Stockholms universitet

SIR epidemics on random graphs with clustering

Carolina Fransson

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2017:3
Matematisk statistik
Maj 2017

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Matematiska institutionen

Stockholm University

Mathematical Statistics
Stockholm University
Master Thesis **2017:3**
http://www.math.su.se

SIR epidemics on random graphs with clustering

Carolina Fransson*

May 2017

Abstract

In this project, we consider SIR (Susceptible → Infectious → Recovered) epidemics on random graphs generated by a version of the configuration model with clustering. Miller (2009) investigated SIR epidemics on graphs of this model under the assumption of homogeneous infectivity. We extend previous results by relaxing this assumption. We use a branching process approximation of the spread of the disease to provide expressions for the probability of a major outbreak and the expected final size. Furthermore, we show that for this particular model, the basic reproduction number, here defined as the rank based geometric growth rate of the epidemic, equals the perfect vaccine-associated reproduction number. Moreover, we use maximal coupling to prove that the branching process approximation is exact in the limit as the population size approaches infinity.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: carolina.fransson@math.su.se. Supervisor: Pieter Trapman.

Acknowledgements

I sincerely thank my supervisor Dr. Pieter Trapman for the many fruitful discussions and for giving me the freedom to explore my ideas: your help and support have been invaluable.

Contents

1 Introduction

Infectious diseases have played a dramatic role in the history of humankind. Among the most lethal pandemics recorded are the Justinian Plague (541-544), the Black Death (1346–1353) and the Spanish flu (1918-1920). The Justinian Plague, named after the Byzantine emperor Justinian, and the Black Death, probably named after the black plaques that appeared on the skin of patients, were caused by the bacterium *Yersinia pestis* (Harbeck et al. 2013). The Justinian Plague hit the Byzantine Empire in 541, and then spread rapidly to North Africa, Italy, Spain, and the French- German border. The Black Death struck the Mediterranean and Europe. It is estimated that the Black Death reduced the European population by 25% and the population of the Muslim Middle East by 30% (Tibayrenc 2007). The 1918 flu pandemic, also called the Spanish flu, was caused by a strain of the H1N1 influenza virus. One third of the world's population is said to have been infected, with a total number of 40-50 million deceased (Taubenberger and Morens 2006), which makes the Spanish flu one of the deadliest pandemics ever recorded. Unlike many other influenza strains, the Spanish flu was especially dangerous for previously healthy young adults. More recent outbreaks of infectious diseases include SARS, MERS, Ebola, Zika and the 2009 H1N1 pandemic.

Finding adequate models for the spread of infectious diseases is of considerable importance to society. In order to take suitable prevention and control measures, adequate understanding of the spread of the disease is required. The practical usefulness of an epidemic model depends on its mathematical tractability and how well it captures the most important mechanisms of the epidemic. In most modelling activities, one faces the classical trade-off of fidelity versus simplicity. One of the most important factors that determine the fate of an outbreak is the contact patterns in the population. The frequency and duration of the contacts between individuals typically depends on the nature of their relationship. To incorporate social structure in an epidemic model, the individuals of the population and the relationship between them may be represented by a graph; the nodes of the graph represent the individuals of the population, and an edge represents a "close" social relationship (Figure 1.1). In other words, an edge represents a potential channel for transmission of the disease.

In this project, we incorporate two main features in the analysis of the spread of an infectious disease; heterogeneity in the infectivity of the individuals of the population, and a feature of the underlying social network called *clustering*.

Heterogeneity in individual infectivity may, for instance, reflect variability in the infectious periods of the individuals who contract the disease, or variability in the type and amount of social interaction that individuals engage in.

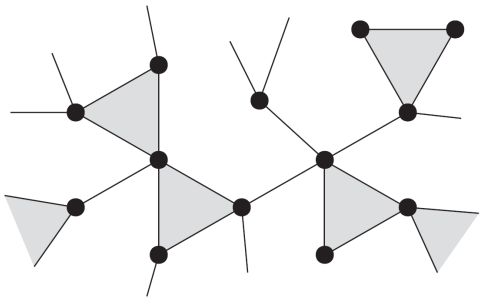

Figure 1.1: Network with clustering representing a social network. Individuals are represented by nodes (black dots), an edge connecting two nodes signifies a social relationship. This picture was retrieved from the paper by Newman (2009).

A network exhibits clustering if it has a high amount of triangles (see Figure 1.1). For a social network, the presence of triangles represents a tendency to have mutual acquaintances. That is, the friends of an individual tend to be friends as well. To incorporate clustering, we use a version of the configuration model[1], proposed by

[1] For an account on the standard configuration model we refer the reader to the book by van der Hofstad (2016, Chapter 7).

Miller (2009) and Newman (2009), independently.

Miller (2009) investigated the spread of a disease on a graph generated by the configuration model with clustering, assuming homogeneous infectivity and susceptibility of infectives. In other words, transmission occurs along each infected-susceptible edge independently with some fixed probability T. In this project, we extend these results by allowing for heterogeneity in individual infectivity. We employ a branching process approximation, which enables us to provide expressions for the probability of a major outbreak (that is, the probability that a non-negligible fraction of the population is infected). We use a similar approach, based on *susceptibility sets and backward branching processes* (see section 3.2) to provide expressions for the expected final size of a major outbreak.

The basic reproduction number R_0, loosely defined as the expected number of infected cases caused by a "typical" infected individual in an otherwise susceptible population, has the threshold property that a major outbreak is possible if and only if $R_0 > 1$. Miller (2009) calculated the basic reproduction number by attributing all secondary and tertiary cases infected by virtue of transmission within a given triangle to the primary case in that triangle, regardless of the true path of transmission. In this thesis, we extend this result by calculating the rank based reproduction number. In addition, we incorporate vaccination in the epidemic model and show that, for the particular model under consideration, vaccinating a fraction $1-\frac{1}{R_0}$ of the population with a perfect vaccine is sufficient to surely prevent a major outbreak.

This report is structured as follows. In section 2, we specify the epidemic model and describe the tools we use to analyse the spread of the disease. Section 3 contains a description of how a branching process approximation can be employed to approximate the probability and size of a major outbreak. In section 4 we analyse an SIR epidemic in discrete time and provide expressions for the probability and size of a major outbreak. The continuous time case is discussed briefly in section 5. In section 6 we incorporate immunity stemming from vaccination in the epidemic model, and show that the basic reproduction number equals the perfect vaccine-associated reproduction number. In section 7 we give a brief summary of the results presented and discuss possible extensions and future work.

A list of frequently used notation and its (usual) meaning, and a list of the assumptions made, are available in Appendix A. In Appendix B we discuss tools to compare the final size and the probability of non-extinction of epidemic models by the infectivity and transmissibility of the models. In Appendix C we describe the concept of coupling, which we use to prove that the branching process approximations are exact in the limit of large population sizes. The proof is presented in Appendix D.

2 Background

In this section, we describe the tools that we will later use to analyse the spread of the disease.

2.1 The SIR model

In this project, we use an SIR model to investigate the dynamics of the spread of an infectious disease. At any given time point, the population is divided into three groups, depending on health status. The groups are susceptible (**S**), infectious (**I**) and recovered (**R**) (Britton 2010). The role of an individual in the epidemic is determined by its group. Individuals of the population make contact with other individuals at (possibly random) points in time. If, at some time point, an infectious individual makes contact with a susceptible individual then the susceptible individual instantaneously becomes infectious. An infectious individual may cease to be contagious after a period of time, which we call the *infectious period* of the individual in question, and is then transferred to the group recovered. Recovered individuals are those that are immune to the disease (or dead). Individuals belonging to this group play no role in the spread of the disease. The population is assumed to be closed. This means that we ignore births, deaths (except for the individuals in the group recovered) and migration.

In the simplest case, we only allow the transitions $S \to I$ and $I \to R$. That is, recovered individuals gain lasting immunity against the disease. This model, the SIR model, is the model considered in this report. One possible extension of this model is to incorporate waning immunity in the model by, in addition, allowing the transition $R \to S$, which results in the SIRS model. These two models are illustrated in Figure 2.1. Another possible extension of the SIR model is to incorporate incubation periods by adding the latent group "exposed" (**E**) to the model. In the SEIR model, a susceptible individual that gets infected resides in the group exposed for a period of time (the incubation period) before transferring to the group infectious. We now turn our attention to the SIR model again.

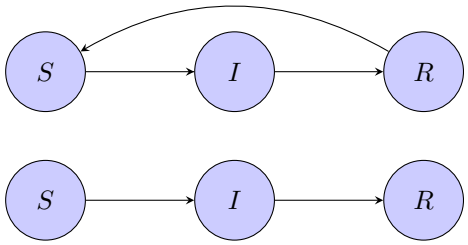

Figure 2.1: Allowed transitions in two different models. Top: An SIRS model with waning immunity. Bottom: An SIR model with lasting immunity. This is the model considered in this report.

More formally, let the set $\mathcal{T}$ be a time index set. We consider epidemics in discrete time, corresponding to $\mathcal{T} = \{0, 1, 2, \ldots\}$, and in continuous time, corresponding to $\mathcal{T} = \mathbb{R}_+$. For each time point $t \in \mathcal{T}$, let $\mathcal{S}(t)$ be the set of susceptible individuals at t. Similarly, let $\mathcal{I}(t)$ be the set of infectious individuals at time t, and $\mathcal{R}(t)$ the set of recovered individuals at t. The number of susceptible, infectious and recovered individuals at t is then given by $|\mathcal{S}(t)|$, $|\mathcal{I}(t)|$ and $|\mathcal{R}(t)|$ respectively.

Throughout, we denote the population size by N. At the start of the epidemic (that is at time $t = 0$), we typically have a relatively small number of infectious and no recovered individuals, say $|\mathcal{I}(0)| = m$ initially infected and $|\mathcal{S}(0)| = N - m$ initially susceptible.

In our analysis, we assume that there is one initially infected individual and that the population size N is large.

We divide the individuals of the populations into the three sets $\mathcal{S}(0)$, $\mathcal{I}(0)$ and $\mathcal{R}(0)$ according to their health status at time 0. We often assume that $\mathcal{R}(0)$ is empty, that is that all individuals except for the initial case(s) are initially susceptible to the disease. If v is an initial case then the time point $t_I^{(v)}$ at which v contracts the disease is set to 0.

The disease is transmitted if an infectious individual makes contact with a susceptible individual. In a real world setting, the contact pattern between individuals typically varies from pair to pair. To capture these heterogeneities we introduce contact processes (not to be confused with an SIS epidemic, which is often referred to as a contact process, see for instance (Liggett 1999)) of the form $C^{(u,v)} = \{C_n^{(u,v)}\}_{n\in\mathbb{N}}$. For each pair of individuals u, v the time points at which u attempts to make contact with v is a point process $C^{(u,v)}$ on $\mathcal{T}$. Note that the time points $C^{(u,v)}$ at which u attempts to make contact with v are in general not the same as the time points $C^{(v,u)}$ at which v attempts to make contact with u. In section 2.3 we incorporate social structure and "friendship" in our model, so that the probability law of $C^{(u,v)}$ is determined by the nature of the relationship between u and v. In the continuous time case, we assume that the attempted contacts between individuals are independent homogeneous Poisson processes. If an infectious individual attempts to make contact with a susceptible individual and succeeds, then the susceptible individual immediately becomes infectious. Whether an attempted contact is successful or not is governed by some probability law, specified in the model. We may, for instance, consider a so called leaky vaccine. Such vaccine provides partial protection against the disease, in the sense that the risk of transmission at an attempted infectious contact is reduced, but not eliminated. Leaky vaccines are briefly discussed in Example B.2, Appendix B.

