Stochastic epidemics on random networks

Abid Ali Lashari
Stochastic epidemics on random networks

Abid Ali Lashari

Academic dissertation for the Degree of Doctor of Philosophy in Mathematical Statistics at Stockholm University to be publicly defended on Thursday 16 May 2019 at 10.00 in sal 14, hus 5, Kräftriket, Roslagsvägen 101.

Abstract
This thesis considers stochastic epidemic models for the spread of epidemics in structured populations. The asymptotic behaviour of the models is analysed by using branching process approximations. The thesis contains four manuscripts.

Paper I is concerned with the study of the spread of sexually transmitted infections, or any other infectious diseases on a dynamic network. The model we investigate is about the spread of an SI (Susceptible → Infectious) type infectious disease in a population where partnerships are dynamic. We derive explicit formulas for the probability of extinction and the threshold parameter $R_0$ using two branching process approximations for the model. In the first approximation some dependencies between infected individuals are ignored while the second branching process approximation is asymptotically exact and only defined if every individual in the population can have at most one partner at a time. By comparing the two approximations, we show that ignoring subtle dependencies in the dynamic epidemic model leads to wrong prediction of the probability of a large outbreak.

In paper II, we study a stochastic SIR (Susceptible → Infectious → Removed) epidemic model for the spread of an epidemic in populations structured through configuration model random graphs. We study the asymptotic (properly scaled) time until the end of an epidemic. This paper heavily relies on the theory of branching processes in continuous time.

In paper III, the effect of vaccination strategies on the duration of an epidemic in a large population is investigated. We consider three vaccination strategies: uniform vaccination, leaky vaccination and acquaintance vaccination.

In paper IV, we present a stochastic model for two successive SIR epidemics in the same network structured population. Individuals infected during the first epidemic might have (partial) immunity for the second one. The first epidemic is analysed through a bond percolation model, while the second epidemic is approximated by a three-type branching process in which the types of individuals depend on their status in the percolation clusters used for the analysis of the first epidemic. This branching process approximation enables us to calculate a threshold parameter and the probability of a large outbreak for the second epidemic. We use two special cases of acquired immunity for further evaluation.

Keywords: Branching process, Configuration model, Random graph, Epidemic process, Final size, Threshold behaviour, Duration of an epidemic, Vaccination.

Stockholm 2019
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-167373


Department of Mathematics
Stockholm University, 106 91 Stockholm
STOCHASTIC EPIDEMICS ON RANDOM NETWORKS
Abid Ali Lashari
Stochastic epidemics on random networks

Abid Ali Lashari
To my Parents, Saba and Musa.
List of Papers

The following papers, referred to in the text by their Roman numerals, are included in this thesis.


PAPER II: Abid Ali Lashari, Ana Serafimović, Pieter Trapman, The duration of an *SIR* epidemic on a configuration model. (Under review)

PAPER III: Abid Ali Lashari, Pieter Trapman, Effect of vaccination on the duration of an *SIR* epidemic in homogeneously mixing and structured populations. (Manuscript)

PAPER IV: Abid Ali Lashari, Frank Ball, David Sirl, Pieter Trapman, Modeling the spread of two successive *SIR* epidemics on a configuration model network. (Manuscript)

Paper I is included with permission from the publisher.
Author’s contribution

PAPER I: “Branching process approach for epidemics in dynamic partnership network.”

This paper is based on an idea of Pieter Trapman. Abid Ali Lashari did the modelling together with Pieter Trapman and wrote most of the manuscript.

PAPER II: “The duration of an SIR epidemic on a configuration model.”

This paper is based on the master thesis of Ana Serafimović on a problem proposed by Pieter Trapman. Ana Serafimović developed the model and analysed the model together with Pieter Trapman. Abid Ali Lashari and Pieter Trapman extended the results and wrote the manuscript.

PAPER III: “Effect of vaccination on the duration of an SIR epidemic in homogeneously mixing and structured populations.”

In this paper, Abid Ali Lashari derived the analytical results and wrote the manuscript with input from Pieter Trapman.

PAPER IV: “Modeling the spread of two successive SIR epidemics on a configuration model network.”

In this paper, Abid Ali Lashari and Pieter Trapman developed the model and wrote the manuscript while David Sirl and Frank Ball contributed with valuable comments and discussions.
Acknowledgements

First, I would like to thank my supervisor Pieter Trapman for his guidance, encouragement and dedicating time to me whenever it was needed. He made me push past my limits to accomplish my goals. Thanks to Tom Britton for discussions and guidance on mathematical modeling of infectious diseases.

