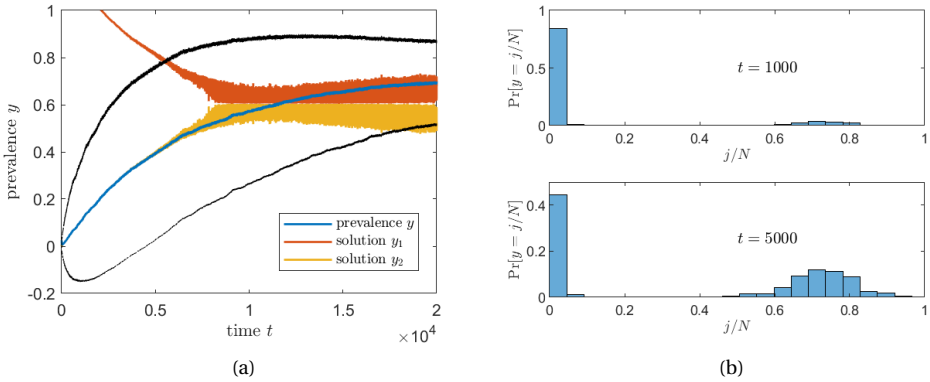


Figure 5.8: An example of very long convergence times in the AID model with $N = 40$, $\tau = 0.18, \zeta = 0.5$ and $\xi = 0.1$. The simulations are initiated with an empty graph and a single infected node and the results are averaged over 1,500 simulations. Subfigure (a) shows the time-dependent prevalence $y(t)$, the computed solutions y_1 and y_2 and the black curves indicate the prevalence $y(t)$ plus/minus one standard deviation $\sigma(t)$. Subfigure (b) shows the prevalence distribution at $t = 1000$ and $t = 5000$.

Unfortunately, the cubic equation cannot be further simplified. A numerical approximation of the steady-state prevalence y_∞ is shown in Figure 5.10. The epidemic threshold τ_c is approximately located at $\tau_c \approx 1/(N-1)$ for all network sizes N, link-breaking rates ζ and link-creation rates ξ.

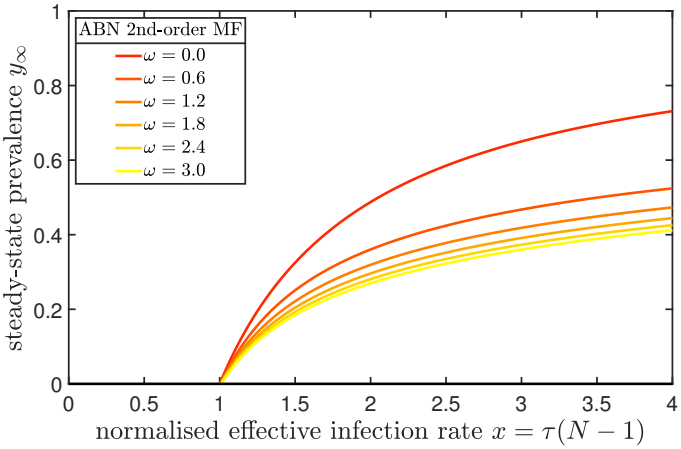


Figure 5.10: The transcritical bifurcation in the second-order mean-field ABN model with $N = 40$ and $\xi = 0.5$. The epidemic threshold is constant for varying ω (different lines) but also for changing ξ (not shown).

5.5. NUMERICAL SIMULATIONS

In this section, we compare the Markovian G-ASIS model with the first-order mean-field approximation from Section 5.3 and second-order mean-field approximation from

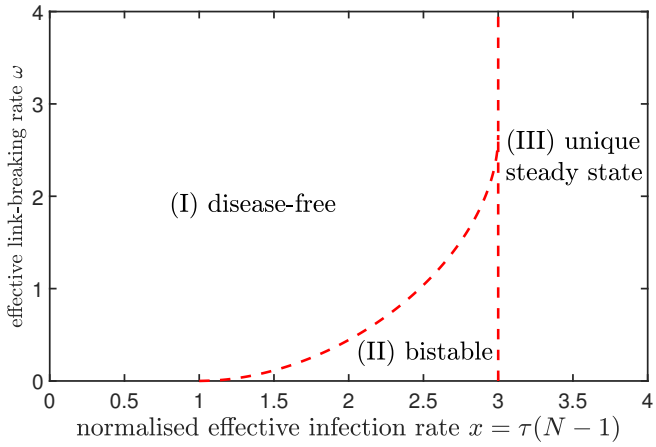


Figure 5.9: The phase diagram for the second-order mean-field AID model with $N = 40$ and $\xi = 0.5$. In region (I) only the disease-free state exists, in (III) there exists a unique stable endemic state (and the disease-free state is unstable) and in (II), there are two non-trivial steady states, of which one is stable and the other unstable. The all-healthy state is stable as well.

Section 5.4. We perform many independent Monte Carlo simulations of the Markovian G-ASIS model, whereby we use the sampled-time Markov chain [44] with time step $\Delta t = 0.05$. At each discrete time step, we compute the probability for each node and each link to change its state. If the probability is larger than a random number between zero and one, the state is changed, and is left unchanged otherwise. We use a small self-infection rate $\varepsilon = 10^{-3}$ for all simulations [16]. Each simulation starts at $t = 0$ and ends at $t = 500$, unless specified otherwise. The metastable prevalence y is determined by averaging over all simulations and over all prevalences from $t = 100$ to $t = 500$. We focus on the relation between the epidemic threshold τ_c and the effective link-breaking rate ω and illustrate this relation for the ASIS, AID and ABN model.

The phase diagram for the ASIS model in Figure 5.11 illustrates that the static SIS model, shown in blue, has a similar accuracy for both the first-order and second-order mean-field approximation. For the ASIS model, however, the simulations are closer to the second-order mean-field approximation than the first-order mean-field approximation, both in terms of the average distance between the two curves as well as the location of the epidemic threshold. If we estimate the epidemic threshold τ_c from the simulations as the smallest effective infection rate τ for which the steady-state prevalence y_∞ exceeds $1/N$, then the estimated threshold τ_c is much closer to the second-order mean-field threshold $\tau_c^{(2),\mathrm{ASIS}}$ than the first-order mean-field threshold $\tau_c^{(1),\mathrm{ASIS}}$.

The phase diagram of the AID model is depicted in Figure 5.12. The inaccuracy of the first-order mean-field approximation is large for the AID model. The first-order mean-field approximation predicts a continuously increasing epidemic threshold τ_c, whereas the second-order mean-field approximation predicts a slightly increasing but strictly bounded threshold. The simulations in Figure 5.12 show that the epidemic threshold τ_c indeed increases a little, but seems to converge to a finite value for increasing effective link-breaking rates ω. The first-order mean field for $\zeta = 25$ is not shown in Figure 5.12,

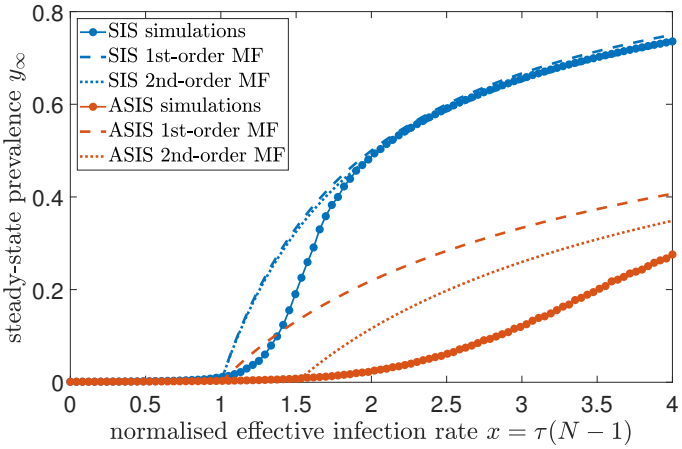


Figure 5.11: The phase diagram for the ASIS model on a complete network with $N = 40$ nodes and $\xi = \zeta = 0.5, \delta = 1$ for the first-order and second-order mean-field approximations and the Markovian result is averaged over 1,000 simulations.

because $\tau_c^{(1),\mathrm{AID}}$ is located approximately at $x = 10$, which is far outside of the figure. Thus for $\zeta = 25$, the second-order approximation is much more accurate than the first-order approximation.

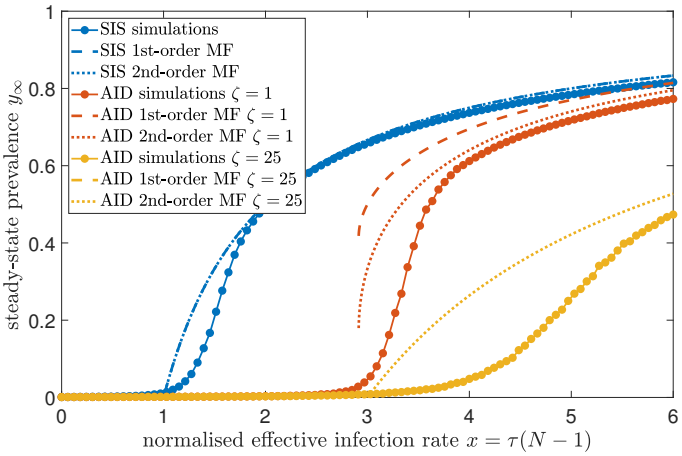


Figure 5.12: The phase diagram for the AID model on a complete network with $N = 40$ nodes and $\xi = 0.5, \delta = 1$ for $\zeta = 1$ and $\zeta = 25$ for the first-order and second-order mean-field approximations and the Markovian result is averaged over 1,000 simulations. The first-order and second-order mean-field for static SIS are indistinguishable.

For the ABN model, both the first-order and second-order mean-field approximation correctly predict a fixed epidemic threshold, which is exemplified in Figure 5.13. In other words, both mean-field approximations correctly capture the epidemic threshold.

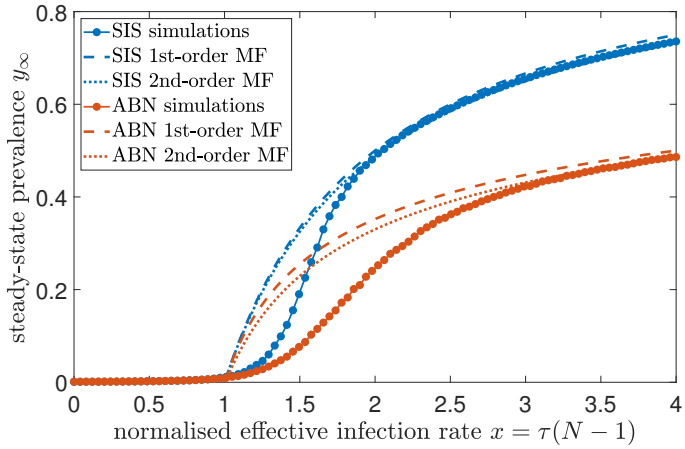


Figure 5.13: The phase diagram for the ABN model on a network with $N = 40$ nodes and $\xi = 0.5, \zeta = 0.5, \delta = 1$ for the first-order and second-order mean-field approximations and the Markovian result is averaged over 1,000 simulations.

5.6. CONCLUSION

In this chapter, we considered various mean-field approximations for the Markovian Generalised Adaptive Susceptible-Infected-Susceptible (G-ASIS) model. We rigorously derived the first-order and second-order mean-field approximations of the G-ASIS model. We discussed two instances of G-ASIS in particular; the ASIS model where nodes prevent themselves from contracting the disease by breaking connections with infected nodes and the AID model that describes the tendency of unaware (healthy) individuals to connect to nodes that are aware (infected) of the gossip or news. We showed that the relation between the epidemic threshold and the effective link-breaking rate is qualitatively captured correctly by the second-order mean-field approximation whereas it is not by the first-order mean-field approximation.

A summary of our results for all possible updating rules for the link-breaking and link-creation mechanisms is presented in Table D.2 in Appendix D.6. In particular, Table D.2 discusses on the relation between the epidemic threshold τ_c and the effective link-breaking rate ω. The first-order approximation is able to predict 6 correct relations and yields 9 wrong relations, whereas the second-order approximation predicts 14 correct relations and only 1 incorrect relation. For the remaining 21 instances, the exact scaling has not been determined in the Markovian model. Overall, we conclude that the second-order approximation is much better at correctly estimating the relation between the epidemic threshold and the effective link-breaking rate, although not for all instances of the G-ASIS model.

Finally, we showed that the Markovian AID model shows opposite behaviour to the classical SIS model in the sense that the convergence time towards the metastable state can be very large, even though the effective infection rate τ is way above the epidemic threshold τ_c. The reason is that both the all-healthy state and the endemic state are stable, causing the system to remain in one of the two metastable states for a significant

amount of time. The bi-metastability phenomenon is in compliance with the second-order mean-field approximation, which also exhibits the bistability phenomenon.

We see several directions for future research on the G-ASIS model. So far, three out of 36 instances of the G-ASIS model have been analysed in detail, but the remaining models remain largely unexplored. Further investigation on the remaining models, especially those models whose relation between the epidemic threshold and the effective link-breaking rate in Table D.2 remains unclear, may enhance our understanding of the interplay between the disease spreading, network topologies and human decision making. Another underestimated research topic is related to the time-varying graph in adaptive epidemics. Most research focusses on the determination of the number of infected cases, but limited attention is devoted to the properties of the time-varying network itself. Even though we believe mean-field methods constitute a powerful tool to enhance our understanding of Markovian epidemic processes, many properties of the underlying network are not captured by mean-field models, complicating the direct analysis of the underlying, time-varying network.

Finally, the G-ASIS model can be extended to include simplicial contagion [72]. Besides the usual spread of an infection over a link, simplicial contagion assumes that the existence of higher-order interactions in the network may speed up or slow down the spread over the links. For example, the 2-simplex considers the interaction between three nodes (a full triangle), of which two nodes are infected and one node is susceptible. The rate at which the susceptible node becomes infected, is no longer the sum of the individual infection rates, but the existence of the triangle creates an additional infection rate, effectively infecting the susceptible node at a higher rate. In addition to extending the SIS equations (4.1) to simplicial contagion, the topology updating rule (4.4) can also be generalised to link-breaking and link-creation mechanisms between D-dimensional simplices.

6

A MINIMAL MODEL FOR ADAPTIVE SIS EPIDEMICS

Even though the first-order mean-field approximation was proven to be a poor approximation of the Markovian G-ASIS model in Chapter 5, the model is astonishingly simple and yet describes the interplay between disease dynamics and personal risk perception. We propose a generalisation of the first-order mean field, called aNIMFA, for general link-breaking and link-creation functional responses. The original aNIMFA model from Chapter 5 was derived from the Markovian G-ASIS model, which imposes natural constraints on the link-breaking and link-creation rules. In this chapter, we drop those constraints and consider the most general version. For simplicity, we focus on the homogeneous case and a complete graph. We derive an explicit form for the basic reproduction number and guarantee the existence of at least one endemic equilibrium, for all possible functional responses. Moreover, we show that for all functional responses, limit cycles do not exist. This means that our minimal model is not able to reproduce consequent waves of an epidemic, and more complex disease or behavioural dynamics are required to reproduce epidemic waves.

This chapter is based on M. A. Achterberg and M. Sensi, *A minimal model for adaptive SIS epidemics*, Nonlinear Dynamics, May 2023 [117].

6.1. INTRODUCTION

Investigating the interplay between disease dynamics and the effect of human responses on the epidemic is in general a complicated task. The G-ASIS model and its mean-field approximations from Chapter 5 illustrated that even under the simplest disease dynamics (a simple SIS model) and extremely simple link-adaptation rules (link-breaking and link-creation) exact analytical results are scarce.

Many models on adaptive epidemics in the literature do also not admit analytical solutions. For example, Sahneh *et al.* [96] developed a multilayer approach, where one layer describes the disease transmission and the second layer describes the awareness of individuals about the disease. Gross *et al.* [81] proposed a rewiring mechanism, which rewires the link between a pair of susceptible-infected nodes to two susceptible nodes. Jolad *et al.* [118] assumes that all individuals have a preferred number of neighbours, subject to random link addition and removals. The preferred degree is taken to be a function of the current number of infected nodes in the network. Brauer [119] discusses an SIR model in which a certain percentage of the links is removed. The removal percentage is larger if the link is connected to infected nodes rather than susceptible nodes. All abovementioned models capture a particular aspect of human behaviour on disease dynamics, but most are so complicated that an exact analysis is completely infeasible. The reason is that the models often involve a large amount of individuals or equations, or the infection probabilities involve the computation of many dependent random variables.

In this chapter, we propose a minimal model consisting of two differential equations, one for the viral prevalence in the population using the NIMFA equations [29], and one for the link density of the contact network. Our model is minimal in terms of the number of equations and parameters, while still capturing key aspects of behavioural disease dynamics. We model the creation and removal of links as an overall increase or decrease of the link density of the contact network. We call the model adaptive NIMFA (aNIMFA), in line with the first-order mean-field approximation from Chapter 5. The novel aspect of aNIMFA are the *functional responses* of individuals to create or break links in the network, based on the current number of infected people. In predator-prey systems like Volterra-Lotka dynamics, Holling introduced functional responses to describe the food intake by predators as a function of the number of available prey [120]. We define similar functional responses to describe the change in the link density, based on the prevalence of the disease in the population.

The aNIMFA model is not limited to modelling epidemic spread, but can be utilised for describing general spreading phenomena, including opinion dynamics, Maki-Thompson rumour spread, innovation spread and epileptic seizures in the human brain. If the dynamics evolve over a network structure and the link density can be modelled by link-breaking and link-creation dynamics, the aNIMFA model can be generalised to such models. In the context of epidemics, one would expect the removal (resp. creation) of links to be directly (resp. inversely) proportional to the prevalence. For other spreading phenomena, such as rumor spreading, this might not be the case, and other choices for the functional responses can be made. The simplicity of the aNIMFA model makes it a promising tool for future generalisations and for the integration of more complex mechanisms.

This chapter is structured as follows. We introduce the aNIMFA model in Section 6.2

and provide a thorough analysis in Section 6.3. Then we consider several examples of functional responses in Section 6.4. Finally, we conclude in Section 6.5.

6.2. THE aNIMFA MODEL

Consider a well-mixed population of N individuals, subject to the spread of a disease. The governing equation for the prevalence y of the SIS process for a well-mixed population is given by

$$\frac{dy}{dt} = -\delta y + \beta y(1-y)z, \tag{6.1}$$

where the curing process is denoted by its rate δ, the infection process by the corresponding rate β and z is the link density of the contact network, i.e. the average fraction of connections between all individuals compared to the total amount of possible connections. In the classical model, the link density z is not varying over time. In the first term on the right-hand side of Eq. (6.1), the prevalence decreases proportional to the current number of infected cases. The second term on the right-hand side in Eq. (6.1) increases the prevalence because of contact between infected y and susceptible $1-y$ nodes. Because of the homogeneous mixing, we multiply with the link density z to obtain the average number of contacts. Equation (6.1) is equivalent to the N-Intertwined Mean-Field Approximation (NIMFA) of a Markovian SIS process on a complete static graph with link weights z, equal initial conditions for all nodes and homogeneous infection and curing rates [29].

Contrary to the static SIS process, we assume that the link density $z(t)$ is varying over time and its dynamics is governed by a link-breaking and a link-creation process. Then the link density $z(t)$ changes over time as

$$\frac{dz}{dt} = -\zeta z f_{\text{br}}(y) + \xi(1-z) f_{\text{cr}}(y), \tag{6.2}$$

where ζ is the link-breaking rate, ξ the link-creation rate and $f_{\text{br}}(y)$ and $f_{\text{cr}}(y)$ are the functional responses to the link-breaking and link-creation process, respectively. The breaking (resp. creation) of links translates into decreasing (increasing) the link density z in Eq. (6.2), implying that f_{br} and f_{cr} must be non-negative. We assume the parameters $\delta, \beta, \zeta, \xi$ to be $\mathcal{O}(1)$ and positive. The link density z has been normalised, such that $z = 1$ is the maximum link density (corresponding to a complete graph) and $z = 0$ corresponds to an empty graph (no connections, so the link density is zero).

Equations (6.1) and (6.2) can be simplified by introducing the scaled time $\tilde{t} = \delta t$. We additionally introduce the *effective infection rate* $\tau = \beta/\delta$. Using the transformations $\tilde{\zeta} = \zeta/\delta$ and $\tilde{\xi} = \xi/\delta$, the well-mixed adaptive NIMFA (aNIMFA) equations are obtained (after dropping the tildes, for ease of notation)

$$\frac{dy}{dt} = -y + \tau y(1-y)z, \tag{6.3a}$$

$$\frac{dz}{dt} = -\zeta z f_{\text{br}}(y) + \xi(1-z) f_{\text{cr}}(y), \tag{6.3b}$$

$$\text{feasible region } 0 \le y \le 1, 0 \le z \le 1.$$

The initial conditions $y(0) \in [0,1]$ and $z(0) \in [0,1]$ describe the initial prevalence and link-density, respectively. We assume that the functional responses $f_{\text{br}}(y)$ and $f_{\text{cr}}(y)$ are non-negative, sufficiently regular functions on the interval $0 \leq y \leq 1$. We exclude the possibility that $f_{\text{br}}(y) = 0$ and $f_{\text{cr}}(y) = 0$ for all y, as in this case, the link density z is not affected by f_{br} and f_{cr} and remains constant over time.

Comparing (6.3) with (5.6), we see the same equations, but here we use general functional responses for the link-breaking and link-creation process, whereas (5.6) is limited to rules derived from the G-ASIS model.

6.3. ANALYSIS OF THE MODEL

Prior to confining ourselves to specific link-breaking and link-creation functions f_{br} and f_{cr}, we first derive general results for the aNIMFA model.

Lemma 6.1 *Consider a solution of system* (6.3) *starting at* $(y(0), z(0)) \in [0,1]^2$. *Recall that* $f_{\text{br}}(y), f_{\text{cr}}(y) \geq 0$ *for all* $y \in [0,1]$. *Then,* $(y(t), z(t)) \in [0,1]^2$ *for all* $t \geq 0$, *meaning* $[0,1]^2$ *is forward invariant for system* (6.3).

Proof. We calculate

$$\left.\frac{dy}{dt}\right|_{y=0} = 0, \quad \left.\frac{dy}{dt}\right|_{y=1} = -1 < 0,$$
$$\left.\frac{dz}{dt}\right|_{z=0} = \xi f_{\text{cr}}(y) \geq 0, \quad \left.\frac{dz}{dt}\right|_{z=1} = -\zeta f_{\text{br}}(y) \leq 0,$$

which proves the forward invariance of $[0,1]^2$. □

6.3.1. DISEASE-FREE EQUILIBRIUM

The aNIMFA model always has one steady state $y_0 = 0$, which corresponds to the situation in which no infected individuals are present in the population. We call this steady state the *disease-free equilibrium* (DFE) or all-healthy state. The DFE of the mean-field equations (6.3) equals

$$y_0 = 0,$$
$$z_0 = \begin{cases} \frac{f_{\text{cr}}(0)}{\omega f_{\text{br}}(0) + f_{\text{cr}}(0)}, & \text{if } f_{\text{br}}(0) \neq 0 \text{ or } f_{\text{cr}}(0) \neq 0 \\ c, & \text{if } f_{\text{br}}(0) = f_{\text{cr}}(0) = 0 \end{cases}$$

for any $c \in [0,1]$ and we introduced the *effective link-breaking rate* $\omega = \zeta/\xi$.

6.3.2. ENDEMIC EQUILIBRIA

Depending on the choice of the functional responses f_{br} and f_{cr}, multiple additional steady states may exist, which are called the *endemic equilibria* (EE). Recall that $\omega = \zeta/\xi$ is the effective link-breaking rate of system (6.3). The endemic equilibria are the solutions of the generally non-linear equation

$$\omega f_{\text{br}}(y_\infty) = (\tau - 1) f_{\text{cr}}(y_\infty) - \tau y_\infty f_{\text{cr}}(y_\infty), \tag{6.4}$$

and the corresponding steady-state link density z_∞ follows as

$$z_\infty = \frac{1}{\tau(1-y_\infty)}. \tag{6.5}$$

We remark that the solution $y_\infty = 1$ is never a valid EE for any functional responses f_{br} and f_{cr}, which follows immediately from substituting $y_\infty = 1$ into Eq. (6.4). Hence, if an EE exist, the steady-state prevalence y_∞ must be in the open interval $(0,1)$. We further investigate the existence of endemic equilibria in Section 6.3.4.

6.3.3. LINEAR STABILITY ANALYSIS

We analyse the linear stability of the steady states by computing the Jacobian of Eq. (6.3) as

$$J = \begin{pmatrix} -1+\tau(1-2y_\infty)z_\infty & \tau y_\infty(1-y_\infty) \\ -\zeta z_\infty f'_{\text{br}}(y_\infty) + \xi(1-z_\infty) f'_{\text{cr}}(y_\infty) & -\zeta f_{\text{br}}(y_\infty) - \xi f_{\text{cr}}(y_\infty) \end{pmatrix}. \tag{6.6}$$

Evaluating (6.6) in the disease-free equilibrium $y_\infty = 0, z_\infty = z_0$, we find

$$J(0,z_0) = \begin{pmatrix} -1+\tau z_0 & 0 \\ -\zeta z_0 f'_{\text{br}}(0) + \xi(1-z_0) f'_{\text{cr}}(0) & -\zeta f_{\text{br}}(0) - \xi f_{\text{cr}}(0) \end{pmatrix}. \tag{6.7}$$

Since the Jacobian for the disease-free equilibrium is lower-triangular, the eigenvalues are $\lambda_1 = -1+\tau z_0$ and $\lambda_2 = -\zeta f_{\text{br}}(0) - \xi f_{\text{cr}}(0)$. The eigenvalues are always real, so (un)stable spirals cannot be observed. We now consider several cases.

1. **Case** $f_{\text{br}}(0) = 0$ **and** $f_{\text{cr}}(0) = 0$
   The eigenvalues are $\lambda_1 = -1+\tau z_0$ and $\lambda_2 = 0$, which makes the stability undeterminable using linear stability analysis.

2. **Case** $f_{\text{br}}(0) = 0$ **and** $f_{\text{cr}}(0) > 0$
   The eigenvalues are $\lambda_1 = -1+\tau$ and $\lambda_2 = -\xi f_{\text{cr}}(0)$. Thus the disease-free equilibrium is a stable node if $\tau < 1$ and an unstable node if $\tau > 1$. For $\tau = 1$, the stability is undetermined. In this case, $z_0 = 1$.

3. **Case** $f_{\text{br}}(0) > 0$ **and** $f_{\text{cr}}(0) = 0$
   The eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = -\zeta f_{\text{br}}(0)$, thus the DFE is a stable node. In this case, $z_0 = 0$.

4. **Case** $f_{\text{br}}(0) > 0$ **and** $f_{\text{cr}}(0) > 0$
   The eigenvalues are $\lambda_1 = -1+\tau\frac{f_{\text{cr}}(0)}{\omega f_{\text{br}}(0)+f_{\text{cr}}(0)}$ and $\lambda_2 = -\zeta f_{\text{br}}(0) - \xi f_{\text{cr}}(0)$. Eigenvalue $\lambda_2 < 0$, thus the stability solely depends on λ_1. The disease-free equilibrium is a stable node if $\tau < \frac{\omega f_{\text{br}}(0)+f_{\text{cr}}(0)}{f_{\text{cr}}(0)}$, an unstable node if $\tau > \frac{\omega f_{\text{br}}(0)+f_{\text{cr}}(0)}{f_{\text{cr}}(0)}$ and is undetermined otherwise.

Unfortunately, we cannot directly analyse the stability of the endemic equilibria, because (i) we do not know y_∞ nor z_∞ and (ii) we require the functions f_{br} and f_{cr} and its derivatives f'_{br} and f'_{cr} to determine the stability. Moreover, the existence of multiple

endemic equilibria rules out the possibility of finding a Lyapunov function to prove the global stability of system (6.3). Nevertheless, for specific functional responses f_{br} and f_{cr} that have only a single EE, one could attempt to construct a Lyapunov function, which is outside the scope of this thesis.

6.3.4. BASIC REPRODUCTION NUMBER

The *basic reproduction number* R_0 is the number of secondary infections produced by one average infected individual in a completely susceptible population. At the point $R_0 = 1$, the disease-free equilibrium loses stability and an endemic equilibrium emerges. The point $R_0 = 1$ coincides with the epidemic threshold τ_c.

We compute the basic reproduction number R_0 using the next generation matrix method, which was first introduced in [121], then generalised in [122] (see also [123]). Even though the compartmental component of system (6.3) is one-dimensional (the equation for the link density z does not count, as the link density z is not related to individuals) and the analysis could also be performed by local stability analysis (see Section 6.3.3), we have chosen for the next generation matrix method due to its widely spread use.

We rewrite the first entry of (6.7) as $J_{11} = M_{11} - V_{11}$, with $M_{11}, V_{11} > 0$. The only such splitting possible, assuming $f_{\mathrm{cr}}(0) > 0$, is

$$M_{11} = \tau z_0, \quad V_{11} = 1.$$

Then, the basic reproduction number R_0 is $M_{11} V_{11}^{-1}$, i.e.

$$R_0 = \tau \frac{f_{\mathrm{cr}}(0)}{\omega f_{\mathrm{br}}(0) + f_{\mathrm{cr}}(0)}. \tag{6.8}$$

For $f_{\mathrm{cr}}(0) = 0$, equation (6.8) does not apply, because the disease-free equilibrium remains stable for all parameter values (recall Section 6.3.3). In case $f_{\mathrm{cr}}(0) = 0$, we will define the basic reproduction number R_0 in an alternative way, which will be specified later.

We use the definition (6.8) of R_0 to prove the following theorem.

Theorem 6.2 *If $f_{\mathrm{cr}}(0) > 0$ and $R_0 > 1$, system (6.3) admits at least one endemic equilibrium.*

Proof. It follows from the second equation of system (6.3) that, at equilibrium,

$$z_\infty = \frac{f_{\mathrm{cr}}(y_\infty)}{\omega f_{\mathrm{br}}(y_\infty) + f_{\mathrm{cr}}(y_\infty)}.$$

Moreover,

$$\left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_{y=1} = -1 < 0.$$

Hence, if we prove that there exists a $\varepsilon > 0$ such that

$$\left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_{y=\varepsilon} > 0,$$

the intermediate value theorem ensures the existence of at least one positive (i.e. endemic) equilibrium value for y. This coincides with requiring

$$-\varepsilon + \tau\varepsilon(1-\varepsilon)\frac{f_{\mathrm{cr}}(\varepsilon)}{\omega f_{\mathrm{br}}(\varepsilon) + f_{\mathrm{cr}}(\varepsilon)} > 0.$$

Simplifying by ε on both sides and rearranging the terms, we obtain

$$\tau\frac{f_{\mathrm{cr}}(\varepsilon)}{\omega f_{\mathrm{br}}(\varepsilon) + f_{\mathrm{cr}}(\varepsilon)} > \frac{1}{1-\varepsilon}. \tag{6.9}$$

Inequality 6.9 coincides with the assumption $R_0 > 1$ for $\varepsilon = 0$; hence, by continuity, there exists an $\varepsilon > 0$ such that the desired inequality is satisfied. This concludes the proof. □

Theorem 6.2 is a more general version of Theorem 5.2 from Chapter 5, where we also proved the existence of at least one endemic equilibrium. The difference is that Theorem 6.2 considers general link-breaking and link-creation functional responses, whereas Theorem 5.2 only considers the functional responses generated by the G-ASIS model.

6.3.5. GLOBAL STABILITY

Before proving global stability, we first consider limit cycles of the aNIMFA model, for which we invoke the Bendixson-Dulac theorem.

Theorem 6.3 (Bendixson-Dulac) *If there exists a C^1-function $\phi(y,z)$ such that the expression*

$$F(y,z) = \frac{\partial(\phi f)}{\partial y} + \frac{\partial(\phi g)}{\partial z} \tag{6.10}$$

has the same sign ($\neq 0$) almost everywhere in a simply connected region R, then the planar autonomous system

$$\frac{dy}{dt} = f(y,z),$$
$$\frac{dz}{dt} = g(y,z),$$

has no non-constant periodic solutions lying entirely within the region R.

A proof of Theorem 6.3 can be found in [124], or in [125] for the N-dimensional case. We now apply Theorem 6.3 to prove that system (6.3) admits no periodic solutions.

Theorem 6.4 *System (6.3) admits no non-trivial periodic solutions.*

Proof. We verify the Bendixson-Dulac criterion using $\phi(y,z) = \frac{1}{yz}$ for our system (6.3) in the region $R = (0,1)^2$. We find

$$\phi f = -\frac{1}{z} + \tau(1-y),$$
$$\phi g = -\zeta\frac{f_{\mathrm{br}}(y)}{y} + \xi\frac{f_{\mathrm{cr}}(y)}{yz} - \xi\frac{f_{\mathrm{cr}}(y)}{y}.$$

Filling in Eq. (6.10) gives

$$F(y,z) = -\tau - \xi \frac{f_{\text{cr}}(y)}{yz^2}.$$

Since $y, z > 0$ and $f_{\text{cr}}(y)$ is a non-negative function, we conclude that $F < 0$ in the whole region $R = (0,1)^2$ and there cannot exist any limit cycles. □

We finalises the theoretical analysis with a statement about global stability of the aNIMFA model. Recall that the DFE is locally (hence, globally) unstable when $R_0 > 1$. From Theorem 6.2, we can conclude the following.

Corollary 6.5 *Assume that $R_0 > 1$ and that the DFE is on the repelling part of the z-axis $\{y = 0, z > \frac{1}{\tau}\}$. Then, the endemic equilibrium, if it is unique, is globally asymptotically stable. If multiple endemic equilibria exist, or the DFE is in the attracting part of the z-axis, i.e. $\{y = 0, z < \frac{1}{\tau}\}$, no general conclusions can be drawn.*

Proof. We can exclude the possibility of homoclinic orbits to the DFE, whose stable manifold is the z-axis. Under our assumptions, the corollary is an immediate consequence of Theorem 6.4. □

6.4. EXAMPLES

In the previous section, we derived several general results for the aNIMFA model. However, certain properties, like the number and stability of the endemic states, could not be determined for general functional responses. Here, we investigate several examples of functional responses f_{br} and f_{cr}, whereby we primarily focus on epidemiological applications. Then, by assumption, the link-breaking rule $f_{\text{br}}(y)$ is likely to be increasing with the prevalence y and the link-creation rule $f_{\text{cr}}(y)$ is exactly opposite. The aNIMFA model is, however, more versatile and can be applied to other spreading phenomena, including opinion dynamics, cascading failures and information transport in the human brain. These spreading phenomena are often more complex than SIS epidemics, thus requiring more complex (maybe even non-monotone) functional responses f_{br} and f_{cr}.

6.4.1. EXAMPLE 1: RANDOM LINK-ACTIVATION DEACTIVATION

Presumably the easiest functional responses are those that are totally unaffected by the current number of infected cases. Then the network density evolves independently of the epidemic prevalence. This model is known as the Random Link-Activation Deactivation (RLAD) model [98]. In the RLAD model, each link in the underlying network can be randomly created or broken, with rates ξ and ζ respectively. Mathematically, we require that the functional responses f_{br} and f_{cr} are constant and for simplicity, we consider $f_{\text{br}}(y) = f_{\text{cr}}(y) = 1$, and hence system (6.3) becomes

$$\frac{dy}{dt} = -y + \tau y(1-y)z, \tag{6.11a}$$

$$\frac{dz}{dt} = -\zeta z + \xi(1-z). \tag{6.11b}$$

Then, the basic reproduction number as defined in Eq. (6.8) is $R_0 = \frac{\tau}{1+\omega}$. In this simple example, the governing equation (6.3b) for the link-density $z(t)$ is decoupled from the

prevalence $y(t)$ and can be solved directly;

$$z(t) = \frac{1}{1+\omega} + \left(z_0 - \frac{1}{1+\omega}\right) e^{-(\xi+\zeta)t},$$

where the effective link-breaking rate $\omega = \zeta/\xi$. If the exponential decays sufficiently fast (i.e. $\xi+\zeta$ is large), the network density quickly converges to $z = 1/(1+\omega)$. Substituting $z = 1/(1+\omega)$ into Eq. (6.3a) and solving yields the famous logistic equation [126] for the prevalence;

$$y(t) = \frac{y_\infty}{1 + e^{-K(t-t_0)}}, \tag{6.12}$$

where $y_\infty = 1 - \frac{1+\omega}{\tau}$ is the steady-state prevalence, $K = \tau - 1$ is the growth rate and $t_0 = \frac{1}{K}\ln\left(\frac{y_\infty}{y_0} - 1\right)$ is the *inflection point*, better known as the *epidemic peak*. Formula (6.12) only holds if $y_0 \neq y_\infty$ and the solution equals $y(t) = y_0$ if $y_0 = y_\infty$.

The time-varying prevalence $y(t)$, given by Eq. (6.12), converges to a unique, non-zero, steady-state prevalence $y_\infty > 0$ if $\tau > 1+\omega$. Otherwise, for $\tau < 1+\omega$, the prevalence decreases exponentially to zero. The same result follows from linear stability analysis. The DFE, given by $(y_\infty, z_\infty) = (0, \frac{1}{1+\omega})$, is asymptotically stable for $\tau < 1+\omega$, unstable for $\tau > 1+\omega$ and undetermined for $\tau = 1+\omega$. The unique endemic equilibrium is given by $(y_\infty, z_\infty) = \left(1 - \frac{1+\omega}{\tau}, \frac{1}{1+\omega}\right)$, which is in the biologically feasible region only if $R_0 > 1$, and coincides with the DFE when $R_0 = 1$. The Jacobian of the EE is

$$J = \begin{pmatrix} 1 - \frac{\tau}{1+\omega} & (1+\omega)\left(1 - \frac{1+\omega}{\tau}\right) \\ 0 & -\zeta - \xi \end{pmatrix}.$$

The eigenvalues are $\lambda_1 = 1 - \frac{\tau}{1+\omega}$ and $\lambda_2 = -\zeta - \xi < 0$. Thus the EE is a stable node if $R_0 > 1$, unstable node if $R_0 < 1$ and undetermined for $R_0 = 1$. As remarked above, the case $R_0 < 1$ leads to $y_\infty < 0$ which is biologically infeasible. The steady states and their behaviour of the RLAD model are shown in Table 6.1.

Table 6.1: The equilibria of Example 1 and their local stability.

Example 1: $f_{\text{br}}(y) = f_{\text{cr}}(y) = 1$	$R_0 \leq 1$	$R_0 > 1$
Disease-free state $\left(0, \frac{1}{1+\omega}\right)$	stable node	unstable node
Endemic equilibrium $\left(1 - \frac{1+\omega}{\tau}, \frac{1}{1+\omega}\right)$	unstable node	stable node

Since the link-dynamics is decoupled from the disease dynamics in the RLAD model, the behaviour of the RLAD model is very similar to the static SIS model and undergoes the usual transcritical bifurcation, except that the basic reproduction number R_0 is a function of the effective link-breaking rate ω. For other functional responses f_{br} and f_{cr}, we expect different behaviour, which will be investigated in the upcoming examples.

6.4.2. EXAMPLE 2: EPIDEMICS: $f_{\text{br}}(y) = y$, $f_{\text{cr}}(y) = 1$

Contrary to the randomly evolving links in Example 1, we expect that genuine epidemic outbreaks affect the number of contacts of people. We consider the simple case where

the link-breaking process $f_{\text{br}}(y) = y$ is a linear function of the prevalence, but the link-creation process remains independent from the total number of infections [$f_{\text{cr}}(y) = 1$]. Then, the governing equations become

$$\frac{dy}{dt} = -y + \tau y(1-y)z, \tag{6.13a}$$

$$\frac{dz}{dt} = -\zeta zy + \xi(1-z). \tag{6.13b}$$

The basic reproduction number as defined in Eq. (6.8) is $R_0 = \tau$. The disease-free equilibrium $(y_\infty, z_\infty) = (0,1)$ is a stable node if $\tau < 1$, an unstable node if $\tau > 1$ and is otherwise undetermined. The unique EE follows from Eq. (6.4) as $(y_\infty, z_\infty) = \left(\frac{\tau-1}{\tau+\omega}, \frac{\tau+\omega}{\tau(1+\omega)}\right)$ and exists in the biologically feasible region for $\tau > 1$.

We now show that the unique EE is locally stable for this specific choice of f_{br} and f_{cr}. The Jacobian around the EE equals

$$J = \begin{pmatrix} -1+\tau(1-2y_\infty)z_\infty & \tau y_\infty(1-y_\infty) \\ -\zeta z_\infty & -\zeta y_\infty - \xi \end{pmatrix} = \begin{pmatrix} -\dfrac{\tau-1}{1+\omega} & \dfrac{\tau(\tau-1)(\omega+1)}{(\tau+\omega)^2} \\ -\zeta\dfrac{\tau+\omega}{\tau(1+\omega)} & -\zeta\dfrac{\tau(\omega+1)}{\omega(\tau+\omega)} \end{pmatrix}.$$

For $\tau > 1$ and $\zeta, \omega > 0$, we have $J_{1,1}, J_{2,1}, J_{2,2} < 0$ and $J_{1,2} > 0$; hence, $\text{tr}(J) < 0$ and $\det(J) > 0$, which implies that the real parts of its eigenvalues are negative. Hence, the EE is locally stable. Following Corollary 6.5, the EE is also globally asymptotically stable for $R_0 > 1$, which is a consequence of the absence of limit cycles guaranteed by Bendixson-Dulac and the fact that the DFE is unstable for $R_0 > 1$.

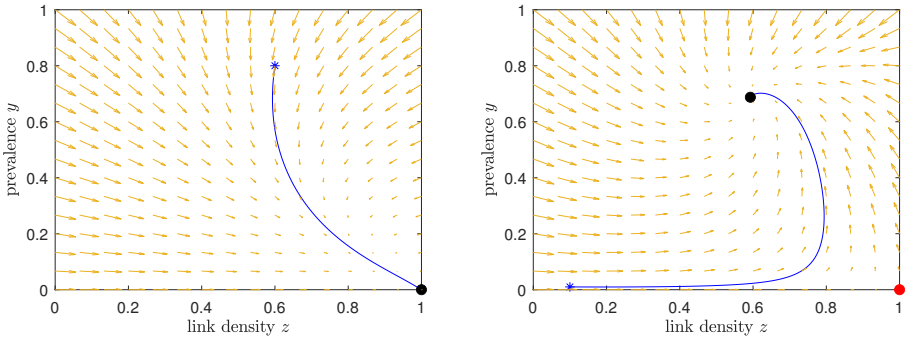
We summarize the stability of the two equilibria in Table 6.2 and present simulations of the two possible behaviours of system (6.13) in Figure 6.1. Comparing this example to Example 1, the behaviour is different in two ways. First, the basic reproduction number $R_0 = \tau$ does not depend on the link-breaking rate ζ and link-creation rate ξ, whereas R_0 is a linear function of $\omega = \zeta/\xi$ in Example 1. Second, the unique endemic equilibrium for $R_0 > 1$ converges in a spiral, whereas Example 1 converges without spiral behaviour.

Table 6.2: The equilibria of Example 2 and their local stability.

Example 2: $f_{\text{br}}(y) = y, f_{\text{cr}}(y) = 1$	$R_0 \leq 1$	$R_0 > 1$
Disease-free state $(0,1)$	stable node	unstable node
Endemic equilibrium $\left(\frac{1-\frac{1}{\tau}}{1+\frac{\omega}{\tau}}, \frac{1+\frac{\omega}{\tau}}{1+\omega}\right)$	unstable spiral	stable spiral

6.4.3. EXAMPLE 3: THE ADAPTIVE SIS MODEL

The adaptive SIS (ASIS) model was explained in Chapter 4 and describes the response of individuals to an on-going pandemic. In particular, the ASIS model assumes that links are broken between susceptible and infected nodes and (re)created between susceptible nodes. The aNIMFA approximation of the ASIS model was already analysed in Chapter 5 and the functional responses were derived as $f_{\text{br}}(y) = 2y(1-y)$ and $f_{\text{cr}}(y) = (1-y)^2$.



(a) $R_0 = 0.8$ (b) $R_0 = 5.4$

Figure 6.1: Dynamics for Example 2. Starting point: asterisk; stable equilibrium: black dot; unstable equilibrium: red dot. (a) If $R_0 < 1$, any initial condition converges to the DFE. (b) If $R_0 > 1$, the unique EE is globally stable. The other parameters are, for simplicity, $\zeta = \xi = 1$.