We now give a description of how individuals transfer from the group infectious to the group susceptible and from susceptible to recovered. Assign to each individual v two numbers, the time point $t_I^{(v)} \in \mathcal{T} \cup \{\infty\}$ at which v gets infected, and the infectious period τ_v of v, where $0 \leq \tau_v \leq \infty$. An individual v is susceptible at time points in $[0, t_I^{(v)})$, infectious in $[t_I^{(v)}, t_I^{(v)} + \tau_v)$ and recovered in $[t_I^{(v)} + \tau_v, \infty)$. If v ultimately escapes infection then $t_I^{(v)} = \infty$. These numbers are assigned according to the following rules.

For any function f with domain $\mathcal{T}$, we define $f(t^-)$ as the limit as t goes to 0 from below. For instance, $\mathcal{S}(t^-)$ is the set of individuals that are susceptible "right before" t. In the discrete time case $\mathcal{S}(t^-) = \mathcal{S}(t-1)$ for each $t \in \mathcal{T}_{>0}$. In the continuous time case, $\mathcal{S}(t^-)$ consists of precisely the individuals u that are susceptible in the non-empty time interval $(t-\delta, t)$ for some $\delta(u) > 0$. In our model, transmission of the disease and recovery happens at time points t^-, $t \in \mathcal{T}_{>0}$. If at some time point $t \in \mathcal{T}$ the individuals u makes contact with v, and $u \in \mathcal{I}(t^-)$ and $v \in \mathcal{S}(t^-)$, then v instantaneously becomes infected, and stays infectious for a time period of length τ_v. That is $v \in \mathcal{I}(s)$ for all $t \leq s < t + \tau_v$ and $v \in \mathcal{R}(s)$ for $s \geq t + \tau_v$, and $t_I^{(v)} = t$.

One of the simplest SIR models is the *standard stochastic SIR model*, in which *homogeneous mixing* of the population is assumed, i.e. any pair of individuals is assumed to be equally likely to interact with each other during any given time period. In this continuous time model, the contact processes of $\{C^{(u,v)}\}$ are independent and all obey the same probability law. That is, the contact processes $\{C^{(u,v)}\}$ are independent homogeneous Poisson processes on $\mathbb{R}_+$ with common intensity. Let λ be the positive number that satisfies that the per-pair contact intensity is $\frac{\lambda}{N-1}$, where N is the population size. For a specific individual u, the rest of the population attempts to make contact with u at a Poisson rate λ. Similarly, at any given time point t, infectious individuals attempts to make contact with a specific susceptible individual at a Poisson rate $\frac{|\mathcal{I}(t)|}{N-1}$. In this

model, the infectious periods of the individuals are assumed to be independent and exponentially distributed with common mean. Note that the standard stochastic SIR model is Markovian, i.e. $(\mathcal{I}(t), \mathcal{S}(t), \mathcal{R}(t)),\ t \in \mathbb{R}_+$, is a Markov process.

2.2 Reproduction numbers

One of the key quantities in the study of epidemics is the basic reproduction number, often denoted by R_0. One of the most important features of R_0 is its threshold properties; a *major outbreak* is only possible if $R_0 > 1$. We say that an outbreak is major if a non-negligible fraction of the population contracts the disease. In section 3.2, we will see that for the models considered in this thesis, the distribution of the proportion of the population that is ultimately infected converges to a two point distribution concentrated on $\{0, f\}$ for some constant $0 < f < 1$ as the population size $N \to \infty$ (Theorem 3.1), provided that $R_0 > 1$. In other words, in the limit of large population sizes, an outbreak may be minor, with a fraction ≈ 0 of the population ultimately infected, or major with a fraction $\approx f$ of the population ultimately infected. Another important feature of R_0 is that it has the interpretation of the expected number of infected cases caused by a typical infected individual in an otherwise susceptible population (Heesterbeek 2002). It has its roots both in epidemiology and ecology/demography. In ecology and demography, R_0 represents the expected number to female offspring born to one female of the population over the course of her lifetime. As we will see in this report, the two interpretations are closely related.

It is in general not straightforward to define R_0 for more complex models, where the interactions between the individuals of the population depend on the underlying social structure established by the individuals of the population. There is a vast number of proposed reproduction numbers for such models, see for instance Ball et al. (2016) for a overview and comparison of a number of reproduction numbers for epidemics among households and households-workplaces.

For many models, including those analysed in this report, the basic reproduction R_0 number may be defined as the *geometric growth rate* (Pellis et al. 2012)

$$\lim_{k\to\infty} \lim_{N\to\infty} E\left(X_N^{(k)}\right)^{1/k}, \tag{2.1}$$

if this limit exists[2], where $X_N^{(k)}$ is the number of cases of *generation* k of an epidemic in a population of size N. For $k \in \mathbb{N}$, an infected case v is said to belong to generation k if the chain of transmission from an initial case to v is of length k. That is, the initial case(s) belongs to generation zero, the cases caused by the initial case(s) belongs to generation one etc.

If the spread of the disease in the early phase of the epidemic is well approximated by a suitably chosen branching process, then the interpretation of R_0 as the expected number of cases cased by the typical individual in the early phase of the epidemic is retained by this definition. The approximating branching process may be a single or multi-type branching process. In the single-type case, the expected number of individuals infected by a typical infected individual during the early stages of the epidemic is given by the expected number of offspring μ of an individual in the approximating branching process. As we will see in section 2.6, the probability of extinction of the branching process is less than one if and only if $\mu > 1$. Thus, $R_0 = \mu$ is a threshold parameter. The multi-type case, which is discussed in more detail in section 3.3, can be given a similar interpretation.

[2] Under mild conditions, the limit in (2.1) exists and has the threshold properties of the basic reproduction number. This is discussed in more detail in section 3.3

In addition to the basic reproduction number R_0, we consider the *perfect vaccine-associated reproduction number* R_V. A vaccine is *perfect* if it provides full and permanent immunity. That is, an individual vaccinated with a perfect vaccine cannot contract the disease. The perfect vaccine-associated reproduction number R_V is defined as (Ball et al. 2016)

$$R_V = \frac{1}{1 - f_{\mathrm{v}}^{(c)}},$$

where $f_{\mathrm{v}}^{(c)}$ is the *critical vaccination coverage*. The critical vaccination coverage is defined as the fraction of the population that has to be vaccinated with a perfect vaccine in order to reduce R_0 to unity. That is, $f_{\mathrm{v}}^{(c)}$ is the fraction necessary to vaccinate in order to be guaranteed to prevent a major outbreak (Britton 2010). Note that if $R_0 \leq 1$ then $f_{\mathrm{v}}^{(c)} = 0$.

In the example of the standard stochastic SIR epidemic model introduced in section 2.1, $R_V = R_0$ (Britton 2010). Ball et al. (2016) showed that, for the households and households-workplaces models $R_V \geq R_0$ if $R_0 > 1$. That is to say, in order to surely prevent a major outbreak, vaccinating a fraction $1 - \frac{1}{R_0}$ with a perfect vaccine is sufficient to surely prevent a major outbreak only if $R_V = R_0$. However, if a fraction strictly less than $1 - \frac{1}{R_0}$ is vaccinated with a perfect vaccine then a major outbreak might still occur.

In section 6, we show that for the models analysed in this report, $R_V = R_0$.

2.3 Graph interpretation of an epidemic

The spread of an epidemic can be given a directed graph interpretation. As we pointed out in section 2.1, the contact processes $\{C^{(u,v)}\}$ typically varies from pair to pair. The contact pattern in a population can be represented by a social network, where the structure of the network is determined by the interactions between the individuals of the population. A link between two individuals in such network represents a relationship. Contacts are often directed in the sense that an individual makes contact with another individual. To capture directed contacts, we represent the social network by a directed graph, where the nodes of the graph represents the individuals of the population and a directed edge represents frequent (directed) contact between the two individuals. In this project, we investigate the spread of SIR epidemics in populations where the social structure can be represented by a graph generated by the configuration model with clustering presented in section 2.5.

Consider a graph G_N consisting of N nodes $v_1, v_2, \ldots, v_N$, each representing an individual of the population. Let $\mathcal{V} = \{v_1, v_2 \ldots, v_N\}$ be the node set. In the graph interpretation, potential infectious contacts between individuals are represented by weighted directed edges. Let $\mathcal{E}$ be a set of ordered triples, or *directed edges*, (v_i, v_j, d_{ij}) such that $v_i, v_j \in \mathcal{V}$ and $d_{ij} \in \mathbb{R}_+ \cup \{\infty\}$. We sometimes write (v_i, v_j) instead of (v_i, v_j, d_{ij}), not making the *edge weight* d_{ij} explicit. Given an edge (v_i, v_j), we refer to v_i as the *tail* of (v_i, v_j) and to v_j as the *head* of (v_i, v_j). In Figure 2.2, a graph representation of an epidemic in a small population is illustrated. Each directed edge is represented by an arrow, pointing from the tail to the head. The edge weight d_{ij} of an edge (v_i, v_j, d_{ij}) is to be interpreted as the transmission time from v_i to v_j, should v_i contract the disease. That is to say, if v_i gets infected at the time point t say, then v_j gets infected at $t + d_{ij}$, provided that v_i transmits the disease to v_j. In the example of Figure 2.2, the disease would not be transmitted along the edge (v_2, v_1) since the edge weight d_{21} is infinite. The disease would, on the other hand, be transmitted along the edge (v_4, v_3) if v_4 were to contract the disease, since the edge weight $d_{42} = 3$ is finite.

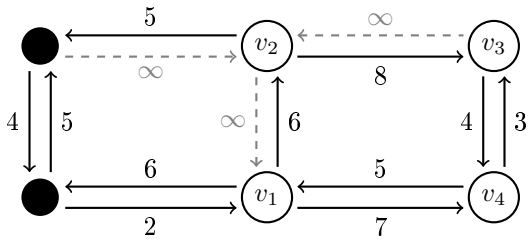

Figure 2.2: Graph representation of an epidemic in a small ($N = 6$) population. A solid arrow (a directed edge) from a node (the tail) to another node (the head) represents that the tail would attempt to infect the head, if infected. Each edge is labeled with an edge weight, which is to be interpreted as the time of disease transmission from the tail to the head.

Let $i_1, i_2, \ldots, i_k$ be indices such that there is a directed edge from the node v_{i_m} to the node $v_{i_{m+1}}$ for $m = 1, 2, \ldots, k-1$. We may then define the *path* $v_{i_1} \to \ldots \to v_{i_k}$ as the ordered sequence $\{(v_{i_m}, v_{i_{m+1}})\}_{m=1}^{k-1}$. In the example of Figure 2.2, the path $v_1 \to v_2 \to v_3$ exists since there is a sequence of directed edges corresponding to $v_1 \to v_2 \to v_3$. However, the path $v_1 \to v_3$ does not exist. We refer to v_{i_1} as the *starting point* of $v_{i_1} \to \ldots \to v_{i_k}$ and to v_{i_k} as the *end point* of $v_{i_1} \to \ldots \to v_{i_k}$. We define the *length* $L(v_{i_1} \to \ldots \to v_{i_k})$ of the path $v_{i_1} \to \ldots \to v_{i_k}$ as

$$L(v_{i_1} \to \ldots \to v_{i_k}) = \sum_{m=1}^{k-1} d_{i_m i_{m+1}}.$$

We make the convention that for any node v, there exists a path $v \to v$ from v to itself of length 0. For $v_i, v_j \in V$ define the *distance* $d(v_i, v_j)$ from the node v_i to the node v_j as the minimum length taken over all paths with starting point v_i and end point v_j. If no such path exists then $d(v_i, v_j)$ is taken to be ∞. Note that the distance between a node and itself is 0, that is $d(v, v) = 0$ for all nodes v. The distance $d(u, v)$ from a node u to another node v represents the transmission time of the disease from u to v, should u contract the disease.