I would also like to thank all colleagues who have made the Division of Mathematical Statistics a great work place. I am grateful for the support of all my PhD fellows, especially Theresa for being a wonderful officemate, Sebastian for being nice and translating the summary of papers in Swedish and Stanislas for being my lunch buddy. I am eternally thankful to my family for their prayers and support throughout my studies. I am grateful for the support and Duas of all my friends, especially Waqas, Habib, Asif and Nabeel, to Mateen for being my Jigar.

I appreciate the financial support from the Swedish Research Council and the Department of Mathematics, Stockholm University.

Most importantly, I am extremely grateful to Almighty Allah for providing me with the ability, understanding and insight to carry out this work and who has been with me from the start till the end of this work (Alhamdulillah).

Last but not least, I would like to express my deep appreciation to my wife Saba who filled a giant hole in my life, for her efforts to make me a better person, for her loving and caring attitude and for giving birth to an adorable and wonderful boy Musa.
Contents

List of Papers i

Author’s contribution iii

Acknowledgements v

1 Introduction 1

2 Epidemic models 3
   2.1 The beginning of epidemic modeling . . . . . . . . . . . . . . 3
   2.2 Branching processes . . . . . . . . . . . . . . . . . . . . . . 5
      2.2.1 Generating functions and the extinction probability . . 5
      2.2.2 Approximating epidemics by branching processes . . . 6
   2.3 Vaccination strategies . . . . . . . . . . . . . . . . . . . . . . 7

3 Spread of epidemics on networks 9
   3.1 Random networks and epidemics . . . . . . . . . . . . . . . . 9
   3.2 The configuration model network . . . . . . . . . . . . . . . . 10
      3.2.1 The SIR epidemic on networks . . . . . . . . . . . . . . 10

4 Overview of the papers 13

5 Sammanfattning 17

References 21
1. Introduction

Infectious diseases such as smallpox, the Black Death and the Spanish Influenza which killed millions have shaped the history of mankind (Morens et al., 2008). In recent years, infectious diseases such as swine flu (influenza type A virus) and the Ebola virus have shown considerable impact and emerged as a serious public health problem to the society (Camacho et al., 2014; Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team, 2009; WHO Ebola Response Team, 2014). Despite of substantial advancements in medical and healthcare technology, infectious diseases such as measles, malaria, and HIV remain major threats to public health in many parts of the world (Lopez and Murray, 1998; Murray et al., 2014). Outbreaks of such epidemics remind us that the battle against infectious disease is not over and the epidemics are an actual problem for health authorities.

Indeed the spread of infectious diseases has declined in most parts of the world, but continuous appearance of outbreaks of new infectious diseases is still a significant problem in the modern world. In infectious disease epidemiology, the three major tasks for health authorities are to first understand the possible causes of the disease, then make predictions of the course of disease, and finally to develop methods for its prevention and control. The usual procedure for this is obtaining and analysing observed data because human experimentation is often not possible or ethical. In such cases, understanding, assessing and predicting the dynamics of the spread of infectious diseases is essential and for this mathematical modeling plays an important role (Heesterbeek et al., 2015).

There has been a significant interest in understanding the infectious disease epidemiology and use of mathematical models in the spread of epidemics (Chubb and Jacobsen, 2010). Mathematical models have played a valuable role in providing insight into most effective control measures such as vaccination strategies or the required infrastructure for the public health decision-making (see, Bacaër, 2011; Hollingsworth, 2009, for an overview of the history of mathematical modeling). However, numerous challenges are involved in developing a complete understanding of epidemics and for an overview of some of the recent important challenges in epidemic modeling we refer the reader to the special issue of epidemics by Lloyd-Smith et al., 2015 and in particular see, e.g. Britton et al., 2015; Eames et al., 2015; Pellis et al., 2015 (all papers
in that issue).

The specific goal of this thesis is to study the spread of epidemics on random networks by developing new stochastic epidemic models. We identify the problems that arise in the approaches used for computing important epidemiological quantities in some existing studies and provide new techniques for the computation of those important quantities. In particular, the main theoretical results provide answers to questions: what is the probability of a large outbreak (the epidemic becomes established and infects a nonnegligible proportion of the population)? What fraction of the population is infected or will be infected if a large outbreak is possible and occurs? What is the duration of a large outbreak? What is the effect of vaccination on the duration of an entire epidemic.