The link-breaking response is similar to Example 2, but the term $1 - y$ was added to account for the fact that for large epidemic outbreaks, the susceptible population may be depleted and the possibility to break links between susceptible and infected individuals decreases, simply because of the lack of susceptible individuals. The factor 2 is a conversion factor from the original Markovian model; we keep this factor for consistency with Chapter 5. The link-creation response f_{cr} is more intuitive; we expect many links to be created if the disease is almost nonexistent. Hence, we are considering the governing equations

$$\frac{dy}{dt} = -y + \tau y(1-y)z, \tag{6.14a}$$

$$\frac{dz}{dt} = -2\zeta z y(1-y) + \xi(1-z)(1-y)^2. \tag{6.14b}$$

The basic reproduction number as defined in Eq. (6.8) is, once again, $R_0 = \tau$. The disease-free equilibrium $(0,1)$ is a stable node for $R_0 < 1$, unstable node for $R_0 > 1$ and is undetermined otherwise. The unique endemic equilibrium follows from (6.4) and has y-coordinate (recall (5.12))

$$y_\infty = 1 - \frac{1-2\omega}{2\tau} - \sqrt{\left(\frac{1-2\omega}{2\tau}\right)^2 + \frac{2\omega}{\tau}},$$

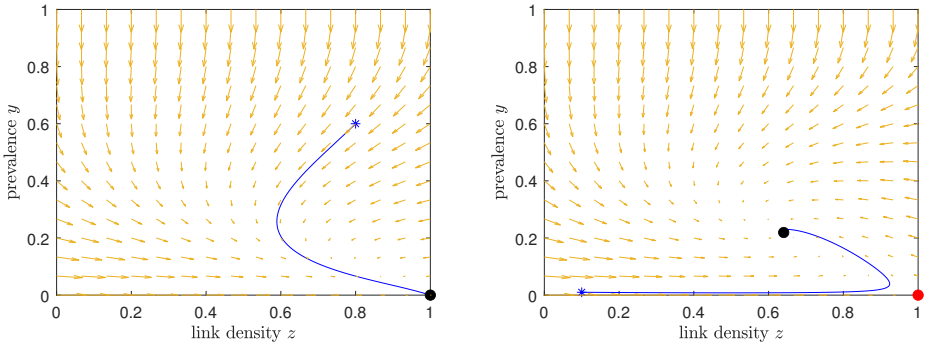
and the EE becomes $\left(y_\infty, \frac{1}{\tau(1-y_\infty)}\right)$. Using basic arithmetic, it can be verified that $\tau > 1$ implies $0 < y_\infty \leq 1 - \frac{1}{\tau}$, which ensures that the EE is contained in the physical region $(0,1)^2$. Thus, the EE exists for $R_0 > 1$.

The calculations needed for the stability of the EE become extremely cumbersome; however, the Bendixson-Dulac theorem, the uniqueness of the EE, the boundedness of solutions (see Lemma 6.1) and the instability of the DFE ensure that the EE is globally asymptotically stable when $R_0 > 1$ (recall Corollary 6.5).

We summarize the stability of the two equilibria in Table 6.3 and present simulations of the two possible behaviours of system (6.14) in Figure 6.2.

Table 6.3: The equilibria of Example 3 and their local stability.

Example 3: $f_{\text{br}}(y) = 2y(1-y), f_{\text{cr}}(y) = (1-y)^2$	$R_0 \leq 1$	$R_0 > 1$
Disease-free state $(0,1)$	stable node	unstable node
Endemic equilibrium $\left(y_\infty, \frac{1}{\tau(1-y_\infty)}\right)$	unstable spiral	stable spiral



(a) $R_0 = 0.8$ (b) $R_0 = 2.0$

Figure 6.2: Dynamics for Example 3. Starting point: asterisk; stable equilibrium: black dot; unstable equilibrium: red dot. (a) If $R_0 < 1$, any initial condition appears to converge to the DFE. (b) If $R_0 > 1$, the unique EE is globally stable. The other parameters are, for simplicity, $\zeta = \xi = 1$.

6.4.4. EXAMPLE 4: INFORMATION SPREAD

Besides epidemiological applications, we also revisit the AID model from Trajanovski *et al.* [97]. The functional responses of the AID model have been derived in Chapter 4, and are given by $f_{\text{br}}(y) = (1-y)^2$ and $f_{\text{cr}}(y) = 2y(1-y)$. The link-breaking response reduces the links in the network when the gossip prevalence is low. On the other hand, the link-creation response increases the link density for low prevalence, based on the fact that gossip is appealing and links are established to spread the news. The factor 2 in f_{cr} is again a conversion factor from the Markovian G-ASIS model. To summarise, the governing equations are given by:

$$\frac{dy}{dt} = -y + \tau y(1-y)z, \tag{6.15a}$$

$$\frac{dz}{dt} = -\zeta z(1-y)^2 + 2\xi y(1-z)(1-y). \tag{6.15b}$$

The disease-free equilibrium $(0,0)$ is stable for all $\tau > 0$. The y-coordinates of the two endemic equilibria follow from (6.4) and are equal to (see Eq. (5.13))

$$(y_\infty)_{1,2} = \frac{2\tau + \omega - 2 \pm \sqrt{(2\tau+\omega-2)^2 - 8\tau\omega}}{4\tau} \tag{6.16}$$

and the EE become $\left(y_\infty, \frac{1}{\tau(1-y_\infty)}\right)$. The basic reproduction number R_0 cannot be determined in the traditional way using Eq. (6.8), as the disease-free equilibrium does not lose stability. Instead, we *define* the basic reproduction number as the point where the two endemic equilibria are born (i.e. where (6.16) has non-complex solutions). Then the basic reproduction number follows as

$$R_0 = \frac{2\tau}{\omega + 2 + \sqrt{8\omega}}. \tag{6.17}$$

The dynamics of the AID model is plotted in Figure 6.3. For $R_0 < 1$, the solution converges to $(0,0)$. For $R_0 > 1$, the solution may converge to the disease-free state $(0,0)$, but can also converge to the stable endemic equilibrium, depending on the initial condition. The dependence of the basic reproduction number R_0 on the effective link-breaking rate ω is non-linear, which contrasts all earlier examples, that were either independent or linearly dependent on the effective link-breaking rate ω. Lastly, since the DFE lies in the attracting part of the z-axis, we cannot rule out the existence of a homoclinic orbit from the DFE.

We summarize the stability of the three equilibria in Table 6.4 and present simulations of the two possible behaviours of system (6.15) in Figure 6.3.

Table 6.4: The equilibria of Example 4 and their local stability.

Example 4: $f_{\text{br}}(y) = (1-y)^2, f_{\text{cr}}(y) = 2y(1-y)$	$R_0 < 1$	$R_0 \geq 1$
Disease-free state $(0,0)$	stable node	stable node
Endemic equilibrium $\left((y_\infty)_1, \frac{1}{\tau(1-(y_\infty)_1)}\right)$	non-existent	unstable node
Endemic equilibrium $\left((y_\infty)_2, \frac{1}{\tau(1-(y_\infty)_2)}\right)$	non-existent	stable node

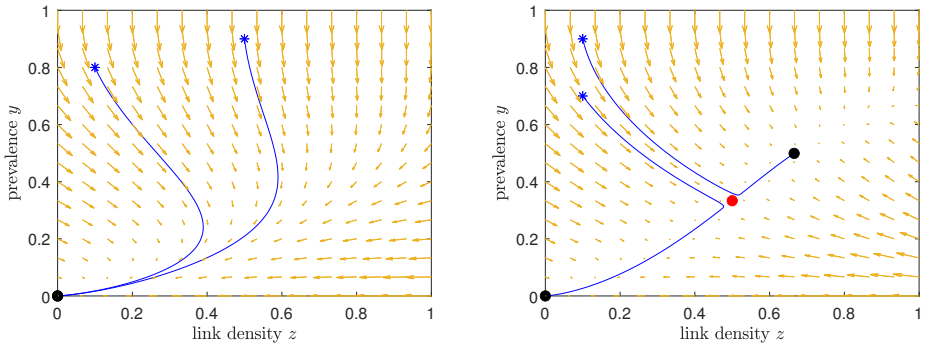
The basin of attraction of each stable equilibrium can be determined using a Lyapunov function. Such Lyapunov functions distinguish for which initial conditions the system will converge to either the DFE or the stable EE. However, up to the best of our knowledge, no exact Lyapunov function can be constructed for system (6.3) nor for most choices of the link-breaking and link-creation mechanisms.

However, the Lyapunov function can be approximated by considering a linearisation around a fixed point. For example, for the DFE $(0,0)$, its Jacobian equals

$$J_{(0,0)} = \begin{pmatrix} -1 & 0 \\ 2\xi & -\zeta \end{pmatrix}.$$

According to Khalil [127, p. 73–80], we can obtain an approximate Lyapunov function $\hat{V}$ by solving for the matrix P in the following matrix equation

$$PJ + J^T P = -I,$$



(a) $R_0 = 0.27$ (b) $R_0 = 1.03$

Figure 6.3: Dynamics for Example 4. Starting points: asterisks; stable equilibrium: black dot; unstable equilibrium: red dot. (a) If $R_0 < 1$, any initial condition converges to the DFE. (b) If $R_0 > 1$, solutions may converge to the stable endemic equilibrium or the disease-free state, depending on the initial condition. The other parameters are, for simplicity, $\zeta = \xi = 1$.

and the Lyapunov function follows as

$$\hat{V}(y,z) = \begin{pmatrix} y \\ z \end{pmatrix}^T P \begin{pmatrix} y \\ z \end{pmatrix}.$$

The estimated Region of Attraction Ω is then determined by the largest $c > 0$ for which

$$\Omega_c := \{(y,z) \in [0,1]^2 \mid \hat{V}(y,z) \leq c\},$$

is such that

$$\Omega := \max_{c>0} \left\{ \frac{d}{dt} \hat{V}(\Omega_c) < 0 \right\}.$$

For Example 4, the estimated Lyapunov function around $(0,0)$ becomes

$$\hat{V}(y,z) = \frac{1}{2\zeta} z^2 + \frac{1}{2} y^2 + \frac{2\xi}{\zeta(1+\zeta)} yz + \frac{2\xi^2}{\zeta(1+\zeta)} y^2$$

which is a tedious formula, but it is clear that $\hat{V} > 0$ in the biologically relevant region $[0,1]^2$. Unfortunately, the derivative $\frac{d}{dt}\hat{V}$ is extremely complicated, even in such a simple case. Hence, we derive the largest possible approximate region of attraction numerically, which is shown in Figure 6.4. The approximate regions of attraction for the disease-free equilibrium $(0,0)$ and the stable endemic equilibrium are shown in orange, whereas the exact boundary separating the two regions, and thus the actual basins of attraction of the two equilibria, is shown as a light-blue curve. The estimated regions of attraction often poorly match with the true regions of attraction [127], which is especially true for the stable EE in Figure 6.4. On the other hand, the region of attraction for the DFE is reasonably accurate.

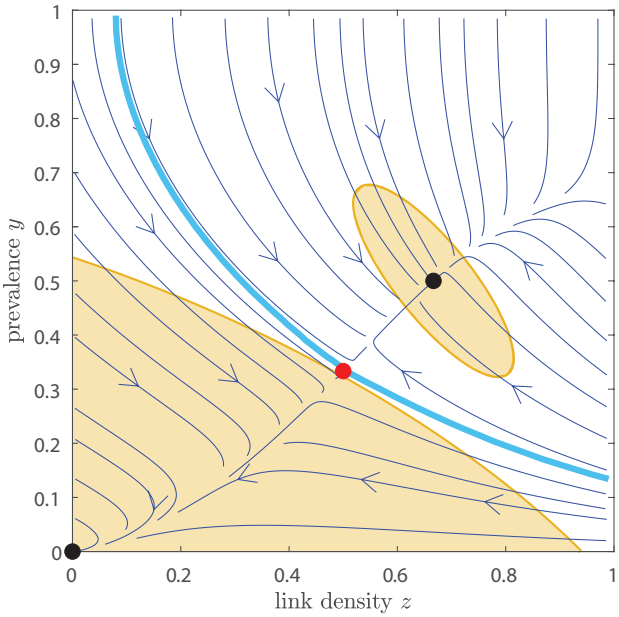


Figure 6.4: Regions of attraction for Example 4. The two attraction regions are separated by the numerically determined light-blue curve, which is the stable manifold of the unstable equilibrium. The approximate regions of attraction for the disease-free equilibrium and the stable endemic equilibrium are shown in orange. Black dots denote the stable equilibria and the red dot the unstable equilibrium. The parameters are, for simplicity, $\tau = 3, \zeta = \xi = 1$.

6.5. CONCLUSION

In this chapter, we developed a minimal model for SIS disease dynamics with personal risk perception, called adaptive NIMFA (aNIMFA). We investigated local and global stability of the model and showed that limit cycles cannot exist. The non-existence of limit cycles implies that epidemic waves cannot occur in the SIS model based on disease and behaviour dynamics alone and time-varying parameters are required to exhibit epidemic waves. Furthermore, we analysed various examples in detail, from epidemic contagion to information spread.

In this chapter, we assumed an homogeneous mixing of the population. In reality, some people contact each other frequently while other people never meet. Thus, the population exhibits very heterogeneous contact patterns. We expect that one can extend the current results to a heterogeneous network with N nodes. While considering a network, one must decide whether the link-breaking and link-creation functional responses act on the local prevalence of the node or on the global prevalence of the whole network. From a modelling perspective, we see possibilities for both approaches, or a mix of these.

We see several other interesting directions for future research. For example, is it possible to provide, besides continuity, conditions on f_{br} and f_{cr} such that we can limit/bound

the number of endemic equilibria from Eq. (6.4)? Can we determine for which f_{br} and f_{cr} the endemic equilibrium is unique?

The aNIMFA model is derived under SIS dynamics, but other compartmental models can be augmented with the link-breaking and link-creation functional responses as well. The extension of aNIMFA towards other, more realistic compartmental models is straightforward and we expect several interesting phenomena to occur based on the link-breaking and link-creation functional responses.

As a final comment, we mention the possibility to include delays into the knowledge about the current prevalence. As the COVID-19 pandemic exemplified, testing an individual typically takes several hours or days before the result is communicated. Moreover, the daily reported cases by governmental agencies typically run a few days behind. One modelling approach is to convert the aNIMFA model into a delay-differential equation, which typically complicates the analysis significantly. We leave these possibilities as an outlook for future works.

III

Practical epidemics: Modelling COVID-19 spread

7

FORECASTING THE SPREAD OF COVID-19

In this last chapter, we turn our focus towards a practical case: forecasting the spread of the Coronavirus Disease 2019 (COVID-19). Researchers from various scientific disciplines have attempted to forecast the spread of COVID-19. The proposed epidemic prediction methods range from basic curve fitting methods and traffic interaction models to machine-learning approaches. If we combine all these approaches, we obtain the Network Inference-based Prediction Algorithm (NIPA). In this chapter, we analyse a diverse set of COVID-19 forecast algorithms, including several NIPA variants. Among the diverse set of algorithms that we evaluated, the original NIPA performs best on forecasting the spread of COVID-19 in Hubei, China and in the Netherlands. In particular, we show that network-based forecasting is superior to any other forecasting algorithm.

This chapter is based on:
B. Prasse, M. A. Achterberg, L. Ma and P. Van Mieghem, *Network-inference-based prediction of the COVID-19 epidemic outbreak in the Chinese province Hubei*, Applied Network Science 5:35, Jul 2020 [128]
and
M. A. Achterberg, B. Prasse, L. Ma, S. Trajanovski, M. Kitsak and P. Van Mieghem, *Comparing the accuracy of several network-based COVID-19 prediction algorithms*, International Journal of Forecasting, 38:2, Apr–Jun 2022 [129].

7.1. INTRODUCTION

In December, 2019, the SARS-CoV-2 virus suddenly emerged in the Chinese province Hubei. The SARS-CoV-2 virus causes the Coronavirus Disease 2019 (COVID-19), which is a contagious respiratory disease. The number of COVID-19 cases in China rose dramatically to almost 80,000 at the end of February 2020. From China, COVID-19 quickly spread over the world, with almost ten million reported cases at the end of June, 2020. Many countries imposed a nation-wide lockdown to slow down the spread of COVID-19. A reliable forecast of the pandemic is key for deploying targeted and effective disease countermeasures, which aim at reducing the impact of the disease on the population.

Just like weather forecasts, predicting the future course of an epidemic outbreak is subject to fundamental limits [130]. First, the availability of reported data is limited, because the data is often reported only on a daily basis and carrying out medical tests on a large scale is challenging, especially if the testing capacity is limited. Second, forecasting the number of infected cases, especially in the starting phase, is highly sensitive to initial perturbations [131].

Nevertheless, many algorithms have been developed for forecasting the spread of COVID-19. One of the simplest approaches fits the number of infections using nonlinear regression to a sigmoid curve, such as the logistic function [126, 132], Hill function [133] or Gompertz function [134]. Other methods have been developed using Kalman filtering [135], Bayesian approaches [136], using aeroplane networks, daily commute traffic or cell phone traffic [137], machine learning algorithms such as adaptive neuro-fuzzy inference system [138] or Long Short-Term Memory (LSTM) [139] or are based on parameter estimation of the SIR [139, 140] or the SEIR model [141].

Many of the aforementioned methods implicitly assume that the number of infected cases is a certain region is independent of the other regions. In reality, the traffic between regions is of vital importance to transmit the disease across regions. Networks are a natural way to capture such interactions. We describe the interactions by a network G with N nodes. Each node i in the network G represents a particular region (country, province, municipality or city) and the link $a_{ij} \in \{0,1\}$ represents the existence of an interaction from region j to region i. The link weight β_{ij} describes the infection probability from region j to region i. The self-infection probability within a region i is given by β_{ii}, which is expected to be dominant over the other infection probabilities β_{ij}, because the interaction within a region is stronger than the interaction with other regions. The $N \times N$ infection probability matrix B, with elements β_{ij} is, however, not known.

We propose the Network-Inference-based Prediction Algorithm (NIPA) which forecasts the spread of the disease in three steps. First, the reported number of infected cases are preprocessed to remove outliers. As a second step, the unknown infection probability matrix B with elements β_{ij} is estimated based on past observations using a machine-learning algorithm. Finally, using the estimated infection probability matrix $\hat{B}$ and an SIR model, NIPA provides short-term forecasts on the number of infected cases.

In this chapter, we compare a diverse set of methods to forecast the spread of COVID-19, ranging from fitting closed-form epidemic curves and comprehensive machine-learning algorithms to network-based approaches, including NIPA. We focus on the spread of COVID-19, but we emphasise that all methods can be applied to general epidemic outbreaks. We focus on the outbreak in the initial stage, i.e. without access to vaccina-

tions and with a single dominant COVID-19 variant. Then we can safely assume that all disease parameters remain constant over the whole time period. We show that pure machine-learning, network-based algorithms or epidemiological models are inferior to algorithms that combine multiple approaches. In particular, the Network Inference-based Prediction Algorithm (NIPA) is shown to be superior to any other algorithm that we have evaluated. Section 7.2 introduces eight different forecast algorithms for predicting the future number of COVID-19 cases. We evaluate the accuracy of each algorithm in Section 7.3 in two selected regions: Hubei (China) and the Netherlands. The strengths and weaknesses of each algorithm is also discussed. Finally, we conclude in Section 7.4.

7.2. PREDICTION ALGORITHMS

The spread of COVID-19 can be measured in terms of the daily number of reported cases. Throughout this chapter, the "number of infected cases" is understood as the "number of cases *reported* by local authorities". The asymptomatic individuals, who do not feel sick and do not even know that they are infected and infectious, are not reported and can infect others unnoticed. To gain understanding of the percentage of asymptomatic cases, one possibility is to test the population at random. For COVID-19, the portion of asymptomatic cases is estimated to be as large as 80% [142]. Since the number of asymptomatic cases cannot be determined on a daily basis, we confine ourselves to the number of reported cases.

At the foundation of our prediction model lies the SIR compartmental model. We denote the discrete time by $k = 1,\dots,n$ where n is the total number of days since the first reported case on day $k = 1$. Contrary to the other chapters in this thesis, we use a discrete-time SIR model instead of a continuous-time process. The main reason is that nearly all authorities report the COVID-19 infections on a daily basis, thus using a time step of 1 day follows as a natural choice. The influence of the chosen time step on the prediction accuracy is investigated in Appendix E.9.

The SIR epidemic model with time-varying spreading parameters is given by:

Definition 7.1 (Discrete-time SIR model [143, 144])
The viral state $v_i[k] = (\mathcal{S}_i[k], \mathcal{I}_i[k], \mathcal{R}_i[k])^T$ *of every region* i *evolves in discrete time* $k = 1,2,\dots,n$ *according to*

$$\mathcal{I}_i[k+1] = (1-\delta_i)\mathcal{I}_i[k] + (1-\mathcal{I}_i[k]-\mathcal{R}_i[k]) \sum_{j=1}^{N} \beta_{ij}[k]\mathcal{I}_j[k], \tag{7.1a}$$

$$\mathcal{R}_i[k+1] = \mathcal{R}_i[k] + \delta_i \mathcal{I}_i[k] \tag{7.1b}$$

and the fraction of susceptible individuals follows as

$$\mathcal{S}_i[k] = 1 - \mathcal{I}_i[k] - \mathcal{R}_i[k], \tag{7.2}$$

where $\beta_{ij}[k] \geq 0$ *denotes the infection probability from region* j *to region* i *at time* k *and* $\delta_i > 0$ *denotes the curing probability of region* i.

We further define $\mathcal{I}_{\text{cum},i}[k] = \sum_{l=1}^{k} \mathcal{I}_i[l]$ as the cumulative fraction of infected cases in region i up to time k. Each forecasting algorithm aims to find the best possible forecast $\hat{\mathcal{I}}_{\text{cum},i}[k]$ for the cumulative number of infected cases $\mathcal{I}_{\text{cum},i}[k]$ for every region i and time k. In the sequel, we discuss the prediction methods in detail.

7.2.1. NIPA

Network-based approaches take into account the interactions between different regions. However, the contact network G is unknown (and consequently also the infection probabilities β_{ij}) and must be inferred from the epidemic outbreak. The methodology behind the Network Inference-based Prediction Algorithm (NIPA) was originally proposed in [144] to predict the future number of infected cases in a theoretical scenario. We apply the NIPA method to predict the outbreak of COVID-19, which consists of three steps. First, the raw data on the number of reported cases is preprocessed to obtain an SIR time series. Second, based on the SIR time series, we obtain estimates $\hat{\delta}_i$ and $\hat{\beta}_{ij}$ of the unknown spreading parameters δ_i and β_{ij} of the SIR model (7.1). In the third step, the estimates $\hat{\delta}_i$ and $\hat{\beta}_{ij}$ are used by iterating the SIR model for k days. The outcome of the SIR model serves as the prediction of the future evolution of the COVID-19 outbreak. We now describe the NIPA method in more detail.

STEP 1: DATA PREPROCESSING

Most governments report the number of infected cases $N_{\text{rep},i}[k]$ in region i at every day k. Using the number of inhabitants N_i of a region i, we obtain the report fraction of infected cases $\mathcal{I}_{\text{rep},i}[k]$. Based on the fraction of reported infections $\mathcal{I}_{\text{rep},i}[k]$ in region i, our goal is to obtain an SIR viral state vector $\nu_i[k] = (\mathcal{S}_i[k], \mathcal{I}_i[k], \mathcal{R}_i[k])^T$ for every region i at any time $k = 1, \ldots, n$. The fraction of susceptible individuals follows as $\mathcal{S}_i[k] = 1 - \mathcal{I}_i[k] - \mathcal{R}_i[k]$ at any time $k \geq 1$. Hence, it suffices to determine the fraction of infectious individuals $\mathcal{I}_i[k]$ and recovered individuals $\mathcal{R}_i[k]$.

The fraction of infectious individuals $\mathcal{I}_i[k]$ follows from the reported fraction of infections $\mathcal{I}_{\text{rep},i}[k]$. To be precise, the reported data is the fraction $\mathcal{I}_{\text{rep},i}[k]$ of individuals that are *detected* to be infected by COVID-19. Upon detection of the infection, the respective individuals are hospitalised or self-quarantined and are therefore no longer able to infect other individuals. We consider the reported fraction of infections $\mathcal{I}_{\text{rep},i}[k]$ as an *approximation* for the number of infectious individuals $\mathcal{I}_i[k]$.

We do not know the fraction of removed individuals $\mathcal{R}_i[k]$. At the initial time $k = 1$, it is realistic to assume that $\mathcal{R}_i[1] = 0$ holds for every region i. For any time $k \geq 2$, the fraction of removed individuals $\mathcal{R}_i[k]$ can be obtained from (7.1), if the curing probabilities δ_i were known. However, we do not know the curing probabilities δ_i either. Instead, we consider 50 equidistant *candidate values* for the curing probabilities δ_i, ranging from $\delta_{\min} = 0.01$ to $\delta_{\max} = 1$. We define the set of candidate values as $\Omega = \{\delta_{\min}, \ldots, \delta_{\max}\}$. For every candidate value $\delta_i \in \Omega$, the fraction of removed individuals $\mathcal{R}_i[k]$ follows from (7.1) at all times $k \geq 2$. Thus, we obtain 50 potential sequences $\mathcal{R}_i[1], \ldots, \mathcal{R}_i[n]$, each of which corresponding to a candidate value $\delta_i \in \Omega$. We estimate the curing probability δ_i, and hence the sequence $\mathcal{R}_i[1], \ldots, \mathcal{R}_i[n]$, as the element in Ω that resulted in the best fit of the SIR model (7.1) to the reported number of infections.

The raw time series $\mathcal{I}_{\text{rep},i}[1], \ldots, \mathcal{I}_{\text{rep},i}[n]$ exhibits erratic fluctuations. To reduce the fluctuations, we apply a moving average, provided by the Matlab command `smoothdata`, to the time series $\mathcal{I}_{\text{rep},i}[1], \ldots, \mathcal{I}_{\text{rep},i}[n]$ of every region i. The preprocessed time series $\mathcal{I}_i[1], \ldots, \mathcal{I}_i[n]$ equals the output of `smoothdata`.

STEP 2: NETWORK INFERENCE

For every region i, the curing probability δ_i is estimated as one of the candidate values in the set Ω, as outlined above. The remaining task is to estimate the infection probabilities β_{ij}. The goal of *network inference* [145, 146, 147] is to estimate the matrix B of infection probabilities from the SIR viral state observations $v_i[1], ..., v_i[n]$. The matrix B can be interpreted as the weighted adjacency matrix of a contact network. We adapt a network inference approach [144], which is based on formulating a set of linear equations and the *least absolute shrinkage and selection operator* (LASSO) [148].

The crucial observation from the SIR governing equations (7.1) is that β_{ij} appears linearly, whereas the state variables $\mathcal{S}_i$, $\mathcal{I}_i$ and $\mathcal{R}_i$ do not. Equation (7.1) can be rewritten in the following linear equation

$$V_i = F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \tag{7.3}$$

for all regions $i = 1, ..., N$. Here, the $(n-1) \times 1$ vector V_i and the $(n-1) \times N$ matrix F_i are defined as

$$V_i = \begin{pmatrix} \mathcal{I}_i[2] - (1-\delta_i)\mathcal{I}_i[1] \\ \vdots \\ \mathcal{I}_i[n] - (1-\delta_i)\mathcal{I}_i[n-1] \end{pmatrix} \tag{7.4}$$

and

$$F_i = \begin{pmatrix} \mathcal{S}_i[1]\mathcal{I}_1[1] & \dots & \mathcal{S}_i[1]\mathcal{I}_N[1] \\ \vdots & \ddots & \vdots \\ \mathcal{S}_i[n-1]\mathcal{I}_1[n-1] & \dots & \mathcal{S}_i[n-1]\mathcal{I}_N[n-1] \end{pmatrix}. \tag{7.5}$$

If the SIR model (7.1) were an exact description of the COVID-19 spread, then the linear system (7.3) would hold with equality. However, the viral state vector $v_i[k]$ in region i does not exactly follow the SIR model (7.1). Instead, the evolution of the viral state vector $v_i[k]$ is described by

$$v_i[k+1] = f_{\text{SIR}}(v_1[k], ..., v_N[k]) + w_i[k],$$

where the 3×1 vector $f_{\text{SIR}}(v_1[k], ..., v_N[k])$ denotes the right-hand sides of the SIR model (7.1), and the 3×1 vector $w_i[k]$ denotes the unknown *model error* of region i at time k. The model errors originate from the fact that (i) cases are only reported once per day, whereas the true epidemic evolves in continuous time and (ii) social distancing, nationwide lockdowns and the availability of vaccinations are not captured by the simple SIR model (7.1).

Due to the existence of model errors $w_i[k]$, the linear system (7.3) only holds approximately. Thus, we resort to estimating the infection probabilities β_{ij} by minimising the deviation of the left side and the right side of (7.3). We infer the network by the LASSO

[148] as follows:

$$\min_{\beta_{i1},\dots,\beta_{iN}} \quad \left\| V_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|_2^2 + \rho_i \sum_{j=1, j\neq i}^{N} \beta_{ij} \qquad (7.6)$$
$$\text{s.t.} \qquad 0 \leq \sum_{j=1}^{N} \beta_{ij} \leq 1, \quad j = 1,\dots,N.$$

The first term in the objective function of (7.6) measures the deviation of the left side and the right side of (7.3). The second term in the objective of (7.6) is an ℓ_1–norm regularisation term which avoids overfitting. We choose to not penalise the probabilities β_{ii}, since we expect the infections among individuals within the same region i to be dominant compared to infections across regions. The regularisation parameter $\rho_i > 0$ is set by cross–validation.

We emphasise that an accurate prediction of an SIR epidemic outbreak does not require accurate network inference [144, 149]. The main reason is that the dynamics evolve over a small subspace of the total N-dimensional space, requiring a few eigenvectors for an accurate reconstruction of the underlying network. For further details on the derivation of (7.6), we refer to Appendix E.3.

STEP 3: PREDICTION

Using the obtained estimates $\hat{\delta}_i$ for the curing probabilities and $\hat{\beta}_{ij}$ for the infection probabilities, we iterate the SIR equations (7.1) to obtained a forecast $\hat{\mathcal{I}}_i[k]$ for the number of reported cases for every region i at time k.

7.2.2. NIPA ON EACH REGION SEPARATELY

As a benchmark model, we apply NIPA on each region separately, which we name *NIPA separate*. NIPA separate is a machine-learning method based on the SIR model, but does not consider the interaction between different regions.

7.2.3. NIPA STATIC PRIOR

The standard NIPA approach considers the most extreme case, where no knowledge is available on the contact network. In practice, estimations of the contact network are often available (but may not be very accurate). We introduce *NIPA static prior*, which uses an estimate of the contact network as a prior distribution in the NIPA algorithm. We use a time-independent traffic network (with the corresponding traffic intensity matrix M) to obtain a prior for the infection probability matrix B as

$$B_{\text{prior}} = \text{diag}\,(c_1,\dots,c_N)\,M. \qquad (7.7)$$

We explain our motivation for the prior infection matrix B_{prior} in Appendix E.4. The positive scalars $c_1,\dots,c_N$ are unknown and are set by cross-validation. We assume that the true infection matrix B is normally distributed around the prior infection matrix B_{prior}. Based on the prior infection matrix B_{prior} and the observations of the COVID-19 spread,

we obtain the Bayesian estimate $\hat{B}$ by solving the optimisation problem

$$\hat{B} = \underset{B}{\operatorname{argmax}} \Pr\left[B \middle| \mathcal{I}[1], ..., \mathcal{I}[n]\right] \tag{7.8}$$
$$\text{s.t.} \sum_{j=1}^{N} \beta_{ij} \leq 1, \quad i = 1, ..., N,$$

where $\mathcal{I}[k]$ is the observed $N \times 1$ infection vector $\mathcal{I}[k] = (\mathcal{I}_1[k], ..., \mathcal{I}_N[k])^T$ at all times $k = 1, ..., n$. Using the estimated infection matrix $\hat{B}$ and the estimated curing probabilities $\hat{\delta}_i$ for every region i, we forecast the outbreak by iterating the SIR model. For details on NIPA static prior, we refer to Appendix E.5.

7.2.4. NIPA DYNAMIC PRIOR

Many governments imposed some kind of lockdown, during which the movement of its citizens is significantly restricted. Thus, the true contact network G varies over time. We use a time-varying traffic matrix $M[k]$ as an approximation for the prior infection matrix $B_{\text{prior}}[k]$, whose entries equal

$$B_{\text{prior}}[k] = \text{diag}\,(c_1, ..., c_N)\, M[k] \tag{7.9}$$

for all times k. The positive scalars $c_1, ..., c_N$ are unknown and are set by hold-out validation. We propose a Bayesian approach called *NIPA dynamic prior* to estimate the true infection matrix $B[k]$ from the time series of infected cases $\mathcal{I}_i[k]$ and the prior infection matrix $B_{\text{prior}}[k]$. Using the estimated time-varying infection matrix $\hat{B}[k]$ and the curing probabilities $\hat{\delta}_i$ for each region i, we forecast the outbreak by iterating the SIR model. Appendix E.6 explains the technical details of NIPA dynamic prior.

A challenge to NIPA dynamic prior is the unavailability of the contact network in the future. Hence, we assume the traffic matrix to remain constant after the last observation point n: $B_{\text{prior}}[n+k] = B_{\text{prior}}[n]$ for all $k > 0$.

7.2.5. SIGMOID CURVES

Epidemiological models such as the SIR model admit exact solutions for the fraction of infected nodes (recall Chapter 3). Such solutions often have the shape of a sigmoid curve. Out of all sigmoid curves, the logistic function is of particular interest, because the logistic function is the (approximate) solution for the number of infected cases in the SIS epidemic model [100, 126, 150] and the cumulative number of infected cases in the SIR epidemic model [131]. The logistic function assumes that the cumulative number of infected cases $\mathcal{I}_{\text{cum},i}[k]$ in region i at time k equals

$$\mathcal{I}_{\text{cum},i}[k] = \frac{\mathcal{I}_{\infty,i}}{1 + e^{-K_i(k - t_{0,i})}}, \tag{7.10}$$

where $\mathcal{I}_{\infty,i}$ is the long-term fraction of infections, K_i is the logistic growth rate and $t_{0,i}$ is the inflection point, also known as the epidemic peak. The parameters $y_{\infty,i}$, K_i and $t_{0,i}$ are estimated for each region separately using a nonlinear curve fitting procedure, which is explained in Appendix E.8. Two other sigmoid curves, namely the Hill function and Gompertz function, are also discussed in Appendix E.8.

7.2.6. LSTM

Recurrent neural networks [151] (RNNs) have been used in various tasks related to sequences [152], time series analysis, forecasting, speech recognition and natural language processing [153]. Long Short-term Memory (LSTM) networks [154] are specific types of RNNs that resolved the long-standing problem for long-term dependencies caused by the difference in input growth which in turns leads to vanishing or exploding gradients in neural networks backpropagation. LSTM introduces additional input, output and optional forget gates as interfaces with additional weights on top of the standard input data and hidden weights in the standard RNN. There are several variations [155, 156] for the LSTM networks, just to mention few: with or without forget gate and a “peephole connection"; that perform better in one or another task [157]. For the internal mechanism between the gates and the exact mathematical relations, we refer to [156] or [158]. In this chapter, we utilize the most common one – an LSTM with a forget gate. In the simulations, we use an LSTM with sequence and hidden sizes both equal to four in a single LSTM layer (e.g. it is possible to stack few LSTM layers which leads to more overfitting), a learning rate of 0.1 and Adam optimizer [159], with mean square error loss in 2000 epochs of training.

Table 7.1: All algorithms discussed in this chapter. *If the algorithm is based on a phenomenological epidemic process, like the SIR model. **If the algorithm is able to forecast small perturbations in the global trend. ***If the spread between different regions is considered.

Algorithm	Epidemiology*	Flexible**	Network***
NIPA	✓	✓	✓
NIPA separate	✓	✓	×
NIPA static prior	✓	✓	✓
NIPA dynamic prior	✓	✓	✓
Logistic function	✓	×	×
Hill function	✓	×	×
Gompertz function	✓	×	×
LSTM	×	✓	×

7.3. EVALUATION OF THE PREDICTION PERFORMANCE

We evaluate the prediction accuracy of the methods discussed in Section 7.2 by forecasting the spread of COVID-19 in a selected number of regions. We set the maximal forecast horizon to six days, because of the difficulty of predicting epidemic outbreaks [131].

Each prediction algorithm produces a forecast $\hat{\mathcal{I}}_{\text{cum},i}[k]$ for the cumulative fraction of infected cases $\mathcal{I}_{\text{cum},i}[k]$ for every region i at time k. To quantify the prediction error at time k, we use the Symmetric Mean Absolute Percentage Error (sMAPE)

$$e_{\text{sMAPE}}[k] = \frac{1}{N}\sum_{i=1}^{N}\frac{|\mathcal{I}_{\text{cum},i}[k]-\hat{\mathcal{I}}_{\text{cum},i}[k]|}{(\mathcal{I}_{\text{cum},i}[k]+\hat{\mathcal{I}}_{\text{cum},i}[k])/2}, \tag{7.11}$$

which is commonly used in forecasting [160]. Furthermore, we quantify the Percentage

Error (PE)

$$e_{\text{PE},i}[k] = \frac{\mathcal{I}_{\text{cum},i}[k] - \hat{\mathcal{I}}_{\text{cum},i}[k]}{\mathcal{I}_{\text{cum},i}[k]}, \tag{7.12}$$

for every region i and time k to investigate over- and underestimations. We consider the spread of COVID-19 in two regions: The cities in Hubei, China and the provinces in the Netherlands. These regions cannot be regarded as full representatives of the spread of COVID-19, let alone for general infectious diseases. Rather, these two regions illustrate the strengths and weaknesses of NIPA and the other algorithms.

7.3.1. HUBEI, CHINA

We first evaluate the prediction accuracy on the Chinese province Hubei. In December 2019, the first case of COVID-19 was detected in Wuhan, the capital of Hubei. The first case outside Wuhan was reported on January 21. From January 24 onwards, the whole province Hubei was under lockdown, prohibiting any non-urgent travels. On February 15, the local government in Hubei changed the diagnosing policy, causing an erratic increase in the number of reported cases on February 15. Therefore we restrict ourselves to the period from January 21 to February 14. The number of reported cases are provided by the Health Commission of Hubei [161] and are shown in Appendix E.2. We normalise the reported cases for each region by the number of inhabitants of each region, which are also reported in Appendix E.2.

The majority of COVID-19 patients were reported in Wuhan, as shown in Figure 7.1. We have removed Shennongjia from our analysis, because of the small number of infections in that region. Moreover, there is an outlier in Wuhan at time $k = 8$ (January 28), which we replaced by $\mathcal{I}_{\text{rep},1}[8] = (\mathcal{I}_{\text{rep},1}[7] + \mathcal{I}_{\text{rep},1}[9])/2$. The outlier is most likely due to the increase in the maximum testing capacity in Wuhan, which increased[1] from 200 to 2000 people per day as of January 27, 2020.

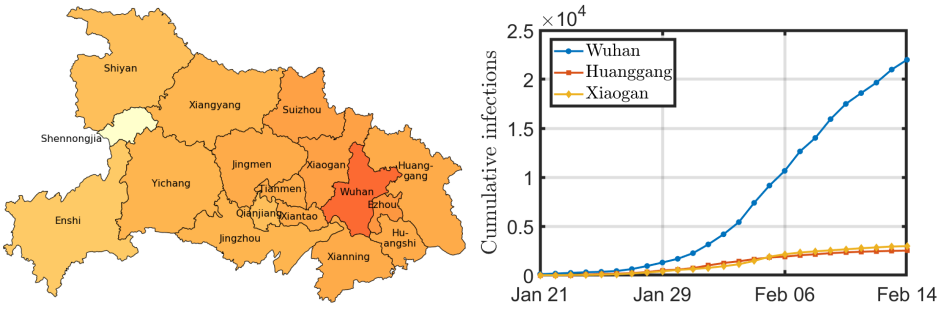


Figure 7.1: The left figure shows the geographical map of Hubei. The darker the region, the more infections per 100,000 inhabitants on during the time period January 21 – February 14. The three cities with the most infections on February 14 are displayed on the right.

For NIPA static prior, we require a traffic network describing the interactions between the regions in Hubei. The Chinese company Baidu provides an estimate of the number

[1] See https://m.chinanews.com/wap/detail/zw/sh/2020/01-28/9071697.shtml (in Chinese).

of commuters between all regions in Hubei on a daily basis [162]. The static prior is set proportional to the traffic network on January 21, which corresponds to day $k = 1$.

Figure 7.2 shows the prediction accuracy over time for different forecast algorithms. The horizontal axis shows the date d. We have forecasted the disease several days ahead in time, using all available information from January 22 until d. For example, the rightmost point in Figure 7.2a includes data from January 22 to February 13 to forecast the situation on February 14.

The sMAPE error in Figure 7.2 tends to decrease as time evolves, because a growing amount of data is available. Furthermore, the total number of infected cases quickly increases, whereas the daily infected cases increase at a lower rate, indicating sub-exponential growth [131, 163]. Sub-exponential growth will inevitably reduce the sMAPE error, because sMAPE is a relative error metric. On the other hand, the prediction accuracy decreases rapidly if the forecast horizon is enlarged. Especially for five and six days ahead in time around February 1 cannot be predicted accurately, which is illustrated by Figure 7.2e and 7.2f, respectively.

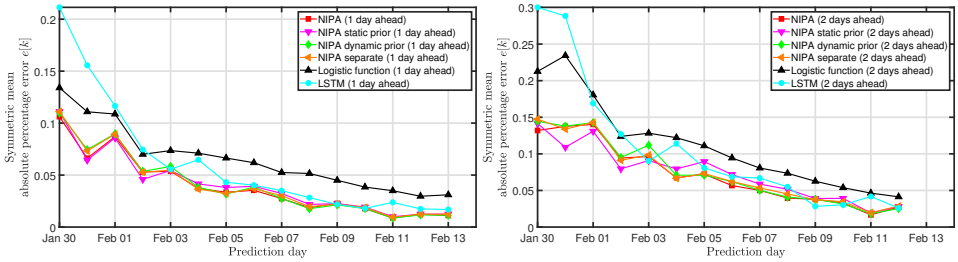
The logistic function performs generally worse than the other algorithms, for which several reasons exist. First, by fitting a logistic curve, we assume the number of cases to follow the SIR model closely [131, 4]. Hence, we do not allow any individual or governmental responses to COVID-19, which typically flattens the (logistic) curve. Second, the logistic function ignores the spread between regions, which further deteriorates the prediction accuracy. Third, the logistic function is symmetric around the epidemic peak at $k = t_0$; the increase and decrease of the number of cases around the peak is equal. Most epidemic outbreaks of COVID-19 show a rapid increase and a more gradual decrease of the daily number of cases. A possible reason is that most lockdowns are enforced immediately, whereas lockdown measures are lifted gradually. Occasionally, the Hill function [133] and Gompertz function [134] are used to predict epidemic outbreaks, because they allow asymmetry around the epidemic peak. We focus in the remainder of this chapter on the logistic function, but report in the conclusion in Table 7.2 on the accuracy of the Hill and Gompertz function.

The performance of LSTM is moderately good, but LSTM fails to find an accurate forecast around January 31. Since the time series is the shortest at the left part of Figure 7.2, less data is available to train LSTM. Pure machine-learning algorithms are known to yield a lower prediction accuracy than other methods if the time series is short [164].

The prediction accuracy of all NIPA methods in Figure 7.2 is similar, although NIPA static prior is considerably worse around February 4 for the prediction of three or more days ahead in time. A possible reason is as follows. The impact of the nation-wide lockdown on January 24 is captured incorrectly by the static prior, whereas the original NIPA method has more freedom to adjust its contact network accordingly and NIPA dynamic prior receives a more tailored prior to the current epidemiological situation. Another reason is that the prior network (dynamic or static) may deviate significantly from the true infection matrix. We emphasise that under ideal circumstances, namely when the epidemic outbreak exactly follows the SIR model, NIPA static prior outperforms NIPA, see Appendix E.7.

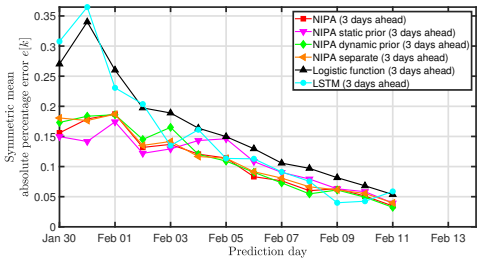
Figure 7.2 also shows that by neglecting the interaction between regions, as done by NIPA seperate, decreases the prediction accuracy compared to the original NIPA. We

conclude that a network-based approach appears beneficial for forecasting. We summarise the results in Section 7.4.

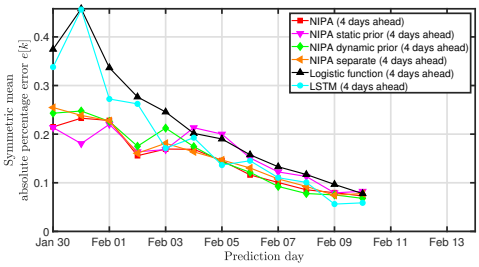


(a) 1 day ahead

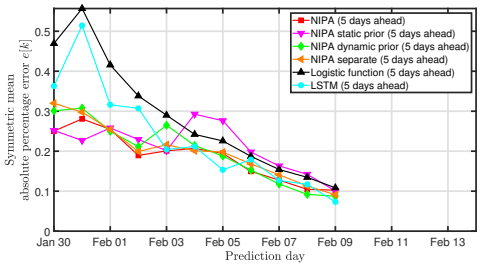
(b) 2 days ahead



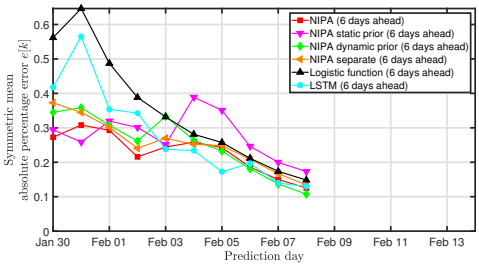
(c) 3 days ahead



(d) 4 days ahead



(e) 5 days ahead

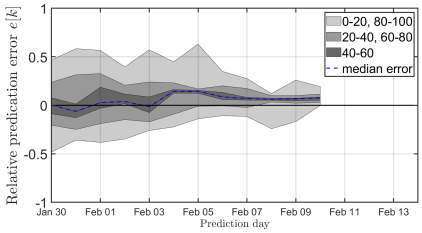


(f) 6 days ahead

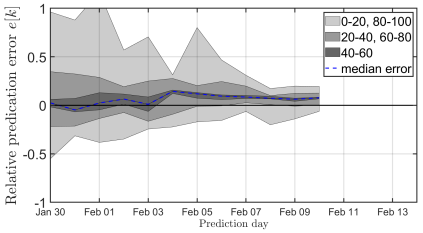
Figure 7.2: The prediction accuracy for the situation in Hubei, China. The subplots show the prediction accuracy for a forecast horizon of (a) 1 day, (b) 2 days, (c) 3 days, (d) 4 days, (e) 5 days and (f) 6 days for the prediction algorithms from Section 7.2.