An edge v gets infected if and only if $d(v_*, v) < \infty$, where v_* is the initial case. It should be noted that an edge with finite weight does not necessarily transmit the disease, since the weight represents the time of transmission, *should the tail get infected*. In the example of Figure 2.2, the disease would not be transmitted along the path $v_1 \to v_2 \to v_3$, since the path $v_1 \to v_4 \to v_3$ is shorter. In the case of a non-atomic weight distribution the path lengths are almost surely distinct, and the paths of transmission are almost surely well defined. We refrain from elaborating on the path of infection in the case of weight distributions with atoms, since it is not relevant for our analysis.

In the setting with random infectious periods described in section 2.1, the graph may be constructed as follows. Recall that $t_I^{(v_i)}$ is the time point at which the node v_i contracts the disease. To each node v_i, assign an infectious period τ_i, and for each directed edge (v_i, v_j) let t_{ij} be the time elapsed between $t_I^{(v_i)}$ and the first time point in $(t_I^{(v_i)}, \infty)$ at which v_i contacts v_j. If v_i does not contact v_j then t_{ij} is taken to be ∞. The edge weight d_{ij} of the edge (v_i, v_j) is then taken to be $d_{ij} = \infty \cdot \mathbb{1}(\tau_i < t_{ij}) + t_{ij}$, where we agree to $\infty \cdot 0 = 0$.

2.4 Clustering

In this project, we analyse epidemics in populations where the structure of the contact network established by the individuals of the population may be represented by a graph

with high degree of clustering. A graph exhibits clustering if nodes with mutual neighbours tends to be neighbours as well (van der Hofstad 2016, Section 1.5). In other words, a network with high degree of clustering contains a relatively high amount of triangles. A high degree of clustering is, for instance, prevalent in social networks.

Let $G = (E, V)$ be an undirected graph with node set V and edge set E. The *clustering coefficient* of $G = (V, E)$ is a measure of the degree of clustering of G. Three nodes $u, v, w \in V$ are said to form an *ordered wedge* (u, v, w) if v is connected both to u and to w. If, in addition, u and w share an edge then the wedge (u, v, w) is called a *triangle*. The ordered wedge/triangle (u, v, w) is said to be *centered on* v. We note that (u, v, w) is an ordered wedge if and only if (w, v, u) is an ordered wedge. Furthermore, if u, v and w constitute a triangle then there exist six ordered wedges with members u, v and w. Let

$$\mathcal{W}^G = \{(u, v, w) \in V^3 : (u, v), (v, w) \in E\}$$

be the set of all ordered wedges of G and

$$\mathcal{W}^G_\Delta = \{(u, v, w) \in V^3 : (u, v), (v, w), (w, u) \in E\} \subset \mathcal{W}$$

be the set of all ordered triangles of G. The clustering coefficient $C(G)$ of a graph G is defined as the fraction of ordered wedges that are also triangles:

Definition 2.1 (Clustering coefficient)**.** *Let $G = (V, E)$ with $|\mathcal{W}^G| > 0$. The* clustering coefficient *of G, $C(G)$, is given by*

$$C(G) = \frac{|\mathcal{W}^G_\Delta|}{|\mathcal{W}^G|}.$$

Let $\overline{G} = \{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs. The sequence $\overline{G}$ is said to be *highly clustered* if

$$\liminf_{n\to\infty} C(G_n) > 0.$$

2.5 The graph model

In this project, we investigate an SIR epidemic on random graphs with clustering, generated by a version of the configuration model (see for instance van der Hofstad (2016, Chapter 7). The configuration model with clustering was independently introduced by Miller (2009) and Newman (2009). For alternative graph models with clustering, see for instance Deijfen and Kets (2009) and Trapman (2007).

A graph generated from this model is constructed as follows. Let

$$\{p(k_S, k_\Delta)\}_{k_S, k_\Delta \in \mathbb{N}_0}$$

be a prescribed joint degree distribution, where k_S denotes the number of single edges, and k_Δ denotes the number of pairs of triangle edges. Throughout, the two-dimensional random vector (S, Δ) is assumed to have the joint degree distribution p. Let $\{(S_i, \Delta_i)\}_{i=1}^N$ be a sequence of independent copies of (S, Δ). In other words, we assume that the elements of this sequence are independent realizations of the degree distribution p. Analogously to the standard configuration model, a graph $G_N = G_N(p)$ of size N is then constructed by first assigning the single degree S_i and the triangle degree Δ_i to the node v_i, $i = 1, 2, \ldots, N$. It may be helpful think of this step in terms of half-edges; to each node v_i, we attach S_i single half-edges and Δ_i pairs of triangle half-edges. The

single half-edges are then matched uniformly, and the triangle half-edges are matched uniformly. The process of joining half-edges is illustrated in Figure 2.3.

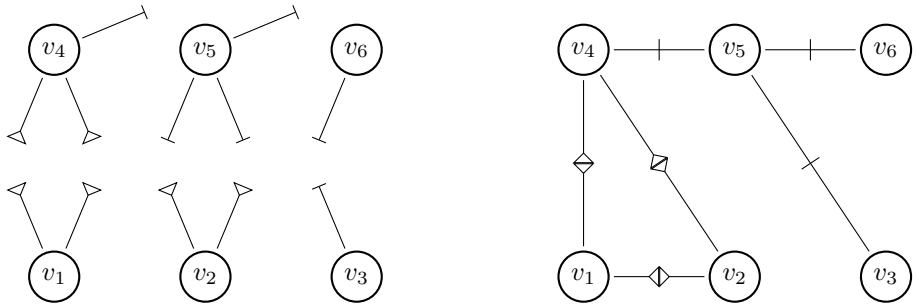

Figure 2.3: Schematic illustration of the construction of a configuration graph with clustering. Triangle half-edges (marked with a triangle) and single half-edges (marked with a perpendicular line) are assigned to the nodes of the graph (left). The half-edges are then matched uniformly at random (right).

In practice, the matching may be carried out as follows. Two lists of nodes, one single degree list and one triangle degree list are created. A node with joint degree (k_S, k_Δ) will appear k_S times in the single list and k_Δ times in the triangle list. The lists are then shuffled uniformly, and the nodes on positions $2m+1$ and $2m+2$ in the single list and positions $3m+1$, $3m+2$ and $3m+3$ in the triangle list are then matched, $m \in \mathbb{N}_0$.

We define the *total single degree* as

$$D_S^{(N)} = \sum_{i=1}^{N} S_i$$

and the *total triangle degree* as

$$D_\Delta^{(N)} = \sum_{i=1}^{N} \Delta_i.$$

If the total single degree (that is, the length of the single degree list) is not even, or if the total triangle edge degree (the length of the triangle degree list) is not a multiple of three, then we erase a single half-edge and/or one or two triangle half-edge pairs. Similarly, we erase self-loops and multiple edges, so that the resulting graph is simple. The impact on the degree distribution is negligible as the number of nodes N tends to infinity, as we will see below.

We assume that S and Δ both have finite second moments. Note that this assumption implies that $S\Delta$ is integrable. This follows from Hölders inequality (Friedman 1982, Theorem 3.2.1), or just from noting that $2|S\Delta| \leq (S^2 + \Delta^2)$. It holds that under the assumption of finite second moments the number of single self loops and single double edges converge in distribution to independent Poisson random variables with means (van der Hofstad 2016, Proposition 7.13)

$$\frac{1}{2}E(S_\bullet^{(s)}) \text{ and } \frac{1}{4}E(S_\bullet^{(s)})^2$$

respectively, where $S_\bullet^{(s)}$ has the downshifted single size biased distribution (defined on page 18), that is $E(S_\bullet^{(s)}) = \left(\frac{E(S^2)}{E(S)} - 1\right)$. One can show that for a configuration model graph with clustering, the number of triangle self loops (that is, the number of unordered "triangles" u, v, w where at least two of the nodes u, v and w are identical), the number of triangle parallel edges and the number of triangle edges that are parallel with a single edge converge in distribution to independent Poisson random variables with finite means. The

proof, which is completely analogous to the proof of Proposition 7.13 in van der Hofstad (2016), is omitted.

Thus, self loops and multiple edges are negligible in the limit as $N \to \infty$. In the remainder of this report, we ignore the small differences in the topology of the graph that arise from erasing multiple edges or self loops. In addition, we ignore the small differences that arise from erasing half-edges so that the number of single and triangle half-edges are multiples of two and three, respectively.

An important feature of the matching process is that it may be carried out sequentially, in a random order. In each step, a half-edge (or a half-edge triangle pair) is chosen according to some arbitrary, possibly random, rule. The chosen half-edge (or pair) is then paired with another half-edge (or to two other triangle half-edge pairs) chosen uniformly at random. The (possibly random) order in which the pairing is carried out does not affect the distribution of the topology of the graph, as long as the matching is done uniformly (van der Hofstad 2016, Lemma 7.6). We say that the order in which the pairing is carried out is *exchangeable*.

This is a useful result, since it allows us to construct the graph by exploring the neighbourhood of a vertex in a random order (van der Hofstad 2016, Section 7.2). In other words, we may construct/explore the graph G_N as the epidemic propagates, by matching the half-edges attached to a node when the node gets infected. The exchangeability of the pairing order plays a crucial role in the proofs presented in Appendix D.

2.5.1 Assumptions on the degree distribution

Recall that (Δ, S) has the joint degree distribution p. We make the following assumptions on p.

A1) $E(\Delta^2) < \infty$ and $E(S^2) < \infty$.

A2) $E(\Delta^2 \log(\Delta)) < \infty$ and $E(S^2 \log(S)) < \infty$

A3) $P(\max(\Delta, S) \geq 2) > 0$ and $E(\Delta S) > 0$.

It is not hard to see that assumption A2 implies assumption A1. Assumption A2 ensures that the condition given in (2.11), section 3, is satisfied for the approximating branching process. Note that the assumption A1 implies $E(S\Delta) < \infty$ by the inequality $S\Delta \leq \frac{S^2+\Delta^2}{2}$. Similarly, assumption A2 implies $E(S\Delta \log(\Delta)) < \infty$ and $E(S\Delta \log(S)) < \infty$.

2.5.2 Clustering coefficient of the graph model

Let the degree sequence $\bar{d} = \{(S_i, \Delta_i)\}_{i \in \mathbb{N}}$ be a given sequence of independent copies of (S, Δ). Consider the graph sequence $\overline{G} = \{G_N\}_{N \in \mathbb{N}}$ of graphs generated by the configuration model with clustering, where the degree sequence of the graph G_N is given by $\bar{d}_N = \{(S_i, \Delta_i)\}_{i=1}^N$ for each N. We now show that the graph sequence $\overline{G}$ is highly clustered for almost every realization of the degree sequence $\bar{d}$.