In the following chapters of the thesis, we provide an intuitive introduction to stochastic models for epidemics on random networks and the application of branching process approximations in analysing the dynamics of the infectious disease spread on those random networks. We give a brief introduction of some basic deterministic and stochastic epidemic models, basic epidemic quantities and the theory of branching processes. Moreover, a short introduction of the random graphs constructed through a configuration model network and its application to the modeling of epidemics on random networks is also provided. In the end, we give an overview/summary of the papers included in this thesis.
2. Epidemic models

In this chapter, first we very briefly give a review of early work in epidemic modeling and define the *SIR* (Susceptible-Infective-Removed) epidemic model. Next, we discuss the basic theory of branching processes and its application to epidemics and provide a brief introduction of some vaccination strategies in epidemic modeling.

2.1 The beginning of epidemic modeling

One of the earliest publications addressing the mathematical modeling of epidemics dates back to 1760 and is written by Daniel Bernoulli (see Bacaër, 2011; Hethcote, 2000). In this paper Daniel Bernoulli created a mathematical model to analyse the impact of vaccination on the spread of smallpox. The mathematical modeling of epidemics has developed systematically and gained popularity with the benchmark paper of Ross (1911) on discovering the mosquito based transmission of malaria and the contribution of Kermack and McKendrick (1927) (see Diekmann et al., 2012, 1995, and references therein). The pioneering work of Kermack and McKendrick established the deterministic compartmental epidemic modeling.

In the widely studied *SIR* epidemic models (which is a special case of the Kermack and McKendrick model), the population is split into three epidemiological classes:

- **Susceptible** individuals who have not yet been infected by the disease but are at risk of becoming infected.

- **Infectious** individuals who have been infected and can transmit the infection to other (susceptible) individuals during their *infectious period*. Infectious individuals stays infectious for some time (random/fixed) after which they eventually recover from the infection.

- **Recovered/Removed** individuals who have had the disease but now are no longer spreading the disease. Recovered individuals are unable to acquire or transmit the infection because they are no longer infectious and are fully immune (or dead).
At any given time (depending upon the disease stage) an individual belongs to one of the above epidemiological classes. The individual can change class by moving either from $S$ to $I$ or from $I$ to $R$. Eventually the epidemic ends when there are no more infectious individuals in the $I$-class. The SIR model is more realistic for diseases leading to lifelong immunity after infection (e.g. chickenpox, measles).

The evolution of the number of susceptible, infectious and recovered individuals over time in Kermack and McKendrick’s special case SIR model is usually studied through three differential equations (see, e.g. Diekmann et al., 2012). In this deterministic model the demography (births, deaths and migration) is ignored and the assumption is that the population size is effectively large and fixed. Another simplifying assumption is that individuals mix homogeneously. With this assumption the infectious contact pattern is ignored and it is assumed that an individual is equally likely to interact with any other individual in the population (regardless of geographic location, age, etc.). Although those simplifying assumptions seem unrealistic, the mathematical models with those assumptions have long been providing important insight and predictive capability. However, the deterministic model is not useful in predicting quantities such as the probability of a large epidemic outbreak. Furthermore, the deterministic models are not useful for studying the start and the end of an epidemic because those models do not handle small numbers of infecteds well (see, e.g. Andersson and Britton, 2012; Diekmann et al., 2012).

As a prime example of the stochastic version of the SIR epidemic model, we consider the Markov SIR epidemic model. In this model, the duration of the infectious period is stochastic and it is assumed to be exponentially distributed, and during the infectious period the infected individuals make close contacts at a constant rate. If the individuals who comes into contact with infectious individuals are susceptible they becomes infectious immediately.

For a more extensive survey of stochastic epidemics and literature reviews, the reader can consult Andersson and Britton, 2012 and Britton, 2010.

The fundamental parameter in both stochastic and deterministic epidemics that governs the epidemic dynamic is the basic reproduction number, usually denoted by $R_0$. It is usually defined as the expected number of secondary infections that a typical newly infected individual produce in an entirely susceptible population during the early phase of an epidemic. The critical value of $R_0 = 1$ is known as the epidemic threshold. This $R_0$ will also play an important role in this thesis.
2.2 Branching processes

In this section, we briefly give an overview of the basic theory of the branching processes that is used in the analysis of the infection spread on random networks in all the papers of this thesis. In the following we define the so-called single-type Galton-Watson process and the probability of extinction of this branching process.