Another interesting topic is *forecast bias*: The tendency to systematically overestimate or underestimate the true number of infected cases. Using the Percentage Error (PE) we estimate the bias for all prediction algorithms for region i at time k. The surface error plots in Figure 7.3 show the Percentage Error as a function of time for a 4-days ahead prediction. The logistic function and LSTM show the largest deviation around the mean, especially around February 1, which is in agreement with Figure 7.2. Furthermore, Figure 7.3 illustrates that the logistic function and LSTM systematically underestimate the true number of cases. On the other hand, NIPA static prior appears to overestimate

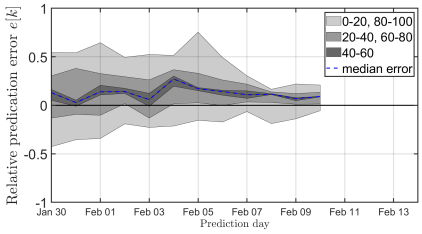
the true number of cases. A possible reason is the following. The static network is taken to be proportional to the traffic flow before the lockdown measures. When the lockdown is introduced, the static prior remains constant, so the algorithm overestimates the true result. After some days, the newly collected data shows evidence that the prior is not very accurate, so NIPA static prior ignores the prior and more closely follows the data instead, which improves the forecast accuracy.



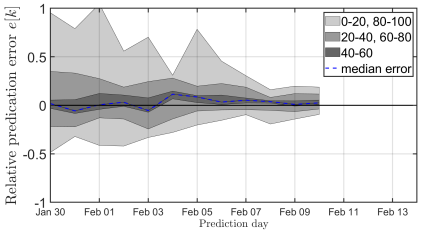
(a) NIPA



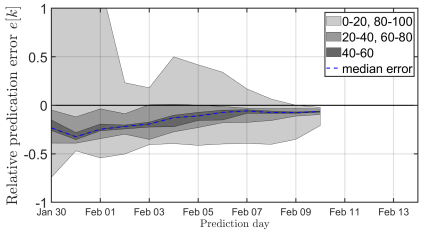
(b) NIPA separate



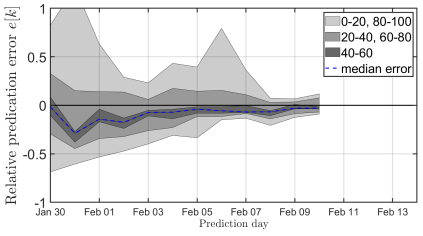
(c) NIPA static prior



(d) NIPA dynamic prior



(e) Logistic function



(f) LSTM

Figure 7.3: The surface plots for the relative prediction error for a 4-days forecast horizon over time. The subfigures show (a) NIPA, (b) NIPA separate, (c) NIPA static prior, (d) NIPA dynamic prior, (e) Logistic function and (f) LSTM.

7.3.2. THE NETHERLANDS

As a second case study, we investigate the spread of COVID-19 in the Netherlands. The first case in the Netherlands was diagnosed on February 27. After February 27, the number of cases grew rapidly, as depicted in Figure 7.4. The epidemic peak was observed at the end of March, and the daily number of cases slowly decreased from that time on-

wards. We consider the spread of COVID-19 on a provincial level, for which data is available from the Rijksinstituut voor Volksgezondheid en Milieu (RIVM) [165]. The Netherlands is subdivided into twelve provinces, for which the RIVM reports the daily number of new infections. Since the number of infected cases increases more gradually in the Netherlands than in Hubei, China, the total epidemic period is longer and more data points are available. A more gradual increase in the number of cases should be beneficial for the prediction accuracy.

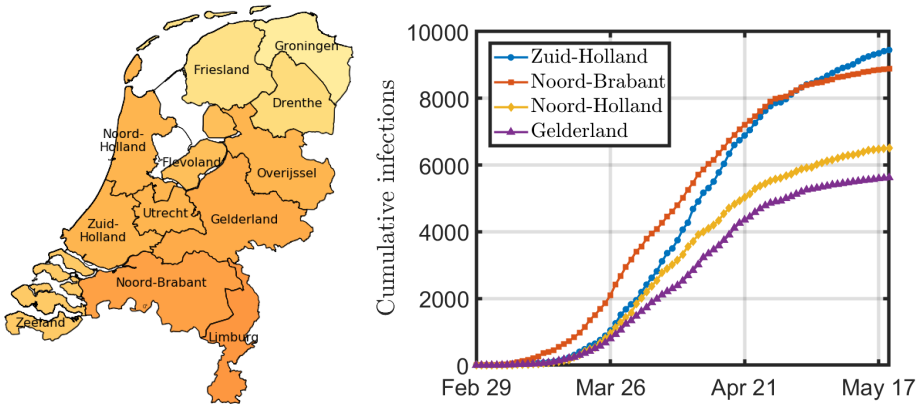


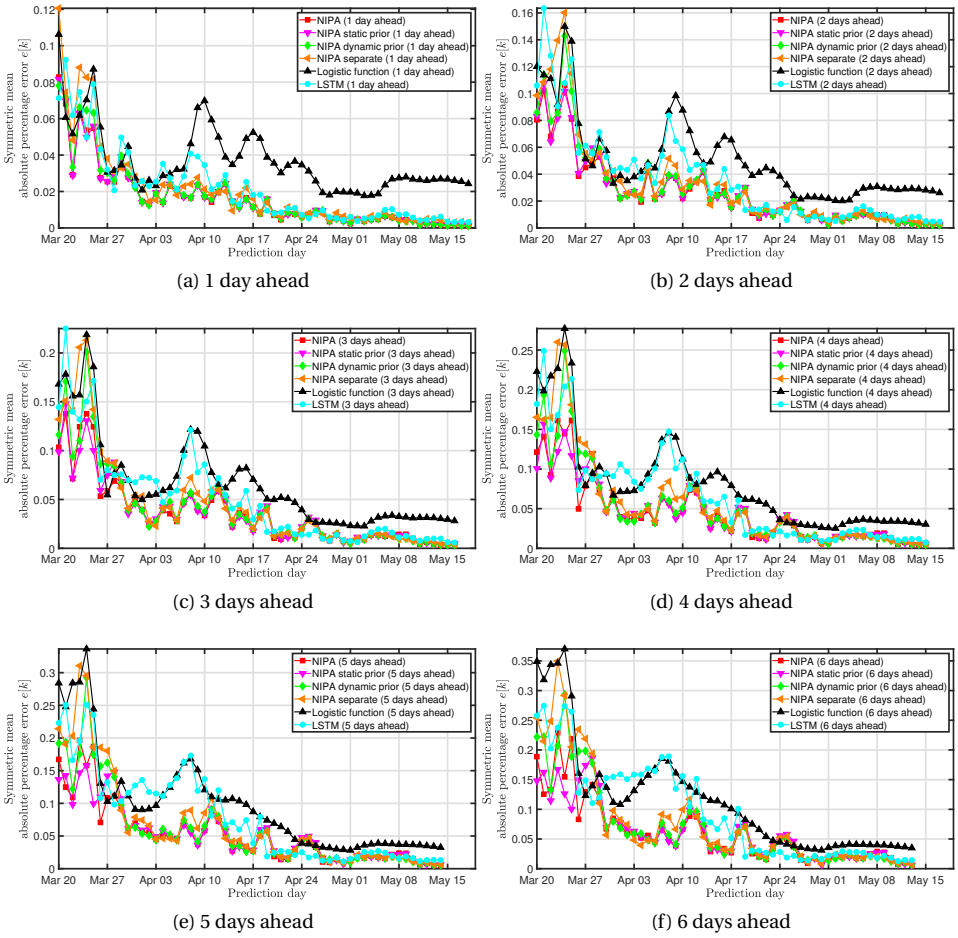
Figure 7.4: The left figure shows the geographical map of the Netherlands. The darker the province, the more infections per 100,000 inhabitants on May 19. The four provinces with the most infections on May 19 are displayed on the right.

For NIPA static prior, we require a traffic network as an approximation for the interaction between the provinces. Statistics Netherlands (Centraal Bureau voor de Statistiek) reports the number of people m_{ij} working in province i and living in province j, averaged over one year [166]. We use the Google Mobility Data “Workplaces" to estimate the time-varying traffic network for each province in the Netherlands [167]. Google reports the percentage decrease of traffic $p_i[k]$ on day k in province i compared to an ordinary day between January 3 and February 6, 2020. During the lockdown, we expect $p_i[k] < 1$ because of the lockdown measures. Then we construct the time-dependent traffic matrix as follows: $m_{ij}[k] = m_{ij} \cdot p_i[k]$.

The prediction accuracy for the Netherlands is outlined in Figure 7.5. Before April 1, the situation in the Netherlands is similar to Hubei, where the NIPA methods perform better, but there exist large deviations in the prediction accuracy. After April 1, the accuracy of all NIPA methods is nearly identical. In other words, the influence of the initial static/dynamic network on the prediction is small. The main reason is that the NIPA algorithms are trained on a growing amount of infection data as time advances, naturally implying that the algorithm focusses more on the data rather than the given initial prior. Among the best performing methods over the whole period are original NIPA and NIPA separate, whereas the logistic function and LSTM show the worst performance.

The prediction accuracy of NIPA separate and NIPA are comparable, except at the left-hand side of Figure 7.5. A possible reason is that the spread of COVID-19 is at the be-

ginning dominated by interprovincial interactions. After imposing the lockdown at the end of March, the interaction between the provinces is lowered significantly, resulting in a spreading process that dominantly takes place within each province.



(a) 1 day ahead

(b) 2 days ahead

(c) 3 days ahead

(d) 4 days ahead

(e) 5 days ahead

(f) 6 days ahead

Figure 7.5: The prediction accuracy for the situation in the Netherlands. The subplots show the prediction accuracy (a) 1 day ahead, (b) 2 days ahead, (c) 3 days ahead, (d) 4 days ahead, (e) 5 days ahead and (f) 6 days ahead.

7.4. CONCLUSION

We have compared the prediction accuracy of eight algorithms to forecast the spread of COVID-19. We summarise the results in Table 7.2. The error in Table 7.2 is obtained by averaging over all sMAPE forecast errors for forecast horizons between one and six days and over all regions and all times. Fitting a sigmoid curve, like the logistic function, performs the worst of all methods. The main reasons for the low prediction accuracy

are the imposed symmetry around the epidemic peak and the negligence of the interaction between regions. Other sigmoid curves, such as the Hill function and the Gompertz function, perform slightly better than the logistic function, but perform worse than most other algorithms. The machine-learning algorithm Long Short-Term Memory (LSTM) is not based on any phenomenological epidemic processes nor considers provincial interactions. Table 7.2 shows that the prediction accuracy of LSTM is comparable to the Hill and Gompertz functions.

The Network Inference-based Prediction Algorithm (NIPA) is a combination of machine learning, phenomenological epidemics (SIR model) and considers the interaction between different regions. After preprocessing the data, NIPA infers the contact network as an intermediate step. Using the obtained contact network, the prediction is computed based on an SIR model. Table 7.2 illustrates that the average prediction accuracy of NIPA is better than any other algorithm. Applying NIPA to each region separately (NIPA separate) yields a forecast error which is comparable to LSTM. We conclude that network-based approaches are beneficial for accurate forecasts. We have also shown that choosing a time-varying or static prior close to the true contact network may improve the forecast accuracy of NIPA. Surprisingly, the inclusion of a time-varying or static prior in NIPA on real infection data is not beneficial for the forecast accuracy for the considered regions. Among several reasons, the chosen prior might be an inaccurate estimate of the true contact network.

Our discrete-time SIR model (7.1) was derived under the assumption that the force-of-infection is small [168, 169]. Although the number of simultaneously infected people during the COVID-19 pandemic is relatively small compared to the total population and the approximation is therefore reasonable, we believe that the application of the genuine force-of-infection may help to ameliorate the forecasting results obtained in this chapter.

In a practical setting, such as the COVID-19 pandemic, policymakers might prefer to anticipate to a worst-case scenario. In that case, an asymmetric error metric that penalises underestimations more significantly than overestimations may be more suitable.

Table 7.2: All algorithms discussed in this chapter. The Netherlands is abbreviated as NL. The sMAPE error is averaged over all forecast horizons, all regions and all times. *As input, each algorithm requires the population size N_i of each region i and a time series of the fraction of reported cases $\mathcal{I}_{\text{rep},i}[k]$ in each region i at every time k.

Algorithm	Additional input*	Error (Hubei)	Error (NL)	Bias
NIPA	-	0.122	0.0381	
NIPA separate	-	0.129	0.0487	
NIPA static prior	static network	0.135	0.0384	over
NIPA dynamic prior	dynamic network	0.129	0.0429	
Logistic function	-	0.186	0.0735	under
Hill function	-	0.142	0.0531	
Gompertz function	-	0.141	0.0528	
LSTM	-	0.160	0.0570	under

8

CONCLUSION

In this thesis, we investigated several aspects of the spread of epidemics on networks. Just as the SIS and SIR models played a central role in the history of epidemiology, they have also played a central role in this thesis. In this conclusion, we first present an overview of our obtained results and subsequently provide several interesting research directions in the field of network epidemiology.

8.1. MAIN CONTRIBUTIONS

We started in Chapter 2 with a thorough investigation of the spectrum of the transition matrix of the Markovian ε-SIS process on the complete graph. The tridiagonal transition matrix of the ε-SIS process was already known, but we derived exact results for the eigenvalues in the limit of large infection rates, large self-infection rates and small self-infection rates. Even though the eigenvalues of the ε-SIS process are strictly distinct, we observed nearly degenerate eigenvalues for some parameter values, indicating that for those parameter values, the effect of the infection process and the self-infection process are almost equally strong. We additionally devised a novel estimate of the epidemic threshold in the Markovian ε-SIS process, which is defined as the effective infection rate for which the third-largest eigenvalue of the transition matrix is minimal.

Chapter 3 focussed on general compartmental models, where the population is split into c compartments. If the compartmental graph does not contain any cycles, we exactly computed the time-varying solution on heterogeneous networks. We exemplified the exact solution in the SI and SIR model and also provided several extensions, e.g. on temporal networks, added self-infection rates, included non-Markovian dynamics and added simplicial contagion. Our novel observation is that the Markov graph, which consists of all possible states in the Markov chain, can be represented by an infinitesimal generator that is a triangular matrix. Using an iterative method, the exact solution can be computed, even for heterogeneous transition rates. Using the exact solution, we investigated one of the key elements in epidemiology: the location of the epidemic peak. Using a second-order Newton-Raphson method, we accurately determined the epidemic peak time with only a few iterations.

Overcoming the unrealistic assumption that contact graphs are static, we proposed the Generalised Adaptive SIS (G-ASIS) model in Chapter 4, which assumes that the contact graph changes based on the presence of the virus in the population. More specifically, the link-breaking and link-creation mechanism describe how nodes break or create links with their neighbours, based on the viral state of their neighbours. We derived the entire set of link-updating rules, leading to 36 instances in the G-ASIS model. We proceeded by deriving lower and upper bounds for the epidemic threshold as a function of the link-breaking rate. Several instances of the G-ASIS model were investigated in detail, including the ASIS (disease), AID (gossip) and ABN (brain activity) model.

The huge state space in the Markovian G-ASIS model prevented us from deriving further properties of the G-ASIS model. In Chapter 5, we derived the first-order and second-order mean-field approximation of the G-ASIS model. We showed that the first-order mean-field approximation performs poorly and is inadequate for extracting properties of the Markov process. On the other hand, the second-order mean-field approximation performs much better and is mostly able to qualitatively recover the epidemic threshold. The main reason is that the second-order approximation only approximates the infection process, but exactly preserves the link-breaking and link-creation processes, whereas the first-order approximation approximates all these processes.

Building on the idea of the G-ASIS model, we extended the mean-field approximation from Chapter 5 to general link-breaking and link-creation rules in Chapter 6. The primary difference is that the mean-field approximation from Chapter 5 is derived from the G-ASIS model, whereas the aNIMFA model from Chapter 6 includes general link-breaking and link-creation rules, which are not necessarily limited to knowledge about the local prevalence around a node, but can also be based on global knowledge of the disease prevalence in the whole population. The link adaptation rules are reminiscent of functional responses in Lotka-Volterra equations, which describe the intake of predators based on the number of available prey. For all possible functional responses, we computed the epidemic threshold and proved that limit cycles do not exist. The non-existence of limit cycles implies that the SIS process with adaptive link dynamics cannot reproduce COVID-19 waves and more complicated models are required to replicate the COVID-19 waves.

In Chapter 7, we shift our focus to modelling the spread of COVID-19. We derived the Network-Inference-based Prediction Algorithm (NIPA), which consists of three steps. First, the daily number of reported cases is preprocessed into an SIR time series. As an intermediate step, the contact graph is estimated from the SIR time series using a machine-learning algorithm. Using the inferred contact graph, a short-term forecast for the number of reported cases is obtained. We also extended NIPA by including static and dynamic prior information on the contact graph, which was based on traffic networks. We compared NIPA and its variants to several well-known prediction models and concluded that NIPA, as a hybrid method including epidemiological, network-based and machine-learning aspects, performs as one of the best methods to forecast the spread of COVID-19.

8.2. DIRECTIONS FOR FUTURE RESEARCH

The exact solution of the SIR model in heterogeneous populations from Chapter 3 is both surprising and intriguing. Unfortunately, the exponential size of the state space prevents computations on networks of realistic sizes. The commonly-used mean-field approximation performed poorly, hence novel approximation methods are required to better understand the SIR process. A promising method involves the clustering of the exponentially large Markov graph by aggregating states with similar properties. Such a clustering drastically reduces the size of the Markov graph, while preserving the most important dynamical properties.

The G-ASIS model from Chapter 4 describes the spread of a disease over an adaptive network. We focussed in this thesis on the derivation of the epidemic threshold, which is related to the spread of the disease. An important, but a less frequently researched topic, involves the time-varying topology itself. By understanding the time-varying contact graph in the G-ASIS model, further inside into the characteristics of adaptive epidemics can be obtained. The first-order and second-order mean-field approximations from Chapter 5 are unsuitable for this task, because the relevant metrics of the time-varying contact graph are not explicitly computed under the mean-field approximation. Hence, research must either focus on the Markovian, stochastic model or novel mean-field approximations must be derived, which are devised for capturing one particular metric of the time-varying topology.

The G-ASIS model was derived to describe the contact avoidance of individuals during an epidemic, in which one type of functional response to the disease is prescribed. Other sociological models for human-disease interaction may be used as an alternative for adaptive disease modelling. Additionally, we mention that we applied contact avoidance to SIS epidemics, but the link-breaking and link-creation mechanisms are model-agnostic and can be applied to any compartmental epidemic model, including models aimed at describing the spread of real-world infectious diseases. We believe novel effects will appear by combining other compartmental models with the link-breaking and link-creation dynamics and subsequent derivations of the epidemic threshold and mean-field approximations will help to shed light on the interplay between disease spread and contact avoidance.

The simple adaptive SIS model from Chapter 6 was only defined on a complete graph, but can easily be extended towards general graphs. Of key interest is the interplay between intrinsically different regions, as occurred during the COVID-19 pandemic. Each country has a different culture, climate and population structure, and had therefore invented its own disease countermeasures against COVID-19. It would be curious to understand how the different decisions of each country affect the propagation of the disease within and across countries. A key question is whether the simple model from Chapter 6 is able to replicate similarities and differences in disease prevalences across heterogeneous populations in an applied setting such as the COVID-19 pandemic.

The COVID-19 pandemic thought us many lessons on the modelling of epidemics. One particular effect is currently underestimated or not even considered in many epidemiological models: the compliance of individuals to the social distancing rules. Besides the spreading of the disease and the breaking and creation of links in the contact graph, we believe a third, relevant process must be considered; the spread of misinfor-

mation regarding the spread of the disease. The propagation of misinformation is hard to measure, but we believe that the effect of compliance is crucial and including misinformation in the epidemic model is ultimately able to ameliorate our understanding and forecast accuracy of real-world epidemic outbreaks.

A

APPENDIX TO CHAPTER 2

A.1. SIMPLIFYING THE TRANSITION MATRIX P

To prove that the (asymmetric) transition matrix P has real eigenvalues, we show that P is similar to a symmetric matrix. Here, we follow [44, Appendix A.6.3]. We omit the tildes on the scaled birth rate Ξ_k and scaled death rate μ_k in this appendix for readability. Let H be a diagonal matrix with values $(h_1, h_2, \ldots, h_{N+1})$ on the diagonal. Then the similarity transform $\tilde{P} = HPH^{-1}$ equals

$$\tilde{P} = \begin{pmatrix} -\Xi_0 & \Xi_0 \frac{h_1}{h_2} & & & & & \\ \mu_1 \frac{h_2}{h_1} & -(\Xi_1 + \mu_1) & \Xi_1 \frac{h_2}{h_3} & & & & \\ & \mu_2 \frac{h_3}{h_2} & -(\Xi_2 + \mu_2) & \ddots & & & \\ & & & \ddots & & & \\ & & & \ddots & & & \\ & & & & \mu_{N-2} \frac{h_N}{h_{N-1}} & -(\Xi_{N-1} + \mu_{N-1}) & \Xi_{N-1} \frac{h_N}{h_{N+1}} \\ & & & & & \mu_{N-1} \frac{h_{N+1}}{h_N} & -\mu_N \end{pmatrix}. \tag{A.1}$$

To assure the symmetry $\tilde{P} = \tilde{P}^T$, we require that $\tilde{P}_{ij} = \tilde{P}_{ji}$ for all $1 \leq i, j \leq N+1$. Then we find the following condition

$$\frac{h_{k+1}}{h_k} \mu_k = \frac{h_k}{h_{k+1}} \Xi_{k-1} \tag{A.2}$$

for all $k = 2, \ldots, N+1$. Solving recurrence relation (A.2) and taking $h_1 = 1$, we find that

$$h_k = \sqrt{\prod_{i=1}^{k-1} \frac{\Xi_{i-1}}{\mu_i}},$$

A

which implies that $\tilde{P} = HPH^{-1}$ is a symmetric matrix. The transformed matrix $\tilde{P}$ equals

$$\tilde{P} = \begin{pmatrix} -\Xi_0 & \sqrt{\Xi_0\mu_1} & & & & & \\ \sqrt{\Xi_0\mu_1} & -(\Xi_1+\mu_1) & \sqrt{\Xi_1\mu_2} & & & & \\ & \sqrt{\Xi_1\mu_2} & -(\Xi_2+\mu_2) & \ddots & & & \\ & & & \ddots & & & \\ & & & \ddots & & & \\ & & & & -(\Xi_{N-1}+\mu_{N-1}) & \sqrt{\Xi_{N-1}\mu_N} & \\ & & & & \sqrt{\Xi_{N-1}\mu_N} & -\mu_N \end{pmatrix}.$$

The transformed matrix $\tilde{P}$ is symmetric and has therefore real eigenvalues. Since the similarity transform preserves the eigenvalues, we conclude that P has real eigenvalues.

For the numerical computation of the eigenvalues of the transition matrix P, the best practice is to consider the transformed matrix $\tilde{P}$ instead of the original matrix P, because symmetric matrices have favourable properties for most eigenvalue algorithms [170]. We further improve the numerical procedure by removing the steady state π corresponding to the eigenvalue $\lambda = 0$. Here, we follow [47, Section 3 and 4]. Given the transition matrix P, we define $\bar{P}$ as the reduced $N \times N$-matrix with entries $\bar{p}_{ij} = p_{ji} - p_{0i}$ for $i, j = 1, 2, \ldots, N$. We now prove that the eigenvalues of $\bar{P}$ are equal to the eigenvalues of P, except for the removed eigenvalue zero. Let λ be an eigenvalue of P and $\mathbf{v} = (v_0, v_1, \ldots, v_N)$ the corresponding eigenvector. Furthermore, we denote $\mathbf{u}$ as the $(N+1) \times 1$ all-ones vector. Since $P\mathbf{u} = \mathbf{0}$, it follows that

$$\bar{P}^T(\bar{\mathbf{v}} - v_0\bar{\mathbf{u}}) = \lambda(\bar{\mathbf{v}} - v_0\bar{\mathbf{u}}),$$

where $\bar{\mathbf{v}} = (v_1, \ldots, v_N)$ and $\bar{\mathbf{u}}$ the $N \times 1$ all-ones vector. Thus λ is also an eigenvalue of $\bar{P}$, unless $\bar{\mathbf{v}} - v_0\bar{\mathbf{u}} = \mathbf{0}$, which is only true if $\bar{\mathbf{v}}$ is constant, thus corresponding to $\lambda = 0$. Hence, the eigenvalues of $\bar{P}$ equal the eigenvalues of P, and the matrix $\bar{P}$ equals

$$\bar{P} = \begin{pmatrix} -(\Xi_0+\Xi_1+\mu_1) & \mu_2-\Xi_0 & -\Xi_0 & \ldots & -\Xi_0 & -\Xi_0 & -\Xi_0 \\ \Xi_1 & -(\Xi_2+\mu_2) & \mu_3 & & & & \\ & & & \ddots & & & \\ & & & \ddots & & & \\ & & & & \Xi_{N-2} & -(\Xi_{N-1}+\mu_{N-1}) & \mu_N \\ & & & & & \Xi_{N-1} & -\mu_N \end{pmatrix}.$$

The reduced matrix $\bar{P}$ can be transformed into a tridiagonal matrix. Let T be the the upper triangular matrix with ones at and above the diagonal, and the remaining terms

are zero. Then

$$T = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ & 1 & 1 & \dots & 1 & 1 \\ & & \ddots & & \vdots & \vdots \\ & & & \ddots & \vdots & \vdots \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix}, \qquad T^{-1} = \begin{pmatrix} 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & \ddots & \ddots & & \\ & & & & 1 & -1 \\ & & & & & 1 \end{pmatrix},$$

such that the reduced, transformed matrix $\bar{\bar{P}} = T\bar{P}T^{-1}$ equals

$$\bar{\bar{P}} = \begin{pmatrix} -(\Xi_0+\mu_1) & \mu_1 & & & & \\ \Xi_1 & -(\Xi_1+\mu_2) & \mu_2 & & & \\ & \Xi_2 & -(\Xi_2+\mu_3) & \ddots & & \\ & & & \ddots & & \\ & & & \ddots & & \\ & & & & -(\Xi_{N-2}+\mu_{N-1}) & \mu_{N-1} \\ & & & & \Xi_{N-1} & -(\Xi_{N-1}+\mu_N) \end{pmatrix}.$$

The reduced, transformed matrix $\bar{\bar{P}}$ is asymmetric and can be converted to a symmetric matrix $\bar{\bar{\bar{P}}}$ using the same transformation matrix H as for the original transition matrix P. The main advantages of the reduced, transformed, symmetric matrix $\bar{\bar{\bar{P}}}$ compared to the original transition matrix P are that numerical methods to obtain eigenvalues (i) are more efficient because the matrix is symmetric, (ii) prevent complex eigenvalues due to symmetry and (iii) are less prone to rounding errors as the zero eigenvalue is removed.

A.2. EQUAL METASTABLE AND STEADY-STATE PREVALENCE

The time-dependent behaviour of the ε-SIS process, which is illustrated in Figure 2.2, shows that the final steady-state prevalence y_∞ can be lower or higher than the prevalence y in the metastable state. For a given network size N and effective infection rate τ, we believe that there always exists some $\varepsilon^* > 0$ for which the prevalence in the metastable state and the steady-state prevalence y_∞ are equal. We numerically determine this self-infection rate ε^* and depict the result in Figure A.1. We only show effective infection rates τ above the epidemic threshold τ_c, because the metastable state does not exist below the threshold. If the effective infection rate τ is sufficiently large, Figure A.1 depicts a power-law decay with exponent $\alpha = -27.6$.

We find an explanation for this result in Theorem 4 by Van Mieghem [21], which states that the epidemic threshold τ_c for sufficiently small ε^* is bounded by

$$\frac{1}{e}\left(\frac{10^{-s}}{\varepsilon^*(N-1)!}\right)^{\frac{1}{N-1}} < \tau_c < \left(\frac{10^{-s}}{\varepsilon^*(N-1)!}\right)^{\frac{1}{N-1}},$$

where 10^{-s} specifies an agreed fraction of infected nodes that determines whether an outbreak has taken place. Considering that the final steady-state prevalence y_∞ is of

A

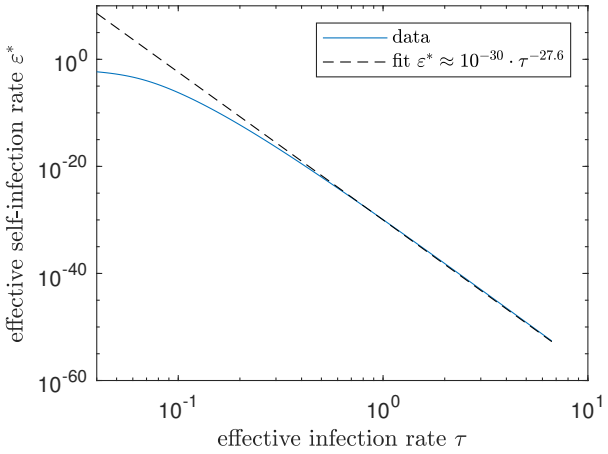


Figure A.1: Illustration of the effective self-infection rate ε^* for which the metastable prevalence y and the steady-state prevalence y_∞ are equal. The power-law tail is observed for $\tau > 0.3 \approx 9\tau_c^{(1)}$ and the results are obtained on a graph with $N = 30$ nodes.

order 1, we choose $s = 0$. Naturally, the equation can be rewritten in terms of the effective self-infection rate ε^*:

$$\frac{1}{e^{1-N}(N-1)!}\tau^{1-N} \leq \varepsilon^* \leq \frac{1}{(N-1)!}\tau^{1-N}. \tag{A.3}$$

Considering the example in Figure A.1 with $N = 30$ nodes, our estimation (A.3) states that

$$10^{-19}\tau^{-29} \leq \varepsilon^* \leq 10^{-31}\tau^{-29}. \tag{A.4}$$

The bounds in (A.4) are very similar to the fit in Figure A.1. Only for larger effective self-infection rates ε^*, the approximated solution (A.4) starts to deviate.

A.3. MEAN-FIELD ε-SIS

The well-known N-Intertwined mean-field approximation (NIMFA) [29] assumes that any two stochastic variables X and Y are uncorrelated: $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$. Applying the mean-field approximation to the ε-SIS process on the complete graph, the governing equations become [44, p. 462]

$$\frac{d\mathbf{s}_{\mathrm{MF}}(t)}{dt} = \varepsilon\mathbf{u} - (\delta+\varepsilon)\mathbf{s}_{\mathrm{MF}}(t) + \mathrm{diag}\left(\mathbf{u} - \mathbf{s}_{\mathrm{MF}}(t)\right)\tilde{B}\mathbf{s}_{\mathrm{MF}}(t) \tag{A.5}$$

where $\mathbf{s}_{\mathrm{MF}}(t) = (s_1(t), ..., s_N(t))^T$ is the $N \times 1$ viral state vector, $\mathbf{u} = (1, ..., 1)^T$ the $N \times 1$ all-ones vector and $\tilde{B} = \beta(\mathbf{u}\mathbf{u}^T - I)$ where I is the $N \times N$ identity matrix. Cator and Van Mieghem [68] showed that NIMFA upperbounds the Markovian SIS process. As an upper bound to Eq. (A.5), we consider the ε-NIMFA process with self-loops, thus $\beta_{ii} = \beta > 0$.

Furthermore, we rescale time $\tilde{t} = \delta t$, and introducing the effective infection rate $\tau = \beta/\delta$, the effective self-infection rate $\varepsilon^* = \varepsilon/\delta$, we find after dropping the tildes

$$\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} = \varepsilon^*\mathbf{u} - (1+\varepsilon^*)\mathbf{s}_{\text{MF}}(t) + \text{diag}\left(\mathbf{u} - \mathbf{s}_{\text{MF}}(t)\right) B\mathbf{s}_{\text{MF}}(t) \tag{A.6}$$

where $B = \tau\mathbf{u}\mathbf{u}^T$. We adopt the approach in [48] to obtain the solution of the ε-SIS process on the complete graph. Since the Markovian ε-SIS process is non-negatively correlated [68] and we introduced self-loops $\beta_{ii} > 0$, the solution of (A.6) is a strict upper bound for Markovian ε-SIS dynamics.

Proof of Theorem 2.4. We prove Theorem 2.4 in three steps. First, in Subsection A.3.1, we show that the viral state $\mathbf{s}_{\text{MF}}(t)$ is in a two-dimensional subspace at every time t. More specifically, we show that $\mathbf{s}_{\text{MF}}(t) = c_1(t)\mathbf{z}_1 + c_2(t)\mathbf{z}_2$ for two $N \times 1$ agitation modes $\mathbf{z}_1, \mathbf{z}_2$ and two scalar functions $c_1(t), c_2(t) \in \mathbb{R}$. Second, in Subsection A.3.2, we obtain the closed-form expression for the function $c_1(t)$. Third, given the function $c_1(t)$, we obtain the function $c_2(t)$ in Subsection A.3.3.

A.3.1. THE VIRAL STATE IS IN A TWO-DIMENSIONAL SUBSPACE

With the definition of the agitation mode $\mathbf{z}_1$ in (2.8), we can write the infection rate matrix B as

$$B = \tau N\mathbf{z}_1\mathbf{z}_1^T.$$

Thus, ε-NIMFA on the complete graph (A.6) becomes

$$\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} = \varepsilon^*\sqrt{N}\mathbf{z}_1 - (1+\varepsilon^*)\mathbf{s}_{\text{MF}}(t) + \text{diag}\left(\sqrt{N}\mathbf{z}_1 - \mathbf{s}_{\text{MF}}(t)\right)\tau N\mathbf{z}_1\mathbf{z}_1^T\mathbf{s}_{\text{MF}}(t).$$

Suppose that (2.7) holds at time t. Then, we obtain that

$$\begin{aligned}\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} =& \varepsilon^*\sqrt{N}\mathbf{z}_1 - (1+\varepsilon^*)c_1(t)\mathbf{z}_1 - (1+\varepsilon^*)c_2(t)\mathbf{z}_2 \\ &+ \text{diag}\left(\left(\sqrt{N} - c_1(t)\right)\mathbf{z}_1 - c_2(t)\mathbf{z}_2\right)\tau N\mathbf{z}_1\mathbf{z}_1^T\left(c_1(t)\mathbf{z}_1 + c_2(t)\mathbf{z}_2\right).\end{aligned}$$

Since $\mathbf{z}_1^T\mathbf{z}_1 = 1$ and $\mathbf{z}_1^T\mathbf{z}_2 = 0$, it follows that

$$\begin{aligned}\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} =& \left(\varepsilon^*\sqrt{N} - (1+\varepsilon^*)c_1(t)\right)\mathbf{z}_1 - (1+\varepsilon^*)c_2(t)\mathbf{z}_2 \\ &+ \tau N c_1(t)\,\text{diag}\left(\left(\sqrt{N} - c_1(t)\right)\mathbf{z}_1 - c_2(t)\mathbf{z}_2\right)\mathbf{z}_1,\end{aligned}$$

which is equivalent to

$$\begin{aligned}\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} =& \left(\varepsilon^*\sqrt{N} - (1+\varepsilon^*)c_1(t)\right)\mathbf{z}_1 - (1+\varepsilon^*)c_2(t)\mathbf{z}_2 \\ &+ \tau N\left(\sqrt{N}c_1(t) - c_1^2(t)\right)\text{diag}\left(\mathbf{z}_1\right)\mathbf{z}_1 - \tau N c_1(t)c_2(t)\,\text{diag}\left(\mathbf{z}_2\right)\mathbf{z}_1.\end{aligned}$$

A

From the definition of the agitation mode $\mathbf{z}_1$ in (2.8), we obtain that $\text{diag}(\mathbf{z}_1)\,\mathbf{z}_1 = \mathbf{z}_1/\sqrt{N}$ and $\text{diag}(\mathbf{z}_2)\,\mathbf{z}_1 = \mathbf{z}_2/\sqrt{N}$. Thus, we arrive at

$$\begin{aligned}\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} &= \Big(\varepsilon^*\sqrt{N} - (1+\varepsilon^*)c_1(t)\Big)\mathbf{z}_1 - (1+\varepsilon^*)c_2(t)\mathbf{z}_2 \\ &\quad + \tau\sqrt{N}\Big(\sqrt{N}c_1(t) - c_1^2(t)\Big)\mathbf{z}_1 - \tau\sqrt{N}c_1(t)c_2(t)\mathbf{z}_2,\end{aligned}$$

which simplifies to

$$\begin{aligned}\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} &= \Big(\varepsilon^*\sqrt{N} + (\tau N - 1 - \varepsilon^*)c_1(t) - \tau\sqrt{N}c_1^2(t)\Big)\mathbf{z}_1 \\ &\quad - \Big(1 + \varepsilon^* + \tau\sqrt{N}c_1(t)\Big)c_2(t)\mathbf{z}_2.\end{aligned} \tag{A.7}$$

Hence, the $N \times 1$ viral state vector $\mathbf{s}_{\text{MF}}(t)$ is equal to the linear combination (2.7) of only two agitation modes $\mathbf{z}_1, \mathbf{z}_2$ at every time t. Thus, solving ε-NIMFA on the complete graph simplifies to obtaining a closed-form expression for the functions $c_1(t)$ and $c_2(t)$.

A.3.2. FIRST AGITATION MODE

Since

$$\frac{dc_l(t)}{dt} = \mathbf{z}_l^T\frac{d\mathbf{s}_{\text{MF}}(t)}{dt} \tag{A.8}$$

for both $l = 1, 2$, we obtain for the scalar function $c_1(t)$ from (A.7) that

$$\frac{dc_1(t)}{dt} = \varepsilon^*\sqrt{N} + (\tau N - 1 - \varepsilon^*)c_1(t) - \tau\sqrt{N}c_1^2(t). \tag{A.9}$$

The differential equation (A.9) is separable,

$$\frac{dc_1(t)}{-\tau\sqrt{N}c_1^2(t) + (\tau N - 1 - \varepsilon^*)c_1(t) + \varepsilon^*\sqrt{N}} = dt.$$

Hence, it follows that

$$\frac{dc_1(t)}{c_1^2(t) + \mu_1 c_1(t) - \mu_2} = -\tau\sqrt{N}dt \tag{A.10}$$

with the constants

$$\mu_1 = \frac{1 + \varepsilon^* - \tau N}{\tau\sqrt{N}} \tag{A.11}$$

and

$$\mu_2 = \frac{\varepsilon^*}{\tau}. \tag{A.12}$$

We obtain from (A.10) that

$$\int \frac{dc_1(t)}{c_1^2(t) + \mu_1 c_1(t) - \mu_2} = -\tau\sqrt{N}t + K(c_1(0))$$

for some constant $K(c_1(0)) \in \mathbb{R}$. By integration, it follows that

$$\frac{2}{\sqrt{-4\mu_2-\mu_1^2}} \tan^{-1}\left(\frac{\mu_1+2c_1(t)}{\sqrt{-4\mu_2-\mu_1^2}}\right) = -\tau\sqrt{N}t + K(c_1(0)),$$

which yields that

$$\frac{\mu_1+2c_1(t)}{\sqrt{-4\mu_2-\mu_1^2}} = \tan\left(\frac{1}{2}\sqrt{-4\mu_2-\mu_1^2}\left(-\tau\sqrt{N}t + K(c_1(0))\right)\right).$$

We isolate for $c_1(t)$ and arrive at

$$c_1(t) = -\frac{1}{2}\mu_1 + \frac{1}{2}\sqrt{-4\mu_2-\mu_1^2}\tan\left(\frac{1}{2}\sqrt{-4\mu_2-\mu_1^2}\left(-\tau\sqrt{N}t + K(c_1(0))\right)\right).$$

With the imaginary unit $\mathrm{i} = \sqrt{-1}$, it follows that

$$c_1(t) = -\frac{1}{2}\mu_1 + \frac{1}{2}\mathrm{i}\sqrt{\mu_1^2+4\mu_2}\tan\left(\frac{1}{2}\mathrm{i}\sqrt{\mu_1^2+4\mu_2}\left(-\tau\sqrt{N}t + K(c_1(0))\right)\right).$$

Finally, with the relation $\mathrm{i}\tan(\mathrm{i}\,x) = -\tanh(x)$ of the tangent and hyperbolic tangent and with $-\tanh(x) = \tanh(-x)$ for all $x \in \mathbb{R}$, we arrive at

$$c_1(t) = -\frac{1}{2}\mu_1 + \frac{1}{2}\sqrt{\mu_1^2+4\mu_2}\tanh\left(\frac{1}{2}\sqrt{\mu_1^2+4\mu_2}\left(\tau\sqrt{N}t - K(c_1(0))\right)\right). \tag{A.13}$$

To further simplify (A.13), we obtain with the definition of the constants μ_1 and μ_2 in (A.11) and (A.12) that

$$\sqrt{\mu_1^2+4\mu_2} = \sqrt{\frac{(1+\varepsilon^*-\tau N)^2}{\tau^2 N} + 4\frac{\varepsilon^*}{\tau}},$$

which simplifies to

$$\sqrt{\mu_1^2+4\mu_2} = \frac{1}{\tau\sqrt{N}} w_{\varepsilon^*}, \tag{A.14}$$

where we define the *viral slope* w_{ε^*} for ε-NIMFA as

$$w_{\varepsilon^*} = \sqrt{(1+\varepsilon^*-\tau N)^2 + 4\varepsilon^*\tau N}.$$

If $\varepsilon^* = 0$, then the viral slope w_{ε^*} equals $w_{\varepsilon^*} = |w|$, where w is the viral slope of the NIMFA model, defined in [150]. With (A.14) and the definition of μ_1 in (A.11), the function $c_1(t)$ in (A.13) becomes

$$c_1(t) = \frac{1}{2}\frac{\tau N - 1 - \varepsilon^*}{\tau\sqrt{N}} + \frac{1}{2}\frac{1}{\tau\sqrt{N}} w_{\varepsilon^*} \tanh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)$$

A

with the constant $\Upsilon_{1,\varepsilon^*}(c_1(0)) = -\frac{1}{2}\frac{1}{\tau\sqrt{N}} w_{\varepsilon^*} K(c_1(0))$. Thus, it holds that

$$c_1(t) = \frac{1}{2\tau\sqrt{N}}\left(\tau N - 1 - \varepsilon^* + w_{\varepsilon^*}\tanh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)\right). \tag{A.15}$$

The initial condition of the projection $c_1(t)$ is given by $c_1(0) = \mathbf{z}_1^T \mathbf{s}_{\mathrm{MF}}(0)$, which yields for the constant $\Upsilon_{\varepsilon^*}(c_1(0))$ that

$$\Upsilon_{1,\varepsilon^*}(c_1(0)) = \operatorname{arctanh}\left(\frac{1}{w_{\varepsilon^*}}\left(2\tau\sqrt{N}\mathbf{z}_1^T\mathbf{s}_{\mathrm{MF}}(0) - \tau N + 1 + \varepsilon^*\right)\right).$$

A.3.3. SECOND AGITATION MODE

From (A.8) for $l = 2$, it follows with (A.7) that the scalar function $c_2(t)$ obeys

$$\frac{dc_2(t)}{dt} = -\left(1 + \varepsilon^* + \tau\sqrt{N}c_1(t)\right)c_2(t). \tag{A.16}$$

The closed-form expression for the function $c_1(t)$ is given in (A.15), and the differential equation (A.16) becomes

$$\begin{aligned}\frac{dc_2(t)}{dt} &= -\left(1+\varepsilon^*\right)c_2(t) - \frac{1}{2}\left(\tau N - 1 - \varepsilon^* + w_{\varepsilon^*}\tanh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)\right)c_2(t)\\ &= -\frac{1+\varepsilon^*+\tau N}{2}c_2(t) - \frac{w_{\varepsilon^*}}{2}\tanh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)c_2(t).\end{aligned}$$

The remaining steps are similar to [48, Proof of Theorem 4]. Since

$$\frac{d\log(c_2(t))}{dt} = \frac{1}{c_2(t)}\frac{dc_2(t)}{dt},$$

we obtain that

$$\frac{d\log(c_2(t))}{dt} = -\frac{1+\varepsilon^*+\tau N}{2} - \frac{w_{\varepsilon^*}}{2}\tanh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right). \tag{A.17}$$

The integral of the hyperbolic tangent equals [171]

$$\int \tanh(\xi)\,d\xi = \log(\cosh(\xi)) + C$$

for an arbitrary constant $C \in \mathbb{R}$, where $\cosh(\xi)$ denotes the hyperbolic cosine. Hence, we obtain with the chain rule from (A.17) that

$$\log(c_2(t)) = -\frac{1+\varepsilon^*+\tau N}{2}t - \frac{w_{\varepsilon^*}}{2}\frac{2}{w_{\varepsilon^*}}\log\left(\cosh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)\right) + C$$

for some constant $C \in \mathbb{R}$. Thus, exponentiation yields that the function $c_2(t)$ equals

$$\begin{aligned}c_2(t) &= \exp(C)\exp\left(-\frac{1+\varepsilon^*+\tau N}{2}t\right)\left(\cosh\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right)\right)^{-1}\\ &= \Upsilon_{2,\varepsilon^*}(c_2(0))\exp\left(-\frac{1+\varepsilon^*+\tau N}{2}t\right)\operatorname{sech}\left(\frac{w_{\varepsilon^*}}{2}t + \Upsilon_{1,\varepsilon^*}(c_1(0))\right),\end{aligned}$$

where we denote the hyperbolic secant by $\operatorname{sech}(\xi) = \cosh(\xi)^{-1}$. The constant $\Upsilon_{2,\varepsilon^*}(c_2(0))$ follows from the initial condition $c_2(0) = \mathbf{z}_2^T \mathbf{s}_{\mathrm{MF}}(0)$ as

$$\Upsilon_{2,\varepsilon^*}(c_2(0)) = \mathbf{z}_2^T \mathbf{s}_{\mathrm{MF}}(0) \cosh\left(\Upsilon_{1,\varepsilon^*}(c_1(0))\right). \tag{A.18}$$

□

A.4. EIGENMODE TRUNCATION OF THE ε-SIS PROCESS

If the metastable state exists, Figure 2.2 depicts roughly three regimes for the time-varying prevalence: (I) initial phase, (II) metastable behaviour and (III) convergence to the steady state. Since the behaviour is rather limited, we expect that the time-dependent dynamics of the ε-SIS process can be accurately approximated. Apart from the largest eigenvalue $\lambda_1 = 0$ and the second-largest eigenvalue λ_2, the majority of the eigenvalues is largely clustered. The $N+1$-sized linear process (2.1) can perhaps be approximated accurately with only $m \ll N+1$ eigenvalues and eigenvectors. Eigenmode approximation has been applied successfully for the mean-field SIS model around the epidemic threshold [150], but to the best of our knowledge, no results have been obtained for the Markovian SIS process.