Ignoring the small changes in the empirical degree distributions of the graphs of $\overline{G}$ that might arise from multiple edges, self loops or total single (triangle) degrees that are not a multiple of two (three), the total number of ordered triangles of G_N is bounded from below by

$$|\mathcal{W}_\Delta^{G_N}| \geq \sum_{i=1}^{N} 2\Delta_i.$$

and the total number of ordered wedges is given by

$$|\mathcal{W}^{G_N}| = \sum_{i=1}^{N} \binom{S_i + 2\Delta_i}{2} 2 = \sum_{i=1}^{N} (S_i + 2\Delta_i)(S_i + 2\Delta_i - 1).$$

Indeed, for a given node u_i of G_N with joint degree (S_i, Δ_i), the number of ordered triangles formed by triangle half-edges centered on u_i is given by $2\Delta_i$. Similarly, the number of ordered wedges centered on u_i is given by $2\binom{S_i+2\Delta_i}{2}$.

By the Strong Law of Large Numbers (Theorem D.1) and the assumption A1 that S and Δ both have finite second moments

$$C(G_N) \geq \frac{\left(\frac{\sum_{i=1}^{N} 2\Delta_i}{N}\right)}{\left(\frac{\sum_{i=1}^{N}(S_i+2\Delta_i)(S_i+2\Delta_i-1)}{N}\right)} \to \frac{E(2\Delta)}{E((2\Delta+S)^2) - E(2\Delta+S)} \tag{2.2}$$

as the population size $N \to \infty$ for almost every realization of the degree sequence $\bar{d}$. Thus, the graph sequence $\overline{G}$ is almost surely highly clustered, provided that $E(\Delta) > 0$.

This bound is tight in the limit as the number of nodes $N \to \infty$ (Newman 2009). Indeed, by the Strong Law of Large Numbers (Theorem D.1), for almost every realization of the degree sequence $\bar{d}$, the empirical second moments of the components of the elements of $\bar{d}_N$ converge to the corresponding expected values as $N \to \infty$. Denote by $\mathcal{W}_s^{G_N}$ the set of ordered triangles of G_N that consists solely of single edges, i.e.

$$\mathcal{W}_s^{G_N} = \{(u, v, w) \in V_N^3 : \ (u, v), (u, w) \text{ and } (v, w) \text{ are single edges}\},$$

where V_N is the node set of G_N. To stress that the degree sequence $\bar{d}$ is regarded as given, we denote the underlying probability measure governing $\overline{G}$ by $P_{\bar{d}}(\cdot)$, and by $E_{\bar{d}}$ the corresponding conditional expectation. For the given degree sequence $\bar{d}$, the expected number of triangles of G_N formed by single edges is (approximately) given by (ignoring self-loops, multiple edges the small change in the set of free half-edges that arise as half-edges are paired)

$$E_{\bar{d}}\left(|\mathcal{W}_s^{G_N}|\right) = \sum_i \left(2\binom{S_i}{2} \sum_{j \neq i} \frac{S_j}{D_S^{(N)}} \left(\sum_{l \neq i,j} \frac{S_l}{D_S^{(N)}} \left(\frac{(S_j - 1)(S_l - 1)}{D_S^{(N)}} \right) \right) \right) \tag{2.3}$$

where the sums run over the integers $1, \ldots, N$. The term $2\binom{S_i}{2}$ arises because, for a node v_i with single degree S_i, there are $\binom{S_i}{2}$ ways to choose two of the S_i single half-edges attached to v_i. The terms $\frac{S_j}{D_S^{(N)}}$ and $\frac{S_l}{D_S^{(N)}}$ result from the fact that the probability that a node is chosen in the pairing procedure is proportional to its degree. The term $\frac{(S_j-1)(S_l-1)}{D_S^{(N)}}$ arises because the probability that the nodes v_j and v_l are neihgbours is approximately $\frac{(S_j-1)(S_l-1)}{D_S^{(N)}}$, given that they are both neighbours of v_i.

By the Strong Law of Large Numbers, dividing by N in (2.3) and letting N approach infinity gives

$$\frac{E_{\bar{d}}\left(|\mathcal{W}_S^{G_N}|\right)}{N} \to 0$$

as $N \to \infty$ for almost every realization of $\bar{d}$. Since convergence in mean implies convergence in probability

$$\frac{|\mathcal{W}_S^{G_N}|}{N} \to 0$$

in $P_{\bar{d}}$-measure as $N \to \infty$ for almost every realization of the degree sequence $\bar{d}$. Repeating this for triangles formed by a combination of triangle and single edges gives

$$C(G_N) \xrightarrow{P_{\bar{d}}} \frac{E(2\Delta)}{E((2\Delta+S)^2) - E(2\Delta+S)}. \tag{2.4}$$

That is, the asymptotic lower bound in 2.2 is attained in the limit as $N \to \infty$.

Remark 2.1. We may also regard the degree sequence $\bar{d}$ as random. By bounded convergence (Friedman 1982, Theorem 2.9.1) and the fact that $P_{\bar{d}}(\cdot) \leq 1$ the convergence in (2.4) holds also in the (unconditional) measure P as $N \to \infty$.

2.6 Branching processes

To analyse the spread of the disease in the early stages of the epidemic, we employ a multi-type branching process approximation. In the early phase of the epidemic, short cycles (except for the triangles formed by triangle half-edges) are unlikely to occur. For this reason, the early spread of the disease is well approximated by a suitably chosen branching process. In Appendix D, we construct a coupling of the epidemic process and a branching process and show that in the limit as the population size $N \to \infty$, the probability of a major outbreak is given by the probability that the approximating branching process avoids extinction. Furthermore, the expected fraction of the population that contracts the disease provided that a major outbreak occurs is given by the probability that the approximating *backward process* described in section 3.2 avoids extinction. A more extensive branching process framework is described in Appendix D.

In this project, we consider multi-type branching processes in discrete time. An s-type Galton-Watson process is a branching process in discrete time, involving s types of individuals. Let $\mathcal{T} = \mathbb{N}_0$ be the time index set. Each individual has a unit life length, and reproduces at age one, independently of the other individuals of the process. Let $\nu = (\nu_1, \ldots, \nu_s)$ be the distributions governing the reproduction of the individuals of the branching process. At age one, a type i individual is transformed into a finite number j_k of individuals of type k, $k = 1, \ldots, s$ with probability $\nu_i(\mathbf{j})$, where $\mathbf{j} = (j_1, \ldots, j_s)^\mathsf{T} \subset \mathbb{N}_0^s$. Denote the total number of type i individuals born at time n by $Z_{n,i}$, and let $\bar{Z}_n = (Z_{n,1}, \ldots, Z_{n,d})^\mathsf{T}$. We refer to n as the *generation* of $\bar{Z}_n$.

Denote the standard basis of $\mathbb{R}^s$ by $\bar{e}_1, \ldots, \bar{e}_s$. That is, $\bar{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^\mathsf{T}$ with the 1 in the ith component. We write $\bar{Z}_n^{(i)} = (Z_{n,1}^{(i)}, \ldots, Z_{n,d}^{(i)})^\mathsf{T}$ to indicate that the ancestor is of type i. That is to say, $\bar{Z}_0^{(i)} = \bar{e}_i$. A Galton-Watson process can be described as a Markov chain $Z = \{\bar{Z}_n;\ n \in \mathbb{N}_0\}$ on the state-space $\mathbb{N}_0^s$. The transition probability

$$P(\bar{Z}_{n+1} = (j_1, \ldots, j_s) | \bar{Z}_n = (i_1, \ldots, i_s))$$

of the Markov chain is the probability that the aggregated offspring vector of the $i_1 + \ldots + i_s$ individuals of generation n is $(j_1, \ldots, j_s)$.

We refer to the individuals of generation 0 as the *ancestors* of Z. In this project, we typically consider a branching process with a single ancestor of a different type than the other individuals of the branching process. If this is the case, the type of the ancestor is taken to be 0 and the type space is taken to be $\{0, 1 \ldots, s\}$. Although the number of types is $s + 1$, we refer to such branching process as an s-type branching process.

Throughout this section, Z denotes a s-type Galton-Watson process. The following theorems and definitions (Jagers 1975, Chapter 4) are central when analysing the asymptotic behaviour of branching processes.

Definition 2.2 (Mean matrix)**.** *The* mean matrix $M = (m_{i,j})_{i,j=1}^s$ *of* Z *is the matrix with entries* $m_{i,j} = E(Z_{1,j}^{(i)})$.

In the remainder of this section, we denote the mean matrix of Z by $M = (m_{i,j})_{i,j=1}^{s}$. Conditioning on the previous generation $n-1$ yields the following recursive expression for the expected number of individuals of generation n for $n \geq 1$ (Haccou et al. 2005)

$$E(\bar{Z}_n^{\mathsf{T}}) = E(\bar{Z}_{n-1}^{\mathsf{T}})M.$$

Thus

$$E(\bar{Z}_n^{\mathsf{T}}) = E(\bar{Z}_0^{\mathsf{T}})M^n \tag{2.5}$$

for $n \in \mathbb{N}$. If the ancestor is of a different type than the other individuals, the expected number of individuals of generation n is given by the slightly modified version of (2.5)

$$E(\bar{Z}_n^{\mathsf{T}}) = E(\bar{Z}_1^{\mathsf{T}})M^{n-1} \tag{2.6}$$

for $n \geq 1$, where $E(\bar{Z}_1^{\mathsf{T}})$ is the expected number of offspring produced by the ancestor.

Definition 2.3 (Positive regularity)**.** *The mean matrix M is said to be* positively regular *if all entries of M are finite and non-negative and if there exists some positive integer n such that all entries of M^n are strictly positive. If M is positively regular then Z is said to be* positively regular.

Theorem 2.1 (Perron-Frobenius theorem, cf. Jagers (1975) Therorem 4.2.1)**.** *Let M be positively regular. Then M has a real-valued eigenvalue $r > 0$ such that for any other eigenvalue λ of M it holds that $|\lambda| < r$, and there exists vectors $\bar{u}, \bar{v}$ with strictly positive coordinates such that*

$$M\bar{u} = \bar{u} \text{ and } \bar{v}^{\mathsf{T}}M = \bar{v}^{\mathsf{T}}.$$

If $\bar{u}$ and $\bar{v}$ are normalized, so that $\bar{u} \cdot \bar{1} = \bar{v} \cdot \bar{u} = 1$, then

$$\frac{1}{r^k}M^k \to \bar{u}\bar{v}^{\mathsf{T}}$$

as $k \to \infty$.

Definition 2.4 (Perron root)**.** *Let M and r be as in Theorem 2.1. The eigenvalue r is then called the Perron root of M.*

If the Perron root r of the mean matrix of Z is strictly less than 1 then Z is said to be *subcritical*, if r equals 1 then Z is said to be *critical* and if r is strictly larger than 1 then Z is said to be *supercritical*.

For two s-dimensional vectors $\bar{a} = (a_1, \ldots, a_s)^{\mathsf{T}}$ and $\bar{b} = (b_1, \ldots, b_s)^{\mathsf{T}}$, we define

$$\bar{a}^{\bar{b}} := a_1^{b_1} \cdot \ldots \cdot a_s^{b_s}.$$

Let $f : [0,1]^s \to \mathbb{R}^s$ be the probability generating function of the offspring distribution of the s types. That is, for $\bar{z} =\in [0,1]^s$, the ith component of the vector-valued function f evaluated at $\bar{z}$ is given by

$$(f(\bar{z}))_i = E\left(\bar{z}^{\bar{\xi}_i}\right)$$

where $\bar{\xi}_i = (\xi_{i,1}, \ldots, \xi_{i,s})$ is distributed as the offspring of a type i individual, $i = 1, 2, \ldots, s$.