Consider a population in which the number of individuals in the $n$th generation is $Z_n$. The difference between male and female is ignored. Each individual produces a (random) number of individuals independently of the other individuals and the numbers of children are identically distributed. The number of individuals in $(n+1)$th generation of the branching process is

$$Z_{n+1} = \sum_{i=1}^{Z_n} \zeta_i,$$

where the $\zeta_i$ are independent and identically distributed random variables with distribution

$$\mathbb{P}(\zeta_i = k) = p_k, \quad k = 0, 1, 2 \cdots.$$

So $\zeta_i$ is the number of offspring of the $i$th individual and $p_k$ is the probability that a single individual produces $k$ offspring.

The branching process is said to be extinct by generation $n$ if $Z_n = 0$. Trivially if $Z_n = 0$, then $Z_{n'} = 0$ for all $n' \geq n$. Let $\zeta_i \sim \zeta$ and $\mu := \mathbb{E}[\zeta]$ denote the expected number of offspring of a single individual. The branching process is said to be subcritical, critical or supercritical according to whether $\mu < 1$, $\mu = 1$ or $\mu > 1$, respectively. The subcritical and critical branching processes almost surely becomes extinct while the supercritical branching process has a positive probability of avoiding extinction (see Harris, 2002; Jagers, 1975). To show this we use generating functions as introduced below.

2.2.1 Generating functions and the extinction probability

We now define the probability generating functions (PGFs) describing the offspring of individuals in the branching process. For $s \in [0, 1]$, the probability generating function $f(s)$ for the branching process with offspring distribution given through $p_k$ for $k \in \mathbb{N}$ is defined by

$$f(s) = \mathbb{E}[s^\zeta] = \sum_{k=0}^{\infty} \mathbb{P}(\zeta = k)s^k = \sum_{k=0}^{\infty} p_k s^k.$$
The generating function of the number of individuals in the \( n \)th generation of the branching process conditioned on a single ancestor is

\[
f_n(s) = \sum_{k=0}^{\infty} \mathbb{P}(Z_n = k) s^k.
\]

Since \( f_1 = f \), and using the independence property of offspring distribution we have that

\[
f_n(s) = \mathbb{E}[s^{Z_n}] = \mathbb{E} \left[ \mathbb{E}[s^{Z_n} \mid Z_1] \right] = \mathbb{E} \left[ \mathbb{E}[s^{Z_n-1}] Z_1 \right] = f \left( f_{n-1}(s) \right).
\]

We now consider the probability of extinction of the branching process. Let \( q \) denote the extinction probability of the above mentioned Galton-Watson branching process initiated by one single ancestor. In order for the offspring of an individual to go extinct, all the offspring of its children should go extinct. Therefore,

\[
q = \sum_{k=0}^{\infty} \mathbb{P}(Z_1 = k) q^k = f(q)
\]

and we know from standard branching process theory that \( q \) is the smallest solution in \([0, 1]\) of \( s = f(s) \). Furthermore, for \( s \in (0, 1) \)

\[
\lim_{n \to \infty} f_n(s) = q.
\]

The above mentioned Galton-Watson process describes a family tree from a generation perspective in which a population only contains individuals of a single generation at a time. Generalizations of the above setup to multitype branching processes and branching processes in real time is well studied. For detailed descriptions of possible generalizations of branching processes with biological applications, we refer the reader to e.g. Haccou et al., 2005; Harris, 2002; Jagers, 1975.

### 2.2.2 Approximating epidemics by branching processes

It is well known in epidemic literature that the beginning of an SIR epidemic (when the fraction infected is still small) in a homogeneously mixing population can be analysed by a suitable branching process (see Andersson and Britton, 2012; Ball and Donnelly, 1995; Britton, 2010, and references therein). In those branching process approximations, infecting someone corresponds to giving birth while death of individuals corresponds to actual death or recovery from infection. Such branching processes have shown to be good approximations to the stochastic epidemic processes during the initial stages of the epidemics, when the initial number of infectives is small and the number of
susceptibles is still large (see e.g. Ball and Donnelly, 1995), and also during the final stages of the epidemics (Ball, 1983; Lashari et al., 2018). The initial stage of the epidemic is the most crucial one because during that stage it can be determined if there is a possibility for the epidemic to take off or the epidemic will die out. The basic reproduction number $R_0$ corresponds to the offspring mean of the particular branching process. In particular, $R_0$ determines whether the infection will die out with probability 1, or whether the epidemic is capable of causing a large epidemic outbreak.