We approximate the solution (2.5) by considering only the m largest eigenvalues and corresponding eigenvectors:

$$\tilde{\mathbf{s}}(t) = \sum_{k=1}^{m} c_k e^{\lambda_k t} \mathbf{w}_k, \tag{A.19}$$

where $c_k = \mathbf{v}_k^T \mathbf{s}(0)$. Approximating the exact solution (2.5) by the approximation (A.19) introduces an error;

$$e(t) = \|\mathbf{s}(t) - \tilde{\mathbf{s}}(t)\| = \left\| \sum_{k=m+1}^{N+1} c_k e^{\lambda_k t} \mathbf{w}_k \right\| \leq \sum_{k=m+1}^{N+1} e^{\lambda_k t} \|c_k \mathbf{w}_k\|,$$

where $\|\cdot\|$ denotes a vector norm and we used the vector inequality $\|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$. Further, since eigenvectors are normalized with norm 1,

$$e(t) \leq \sum_{k=m+1}^{N+1} e^{\lambda_k t} \|c_k \mathbf{w}_k\| < e^{\lambda_{m+1} t} \sum_{k=m+1}^{N+1} |c_k|.$$

Thus the error scales as $e(t) = \mathcal{O}(e^{\lambda_{m+1} t})$. Figure A.2 shows the solution (2.5) and the approximation (A.19) for various choices of m. The original solution (2.5) is recovered accurately using only $m = 8$ from the total of 31 eigenvalues.

We define the critical number of eigenvalues m^* as the smallest m for which $|y(0) - \tilde{y}(0)| \leq 10^{-3}$, where $y(0)$ and $\tilde{y}(0)$ describe the initial prevalence of the exact solution (2.5) and approximated solution (A.19), respectively. The critical number m^* is an integer between 1 and $N+1$. Figure A.3 shows an apparent linear relationship between the critical number m^* and the network size N. Even though an accurate approximation is possible, the number of required eigenvalues grows linearly with the network size N, effectively rendering the approximation method infeasible for large networks. This contrasts results for mean-field ε-SIS process with arbitrary initial conditions but homogeneous parameters, where the number of required equations reduces to only two (see Appendix A.3).

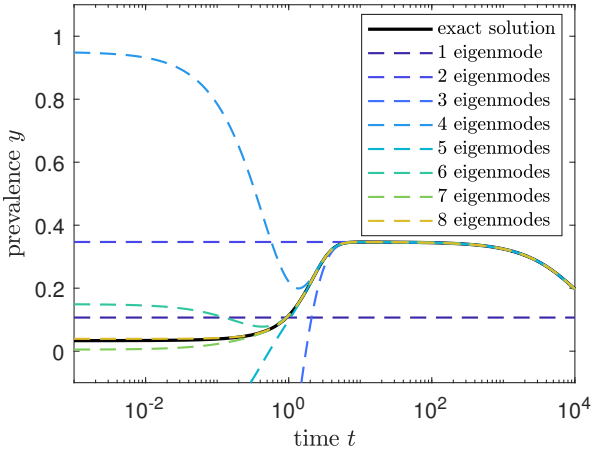


Figure A.2: The time-varying prevalence $y(t)$ for the complete graph with $N = 30$ nodes, effective infection rate $\tau = 2.5\tau_c^{(1)}$ and effective self-infection rate $\varepsilon^* = 10^{-6}$. The solid line indicates the solution (2.5) and the dashed lines show the approximation (A.19) for various values of m.

A.5. EIGENVALUE APPROXIMATIONS AND BOUNDS

This section contains the proofs for the eigenvalues λ_k of the transition matrix P in equation (2.2) in several parameters limits. Most proofs use the symmetric transition matrix $\tilde{P}$, which is derived in Appendix A.1. We start by presenting Theorem A.1.

Theorem A.1 (Based on [172, pp. 366 - 372] and [173, pp. 303-321]) *Given an Hermitian matrix $P = A + \alpha B$ where A and B are Hermitian matrices and α is a small parameter, such that the element a_{ij} is strictly larger than αb_{ij} for all α and all $1 \leq i, j \leq N$. We assume that the eigenvalues $\lambda_k^{(0)}$ and eigenvectors $\boldsymbol{x}_k^{(0)}$ of the unperturbed matrix A can be computed easily.*

*(a) If an eigenvalue $\lambda_k^{(0)}$ of A is **simple**, then the eigenvalue λ_k of P is up to fourth order;*

$$\lambda_k = \lambda_k^{(0)} + \alpha W_{kk} + \alpha^2 \sum_{\substack{l=1 \\ l\neq k}}^{N} \frac{|W_{lk}|^2}{\lambda_k^{(0)} - \lambda_l^{(0)}} + \alpha^3 \sum_{\substack{l=1 \\ l\neq k}}^{N} (W_{ll} - W_{kk}) \left(\frac{W_{kl}}{\lambda_k^{(0)} - \lambda_l^{(0)}} \right)^2 + \alpha^3 \sum_{\substack{l=1 \\ l\neq k}}^{N} \sum_{\substack{m=1 \\ m\neq k \\ m\neq l}}^{N} \frac{W_{kl} W_{lm} W_{mk}}{(\lambda_k^{(0)} - \lambda_l^{(0)})(\lambda_k^{(0)} - \lambda_m^{(0)})} + \mathcal{O}(\alpha^4),$$

where $W_{ij} = (\boldsymbol{x}_i^{(0)})^T B \boldsymbol{x}_j^{(0)}$.

*(b) If an eigenvalue $\lambda_k^{(0)}$ of A is **2-fold single degenerate**[1], then the eigenvalue λ_k of P*

[1] An eigenvalue λ of the matrix A is **2-fold single degenerate** if the eigenvalue λ appears twice in the spectrum of A and is no longer degenerate after adding the first correction term.

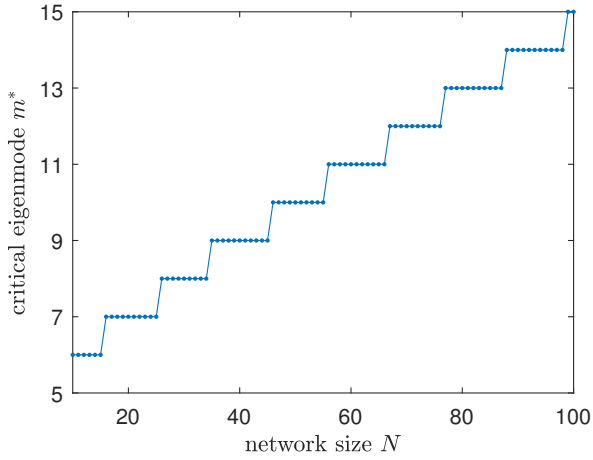


Figure A.3: The relation between the critical number of eigenvalues m^* and the number of nodes N in the graph. The critical number m^* is defined as the smallest integer m for which the initial prevalence $\tilde{y}(0)$ of the approximated solution (A.19) differs at most 10^{-3} from the initial prevalence $y(0)$ of the exact solution (2.1).

is up to third order;

$$\lambda_k = \lambda_k^{(0)} + \alpha\lambda_k^{(1)} + \alpha^2 \sum_{\substack{l=1 \\ l\neq k_1 \\ l\neq k_2}}^{N} \frac{|W'_{lk}|^2}{\lambda_k^{(0)} - \lambda_l^{(0)}} + \mathcal{O}(\alpha^3),$$

where the eigenvalues $\lambda_k^{(1)}$ are determined from the eigenvalue equation

$$\begin{pmatrix} W_{k_1,k_1} & W_{k_1,k_2} \\ W_{k_2,k_1} & W_{k_2,k_2} \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} = \lambda_k^{(1)} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix}, \tag{A.20}$$

where $W_{ij} = (\boldsymbol{x}_i^{(0)})^T B \boldsymbol{x}_j^{(0)}$. The indices k_1 and k_2 correspond to the degenerate eigenvalues $\lambda_k^{(0)} = \lambda_{k_1}^{(0)} = \lambda_{k_2}^{(0)}$. The corrected zero-order eigenvectors are $\boldsymbol{x}_k'^{(0)} = \alpha_{k_1}\boldsymbol{x}_{k_1}^{(0)} + \beta_{k_2}\boldsymbol{x}_{k_2}^{(0)}$. The second-order correction term requires the W' matrix, which has elements $W'_{ij} = (\boldsymbol{x}_i'^{(0)})^T B \boldsymbol{x}_j'^{(0)}$.

(c) If an eigenvalue $\lambda_k^{(0)}$ of A is ***2-fold double degenerate***[2]*, then the eigenvalue λ_k of P is up to third order;*

$$\lambda_k = \lambda_k^{(0)} + \frac{\alpha}{2}\left(W_{k_1,k_1} + W_{k_2,k_2} \pm \sqrt{(W_{k_1,k_1} - W_{k_2,k_2})^2 + 4W_{k_1,k_2}W_{k_2,k_1}}\right)$$
$$+ \frac{\alpha^2}{2}\left(M_{k_1,k_1} + M_{k_2,k_2} \pm \sqrt{(M_{k_1,k_1} - M_{k_2,k_2})^2 + 4M_{k_1,k_2}M_{k_2,k_1}}\right) + \mathcal{O}(\alpha^3),$$

[2] An eigenvalue λ of the matrix A is **2-fold double degenerate** if the eigenvalue λ appears twice in the spectrum of A and is no longer degenerate after adding the second-order correction term.

where $W_{ij} = (\boldsymbol{x}_i^{(0)})^T B \boldsymbol{x}_j^{(0)}$ *and*

$$M_{ij} = \sum_{\substack{m=1 \\ m\neq i \\ m\neq j}}^{N} \frac{W_{i,m} W_{m,j}}{\lambda_i^{(0)} - \lambda_m^{(0)}}.$$

Sketch of the proof. Instead of providing a proof, we sketch the idea of the proof here. An actual proof can be based on [172] and/or [173].

The primary goal of this theorem is to approximate the eigenvalues of a matrix P, for which exact computations are generally infeasible. Defining a small variable α, we split up the Hermitian (or here, symmetric) matrix P into $P = A + \alpha B$, where A is a diagonal matrix and B contains the remaining, symmetric terms. We emphasise that both A and B may contain functions of α, as long as the element a_{ij} is larger than αb_{ij} for all i, j in the limit $\alpha \to 0$. The division of P into A and B is generally not unique, but the number of choices is heavily restricted by the small parameter α.

If the eigenvalue λ_k of the matrix A is non-degenerate, one can perform an ordinary eigenvalue expansion which can be found in any textbook covering perturbation expansions of linear operators (part (a)). Some eigenvalues may appear multiple times in the spectrum of A, which complicates the analysis. In our case, each eigenvalue exists at most twice, thus we confine ourselves to two-fold degeneracy. The procedure can be easily generalised to n-fold degeneracy.

If the eigenvalues $\lambda_k^{(0)}$ are two-fold degenerate, the corresponding eigenvectors are not determined up to a scalar value, but are only known to be in the span of two vectors. We can choose the eigenvectors freely, as long as they belong to the span and are orthonormal to each other. If the eigenvalues are distinct at the first order, that is, the degeneracy is lifted at the first order (part (b)), we can determine the zeroth-order eigenvectors and use that basis as if we would perform a regular expansion as in part (a), with the exception that the summation over all terms excludes both the current index k as well as the index k' corresponding to the same zero-order eigenvalue $\lambda_k^{(0)} = \lambda_{k'}^{(0)}$ and that we use the zeroth-order eigenvectors $\boldsymbol{x}'^{(0)}_k$ instead of the original $\boldsymbol{x}_k^{(0)}$.

If the first-order correction of the eigenvalue still maintains the degeneracy, the second-order correction of the eigenvalues must be computed (part (c)). As before, if the eigenvalues are no longer degenerate after adding the second-order correction, the zeroth-order eigenvectors can be determined. The procedure can be repeated up to higher orders, which is outside of the scope of this theorem. □

A.5.1. THE LIMIT $\varepsilon^* \to \infty$

Proof. The symmetric transition matrix $\tilde{P}$ from equation (A.1) can be rewritten as

$$\tilde{P} = \tilde{A} + \tilde{B}$$

where $\tilde{A}$ is a diagonal matrix with elements

$$\tilde{A}_{kk} = -\varepsilon^*(N-k), \qquad \text{for } k = 0, 1, \ldots, N,$$

and $\tilde{B}$ is a symmetric, tridiagonal matrix with elements

$$\tilde{B}_{kk} = -\tau k(N-k) - k,$$
$$\tilde{B}_{k-1,k} = \sqrt{k(N-k+1)((k-1)\tau + \varepsilon^*)}.$$

We define the small parameter $\alpha = \frac{1}{\sqrt{\varepsilon^*}}$ (which is small, because $\varepsilon^* \to \infty$), such that $\alpha^2 \tilde{P} = A + \alpha B$, where we defined $A = \alpha^2 \tilde{A}$ and $B = \alpha \tilde{B}$. The matrix A is a diagonal matrix with elements

$$A_{kk} = -(N-k)$$

and the matrix B is then tridiagonal with elements

$$B_{kk} = -\alpha k(\tau(N-k)+1),$$
$$B_{k-1,k} = \sqrt{k(N-k+1)((k-1)\tau\alpha^2 + 1)}.$$

The eigenvalues of A are simply $a_k = -(N-k)$ for $k = 0, 1, \ldots, N$ and the corresponding eigenvector $\mathbf{x}_k = \mathbf{e}_k$, where $\mathbf{e}_k$ is the $(N+1)\times 1$ all-zeros vector, except at entry k where it is one. Since $\mathbf{x}_k = \mathbf{e}_k$, we immediately find that $W_{lk} = \mathbf{x}_l^T B \mathbf{x}_k = B_{kl}$. In the same manner, we find $a_k - a_l = k - l$. Given the uniqueness of the eigenvalues a_k, we follow part (a) of Theorem A.1 to find

$$\alpha^2 \lambda_k = -(N-k) + \alpha B_{kk} + \alpha^2 \sum_{\substack{l=1 \\ l\neq k}}^{N} \frac{(B_{kl})^2}{k-l} + \alpha^3 \sum_{\substack{l=1 \\ l\neq k}}^{N} (B_{ll} - B_{kk}) \left(\frac{B_{lk}}{k-l}\right)^2$$
$$+ \alpha^3 \sum_{\substack{l=1 \\ l\neq k}}^{N} \sum_{\substack{m=1 \\ m\neq k \\ m\neq l}}^{N} \frac{B_{lk}B_{ml}B_{km}}{(k-l)(k-m)} + \mathcal{O}(\alpha^4).$$

The value B_{kl} is only non-zero if $l \in \{k-1, k, k+1\}$. We conclude that $B_{lk}B_{ml}B_{km} = 0$ if $k \neq l \neq m$. Thus

$$\alpha^2 \lambda_k = -(N-k) + \alpha B_{kk} + \alpha^2 \sum_{\substack{l=1 \\ l\neq k}}^{N} \frac{(B_{kl})^2}{k-l} + \alpha^3 \sum_{\substack{l=1 \\ l\neq k}}^{N} (B_{ll} - B_{kk}) \left(\frac{B_{lk}}{k-l}\right)^2 + \mathcal{O}(\alpha^4).$$

The α^2-term can be worked out as follows

$$\sum_{\substack{l=1 \\ l\neq k}}^{N} \frac{(B_{kl})^2}{k-l} = \left(\frac{B_{k,k-1}^2}{k-(k-1)} + \frac{B_{k,k+1}^2}{k-(k+1)}\right)$$
$$= (\tau\alpha^2(k-1)+1)(N-k+1)k - (\tau\alpha^2 k + 1)(N-k)(k+1)$$
$$= \tau\alpha^2\left((k-1)(N-k+1)k - k(N-k)(k+1)\right) + \left((N-k+1)k - (N-k)(k+1)\right)$$
$$= \tau\alpha^2 k(3k-2N-1) + (2k-N).$$

A

The α^3-term can be worked out as follows

$$\sum_{\substack{l=1\\l\neq k}}^{N}(B_{ll}-B_{kk})\left(\frac{B_{lk}}{k-l}\right)^2=\alpha\left((2\tau k-\tau N-\tau-1)\frac{B_{k,k-1}^2}{(k-(k-1))^2}+(2\tau k-\tau N+\tau-1)\frac{B_{k,k+1}^2}{(k-(k+1))^2}\right)$$

$$\begin{aligned}
&=\alpha(2\tau k-\tau N-\tau-1)(\tau\alpha^2(k-1)+1)(N-k+1)k\\
&\quad+\alpha(2\tau k-\tau N+\tau-1)(\tau\alpha^2 k+1)(N-k)(k+1)\\
&=-8\alpha^3k^3\tau^2+3\alpha^3k^2\tau^2+3\alpha^3k^2\tau-6\alpha k^2\tau+9\alpha^3k^2N\tau^2-\alpha^3k\tau^2-\alpha^3k\tau\\
&\quad+2\alpha k-2\alpha^3kN^2\tau^2-\alpha^3kN\tau^2-2\alpha^3kN\tau+6\alpha kN\tau-\alpha N^2\tau+\alpha N\tau-\alpha N\\
&=-6\alpha k^2\tau+2\alpha k+6\alpha kN\tau-\alpha N^2\tau+\alpha N\tau-\alpha N+\mathcal{O}(\alpha^3).
\end{aligned}$$

Assembling the results, we find

$$\begin{aligned}
\alpha^2\lambda_k=&-(N-k)+\alpha^2\left(-\tau k(N-k)-k\right)+\alpha^2\left(\tau\alpha^2k(3k-2N-1)+(2k-N)\right)\\
&+\alpha^4\left(-6k^2\tau+2k+6kN\tau-N^2\tau+N\tau-N\right)+\mathcal{O}(\alpha^4).
\end{aligned}$$

Using $\alpha^2=1/\varepsilon^*$, we find

$$\begin{aligned}
\lambda_k=&-\varepsilon^*(N-k)-(\tau k(N-k)+k)+\left(\frac{1}{\varepsilon^*}\tau k(3k-2N-1)+(2k-N)\right)\\
&+\frac{1}{\varepsilon^*}\left(-6k^2\tau+2k+6kN\tau-N^2\tau+N\tau-N\right)+\mathcal{O}\left(\frac{1}{\varepsilon^*}\right).
\end{aligned}$$

The index $k=0,1,\ldots,N$ can be transformed as to make sure that the eigenvalues λ_k are descending: $0=\lambda_1>\lambda_2>\ldots>\lambda_{N+1}$. We define the index $\hat{k}=N-k+1$, which takes values in $\hat{k}=1,2,\ldots,N+1$ such that we recover

$$\begin{aligned}
\lambda_{\hat{k}}=&-(\hat{k}-1)\varepsilon^*\\
&-(\hat{k}-1)(\tau(N+1-\hat{k})+1)\\
&+\left(\tau(2N+2\hat{k}N-2-5\hat{k}-3\hat{k}^2)+N-2-2\hat{k}\right)\frac{1}{\varepsilon^*}\\
&+\mathcal{O}\left(\frac{1}{\varepsilon^*}\right),
\end{aligned}$$

which concludes our proof. Presumably the $\mathcal{O}(\alpha^4)$-terms will bring additional terms for the $\mathcal{O}\left(\frac{1}{\varepsilon^*}\right)$-term in the final solution, thus our estimate is only correct up to $\mathcal{O}\left(\frac{1}{\varepsilon^*}\right)$. □

A.5.2. THE LIMIT $\tau\to\infty$

If the effective infection rate τ tends to infinity, then the dynamics of the ε-SIS process simplifies to an SI process. Metastability cannot be observed in the SI process, because the number of infected nodes only increases, until all nodes are infected. Further information on the SI process is provided in Chapter 3.

Proof. The symmetric transition matrix $\tilde{P}$ from equation (A.1) can be rewritten as

$$\tilde{P}=\tilde{A}+\tilde{B}$$

where $\tilde{A}$ is a diagonal matrix with elements

$$\tilde{A}_{kk} = -\tau k(N-k), \qquad \text{for } k = 0, 1, \ldots, N,$$

and $\tilde{B}$ is a symmetric, tridiagonal matrix with elements

$$\tilde{B}_{kk} = -\varepsilon^*(N-k) - k,$$
$$\tilde{B}_{k-1,k} = \sqrt{k(N-k+1)((k-1)\tau + \varepsilon^*)}.$$

We define the small parameter $\alpha = \frac{1}{\sqrt{\tau}}$ (which is small, because $\tau \to \infty$), such that $\alpha^2 \tilde{P} = A + \alpha B$, where we defined $A = \alpha^2 \tilde{A}$ and $B = \alpha \tilde{B}$. The matrix A is then diagonal with elements

$$A_{kk} = -k(N-k)$$

and the matrix B is then tridiagonal with elements

$$B_{kk} = -\alpha(\varepsilon^*(N-k) + k),$$
$$B_{k-1,k} = \sqrt{k(N-k+1)(k-1+\varepsilon^*\alpha^2)}.$$

The eigenvalues of A are simply $a_k = -k(N-k)$ for $k = 0, 1, \ldots, N$ and the corresponding eigenvector $\mathbf{x}_k = \mathbf{e}_k$, where $\mathbf{e}_k$ is the $(N+1) \times 1$ all-zeros vector, except at entry k where it is one. We distinguish between networks with even and odd sizes and treat the case $\varepsilon^* = 1$ with special care.

Case 1: Even size N

Consider a graph with an even number of nodes N. Then the matrix A has one simple eigenvalue a_k with index $k = N/2$. The remaining eigenvalues are 2-fold degenerate. Let a_k and a_{N-k} be a pair of degenerate eigenvalues. Then the eigenvalue equation is according to part (b) from Theorem A.1;

$$\begin{pmatrix} B_{k,k} & B_{k,N-k} \\ B_{N-k,k} & B_{N-k,N-k} \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix} = \lambda_k^{(1)} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix}. \tag{A.21}$$

Since the number of nodes N is even, we know that $B_{k,N-k} = B_{N-k,k} = 0$ for all $k \neq N/2$. Thus the eigenvalue correction equals

$$\lambda_k^{(1)} = B_{k,k} = -\alpha(\varepsilon^*(N-k) + k),$$
$$\lambda_{N-k}^{(1)} = B_{N-k,N-k} = -\alpha(\varepsilon^* k + (N-k)),$$

which are different for all $k \neq \frac{N}{2}$ and all even network sizes N, except when $\varepsilon^* = 1$. If $\varepsilon^* \neq 1$, the eigenvectors $(\alpha_k \ \beta_k)^T$ equal the elementary vectors $\mathbf{e}_k$, which implies[3] that part (a) from Theorem A.1 can be used instead of part (b).

[3] Please consult the sketch of the proof of Theorem A.1 for the complete reasoning.

A

Applying part (a) from Theorem A.1, we find

$$\begin{aligned}\alpha^2\lambda_k &= a_k + \alpha B_{kk} + \alpha^2 \sum_{\substack{l=1\\ l\neq k}}^{N} \frac{(B_{lk})^2}{a_k - a_l} + \mathcal{O}(\alpha^3)\\ &= -k(N-k) - \alpha^2(\varepsilon^*(N-k)+k) + \alpha^2\left(\frac{(B_{k-1,k})^2}{a_k - a_{k-1}} + \frac{(B_{k+1,k})^2}{a_k - a_{k+1}}\right) + \mathcal{O}(\alpha^3)\\ &= -k(N-k) + \alpha^2\left(-\varepsilon^*(N-k) + \frac{k(N+1)(N-k)}{(2k-N-1)(2k-N+1)}\right) + \mathcal{O}(\alpha^3).\end{aligned}$$

Using $\alpha^2 = \frac{1}{\tau}$, we find the following relationship for the eigenvalues;

$$\lambda_k = -\tau k(N-k) + \left(-\varepsilon^*(N-k) + \frac{k(N+1)(N-k)}{(2k-N-1)(2k-N+1)}\right) + \mathcal{O}\left(\frac{1}{\sqrt{\tau}}\right). \tag{A.22}$$

For $k = \frac{N}{2}$ we simply find

$$\lambda_{N/2} = -\tau\frac{N^2}{4} - \frac{N}{2}\left(\varepsilon^* + \frac{N}{2}(N+1)\right) + \mathcal{O}\left(\frac{1}{\sqrt{\tau}}\right). \tag{A.23}$$

From Eq. (A.23), we may conclude that (A.22) is only valid if $\tau > \frac{2\varepsilon^*}{N}$ and $\tau > N+1$.

Case 2: Odd size N and $\varepsilon^* \neq 1$

For odd network sizes N, all eigenvalues a_k are two-fold degenerate. All eigenvalues can be computed using (A.22) provided that $\varepsilon^* \neq 1$. However, special attention is required for eigenvalues a_k with indices $k = \frac{N+1}{2}$ and $k = \frac{N-1}{2}$. In that case, the eigenvalue equation (A.20) becomes

$$\begin{pmatrix} -\frac{\alpha}{2}(\varepsilon^*(N+1)+N-1) & \frac{N+1}{2}\sqrt{\frac{N-1}{2}+\varepsilon^*\alpha^2} \\ \frac{N+1}{2}\sqrt{\frac{N-1}{2}+\varepsilon^*\alpha^2} & -\frac{\alpha}{2}(\varepsilon^*(N-1)+N+1)\end{pmatrix}\begin{pmatrix}\alpha_{(N-1)/2}\\ \beta_{(N-1)/2}\end{pmatrix} = \lambda^{(1)}_{(N-1)/2}\begin{pmatrix}\alpha_{(N-1)/2}\\ \beta_{(N-1)/2}\end{pmatrix}$$

whose eigenvalues are distinct:

$$\lambda^{(1)}_{(N-1)/2} = \frac{\alpha(N-\varepsilon^*)}{2} \pm \sqrt{\left(\frac{\alpha(N-\varepsilon^*)}{2}\right)^2 + \frac{1}{8}(N+1)^2(N-1)}.$$

The corresponding eigenvectors are

$$\begin{pmatrix}\alpha_{(N-1)/2}\\ \beta_{(N-1)/2}\end{pmatrix} = \begin{pmatrix}\alpha(N-\varepsilon^*) \pm \sqrt{\alpha^2(N-\varepsilon^*)^2 + (N+1)^2\left(\frac{N-1}{2}+\varepsilon^*\alpha^2\right)}\\ (N+1)\sqrt{\frac{N-1}{2}+\varepsilon^*\alpha^2}\end{pmatrix}.$$

We will not continue our analysis for the $\mathcal{O}(\alpha^2)$ terms, because the computations are tedious. Our final result is

$$\alpha^2\lambda_{(N-1)/2} = -\frac{N-1}{2}\frac{N+1}{2} + \alpha\frac{\alpha(N-\varepsilon^*)}{2} \pm \sqrt{\left(\frac{\alpha(N-\varepsilon^*)}{2}\right)^2 + \frac{1}{8}(N+1)^2(N-1)} + \mathcal{O}(\alpha^2).$$

Using $\alpha^2 = \frac{1}{\tau}$, we find the following relationship for the eigenvalues;

$$\lambda_{(N-1)/2} = -\frac{1}{4}\tau(N-1)(N+1) \pm \sqrt{\tau}(N+1)\sqrt{\frac{N-1}{2}} + \mathcal{O}(1). \tag{A.24}$$

The relation (A.24) is a valid perturbation expansion if $\tau > \frac{8}{N-1}$. Since the computations of the second-order terms are tedious, we have omitted them here. The key observation is that the eigenvalue $\lambda_{(N-1)/2}$ scales with an $\mathcal{O}(\sqrt{\tau})$ term, which is not the case for even-sized networks. We have not found any physical or intuitive reasoning why this is the case.

Case 3: Even size N and $\varepsilon^* = 1$

We construct the M-matrix for $k \neq N/2$:

$$M = \begin{pmatrix} M_{kk} & M_{k,N-k} \\ M_{N-k,k} & M_{N-k,N-k} \end{pmatrix}$$

where

$$M_{k,N-k} = \sum_{\substack{m=1 \\ m\neq k \\ m\neq N-k}}^{N} \frac{W_{k,m}W_{m,N-k}}{\lambda_k^{(0)} - \lambda_m^{(0)}}.$$

We find

$$M_{kk} = \frac{k(N-k)(N+1)}{(2k-N-1)(2k-N+1)} + k + \alpha^2 \frac{k^2 + (N-k)^2 + N}{(2k-N-1)(2k-N+1)}.$$

In most cases $M_{k,N-k} = 0$ because the product $W_{k,m}W_{m,N-k}$ is only non-zero for $m = k-1$ or $m = k+1$, and $m = N-k+1$ or $m = N-k-1$. Hence, the product is only non-zero if $k = \frac{N}{2} - 1$. Thus, for $k \neq \frac{N}{2} - 1$, the second-order correction of the eigenvalues is unique and the eigenvalues follow as

$$\begin{aligned}\alpha^2\lambda_k = &-k(N-k) - \alpha^2 N + \\ &\alpha^2\left(\frac{k(N-k)(N+1)}{(2k-N-1)(2k-N+1)} + k + \alpha^2 \frac{k^2 + (N-k)^2 + N}{(2k-N-1)(2k-N+1)}\right) + \mathcal{O}(\alpha^3)\end{aligned}$$

Using $\alpha^2 = \frac{1}{\tau}$, we find the following relationship for the eigenvalues;

$$\lambda_k = -\tau k(N-k) + k - N + \frac{k(N-k)(N+1)}{(2k-N-1)(2k-N+1)} + \mathcal{O}\left(\frac{1}{\sqrt{\tau}}\right). \tag{A.25}$$

For $k = \frac{N}{2} - 1$, additional work is required because the off-diagonal terms are non-zero. The M-matrix is in this case

$$M = \begin{pmatrix} \frac{1}{12}(N-2)(N^2+3N+8) & \frac{N}{2}\left(\frac{N}{2}+1\right)\sqrt{\left(\frac{N}{2}-1+\alpha^2\right)\left(\frac{N}{2}+\alpha^2\right)} \\ \frac{N}{2}\left(\frac{N}{2}+1\right)\sqrt{\left(\frac{N}{2}-1+\alpha^2\right)\left(\frac{N}{2}+\alpha^2\right)} & \frac{1}{12}(N+2)(N^2-N+4) \end{pmatrix},$$

whose eigenvalues are

$$\lambda_k^{(2)} = \frac{N^3}{12} + \frac{N^2}{12} + \frac{N}{6} - \frac{1}{3} \pm \sqrt{1 + \left(\frac{N}{2}\right)^3 \left(\frac{N}{2}+1\right)^2 \left(\frac{N}{2}-1\right)}.$$

A

Thus the total eigenvalue expansion equals

$$\lambda_k = -\tau\left(\frac{N}{2}-1\right)\left(\frac{N}{2}+1\right)+\frac{N^3}{12}+\frac{N^2}{12}-\frac{5N}{6}-\frac{1}{3}\pm\sqrt{1+\left(\frac{N}{2}\right)^3\left(\frac{N}{2}+1\right)^2\left(\frac{N}{2}-1\right)}.$$

Case 4: Odd size N and $\varepsilon^* = 1$

For the case $k \neq \frac{N-1}{2}$, the eigenvalues follow from (A.25) and the result for $k = \frac{N-1}{2}$ is described in (A.24).

Finally introducing $\tilde{k} = k+1$, such that the index $\tilde{k}$ runs from 1 to $N+1$ finalises the proof. □

A.5.3. LIMIT $\varepsilon^* \to 0$

For self-infection rates $\varepsilon^* < \frac{1}{N}$, Gershgorin's circle theorem leads to a tight bound for the smallest eigenvalue λ_{N+1}. The case $\tau > \tau_c$ appeared earlier in [56, Corollary 3].

Proof. We denote the (not necessarily ordered) eigenvalues $\lambda_1, \ldots, \lambda_{N+1}$ of the transition matrix P from equation (2.2). Given the scaled birth rate $\tilde{\Xi}_k$ and death rate $\tilde{\mu}_k$, Gershgorin's circle theorem provides the following bounds for the eigenvalues

$$|\lambda_k + \tilde{\Xi}_k + \tilde{\mu}_k| \leq \tilde{\Xi}_{k-1} + \tilde{\mu}_{k+1}.$$

Substituting the scaled birth rate $\tilde{\Xi}_k = (\tau k+1)(N-k)$ and scaled death rate $\tilde{\mu}_k = k$, we find

$$f(k) \leq \lambda_k \leq 1 + \varepsilon^* + \tau(2k-N-1), \tag{A.26}$$

where the lower bound $f(k)$ equals

$$f(k) = -\tau\left[(k-1)(N-k+1)+k(N-k)\right] - \varepsilon^*\left[2N-2k+1\right] - \left[2k+1\right]. \tag{A.27}$$

The lower bound $f(k)$ is negative for all $k = 1, \ldots, N+1$. The upper bound is most negative for $k = 1$, hence $\min_k \lambda_k \leq 1 + \varepsilon^* - \tau(N-1)$. Using $\tau = x\tau_c^{(1)} = \frac{x}{N-1}$, we find $\min_k \lambda_k \leq 1 - x + \varepsilon^* < 0$ above the epidemic threshold ($x > 1$). For the remaining eigenvalues λ_k, the upper bound is larger than zero, which is not confining. Simulations also indicate that the upper bound is very loose. Instead, we focus on the lower bound $f(k)$.

ABOVE THE EPIDEMIC THRESHOLD

The lower bound $f(k)$ is the smallest when $\frac{df}{dk} = 0$, and we find[4]

$$\frac{df}{dk}(\hat{k}) = -\tau\left(2N-4\hat{k}+2\right)+2\varepsilon^*-2 = 0. \tag{A.28}$$

Solving for $\hat{k}$ gives

$$\hat{k} = \frac{1-\varepsilon^*}{2\tau} + \frac{N+1}{2}.$$

[4] We assume here that the function $f(k)$ is varying slowly, which means that the difference between $f(k+1)$ and $f(k)$ is small for all k. Since the function $f(k)$ is a quadratic function in k, the function $f(k)$ indeed varies sufficiently slow.

Since we consider the case $\varepsilon^* < \frac{1}{N}$, the effective self-infection rate ε^* can be neglected;

$$\hat{k} = \frac{1}{2\tau} + \frac{N+1}{2}. \tag{A.29}$$

Hence, $\hat{k}$ is at least larger than $N/2$. The lower bound becomes

$$f(\hat{k}) = -\frac{1}{2\tau} - \frac{3}{2} - N - \frac{\tau N(N-1)}{2} + \varepsilon^* \left[\frac{1}{\tau} - N\right].$$

Using the normalised effective infection rate $x = \tau(N-1)$, we approximately find

$$f(\hat{k}) \approx -\left(\frac{1}{2x} + 1 + \frac{x}{2}\right) N + \left(\frac{1}{2x} - \frac{3}{2}\right). \tag{A.30}$$

The lower bound (A.30) holds for all eigenvalues λ_k where $1 \leq k \leq N+1$. Thus we may conclude that

$$\lambda_{N+1} \gtrsim -\left(\frac{1}{2x} + 1 + \frac{x}{2}\right) N.$$

BELOW THE EPIDEMIC THRESHOLD

If $\tau < \tau_c^{(1)} = \frac{1}{N-1}$, the lower bound $f(k)$ is the smallest for $\hat{k} = N-1$. We find

$$f(\hat{k}) = -\tau(3N-5) - 2N + 1.$$

Using $\tau = \frac{x}{N-1}$ and $x < 1$, we find approximately

$$f(\hat{k}) \approx -2N + 1 - 3x. \tag{A.31}$$

Equation (A.31) holds for all eigenvalues λ_k, with $1 \leq k \leq N+1$, thus we conclude that

$$\lambda_{N+1} \gtrsim -2N.$$

□

A.5.4. THE LIMIT $\tau \to 0$

If the effective infection rate $\tau = 0$, then the ε-SIS model reduces to a birth and death process with linear rates, that can be solved exactly for any time t and any number of initially infected nodes [33]. The eigenvalue ratio ρ in the limit of small effective infection rates τ is approximately $\rho = \lambda_3/\lambda_2 \approx 3/2$, because the second-largest eigenvalue $\lambda_2 \approx -2$ and $\lambda_3 \approx -3$. Thus, metastability cannot be observed for small effective infection rates τ.

A.5.5. FINAL CONSIDERATIONS

Apart from limit cases and bounds, the second-largest eigenvalue λ_2 can be computed using the following approach.

Theorem A.2 (Van Doorn *et al.* [47]) *The convergence rate $-\lambda_2$ equals*

$$\max_{\mathbf{d}>\mathbf{0}} \min_{1\leq k\leq N} \alpha_k = -\lambda_2 = \min_{\mathbf{d}>\mathbf{0}} \max_{1\leq k\leq N} \alpha_k \tag{A.32}$$

where $\mathbf{d} = (d_1, \ldots, d_N)$, *and* $d_i > 0$ *for* $i = 1, \ldots, N$ *and*

$$\alpha_k = \tau \left[\left(\frac{d_{k+1}}{d_k} - 1 \right) k^2 + \left(1 - \frac{d_{k+1}}{d_k} \right) kN + 2k - N - 1 \right] + \varepsilon^* \left[\left(\frac{d_{k+1}}{d_k} - 1 \right) k + \left(1 - \frac{d_{k+1}}{d_k} \right) N + 1 \right] + \left[\left(1 - \frac{d_{k-1}}{d_k} \right) k + \frac{d_{k-1}}{d_k} \right]$$

for $k = 1, \ldots, N$ *and* $d_0 = d_{N+1} = 0$.

Theorem A.2 associates the computation of the second-largest eigenvalue λ_2 of the transition matrix P to the finding of a suitable, positive vector $\mathbf{d}$. Equation (A.32) illustrates that choosing any positive vector $\mathbf{d}$ directly provides lower and upper bounds for the convergence rate λ_2. Unfortunately, simply generating random values for the vector $\mathbf{d}$ does not provide sharp bounds for the second-largest eigenvalue λ_2.

As an example, we consider $\mathbf{d} = \mathbf{u}$ where $\mathbf{u}$ is the $N \times 1$ all-ones vector. According to Theorem A.2, the convergence rate λ_2 is then bounded by

$$-\tau(N-1) + 1 + \varepsilon^* \leq -\lambda_2 \leq \tau(N-1) + 1 + \varepsilon^* \tag{A.33}$$

If the effective infection rate τ is larger than the mean-field epidemic threshold $\tau_c^{(1)} = \frac{1}{N-1}$, the lower bound in (A.33) is negative and therefore not confining. For $\tau < \tau_c$, the lower bound in (A.33) appears to be a loose bound. In the limit $\tau \to 0$, the convergence rate equals $-\lambda_2 = 1 + \varepsilon^*$, which agrees with Theorem 2.5. The upper bound is positive in both cases, and is at least one, but appears to be a loose bound as well.

B

APPENDIX TO CHAPTER 3

B.1. NEWTON'S ROOT-FINDING METHOD

We shortly explain Newton's method of finding the root t^* of a function $f(t)$. In our case, we are interested in the root of $y'(t)$ at t_{peak}, but we write $f(t)$ and t^* for generality.

B.1.1. FIRST-ORDER NEWTON-RAPHSON

The Taylor expansion of the function $f(t^*)$ around $t^* = t$ equals

$$f(t^*) = f_0(t) + (t^* - t)f_1(t) + (t^* - t)f_2(t) + \mathcal{O}((t^* - t)^3), \tag{B.1}$$

where we defined

$$f_k(t) = \frac{f^{(k)}(t)}{k!}.$$

Since t^* is a zero of $f(t)$, we find $f(t^*) = 0$. Only using terms up to first order in (B.1), we find

$$0 = f_0(t) + (t^* - t)f_1(t) + \mathcal{O}((t^* - t)^2),$$

which can be rewritten as

$$t^* = t - \frac{f_0(t)}{f_1(t)} + \mathcal{O}((t - t^*)^2).$$

Taking $t^* = \tilde{t}_{k+1}, t = \tilde{t}_k$ and neglecting the second-order terms, we find the Newton-Raphson iterative scheme:

$$\tilde{t}_{k+1} = \tilde{t}_k - \frac{f_0(\tilde{t}_k)}{f_1(\tilde{t}_k)}.$$

The Newton-Raphson method is known to have quadratic convergence.

B

B.1.2. SECOND-ORDER NEWTON-RAPHSON

Instead of using only first-order terms, we can also use second-order terms in (B.1). Again using $f(t^*) = 0$, we find

$$0 = f(t) + (t^* - t) f_1(t) + (t^* - t)^2 f_2(t) + \mathcal{O}((t^* - t)^3).$$

Neglecting the third-order terms, we can solve for t^*;

$$t^* = t + \frac{-f_1(t) \pm \sqrt{f_1(t)^2 - 4 f_0(t) f_2(t)}}{2 f_2(t)}.$$

Using $t^* = \tilde{t}_{k+1}$ and $t = \tilde{t}_k$, we find the second-order Newton-Raphson iterative scheme:

$$\tilde{t}_{k+1} = \tilde{t}_k + \frac{-f_1(\tilde{t}_k) \pm \sqrt{f_1(\tilde{t}_k)^2 - 4 f_0(\tilde{t}_k) f_2(\tilde{t}_k)}}{2 f_2(\tilde{t}_k)}.$$

Substituting $f(t) = y'(t)$ gives Eq. (3.23). The major advantage of the second-order Newton-Raphson method is that the radius of convergence seems larger than Newton-Raphson and the speed of convergence is faster.

B.2. MEAN-FIELD EQUATIONS

The heterogeneous mean-field approximation of the stochastic SI(S) process is known as the N-Intertwined Mean-Field Approximation (NIMFA) [29] and equals

$$\frac{d\, v_i}{dt} = -\delta_i v_i + (1 - v_i) \sum_{j=1}^{N} v_j \tilde{\beta}_{ij}$$

where $v_i(t)$ is the probability that node i is infected at time t and $\tilde{\beta}_{ij} = \beta_{ij} a_{ij}$.

Similarly, the heterogeneous individual mean-field approximation for the SIR process is given by [143]

$$\begin{aligned} \frac{d\, s_i}{dt} &= -\sum_{j=1}^{N} \tilde{\beta}_{ij} s_i v_j, \\ \frac{d\, v_i}{dt} &= \sum_{j=1}^{N} \tilde{\beta}_{ij} s_i v_j - \delta_i v_i, \end{aligned} \tag{B.2}$$

where $s_i(t)$ and $v_i(t)$ represent the probabilities that node i is susceptible or infected at time t, respectively.

The mean-field SIS and SIR threshold is [122]

$$\tau_c^{(1)} = \lambda_{\max}(S^{-1} B), \tag{B.3}$$

where $S = \mathrm{diag}(\delta_1, \ldots, \delta_N)$ is the $N \times N$ curing rate matrix and B is the $N \times N$ infection rate matrix, whose elements are $\tilde{\beta}_{ij}$.