The probability that a process started by a type i individual, $i = 1, \ldots, s$, goes extinct is given by q_i, where $\bar{q} = (q_1, q_2, \ldots, q_s)^\mathsf{T}$ is a solution of (Bellman and Kalaba 1967, Theorem 7.1)

$$\bar{q} = f(\bar{q}).$$

Definition 2.5 (Singular branching process)**.** *If the probability generating function f of the offspring distribution of Z is given by $f(\bar{z}) = M\bar{z}$ then Z is said to be* singular.

That is, Z is singular if almost surely each individual of Z gives birth to precisely one individual over the course of its life time (Athreya and Ney 1972, page 183).

Theorem 2.2 (Extinction probability)**.** *Let q_i be the probability of extinction of Z, conditioned on $\bar{Z}_0 = \bar{e}_i$. If Z is positively regular and non-singular then $\bar{q} = (q_1, \ldots, q_d)^\mathsf{T}$ is the unique solution of*

$$f(\bar{q}) = \bar{q}$$

that lies closest to the origin in the unit cube, where f is the probability generating function of the offspring distribution of Z. Moreover, if Z is critical or subcritical then $\bar{q} = \bar{1}$. If Z is supercritical then $q_i < 1$, $i = 1, \ldots, d$.

One can show that if Z is non-singular and positively regular then the extinction probability vector $\bar{q}$ and $\bar{1}$ are the only fixed points of f in the unit cube $[0, 1]^s$ (Athreya and Ney 1972, page 186).

For $n \in \mathbb{N}_0$, let $f_n : [0, 1]^s \to \mathbb{R}^s$ be the probability generating function of generation n in a branching process descending from a type $1, \ldots, s$ individual. That is, for $\bar{a} \in [0, 1]^s$ the ith component, $i = 1, \ldots, s$, of the vector-valued function f_n evaluated at $\bar{a}$ is given by

$$(f_n(\bar{a}))_i = E\left(\bar{a}^{\bar{Z}_n^{(i)}}\right). \tag{2.7}$$

Note that $f_0(\bar{z}) = \bar{z}$ and $f_1(\bar{z}) = f(\bar{z})$ for all $\bar{z} \in [0, 1]^s$. By conditioning on the individuals $\bar{Z}_{n-1}^{(i)}$, $i = 1, \ldots, s$, of the preceding generation $n - 1$ we obtain the following recursive expression for f_n

$$f_n(\bar{z}) = f(f_{n-1}(\bar{z})) \tag{2.8}$$

for $n \in \mathbb{N}$. This implies

$$f_n = f^{\circ n} \tag{2.9}$$

where $f^{\circ n}$ is the composition of f with itself n times, and $f^{\circ 0}$ is the identity operator. The event $\{\bar{Z}_n^{(i)} = 0\}$ that a branching process descending from an individual of type i has gone extinct in generation n is monotonically increasing in n and it is readily checked that

$$P(\bar{Z}_n^{(i)} = \bar{0}) = (f_n(\bar{0}))_i$$

for $i = 1, \ldots, s$. Since the sample space has measure $1 < \infty$ it follows that (Friedman 1982, Theorem 1.2.1)

$$\begin{aligned} q_i &= P\left(\cup_n \{\bar{Z}_n^{(i)} = \bar{0}\}\right) \\ &= \lim_{n\to\infty} P\left(\{\bar{Z}_n^{(i)} = \bar{0}\right) \\ &= \lim_{n\to\infty} (f_n(\bar{0}))_i \\ &= \lim_{n\to\infty} f^{\circ n}(\bar{0})_i. \end{aligned} \tag{2.10}$$

Thus, the extinction probabilities $(q_1, \ldots, q_s)$ can be approximated by repeated application of f until convergence.

In our analysis, the offspring distribution of the ancestor of Z typically differs from the offspring distribution of the s types, since we assume that the initial case is chosen uniformly at random (see section 3.1). Let $f^* : [0,1]^s \to \mathbb{R}$ be the probability generating function of the offspring distribution of the ancestor. If $\bar{\xi}_* = (\xi_{*,1}, \xi_{*,2}, \ldots, \xi_{*,s})$ is distributed as the offspring of the ancestor, then f^* is given by

$$f^*(\bar{z}) = E\left(\bar{z}^{\bar{\xi}_*}\right).$$

Since Z becomes extinct if and only if each of the processes started by the children of the ancestor dies out, the probability of extinction is given by

$$f^*(\bar{q}).$$

As the following theorem asserts, the asymptotic growth rate of a supercritical branching process Z is given by the Perron root of the corresponding mean matrix M if and only if the following inequality is satisfied

$$E(Z_{1,k}^{(i)} \log Z_{1,k}^{(i)}) < \infty \tag{2.11}$$

for $i, k = 1, \ldots, s$.

Theorem 2.3 (cf. Jagers (1975) Therorem 4.2.6)**.** *Let $Z^{(i)}$, $1 \le i \le s$, be a supercritical positively regular process. If Z satisfies (2.11) for $i, k = 1, \ldots, s$ then for each i there is a random variable W_i, $E(W_i) = u_i$, such that*

$$\frac{\bar{Z}_n^{(i)}}{r^n} \xrightarrow{a.s.} W_i \bar{v} \text{ as } n \to \infty,$$

where r is the Perron root of M and $\bar{u}$ and $\bar{v}$ are the eigenvectors corresponding to r. Moreover, $\{W_i = 0\} \stackrel{P}{=} \{\bar{Z}^{(i)} \to 0\}$.

If Z does not satisfy (2.11) then

$$\frac{\bar{Z}_n^{(i)}}{r^n} \xrightarrow{a.s.} \bar{0} \text{ as } n \to \infty,$$

Corollary 2.3.1 (cf. Jagers (1975) Corollary 4.2.7)**.** *Under the conditions of Theorem 2.3 it holds for $i, k = 1, \ldots, k$ that*

$$\frac{Z_{n,k}^{(i)}}{Z_{n,1}^{(i)} + \ldots + Z_{n,s}^{(i)}} \xrightarrow{a.s.} \frac{v_k}{v_1 + \ldots + v_s} \text{ as } n \to \infty$$

on $\{Z^{(i)} \not\to 0\}$.

Now assume that the ancestor of Z is of type 0 with offspring $\bar{\xi}^* = (\xi_1^*, \ldots, \xi_s^*)^\mathsf{T}$. From Theorem 2.3 we deduce that if $Z^{(i)}$ is supercritical and positively regular for $i = 1, \ldots, s$, then

$$\frac{\bar{Z}_n}{r^n} \xrightarrow{a.s.} W\bar{v} \text{ as } n \to \infty, \tag{2.12}$$

where W is distributed as the sum of ξ_1^* independent copies of W_1, ξ_2^* independent copies of W_2..., and ξ_s^* independent copies of W_s.

3 Branching process approximation

This section serves as an informal description of how branching processes can be used to approximate the spread of the disease in the early phase of the epidemic. This approximation enables us to obtain expressions for the probability that a major outbreak occurs. By reversing the direction of the edges in the directed graph representation of the epidemic, a similar technique can be used to obtain expressions for the expected final size of a major outbreak. Proofs of these results are available in Appendix D. This section draws heavily on Ball et al. (2009, 2014).

3.1 The early phase of the epidemic

As described in section 2.5, the graph G_N may be constructed by joining the half-edges in a random order. In particular, the graph G_N may be constructed (or explored) as the epidemic progresses. Starting with the initial infected case u^*, we sequentially join half-edges. An infected individual u is said to belong to generation n if the chain of infection from the initial case v_* to u consists of n edges. That is, the initial case v_*, which belongs to generation 0, transmits the disease to the individuals of generation 1. The individuals of generation 1 then transmit the disease to the individuals of generation 2 etc. The graph G_N is explored generation-wise. That is, each of the half-edges attached to an explored node of generation n is paired with some unpaired, or *free*, half-edge/half-edge pair chosen uniformly at random before the half-edges attached to nodes of subsequent generations are paired. For a epidemic in discrete time with unit infectious period, the individuals of generation n are the members of $\mathcal{I}(n)$. That is, the individuals of generation n are those who are infectious in the time step n.

If only a small part of the graph G_N is explored, short loops (except for the triangles formed by triangle edges) and multiple edges are unlikely to occur. This is illustrated in Figure 3.1. If the spread of the disease on a treelike graph such as the graph showed in Figure 3.1 can be described by a suitably chosen branching process, then we may use this branching process to approximate the spread of the disease in the early stages of the epidemic.

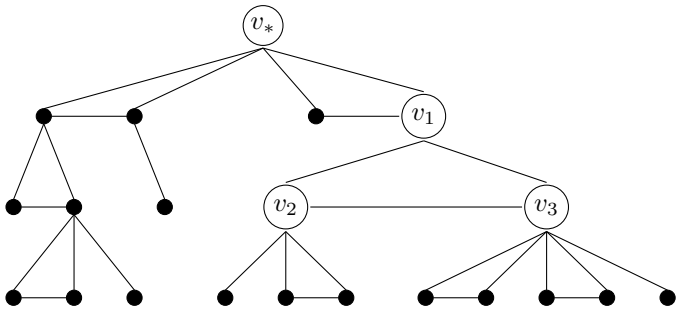

Figure 3.1: The local structure of G_N around the initial case v_* is treelike, except for the triangles formed by triangle edges. We may therefore ignore paths from v_* to v_2 and v_3 that do not go through v_1 in the early stage of the epidemic.

We denote the empirical degree distribution of G_N by $p^{(N)}$. Since half-edges are chosen uniformly at random, the probability to choose a specific node is proportional to the number of free half-edges attached to the node in question. That is, if we pair a single half-edge, the probability of choosing a specific node with k_S unpaired single half-edges is proportional to k_S. For this reason, the degree distribution a node explored by joining a single half-edge in the early phase of the epidemic is well approximated by the *single*

size biased empirical degree distribution $p_\circ^{(N,s)}$

$$p_\circ^{(N,s)}(k_\Delta, k_S) = \frac{k_S p^{(N)}(k_\Delta, k_S)}{E(S^{(N)})} \tag{3.1}$$

where $(\Delta^{(N)}, S^{(N)}) \sim p^{(N)}$.

Similarly, the degree distribution of the nodes explored by joining three triangle half-edge pairs in the early phase of the epidemic is close to the *triangle size biased* empirical degree distribution $p_\circ^{(N,\Delta)}$

$$p_\circ^{(N,\Delta)}(k_\Delta, k_S) = \frac{k_\Delta p^{(N)}(k_\Delta, k_S)}{E(\Delta^{(N)})}. \tag{3.2}$$

The *single size biased* degree distribution $p_\circ^{(s)}$ and the *triangle size biased* degree distribution $p_\circ^{(\Delta)}$ are defined analogously:

$$p_\circ^{(s)}(k_\Delta, k_S) = \frac{k_S p(k_\Delta, k_S)}{E(S)} \tag{3.3}$$

and

$$p_\circ^{(\Delta)}(k_\Delta, k_S) = \frac{k_\Delta p(k_\Delta, k_S)}{E(\Delta)}. \tag{3.4}$$

In the early stage of the epidemic, the degree distribution of the unexplored nodes is close to the empirical distribution $p^{(N)}$, and if the population size N is large, the empirical size biased degree distributions (3.1) and (3.2) are close to the size biased degree distributions (3.3) and (3.4). For this reason, we approximate the degree distributions of nodes infected along single and triangle edges by the size biased distributions (3.3) and (3.4), respectively.