In many scenarios one may be interested in the real-time development of the number of infected individuals in a population. For example, in epidemiology, the real-time growth rate during the initial phase of the epidemic is an important parameter with a clear biological implications (Trapman et al., 2016) as we will show in the second and third paper of this thesis. The epidemic/branching process growth/decay rate $\alpha$ corresponds to the so-called Malthusian parameter for the population growth/decay rate and is the solution of

$$1 = \int_0^\infty e^{-\alpha t} \mu(t) dt,$$

where $\mu(t)$ represent the expected rate at which infected individual make infectious contacts with other individuals $t$ time units after it was infected. Note that

$$R_0 := \int_0^\infty \mu(t) dt.$$

If $R_0 > 1$, the epidemic/branching process is said to be supercritical and $\alpha > 0$ exists while if $R_0 = 1$, the epidemic/branching process is critical and in this case $\alpha = 0$. If $R_0 < 1$, the epidemic/branching process is said to be subcritical and $\alpha$ might exist and if it does then $\alpha < 0$.

### 2.3 Vaccination strategies

An important application of modeling the infectious disease spread is to inform public health decision-makers to how to reduce the impact of the diseases. In order to prevent or reduce the impact of a disease, it is sometimes possible to vaccinate individuals prior to the introduction of the disease. In paper III, we consider the effect of three vaccination strategies (defined below) to control the spread of a disease.

- **Uniform vaccination**: in this strategy individuals are chosen uniformly at random from the population and then they are vaccinated.

- **Leaky vaccination**: a vaccine that reduces the susceptibility and infectivity of every vaccinated individual but does not eliminate the potential
for infection is called a leaky vaccine.

- *Acquaintance vaccination:* in this strategy some “acquaintances” of randomly selected individuals are vaccinated.

If the vaccine is *perfect* then the *critical vaccination coverage* required to prevent a large outbreak of an epidemic in a randomly mixing population is given by $1 - 1/R_0$. 
3. Spread of epidemics on networks

In this chapter we define the random graph models based on the configuration model and the spread of SIR epidemics on the network. Moreover, we briefly describe branching process approximations of an epidemic on random graphs.

3.1 Random networks and epidemics

In recent years, there has been a lot of research on modeling stochastic epidemics spreading through networks of individuals (Davis, 2017; Miller and Kiss, 2014). Diseases like HIV are clearly dependent on the contact structure of the infected individuals and indeed for all sexually transmitted disease the simplifying random mixing assumption is not valid.

The network is typically represented by a random graph (Newman, 2002). The graph can be viewed as a collection of vertices/nodes that represent individuals and edges represent connections/relationships of individuals (for example, social contacts with family members, work colleagues, and friends etc.). The degree of a node is the number of edges attached to that node. Two nodes are neighbors if they share a common edge.

The network used to analyse the spread of an infectious disease can either be static or dynamic. In paper I (Lashari and Trapman, 2018) the network we consider is dynamic while in all other papers of this thesis static network is considered. The epidemic is modeled in populations having a “Configuration Model” structure (defined in the next section). This special model structure is chosen since an epidemic on a general network is too difficult to analyse. This construction allows to compute important quantities of interest such as the reproduction number, the probability of a major epidemic outbreak, the final size of an epidemic and the duration of an epidemic on the random network. Below we formally introduce the random graph and the epidemic model. It is important to note that all the results are asymptotic as the population size $n \to \infty$. 