B.3. ABEL SUMMATION

A finite-n-analysis on the sum $\sum_{k=m}^{n} a_k\, b_k$, where $n \geq m$, illustrates interesting formal manipulations dealing with partial sums, $s_k = \sum_{l=m}^{k} a_l$ where $k \geq m$, attributed to Niels Abel.

The basic observation is that $a_k = s_k - s_{k-1}$ for $k > m$. Thus,

$$\begin{aligned}\sum_{k=m}^{n} a_k\, b_k &= a_m\, b_m + \sum_{k=m+1}^{n} (s_k - s_{k-1})\, b_k \\ &= a_m\, b_m + \sum_{k=m+1}^{n} s_k\, b_k - \sum_{k=m}^{n-1} s_k\, b_{k+1} \\ &= a_m\, b_m + \sum_{k=m+1}^{n-1} s_k\, (b_k - b_{k+1}) + s_n\, b_n - s_m\, b_{m+1}.\end{aligned}$$

Since $s_m = a_m$, we arrive at Abel's partial summation valid for any integer $m \leq n$,

$$\sum_{k=m}^{n} a_k\, b_k = \sum_{k=m}^{n-1} \left(\sum_{l=m}^{k} a_l \right) (b_k - b_{k+1}) + b_n \left(\sum_{l=m}^{n} a_l \right). \tag{B.4}$$

A reversal of the k- and l-sum again returns to the sum at left-hand side.

As an application of Abel summation, we consider the sum $h_n(t) = \sum_{k=0}^{n} a_k\, e^{-\lambda_k t}$, where $0 \leq \mathrm{Re}\,(\lambda_0) \leq \mathrm{Re}\,(\lambda_1) \leq \cdots \leq \mathrm{Re}\,(\lambda_n)$, that can appear as a solution of $n+1$-th order linear differential equation. Abel summation (B.4) yields

$$h_n(t) = \sum_{k=0}^{n-1} \left(\sum_{l=0}^{k} a_l \right) (e^{-\lambda_k t} - e^{-\lambda_{k+1} t}) + e^{-\lambda_n t} \left(\sum_{l=0}^{n} a_l \right).$$

Invoking $-t \int_{\lambda_{k+1}}^{\lambda_k} e^{-xt}\, dx = e^{-\lambda_k t} - e^{-\lambda_{k+1} t}$, we obtain

$$h_n(t) = t \sum_{k=0}^{n-1} \int_{\lambda_k}^{\lambda_{k+1}} \left(\sum_{l=0}^{k} a_l \right) e^{-xt}\, dx + e^{-\lambda_n t} \left(\sum_{l=0}^{n} a_l \right).$$

Let the function $y = \lambda(x)$ define the sequence $\{\lambda_k\}_{0 \leq k \leq n}$ by $\lambda(k) = \lambda_k$ at integer values of x. The inverse function $x = \lambda^{-1}(y)$ maps the index to $k = \lambda^{-1}(\lambda_k)$. Since $k \leq \lambda^{-1}(x) < k+1$ for $\lambda_k \leq x < \lambda_{k+1}$, the integer part $[.]$ of $\left[\lambda^{-1}(x)\right] = k$ for $\lambda_k \leq x < \lambda_{k+1}$ and we have that

$$\sum_{k=0}^{n-1} \int_{\lambda_k}^{\lambda_{k+1}} \left(\sum_{l=0}^{k} a_l \right) e^{-xt}\, dx = \sum_{k=0}^{n-1} \int_{\lambda_k}^{\lambda_{k+1}} e^{-xt} \sum_{l=0}^{[\lambda^{-1}(x)]} a_l\, dx = \int_{\lambda_0}^{\lambda_n} e^{-xt} \sum_{l=0}^{[\lambda^{-1}(x)]} a_l\, dx,$$

which leads to

$$h_n(t) = t \int_{\lambda_0}^{\lambda_n} e^{-xt} \sum_{l=0}^{[\lambda^{-1}(x)]} a_l\, dx + e^{-\lambda_n t} \left(\sum_{l=0}^{n} a_l \right). \tag{B.5}$$

Relation (B.5) shows for $t = 0$ that $h_n(0) = \sum_{l=0}^{n} a_l$ for all n (also when $n \to \infty$).

B

However, if $\lambda_n \to \infty$ for $n \to \infty$ and $\lambda_0 = 0$, then the last sum in (B.5) vanishes for all positive $t > 0$ and

$$\frac{h_\infty(t)}{t} = \frac{1}{t}\sum_{k=0}^{\infty} a_k\, e^{-\lambda_k t} = \int_0^\infty e^{-xt}\, g(x)\, dx, \qquad (t > 0) \tag{B.6}$$

is the Laplace transform of the sum $g(x) = \sum_{l=0}^{[\lambda^{-1}(x)]} a_l$. When t tends to zero as $t = \frac{c}{\lambda_n}$ in the limit $n \to \infty$, then $e^{-\lambda_n t} = e^{-c}$, which is a constant, and the first term in (B.5) vanishes, but the second term remains and $e^{-c} = 1$ so that $h_\infty(0) = \sum_{l=0}^{\infty} a_l$, which is a prerequisite for continuity for $t \geq 0$ as well. Hence, the point $t = 0$ needs care[1].

Observe from $g(x) = \sum_{l=0}^{[\lambda^{-1}(x)]} a_l$ in the limit $n \to \infty$ where $\lambda_n \to \infty$, that $g(0) = a_0 = \lim_{t\to\infty} h_\infty(t)$, while $\lim_{x\to\infty} g(x) = \sum_{l=0}^{\infty} a_l = h_\infty(0)$. The inverse Laplace transform is

$$g(x) = \sum_{l=0}^{[\lambda^{-1}(x)]} a_l = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{h_\infty(t)}{t} e^{xt}\, dt. \qquad (c > 0)$$

The infinitesimal generator Q is minus a weighted Laplacian matrix and the smallest eigenvalue of any Laplacian matrix is zero. Hence, $\lambda_0 = 0$ is satisfied, but λ_n is always finite for any finite graph with n nodes. Only if the graph is sufficiently large, then the Laplace transform (B.6) is exact for $t > 0$, while $h_n(0) = \sum_{l=0}^{n} a_l$ for all n.

[1] If $h_\infty(0) = 0$, then the limit $\lim_{t\to 0} \frac{h_\infty(t)}{t} = h'_\infty(0)$ and (B.6) shows that $\int_0^\infty g(x)\, dx = h'_\infty(0)$.

C

APPENDIX TO CHAPTER 4

C.1. PROOF OF THEOREM 4.1

We follow the method of [101]. The governing equation of $\mathrm{E}[a_{ij}X_i]$ can be computed analogous to (4.1) and (4.2) and equals

$$\frac{d\,\mathrm{E}[a_{ij}X_i]}{dt} = a_{ij}(0)\Big((a_{\mathrm{cr}}+b_{\mathrm{cr}})\xi\mathrm{E}[X_i] - (a_{\mathrm{br}}\zeta + b_{\mathrm{br}}\zeta + a_{\mathrm{cr}}\xi + b_{\mathrm{cr}}\xi + \delta)\mathrm{E}[a_{ij}X_i] \\ + (b_{\mathrm{cr}}+c_{\mathrm{cr}})\xi\mathrm{E}[(1-a_{ij})X_iX_j] - (b_{\mathrm{br}}+c_{\mathrm{br}})\zeta\mathrm{E}[a_{ij}X_iX_j] + \beta\mathrm{E}\left[(1-X_i)a_{ij}\sum_{k=1}^{N}a_{ik}X_k\right]\Big), \tag{C.1}$$

where we used the notation that $a_{ij}(0) = 1$ if $(i,j) \in \mathcal{L}_{\mathrm{adaptive}}$ and $a_{ij}(0) = 0$ otherwise. In this proof, the governing equations of $\mathrm{E}[X_i]$ from (4.1) and $\mathrm{E}[a_{ij}X_i]$ from (C.1) are used. The governing equations (4.1) and (C.1) are rewritten in terms of $\mathrm{E}[X_i]$ and $\mathrm{E}[a_{ij}X_i]$, and the remaining terms are denoted by W. Our goal is to define W such that it is negative. For equation (C.1), we rewrite the infection term with coefficient β as

$$\beta\mathrm{E}\left[(1-X_i)a_{ij}\sum_{k=1}^{N}a_{ik}X_k\right] = \beta\sum_{k=1}^{N}\mathrm{E}\Big[a_{ik}X_k - X_ia_{ik}X_ka_{ij} - (1-a_{ij})a_{ik}X_k\Big].$$

Then equation (C.1) can be rewritten as

$$\frac{d\,\mathrm{E}[a_{ij}X_i]}{dt} = a_{ij}(0)\Big((a_{\mathrm{cr}}+b_{\mathrm{cr}})\xi\mathrm{E}[X_i] - (a_{\mathrm{br}}\zeta + b_{\mathrm{br}}\zeta + a_{\mathrm{cr}}\xi + b_{\mathrm{cr}}\xi + \delta)\mathrm{E}[a_{ij}X_i] \\ + \beta\mathrm{E}\left[\sum_{k=1}^{N}a_{ik}X_k\right] + W_A\Big), \tag{C.2}$$

where the remaining terms of the network W_A are

$$W_A = (b_{\text{cr}} + c_{\text{cr}})\xi \text{E}[(1 - a_{ij})X_i X_j] - (b_{\text{br}} + c_{\text{br}})\zeta \text{E}[a_{ij} X_i X_j] \\ - \beta \sum_{k=1}^{N} \text{E}\Big[X_i X_k a_{ik} a_{ij} + (1 - a_{ij}) a_{ik} X_k \Big].$$

Similarly for $\text{E}[X_i]$:

$$\frac{d\,\text{E}[X_i]}{dt} = -\delta \text{E}[X_i] + \beta \sum_{k=1}^{N} \text{E}[a_{ik} X_k] + W_X, \tag{C.3}$$

where the remaining terms for the nodes W_X are

$$W_X = -\beta \sum_{k=1}^{N} \text{E}[X_i a_{ik} X_j].$$

The remaining term W_X is always negative, whereas W_A is only negative in some cases. Each positive term in W_A is merged with other terms to ensure that W_A is negative. The term W_A is surely negative when each of the individual components is negative. The term with infection rate β is negative. For the link-breaking rate ζ, the case $b_{\text{br}} + c_{\text{br}} = -1$ is a potential problem. By applying $b_{\text{br}} + c_{\text{br}} = -1$, Table 4.1 illustrates that $a_{\text{br}} + b_{\text{br}} = 1$. Then we combine terms from (C.2) with W_A in the following way:

$$-(a_{\text{br}} + b_{\text{br}})\zeta \text{E}[a_{ij} X_i] - (b_{\text{br}} + c_{\text{br}})\zeta \text{E}[a_{ij} X_i X_j] = -\zeta \text{E}[a_{ij} X_i (1 - X_j)], \qquad \text{if } b_{\text{br}} + c_{\text{br}} = -1.$$

Therefore we propose the following changes to ensure that W_A is negative for the link-breaking coefficients ζ:

$$\begin{aligned} \text{In (C.2):} \quad & -(a_{\text{br}} + b_{\text{br}})\zeta \text{E}[a_{ij} X_i] \quad \to \quad -\zeta \mathbf{1}_{\{a_{\text{br}}=0, b_{\text{br}}=1, c_{\text{br}}=-1\}} \text{E}[a_{ij} X_i], \\ \text{In } W_A\text{:} \quad & -(b_{\text{br}} + c_{\text{br}})\zeta \text{E}[a_{ij} X_i X_j] \quad \to \quad -\zeta (\mathbf{1}_{\{a_{\text{br}}=0, b_{\text{br}}=0, c_{\text{br}}=1\}} + \mathbf{1}_{\{a_{\text{br}}=1, b_{\text{br}}=-1, c_{\text{br}}=2\}}) \text{E}[a_{ij} X_i X_j] \\ & \qquad -\zeta (\mathbf{1}_{\{a_{\text{br}}=1, b_{\text{br}}=0, c_{\text{br}}=-1\}} + \mathbf{1}_{\{a_{\text{br}}=0, b_{\text{br}}=1, c_{\text{br}}=-2\}}) \text{E}[a_{ij} X_i (1 - X_j)]. \end{aligned}$$

where $\mathbf{1}_f$ is the indicator function, which is 1 if the link-breaking rule f_{br} satisfies f and is zero otherwise. We repeat the procedure for the link-creation term ξ, except that we apply an extra trick: we add zero to equation (C.2), where $\varepsilon > 0$ is small:

$$-\varepsilon \text{E}[X_i (1 - a_{ij})] + \varepsilon \text{E}[X_i] - \varepsilon \text{E}[a_{ij} X_i] = 0,$$

such that

$$\begin{aligned} \text{In (C.2):} \quad & (a_{\text{cr}} + b_{\text{cr}})\xi \text{E}[X_i], \quad \to \quad \{\xi (1 - \mathbf{1}_{\{a_{\text{cr}}=1, b_{\text{cr}}=-1, c_{\text{cr}}=1\}}) + \varepsilon\} \text{E}[X_i] \\ \text{In } W_A\text{:} \quad & \\ (b_{\text{cr}} + c_{\text{cr}})\xi \text{E}[(1 - a_{ij})X_i X_j] \quad & \to \quad -\xi (\mathbf{1}_{\{a_{\text{cr}}=1, b_{\text{cr}}=0, c_{\text{cr}}=-1\}} + \mathbf{1}_{\{a_{\text{cr}}=0, b_{\text{cr}}=1, c_{\text{cr}}=-2\}}) \text{E}[(1 - a_{ij})X_i X_j] \\ & \qquad -\xi (\mathbf{1}_{\{a_{\text{cr}}=0, b_{\text{cr}}=0, c_{\text{cr}}=1\}} + \mathbf{1}_{\{a_{\text{cr}}=1, b_{\text{cr}}=-1, c_{\text{cr}}=2\}}) \text{E}[a_{ij} X_i X_j] \\ & \qquad -\xi (\mathbf{1}_{\{a_{\text{cr}}=0, b_{\text{cr}}=0, c_{\text{cr}}=1\}} + \mathbf{1}_{\{a_{\text{cr}}=1, b_{\text{cr}}=-1, c_{\text{cr}}=2\}}) \text{E}[X_i (1 - X_j)] \\ & \qquad -\varepsilon \text{E}[X_i (1 - a_{ij})], \\ 0 \cdot \text{E}[a_{ij} X_i] \quad & \to \quad -\varepsilon \text{E}[a_{ij} X_i]. \end{aligned}$$

The differential equations (C.2) and (C.3) can be written in matrix notation. Given a sequence of matrices $A_1, \ldots, A_n$, define $A = \oplus_{i=1}^{n} A_i$ to be the block diagonal matrix with A_i on its block diagonals; A_i does not necessarily have to be square. The matrix A can be visualised as

$$A = \begin{pmatrix} A_1 & 0 & 0 & \ldots \\ 0 & A_2 & 0 & \ldots \\ \vdots & \ddots & \ddots & \ddots \\ \ldots & 0 & 0 & A_n \end{pmatrix}.$$

Define the vectors $q_i = \underset{j:a_{ij}(0)=1}{\text{col}}\left(\mathrm{E}[a_{ij}X_i]\right)$ and $q = \underset{1\le i\le N}{\text{col}}\left(q_i\right)$. Moreover, we define T_i as the row vector satisfying

$$T_i q = \sum_{k:a_{ik}(0)=1} \mathrm{E}[a_{ki}X_k].$$

Here T_i is a boolean row vector containing ones when an initial link is present between node i and node j (where j is the j'th element of T_i) and zero otherwise. The dimension of T_i is therefore $1 \times 2L_{\text{adaptive}}$ where L_{adaptive} is the number of links in the adaptive graph G_{adaptive}. Then define the matrix $T = \underset{1\le i\le N}{\text{col}}\left(T_i\right)$. Also define the matrix $J = \oplus_{i=1}^{n} \mathbf{1}_{d_i}$ where d_i is the number of degrees of node i in the initial network. Finally, define the matrix $S = \underset{1\le i\le N}{\text{col}}\left(\mathbf{1}_{d_i} \otimes T_i\right)$ where $\otimes$ is the Kronecker product. To summarise, the following parameters have been defined;

$$\begin{aligned} q_i &= \underset{j:a_{ij}(0)=1}{\text{col}}\left(\mathrm{E}[a_{ij}X_i]\right), \\ q &= \underset{1\le i\le N}{\text{col}}\left(q_i\right), \\ T_i q &= \sum_{k:a_{ik}(0)=1} \mathrm{E}[a_{ki}X_k], \\ T &= \underset{1\le i\le N}{\text{col}}\left(T_i\right), \\ J &= \oplus_{i=1}^{n} \mathbf{1}_{d_i}, \\ S &= \underset{1\le i\le N}{\text{col}}\left(\mathbf{1}_{d_i} \otimes T_i\right). \end{aligned}$$

The differential equations (C.2) and (C.3) can be formulated in the following way:

$$\frac{d}{dt}\begin{pmatrix} \mathrm{E}[X_i] \\ \mathrm{E}[a_{ij}X_i] \end{pmatrix} = M\begin{pmatrix} \mathrm{E}[X_i] \\ \mathrm{E}[a_{ij}X_i] \end{pmatrix} + \begin{pmatrix} W_X \\ W_A \end{pmatrix}, \tag{C.4}$$

where

$$M = \begin{pmatrix} -\delta I & \beta T \\ \{\xi(1 - \mathbf{1}_{\{a_{\text{cr}}=1,b_{\text{cr}}=-1,c_{\text{cr}}=1\}}) + \varepsilon\}J & \beta S - (\zeta \mathbf{1}_{\{a_{\text{br}}=0,b_{\text{br}}=1,c_{\text{br}}=-1\}} + a_{\text{cr}}\xi + b_{\text{cr}}\xi + \delta + \varepsilon)I \end{pmatrix}. \tag{C.5}$$

Since the remaining terms W_X and W_A are negative by construction, it follows that

$$\frac{d}{dt}\begin{pmatrix} \mathrm{E}[X_i] \\ \mathrm{E}[a_{ij}X_i] \end{pmatrix} \le M\begin{pmatrix} \mathrm{E}[X_i] \\ \mathrm{E}[a_{ij}X_i] \end{pmatrix}. \tag{C.6}$$

If the eigenvalues of the matrix M are smaller than zero, the solution is bounded by an exponentially decaying function. Then the solution dies out over sufficiently large time. The point where one of the eigenvalues of M becomes zero, changes the solution from an exponentially decaying function to an exponentially growing function. This bifurcation point is commonly known as the epidemic threshold. To derive a bound for the epidemic threshold, the eigenvalues of M are investigated. Specifically, the largest (real) eigenvalue is of interest and can be determined by using the Perron-Fröbenius theory.

Lemma C.1 *Given a positive eigenvector $\boldsymbol{x}$ of M, its corresponding eigenvalue is the largest eigenvalue of M.*

Proof. The adaptive network G_{adaptive} was taken to be connected. Since the network is undirected, it is also strongly connected. Ogura and Preciado (2016) proved that the matrix M is irreducible when the initial network is strongly connected [101, Appendix A]. Then, by Perron-Fröbenius theory for irreducible matrices, the statement follows [174, Theorem 8.4.4]. □

Based on Lemma C.1, our approach is to construct a positive eigenvector for the matrix M. Using the positive eigenvector, a lower bound for the epidemic threshold is computed.

Proof of Theorem 4.1. First a positive eigenvector is constructed for the matrix M. Since the adaptive network is strongly connected, there exists a positive eigenvector $\mathbf{v}$ corresponding to eigenvalue λ_1 (the spectral radius) [101]. We define the vector $\mathbf{w} = \underset{1\le i\le N}{\text{col}}\left(v_i \mathbf{1}_{d_i}\right)$. Using the definition of T_i, it follows that

$$T_i\mathbf{w} = \sum_{k:\,a_{ik}(0)=1} w_{ki} = \sum_{k:\,a_{ik}(0)=1} v_k = (Av)_i = \lambda_1 v_i. \tag{C.7}$$

So $T\mathbf{w} = \lambda_1\mathbf{v}$. Equivalently, it follows that $S\mathbf{w} = \lambda_1\mathbf{w}$ and $J\mathbf{v} = \mathbf{w}$.

Define the vector $\boldsymbol{x} = \begin{pmatrix} z\mathbf{v} \\ \mathbf{w} \end{pmatrix}$ where $z \in \mathbb{R}$, which is an eigenvector of M. Indeed,

$$\begin{aligned} M\begin{pmatrix} z\mathbf{v} \\ \mathbf{w} \end{pmatrix} &= \begin{pmatrix} -\delta I & \beta T \\ \{\xi(1-\mathbf{1}_{\{a_{\text{cr}}=1,b_{\text{cr}}=-1,c_{\text{cr}}=1\}})+\varepsilon\}J & \beta S-(\mathbf{1}_{\{a_{\text{br}}=0,b_{\text{br}}=1,c_{\text{br}}=-1\}}\zeta+a_{\text{cr}}\xi+b_{\text{cr}}\xi+\delta)I \end{pmatrix}\begin{pmatrix} z\mathbf{v} \\ \mathbf{w} \end{pmatrix} \\ &= \begin{pmatrix} (\beta\lambda_1 - z\delta)\mathbf{v} \\ \{\xi(1-\mathbf{1}_{\{a_{\text{cr}}=1,b_{\text{cr}}=-1,c_{\text{cr}}=1\}})+\varepsilon\}z+\beta\lambda_1-(\mathbf{1}_{\{a_{\text{br}}=0,b_{\text{br}}=1,c_{\text{br}}=-1\}}\zeta+a_{\text{cr}}\xi+b_{\text{cr}}\xi+\delta))\mathbf{w} \end{pmatrix} \\ &= \mu\begin{pmatrix} z\mathbf{v} \\ \mathbf{w} \end{pmatrix}, \end{aligned}$$

where the eigenvalue μ corresponds to the eigenvector $\boldsymbol{x}$. Since $\mathbf{v}$ and $\mathbf{w}$ are positive, the eigenvector $\boldsymbol{x}$ is positive if and only if $z > 0$. To guarantee that $z > 0$, a system of equations for z and μ is obtained;

$$z\mu = \beta\lambda_1 - z\delta, \tag{C.8a}$$

$$\mu = \{\xi(1-\mathbf{1}_{\{a_{\text{cr}}=1,b_{\text{cr}}=-1,c_{\text{cr}}=1\}})+\varepsilon\}z+\beta\lambda_1-(\mathbf{1}_{\{a_{\text{br}}=0,b_{\text{br}}=1,c_{\text{br}}=-1\}}\zeta+a_{\text{cr}}\xi+b_{\text{cr}}\xi+\delta+\varepsilon). \tag{C.8b}$$

Define $X = \xi(1-\mathbf{1}_{\{a_{cr}=1,b_{cr}=-1,c_{cr}=1\}})+\varepsilon$ and $Y = -\beta\lambda_1+\mathbf{1}_{\{a_{br}=0,b_{br}=1,c_{br}=-1\}}\zeta+(a_{cr}+b_{cr})\xi+\delta+\varepsilon$ and notice that $X>0$. Then (C.8a) and (C.8a) simplify to

$$z(\mu+\delta) = \beta\lambda_1, \tag{C.9a}$$

$$\mu = zX - Y. \tag{C.9b}$$

Inserting (C.9a) into (C.9b), we find a quadratic equation for z:

$$Xz^2 + (\delta - Y)z - \beta\lambda_1 = 0. \tag{C.10}$$

Based on (C.10), we find that $z_1 < 0, z_2 > 0$. The corresponding values for μ can be obtained using (C.9a), which can be rewritten as

$$\mu = \frac{\beta\lambda_1}{z} - \delta.$$

Since $\beta, \delta, \lambda_1 > 0$, for $z_1 < 0$ it follows that $\mu_1 < 0$. For $z_2 > 0$, the sign of μ cannot be determined. However, we require that $z_2 > 0$ to have a positive eigenvector and we require $\mu_2 < 0$ for stability. From the system given by (C.9a) and (C.9b), the quadratic equation for μ can be derived:

$$\mu^2 + (\delta+Y)\mu + \underbrace{(\delta Y - \beta\lambda_1 X)}_{\text{constant term}} = 0.$$

We have concluded earlier that $\mu_1 < 0$. The eigenvalues of M are required to be negative, hence $\mu_2 < 0$. When μ_1, μ_2 are negative, the constant term of the quadratic equation is positive, which leads to the condition

$$\delta Y - \beta\lambda_1 X > 0.$$

Substitution of the definition of X and Y and rewriting yields

$$\frac{\beta}{\delta} < \frac{\mathbf{1}_{\{a_{br}=0,b_{br}=1,c_{br}=-1\}}\zeta + a_{cr}\xi + b_{cr}\xi + \delta + \varepsilon}{\lambda_1(\xi(1-\mathbf{1}_{\{a_{cr}=1,b_{cr}=-1,c_{cr}=1\}})+\delta+\varepsilon)},$$

such that the final form becomes

$$\tau < \frac{1}{\lambda_1}\left(1 + \frac{\mathbf{1}_{\{a_{br}=0,b_{br}=1,c_{br}=-1\}}\omega - (\mathbf{1}_{\{a_{cr}=1,b_{cr}=0,c_{cr}=-1\}\cup\{a_{cr}=0,b_{cr}=1,c_{cr}=-2\}})}{(1-\mathbf{1}_{\{a_{cr}=1,b_{cr}=-1,c_{cr}=1\}})+\delta/\xi+\varepsilon/\xi}\right). \tag{C.11}$$

Since we did not assume any value for ε, we take $\lim_{\varepsilon\to 0}$. Eq. (C.11) is a required condition for the process to exponentially decay to zero over sufficiently large time. Therefore the epidemic threshold τ_c needs to be larger than those τ-values, which proves Theorem 4.1. □

C.2. PROOF OF THEOREM 4.2

We follow the method of [52]. Using (4.2) and the general formulation of any updating rule of the G-ASIS model in (4.4), we find

$$\frac{d\,\mathrm{E}[a_{ij}]}{dt} = \mathrm{E}\Big[-\zeta a_{ij}(a_{br}+b_{br}(X_i+X_j)+c_{br}X_iX_j)+\xi(1-a_{ij})(a_{cr}+b_{cr}(X_i+X_j)+c_{cr}X_iX_j)\Big].$$

Using that the expectation operator is a linear operator, we find

$$\frac{d\,\mathrm{E}[a_{ij}]}{dt} = a_{\mathrm{cr}}\xi + b_{\mathrm{cr}}\xi\mathrm{E}[X_i] + b_{\mathrm{cr}}\xi\mathrm{E}[X_j] - (a_{\mathrm{br}}\zeta + a_{\mathrm{cr}}\xi)\mathrm{E}[a_{ij}] + c_{\mathrm{cr}}\xi\mathrm{E}[X_iX_j] \\ - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}[a_{ij}X_i] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}[a_{ij}X_j] - (c_{\mathrm{br}}\zeta + c_{\mathrm{cr}}\xi)\mathrm{E}[a_{ij}X_iX_j].$$

Taking the sum over all $j \neq i$ and using the degree $d_i = \sum_{j=1,j\neq i}^{N} a_{ij}$ and $a_{ii} = 0$, we obtain

$$\frac{d\,\mathrm{E}[d_i]}{dt} = a_{\mathrm{cr}}\xi(N-1) + b_{\mathrm{cr}}\xi(N-1)\mathrm{E}[X_i] + b_{\mathrm{cr}}\xi\sum_{j=1,j\neq i}^{N}\mathrm{E}[X_j] - (a_{\mathrm{br}}\zeta + a_{\mathrm{cr}}\xi)\mathrm{E}[d_i] \\ + c_{\mathrm{cr}}\xi\mathrm{E}\left[X_i\sum_{j=1,j\neq i}^{N}X_j\right] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}[d_iX_i] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_j\right] \\ - (c_{\mathrm{br}}\zeta + c_{\mathrm{cr}}\xi)\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_iX_j\right].$$

Two terms need to be investigated in more detail. The following relations hold:

$$(N-1)\mathrm{E}[X_i] + \sum_{j=1,j\neq i}^{N}\mathrm{E}[X_j] = ((N-2))\,\mathrm{E}[X_i] + \sum_{j=1}^{N}\mathrm{E}[X_j],$$
$$\mathrm{E}\left[X_i\sum_{j=1,j\neq i}^{N}X_j\right] = \mathrm{E}\left[X_i\left(\sum_{j=1}^{N}X_j - X_i\right)\right] = \mathrm{E}\left[X_i\sum_{j=1}^{N}X_j\right] - \mathrm{E}[X_i],$$

where in the last equation, for the last equality, we used the Bernoulli property $E[X_i^2] = E[X_i]$. Substituting these back yields

$$\frac{d\,\mathrm{E}[d_i]}{dt} = a_{\mathrm{cr}}\xi(N-1) + (b_{\mathrm{cr}}\xi(N-2) - c_{\mathrm{cr}}\xi)\,\mathrm{E}[X_i] + b_{\mathrm{cr}}\xi\sum_{j=1}^{N}\mathrm{E}[X_j] - (a_{\mathrm{br}}\zeta + a_{\mathrm{cr}}\xi)\mathrm{E}[d_i] \\ + c_{\mathrm{cr}}\xi\mathrm{E}\left[X_i\sum_{j=1}^{N}X_j\right] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}[d_iX_i] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_j\right] \\ - (c_{\mathrm{br}}\zeta + c_{\mathrm{cr}}\xi)\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_iX_j\right].$$

Up to now only the network equations from (4.2) have been used. We intend to use the epidemic equations in (4.1) to remove the largest correlation term. Hence, we rewrite (4.1) as

$$\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_iX_j\right] = -\frac{1}{\beta}\frac{d\,\mathrm{E}[X_i]}{dt} - \frac{1}{\tau}\mathrm{E}[X_i] + \mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_j\right],$$

where $\tau = \beta/\delta$ is the effective infection rate. Inserting this back into the previous result gives

$$\begin{aligned}\frac{d\,\mathrm{E}[d_i]}{dt} &= a_{\mathrm{cr}}\xi(N-1) + (b_{\mathrm{cr}}\xi(N-2) - c_{\mathrm{cr}}\xi)\,\mathrm{E}[X_i] + b_{\mathrm{cr}}\xi\sum_{j=1}^{N}\mathrm{E}[X_j] - (a_{\mathrm{br}}\zeta + a_{\mathrm{cr}}\xi)\mathrm{E}[d_i] \\ &\quad + c_{\mathrm{cr}}\xi\mathrm{E}\left[X_i\sum_{j=1}^{N}X_j\right] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}[d_iX_i] - (b_{\mathrm{br}}\zeta + b_{\mathrm{cr}}\xi)\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_j\right] \\ &\quad - (c_{\mathrm{br}}\zeta + c_{\mathrm{cr}}\xi)\left(-\frac{1}{\beta}\frac{d\,\mathrm{E}[X_i]}{dt} - \frac{1}{\tau}\mathrm{E}[X_i] + \mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_j\right]\right).\end{aligned}$$

Taking all time-derivatives to the left and dividing every term by ζ, we obtain

$$\begin{aligned}&\frac{d}{dt}\mathrm{E}\left[\frac{d_i}{\zeta} - \frac{c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1}}{\beta}X_i\right] \\ &= a_{\mathrm{cr}}\omega^{-1}(N-1) + \left(b_{\mathrm{cr}}\omega^{-1}(N-2) - c_{\mathrm{cr}}\omega^{-1} + \frac{c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1}}{\tau}\right)\mathrm{E}[X_i] \\ &\quad + b_{\mathrm{cr}}\omega^{-1}\sum_{j=1}^{N}\mathrm{E}[X_j] - (a_{\mathrm{br}} + a_{\mathrm{cr}}\omega^{-1})\mathrm{E}[d_i] + c_{\mathrm{cr}}\omega^{-1}\mathrm{E}\left[X_i\sum_{j=1}^{N}X_j\right] \\ &\quad - (b_{\mathrm{br}} + b_{\mathrm{cr}}\omega^{-1})\mathrm{E}[d_iX_i] - (b_{\mathrm{br}} + b_{\mathrm{cr}}\omega^{-1} + c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{j=1}^{N}a_{ij}X_j\right].\end{aligned}$$

Using $2L = \sum_{i=1}^{N} d_i$ where L is the number of links, we sum over all $1 \le i \le N$ to find

$$\begin{aligned}&\frac{d}{dt}\mathrm{E}\left[\frac{2L}{\zeta} - \frac{c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1}}{\beta}\sum_{i=1}^{N}X_i\right] \\ &= a_{\mathrm{cr}}\omega^{-1}N(N-1) + \left(b_{\mathrm{cr}}\omega^{-1}(N-2) - c_{\mathrm{cr}}\omega^{-1} + \frac{c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1}}{\tau}\right)\sum_{i=1}^{N}\mathrm{E}[X_i] \\ &\quad + b_{\mathrm{cr}}\omega^{-1}N\sum_{j=1}^{N}\mathrm{E}[X_j] - (a_{\mathrm{br}} + a_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{i=1}^{N}d_i\right] + c_{\mathrm{cr}}\omega^{-1}\mathrm{E}\left[\sum_{i=1}^{N}X_i\sum_{j=1}^{N}X_j\right] \\ &\quad - (b_{\mathrm{br}} + b_{\mathrm{cr}}\omega^{-1})\sum_{i=1}^{N}\mathrm{E}[d_iX_i] - (b_{\mathrm{br}} + b_{\mathrm{cr}}\omega^{-1} + c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{j=1}^{N}d_jX_j\right].\end{aligned}$$

Using the fraction of infected nodes $Z = \frac{1}{N}\sum_{i=1}^{N} X_i$, we can simplify this to

$$\begin{aligned}
&\frac{d}{dt}\mathrm{E}\left[\frac{2L}{\zeta} - \frac{c_{\mathrm{br}}N + c_{\mathrm{cr}}\omega^{-1}N}{\beta}Z\right] \\
&= a_{\mathrm{cr}}\omega^{-1}N(N-1) + \left(b_{\mathrm{cr}}\omega^{-1}N(N-2) - c_{\mathrm{cr}}N\omega^{-1} + \frac{c_{\mathrm{br}}N + c_{\mathrm{cr}}\omega^{-1}N}{\tau}\right)\mathrm{E}[Z] \\
&+ b_{\mathrm{cr}}\omega^{-1}N^2\mathrm{E}[Z] - (a_{\mathrm{br}} + a_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{i=1}^{N} d_i\right] + c_{\mathrm{cr}}\omega^{-1}N^2\mathrm{E}[Z^2] \\
&- (b_{\mathrm{br}} + b_{\mathrm{cr}}\omega^{-1} + b_{\mathrm{br}} + b_{\mathrm{cr}}\omega^{-1} + c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{j=1}^{N} d_j X_j\right].
\end{aligned}$$

When the derivative on the left-hand side vanishes (in the metastable state, which we denote by as asterisk $*$) we have

$$\begin{aligned}
&\left(b_{\mathrm{cr}}\omega^{-1}N(N-1) - b_{\mathrm{cr}}N\omega^{-1} + b_{\mathrm{cr}}\omega^{-1}N^2 - c_{\mathrm{cr}}N\omega^{-1} + \frac{c_{\mathrm{br}}N + c_{\mathrm{cr}}\omega^{-1}N}{\tau}\right)\mathrm{E}[Z^*] \\
&\quad - (a_{\mathrm{br}} + a_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{i=1}^{N} d_i^*\right] + c_{\mathrm{cr}}\omega^{-1}N^2\mathrm{E}[(Z^*)^2] + a_{\mathrm{cr}}\omega^{-1}N(N-1) \\
&\quad - (2b_{\mathrm{br}} + 2b_{\mathrm{cr}}\omega^{-1} + c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{j=1}^{N} d_j^* X_j^*\right] = 0.
\end{aligned}$$

Using $\mathrm{Var}[Z^*] = \mathrm{E}[(Z^*)^2] - \mathrm{E}[Z^*]^2$ and the prevalence $y = \mathrm{E}[Z^*]$, we finally find

$$\begin{aligned}
&c_{\mathrm{cr}}\omega^{-1}N^2y^2 + \left(2b_{\mathrm{cr}}\omega^{-1}N(N-1) - c_{\mathrm{cr}}N\omega^{-1} + \frac{c_{\mathrm{br}}N + c_{\mathrm{cr}}\omega^{-1}N}{\tau}\right)y \\
&+ a_{\mathrm{cr}}\omega^{-1}N(N-1) - (a_{\mathrm{br}} + a_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{i=1}^{N} d_i^*\right] + c_{\mathrm{cr}}\omega^{-1}N^2\mathrm{Var}[Z^*] \\
&- (2b_{\mathrm{br}} + 2b_{\mathrm{cr}}\omega^{-1} + c_{\mathrm{br}} + c_{\mathrm{cr}}\omega^{-1})\mathrm{E}\left[\sum_{i=1}^{N} d_i^* X_i^*\right] = 0
\end{aligned}$$

which is a quadratic equation in y. Since c_{cr} is never zero, every term can be multiplied by $\frac{\omega}{c_{\mathrm{cr}}N^2}$, which proves Theorem 4.2. □

C.3. PROOF OF THEOREM 4.3

The quadratic equation for the prevalence y from (4.6) can be rewritten in more compact form by defining

$$V = -\frac{2b_{\mathrm{cr}}N\tau - (2b_{\mathrm{cr}} + c_{\mathrm{cr}})\tau + c_{\mathrm{br}}\omega + c_{\mathrm{cr}}}{2c_{\mathrm{cr}}N\tau}, \tag{C.12}$$

$$\begin{aligned} H = \frac{(N-1)a_{\mathrm{cr}}}{c_{\mathrm{cr}}N} - \frac{a_{\mathrm{br}}\omega + a_{\mathrm{cr}}}{c_{\mathrm{cr}}N^2}\mathrm{E}\left[\sum_{i=1}^{N} d_i^*\right] + \mathrm{Var}[Z^*] \\ - \frac{(2b_{\mathrm{br}} + c_{\mathrm{br}})\omega + 2b_{\mathrm{cr}} + c_{\mathrm{cr}}}{c_{\mathrm{cr}}N^2}\mathrm{E}\left[\sum_{i=1}^{N} d_i^* X_i^*\right], \end{aligned} \tag{C.13}$$

such that (4.6) can be written as

$$y^2 - 2Vy + H = 0. \tag{C.14}$$

The two possible solutions are

$$y = V \pm \sqrt{V^2 - H}. \tag{C.15}$$

The quadratic equation (C.14) for the prevalence y can be rewritten as

$$V = \frac{1}{2}\left(\frac{H}{y} + y\right). \tag{C.16}$$

Using the definition of V from (C.12), (C.16) can be rewritten as

$$-\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}} + \frac{(2b_{\mathrm{cr}} + c_{\mathrm{cr}})}{2c_{\mathrm{cr}}N} - \frac{c_{\mathrm{br}}\omega + c_{\mathrm{cr}}}{2c_{\mathrm{cr}}N\tau} = \frac{1}{2}\left(\frac{H}{y} + y\right),$$

which can be rearranged to

$$\tau = \frac{c_{\mathrm{br}}\omega + c_{\mathrm{cr}}}{2c_{\mathrm{cr}}N\left(-\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}} + \frac{(2b_{\mathrm{cr}}+c_{\mathrm{cr}})}{2c_{\mathrm{cr}}N} - \frac{1}{2}\left(\frac{H}{y} + y\right)\right)}. \tag{C.17}$$

Taking[1] the limit $y \to 0$, we find an implicit relationship for the epidemic threshold;

$$\tau_c = \frac{\frac{c_{\mathrm{br}}}{c_{\mathrm{cr}}}\omega + 1}{2\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}}(1-N) + 1 - N\lim_{y\downarrow 0}\frac{H}{y}}. \tag{C.18}$$

Since $c_{\mathrm{br}}, c_{\mathrm{cr}} \neq 0$, equation (C.18) is an explicit relation between the epidemic threshold τ_c and the effective link-breaking rate ω. The function H defined in (C.13) depends on ω, ξ and τ and $H = 0$ zero if $y = 0$. Since we have taken $\lim_{y\downarrow 0}$, we have $H(\omega, \xi, \tau_c)$,

[1] We implicitly assume here that the metastable state exists, because we consider the limit of $y \to 0$ from above. During numerical simulations, we have observed that the quadratic equation for the prevalence may have one or two solutions, depending on the choice of the link-breaking and link-creation mechanisms. Thus, taking the limit $y \to 0$ is allowed, because there always exist non-zero solutions for the metastable prevalence y.

which makes equation (C.18) an implicit relation for the epidemic threshold τ_c. Our main effort will be to show the dependence of the epidemic threshold τ_c on the effective link-breaking rate ω by bounding $H(\omega,\xi,\tau_c)$. Due to the continuity of H, we may define

$$h(\omega,\xi,\tau_c) \equiv \lim_{y\downarrow 0} \frac{H}{y}, \tag{C.19}$$

such that the epidemic threshold τ_c becomes

$$\tau_c = \frac{\frac{c_{\mathrm{br}}}{c_{\mathrm{cr}}}\omega + 1}{2\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}}(1-N)+1-Nh(\omega,\xi,\tau_c)}, \tag{C.20}$$

which proves the first part of Theorem 4.3. Next the function $h(\omega,\xi,\tau_c)$ must be bounded. Based on the sign of H, we split up the remainder of the proof of Theorem 4.3 into Lemma C.2 and Lemma C.3.

Lemma C.2 *Let τ_c be the epidemic threshold from (C.20) and assume $H \geq 0$. Then τ_c is bounded by a linear function in ω or by a constant.*

Proof. The only instances of G-ASIS which do not satisfy $H \geq 0$, are instances satisfying $\frac{a_{\mathrm{cr}}}{c_{\mathrm{cr}}} < 0$ and correspond to the link-creation rule $f_{\mathrm{cr}} = 1 - X_i X_j$. These instances are not included in this lemma, but are taken care of by Lemma C.3. This means 30 out of 36 instances of G-ASIS are treated in this lemma. We follow the approach of [97].

Step 1. The prevalence y is real.
The solutions of the quadratic equation (C.14) for the prevalence y need to have a positive discriminant in order to be real solutions. From (C.15), it is required that $H \leq V^2$. Since $H \geq 0$, it is sufficient to show that $\sqrt{H} \leq V$. Inserting the definition of V from (C.12) brings

$$-\sqrt{H} \geq \frac{2b_{\mathrm{cr}}N\tau - (2b_{\mathrm{cr}}+c_{\mathrm{cr}})\tau + c_{\mathrm{br}}\omega + c_{\mathrm{cr}}}{2c_{\mathrm{cr}}N\tau},$$

which can be rearranged as

$$\frac{c_{\mathrm{br}}\omega + c_{\mathrm{cr}}}{2c_{\mathrm{cr}}\tau N} \leq \frac{2\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}}+1}{2N} - \frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}} - \sqrt{H}. \tag{C.21}$$

In the metastable state, the right-hand side of (C.21) is positive, such that

$$\tau \geq \frac{\frac{c_{\mathrm{br}}}{c_{\mathrm{cr}}}\omega + 1}{2\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}}(1-N)+1-2N\sqrt{H}}.$$

This holds for the metastable state, i.e. for all $\tau \geq \tau_c$. Hence

$$\frac{\frac{c_{\mathrm{br}}}{c_{\mathrm{cr}}}\omega + 1}{2\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}}(1-N)+1-Nh(\omega,\xi)} = \tau_c \geq \tau^* = \frac{\frac{c_{\mathrm{br}}}{c_{\mathrm{cr}}}\omega + 1}{2\frac{b_{\mathrm{cr}}}{c_{\mathrm{cr}}}(1-N)+1-2N\sqrt{H}}. \tag{C.22}$$

We conclude that

$$0 \leq 2\sqrt{H} \leq h(\omega,\xi).$$

Furthermore, since τ_c is bounded for $\frac{c_{br}}{c_{cr}}\omega + 1 > 0$, the denominator of (C.20) should be non-zero. In other words,

$$0 \leq 2\sqrt{H} \leq h(\omega,\xi) < \frac{b_{cr}}{c_{cr}}\frac{1-N}{N} + \frac{1}{2N}, \qquad \text{for } \frac{c_{br}}{c_{cr}}\omega + 1 > 0. \tag{C.23}$$

In (C.23), the function $h(\omega,\xi)$ is bounded for some ω values, but not all. The remaining ω values are taken care of by Step 2.

Step 2. Bounding $h(\omega,\xi)$ for the other ω values.

Step 2A. Case $\frac{c_{br}}{c_{cr}} < 0$. (ASIS,AID)
Out of the 30 instances considered in this lemma, 15 are part of this case.
For the limit of $\omega \uparrow -\frac{c_{cr}}{c_{br}}$, the epidemic threshold τ_c given in (C.20) should still be non-negative, or at least not suddenly become zero. This can only be assured when the denominator in (C.20) becomes zero as well. This continuity argument shows that equality holds for (C.23), which is

$$\lim_{\omega\uparrow -\frac{c_{cr}}{c_{br}}} h(\omega,\xi) = \frac{b_{cr}}{c_{cr}}\frac{1-N}{N} + \frac{1}{2N}.$$

For $\omega > -\frac{c_{cr}}{c_{br}}$, the epidemic threshold τ_c in (C.20) should be positive as well, so using (C.22) one finds

$$h(\omega,\xi) > \frac{b_{cr}}{c_{cr}}\frac{1-N}{N} + \frac{1}{2N} \qquad \text{for } \omega > -\frac{c_{cr}}{c_{br}}.$$

We now consider the situation where the effective link-breaking rate ω increases up to infinity. Suppose a node i is infected. The link between node i and its neighbours j is removed (as ω is high) when the link-breaking rule allows for that. The link can be recreated only when (I) the link-creation rule f_{cr} creates the link between node i and j when either i or j is infected (these updating rules are $f_{cr} = 1 - X_iX_j$, $f_{cr} = 1 - (1 - X_i)(1 - X_j)$ and $f_{cr} = (X_i - X_j)^2$) and (II) the link-breaking rule f_{br} does not break the link between susceptible and infected nodes (these updating rules are $f_{br} = X_iX_j$, $f_{br} = (1 - X_i)(1 - X_j)$ and $f_{br} = 1 - (X_i - X_j)^2$). Only when (I) and (II) are satisfied, spreading in the network continues despite the link-breaking rate ω increasing up to infinity. This allows for a split-up into two classes: Class A and B.