To account for the fact that an infected individual has at least one non-susceptible neighbour (namely the source of its infection) we introduce the *downshifted* size biased degree distributions $p_\bullet^{(N,s)}$, $p_\bullet^{(N,\Delta)}$, $p_\bullet^{(s)}$ and $p_\bullet^{(\Delta)}$. To fix ideas, assume that the first half-edge to be paired is a single half-edge. The number of susceptible neighbours/triangle pairs of susceptible neighbours of the first node, w_1 say, explored by joining this first single half-edge has (approximately) the distribution $p_\circ^{(N,s)}(k_\Delta, k_S)$ shifted down by one degree in k_S, since the node that transmitted the disease to w_1 is not susceptible. That is, the distribution of the number of susceptible neighbours of w_1 is approximately

$$\sim 2\Delta_\circ^{(s)} + (S_\circ^{(s)} - 1)$$

where $(\Delta_\circ^{(s)}, S_\circ^{(s)})$ has the single size biased degree distribution $p_\circ^{(s)}$, provided that the population size N is large.

Similarly, the number of susceptible single neighbours/triangle pairs of susceptible neighbours of the first nodes explored by joining the first triangle half-edge pair has distribution $p_\circ^{(N,\Delta)}(k_\Delta, k_S)$ shifted down by one degree in k_Δ.

The downshifted size biased degree distributions $p_\bullet^{(N,s)}$, $p_\bullet^{(N,\Delta)}$, $p_\bullet^{(s)}$ and $p_\bullet^{(\Delta)}$ are given by

$$\begin{aligned}
p_\bullet^{(N,s)}(k_\Delta, k_S) &= p_\circ^{(N,s)}(k_\Delta, k_S + 1) \\
p_\bullet^{(N,\Delta)}(k_\Delta, k_S) &= p_\circ^{(N,\Delta)}(k_\Delta + 1, k_S) \\
p_\bullet^{(s)}(k_\Delta, k_S) &= p_\circ^{(s)}(k_\Delta, k_S + 1) \\
p_\bullet^{(\Delta)}(k_\Delta, k_S) &= p_\circ^{(\Delta)}(k_\Delta + 1, k_S).
\end{aligned} \tag{3.5}$$

Throughout, we will make frequent reference to the following random vectors

$$\begin{aligned}
(\Delta, S) &\sim p & (\Delta^{(N)}, S^{(N)}) &\sim p^{(N)} \\
(\Delta_\circ^{(s)}, S_\circ^{(s)}) &\sim p_\circ^{(s)} & (\Delta_\circ^{(N,s)}, S_\circ^{(N,s)}) &\sim p_\circ^{(N,s)} \\
(\Delta_\circ^{(\Delta)}, S_\circ^{(\Delta)}) &\sim p_\circ^{(\Delta)} & (\Delta_\circ^{(N,\Delta)}, S_\circ^{(N,\Delta)}) &\sim p_\circ^{(N,\Delta)} \\
(\Delta_\bullet^{(s)}, S_\bullet^{(s)}) &\sim p_\bullet^{(s)} & (\Delta_\bullet^{(N,s)}, S_\bullet^{(N,s)}) &\sim p_\bullet^{(N,s)} \\
(\Delta_\bullet^{(\Delta)}, S_\bullet^{(\Delta)}) &\sim p_\bullet^{(\Delta)} & (\Delta_\bullet^{(N,\Delta)}, S_\bullet^{(N,\Delta)}) &\sim p_\bullet^{(N,\Delta)}
\end{aligned} \tag{3.6}$$

and their expected values

$$\begin{aligned}
E(\Delta_\bullet^{(s)}) &= \frac{E(S\Delta)}{E(S)} & E(S_\bullet^{(s)}) &= \frac{E(S^2)}{E(S)} - 1 \\
E(\Delta_\bullet^{(\Delta)}) &= \frac{E(\Delta^2)}{E(\Delta)} - 1 & E(S_\bullet^{(\Delta)}) &= \frac{E(S\Delta)}{E(\Delta)}.
\end{aligned} \tag{3.7}$$

We approximate the distribution of the number of susceptible neighbours of a node infected along a triangle edge by $p_\bullet^{(N,\Delta)}$. Similarly, the distribution of the number of susceptible neighbours of a node infected along a single is approximated by $p_\bullet^{(N,s)}$ As the population size $N \to \infty$, the empirical downshifted size biased distributions converge almost surely to the corresponding downshifted size biased distributions by the Strong Law of Large Numbers (Theorem D.1) and assumption A1.

As half-edges are paired, the set of free half-edges changes. For this reason, the accuracy of the branching process approximation described above decreases as the epidemic propagates. By a birthday problem (see for instance Mosteller (1962)) type of argument, one can show that in the limit as the population size $N \to \infty$, the coupling breaks down when the epidemic reaches a size of order $\sqrt{N}$.

3.2 Susceptibility sets and backward processes

In section 2.3 we described the graph representation of an epidemic; an individual contracts the disease if and only if there is a path of finite length from the initial case to the node representing the individual in question. In the present section, we describe how a branching process approximation of the behaviours of the *susceptibility sets* of the nodes can be used to approximate the expected final size of the epidemic, provided that it does not go extinct at an early stage, see for instance Ball et al. (2009, 2010, 2014) and Miller (2007).

Definition 3.1 (Susceptibility sets)**.** *Let v be some node of the graph G_N. The* susceptibility set $\mathfrak{S}(v)$ *of v is the set of nodes v_* of G_N such that*

$$d(v_*, v) < \infty.$$

Figure (3.2) shows a schematic illustration of a susceptibility set. We sometimes write $\mathfrak{S}_N(v)$ to make the population size N explicit.

Let v be some node of G_N chosen uniformly at random. The individual represented by v contracts the disease if and only if the initial case v_* belongs to $\mathfrak{S}(v)$. In terms of

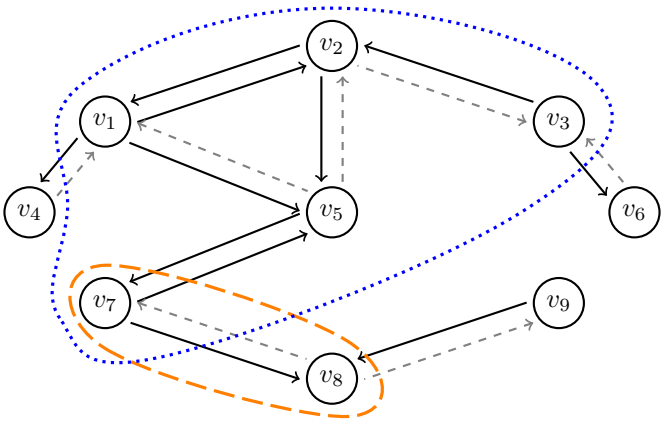

Figure 3.2: Graph representation of an epidemic in a small ($N = 9$) population. As in Figure 2.2, the grey dashed edges have infinite transmission weight. The nodes in the susceptibility set $\mathfrak{S}(v_5) = \{v_1.v_2, v_3, v_5, v_7\}$ of v_5 are enclosed by the blue dotted line. The nodes that v_5 would infect if infected, directly or through other nodes, are enclosed by the orange dashed line. The transmission weights are omitted for ease of presentation.

the graph representation of the epidemic, $\mathfrak{S}(v)$ consists of precisely the nodes that can be reached from v by tracing a path of finite length backwards. If the initial case v_* is chosen uniformly at random, then the probability that v is ultimately infected is the expected fraction of the nodes of G_N that are members of $\mathfrak{S}(v)$,

$$E\left(\frac{|\mathfrak{S}_N(v)|}{N}\right). \tag{3.8}$$

By exchangeability of the nodes, the expected value in (3.8) is also the expected proportion of the population ultimately infected.

By reversing the direction of the edges of the graph representation described in section 2.3, but keeping the weights, the expected final fraction of the population infected in a major outbreak and the probability of a major outbreak are interchanged (Miller 2008), provided that the initial case is chosen uniformly at random. The process so obtained is called the *backward epidemic process* of the node v. If the underlying epidemic model is such that the backward epidemic process can be described as a branching process, then we can use the techniques described in section 3.1 to compute the asymptotic distribution of the proportion $\frac{|\mathcal{S}_N(\infty)|}{N}$ of the population that ultimately escapes infection. This is made precise in the following theorem, due to Ball et al. (2014, Theorem 3.5), who proved the theorem for the related case of random intersection graphs.

Theorem 3.1. *Let q and q_b be the extinction probabilities of the forward and backward approximating branching processes, respectively, and let*

$$\frac{|\mathcal{S}_N(\infty)|}{N}$$

be the proportion of the population that ultimately escapes an epidemic in a population of size N. Then

$$\frac{|\mathcal{S}_N(\infty)|}{N} \xrightarrow{d} \mathfrak{E}$$

as $N \to \infty$ where $P(\mathfrak{E} = 1) = 1 - P(\mathfrak{E} = q_b) = q$.

In other words, in the limit of large population sizes, the epidemic "takes off" with probability $1 - q$, and if this happens a fraction $1 - q_b$ of the population is ultimately infected.

3.3 Geometric growth rate

We now turn our attention back to the definition of the basic reproduction number as the geometric growth rate. Let the graph sequence $\overline{G}$ and degree sequences $\bar{d}$ and $\bar{d}_N$ be as in section 2.5.2. That is, the graph G_N is generated by the configuration model with clustering based on the degree sequence $\bar{d}_N = \{(S_i, \Delta_i)\}_{i=1}^N$, where $\bar{d}_N$ consists of the first N elements of the infinite degree sequence $\bar{d} = \{(S_i, \Delta_i)\}_{N \in \mathbb{N}}$, and $\bar{d}$ is a sequence of independent copies of (S, Δ).

By the Strong Law of Large Numbers (Theorem D.1), for almost every realization of the degree sequence $\bar{d}$ the following regularity conditions are satisfied (cf. van der Hofstad (2016, Condition 7.8 a-c))

$C1)$ $p_\circ^{(N)} \to p_\circ$ as $N \to \infty$

$C2)$ $\frac{\sum_{i=1}^N S}{N} \to E(S)$ and $\frac{\sum_{i=1}^N \Delta_i}{N} \to E(\Delta)$ as $N \to \infty$

$C3)$ $\frac{\sum_{i=1}^N S_i^2}{N} \to E(S^2)$, $\frac{\sum_{i=1}^N \Delta_i^2}{N} \to E(\Delta^2)$ and $\frac{\sum_{i=1}^N S\Delta_i}{N} \to E(S\Delta)$ as $N \to \infty$

We consider epidemics on the graphs of $\overline{G}$. By the regularity conditions $C1$-$C3$ and the Perron-Frobenius Theorem, the limiting geometric growth rate as the population size $N \to \infty$ is given by the Perron root of the mean matrix M of the approximating branching process Z. Thus, the geometric growth rate has the threshold properties of R_0 for the limiting approximating branching process in the limit as $N \to \infty$. It follows from the proof presented in Appendix D that the geometric growth rate has the threshold properties of R_0 for the epidemic process in the limit as $N \to \infty$.