In this chapter we define the random graph models based on the configuration model and the spread of SIR epidemics on the network. Moreover, we briefly describe branching process approximations of an epidemic on random graphs.
3.2 The configuration model network

The models we investigate consider the spread of an infectious disease through a population of \( n \) individuals. The population is assigned a network structure according to the configuration model (see, e.g. Durrett, 2007, chap. 7 or Van der Hofstad, 2016, chap. 9). In order to create the graph describing possible relations/contacts, we assign a random number of so called half-edges (edges with just one endpoint attached to a vertex) to each individual/vertex in the population according to independent draws from a given distribution. Denote this distribution by \( D \), having mass function \( p_k = \mathbb{P}(D = k), k = 1, 2, 3 \ldots \). We assume that \( \mathbb{E}(D) < \infty \). The number of half-edges are independent and identically distributed according to \( D \). The half-edges are then paired uniformly at random which in turn form the edges of the random graph. In this network construction, every vertex has the right degree, although some problems such as self-loops or parallel edges might arise. The degree of a uniformly chosen individual is distributed according to \( D \). By the construction of the graph, the degrees of neighbors of the uniformly chosen individual, or neighbors of any given individual do not have degree distribution \( D \) but have the size biased degree distribution \( \tilde{D} \) given by

\[
\tilde{p}_k = \mathbb{P}(\tilde{D} = k) = \frac{k p_k}{\mathbb{E}(D)}, \quad k = 1, 2, 3 \ldots
\]

Usually we want graphs to be simple, which means that there are no parallel edges (several edges connecting the same two vertices) and self-loops (edge connecting a vertex to itself). The assumption \( \mathbb{E}(\tilde{D}) < \infty \) (or equivalently \( \mathbb{E}[\tilde{D}^2] < \infty \)) ensures that the probability of obtaining a configuration model with no parallel edges and self-loops is bounded away from 0 as \( n \to \infty \) (see, e.g. Van der Hofstad, 2016, p. 219). Furthermore, the graph generated through a configuration model is locally tree-like (as \( n \to \infty \) with probability tending to 1, the “neighborhood” of a given vertex contain no short circuits). The tree-like structure allows branching process approximations for the spread of epidemics on graphs (see, e.g. Andersson, 1998; Britton et al., 2007).

3.2.1 The SIR epidemic on networks

We consider the evolution of SIR epidemics on graphs. A vertex is said be susceptible, infectious, or recovered if the individual it represents is in this specific “infection state.” Infectious vertex stay so for a random period of time (the length of its infectious period) during which it makes infectious contacts with any given neighbor according to independent homogeneous Poisson processes. A contact made by an infectious individual with any other individual is
called an infectious contact. When the infectious period of an infectious individual ends it becomes removed and then plays no further role in the spread of an epidemic. A susceptible individual becomes immediately infectious (birth in the branching process) when it is contacted by an infected individual. The epidemic is assumed to starts at time \( t = 0 \) when all individuals are susceptible except for a single infectious individual chosen uniformly at random from the population. The infectious periods and the Poisson processes describing the contact processes are all independent. The epidemic stops when there is no infected individual left in the population.

The early stages of an epidemic spreading across the network is analysed by a (forward) branching process. This branching process approximates the number of infected individuals by taking into account who an individual will infect forward in time. The initial generation in this branching process corresponds to the single infectious individual chosen uniformly at random from the population. In general, the generation \( k \) consists of individuals who were previously uninfected but becomes infected by an infectious contact with individuals in generation \( k - 1 \). In the early stages of the epidemic as \( n \to \infty \) it is more likely that each of the contact made by an infective individual will be with uninfected individuals apart from contacts with the “infector.” That is why the early spread of an epidemic is well approximated by a (forward) branching process (see, e.g. Ball and Sirl, 2013; Ball et al., 2009, for rigorous treatments of the forward branching process). The threshold parameter that determines whether a major epidemic outbreak is possible is determined by whether or not the (forward) branching process avoid extinction. For convergence of the probability of a major outbreak to the survival probability of the forward branching process see, Ball and Sirl, 2013; Ball et al., 2009.

Similarly, the expected final size of a major outbreak (proportion of individuals who ultimately become recovered) corresponds to the survival probability of an approximating (backward) branching process. This branching process is a good approximation of a generation perspective on an individual’s susceptibility set which is a key tool for finding the fraction of the population that ultimately becomes infected by a major outbreak.

Let \( \mathcal{V} = \{v_1, v_2, \ldots, v_n\} \) denote the set of susceptible vertices of the undirected graph generated through the configuration model. We construct the directed graph on \( \mathcal{V} \), in which for any \( v_i \) and \( v_j \) a directed edge from \( v_i \) to \( v_j \) is present if and only if \( v_i \), if infected during the epidemic, would make contact with \( v_j \) during its infectious period. Then, the susceptibility set of \( v_j \) is the collection of all those vertices in the directed graph (including \( v_j \) itself) that can be reached from \( v_j \) by following a path backwards. Note that any given vertex \( v_i \) becomes ultimately infected in the epidemic if and only if its susceptibility set contains the initial infective (see, e.g. Paper II, Section 3 or Ball and
Neal, 2002 for details of susceptibility sets). The schematic illustration of the susceptibility set is shown in Figure 3.1.