(Class A) (AID) The epidemic threshold remains constant.
The only eligible instances for this class have been listed above. Some of these are still invalid, because they do not obey $H \geq 0$ (e.g. the link-creation rule $f_{cr} = 1 - X_iX_j$) or do not obey $\frac{c_{br}}{c_{cr}} \geq 0$ (which is Step 2B). These constraints yield six instances having any combination of the following link-breaking rules: $f_{br} = X_iX_j$, $f_{br} = (1 - X_i)(1 - X_j)$ and $f_{br} = 1 - (X_i - X_j)^2$ and for the link-creation rules: $f_{cr} = 1 - (1 - X_i)(1 - X_j)$ and $f_{cr} = (X_i - X_j)^2$. These instances have in common that, whilst increasing ω, the epidemic threshold τ_c barely increases. In other words, the limit of $\omega \to \infty$ of τ_c is finite. So define

$$\lim_{\omega\to\infty} \tau_c(\omega,\xi) = C_1 > 0.$$

We continue to prove that $h(\omega,\xi)$ is linear in ω for large ω. The epidemic threshold can be rewritten in terms of h;

$$h(\omega,\xi) = 2\frac{b_{\text{cr}}}{c_{\text{cr}}}\frac{1-N}{N} + \frac{1}{N} - \frac{\frac{c_{\text{br}}}{c_{\text{cr}}}\omega + 1}{N\tau_c(\omega,\xi)}. \tag{C.24}$$

Then we may compute the following

$$\frac{1}{NC_1} = \frac{1}{N\lim_{\omega\to\infty}\tau_c(\omega,\xi)} = \lim_{\omega\to\infty}\frac{1}{N\tau_c(\omega,\xi)} \overset{def.}{=} \lim_{\omega\to\infty}\frac{-h(\omega,\xi) + \frac{1}{N} + 2\frac{b_{\text{cr}}}{c_{\text{cr}}}\frac{1-N}{N}}{\frac{c_{\text{br}}}{c_{\text{cr}}}\omega + 1}$$
$$\overset{\text{l'hôpital}}{=} -\frac{c_{\text{cr}}}{c_{\text{br}}}\frac{\partial h}{\partial\omega}\bigg|_{\omega\to\infty} = C_2.$$

Since $C_1 > 0$, we conclude $C_2 > 0$. Hence, $h(\omega,\xi)$ is a linear function in ω for all instances in Class A.

(Class B) (ASIS) The epidemic threshold scales linearly in ω.
The remaining $15-6=9$ instances not belonging to Class A are part of this class. For each instance in this class, the link-breaking rule is dominant in the sense that spreading between susceptible and infected nodes cannot take place (for $\omega\to\infty$ and fixed τ,ξ) because the link between susceptible and infected nodes is removed immediately. Hence the epidemic threshold τ_c must increase along ω to keep spreading the disease (in the limit of $\omega\to\infty$). This proves that the epidemic threshold scales linearly in ω.

Step 2B. Case $\frac{c_{\text{br}}}{c_{\text{cr}}} \geq 0$.
(Class C) (ABN) The remaining $30-15=15$ instances of G-ASIS follow this constraint. Table 4.1 shows that the ratio $\frac{c_{\text{br}}}{c_{\text{cr}}}$ is strictly positive. Therefore the relation in (C.23) holds for all $\omega \geq 0$. So $h(\omega,\xi)$ is strictly bounded by a constant, also in the limit of $\omega\to\infty$. Then the epidemic threshold τ_c scales linearly in ω, which proves the lemma. □

Lemma C.3 *Let τ_c be the epidemic threshold from (C.20) and assume $H<0$. Then $h(\omega,\xi)$ is bounded by a linear function in ω or by a constant.*

Proof. The only instances of G-ASIS satisfying $H<0$ are instances which have link-creation rule $f_{\text{cr}} = 1 - X_iX_j$. Therefore we substitute $a_{\text{cr}}=1, b_{\text{cr}}=0, c_{\text{cr}}=-1$ in (C.13) to find

$$H = \frac{1}{N} - 1 + \frac{a_{\text{br}}\omega+1}{N^2}\mathrm{E}\left[\sum_{i=1}^{N} d_i^*\right] + \mathrm{Var}[Z^*] + \frac{(2b_{\text{br}}+c_{\text{br}})\omega-1}{N^2}\mathrm{E}\left[\sum_{i=1}^{N} d_i^* X_i^*\right].$$

We have derived that $y = V \pm \sqrt{V^2-H}$. Since $H<0$, the prevalence y has two solutions: $y_1>0$ and $y_2<0$. Our focus lies on the physical solution y_1.

Step 1. The prevalence y_1 is bounded by 1.
This provides the constraint

$$V + \sqrt{V^2-H} \leq 1. \tag{C.25}$$

The sign of V is not determined. The equation above can be rewritten, so

$$V^2 - H \leq 1 - 2V + V^2.$$

Removing V^2 and inserting the definition for V from (C.12), we find

$$-H \leq 1 - 2\left(\frac{1}{2N} + \frac{c_{\text{br}}\omega - 1}{2N\tau}\right),$$

where the values of $a_{\text{cr}}, b_{\text{cr}}$ and c_{cr} have been substituted already. The last equation can be rewritten as

$$\frac{1 - c_{\text{br}}\omega}{1 - Nh(\omega,\xi)} = \tau_c \geq \tau^* = \frac{1 - c_{\text{br}}\omega}{1 - N - NH}. \tag{C.26}$$

The denominator is positive, so this constraint is confining for $c_{\text{br}} < 0$ and for ($c_{\text{br}} > 0$ and $\omega < \frac{1}{c_{\text{br}}}$). In either case, we conclude that

$$0 \leq 1 + H \leq h(\omega,\xi) < \frac{1}{N}. \tag{C.27}$$

implying that $h(\omega,\xi)$ is always positive.

Step 2. Bounding $h(\omega,\xi)$ for all other ω-values. This step is analogous to step 2 from Lemma C.2.

Step 2A. Case $c_{\text{br}} > 0$. Since τ_c is positive for $\omega < \frac{1}{c_{\text{br}}}$, by taking $\lim_{\omega \uparrow \frac{1}{c_{\text{br}}}}$, the limit must be finite and non-zero. This implies

$$\lim_{\omega \uparrow \frac{1}{c_{\text{br}}}} h(\omega,\xi) = \frac{1}{N}.$$

For the relation in (C.26) to be meaningful (τ_c should be non-negative) for $\omega > \frac{1}{c_{\text{br}}}$, it is required that

$$h(\omega,\xi) > \frac{1}{N}, \qquad \text{for } \omega > \frac{1}{c_{\text{br}}}.$$

It remains to analyse the behaviour of $h(\omega,\xi)$ when the effective link-breaking rate ω approaches infinity. For increasing ω and fixed τ, the link-breaking process occurs almost immediately. Nodes become isolated and cure without having any links. There is, however, another possibility. The link-creation mechanism is here $f_{\text{cr}} = 1 - X_i X_j$, which implies a link is created between a susceptible and an infected node. In case the link-breaking process does not include the updating rule where the link is broken between a susceptible and an infected node, the spreading continues despite $\omega \to \infty$. There are two possibilities.

(Class D) The epidemic threshold remains constant.
The link-creation process is $f_{\text{cr}} = 1 - X_i X_j$ and as link-breaking rule, we require either $f_{\text{br}} = X_i X_j, f_{\text{br}} = (1 - X_i)(1 - X_j)$ or $f_{\text{br}} = 1 - (X_i - X_j)^2$. These updating rules have in common that, whilst increasing ω, the epidemic threshold barely increases. In other words, the following limit holds;

$$\lim_{\omega \to \infty} \tau_c(\omega,\xi) = C_1 > 0.$$

Then $h(\omega,\xi)$ scales linearly with ω in a way analogous to Lemma C.2, Step 2A, Class A.

(Class E) The epidemic threshold is constant.
No updating rules belong to this class, as the only eligible do not comply with $c_{\text{br}} > 0$.

Step 2B. Case $c_{\text{br}} < 0$.
(Class F) In this case the relationship (C.27) holds for all ω. So $h(\omega,\xi)$ is bounded by a constant for all ω. This implies τ_c scales linearly in ω, which proves the lemma. □

Combining the result of Lemma C.2 and C.3 proves Theorem 4.3.

D

APPENDIX TO CHAPTER 5

D.1. PROOF OF THEOREM 5.1

Proof. The steady states of any dynamical system can be computed by setting the time derivatives to zero. This reduces (5.6a) and (5.6b) to

$$y_\infty = x(1-y_\infty)y_\infty z_\infty, \tag{D.1a}$$

$$\zeta z_\infty \left(a_{\mathrm{br}} + 2b_{\mathrm{br}} y_\infty + c_{\mathrm{br}} y_\infty^2\right) = \xi(1-z_\infty)\left(a_{\mathrm{cr}} + 2b_{\mathrm{cr}} y_\infty + c_{\mathrm{cr}} y_\infty^2\right). \tag{D.1b}$$

Equation (D.1a) shows that $y_\infty = 0$ is a solution. Inserting $y_\infty = 0$ into Eq. (D.1b), we obtain

$$\zeta z_\infty a_{\mathrm{br}} = \xi(1-z_\infty)a_{\mathrm{cr}}.$$

If $a_{\mathrm{cr}} = a_{\mathrm{br}} = 0$, any value for z_∞ is a steady-state solution. Otherwise, we find

$$z_\infty = \frac{a_{\mathrm{cr}}}{a_{\mathrm{br}}\omega + a_{\mathrm{cr}}}.$$

Now suppose that $y_\infty > 0$, so y_∞ can be removed from Eq. (D.1a). Using the effective link-breaking rate $\omega = \zeta/\xi$, (D.1a) and (D.1b) become

$$1 = x(1-y_\infty)z_\infty, \tag{D.2a}$$

$$\omega z_\infty(a_{\mathrm{br}} + 2b_{\mathrm{br}} y_\infty + c_{\mathrm{br}} y_\infty^2) = (1-z_\infty)(a_{\mathrm{cr}} + 2b_{\mathrm{cr}} y_\infty + c_{\mathrm{cr}} y_\infty^2). \tag{D.2b}$$

Equation (D.2a) shows that $y_\infty \neq 1$. Before making any further claims about y_∞, we present the following Lemma.

Lemma D.1 *Consider an updating rule f with corresponding parameters a, b and c from Table D.1. Then the following function is strictly positive on the interval $0 < y_\infty < 1$:*

$$g(y_\infty) = a + 2by_\infty + cy_\infty^2.$$

Table D.1: All G-ASIS updating rules.

rule f	a	b	c	quadratic form	zeros	gate
X_iX_j	0	0	1	y_∞^2	$y_\infty = 0$ (2x)	AND
$1-X_iX_j$	1	0	-1	$(1-y_\infty)(1+y_\infty)$	$y_\infty = 1, y_\infty = -1$	NAND
$(1-X_i)(1-X_j)$	1	-1	1	$(1-y_\infty)^2$	$y_\infty = 1$ (2x)	NOR
$1-(1-X_i)(1-X_j)$	0	1	-1	$y_\infty(2-y_\infty)$	$y_\infty = 0, y_\infty = 2$	OR
$(X_i-X_j)^2$	0	1	-2	$2y_\infty(1-y_\infty)$	$y_\infty = 0, y_\infty = 1$	XOR
$1-(X_i-X_j)^2$	1	-1	2	$y_\infty^2+(1-y_\infty)^2$	$y_\infty \in \mathbb{C}$	XNOR

Proof. There are six updating rules for the link-creation mechanism f_{cr} and six for the link-breaking mechanism f_{br}, which are listed in Table D.1. Each of these updating rules can be written in a quadratic form. Since $0 < y_\infty < 1$, all quadratic forms are strictly positive. □

Based on Lemma D.1, we may rewrite (D.2b) in terms of z_∞;

$$z_\infty = \frac{a_{\text{cr}} + 2b_{\text{cr}}y_\infty + c_{\text{cr}}y_\infty^2}{a_{\text{cr}} + 2b_{\text{cr}}y_\infty + c_{\text{cr}}y_\infty^2 + \omega(a_{\text{br}} + 2b_{\text{br}}y_\infty + c_{\text{br}}y_\infty^2)}. \tag{D.3}$$

Substituting (D.3) into (D.2a) gives

$$1 = x(1-y_\infty)\frac{a_{\text{cr}} + 2b_{\text{cr}}y_\infty + c_{\text{cr}}y_\infty^2}{a_{\text{cr}} + 2b_{\text{cr}}y_\infty + c_{\text{cr}}y_\infty^2 + \omega(a_{\text{br}} + 2b_{\text{br}}y_\infty + c_{\text{br}}y_\infty^2)}. \tag{D.4}$$

Rewriting (D.4) gives equation (5.7). □

D.2. PROOF OF THEOREM 5.2

We present two proofs. The first proof is specifically tailored towards the G-ASIS model and results in Corollary 5.3, whereas the second proof is more general and encompasses a larger class of spreading processes.

Proof 1. We show that Eq. (5.7) has at least one solution. We split up the proof in two parts.

(i) Consider Eq. (D.3) in Appendix D.1. Equation (5.7) can be simplified if the zeros of the link-breaking and link-creation mechanisms coincide. The coinciding zeros are either unphysical (e.g. $y_\infty = -1$ or complex y_∞) or it is redundant (the solution $y_\infty = 0$ was already provided in Theorem 5.1) or the solution is invalid (e.g. $y_\infty = 1$, which follows by inserting the solution $y_\infty = 1$ into equation (5.6a)). The resulting equation is quadratic in y_∞ and can be readily solved. Further working out the details, one can prove that at least one of the solutions is valid in a certain (x,ω)-region.

(ii) For the remaining instances without coinciding zeros, we define the function $G(y_\infty)$ as the expression on the left-hand side in Eq. (5.7). Filling in the two limit cases $y_\infty = 0$ and $y_\infty = 1$, we find

$$G(0) = a_{\text{cr}} + a_{\text{br}}\omega - a_{\text{cr}}x,$$
$$G(1) = \omega\,(a_{\text{br}} + 2b_{\text{br}} + c_{\text{br}}) + (a_{\text{cr}} + 2b_{\text{cr}} + c_{\text{cr}})\,.$$

The expression $G(1)$ is positive, provided that at least one of the two terms is non-zero. The link-creation and link-breaking mechanisms that violate $G(1) > 0$ are actually contained in case (i), thus we may safely assume that $G(1) > 0$. Furthermore, $G(0) < 0$ if

$$a_{\text{cr}} + a_{\text{br}}\omega < a_{\text{cr}}x.$$

The case $a_{\text{cr}} = 0$ is contained in case (i), thus we assume $a_{\text{cr}} \neq 0$. Thus we can write

$$x > \frac{a_{\text{cr}} + a_{\text{br}}\omega}{a_{\text{cr}}}. \tag{D.5}$$

Equation (D.5) describes the condition under which $G(0) < 0$. Given that $G(1) > 0$, the Intermediate Value Theorem states that there must be some $0 < y_\infty < 1$ for which $G(y_\infty) = 0$, which proves the theorem. □

Proof 2. We prove Theorem 5.2 by showing that the reverse cannot hold, i.e. we look for functions f_{br} and f_{cr} for which no solution exists for all (x, ω)-values. We define

$$h(y_\infty) = \omega f_{\text{br}}(y_\infty) + f_{\text{cr}}(y_\infty) - x(1 - y_\infty) f_{\text{cr}}(y_\infty),$$

where $f_{\text{br}}(y_\infty)$ is the link-breaking rule and $f_{\text{cr}}(y_\infty)$ is the link-creation rule. If $h(y_\infty) = 0$, the equation simplifies to Eq. (5.7). The function h is differentiable, because h is the composite of such functions f_{br} and f_{cr}. According to the Intermediate Value Theorem, if there exists y_∞ and y'_∞ such that $h(y_\infty) > 0$ and $h(y'_\infty) < 0$, then there must exists some y''_∞ for which holds that $h(y''_\infty) = 0$. To guarantee that solutions do not exist, we must prove that either $h(y_\infty) > 0$ or $h(y_\infty) < 0$ for all y_∞. We focus on the first case, the other case goes analogously. The function f_{cr} is non-negative and non-trivial (see Table D.1), thus there must exist some y'_∞ such that $f_{\text{cr}}(y'_\infty) > 0$. To ensure that $h(y'_\infty) > 0$, we find the condition

$$\omega f_{\text{br}}(y'_\infty) + f_{\text{cr}}(y'_\infty) > x(1 - y'_\infty) f_{\text{cr}}(y'_\infty).$$

Suppose $f_{\text{br}}(y'_\infty) = 0$, then the equation simplifies to

$$1 > x(1 - y'_\infty),$$

which is not satisfied unconditionally, that is, for all values of x, except if we would allow $y'_\infty = 1$ as a solution (which is, fortunately, excluded as a steady-state solution, see (5.6)). If $f_{\text{br}}(y'_\infty) > 0$, the condition can also not be satisfied unconditionally. We conclude that there is always a non-empty (x, ω)-region where at least one solution exists. □

D.3. DERIVATION SECOND-ORDER MEAN-FIELD APPROXIMATION

Prior to the derivation of the second-order mean-field equation, we define the following variables:

$$
\begin{aligned}
y &= \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}[X_i], \\
z_{\mathrm{SS}} &= \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathrm{E}[(1-X_i)(1-X_j)a_{ij}], \\
z_{\mathrm{SI}} &= \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathrm{E}[(X_i(1-X_j)+(1-X_i)X_j)a_{ij}], \\
z_{\mathrm{II}} &= \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathrm{E}[X_i X_j a_{ij}], \\
z_{\mathrm{ISI}} &= \frac{2}{N(N-1)(N-2)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{\substack{k=1\\ k\neq i\\ k\neq j}}^{N} \mathrm{E}[(1-X_i)X_j X_k a_{ij}a_{ik}], \\
z_{\mathrm{SSI}} &= \frac{2}{N(N-1)(N-2)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{\substack{k=1\\ k\neq i\\ k\neq j}}^{N} \mathrm{E}[((1-X_i)(1-X_j)X_k + X_i(1-X_j)(1-X_k))a_{ij}a_{ik}],
\end{aligned}
\tag{D.6}
$$

where y denotes the fraction of infected nodes, $z_{\mathrm{SS}}, z_{\mathrm{SI}}$ and z_{II} denote the fraction of links in the graph between susceptible-susceptible (S-S), susceptible-infected (S-I) and infected-infected (I-I) pairs of nodes, respectively. Finally, z_{SSI} and z_{ISI} denote the fraction of connected S-S-I and I-S-I triples in the graph, respectively. Any other triples are irrelevant, because S-S pairs and S-I pairs can be infected by an external infected node I. The external, infected node must be connected to one of the susceptible nodes in the original node pair, leading to the triplet S-S-I or I-S-I.

Using the definitions (D.6), the average fraction of links z in the graph is given by

$$
z = z_{\mathrm{SS}} + z_{\mathrm{SI}} + z_{\mathrm{II}}. \tag{D.7}
$$

We derive the governing equation for z_{II}; the remaining equations can be derived analogously. The governing equation for $\mathrm{E}[X_i X_j a_{ij}]$ is given by

$$
\begin{aligned}
\frac{d\,\mathrm{E}[X_i X_j a_{ij}]}{dt} &= -2\delta \mathrm{E}[X_i X_j a_{ij}] + \beta \mathrm{E}[(1-X_i)X_j a_{ij}] + \beta \mathrm{E}[X_i(1-X_j)a_{ij}] \\
&\quad + \beta \sum_{\substack{k=1\\ k\neq i\\ k\neq j}}^{N} \mathrm{E}[(1-X_i)X_j a_{ij} X_k a_{ik}] + \beta \sum_{\substack{k=1\\ k\neq i\\ k\neq j}}^{N} \mathrm{E}[X_i(1-X_j)a_{ij}X_k a_{jk}] \\
&\quad - \zeta_{\mathrm{II}} \mathrm{E}[X_i X_j a_{ij}] + \xi_{\mathrm{II}} \mathrm{E}[X_i X_j (1-a_{ij})].
\end{aligned}
\tag{D.8}
$$

The G-ASIS model is a Markov chain, where each state encodes which nodes are infected or susceptible and which links are existent or non-existent in the graph. The possible transitions to and from state $X_i = 1, X_j = 1, a_{ij} = 1$ are as follows: One of the two infected nodes cures (first term on the right-hand side of Eq. (D.8)), node j infects node i (second term), node i infects node j (third term), there is an infection from outside (term four and five), the link is broken between node i and node j (term six) or the link is created between node i and node j (term seven) are the possible transitions to and from the state $X_i = 1, X_j = 1, a_{ij} = 1$. The variable ζ_{II} is defined as $\zeta_{\text{II}} = \zeta(a_{\text{br}} + 2b_{\text{br}} + c_{\text{br}})$ and indicates the link-breaking rate between two infected nodes. Similarly, $\xi_{\text{II}} = \xi(a_{\text{cr}} + 2b_{\text{cr}} + c_{\text{cr}})$ indicates the link-creation rate between two infected nodes. Analogously, we define:

delete S-S link	$\zeta_{\text{SS}} = \zeta a_{\text{br}}$	create S-S link	$\xi_{\text{SS}} = \xi a_{\text{cr}}$
delete S-I link	$\zeta_{\text{SI}} = \zeta(a_{\text{br}} + b_{\text{br}})$	create S-I link	$\xi_{\text{SI}} = \xi(a_{\text{cr}} + b_{\text{cr}})$
delete I-I link	$\zeta_{\text{II}} = \zeta(a_{\text{br}} + 2b_{\text{br}} + c_{\text{br}})$	create I-I link	$\xi_{\text{II}} = \xi(a_{\text{cr}} + 2b_{\text{cr}} + c_{\text{cr}})$

By summing over i and $j \neq i$ in equation (D.8), multiplying all terms by $\frac{1}{N(N-1)}$ and using the definitions (D.6), we find

$$\frac{d\,z_{\text{II}}}{dt} = -2\delta z_{\text{II}} + \beta z_{\text{SI}} + \beta(N-2)z_{\text{ISI}} - \zeta_{\text{II}} z_{\text{II}} - \xi_{\text{II}} z_{\text{II}} + \xi_{\text{II}} \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \text{E}[X_i X_j].$$

The last term needs to be rewritten;

$$\begin{aligned}
\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \text{E}[X_i X_j] &= \sum_{i=1}^{N} \sum_{j=1}^{N} \text{E}[X_i X_j] - \sum_{i=1}^{N} \text{E}[X_i^2] \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \Big(\text{Cov}[X_i, X_j] + \text{E}[X_i]\text{E}[X_j] \Big) - \sum_{i=1}^{N} \text{E}[X_i] \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \text{Cov}[X_i, X_j] + N^2 y^2 - Ny \\
&\approx N^2 y^2 - Ny,
\end{aligned}$$

where we have made the approximation that the covariance between the state X_i and X_j is zero. Then we finally obtain

$$\frac{d\,z_{\text{II}}}{dt} = -2\delta z_{\text{II}} + \beta z_{\text{SI}} + \beta(N-2)z_{\text{ISI}} - \zeta_{\text{II}} z_{\text{II}} + \xi_{\text{II}} \left(\frac{N}{N-1} y^2 - \frac{1}{N-1} y - z_{\text{II}} \right).$$

Finally, we rescale time $\tilde{t} = \delta t$ and using $\tau = \beta/\delta, \tilde{\zeta} = \zeta/\delta, \tilde{\xi} = \xi/\delta$, we obtain

$$\frac{d\,z_{\text{II}}}{d\tilde{t}} = -2 z_{\text{II}} + \tau z_{\text{SI}} + \tau(N-2)z_{\text{ISI}} - \tilde{\zeta}_{\text{II}} z_{\text{II}} + \tilde{\xi}_{\text{II}} \left(\frac{N}{N-1} y^2 - \frac{1}{N-1} y - z_{\text{II}} \right).$$

The governing equations for y, z_{SS} and z_{SI} are derived in a similar manner. After dropping the tildes for the time t, link-breaking rate ζ and link-creation rate ξ, we obtain Eq. (5.20).

D.4. PROOF OF THEOREM 5.4

The steady states of equations (5.22) are computed by setting the derivatives to zero, such that

$$y_\infty = \tau \frac{N-1}{2} z_{\mathrm{SI},\infty}, \tag{D.9a}$$

$$(N-1)\tau \frac{z_{\mathrm{SS},\infty} z_{\mathrm{SI},\infty}}{1-y_\infty} = z_{\mathrm{SI},\infty} + \xi_{\mathrm{SS}} \left(\frac{N}{N-1}(1-y_\infty)^2 - \frac{1}{N-1}(1-y_\infty) - z_{\mathrm{SS},\infty} \right) - \zeta_{\mathrm{SS}} z_{\mathrm{SS},\infty}, \tag{D.9b}$$

$$2 z_{\mathrm{II},\infty} = \tau z_{\mathrm{SI},\infty} \left(1 + \frac{N-1}{2} \frac{z_{\mathrm{SI},\infty}}{1-y_\infty} \right) + \xi_{\mathrm{II}} \left(\frac{N}{N-1} y_\infty^2 - \frac{1}{N-1} y_\infty - z_{\mathrm{II},\infty} \right) - \zeta_{\mathrm{II}} z_{\mathrm{II},\infty}, \tag{D.9c}$$

$$\begin{aligned}(1+\tau) z_{\mathrm{SI},\infty} = \tau (N-1) \frac{z_{\mathrm{SI},\infty}}{1-y_\infty} \left(z_{\mathrm{SS},\infty} - \frac{1}{2} z_{\mathrm{SI},\infty} \right) + 2 z_{\mathrm{II},\infty}, \\ + \xi_{\mathrm{SI}} \left(\frac{2N}{N-1} y_\infty (1-y_\infty) - z_{\mathrm{SI},\infty} \right) - \zeta_{\mathrm{SI}} z_{\mathrm{SI},\infty}.\end{aligned} \tag{D.9d}$$

By taking all $z_{\mathrm{II},\infty}$-terms in (D.9c), substituting $z_{\mathrm{II},\infty}$ into (D.9d), we obtain

$$y_\infty = \tau \frac{N-1}{2} z_{\mathrm{SI},\infty}, \tag{D.10a}$$

$$(N-1)\tau \frac{z_{\mathrm{SS},\infty} z_{\mathrm{SI},\infty}}{1-y_\infty} = z_{\mathrm{SI},\infty} + \xi_{\mathrm{SS}} \left(\frac{N}{N-1}(1-y_\infty)^2 - \frac{1}{N-1}(1-y_\infty) - z_{\mathrm{SS},\infty} \right) - \zeta_{\mathrm{SS}} z_{\mathrm{SS},\infty}, \tag{D.10b}$$

$$\begin{aligned}(1+\tau) z_{\mathrm{SI},\infty} = \tau (N-1) \frac{z_{\mathrm{SI},\infty}}{1-y_\infty} \left(z_{\mathrm{SS},\infty} - \frac{1}{2} z_{\mathrm{SI},\infty} \right) \\ + 2 \frac{\tau z_{\mathrm{SI},\infty} \left(1 + \frac{N-1}{2} \frac{z_{\mathrm{SI},\infty}}{1-y_\infty} \right) + \xi_{\mathrm{II}} \left(\frac{N}{N-1} y_\infty^2 - \frac{1}{N-1} y_\infty \right)}{2 + \xi_{\mathrm{II}} + \zeta_{\mathrm{II}}} \\ + \xi_{\mathrm{SI}} \left(\frac{2N}{N-1} y_\infty (1-y_\infty) - z_{\mathrm{SI},\infty} \right) - \zeta_{\mathrm{SI}} z_{\mathrm{SI},\infty}.\end{aligned} \tag{D.10c}$$

For readability, we define the positive constant $\alpha_1 = \frac{2}{2+\xi_{\mathrm{II}}+\zeta_{\mathrm{II}}}$. We substitute (D.10a) into the other equations, such that

$$\begin{aligned}(N-1)\tau \frac{z_{\mathrm{SS},\infty} z_{\mathrm{SI},\infty}}{1-\tau \frac{N-1}{2} z_{\mathrm{SI},\infty}} = \xi_{\mathrm{SS}} \left(\frac{N}{N-1} \left(1 - \tau \frac{N-1}{2} z_{\mathrm{SI},\infty} \right)^2 - \frac{1}{N-1} \left(1 - \tau \frac{N-1}{2} z_{\mathrm{SI},\infty} \right) - z_{\mathrm{SS},\infty} \right) \\ + z_{\mathrm{SI},\infty} - \zeta_{\mathrm{SS}} z_{\mathrm{SS},\infty},\end{aligned} \tag{D.11a}$$

$$\begin{aligned}(1+\tau) z_{\mathrm{SI},\infty} = \tau (N-1) \frac{z_{\mathrm{SI},\infty}}{1-\tau \frac{N-1}{2} z_{\mathrm{SI},\infty}} \left(z_{\mathrm{SS},\infty} - \frac{1}{2} z_{\mathrm{SI},\infty} \right) + \alpha_1 \tau z_{\mathrm{SI},\infty} \left(1 + \frac{N-1}{2} \frac{z_{\mathrm{SI},\infty}}{1-\tau \frac{N-1}{2} z_{\mathrm{SI},\infty}} \right) \\ + \alpha_1 \xi_{\mathrm{II}} \left(\tau^2 \frac{N(N-1)}{4} z_{\mathrm{SI},\infty}^2 - \frac{\tau}{2} z_{\mathrm{SI},\infty} \right) + \xi_{\mathrm{SI}} \left(\tau N z_{\mathrm{SI},\infty} \left(1 - \tau \frac{N-1}{2} z_{\mathrm{SI},\infty} \right) - z_{\mathrm{SI},\infty} \right) - \zeta_{\mathrm{SI}} z_{\mathrm{SI},\infty}.\end{aligned} \tag{D.11b}$$

One solution of equation (D.11b) is the all-healthy state $y_\infty = z_{\mathrm{SI},\infty} = z_{\mathrm{II},\infty} = 0$. By inserting the all-healthy state into the original equations (D.9), we obtain the steady-state

fraction of S-S links $z_{SS,\infty}$ given in the theorem. To remove the all-healthy solution, we divide Eq. (D.11b) by $z_{SI,\infty}$, such that

$$(N-1)\tau\frac{z_{SS,\infty}z_{SI,\infty}}{1-\tau\frac{N-1}{2}z_{SI,\infty}} = \xi_{SS}\left(\frac{N}{N-1}\left(1-\tau\frac{N-1}{2}z_{SI,\infty}\right)^2-\frac{1}{N-1}\left(1-\tau\frac{N-1}{2}z_{SI,\infty}\right)-z_{SS,\infty}\right) + z_{SI,\infty}-\zeta_{SS}z_{SS,\infty}, \tag{D.12a}$$

$$(1+\tau) = \tau(N-1)\frac{1}{1-\tau\frac{N-1}{2}z_{SI,\infty}}\left(z_{SS,\infty}-\frac{1}{2}z_{SI,\infty}\right)+\alpha_1\tau\left(1+\frac{N-1}{2}\frac{z_{SI,\infty}}{1-\tau\frac{N-1}{2}z_{SI,\infty}}\right) +\alpha_1\xi_{II}\left(\tau^2\frac{N(N-1)}{4}z_{SI,\infty}-\frac{\tau}{2}\right)+\xi_{SI}\left(\tau N\left(1-\tau\frac{N-1}{2}z_{SI,\infty}\right)-1\right)-\zeta_{SI}. \tag{D.12b}$$

Rewriting (D.12a) in terms of $z_{SS,\infty}$ and rearranging (D.12b) while introducing $\alpha_2 = (1+\zeta_{SI}+\xi_{SI})+(1-\alpha_1+\frac{\alpha_1}{2}\xi_{II}-N\xi_{SI})\tau, \alpha_3 = \xi_{SS}+\zeta_{SS}$, we find

$$z_{SS,\infty} = \frac{z_{SI,\infty}+\xi_{SS}\left(\frac{N}{N-1}(1-\tau\frac{N-1}{2}z_{SI,\infty})^2-\frac{1}{N-1}(1-\tau\frac{N-1}{2}z_{SI,\infty})\right)}{(N-1)\tau z_{SI,\infty}+\alpha_3(1-\tau\frac{N-1}{2}z_{SI,\infty})}\left(1-\tau\frac{N-1}{2}z_{SI,\infty}\right), \tag{D.13a}$$

$$\alpha_2 = \tau(N-1)\frac{z_{SS,\infty}}{1-\tau\frac{N-1}{2}z_{SI,\infty}}+(\alpha_1-1)\tau\frac{N-1}{2}\frac{z_{SI,\infty}}{1-\tau\frac{N-1}{2}z_{SI,\infty}} + \frac{\tau^2N(N-1)}{4}(\alpha_1\xi_{II}-2\xi_{SI})z_{SI,\infty}. \tag{D.13b}$$

Substituting (D.13a) into (D.13b), we obtain

$$\alpha_2 = \tau(N-1)\frac{\xi_{SS}+\left(\frac{\tau}{2}\xi_{SS}-N\tau\xi_{SS}+1\right)z_{SI,\infty}+\xi_{SS}\frac{(N-1)^2}{4}\tau^2z_{SI,\infty}^2}{\alpha_3+(1-\frac{1}{2}\alpha_3)\tau(N-1)z_{SI,\infty}} + (\alpha_1-1)\tau\frac{N-1}{2}\frac{z_{SI,\infty}}{1-\tau\frac{N-1}{2}z_{SI,\infty}}+\frac{\tau^2N(N-1)}{4}(\alpha_1\xi_{II}-2\xi_{SI})z_{SI,\infty}. \tag{D.14}$$

Defining $\alpha_4 = (N-1)\tau\xi_{SS}, \alpha_5 = \tau(N-1)\left(\frac{\tau}{2}-N\tau\right)\xi_{SS}+(N-1)\tau, \alpha_6 = \xi_{SS}\frac{(N-1)^3}{4}\tau^3, \alpha_7 = \frac{\tau^2N(N-1)}{4}(\alpha_1\xi_{II}-2\xi_{SI})$ and $\alpha_8 = (1-\frac{1}{2}\alpha_3)\tau(N-1)$, we find

$$\alpha_2 = \frac{\alpha_4+\alpha_5z_{SI,\infty}+\alpha_6z_{SI,\infty}^2}{\alpha_3+\alpha_8z_{SI,\infty}}+(\alpha_1-1)\tau\frac{N-1}{2}\frac{z_{SI,\infty}}{1-\tau\frac{N-1}{2}z_{SI,\infty}}+\alpha_7z_{SI,\infty}. \tag{D.15}$$

Multiplying (D.15) with $1-\tau\frac{N-1}{2}z_{SI,\infty}$ gives

$$\alpha_2 = \frac{\alpha_4+(\alpha_5-\tau\frac{N-1}{2}\alpha_4)z_{SI,\infty}+(\alpha_6-\tau\frac{N-1}{2}\alpha_5)z_{SI,\infty}^2-\tau\frac{N-1}{2}\alpha_6z_{SI,\infty}^3}{\alpha_3+\alpha_8z_{SI,\infty}} + \left(\tau\frac{N-1}{2}(\alpha_1-1+\alpha_2)+\alpha_7\right)z_{SI,\infty}-\alpha_7\tau\frac{N-1}{2}z_{SI,\infty}^2. \tag{D.16}$$

Multiplying (D.16) with $\alpha_3 + \alpha_8 z_{\mathrm{SI},\infty}$ and rearranging, gives us the cubic equation

$$\begin{aligned} 0 = &\left(\tau\frac{N-1}{2}\alpha_6 + \alpha_7\alpha_8\tau\frac{N-1}{2}\right) z_{\mathrm{SI},\infty}^3 \\ &+ \left(\tau\frac{N-1}{2}\alpha_5 - \alpha_6 - \alpha_8\left(\tau\frac{N-1}{2}(\alpha_1 - 1 + \alpha_2) + \alpha_7\right) + \alpha_3\alpha_7\tau\frac{N-1}{2}\right) z_{\mathrm{SI},\infty}^2 \\ &+ \left(\alpha_2\alpha_8 + \tau\frac{N-1}{2}\alpha_4 - \alpha_5 - \alpha_3\left(\tau\frac{N-1}{2}(\alpha_1 - 1 + \alpha_2) + \alpha_7\right)\right) z_{\mathrm{SI},\infty} \\ &+ (\alpha_2\alpha_3 - \alpha_4)\,. \end{aligned} \tag{D.17}$$

Using the identity $y_\infty = \tau\frac{N-1}{2} z_{\mathrm{SI},\infty}$, we obtain the required cubic equation for y_∞. □

D

D.5. TWO METASTABLE STATES IN THE MARKOVIAN AID MODEL

Trajanovski *et al.* [97] proved that the metastable state of the Markovian AID model does not exist if $\mathrm{Var}[Z^*] > 1/4$, where Z^* is the metastable fraction of infected nodes (see Eq. (4) in [97]). Even though $\mathrm{Var}[Z^*] > 1/4$ is a sufficient condition for the non-existence of the metastable state, Lemma D.2 demonstrates that $\mathrm{Var}[Z^*] > 1/4$ is never satisfied.

Lemma D.2 *For the static and adaptive Markovian SIS model, it holds that* $\mathrm{Var}[Z^*] < \frac{1}{4}$, *where* Z^* *is the metastable fraction of infected nodes.*

Proof.

$$\begin{aligned} \mathrm{Var}[Z^*] = \mathrm{E}[(Z^*)^2] - \mathrm{E}[Z^*]^2 &= \frac{1}{N^2}\mathrm{E}\left[\left(\sum_{i=1}^N X_i^*\right)^2\right] - y^2 = \frac{1}{N^2}\mathrm{E}\left[\sum_{i=1}^N (X_i^*)^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N X_i^* X_j^*\right] - y^2 \\ &= \frac{1}{N^2}\mathrm{E}\left[\sum_{i=1}^N X_i^*\right] + \frac{1}{N^2}\mathrm{E}\left[\sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N X_i^* X_j^*\right] - y^2 = \frac{y}{N} - y^2 + \frac{1}{N^2}\sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N \mathrm{E}\left[X_i^* X_j^*\right] \\ &\overset{\Delta}{\leq} \frac{y}{N} - y^2 + \frac{1}{N^2}\sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N \mathrm{E}\left[X_i^*\right] = \frac{y}{N} - y^2 + \frac{N-1}{N^2}\sum_{i=1}^N \mathrm{E}\left[X_i^*\right] \\ &= y(1-y), \end{aligned}$$

where the inequality Δ follows from

$$\mathrm{E}[X_i^* X_j^*] = \Pr[X_i^* = 1 \cap X_j^* = 1] = \Pr[X_j^* = 1 | X_i^* = 1]\Pr[X_i^* = 1] \leq \Pr[X_i^* = 1] = \mathrm{E}[X_i^*].$$

Since the prevalence y is bounded between zero and one, $\mathrm{Var}[Z^*]$ is bounded between 0 and $\frac{1}{4}$, thus proving our claim. □

The hypothesis in [97] of the non-existence of the metastable state and the non-convergence of the time-varying prevalence $y(t)$ was implicitly based on the assumption of a fast convergence towards the metastable state. Thus, we believe that Figure 2(b) in [97] is merely a result of a short simulation period, in which the process has not yet converged to the steady state.

D.6. TABLE OF ALL G-ASIS INSTANCES

Table D.2: All 36 instances from the G-ASIS model and the relation between the epidemic threshold τ_c and the effective link-breaking rate ω in the Markovian, stochastic model. The lower bound is obtained from Theorem 4.1 and the upper bound follows from Theorem 4.3. The columns MF 1 and MF 2 represents the relation under the first-order and second-order mean-field approximation, respectively. The first-order mean-field approximation predicts 6 correct relations and 9 wrong relations, whereas the second-order mean-field approximation predicts 14 correct relations and 1 incorrect relation. Unfortunately, the remaining 21 relations could not be determined. If a relation is specified by "other", further details on the mean-field approximation results are given in Appendix D.7.

Rules (link-breaking) (link-creation)	Model name	Lower bound epidemic threshold τ_c	Upper bound epidemic threshold τ_c	MF 1	MF 2
X_iX_j X_iX_j	ABN model	$\frac{1}{\lambda_1}$	Linear	Linear	Other
X_iX_j $1-X_iX_j$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant	Constant	Constant
X_iX_j $(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Linear	Constant	Constant
X_iX_j $1-(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Constant	Constant	Constant
X_iX_j $(X_i-X_j)^2$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant	Constant	Constant
$1-X_iX_j$ $1-(X_i-X_j)^2$		$\frac{1}{\lambda_1}$	Linear	Constant	Constant
$1-X_iX_j$ X_iX_j	AFND model	$\frac{1}{\lambda_1}$	Linear	Linear	Linear
$1-X_iX_j$ $1-X_iX_j$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Linear	Linear	Linear
$1-X_iX_j$ $(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Linear	Linear	Other
$1-X_iX_j$ $1-(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Linear	Linear	Linear
$1-X_iX_j$ $(X_i-X_j)^2$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Linear	Linear	Linear
$1-X_iX_j$ $1-(X_i-X_j)^2$		$\frac{1}{\lambda_1}$	Linear	Linear	Linear

(Continues on next page)

Rules (link-breaking) (link-creation)	Model name and appearance in literature	Lower bound epidemic threshold τ_c	Upper bound epidemic threshold τ_c	MF 1	MF 2
$(1-X_i)(1-X_j)$ X_iX_j	SCM model	$\frac{1}{\lambda_1}$	Linear	Other	Constant
$(1-X_i)(1-X_j)$ $1-X_iX_j$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant	Linear	Constant
$(1-X_i)(1-X_j)$ $(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Linear	Linear	Linear
$(1-X_i)(1-X_j)$ $1-(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Constant	Other	Constant
$(1-X_i)(1-X_j)$ $(X_i-X_j)^2$	AID model [97]	$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant	Linear	Constant
$(1-X_i)(1-X_j)$ $1-(X_i-X_j)^2$		$\frac{1}{\lambda_1}$	Linear	Other	Constant
$1-(1-X_i)(1-X_j)$ X_iX_j		$\frac{1}{\lambda_1}\left(1+\frac{\omega}{1+1/\xi}\right)$	Linear	Linear	Other
$1-(1-X_i)(1-X_j)$ $1-X_iX_j$		$\frac{1}{\lambda_1}\left(1+\frac{\omega-1}{1+1/\xi}\right)$	Linear	Constant	Linear
$1-(1-X_i)(1-X_j)$ $(1-X_i)(1-X_j)$	ACSIS model	$\frac{1}{\lambda_1}(1+\omega\xi)$	Linear	Constant	Linear
$1-(1-X_i)(1-X_j)$ $1-(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}\left(1+\frac{\omega}{1+1/\xi}\right)$	Linear	Linear	Linear
$1-(1-X_i)(1-X_j)$ $(X_i-X_j)^2$		$\frac{1}{\lambda_1}\left(1+\frac{\omega-1}{1+1/\xi}\right)$	Linear	Linear	Linear
$1-(1-X_i)(1-X_j)$ $1-(X_i-X_j)^2$		$\frac{1}{\lambda_1}\left(1+\frac{\omega}{1+1/\xi}\right)$	Linear	Constant	Linear

(Continues on next page)

Rules (link-breaking) (link-creation)	Model name and appearance in literature	Lower bound epidemic threshold τ_c	Upper bound epidemic threshold τ_c	MF 1	MF 2
$(X_i - X_j)^2$ $X_i X_j$		$\frac{1}{\lambda_1}$	Linear	Linear	Linear
$(X_i - X_j)^2$ $1 - X_i X_j$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Linear	Constant	Linear
$(X_i - X_j)^2$ $(1-X_i)(1-X_j)$	ASIS model [7, 52, 97]	$\frac{1}{\lambda_1}$	Linear	Constant	Linear
$(X_i - X_j)^2$ $1-(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Linear	Linear	Linear
$(X_i - X_j)^2$ $(X_i - X_j)^2$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Linear	Linear	Linear
$(X_i - X_j)^2$ $1 - (X_i - X_j)^2$		$\frac{1}{\lambda_1}$	Linear	Constant	Linear
$1 - (X_i - X_j)^2$ $X_i X_j$		$\frac{1}{\lambda_1}$	Linear	Linear	Other
$1 - (X_i - X_j)^2$ $1 - X_i X_j$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant	Linear	Constant
$1 - (X_i - X_j)^2$ $(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Linear	Linear	Constant
$1 - (X_i - X_j)^2$ $1-(1-X_i)(1-X_j)$		$\frac{1}{\lambda_1}$	Constant	Linear	Constant
$1 - (X_i - X_j)^2$ $(X_i - X_j)^2$		$\frac{1}{\lambda_1}\left(\frac{1}{1+\xi}\right)$	Constant	Linear	Constant
$1 - (X_i - X_j)^2$ $1 - (X_i - X_j)^2$		$\frac{1}{\lambda_1}$	Linear	Linear	Constant

D.7. MEAN-FIELD RESULTS

Figure D.1 and D.2 show the relation between the epidemic threshold τ_c and the effective link-breaking rate ω under the first-order and second-order mean-field approximation. Figure D.1 and D.2 are based on a preliminary analysis and have been included in this thesis to serve as an initial step towards understanding all 36 instances of the G-ASIS model, but may not be completely accurate in describing the rich, versatile behaviour of every individual instance in the G-ASIS model.