Recall that, for simpler models, the basic reproduction number is loosely defined as the expected number of infected cases caused by a "typical" infected individual in an otherwise susceptible population. By Corollary 2.3.1, the type composition of the branching process population converges to a stable, asymptotic composition as the generation n approaches infinity, provided that the branching process avoids extinction. That is, the asymptotic population is a combination of the types of the branching process, in proportions given by the left eigenvector $\bar{v}$ corresponding to the Perron root of the mean matrix M. By letting the population size N and the generation n of the epidemic process approach infinity in an appropriate manner, the type composition of the nth generation of the epidemic process converge in distribution to the type composition of the nth generation of the approximating branching process (see Appendix D for further details). Thus, the "typical case" of the epidemic can be interpreted as a weighted combination of the types of the approximating branching process, where the weights are proportional to the elements of the left eigenvector $\bar{v}$, and the Perron root has the interpretation of the average number of offspring produced by a typical individual.

4 An SIR epidemic in discrete time

We now have the tools to derive expressions for the probability of extinction and the expected final size of a major outbreak. Miller (2009) analysed an SIR epidemic on a configuration model graph with clustering under the assumption of homogeneous infectivity. In the present section, we extend the results presented by Miller (2009) by relaxing this assumption. A straightforward interpretation of the results presented in the present section is an SIR epidemic in continuous time with general infectious periods, see section 4.1 for further details.

We consider an SIR epidemic in discrete time (that is $\mathcal{T} = \{0, 1, 2, \ldots\}$) on the clustered graph G_N. We assume heterogeneity in infectivity. That is, some infected individuals are more contagious than others. To this end, let T be a random variable with support in $[0, 1]$, and let $\{T_i\}_{i=1}^N$ be independent random variables distributed as T. To each node v_i of G_N, we assign the *transmission weight* T_i of v_i. If v_i gets infected, then each susceptible neighbour of v_i gets infected independently in the next time step with probability T_i (conditioned on $\{T_i\}_i$). The node v_i thereafter becomes immune, playing no further role in the epidemic.

In other words, the marginal probability of transmission along an edge is $E(T)$, provided that the tail contracts the disease. An infected node transmits the disease independently of the transmissions from other infected nodes. An infected node does not, however, transmit the disease to its neighbours independently, unless the distribution of T is degenerate. Conditioned on the transmission weights $\{T_i\}_i$ and the structure of G_N, the number of neighbours that an infected node v_i makes contact with while infectious has a binomial distribution with parameters d_i and T_i, where d_i is the number of edges with tail v_i. We assume that susceptible individuals are fully susceptible, so that each contact with a susceptible individual made by an infectious individual results in transmission.

The spread of this epidemic can be fully captured by a directed graph (see section 2.3). To construct such directed graph from the undirected configuration model graph described in section 2.5, we replace each undirected edge of G_N by two parallel directed edges, pointing in the opposite direction. The weight of an edge (v_i, v_j) is taken to be 1 if v_i would make infectious contact with v_j if infected, and ∞ otherwise. That is, conditioned on T_i, the edge weights $\{d_{ij}\}_j$ are independent and have a two-point distribution.

If the population size N is large, the graph G_N is locally treelike, except for triangles formed by triangle edges. Therefore, for large N, short cycles (except for the triangles) are negligible in the early phase of the epidemic. For a given triangle u, v, w, where u is the first individual to be infected in the triangle u, v, w, we refer to v and w as *brothers*. The spread of the disease in the early phase may be approximated by a multi-type branching process consisting of the following three types (except for the initial case):

- Type I: A node infected along a triangle edge that does not have susceptible brother at the time point of infection
- Type II: A node infected along a triangle edge that has a susceptible brother at the time point of infection
- Type III: A node infected along a single edge

Figure 4.1 shows three examples of possible paths of transmission within a triangle giving rise to type I and II individuals in the approximating branching process.

Nodes represented by individuals of type I or II have the triangle size biased degree distribution $p_\circ^{(\Delta)}$ defined in (3.4). Similarly, a node represented by a type III individual has the single size biased degree distribution $p_\circ^{(s)}$ defined in (3.3) We assume that the initial case is chosen uniformly at random. The node representing the initial case then has degree distribution p. That is, the ancestor of the approximating branching process, which represents the initial case, is of a different type than the other individuals.

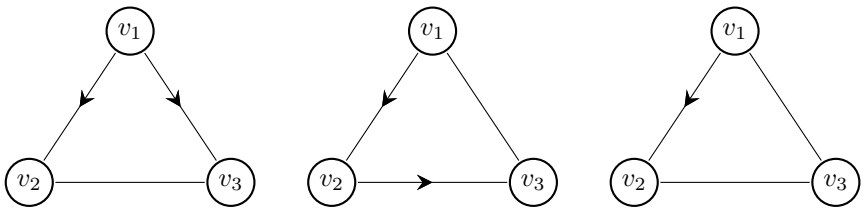

Figure 4.1: Three examples of possible paths of transmission in a triangle v_1, v_2, v_3, where v_1 is the first node to be infected. Left: v_1 infects both v_2 and v_3. Both v_2 and v_3 are represented by type I individuals in the approximating branching process. Center: v_1 infects v_2 and v_2 infects v_3. Then v_3 and v_2 are represented by type I and type II individuals, respectively. Right: v_1 infects v_2. Then v_2 is represented by a type II individual.

Denote by

$$M_f = (m_{ij})_{i,j=1}^3$$

the mean matrix of the above described branching process. Suppose that v_1 is the first individual to be infected in the triangle v_1, v_2, v_3. The probability that v_1 transmits the disease both to v_2 and v_3 is

$$E(T^2).$$

Similarly, the probability that v_1 transmits the disease to either v_2 or v_3, but not to both, is

$$2E(T(1-T)).$$

By linearity of expectation

$$M_f = \begin{pmatrix} 2E(T^2)E(\Delta_\bullet^{(\Delta)}) & 2E(T(1-T))E(\Delta_\bullet^{(\Delta)}) & E(T)E(S_\bullet^{(\Delta)}) \\ 2E(T^2)E(\Delta_\bullet^{(\Delta)}) + E(T) & 2E(T(1-T))E(\Delta_\bullet^{(\Delta)}) & E(T)E(S_\bullet^{(\Delta)}) \\ 2E(T^2)E(\Delta_\bullet^{(S)}) & 2E(T(1-T))E(\Delta_\bullet^{(S)}) & E(T)E(S_\bullet^{(S)}) \end{pmatrix} \tag{4.1}$$

since the distribution of the susceptible neighbours of infected nodes in the early phase of the epidemic is given by the downshifted degree distributions in (3.5). Recall that the random variables $\Delta_\bullet^{(\Delta)}$, $\Delta_\bullet^{(s)}$, $S\bullet^{(\Delta)}$ and $S\bullet^{(s)}$ defined in (3.6) have the downshifted size biased distributions. We note that all entries of M_f are finite and by assumption A1 that S and Δ both have finite second moments. This follows from the inequality $2|ab| \leq a^2 + b^2$, $a, b \in \mathbb{R}$.

With a little effort, one can use the expected values provided in (3.7) to show that necessary and sufficient conditions for M_f to be positively regular are that assumption A3 holds and that $0 < E(T) < 1$. If some of these conditions are not satisfied, we may analyse the spread of the disease by reducing the number of types of the approximating branching process to one or two.

We refrain from giving an explicit expression for the Perron root of M_f, since we do not expect it to provide further insight.

4.1 The rank based geometric growth rate

A straightforward interpretation of the above discussed results is an epidemic in continuous time with random infectious periods. Miller (2009) calculated a basic reproduction number by attributing all cases in a given triangle to the primary case in that triangle, regardless of the true paths of transmission. This means that, in the example of Figure 4.2, v_2 and v_3 are attributed to v_1, and w_2 and w_3 are attributed to w_1, despite the fact that v_1 and w_1 do not transmit the disease to v_3 and w_3, respectively. In this project, we calculate the *rank based geometric growth rate* (defined below).

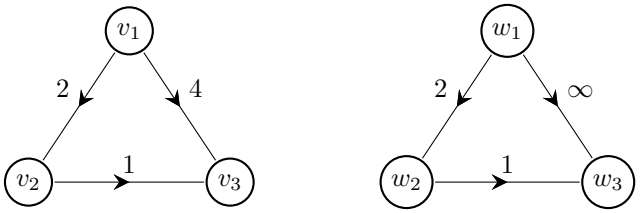

Figure 4.2: The difference between rank based generations, true generations and the method used by Miller (2009). Left: The length 4 of the path $v_1 \to v_3$ exceeds the length 3 of the path $v_1 \to v_2 \to v_3$. Therefore, the true path of transmission is $v_1 \to v_2 \to v_3$. In the rank based approach, however, v_3 is attributed to v_1. In the method proposed by Miller v_2 and v_3 are attributed to v_1. Right: The true path of transmission is $w_1 \to w_2 \to w_3$. In this case, the rank based generations and the true generations coincide. In the method used by Miller w_2 and w_3 are attributed to w_1.

Suppose that the infectious period is distributed as the random variable τ, $\tau \sim F$, and independent for different nodes. Suppose further that a node makes contact with each neighbour independently at a Poisson rate β while infected, and that susceptible individuals are fully susceptible, so that each infectious-susceptible contact results in transmission. Without loss of generality we assume $\beta = 1$, since we may rescale time (and F accordingly). The transmission weight T is then distributed as $1 - e^{-\tau}$.

Denote the initial case by v_*. The *rank* of a node v in G_N is the distance from v_* to v, if every edge along which the disease would be transmitted is assigned the edge weight 1, and every other edge is assigned the edge weight ∞. That is, the rank of v is the smallest number of directed edges that have to be traversed in order to follow a path of (potential) transmission from v_* to v. We may then calculate the rank based geometric growth rate, by letting generation n of the epidemic process consist of the individuals of rank n. If, for instance, v_1 is the first node in a triangle consisting of the nodes v_1, v_2, v_3 to be infected, and v_1 infects v_2 and thereafter attempts to infect v_3, then v_3 is attributed to v_1 regardless of whether v_1 or v_2 infected v_3. This is illustrated in Figure 4.2.

Let $\mathcal{L}(z) = \int_{\mathbb{R}_+} e^{-zx} dF(x)$ be the Laplace transform of the infectious period τ. Then $E(T) = 1 - \mathcal{L}(1)$ and $E(T(1-T)) = \mathcal{L}(1) - \mathcal{L}(2)$. The rank based geometric growth rate is obtained by substituting these identities into the mean matrix M_f in (4.1).

4.2 Probability of a major outbreak

Let $f : [0,1]^3 \to \mathbb{R}^3$ be the probability generating function of the offspring distribution of the three types. That is, for $\bar{z} = (z_1, z_2, z_3)^\mathsf{T} \in [0,1]^3$

$$f(\bar{z})^\mathsf{T} = (E(z_1^{\xi_{1,1}} z_2^{\xi_{1,2}} z_3^{\xi_{1,3}}), E(z_1^{\xi_{2,1}} z_2^{\xi_{2,2}} z_3^{\xi_{2,3}}), E(z_1^{\xi_{3,1}} z_2^{\xi_{3,2}} z_3^{\xi_{3,3}})) \tag{4.2}$$

where $(\xi_{i,1}, \xi_{i,2}, \xi_{i,3})$ is distributed as the offspring of a type i individual, $i = 1, 2, 3$.