Figure 3.1: The solid red nodes represent the susceptibility set of node $u$ in the directed random graph representation of an epidemic.

The susceptibility set is approximated by a (backward) branching process, which we now briefly describe. Let $v_i$ denote a uniformly at random chosen initial individual. In the (backward) branching process, we consider how many neighbors make infectious contact with that individual. The 0-th generation in this branching process consist of individual $v_i$. In general, the generation $k$ comprises individuals who enter the susceptibility set by making infectious contact with an individual in generation $k - 1$ (for details, see Ball and Sirl, 2013; Ball et al., 2009).

For an overview of the asymptotic results that relate expected final size of an epidemic outbreak to the survival probability of the (backward) branching process see, e.g. Ball and Neal, 2002, 2008; Ball et al., 2009.
4. Overview of the papers

The understanding of the structure of a network in the absence of infection is very important for modeling the spread of the infectious diseases on networks. There has been a considerable interest in epidemics on networks as a way to represent the complex structure of contacts that are capable of spreading infection in a population (see Ball and Neal, 2008; Durrett, 2007; Newman, 2002, and references therein). Leung et al. (2015) consider a (deterministic) model for the spread of an \(SI\) (Susceptible \(\rightarrow\) Infectious) type epidemic on a network of dynamic partnerships. The network is dynamic in the sense that individuals enter the population (by birth) and leave the population (by death) and the partnerships (edges in the network) are forming and dissolving while the epidemic spreads through the population. A key ingredient in the model of Leung et al. (2015) is the “mean field at distance one” assumption which is a (non-exact) approximation of the distribution of the number of partners of partners of a newly infected individual. In dynamic network model with demography, the degrees of partners are not independent (Leung et al., 2017) however Leung et al. (2015) ignored the degree dependence between partners. This approximation allowed them to write down a system of differential equations to obtain analytic expressions for the basic reproduction number \(R_0\).

In paper I, we extend the model of Leung et al. (2015) by considering epidemic and population dynamics as stochastic processes. We discuss the complications that come by ignoring the degree dependencies between the partners while modeling the spread of infectious disease on a dynamic network. We propose a new scheme of modeling demography with the assumption that the new individuals have a stationary distribution (degree distribution) for the number of partners. That is to say that the new individuals immediately form a (random) number of partnerships upon their arrival in the population with individuals already in the population. Under this assumption the dependencies between the number of partners of an individual are absent and the ‘mean-field at distance one’ of (Leung et al., 2015) is no longer required.

The early stages of a stochastic \(SI\) type epidemic is approximated by a suitable branching process as the network size tends to infinity. In particular, we derive the expressions for two of the most fundamental quantities in the theory of epidemiology: the reproduction number \(R_0\) that determines whether or not an epidemic with few initial infectives can become established and the
probability of a large outbreak of a disease that becomes established. These quantities have importance for health officials in understanding and controlling the spread of epidemics. We explored different strategies to derive explicit expressions for these two quantities for an SI type epidemic that spread across a dynamic network.

Another area of research is to investigate the duration of the epidemic. This is not studied very often (but see Barbour and Reinert, 2013; Barbour, 1975). In case of a large outbreak, the duration of an epidemic is important, for instance, during the outbreak travel/trade bans are imposed on the countries within which the disease is spreading. The longer durations of epidemics might lead to severe economic costs for those countries. Therefore, in the second paper we focus on the duration of an SIR (Susceptible, Infectious, Recovered) epidemic in a population structured through a random graph, generated by the configuration model. We focus especially on the final stages of the large epidemic outbreak and analyse the large graph limit of an SIR epidemic. The aim is to derive the asymptotic form of the time needed for an infection to completely disappear from the population as the size of the graph tends to infinity. In order to achieve this we investigate the so-called epidemic-generated graphs. To obtain the proposed results, we couple both the start and the end of the epidemic with suitable branching processes where we use a supercritical branching process for the start of an epidemic and a subcritical one for the end of an epidemic.