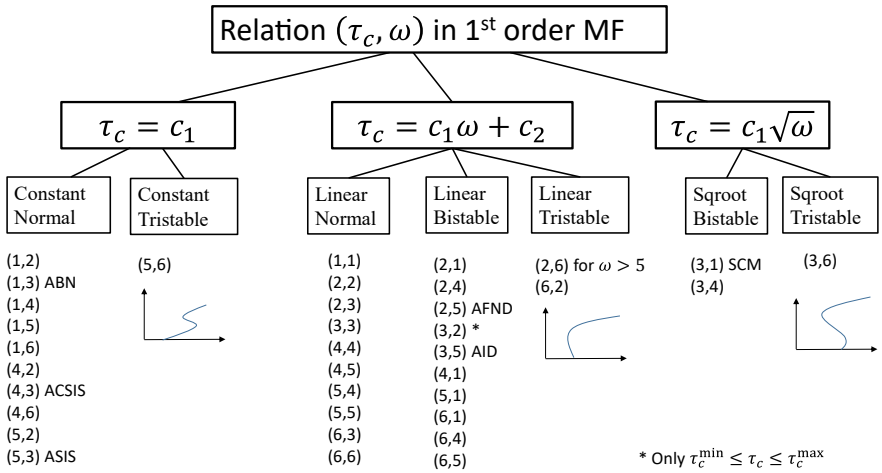


Figure D.1: The relation between the epidemic threshold τ_c and the effective link-breaking rate ω under the first-order mean-field approximation for all 36 instances of the G-ASIS model. We categorise the instances under three main pillars: Constant (left), Linear (middle) and Other (right). Within each pillar, several further distinctions are given. Each G-ASIS instance is specified by the pair (i, j), where i specifies the link-breaking rule and j represents the link-creation rule. The number 1 represents the link-updating rule from the 1st row from Table 4.1, the number 2 the second row, etc.

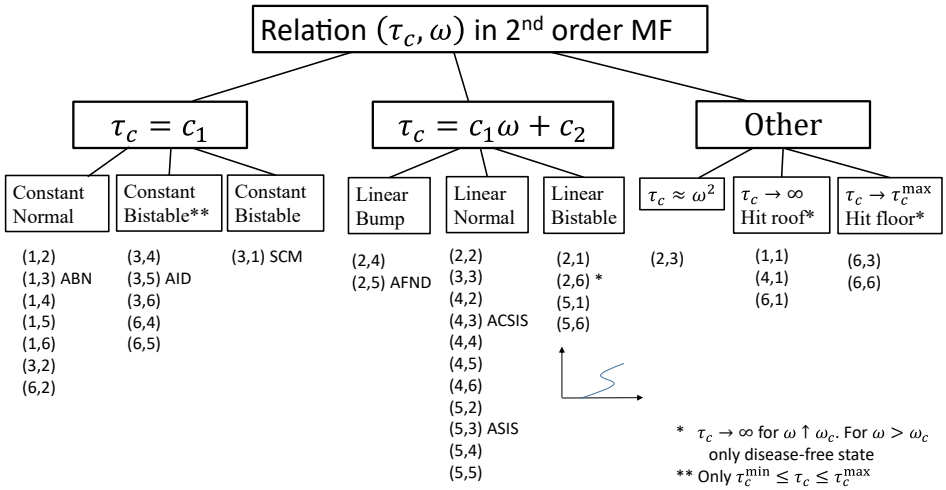


Figure D.2: The relation between the epidemic threshold τ_c and the effective link-breaking rate ω under the second-order mean-field approximation for all 36 instances of the G-ASIS model. We categorise the instances under three main pillars: Constant (left), Linear (middle) and Other (right). Within each pillar, several further distinctions are given. Each G-ASIS instance is specified by the pair (i, j), where i specifies the link-breaking rule and j represents the link-creation rule. The number 1 represents the link-updating rule from the 1st row from Table 4.1, the number 2 the second row, etc.

E

APPENDIX TO CHAPTER 7

E.1. SIR EPIDEMIC MODEL

The SIR epidemic model is defined in Definition 7.1. The COVID-19 pandemic does not exactly follow the SIR epidemic model. Instead, at every time k, the fraction of COVID-19 infections in region i obeys

$$\mathcal{I}_i[k+1] = (1-\delta_i)\,\mathcal{I}_i[k] + \mathcal{S}_i[k] \sum_{j=1}^{N} \beta_{ij}[k]\mathcal{I}_j[k] + w_i[k] \tag{E.1}$$

where $w_i[k]$ denotes the *model error* of region i at time k. Under Assumption E.1, the model errors $w_i[k]$ are identically distributed at all times k and for every region i:

Assumption E.1 *The model error $w_i[k]$ is normally distributed as*

$$w_i[k] \sim \mathcal{N}\left(0, \sigma_w^2\right). \tag{E.2}$$

Furthermore, the model errors $w_i[k]$, $w_j[\tilde{k}]$ are stochastically independent for all times $k \neq \tilde{k}$ and regions $i \neq j$.

Assumption E.2 *For every node i, the curing probabilities satisfy $\delta_i \leq 1$, and, at every time $k \in \mathbb{N}$, the infection probabilities $\beta_{ij}[k]$ satisfy*

$$\sum_{j=1}^{N} \beta_{ij}[k] \leq 1. \tag{E.3}$$

Under Assumption E.2, the fractions $\mathcal{S}_i[k]$, $\mathcal{I}_i[k]$ and $\mathcal{R}_i[k]$ remain in $[0,1]$ at every time k as stated by Lemma E.3, which is inspired by [175, Lemma 1].

Lemma E.3 *Suppose that $\mathcal{I}_i[1] \geq 0$, $\mathcal{R}_i[1] \geq 0$ and $\mathcal{I}_i[1]+\mathcal{R}_i[1] \leq 1$ for every node i. Then, under Assumption E.2, it holds that $\mathcal{I}_i[k] \geq 0$, $\mathcal{R}_i[k] \geq 0$ and $\mathcal{I}_i[k]+\mathcal{R}_i[k] \leq 1$ at every time $k \in \mathbb{N}$ for every node i.*

Proof. We prove Lemma E.3 by induction with respect to time k. The base case at the initial time $k = 1$, i.e. $\mathcal{I}_i[1] \geq 0$, $\mathcal{R}_i[1] \geq 0$ and $\mathcal{S}_i[1] \geq 0$, is satisfied by assumption.

For the inductive step, suppose that at time k it holds that $\mathcal{I}_i[k] \geq 0$, $\mathcal{R}_i[k] \geq 0$ and $\mathcal{S}_i[k] \geq 0$ for all nodes i. Since $\mathcal{S}_i[k] = 1 - \mathcal{I}_i[k] - \mathcal{R}_i[k]$, we obtain

$$0 \leq \mathcal{I}_i[k] + \mathcal{R}_i[k] \leq 1. \tag{E.4}$$

Under Assumption E.2, it holds that $0 \leq \delta_i \leq 1$ and $\beta_{ij}[k] \geq 0$. Thus, we obtain from the SIR governing equations (7.1) and from (E.4) that both $\mathcal{I}_i[k+1]$ and $\mathcal{R}_i[k+1]$ equal a sum of merely positive terms, implying that $\mathcal{I}_i[k+1] \geq 0$ and $\mathcal{R}_i[k+1] \geq 0$ for all nodes i. Then it remains to show that $\mathcal{S}_i[k+1] \geq 0$.

Adding the two SIR equations in (7.1), we find that

$$\mathcal{I}_i[k+1] + \mathcal{R}_i[k+1] = \mathcal{I}_i[k] + \mathcal{R}_i[k] + (1 - \mathcal{I}_i[k] - \mathcal{R}_i[k]) \sum_{j=1}^{N} \beta_{ij}[k]\mathcal{I}_j[k] \tag{E.5}$$

for every node i. From (E.4) and $\mathcal{R}_i[k] \geq 0$, we find that $\mathcal{I}_i[k] \leq 1$, such that

$$\sum_{j=1}^{N} \beta_{ij}[k]\mathcal{I}_j[k] \leq \sum_{j=1}^{N} \beta_{ij}[k] \leq 1 \tag{E.6}$$

where the last inequality follows from Assumption E.2. We conclude that $\mathcal{I}_i[k+1] + \mathcal{R}_i[k+1] \leq 1$, since the right side of (E.5) is a convex combination of 1 and $\mathcal{I}_i[k] + \mathcal{R}_i[k]$ with coefficient $\sum_{j=1}^{N} \beta_{ij}[k]\mathcal{I}_j[k] \in [0,1]$. □

E.2. DATA OF THE COVID-19 OUTBREAK IN HUBEI

Table E.1 shows the cities of the province Hubei and the population size N_i for every city i, which is obtained from the Hubei Statistical Yearbook [176]. The time series of the reported number of infections $N_{\text{rep},i}[k]$ is shown in Table E.2.

E.3. DETAILS OF NIPA

Algorithm 1 describes the NIPA prediction method in pseudocode. In line 3, the Matlab command `smoothdata` removes erratic fluctuations of the raw data $\mathcal{I}_{\text{rep},i}[k]$. We denote the $N \times 1$ infection state vector by $\mathcal{I}[k] = (\mathcal{I}_1[k], ..., \mathcal{I}_N[k])^T$ at any time k. The loop starting in line 7 iterates over all candidate values of the curing probability δ_i in the set Ω. Algorithm 1 calls the `Network_inference` method, whose pseudocode is given in Algorithm 2. For a fixed curing probability δ_i, the network inference in line 11 returns an estimate for the infection probabilities $\beta_{ij}[k](\delta_i)$ for all $j = 1, ..., N$. Furthermore, the network inference returns the mean squared error $\text{MSE}(\delta_i)$, which corresponds to the first term in the objective of (7.6). The smaller the mean squared error $\text{MSE}(\delta_i)$, the better the fit of the SIR model (7.1) to the data $\mathcal{I}_i[1], ..., \mathcal{I}_i[n]$. In line 13, the final estimate $\hat{\delta}_i$ for the curing probability is obtained as the minimiser of the mean squared error $\text{MSE}(\delta_i)$. The estimate $\hat{\delta}_i$ determines the final estimates $\hat{\beta}_{i1}$, ..., $\hat{\beta}_{iN}$ for the infection probabilities in line 14. From line 16 to line 26, the SIR model (7.1) is iterated, which results in the predicted fraction of infections $\hat{\mathcal{I}}_i[n+1], ..., \hat{\mathcal{I}}_i[n + n_{\text{pred}}]$ for all nodes i.

Algorithm 1 `Network-Inference-based Prediction Algorithm (NIPA)`

Input: reported fraction of infections $\mathcal{I}_{\text{rep},i}[1],...,\mathcal{I}_{\text{rep},i}[n]$ for all nodes i; prediction time n_{pred}

Output: predicted fraction of infections $\hat{\mathcal{I}}_i[n+1],...,\hat{\mathcal{I}}_i[n+n_{\text{pred}}]$ for all nodes i

Step 1 - Data preprocessing

$\mathcal{I}_i[1],...,\mathcal{I}_i[n] \leftarrow$ `smoothdata`$(\mathcal{I}_{\text{rep},i}[1],...,\mathcal{I}_{\text{rep},i}[n])$ for all nodes $i = 1,\dots,N$

$\mathcal{I}[k] \leftarrow (\mathcal{I}_1[k],...,\mathcal{I}_N[k])^T$ for all $k = 1,...,n$

Step 2 - Network inference

for $i = 1,...,N$ **do**
  $\mathcal{R}_i[1] \leftarrow 0$
  for $\delta_i \in \Omega$ **do**
    $\mathcal{R}_i[k] \leftarrow \mathcal{R}_i[k-1] + \delta_i \mathcal{I}_i[k-1]$ for all $k = 2,...,n$
    $\mathcal{S}_i[k] \leftarrow 1 - \mathcal{I}_i[k] - \mathcal{R}_i[k]$ for all $k = 1,...,n$
    $v_i[k] \leftarrow (\mathcal{S}_i[k], \mathcal{I}_i[k], \mathcal{R}_i[k])^T$ for all $k = 1,...,n$
    $\big(\beta_{i1}[k](\delta_i),...,\beta_{iN}[k](\delta_i), \text{MSE}(\delta_i)\big) \leftarrow$ `Network_inference`$\big(\delta_i, v_i[1],...,v_i[n], \mathcal{I}[1],...,\mathcal{I}[n]\big)$
  end for
  $\hat{\delta}_i \leftarrow \underset{\delta_i \in \Omega}{\text{argmin}}\ \text{MSE}(\delta_i)$
  $(\hat{\beta}_{i1}[k],...,\hat{\beta}_{iN}[k]) \leftarrow \beta_{i1}[k]\big(\hat{\delta}_i\big),...,\beta_{iN}[k]\big(\hat{\delta}_i\big)$
end for

Step 3 - Iterating SIR model

for $i = 1,...,N$ **do**
  $\hat{\mathcal{I}}_i[n] \leftarrow \mathcal{I}_i[n]$
  $\hat{\mathcal{R}}_i[1] \leftarrow 0$
  $\hat{\mathcal{R}}_i[k] \leftarrow \hat{\mathcal{R}}_i[k-1] + \hat{\delta}_i \mathcal{I}_i[k-1]$ for all $k = 2,...,n$
end for
for $k = n+1,...,n+n_{\text{pred}}$ **do**
  for $i = 1,...,N$ **do**
    $\hat{\mathcal{I}}_i[k] \leftarrow \big(1-\hat{\delta}_i\big)\hat{\mathcal{I}}_i[k-1] + \big(1-\hat{\mathcal{I}}_i[k-1]-\hat{\mathcal{R}}_i[k-1]\big)\sum_{j=1}^{N}\hat{\beta}_{ij}[k-1]\hat{\mathcal{I}}_j[k-1]$
    $\hat{\mathcal{R}}_i[k] \leftarrow \hat{\mathcal{R}}_i[k-1] + \hat{\delta}_i\hat{\mathcal{I}}_i[k-1]$
  end for
end for

Identifier i	City	Population N_i
1	Wuhan	10,607,700
2	Huanggang	6,291,000
3	Jingzhou	5,705,900
4	Xiangyang	5,614,000
5	Xiaogan	4,878,000
6	Xiantao	1,155,000
7	Yichang	4,115,000
8	Shiyan	3,383,000
9	Enshi (autonomous prefecture)	3,327,000
10	Jingmen	2,896,300
11	Xianning	2,507,000
12	Huangshi	2,458,000
13	Suizhou	2,190,800
14	Ezhou	1,059,500
15	Tianmen	1,292,000
16	Qianjiang	958,000

Table E.1: The regions (prefecture-level divisions) in the province Hubei. We do not consider the city Shennongjia, since the number of SARS-CoV-2 infections in Shennongjia is very small.

To determine the regularisation parameter ρ_i in the LASSO (7.6), we consider 100 candidate values, specified by the set $\Theta_i = \{\rho_{\min,i}, \dots, \rho_{\max,i}\}$. In line 4 of Algorithm 2, the maximum value is set to $\rho_{\max,i} = 2\|F_i^T V_i\|_\infty$. If $\rho_i > \rho_{\max,i}$, then the solution to the LASSO (7.6) is $\beta_{ij} = 0$ for all regions j. For every value of the regularisation parameter $\rho_i \in \Theta_i$, we compute the mean square error MSE (δ_i, ρ_i) by 3-fold cross-validation [177]. For every fold, the rows of the matrix F_i and the vector V_i are divided into a training set $F_{i,\text{train}}, V_{i,\text{train}}$ and a validation set $F_{i,\text{val}}, V_{i,\text{val}}$. We compute the solution $\beta_{i1}, \dots, \beta_{iN}$ to the LASSO (7.6) on the training set of every fold $F_{i,\text{train}}, V_{i,\text{train}}$. The mean squared error MSE (δ_i, ρ_i) then equals

$$\left\| V_{i,\text{val}} - F_{i,\text{val}} \begin{pmatrix} \beta_{i1} \\ \dots \\ \beta_{iN} \end{pmatrix} \right\|_2^2,$$

averaged over all folds. Finally, we set the regularisation parameter ρ_i equal to the case where the MSE (δ_i, ρ_i) is minimised. The final estimate $\beta_{i1}(\delta_i), \dots, \beta_{iN}(\delta_i)$ for the infection probabilities is obtained by solving the LASSO (7.6) on the whole matrix F_i and vector V_i. To solve the LASSO (7.6) numerically, we use the Matlab command `quadprog`.

E.4. MOTIVATION FOR THE STATIC AND DYNAMIC PRIOR

We intend to give a short motivation for (7.7). Suppose that each person has on average $\langle d \rangle$ contacts (here $\langle \cdot \rangle$ denotes the average) in the population. If a person is infected and its contacts are healthy, the person can infect any of its contacts independently with

City	21-1	22-1	23-1	24-1	25-1	26-1	27-1	28-1	29-1	30-1	31-1	1-2	2-2	3-2	4-2	5-2	6-2	7-2	8-2	9-2	10-2	11-2
Wuhan	105	62	70	65	46	80	892	315	356	378	576	894	1033	1242	1967	1766	1501	1985	1378	1921	1552	1104
Huanggang	0	0	0	64	58	32	59	111	172	77	153	276	244	176	223	162	90	144	96	115	80	66
Jingzhou	0	6	2	2	23	14	24	30	50	70	66	46	166	114	100	88	84	56	56	48	30	35
Xiangyang	0	0	0	0	2	34	34	61	32	123	61	94	107	84	103	52	51	69	55	57	44	25
Xiaogan	0	0	22	4	29	45	73	101	125	142	87	121	169	202	342	424	255	172	123	105	101	109
Xiaotao	0	0	2	8	1	1	15	5	23	35	7	43	29	19	37	40	42	52	20	37	22	22
Yichang	0	0	1	0	19	11	20	12	54	50	109	77	39	60	44	67	47	23	71	45	23	12
Shiyan	0	0	1	4	15	20	25	23	31	31	27	35	44	35	27	35	42	43	29	14	24	31
Enshi	0	0	0	11	6	8	13	13	15	9	12	18	6	12	15	6	13	3	6	21	8	8
Jingmen	0	1	0	20	17	52	24	28	49	36	24	78	16	55	22	86	45	35	41	12	15	40
Xianning	0	0	0	0	43	21	27	21	18	36	40	40	50	52	36	15	44	33	17	14	8	10
Huangshi	0	0	0	0	31	5	17	33	27	55	41	43	82	71	104	57	69	68	50	52	30	39
Suizhou	0	0	0	5	31	16	18	46	27	85	76	80	74	183	65	128	81	38	31	65	46	34
Ezhou	0	0	0	1	0	19	37	27	39	66	38	51	28	26	50	41	48	98	67	89	65	71
Tianmen	0	0	0	3	2	8	10	11	10	23	15	17	16	2	11	10	25	16	18	20	44	32
Shennongjia	0	0	0	0	0	0	1	2	2	2	0	0	0	3	0	0	0	0	0	0	0	0
Qianjiang	0	0	0	0	0	5	2	1	2	2	15	8	0	9	10	10	10	6	2	3	5	0

Table E.2: The time series of the reported number of infections $N_{\text{rep},i}[k]$ for every city i in Hubei.

Algorithm 2 `Network inference for NIPA`

Input: curing probability δ_i; viral state $v_i[k]$ for $k = 1,...,n$; infection state vector $\mathcal{I}[k]$ for $k = 1,...,n$
Output: infection probability estimates $\beta_{i1}(\delta_i),...,\beta_{iN}(\delta_i)$; mean squared error $\mathrm{MSE}(\delta_i)$
Compute V_i and F_i by Eqs. (7.4) and (7.5)
$\rho_{\max,i} \leftarrow 2\|F_i^T V_i\|_\infty$
$\rho_{\min,i} \leftarrow 10^{-4}\rho_{\max,i}$
$\Theta_i \leftarrow 100$ logarithmically equidistant values from $\rho_{\min,i}$ to $\rho_{\max,i}$
for $\rho_i \in \Theta_i$ **do**
    estimate $\mathrm{MSE}(\delta_i,\rho_i)$ by 3-fold cross-validation on F_i, V_i and solving (7.6) on the respective training set
end for
$\rho_{\mathrm{opt},i} \leftarrow \underset{\rho_i\in\Theta_i}{\operatorname{argmin}}\ \mathrm{MSE}\left(\delta_i,\rho_i\right)$
$(\beta_{i1}(\delta_i),...,\beta_{iN}(\delta_i)) \leftarrow$ the solution to (7.6) on the whole data set F_i, V_i for $\rho_i = \rho_{\mathrm{opt},i}$
$\mathrm{MSE}(\delta_i) \leftarrow \mathrm{MSE}(\delta_i,\rho_{\mathrm{opt},i})$

probability p. Hence, the total number of infections follows a Bernoulli distribution

$$\Pr[m] = \binom{\langle d\rangle}{m} p^m (1-p)^{\langle d\rangle - m}. \tag{E.7}$$

In case $\langle d\rangle$ is large and $\lambda \equiv p\langle d\rangle$ is small, we can approximate (E.7) by a Poisson distribution

$$\Pr[m] = e^{-\lambda}\frac{\lambda^m}{m!}. \tag{E.8}$$

If there are N visiting, infected individuals, which may all infect the population independently, the resulting distribution is the sum of independent, identically distributed Poisson distributions, which is again a Poisson distribution with $\langle m\rangle = N\lambda$.

We denote the number of people living in region j and travelling for work to region i by m_{ij}. Each individual has $\langle d\rangle$ contacts and can infect each individual with probability p. Then region j has on average $m_{ij}\langle d\rangle p$ new infections, provided that no two individuals who visit the same region j have contact to the same people. In particular, the fraction of new infections that region i gets from region j is given by

$$\beta_{ij} = \frac{m_{ij}\langle d\rangle p}{N_i}. \tag{E.9}$$

If we define $c_i = \frac{\langle d\rangle p}{N_i}$, we obtain equation (7.7).

E.5. DETAILS ON NIPA STATIC PRIOR

We assume that the infection matrix B is normally distributed around the prior B_{prior}, whose elements equal $b_{\mathrm{prior},ij} = c_i m_{ij}$:

Assumption E.4 *Every non-diagonal element β_{ij}, where $i \neq j$, of the matrix B is normally distributed as*

$$\Pr[\beta_{ij}] = \begin{cases} \alpha_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}\left(\beta_{ij} - c_i m_{ij}\right)^2\right), & \text{if} \quad 0 \leq \beta_{ij} \leq 1 \\ 0. & \textit{otherwise} \end{cases} \tag{E.10}$$

Here c_i denotes the proportionality constant, *and the constant α_i is set such that*

$$\int_{\mathbb{R}} \Pr[\beta_{ij}]\, d\beta_{ij} = 1. \tag{E.11}$$

The normal distribution (E.10) is cut off for values outside of $[0,1]$, since the infection probability β_{ij} cannot be outside the interval $[0,1]$. The standard deviation σ_i is a measure for the accuracy of the prior distribution (E.10). Both the proportionality constant c_i and the standard deviation σ_i are unknown. Assumption E.4 does not hold for the diagonal elements β_{ii} of the matrix B. Instead, we assume that β_{ii} are uniformly distributed in the interval $[0,1]$.

We obtain the estimate $B_{\text{posterior}}$ of the contact network by a Bayesian approach. Given the observed $N \times 1$ infection vector $\mathcal{I}[k] = (\mathcal{I}_1[k], ..., \mathcal{I}_N[k])^T$ at all times $k = 1, ..., n$, we pose the optimisation problem

$$\begin{aligned} \hat{B} = \underset{B}{\operatorname{argmax}} \ & \Pr\left[B \middle| \mathcal{I}[1], ..., \mathcal{I}[n]\right] \\ \text{s.t.} \ & \sum_{j=1}^{N} \beta_{ij} \leq 1, \quad i = 1, ..., N. \end{aligned} \tag{E.12}$$

With the constraint in (E.12), we ensure that the predictions of the infections satisfy $0 \leq \mathcal{I}_i[k] \leq 1$, see Lemma E.3 in Appendix E.1.

The Bayesian estimate $\hat{B}$ is obtained by solving a constrained linear least-squares problem defined in Proposition E.5.

Proposition E.5 *Under Assumptions E.1 and E.4, the Bayesian estimation problem (E.12) is equivalent to solving the optimisation problem*

$$\begin{aligned} \min_{\beta_{i1},...,\beta_{iN}} \quad & \left\| V_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|_2^2 + \rho_i \sum_{j=1, j\neq i}^{N} \left(\beta_{ij} - c_i m_{ij}\right)^2 \\ \text{s.t.} \quad & 0 \leq \beta_{ij} \leq 1, \quad j = 1, ..., N, \\ & \sum_{j=1}^{N} \beta_{ij} \leq 1, \end{aligned} \tag{E.13}$$

for every region i, where the penalisation parameter equals $\rho_i = \sigma_w^2/\sigma_i^2$ and V_i and F_i are defined in Eqs. (7.4) and (7.5).

Proof. The objective function of the optimisation problem (E.12) is equivalent to

$$\hat{B} = \underset{B}{\operatorname{argmax}} \log\left(\Pr\left[B\right]\right) + \sum_{k=2}^{n} \log\left(\Pr\left[\mathcal{I}[k] \middle| \mathcal{I}[k-1], B\right]\right). \tag{E.14}$$

In the following, we rewrite the two terms in (E.14). First, with (E.10), it holds that

$$\log\left(\Pr\left[B\right]\right) = \begin{cases} \sum_{i=1}^{N}\sum_{j=1}^{N} \log\left(\alpha_i\right) - \log\left(\sqrt{2\pi}\sigma_i\right) - \dfrac{1}{2\sigma_i^2}\left(\beta_{ij} - c_i m_{ij}\right)^2, & \text{if } 0 \le \beta_{ij} \le 1 \\ -\infty. & \text{otherwise} \end{cases} \tag{E.15}$$

Neither the term $\log\left(\alpha_i\right)$ nor the term $\log\left(\sqrt{2\pi}\sigma_i\right)$ depend on the matrix B. Furthermore, the prior $\log\left(\Pr\left[B\right]\right)$ is finite only if $0 \le \beta_{ij} \le 1$ for all regions i, j. Thus, the optimisation problem (E.14) is equivalent to

$$\begin{aligned} \hat{B} = \underset{B}{\operatorname{argmax}} \quad & \sum_{i=1}^{N}\sum_{j=1}^{N} -\frac{1}{2\sigma_i^2}\left(\beta_{ij} - c_i m_{ij}\right)^2 + \sum_{k=2}^{n} \log\left(\Pr\left[\mathcal{I}[k] \middle| \mathcal{I}[k-1], B\right]\right) \\ \text{s.t.} \quad & 0 \le \beta_{ij} \le 1, \quad i = 1, \ldots, N, \quad j = 1, \ldots, N. \end{aligned} \tag{E.16}$$

Since the model errors $w_i[k]$ are stochastically independent for different regions i, we can rewrite the second term in the objective of (E.16) as

$$\log\left(\Pr\left[\mathcal{I}[k] \middle| \mathcal{I}[k-1], B\right]\right) = \sum_{i=1}^{N} \log\left(\Pr\left[\mathcal{I}_i[k] \middle| \mathcal{I}[k-1], B\right]\right) \tag{E.17}$$

$$= \sum_{i=1}^{N} \log\left(\Pr\left[w_i[k] = \Delta_i[k]\right]\right), \tag{E.18}$$

where the second equality follows from (E.1) and by defining

$$\Delta_i[k] = \mathcal{I}_i[k] - (1 - \delta_i)\mathcal{I}_i[k-1] + \mathcal{S}_i[k-1] \sum_{j=1}^{N} \beta_{ij}\mathcal{I}_j[k-1]. \tag{E.19}$$

Under Assumption E.1, the model error $w_i[k]$ follows the normal distribution. Thus, it holds that

$$\log\left(\Pr\left[w_i[k] = \Delta_i[k]\right]\right) = -\log\left(\sqrt{2\pi}\sigma_w\right) - \frac{1}{2\sigma_w^2}\Delta_i^2[k]. \tag{E.20}$$

The term $\log\left(\sqrt{2\pi}\sigma_w\right)$ is independent of the matrix B. Thus, it follows from (E.17) and (E.20) that the second term in the objective of (E.16) can be replaced by

$$\sum_{i=1}^{N}\sum_{k=2}^{n} \frac{1}{2\sigma_w^2}\Delta_i^2[k] = \sum_{i=1}^{N} \frac{1}{2\sigma_w^2} \left\| V_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|_2^2, \tag{E.21}$$

where the equality follows from the definition of the vector V_i and the matrix F_i in (7.4) and (7.5), respectively. Hence, the optimisation problem (E.16) becomes

$$\begin{aligned} \hat{B} = \underset{B}{\operatorname{argmin}} \quad & \sum_{i=1}^{N} \frac{1}{2\sigma_w^2} \left\| V_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|_2^2 + \sum_{i=1}^{N} \frac{1}{2\sigma_i^2} \sum_{j=1}^{N} \left(\beta_{ij} - c_i m_{ij}\right)^2 \\ \text{s.t.} \quad & 0 \leq \beta_{ij} \leq 1, \quad i = 1, ..., N, \quad j = 1, ..., N. \end{aligned} \tag{E.22}$$

The problem (E.22) can be optimised independently for every region i. Thus, we obtain, after multiplication with $2\sigma_w^2$, the equivalent optimisation problems for every region i as

$$\begin{aligned} \min_{\beta_{i1},...,\beta_{iN}} \quad & \left\| V_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|_2^2 + \frac{\sigma_w^2}{\sigma_i^2} \sum_{j=1}^{N} \left(\beta_{ij} - c_i m_{ij}\right)^2 \\ \text{s.t.} \quad & 0 \leq \beta_{ij} \leq 1, \quad j = 1, ..., N. \end{aligned} \tag{E.23}$$

By introducing $\rho_i = \sigma_w^2/\sigma_i^2$, we obtain that (E.23) with the constraint $\sum_{j=1}^{N} \beta_{ij} \leq 1$ is equivalent to the constrained linear least-squares problem (E.13). □

The first term in the objective of (E.13) measures the fit to the observed epidemic data. The second term measures the deviation of the infection probabilities β_{ij} from the prior (E.10). The scalar parameter ρ_i balances the two terms: if the prior (E.10) is very accurate or the model errors $w_i[k]$ are large, then ρ_i should be large. The optimal value of the parameter ρ_i equals to the ratio of the unknown variances σ_w^2 and σ_i^2 of the model errors $w_i[k]$ and the prior (E.10), respectively. The optimisation problem (E.13) is convex and can be solved efficiently [178]. To obtain the solution to (E.13) numerically, we use the Matlab command `lsqlin`.

We stress the similarity of the optimisation problem (E.13) to the LASSO [148], which is the basis of original NIPA (recall Appendix E.3). Original NIPA and NIPA static prior can be interpreted as Bayesian estimation approaches [149]. The main difference is the inclusion of the prior information on the network in NIPA static prior, whereas original NIPA favours a sparse graph. The second difference is that the second term in the objective of (E.13) contains a least-square term, whereas original NIPA contains an ℓ_1-norm penalisation term of the form

$$\rho_i \sum_{j=1, j\neq i}^{N} \left|\beta_{ij}\right|. \tag{E.24}$$

E.5.1. PSEUDOCODE

To solve the optimisation problem (E.13) for the infection probabilities β_{i1}, ..., β_{iN}, we must specify three unknown variables, namely the curing probability δ_i, the parameter ρ_i and the proportionality constant c_i for every region i. The three unknown variables δ_i, ρ_i and c_i are all determined using cross-validation.

Algorithmically, NIPA static prior is similar to original NIPA, except for two alterations. First, we solve the constrained linear least-squares problem (E.13) instead of

LASSO. Second, additionally to the parameter ρ_i and the curing probability δ_i, for NIPA static prior there is one more unknown variable, namely the proportionality constant c_i, which is a parameter of the prior distribution (E.10). To determine the constant c_i, we consider 50 logarithmically equidistant candidate values in the set $\Psi = \{c_{\min}, ..., c_{\max}\}$. The minimal and the maximal values are set to $c_{\min} = 0.01$ and $c_{\max} = 100$, respectively. To obtain the epidemic outbreak prediction of Bayesian NIPA, we execute Algorithm 1, but instead of using the NIPA network inference algorithm 2, we use Algorithm 3.

Algorithm 3 `Network inference for NIPA static prior`

Input: curing probability δ_i; viral state $v_i[k]$ for $k = 1, ..., n$; infection state vector $\mathcal{I}[k]$ for $k = 1, ..., n$
Output: infection probability estimates $\beta_{i1}(\delta_i), ..., \beta_{iN}(\delta_i)$; mean squared error $\text{MSE}(\delta_i)$
Compute V_i and F_i
$\rho_{\max,i} \leftarrow 2\|F_i^T V_i\|_\infty$
$\rho_{\min,i} \leftarrow 10^{-4}\rho_{\max,i}$
$\Theta_i \leftarrow$ 100 logarithmically equidistant values from $\rho_{\min,i}$ to $\rho_{\max,i}$
$\Psi \leftarrow$ 50 logarithmically equidistant values from $c_{\min} = 0.01$ to $c_{\max} = 100$
for $\rho_i \in \Theta_i$ **do**
    for $c_i \in \Psi$ **do**
        estimate $\text{MSE}(\delta_i, \rho_i, c_i)$ by 3-fold cross–validation on F_i, V_i and solving (E.13) on the respective training set
    end for
end for
$(\rho_{\text{opt},i}, c_{\text{opt},i}) \leftarrow \underset{\rho_i \in \Theta_i, c_i \in \Psi}{\text{argmin}} \text{MSE}(\delta_i, \rho_i, c_i)$
$(\beta_{i1}(\delta_i), ..., \beta_{iN}(\delta_i)) \leftarrow$ the solution to (E.13) on the whole data set F_i, V_i for $\rho_i = \rho_{\text{opt},i}$ and $c_i = c_{\text{opt},i}$
$\text{MSE}(\delta_i) \leftarrow \text{MSE}(\delta_i, \rho_{\text{opt},i}, c_{\text{opt},i})$

E.6. DETAILS ON NIPA DYNAMIC PRIOR

We assume that the time-varying infection probabilities $\beta_{ij}[k]$ are proportional to the known population flow $m_{ij}[k]$. More precisely, we assume that the infection probabilities $\beta_{ij}[k]$ for all regions i, j, when $i \neq j$, equal

$$\beta_{ij}[k] = c_i m_{ij}[k] \tag{E.25}$$

for some unknown proportionality constant $c_i > 0$. Furthermore, we assume the self-infection probabilities β_{ii} do *not* change over time k. With (E.25), the SIR model in Definition 7.1 becomes

$$\mathcal{I}_i[k+1] = (1-\delta_i)\mathcal{I}_i[k] + \beta_{ii}\mathcal{S}_i[k]\mathcal{I}_i[k] + c_i\mathcal{S}_i[k]\sum_{j=1, j\neq i}^{N} m_{ij}[k]\mathcal{I}_j[k] + w_i[k]. \tag{E.26}$$

E.6.1. MAXIMUM-LIKELIHOOD ESTIMATION

To predict the infectious state $\mathcal{I}_i[k]$ with (E.26), we must estimate the constants c_i, the self-infection probabilities β_{ii} and the curing probabilities δ_i. We define the $N \times 1$ vectors $c = (c_1, ..., c_N)^T$ and $b = (\beta_{11}, ..., \beta_{NN})^T$. We pose the estimation problem in a maximum-likelihood sense as

$$\begin{aligned} \max_{c,b} \quad & \Pr\left[\mathcal{I}[1], ..., \mathcal{I}[n] \middle| c, b\right] \\ \text{s.t.} \quad & c_i \geq 0, \quad i = 1, ..., N, \\ & \beta_{ii} \geq 0, \quad i = 1, ..., N, \\ & \beta_{ii} + c_i \sum_{j=1, j\neq i}^{N} m_{ij}[k] \leq 1 \quad i = 1, ..., N, k = 1, ..., n. \end{aligned} \tag{E.27}$$

The last constraint in (E.27) ensures that the predictions of the infections satisfy $\mathcal{I}_i[k] \leq 1$, see Lemma E.3. From the maximum likelihood problem (E.27), we derive in a similar fashion as Proposition E.5 the LASSO optimisation problem as

$$\begin{aligned} \min_{c_i, \beta_{ii}} \quad & \sum_{k=1}^{n-1} \left(\mathcal{I}_i[k+1] - (1-\delta_i)\mathcal{I}_i[k] - \beta_{ii}\mathcal{S}_i[k]\mathcal{I}_i[k] - c_i\mathcal{S}_i[k] \sum_{j=1, j\neq i}^{N} m_{ij}[k]\mathcal{I}_j[k] \right)^2 \\ & + \rho_i(\beta_{ii} + c_i) \\ \text{s.t.} \quad & c_i \geq 0, \\ & \beta_{ii} \geq 0, \\ & \beta_{ii} + c_i \sum_{j=1, j\neq i}^{N} m_{ij}[k] \leq 1, \quad k = 1, ..., n \end{aligned} \tag{E.28}$$

for every region i. Here, we denote the regularisation parameter by $\rho_i \geq 0$, which aims to avoid overfitting. The greater the parameter ρ_i, the smaller the estimates of the coefficients β_{ii}, c_i. If the regularisation parameter $\rho_i = 0$, then solving the LASSO (E.28) for every node i is equivalent to solving the maximum-likelihood problem (E.27).

To solve the optimisation problem (E.28) for the constants c_i and β_{ii}, we must specify two unknown variables. First, the curing probability δ_i of region i, which determines the fractions $\mathcal{S}_i[k]$ and $\mathcal{R}_i[k]$ of susceptible and recovered individuals, respectively. Second, we must specify the parameter ρ_i. We perform hold-out cross-validation to set the unknown variables δ_i and ρ_i: The training set follows from the first 80% of the observations, and the validation set equals the last 20% of the observations. In pseudocode, NIPA dynamic prior is given by Algorithm 4.

E.7. NIPA STATIC PRIOR UNDER PERFECT CONDITIONS

The original NIPA method is known to provide accurate predictions when the epidemic perfectly follows the SIR model [128, Supplementary Material 1]. Here, we intend to show that NIPA static prior performs even better if the prior matrix is close to the real infection matrix.

Algorithm 4 `Network inference for NIPA dynamic prior`

Input: curing probability δ_i; viral state $v_i[k]$ for $k = 1,...,n$; infection state vector $\mathcal{I}[k]$ for $k = 1,...,n$
Output: infection probability estimates $\beta_{i1}(\delta_i),...,\beta_{iN}(\delta_i)$; mean squared error $\text{MSE}(\delta_i)$
Compute V_i and F_i
$\rho_{\max,i} \leftarrow 2\|F_i^T V_i\|_\infty$
$\rho_{\min,i} \leftarrow 10^{-4}\rho_{\max,i}$
$\Theta_i \leftarrow 100$ logarithmically equidistant values from $\rho_{\min,i}$ to $\rho_{\max,i}$
for $\rho_i \in \Theta_i$ **do**
    estimate $\text{MSE}(\delta_i,\rho_i)$ by hold–out cross–validation on F_i, V_i and solving (E.28) on the respective training set
end for
$\rho_{\text{opt},i} \leftarrow \underset{\rho_i\in\Theta_i}{\text{argmin}}\,\text{MSE}\left(\delta_i,\rho_i\right)$
$(\beta_{i1}(\delta_i),...,\beta_{iN}(\delta_i)) \leftarrow$ the solution to (E.28) on the whole data set F_i, V_i for $\rho_i = \rho_{\text{opt},i}$
$\text{MSE}(\delta_i) \leftarrow \text{MSE}(\delta_i,\rho_{\text{opt},i})$

We generate data from an SIR epidemic as in Definition 7.1 on a network with $N = 10$ nodes with the same curing probability $\delta_i = 0.2$ for every node i. We set the curing probability δ_i in the NIPA algorithms equal to the exact curing probabilities $\delta_i = 0.2$, such that both NIPA and NIPA static prior always exactly recover the exact curing probabilities. We generate uniformly distributed infection probabilities β_{ij} in the interval $(0,1)$. The effective reproduction number R_0 can be computed as [122]

$$R_0 = \text{maximum eigenvalue of } \left(B \cdot \text{diag}\left(\frac{1}{\delta_1},...,\frac{1}{\delta_N}\right)\right). \tag{E.29}$$

We normalise B element-wise such that the basic reproduction number R_0 equals 2.0. Furthermore, we set the population size N_i for each region i equal to a uniformly distributed number in the interval $[10^5, 10^6]$ and start with initially 100 infected cases in node 1 and all other nodes are healthy. Most importantly, we set the prior infection matrix B_{prior} equal to the exact infection matrix B, multiplied by some noise

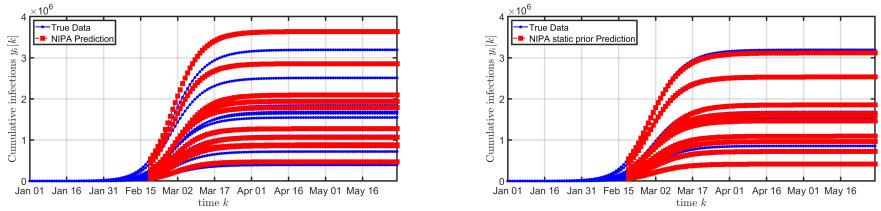
$$B_{\text{prior},ij} = \beta_{ij} w_{ij}. \tag{E.30}$$

Here, w_{ij} is uniformly distributed in the interval $[1,2]$. The other parameters are the same as in the main article.

The result in Figure E.1 is clear: NIPA static prior is able to capture the dynamics much better than NIPA. Hence, we conclude that NIPA static prior in combination with a good prior yields a better prediction accuracy than the original NIPA method.

E.8. SIGMOID CURVES

In epidemiology, sigmoid curves are commonly used to forecast the future number of infected cases.



(a) NIPA (b) NIPA static prior

Figure E.1: The prediction for (a) NIPA and (b) NIPA static prior with generated SIR data based on Definition 7.1 on a 10-node network.

The logistic function was developed by Verhulst in 1845 to explain population growth in a specific region [126]. The logistic function is the most commonly used sigmoid curve in epidemiology, because the logistic function is the (approximate) solution of the SIS and SIR model [131]. The logistic function assumes the cumulative fraction of infected cases $\mathcal{I}_{\text{cum},i}[k]$ in region i and at time k to follow

$$\mathcal{I}_{\text{cum},i}[k] = \frac{\mathcal{I}_{\infty,i}}{1+e^{-K_i(k-t_{0,i})}}, \tag{E.31}$$

where $\mathcal{I}_{\infty,i}$ is the long-term fraction of infections, K_i is the logistic growth rate and $t_{0,i}$ is the inflection point, which is also known as the epidemic peak.

The Hill function was introduced in 1910 to describe the binding of molecules on surfaces [133]. Later, it was successfully applied to describe the spread of epidemics [179]. The Hill function assumes the cumulative fraction of infected cases $\mathcal{I}_i[k]$ in region i at time k to follow

$$\mathcal{I}_{\text{cum},i}[k] = \frac{\mathcal{I}_{\infty,i}}{1+\left(\frac{K_i}{k-t_{0,i}}\right)^{n_i}}, \tag{E.32}$$

where $\mathcal{I}_{\infty,i}$ is the long-term fraction of infections, K_i is the Hill growth rate, n_i is the Hill coefficient and $t_{0,i}$ is the inflection point, also known as the epidemic peak.

The Gompertz function was introduced in 1825 to describe mortality in human populations [134]. The Gompertz function was also successfully used to describe the spread of epidemics [180]. The Gompertz function assumes the cumulative fraction of infected cases $\mathcal{I}_{\text{cum},i}[k]$ in region i at time k to follow

$$\mathcal{I}_{\text{cum},i}[k] = \mathcal{I}_{\infty,i}e^{-c_i e^{-a_i k}}, \tag{E.33}$$

where $\mathcal{I}_{\infty,i}$ is the long-term fraction of infections, c_i is a displacement factor (comparable to the inflection point) and a_i is the Gompertz growth rate.