The main result of paper II (Theorem 2.2) states that the (properly scaled) duration of the epidemic is of the order of $\log n$, where the involved constants are expressed as functions of the Malthusian parameters of suitable continuous time branching processes. This result then generalizes the results of Barbour and Reinert, 2013, who do not study the end of the epidemic, to a wider range of degree distributions and infectious time distributions. The result is proven essentially in two steps: first showing that the first time the number of infectious individuals is of order $n$ is of order $\log n$ (Lemma 2.6), and then investigating the process between the latter time and the extinction time is also of order $\log n$ (Lemma 2.7). The proof of the first result heavily relies on existing results on branching processes (Jagers, 1975), and on a generalization of the results in (Barbour and Reinert, 2013) to degree distributions and infectious time with unbounded supports.

In paper III we use the model presented in paper II, by including vaccination in the model, and explore the effect of various vaccination strategies on the duration of the end of an SIR epidemic. We consider the effect of a uniform and a leaky vaccine in both homogeneous mixing and structured populations. We show that the uniform vaccine in homogeneously mixing populations and structured populations can have different effects on the duration of the end of
an SIR epidemic. In uniformly/homogeneous mixing population, we find that uniform vaccination with a perfect vaccine either prevent the large outbreak or it increase the duration of the epidemics whereas uniform vaccination in structured populations may lead to both an increase and decrease in the duration of the epidemic. Furthermore, the leaky and uniform vaccine have the same effect on the duration of the epidemic in homogeneously mixing populations whereas we conjecture that a leaky vaccine always increases the duration of the epidemic in structured populations. For acquaintance vaccination, we provide a recipe for calculating the asymptotic duration of the epidemic. We show that acquaintance vaccination can both increase and decrease the asymptotic duration of an epidemic.

In the last paper, we study a model in which two SIR type epidemics spreads successively through a single population modeled by a configuration model network. The infection with the second epidemic might depend on prior infection with the other epidemic. We assume that the first disease spreads through the network and causes a major epidemic outbreak infecting a non-negligible fraction of the population. Individuals infected during the first epidemic might becomes partially or fully immune for the second one. The spread of the first epidemic is analysed through a bond percolation model while the early stages of the second epidemic is approximated by a multitype branching process. The results are illustrated through two examples. The first example is for the case in which individuals infected in the first epidemic are independently either completely susceptible or completely immune to the second epidemic. The second example is for the case in which the individuals who have recovered from the first epidemic have reduced susceptibility and infectivity for the second epidemic.

A three type branching process is used to calculate a threshold parameter, $R_0^{(2)}$, that determines whether a major outbreak of the second epidemic can occur or not. A recipe of computing the probability of a large epidemic outbreak of the second epidemic is also provided. Furthermore, our paper generalizes some epidemic models discussed in previous literature (Bansal and Meyers, 2012; Newman, 2005; Newman and Ferrario, 2013)). That is to say that the models in (Newman, 2005; Newman and Ferrario, 2013) and (Bansal and Meyers, 2012) are special cases of our model.
5. Sammanfattning


I paper I, utökar vi modellen i Leung et al. (2015) genom att betrakta epidemin och populationen som stokastiska processer. Vi diskuterar hur antagandet om oberoende gradtal mellan individer (ett felaktigt antagande) påverkar resultaten om modellen. Vi föreslår även ett nytt sätt att modellera demografin under antagandet att nya individer har en stationär gradfördelning av partners. Alltså, nya individer formar direkt vid ankomst ett slumpmässigt antal partnerskap med individer redan i populationen, och detta antal beror ej på tiden av ankomst (stationärt). Under detta antagande försvinner beroendet mellan individers gradtal och antagandet om $mean\ field\ at\ distance\ one$ behövs inte längre.


Huvudresultatet i paper II säger att varaktigheten av epidemin (lämpligt skalad) är av ordning log\(n\). Detta resultat generaliserar resultatet av Barbour och Reinert 2013 (som inte studerar slutet på epidemin), till en bredare klass av gradfördelningar och infektionstider. Resultatet bevisas väsentligen i två steg: först visar vi att tiden då det finns ungefär \(n\) smittade individer är av ordning log\(n\) (Lemma 2.6), sedan visar vi att denna tid och tiden till epidemin bör ut är också av ordning log\(n\) (Lemma 2.7). Beviset för första resultatet bygger på resultat för förödelserprocesser (Jagers, 1975), och på en generalisering av resultaten ur (Barbour and Reinert, 2013).


References