We describe the curve-fitting procedure here for the logistic function, but the parameters for any curve can be estimated analogously. Suppose that we have a time series of

the cumulative fraction of reported cases $\mathcal{I}_{\text{rep},i}[k]$ for $k = 1,\dots,n$ and for every region i. Then we minimise the Mean Square Error for each region i separately;

$$(\hat{y}_{\infty,i}, \hat{K}_i, \hat{t}_{0,i}) = \min_{(y_{\infty,i},K_i,t_{0,i})} \sum_{k=1}^{n} \left(\mathcal{I}_{\text{rep},i}[k] - \frac{\mathcal{I}_{\infty,i}}{1+e^{-K_i(k-t_{0,i})}} \right)^2, \quad \text{s.t.} \quad \begin{aligned} &0 \le \mathcal{I}_{\infty,i} \le 1, \\ &K_i \ge 0, \\ &t_{0,i} \ge 0. \end{aligned} \tag{E.34}$$

Naturally, we have imposed that the cumulative fraction of reported cases $\mathcal{I}_i[k]$ for every region i at time k is always smaller than the total normalised population for every region i. We evaluate the nonlinear minimisation problem (E.34) by the command `GlobalSearch` in Matlab. As initial conditions, we set $\mathcal{I}_{\infty,i} = \mathcal{I}_{\text{cum},i}[t_{\text{obs}}], K_i = 1, t_{0,i} = t_{\text{obs}}$ where t_{obs} is the observation time, i.e. the time at which the forecast is made. The parameters $(y_{\infty,i}, K_i, n_i, t_{0,i})$ for the Hill function and $(y_{\infty,i}, c_i, a_i)$ for the Gompertz function are estimated analogously.

E

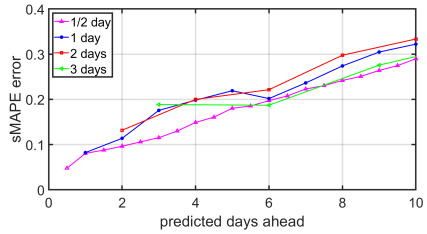
E.9. THE INFLUENCE OF THE TIME STEP ON THE ACCURACY

In the discrete-time SIR model (7.1), we use the time step $\Delta t = 1$ day. By approximating a continuous-time process (the COVID-19 pandemic) by a discrete-time process (SIR model (7.1)) we make a model error. We investigate the influence of the time step on the prediction accuracy, by comparing the NIPA prediction accuracy for various time steps, ranging from $\Delta t = 0.5$ days to $\Delta t = 3$ days. Since the number of infected cases is reported once a day, the data for the time step $\Delta t = 0.5$ days is obtained by linearly interpolating the number of cumulative cases $\mathcal{I}_{\text{cum},i}[k]$. For a time step $\Delta t = 1$ day and $\Delta t = 0.5$ days, we smooth the raw data before calling the NIPA algorithm.

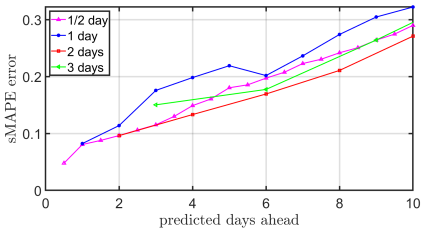
For the time steps $\Delta t = 2$ days and $\Delta t = 3$ days, there are two possible methods. Method (A) assumes that the cumulative number of cases $\mathcal{I}_{\text{cum},i}[k]$ is reported every two (or three) days and is unreported on the intermediate days. Then we smooth the remaining data, whereafter the NIPA algorithm is used. Thus, we simply omitted the data on the intermediate days. In contrast, method (B) first smooths all raw data. Thereafter, we only use the cumulative number of cases $\mathcal{I}_i[k]$ every two or three days for a time step of two or three days, respectively. The main difference is that method (A) completely neglects the data on intermediate days, whereas method (B) first applies a smoother and then neglects the intermediate data. We emphasise that the procedure for the time step $\Delta t = 1$ day and $\Delta t = 0.5$ days is equal for method (A) and (B).

Figure E.2 shows an exemplary situation from the Netherlands for three initial dates. In the beginning of the COVID-19 outbreak, as shown in Figure E.2a and E.2b, the prediction accuracy is similar for all time steps. The small amount of available data and the rapidly increasing number of cases hampers accurate forecasting. As the epidemic evolves, method (A) and method (B) start to deviate. By omitting data as in method (A), the sMAPE error in Figure E.2c and E.2e increases for two and three days quicker than for smaller time steps. Hence, removing data causes the prediction accuracy to decrease. On the other hand, method (B) in Figures E.2d and E.2f shows similar behaviour for all time

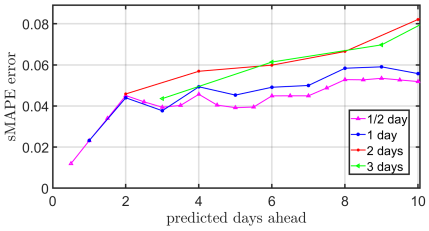
steps. We conclude that the choice of the time step has limited effect on the prediction accuracy and we keep the time step $\Delta t = 1$ for convenience.



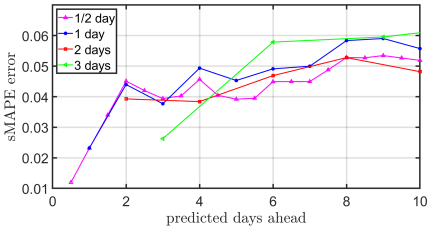
(a) Method A, 18 March



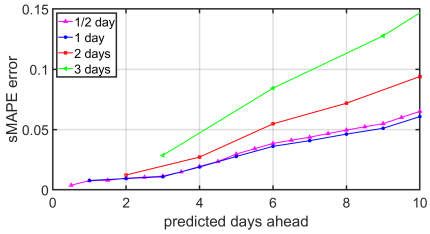
(b) Method B, 18 March



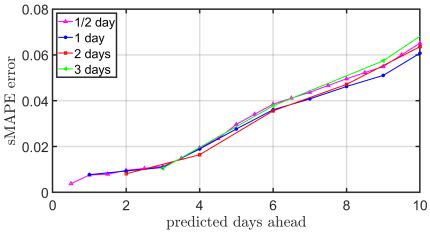
(c) Method A, April 5



(d) Method B, April 5



(e) Method A, April 23



(f) Method B, April 23

Figure E.2: The NIPA prediction accuracy for the situation in the Netherlands for varying time steps Δt. The subplots show the forecast for (a-b) March 18, (c-d) April 5 and (d-e) April 23 and show the situation for method (A) in (a,c,e) and method (B) in (b,d,f). For the time step $\Delta t = 2$ days or $\Delta t = 3$ days, the data is first smoothed and then removed.

BIBLIOGRAPHY

[1] J. P. Byrne and J. N. Hays. *Epidemics and pandemics: From ancient plagues to modern-day threats*. Greenwood, an imprint of ABC-CLIO, LLC, first edition, 2021.

[2] P. Cirillo and N.N. Taleb. Tail risk of contagious diseases. *Nat. Phys.*, 16:606–613, 2020.

[3] D. Bernoulli. Essai d'une nouvelle analyse de la mortalite causee par la petite verole. *Mem. Math. Phys. Acad. Roy. Sci.*, 1, 1766. Reprinted in: L.P. Bouckaert, B.L. van der Waerden (Eds.), Die Werke von Daniel Bernoulli, Bd. 2 Analysis und Wahrscheinlichkeitsrechnung, Birkhäuser, Basel, 1982, p. 235. English translation entitled 'An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it' in: L. Bradley, Smallpox Inoculation: An Eighteenth Century Mathematical Controversy, Adult Education Department, Nottingham, 1971, p. 21. Reprinted in: S. Haberman, T.A. Sibbett (Eds.) History of Actuarial Science, vol. VIII, Multiple Decrement and Multiple State Models, William Pickering, London, 1995, p. 1.

[4] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*, 115:700–721, Aug 1927.

[5] P. Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, first edition, Mar 2011.

[6] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, Aug 2015.

[7] I. Tunc, M. S. Shkarayev, and L. B. Shaw. Epidemics in Adaptive Social Networks with Temporary Link Deactivation. *J. Stat. Phys.*, 151:355–366, Apr 2013.

[8] R. Pastor-Satorras and A. Vespignani. Epidemic Spreading in Scale-Free Networks. *Phys. Rev. Lett.*, 86:3200–3203, Apr 2001.

[9] J. Meier, X. Zhou, A. Hillebrand, P. Tewarie, C. J. Stam, and P. Van Mieghem. The epidemic spreading model and the direction of information flow in brain networks. *NeuroImage*, 152:639 – 646, 2017.

[10] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen. The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106(16):6872–6877, 2009.

[11] F. Jin, E. Dougherty, P. Saraf, Y. Cao, and N. Ramakrishnan. Epidemiological Modeling of News and Rumors on Twitter. In *Proceedings of the 7th Workshop on Social*

Network Mining and Analysis, SNAKDD '13, New York, NY, USA, 2013. Association for Computing Machinery.

[12] F. M. Bass. A New Product Growth for Model Consumer Durables. *Management Science*, 15(5):215–227, 1969.

[13] I. Seoane, E. Calle, J. A. Hernández, J. Segovia, R. Romeral, P. Vilà, M. Urueña, and M. Manzano. Failure propagation in GMPLS optical rings: CTMC model and performance analysis. *Optical Switching and Networking*, 9(1):39–51, 2012.

[14] A. Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.*, 8:32–39, Dec 2011.

[15] M. A. Achterberg, B. Prasse, and P. Van Mieghem. Analysis of continuous-time Markovian ε-SIS epidemics on networks. *Phys. Rev. E*, 105:054305, May 2022.

[16] P. Van Mieghem and E. Cator. Epidemics in networks with nodal self-infection and the epidemic threshold. *Phys. Rev. E*, 86:016116, Jul 2012.

[17] W. Dhouib, J. Maatoug, I. Ayouni, N. Zammit, R. Ghammem, S. Ben Fredj, and H. Ghannem. The incubation period during the pandemic of COVID-19: a systematic review and meta-analysis. *Systematic Reviews*, 10(101), 2021.

[18] W. van der Toorn, D. Oh, D. Bourquain, J. Michel, E. Krause, A. Nitsche, and M. von Kleist. An intra-host SARS-CoV-2 dynamics model to assess testing and quarantine strategies for incoming travelers, contact management, and de-isolation. *Patterns*, 2(6):100262, 2021.

[19] P. Van Mieghem and Q. Liu. Explicit non-Markovian susceptible-infected-susceptible mean-field epidemic threshold for Weibull and Gamma infections but Poisson curings. *Phys. Rev. E*, 100:022317, Aug 2019.

[20] I. Nåsell. *Extinction and Quasi-Stationarity in the Stochastic Logistic SIS Model*. Springer-Verlag, Berlin, first edition, 2011.

[21] P. Van Mieghem. Explosive phase transition in susceptible-infected-susceptible epidemics with arbitrary small but nonzero self-infection rate. *Phys. Rev. E*, 101:032303, Mar 2020.

[22] J. A. Jacquez and C. P. Simon. The stochastic SI model with recruitment and deaths I. comparison with the closed SIS model. *Mathematical Biosciences*, 117(1):77–125, 1993.

[23] E. Cator and P. Van Mieghem. Susceptible-infected-susceptible epidemics on the complete graph and the star graph: Exact analysis. *Phys. Rev. E*, 87:012811, Jan 2013.

[24] M. M. de Oliveira and R. Dickman. How to simulate the quasistationary state. *Phys. Rev. E*, 71:016129, Jan 2005.

[25] M.J. Keeling and J.V. Ross. On methods for studying stochastic disease dynamics. *J. R. Soc. Interface*, 5:171–181, 2008.

[26] A. L. Hill, D. G. Rand, M. A. Nowak, and N. A. Christakis. Emotions as infectious diseases in a large social network: the SISa model. *Proc. R. Soc. B.*, 277:3827–3835, 2010.

[27] P.L. Simon, M. Taylor, and I.Z. Kiss. Exact epidemic models on graphs using graph-automorphism driven lumping. *J. Math. Bio.*, 62:479–508, Apr 2011.

[28] A. Lajmanovich and J. A. Yorke. A deterministic model for gonorrhea in a nonhomogeneous population. *Mathematical Biosciences*, 28(3):221–236, 1976.

[29] P. Van Mieghem. The N-intertwined SIS epidemic network model. *Computing*, 93:147–169, Dec 2011.

[30] K. Devriendt and P. Van Mieghem. Unified mean-field framework for susceptible-infected-susceptible epidemics on networks, based on graph partitioning and the isoperimetric inequality. *Phys. Rev. E*, 96:052314, Nov 2017.

[31] R. van de Bovenkamp and P. Van Mieghem. Survival time of the susceptible-infected-susceptible infection process on a graph. *Phys. Rev. E*, 92:032806, Sep 2015.

[32] P. Van Mieghem and R. van de Bovenkamp. Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks. *Phys. Rev. E*, 91:032812, Mar 2015.

[33] D. G. Kendall. On the Generalized “Birth-and-Death" Process. *The Annals of Mathematical Statistics*, 19(1):1–15, 1948.

[34] M. Assaf and B. Meerson. Spectral Theory of Metastability and Extinction in Birth-Death Systems. *Phys. Rev. Lett.*, 97:200602, Nov 2006.

[35] P. Van Mieghem and F. Wang. Time dependence of susceptible-infected-susceptible epidemics on networks with nodal self-infections. *Phys. Rev. E*, 101:052310, May 2020.

[36] S. Karlin and J. McGregor. Coincidence probabilities. *Pac. J. of Math.*, 9(4):1141 – 1164, 1959.

[37] C. R. Doering, K. V. Sargsyan, and L. M. Sander. Extinction Times for Birth-Death Processes: Exact Results, Continuum Asymptotics, and the Failure of the Fokker–Planck Approximation. *Multiscale Modeling & Simulation*, 3(2):283–299, 2005.

[38] J. R. Artalejo. On the time to extinction from quasi-stationarity: A unified approach. *Physica A: Statistical Mechanics and its Applications*, 391(19):4483–4486, 2012.

[39] P. Holme and L. Tupikina. Epidemic extinction in networks: insights from the 12 110 smallest graphs. *New Journal of Physics*, 20(11):113042, Nov 2018.

[40] P. Van Mieghem. Decay towards the overall-healthy state in SIS epidemics on networks, 2013. ArXiv preprint: https://arxiv.org/abs/1310.3980.

[41] H. Andersson and B. Djehiche. A Threshold Limit Theorem for the Stochastic Logistic Epidemic. *Journal of Applied Probability*, 35(3):662–670, 1998.

[42] R. van de Bovenkamp and P. Van Mieghem. Time to Metastable State in SIS Epidemics on Graphs. In *2014 Tenth International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 347–354, Los Alamitos, CA, USA, Nov 2014. IEEE Computer Society.

[43] Z. He and P. Van Mieghem. The spreading time in SIS epidemics on networks. *Physica A: Statistical Mechanics and its Applications*, 494:317 – 330, 2018.

[44] P. Van Mieghem. *Performance Analysis of Complex Networks and Systems*. Cambridge University Press, Cambridge, United Kingdom, 2014.

[45] B. N. Parlett. *The Symmetric Eigenvalue Problem*. Society for Industrial and Applied Mathematics, first edition, 1980.

[46] A. Ganesh, L. Massoulié, and D. Towsley. The effect of network topology on the spread of epidemics. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, volume 2, pages 1455–1466 vol. 2, 2005.

[47] E. A. van Doorn, A. I. Zeifman, and T. L. Panfilova. Bounds and Asymptotics for the Rate of Convergence of Birth-Death Processes. *Theory of Probability & Its Applications*, 54(1):97–113, 2010.

[48] B. Prasse, K. Devriendt, and P. Van Mieghem. Clustering for epidemics on networks: A geometric approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(6):063115, 2021.

[49] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint. In *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.*, pages 25–34, 2003.

[50] P. L. Simon and I. Z. Kiss. From exact stochastic to mean-field ODE models: a new approach to prove convergence results. *IMA Journal of Applied Mathematics*, 78(5):945–964, 2013.

[51] P. Van Mieghem and R. van de Bovenkamp. Non-Markovian Infection Spread Dramatically Alters the Susceptible-Infected-Susceptible Epidemic Threshold in Networks. *Phys. Rev. Lett.*, 110:108701, Mar 2013.

[52] D. Guo, S. Trajanovski, R. van de Bovenkamp, H. Wang, and P. Van Mieghem. Epidemic threshold and topological structure of susceptible-infectious-susceptible epidemics in adaptive networks. *Phys. Rev. E*, 88:042802, Oct 2013.

[53] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem. Generalized Epidemic Mean-Field Model for Spreading Processes Over Multilayer Complex Networks. *IEEE/ACM Transactions on Networking*, 21(5):1609–1620, 2013.

[54] R. M. Anderson and R. M. May. *Infectious diseases of humans dynamics and Control*. Oxford Univ. Press, 1992.

[55] W. Cota and S. C. Ferreira. Optimized Gillespie algorithms for the simulation of Markovian epidemic processes on large and heterogeneous networks. *Computer Physics Communications*, 219:303–312, 2017.

[56] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14, Feb 2009.

[57] A. Economou, A. Gómez-Corral, and M. López-García. A stochastic SIS epidemic model with heterogeneous contacts. *Physica A: Statistical Mechanics and its Applications*, 421:78–97, 2015.

[58] M. López-García. Stochastic descriptors in an SIR epidemic model for heterogeneous individuals in small networks. *Mathematical Biosciences*, 271:42–61, 2016.

[59] W. Merbis. Exact epidemic models from a tensor product formulation, 2021. arXiv.

[60] W. Merbis and I. Lodato. Logistic growth on networks: Exact solutions for the susceptible-infected model. *Phys. Rev. E*, 105:044303, Apr 2022.

[61] S. Dolgov and D. Savostyanov. Tensor product approach to modelling epidemics on networks. *Applied Mathematics and Computation*, 460:128290, jan 2024.

[62] N. T. J. Bailey. A Simple Stochastic Epidemic. *Biometrika*, 37(3/4):193–202, Dec 1950.

[63] A. P. Millán, E. C. W. van Straaten, C. J. Stam, I. A. Nissen, S. Idema, J. C. Baayen, P. Van Mieghem, and A. Hillebrand. Epidemic models characterize seizure propagation and the effects of epilepsy surgery in individualized brain networks based on MEG and invasive EEG recordings. *Scientific Reports*, 12, 2022.

[64] D. Juher, J. Ripoll, and J. Saldaña. Analysis and Monte Carlo simulations of a model for the spread of infectious diseases in heterogeneous metapopulations. *Phys. Rev. E*, 80:041920, Oct 2009.

[65] M. Stoer and F. Wagner. A Simple Min-Cut Algorithm. *J. ACM*, 44(4):585–591, Jul 1997.

[66] J. A. Fill. The Passage Time Distribution for a Birth-and-Death Chain: Strong Stationary Duality Gives a First Stochastic Proof. *Journal of Theoretical Probability*, 22, Jun 2009.

[67] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61, 1960.

[68] E. Cator and P. Van Mieghem. Nodal infection in Markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated. *Phys. Rev. E*, 89:052802, May 2014.

[69] P. Van Mieghem. Origin of the fractional derivative and fractional non-Markovian continuous-time processes. *Phys. Rev. Research*, 4:023242, Jun 2022.

[70] M. G. Mittag-Leffler. Sur la nouvelle function $E_\alpha(x)$. *Comptes Rendus hebdomadaires de Séances de l'Academié des Sciences, Paris*, 137:554–558, 1903.

[71] D. Centola and M. Macy. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702–734, 2007.

[72] I. Iacopini, G. Petri, A. Barrat, and V. Latora. Simplicial models of social contagion. *Nat. Commun.*, 10:2485, Jun 2019.

[73] M. Alutto, L. Cianfanelli, G. Como, and F. Fagnani. Multiple peaks in network SIR epidemic models. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 5614–5619, 2022.

[74] C. Li, R. van de Bovenkamp, and P. Van Mieghem. Susceptible-infected-susceptible model: A comparison of N-intertwined and heterogeneous mean-field approximations. *Phys. Rev. E*, 86:026116, Aug 2012.

[75] M. A. Achterberg, J. L. A. Dubbeldam, C. J. Stam, and P. Van Mieghem. Classification of link-breaking and link-creation updating rules in susceptible-infected-susceptible epidemics on adaptive networks. *Phys. Rev. E*, 101:052302, May 2020.

[76] T. Gross and B. Blasius. Adaptive coevolutionary networks: a review. *Journal of The Royal Society Interface*, 5(20):259–271, 2008.

[77] I. B. Schwartz and L. B. Shaw. Rewiring for adaptation. *Physics*, 3, 2010.

[78] F. G. Hillary and J. H. Grafman. Injured Brains and Adaptive Networks: The Benefits and Costs of Hyperconnectivity. *Trends in Cognitive Sciences*, 21(5):385 – 401, 2017.

[79] I. J. Benczik, S. Z. Benczik, B. Schmittmann, and R. K. P. Zia. Opinion dynamics on an adaptive random network. *Phys. Rev. E*, 79:046104, Apr 2009.

[80] F. D. Sahneh, A. Vajdi, J. Melander, and C. M. Scoglio. Contact adaption during epidemics: A multilayer network formulation approach. *IEEE Transactions on Network Science and Engineering*, 6(1):16–30, Jan 2019.

[81] T. Gross, C. J. D. D'Lima, and B. Blasius. Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.*, 96:208701, May 2006.

[82] Á. Bodó and P.L. Simon. Analytic Study of Bifurcations of the Pairwise Model for SIS Epidemic Propagation on an Adaptive Network. *Differ. Equ. Dyn. Syst.*, 28:807–826, 2020.

[83] X. Zhang, C. Shan, Z. Jin, and H. Zhu. Complex dynamics of epidemic models on adaptive networks. *Journal of Differential Equations*, 266(1):803 – 832, 2019.

[84] J. Hindes, I. B. Schwartz, and L. B. Shaw. Enhancement of large fluctuations to extinction in adaptive networks. *Phys. Rev. E*, 97:012308, Jan 2018.

[85] V. Marceau, P. Noël, L. Hébert-Dufresne, A. Allard, and L. J. Dubé. Adaptive networks: Coevolution of disease and topology. *Phys. Rev. E*, 82:036116, Sep 2010.

[86] S. Risau-Gusman and D. H. Zanette. Contact switching as a control strategy for epidemic outbreaks. *Journal of Theoretical Biology*, 257(1):52 – 60, 2009.

[87] T. Britton, D. Juher, and J. Saldaña. A Network Epidemic Model with Preventive Rewiring: Comparative Analysis of the Initial Phase. *Bull Math Biol*, 78(2):2427–2454, 2016.

[88] N. Sherborne, K. B. Blyuss, and I. Z. Kiss. Bursting endemic bubbles in an adaptive network. *Phys. Rev. E*, 97:042306, Apr 2018.

[89] C. Liu, N. Zhou, X. Zhan, G. Sun, and Z. Zhang. Markov-based solution for information diffusion on adaptive social networks. *Applied Mathematics and Computation*, 380:125286, 2020.

[90] S. Van Segbroeck, F. C. Santos, and J. M. Pacheco. Adaptive Contact Networks Change Effective Disease Infectiousness and Dynamics. *PLOS Computational Biology*, 6(8):1–10, 08 2010.

[91] F. Ball, T. Britton, K.Y. Leung, and D. Sirl. A stochastic SIR network epidemic model with preventive dropping of edges. *J. Math. Biol.*, 78:1875–1951, Mar 2019.

[92] C. Lagorio, M. Dickison, F. Vazquez, L. A. Braunstein, P. A. Macri, M. V. Migueles, S. Havlin, and H. E. Stanley. Quarantine-generated phase transition in epidemic spreading. *Phys. Rev. E*, 83:026102, Feb 2011.

[93] L. B. Shaw and I. B. Schwartz. Fluctuating epidemics on adaptive networks. *Phys. Rev. E*, 77:066101, Jun 2008.

[94] G. Demirel, E. Barter, and T. Gross. Dynamics of epidemic diseases on a growing adaptive network. *Sci. Rep.*, 7:42352, 2017.

[95] J. Zhou, G. Xiao, S. A. Cheong, X. Fu, L. Wong, S. Ma, and T. H. Cheng. Epidemic reemergence in adaptive complex networks. *Phys. Rev. E*, 85:036107, Mar 2012.

[96] F. D. Sahneh, A. Vajdi, J. Melander, and C. M. Scoglio. Contact Adaption During Epidemics: A Multilayer Network Formulation Approach. *IEEE Transactions on Network Science and Engineering*, 6(1):16–30, 2019.

[97] S. Trajanovski, D. Guo, and P. Van Mieghem. From epidemics to information propagation: Striking differences in structurally similar adaptive network models. *Phys. Rev. E*, 92:030801, Sep 2015.

[98] I. Z. Kiss, L. Berthouze, T. J. Taylor, and P. L. Simon. Modelling approaches for simple dynamic networks and applications to disease transmission models. *Proc. R. Soc. A*, 468:1332–1355, 2012.

[99] L. Zhang, C. Guo, and M. Feng. Effect of local and global information on the dynamical interplay between awareness and epidemic transmission in multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(8):083138, 2022.

[100] P. Van Mieghem. Approximate formula and bounds for the time-varying susceptible-infected-susceptible prevalence in networks. *Phys. Rev. E*, 93:052312, May 2016.

[101] M. Ogura and V. M. Preciado. Epidemic processes over adaptive state-dependent networks. *Phys. Rev. E*, 93:062316, Jun 2016.

[102] M. Fauth and C. Tetzlaff. Opposing Effects of Neuronal Activity on Structural Plasticity. *Frontiers in Neuroanatomy*, 10, 2016.

[103] M. A. Achterberg and P. Van Mieghem. Moment closure approximations of susceptible-infected-susceptible epidemics on adaptive networks. *Phys. Rev. E*, 106:014308, Jul 2022.

[104] W. Wang, M. Tang, H. E. Stanley, and L. A. Braunstein. Unification of theoretical approaches for epidemic spreading on complex networks. *Rep. Prog. Phys.*, 80:036603, 2017.

[105] M. J. Keeling and K. T. D. Eames. Networks and epidemic models. *Journal of The Royal Society Interface*, 2(4):295–307, 2005.

[106] M. Taylor, P. L. Simon, D. M. Green, T. House, and I. Z. Kiss. From Markovian to pairwise epidemic models and the performance of moment closure approximations. *J. Math. Biol.*, 64:1021–1042, 2012.

[107] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha. Accuracy of mean-field theory for dynamics on real-world networks. *Phys. Rev. E*, 85:026106, Feb 2012.

[108] A. Szabó, P. L. Simon, and I. Z. Kiss. Detailed study of bifurcations in an epidemic model on a dynamic network. *Diff. Eq. and Appl.*, 4(2):277–296, 2012.

[109] A. Szabó-Solticzky, L. Berthouze, I. Z. Kiss, and P. L. Simon. Oscillating epidemics in a dynamic network model: stochastic and mean-field analysis. *J. Math. Bio.*, 72:1153–1176, 2016.

[110] M. A. Achterberg and P. Van Mieghem. An exact reduction of the continuous-time SIS process on adaptive networks is infeasible for most graphs. 2022. unpublished.

[111] S. Bonaccorsi, S. Ottaviano, D. Mugnolo, and F. De Pellegrini. Epidemic Outbreaks in Networks with Equitable or Almost-Equitable Partitions. *SIAM Journal on Applied Mathematics*, 75(6):2421–2443, 2015.

[112] C. Kuehn. *Moment Closure - A Brief Review*, chapter 13, pages 253–271. Springer, Cham, 2016.

[113] P. Van Mieghem. Universality of the SIS prevalence in networks, 2016. ArXiv preprint: https://arxiv.org/abs/1612.01386.

[114] F. Ball and T. Britton. Epidemics on networks with preventive rewiring. *Random Structures & Algorithms*, 61(2):250–297, 2022.

[115] J. A. P. Heesterbeek and J. A. J. Metz. The saturating contact rate in marriage- and epidemic models. *J. Math. Bio.*, 31:529–539, 1993.

[116] K. Y. Keung and O. Diekmann. Dangerous connections: on binding site models of infectious disease dynamics. *J. Math. Bio.*, 74:619–671, 2017.

[117] M. A. Achterberg and M. Sensi. A minimal model for adaptive SIS epidemics. *Nonlinear Dynamics*, 111:12657–12670, 2023.

[118] S. Jolad, W. Liu, B. Schmittmann, and R. K. P. Zia. Epidemic Spreading on Preferred Degree Adaptive Networks. *PLOS ONE*, 7(11):1–11, 11 2012.

[119] F. Brauer. A simple model for behaviour change in epidemics. *BMC Public Health*, 11(1):1–5, 2011.

[120] C. S. Holling. The Components of Predation as Revealed by a Study of Small-Mammal Predation of the European Pine Sawfly. *The Canadian Entomologist*, 91(5):293—320, 1959.

[121] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28(4):365–382, 1990.

[122] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1):29 – 48, 2002.

[123] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts. The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society interface*, 7(47):873–885, 2010.

[124] I. Bendixson. Sur les courbes définies par des équations différentielles. *Acta Mathematica*, 24:1–88, 1901.

[125] Y. Li and J. S. Muldowney. On Bendixson's Criterion. *Journal of Differential Equations*, 106(1):27–39, 1993.

[126] P. F. Verhulst. Recherches mathématiques sur la loi d'accroissement de la population. pages 1–45, 1845.

[127] H. K. Khalil. *Nonlinear Control*. Pearson Education, Essex, England, global edition, 2015.

[128] B. Prasse, M. A. Achterberg, L. Ma, and P. Van Mieghem. Network-inference-based prediction of the COVID-19 epidemic outbreak in the Chinese province Hubei. *Applied Network Science*, 5(35), 2020.

[129] M. A. Achterberg, B. Prasse, L. Ma, S. Trajanovski, M. Kitsak, and P. Van Mieghem. Comparing the accuracy of several network-based COVID-19 prediction algorithms. *International Journal of Forecasting*, 38(2):489–504, 2022.

[130] K. R. Moran, G. Fairchild, N. Generous, K. Hickmann, D. Osthus, R. Priedhorsky, J. Hyman, and S. Y. Del Valle. Epidemic Forecasting is Messier Than Weather Forecasting: The Role of Human Behavior and Internet Data Streams in Epidemic Forecast. *The Journal of Infectious Diseases*, 214(suppl4):S404–S408, 11 2016.

[131] B. Prasse, M. A. Achterberg, and P. Van Mieghem. Accuracy of predicting epidemic outbreaks. *Phys. Rev. E*, 105:014302, Jan 2022.

[132] K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J.M. Hyman, P. Yan, and G. Chowell. Short-term Forecasts of the COVID-19 Epidemic in Guangdong and Zhejiang, China: February 13–23, 2020. *J. Clin. Med.*, 9:596, 2020.

[133] A.V. Hill. Proceedings of the physiological society: January 22, 1910. *The Journal of Physiology*, 40(suppl):i–vii, 1910.

[134] B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115:513–583, 1825.

[135] Q. Yang, C. Yi, A. Vajdi, L. W. Cohnstaedt, H. Wu, X. Guo, and C. M. Scoglio. Short-term forecasts and long-term mitigation evaluations for the COVID-19 epidemic in Hubei Province, China. *Infectious Disease Modelling*, 5:563–574, 2020.

[136] L. Lorch, W. Trouleau, S. Tsirtsis, A. Szanto, B. Schölkopf, and M. Gomez-Rodriguez. A Spatiotemporal Epidemic Model to Quantify the Effects of Contact Tracing, Testing, and Containment, 2020. arXiv preprint.

[137] S. Y. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, and J. Leskovec. Mobility network modeling of COVID-19 explain inequities and informs reopening. *Nature*, 589:82–87, 2021.

[138] M.A.A. Al-qaness, A.A. Ewees, H. Fan, and M. Abd El Aziz. Optimization Method for Forecasting Confirmed Cases of COVID-19 in China. *J. Clin. Med.*, 9:674, 2020.

[139] Z. Yang, Z. Zeng, K. Wang, S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, Z. Mai, J. Liang, X. Liu, S. Li, Y. Li, F. Ye, W. Guan, Y. Yang, F. Li, S. Luo, Y. Xie, B. Liu, Z. Wang, S. Zhang, Y. Wang, N. Zhong, and J. He. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease*, 12(3), 2020.

[140] A. Kergassner, C. Burkhardt, D. Lippold, S. Nistler, M. Kergassner, P. Steinmann, D. Budday, and S. Budday. Meso-scale modeling of COVID-19 spatio-temporal outbreak dynamics in Germany. *medRxiv*, 2020.

[141] S. He, Y. Peng, and K. Sun. SEIR modeling of the COVID-19 and its dynamics. *Nonlinear Dynamics*, 101:1667–1680, 2020.

[142] Michael Day. Covid-19: four fifths of cases are asymptomatic, China figures indicate. *BMJ*, 369, 2020.

[143] M. Youssef and C. Scoglio. An individual-based approach to SIR epidemics in contact networks. *Journal of Theoretical Biology*, 283(1):136 – 144, 2011.

[144] B. Prasse and P. Van Mieghem. Network Reconstruction and Prediction of Epidemic Outbreaks for General Group-Based Compartmental Epidemic Models. *IEEE Transactions on Network Science and Engineering*, 7(4):2755–2764, 2020.

[145] T. P. Peixoto. Network Reconstruction and Community Detection from Dynamics. *Physical Review Letters*, 123:128301, Sep 2019.

[146] L. Ma, Q. Liu, and P. Van Mieghem. Inferring network properties based on the epidemic prevalence. *Applied Network Science*, 4(1):93, 2019.

[147] F. Di Lauro, J.-C. Croix, M. Dashti, L. Berthouze, and I. Z. Kiss. Network inference from population-level observation of epidemics. *Scientific Reports*, 10:18779, 2020.

[148] R. Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

[149] B. Prasse and P. Van Mieghem. Predicting network dynamics without requiring the knowledge of the interaction graph. *Proceedings of the National Academy of Sciences*, 119(44):e2205517119, 2022.

[150] B. Prasse and P. Van Mieghem. Time-dependent solution of the NIMFA equations around the epidemic threshold. *J. Math. Bio.*, 81:1299–1355, 2020.

[151] J. L. Elman. Finding Structure in Time. *Cognitive Science*, 14(2):179–211, 1990.

[152] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

[153] T. Young, D. Hazarika, S. Poria, and E. Cambria. Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Computational Intelligence Magazine*, 13:55–75, 2018.

[154] S. Hochreiter and J. Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, 1997.

[155] F. A. Gers and J. Schmidhuber. LSTM recurrent networks learn simple context-free and context-sensitive languages. *IEEE Transactions on Neural Networks*, 12 6:1333–40, 2001.

[156] F. A. Gers, J. Schmidhuber, and F. Cummins. Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, 12(10):2451–2471, 2000.

[157] R. Jozefowicz, W. Zaremba, and I. Sutskever. An Empirical Exploration of Recurrent Network Architectures. In Francis Bach and David Blei, editors, *In Proc. of ICML (32nd International Conference on Machine Learning)*, volume 37 of *PMLR*, pages 2342–2350, Lille, France, 07–09 Jul 2015.

[158] Y. Yu, X. Si, C. Hu, and J. Zhang. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*, 31(7):1235–1270, 2019. PMID: 31113301.

[159] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *In Proc of ICLR (International Conference for Learning Representations), San Diego, 2015*, abs/1412.6980, 2015.

[160] R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006.

[161] News from the Health Commission of Hubei, 2020. Retrieved on February 16, 2020 from http://wjw.hubei.gov.cn/fbjd/dtyw.

[162] Baidu Migration website, 2020. Retrieved on February 16, 2020 from https://qianxi.baidu.com/2020/.

[163] B. F. Maier and D. Brockmann. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*, 368(6492):742–746, 2020.

[164] S. Makridakis, E. Spiliotis, and V. Assimakopoulos. The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1):54 – 74, 2020. M4 Competition.

[165] RIVM. Actuele informatie over het nieuwe coronavirus (COVID-19), 2020. Retrieved on May 25, 2020 from https://www.rivm.nl/coronavirus-covid-19/actueel.

[166] CBS. Banen van werknemers naar woon- en werkregio, 2018. Retrieved on May 29, 2020 from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83628NED/table?ts=1583844319444.

[167] Google LLC. Google COVID-19 Community Mobility Reports, 2020. Retrieved on May 25, 2020 from https://www.google.com/covid19/mobility/.

[168] O. Diekmann, H. G. Othmer, R. Planqué, and M. C. J. Bootsma. The discrete-time Kermack–McKendrick model: A versatile and computationally attractive framework for modeling epidemics. *Proceedings of the National Academy of Sciences*, 118(39):e2106332118, 2021.

[169] N. Hernandez-Ceron, Z. Feng, and C. Castillo-Chavez. Discrete Epidemic Models with Arbitrary Stage Distributions and Applications to Disease Control. *Bull. Math. Bio.*, 75:1716–1746, 2013.

[170] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.

[171] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, volume 55. Dover Publications, nineth edition, 1965.

[172] D. J. Griffiths and D. F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, third edition, 2018.

[173] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, second edition, 2011.

[174] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, United Kingdom, second edition, 2012.

[175] P. E. Paré, J. Liu, C. L. Beck, B. E. Kirwan, and T. Başar. Analysis, Estimation, and Validation of Discrete-Time Epidemic Processes. *IEEE Transactions on Control Systems Technology*, 28(1):79–93, 2020.

[176] Tuanzhong Li and Xiaohong Xu. *Hubei Statistical yearbook*. China Statistics Press, 2016.

[177] T. Hastie, R. Tibshirani, and M. Wainwright. *Statistical learning with sparsity: the lasso and generalizations*. CRC press, 2015.

[178] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[179] M. Kiskowski and G. Chowell. Modeling household and community transmission of Ebola virus disease: Epidemic growth, spatial dynamics and insights for epidemic control. *Virulence*, 7(2):163–173, 2016. PMID: 26399855.

[180] C. P. Winsor. The Gompertz Curve as a Growth Curve. *Proceedings of the National Academy of Sciences*, 18(1):1–8, 1932.

ACKNOWLEDGEMENTS

My journey in the wonderful field of network epidemiology ends with finishing this thesis, after four years of research. Starting my PhD in November 2019, I could not have imagined how the upcoming coronavirus would change my research completely. But it was real: my research suddenly transformed from an abstract notion in a textbook to a vivid, painful reality. With utmost interest, I followed the discussions, talk shows and research papers about the pandemic. While most people thought the virus would disappear within one year (including myself), it lasted for over two years. Being a theoretical scientist with a preference for fundamental aspects of models and equations, COVID-19 proved to be tough. Tough in the sense that modelling real epidemics involves thousands of uncertainties and unknowns, which contrasts the more "hard" physics, like particle physics and dynamical systems, which can be exactly described by physical laws and mathematical formulas. My thesis is therefore a compromise – partially focussing on modelling the spread of COVID-19 and partially on theoretical work. The theoretical work is the largest portion, which is partially due to the fact that I already performed part of the research during my master thesis project.

Writing this thesis and completing my four years of research at TU Delft was impossible without the support and encouragement of many friends, family members and colleagues. First, I like to thank my promotor Piet Van Mieghem, for his effortless reading of my half-baked manuscripts, our highly irregular, but always useful meetings, and his enormous passion for science, both in vision and in solving problems at hand. At least half of this thesis was originally his idea, especially Chapter 3 and Chapter 4. Piet, hartelijk bedankt voor alle support de afgelopen jaren!

My work would have never been so amazing without my other collaborators. I'm grateful that Prof. Rob Kooij shared many ideas and puzzles with me regarding the effective graph resistance and submodularity (whose results are not included in this thesis). I would also like to thank Bastian Prasse for his willingness to read my work which greatly improved my writing skills and for giving a lot of ideas and inspiration. I truly enjoyed working together with Mattia Sensi as well, both in terms of the scientific content as well as on a personal level. Long, Stojan, Maksim, Brian and Liufei, thank you for your dedication and team work on the modelling of the spread of COVID-19. It has been a rough period, but our collaboration turned out to be very nice and fruitful. Finally, I would like to thank Johan Dubbeldam for introducing me to Piet. Without this first rendez-vous, I would have remained disconnected from the NAS network for the rest of my life.

During my work, I have had the pleasure to (co-)supervise several master students. Xinhan, Tian, Brian and Liufei, it was an amazing experience for me to guide you from the start of your thesis to your final presentation. Your input, enthusiasm and ideas inspired and motivated me, both in research and personal life.

Now I would like to take a moment to thank my room mates at work: Zhihao, Yinghue, Maria, Bastian and Long, it has been a great pleasure sharing a room, playing badminton,

having discussions on politics, environment, university issues, cultural similarities and differences, learning Dutch, and many other things. Our office will always be in my mind as a nice welcoming place like home.

I also owe my debt to the other colleagues in the NAS department: Edgar, Remco, Eric, Rogier, Hannah, Scott, Sergey, Yuxuan, Zhidong, Peng, Ivan, Gaby, Fenghua, Brian, Robin, Matteo and Lisa. It has been a great pleasure to be part of the NAS family, which would not have been as great without your presence.

Vervolgens wil ik mijn ouders, broer en zus bedanken voor hun jarenlange support. Hun interesse en ondersteuning hebben mijn tijd op de universiteit tot een prachtige tijd gemaakt. Verder bedank ik ook mijn huisgenoten Mike, Marco en Floris, die de corona-periode tot een aangename tijd hebben gemaakt. En boven al deze mensen bedank ik mijn liefste Margreth voor haar aanwezigheid in mijn leven.

Tenslotte, dit proefschrift zou niet tot stand kunnen zijn gekomen zonder de ware God, die mij al die jaren ondersteund heeft. Daarom sluit ik af met het volgende Bijbelvers:

Psalm 16 : 7–8

Ik loof de HEERE, die mij raad gegeven heeft,
zelfs 's nachts onderwijzen mij mijn nieren.
Ik stel mij de HEERE voortdurend voor ogen,
omdat Hij aan mijn rechterhand is, wankel ik niet.

I bless the Lord who gives me counsel;
in the night also my heart instructs me.
I have set the Lord always before me;
because He is at my right hand, I shall not be shaken.

CURRICULUM VITÆ

Massimo Alessandro ACHTERBERG

22-06-1996 Born in Rotterdam, the Netherlands

EDUCATION

2014–2017 Double Bachelor Applied Mathematics & Applied Physics
TU Delft, The Netherlands

2017–2019 Master Applied Mathematics, *cum laude*
Specialisation: Mathematical Physics
TU Delft, The Netherlands

2019–2024 PhD in Network Epidemiology
Network Architectures and Services group
TU Delft, The Netherlands
Thesis: Epidemics on Static and Adaptive Networks
Promotor: Prof. dr. ir. P. F. A. Van Mieghem

WORK EXPERIENCE

2015–2019 Student Assistant for assignment and exam assessment
and teaching in small groups
TU Delft, The Netherlands

Sep 2018
– Dec 2018 Internship at Deltares, institute for water research
Research on algorithms of water level predictions in the North Sea
Supervisor: Prof. dr. ir. M. Verlaan

AWARDS

2022 Best contributed talk at the SIAM Virtual Workshop on Network Science

List of Publications

Journal publications

8. R. E. Kooij and **M. A. Achterberg**, *Minimizing the effective graph resistance by adding links is NP-hard*, Operational Research Letters 51:6, 601–604, Nov 2023.

7. **M. A. Achterberg** and M. Sensi, *A minimal model for adaptive SIS epidemics*, Nonlinear Dynamics 111, 12657–12670, May 2023.

6. **M. A. Achterberg** and P. Van Mieghem, *Moment closure approximations of susceptible-infected-susceptible epidemics on adaptive networks*, Physical Review E 106, 014308, Jul 2022.

5. **M. A. Achterberg**, B. Prasse and P. Van Mieghem, *Analysis of continuous-time Markovian ε-SIS epidemics on networks*, Physical Review E 105, 054305, May 2022.

4. B. Prasse, **M. A. Achterberg** and P. Van Mieghem, *Accuracy of predicting epidemic outbreaks*, Physical Review E 105, 014302, Jan 2022.

3. **M. A. Achterberg**, B. Prasse, L. Ma, S. Trajanovski, M. Kitsak and P. Van Mieghem, *Comparing the accuracy of several network-based COVID-19 prediction algorithms*, International Journal of Forecasting 38:2, Apr–Jun 2022.

2. B. Prasse, **M. A. Achterberg**, L. Ma and P. Van Mieghem, *Network-inference-based prediction of the COVID-19 epidemic outbreak in the Chinese province Hubei*, Applied Network Science 5:35, Jul 2020.

1. **M. A. Achterberg**, J. L. A. Dubbeldam, C. J. Stam and P. Van Mieghem, *Classification of link-breaking and link-creation updating rules in susceptible-infected-susceptible epidemics on adaptive networks*, Physical Review E 101, 052302, May 2020.

Conference publications

2. B. Chang, L. Yang, M. Sensi, **M. A. Achterberg**, F. Wang, M. Rinaldi and P. Van Mieghem, *Markov Modulated Process to model human mobility*, Proceedings of the Tenth International Conference on Complex Networks and Their Applications, Complex Networks 2021, 30 November – 2 December 2021, Madrid, Spain.

1. X. Liu, **M. A. Achterberg** and R. E. Kooij, *Non-Consensus Opinion Model with Byzantine Nodes*, Proceedings of the Twelfth International Workshop on Resilient Networks Design and Modeling, 19 – 21 September 2022, Compiègne, France.

Preprints / submitted works

3. **M. A. Achterberg** and P. Van Mieghem, *Analytic solution of Markovian epidemics without re-infections on heterogeneous networks*, Oct 2023.

2. X. Liu, **M. A. Achterberg** and R. E. Kooij, *Non-Consensus Opinion Models with Malicious Nodes*, SSRN preprint, Apr 2023.

1. **M. A. Achterberg** and R. E. Kooij, *On the non-submodularity of the problem of adding links to minimize the effective graph resistance*, Mar 2023.

INDEX