Similarly, let $f^* : [0,1]^3 \to \mathbb{R}$ be the probability generating function of the offspring distribution of the initial case. If $(\xi_{*,1}, \xi_{*,2}, \xi_{*,3})$ is distributed as the offspring of the initial case, then f^* is given by

$$f^*(\bar{z})^{\mathsf{T}} = E(z_1^{\xi_{*,1}} z_2^{\xi_{*,2}} z_3^{\xi_{*,3}}).$$

For $i = 1, 2, 3$, let $(S^{(i)}, \Delta^{(i)})$ be distributed as the joint degree of a type i case with offspring $(\xi_{i,1}, \xi_{i,2}, \xi_{i,3})$. That is,

$$(S^{(1)}, \Delta^{(1)}) \stackrel{d}{=} (S^{(2)}, \Delta^{(2)}) \stackrel{d}{=} (S_\circ^{(\Delta)}, \Delta_\circ^{(\Delta)})$$

and

$$(S^{(3)}, \Delta^{(3)}) \stackrel{d}{=} (S_\circ^{(s)}, \Delta_\circ^{(s)}).$$

By conditional independence we have

$$E(z_1^{\xi_{i,1}} z_2^{\xi_{i,2}} z_3^{\xi_{i,3}}) = E(E(z_3^{\xi_{i,3}} | T, S^{(i)}, \Delta^{(i)}) E(z_1^{\xi_{i,1}} z_2^{\xi_{i,2}} | T, S^{(i)}, \Delta^{(i)})).$$

Conditioned on the transmission weight T and the single degree $S^{(1)}$, $\xi_{1,3}$ has a binomial distribution with parameters $S^{(1)}$ and T. Thus

$$\begin{aligned} E(z_3^{\xi_{1,3}} | T, S^{(1)}, \Delta^{(1)}) &= \sum_{k_0+k_1=S^{(1)}} \binom{S^{(1)}}{k_1} (Tz_3)^{k_1} (1-T)^{k_0} \\ &= (Tz_3 + 1 - T)^{S^{(1)}}. \end{aligned}$$

Similarly

$$\begin{aligned} E(z_1^{\xi_{1,1}} z_2^{\xi_{1,2}} | T, S^{(1)}, \Delta^{(1)}) &= \sum_{k_0+k_1+k_2=\Delta^{(1)}-1} \binom{\Delta^{(1)}-1}{k_0, k_1, k_2} (1-T)^{2k_0} (2(1-T)Tz_1)^{k_1} (Tz_2)^{2k_2} \\ &= ((1-T)^2 + 2T(1-T)z_1 + T^2 z_2^2)^{\Delta^{(1)}-1}. \end{aligned}$$

Thus

$$\begin{aligned} &E(z_1^{\xi_{1,1}} z_2^{\xi_{1,2}} z_3^{\xi_{1,3}}) \\ &= E((Tz_3 + 1 - T)^{S_\bullet^{(\Delta)}} ((1-T)^2 + 2T(1-T)z_1 + T^2 z_2^2)^{\Delta_\bullet^{(\Delta)}}) \end{aligned} \tag{4.3}$$

where $(\Delta_\bullet^{(\Delta)}, S_\bullet^{(\Delta)})$ is independent of T.

Since the conditional offspring distribution of a type II individual is identical to the offspring distribution of a type I individual except that a type II individual may give birth to one additional type I individual with probability T, we have

$$\begin{aligned} &E(z_1^{\xi_{2,1}} z_2^{\xi_{2,2}} z_3^{\xi_{2,3}}) \\ &= E((Tz_3 + 1 - T)^{S_\bullet^{(\Delta)}} ((1-T)^2 + 2T(1-T)z_1 + T^2 z_2^2)^{\Delta_\bullet^{(\Delta)}} (Tz_1 + 1 - T)). \end{aligned} \tag{4.4}$$

Similarly

$$E(z_1^{\xi_{3,1}} z_2^{\xi_{3,2}} z_3^{\xi_{3,3}}) = E((Tz_3 + 1 - T)^{S_\bullet^{(s)}}((1-T)^2 + 2T(1-T)z_1 + T^2 z_2^2)^{\Delta_\bullet^{(s)}}). \tag{4.5}$$

Substituting (4.3)-(4.5) into (4.2) gives an expression for f.

By Theorem 2.2, the extinction probability of a process descending from a type i individual, $i = 1, 2, 3$, is given by q_i, where $\bar{q} = (q_1, q_2, q_3)^\mathsf{T}$ is a solution of

$$\bar{q} = f(\bar{q}).$$

As we saw in section 2.6, the extinction probabilities $\bar{q} = (q_1, q_2, q_3)^\mathsf{T}$ are given by

$$\bar{q} = \lim_{n\to\infty} f^{\circ n}(\bar{0}). \tag{4.6}$$

Since the approximating branching process dies out if and only if each of the processes started by the children of the initial case dies out, the probability of extinction is given by

$$f^*(\bar{q}).$$

After some calculations, similar to the calculations that led to (4.3)-(4.5), we find that the probability of extinction is given by

$$f^*(\bar{q}) = E\left((Tq_3 + 1 - T)^S((1-T)^2 + 2T(1-T)q_1 + T^2 q_2^2)^\Delta\right)$$

where (S, Δ) is independent of T.

We conclude that the probability of a major outbreak is given by

$$1 - f^*(\bar{q}),$$

where $\bar{q}$ is the limit in (4.6).

4.3 The backward process

We now turn our attention to backward processes. Let w be a given node of G_N, chosen uniformly at random. We use the results presented in section 3.2 to approximate the probability that a w contracts the disease, which by an exchangeability argument equals the expected final size of a major outbreak. To this end, we classify each member v of the susceptibility set of the node w by the type of the first edge in the shortest (in terms of the rank based-distance) path from v to w. In the limit as the population size goes to infinity, the possibility that more than one path will attain the distance from v to w is negligible for short distances. The members of the susceptibility set are divided into the following two groups.

Type I: Included in the susceptibility set by virtue of potential transmission along a single edge

Type II: Included in the susceptibility set by virtue of potential transmission along a triangle edge

The offspring of an individual v in the backward process are the individuals that would have infected v, if infected. In the limit as $N \to \infty$, the degree of a node represented by a type I individual has the single size biased degree distribution $p_\circ^{(s)}$, and the degree of a node represented by a type II individual has the triangle size biased degree distribution $p_\circ^{(\Delta)}$.

We assign kinship as follows. The children of type I of an individual v are the individuals included in the susceptibility set due to potential transmission along a single edge. The children of type II of v are the individuals included in the susceptibility set due to potential transmission of the disease to v, within a triangle of which v is a member. We note that, given a triangle v, v_1, v_2, both v_1 and v_2 will be members of the susceptibility set of v by virtue of transmissions within the triangle if and only if at least one of the following events happens:

E_1) v_1 and v_2 both infects v

E_2) v_1 infects v and v_2 infects v_1

E_3) v_2 infects v and v_1 infects v_2

The events E_1-E_3 are illustrated in Figure (4.3).

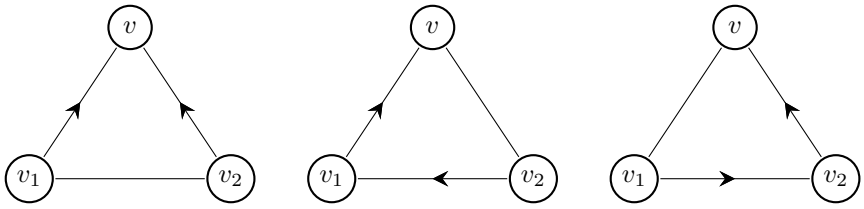

Figure 4.3: The individuals v_1 and v_2 are both in the susceptibility set $\mathfrak{S}(v)$ of v by virtue of transmission within the triangle v, v_1, v_2 if and only if at least one of the events E_1 (left), E_2 (center) or E_3 (right) happens.

Standard calculations gives that the probability of the union of the events E_1-E_3 is given by $p_2 = 3E(T)^2 - 2E(T)E(T^2)$. Similarly, the probability that neither v_1 nor v_2 will be members of the susceptibility set of v by transmissions within the triangle is given by $p_0 = (1 - E(T))^2$. Let $p_1 = 1 - p_0 - p_2$. The expected number of type II offspring produced by a type I individual is then given by

$$(2p_2 + p_1)E(\Delta_\bullet^{(S)}) = 2E(T)(1 + E(T) - E(T^2))E(\Delta_\bullet^{(S)})$$

and the expected number of type II offspring produced by a type II individual is

$$2E(T)(1 + E(T) - E(T^2))E(\Delta_\bullet^{(\Delta)}).$$

The probability that a single edge has finite edge weight is $E(T)$. Thus the expected number of type I individuals produced by a type I individual is given by

$$E(T)E(S_\bullet^{(S)}).$$

Similarly, the expected number of type I individuals produced by a type II individual is given by

$$E(T)E(S_\bullet^{(\Delta)}).$$

Combining these results yields the mean matrix

$$M_b = E(T)\begin{pmatrix} E(S_\bullet^{(S)}) & 2(1+E(T)-E(T^2))E(\Delta_\bullet^{(\Delta)}) \\ E(S_\bullet^{(\Delta)}) & 2(1+E(T)-E(T^2))E(\Delta_\bullet^{(\Delta)}) \end{pmatrix}.$$

If $E(T) > 0$ and assumption A3 holds then M_b is positively regular. We refrain from giving the explicit expression for the Perron root of M_b.

4.4 Expected final size of a major outbreak

Let b be the probability generating function of the offspring distribution of the two types of the approximating backward branching process. That is

$$b(\bar{z}) = (b_1(z_1,z_2), b_2(z_1,z_2)) = (E(z_1^{\xi^b_{1,1}} z_2^{\xi^b_{1,2}}), E(z_1^{\xi^b_{2,1}} z_2^{\xi^b_{2,2}}))$$

where $(\xi^b_{i,1}, \xi^b_{i,2})$ is distributed as the offspring of a type $i, i = 1, 2$, individual. Let further b^* be the probability generating function of the offspring distribution of the initial case w, i. e.

$$b^*(\bar{z}) = E(z_1^{\xi^b_{*,1}} z_2^{\xi^b_{*,2}})$$

where $(\xi^b_{*,1}, \xi^b_{*,2})$ is distributed as the offspring of the ancestor.

Analogously to the forward process, the probability that the bloodline started by a type i, $i = 1, 2$, individual will become extinct is given by q_i^b, where $\bar{q}_b = (q_1^b, q_2^b)^\mathsf{T}$ is the solution of

$$\bar{q}_b = b(\bar{q}_b)$$

that satisfies

$$\bar{q}_b = \lim_{n\to\infty} b^{\circ n}(\bar{0}). \tag{4.7}$$

The probability of extinction is then given by

$$b^*(\bar{q}_b).$$

Similar calculations as in Section 4.2 yields

$$(b(z_1,z_2))_1 = E((E(T)z_1 + 1 - E(T))^{S_\bullet^{(s)}}(p_0 + p_1 z_2 + p_2 z_2^2)^{\Delta_\bullet^{(s)}})$$

where p_1, p_2 and p_3 are as in section 4.3. Similarly

$$(b(z_1,z_2))_2 = E((E(T)z_1 + 1 - E(T))^{S_\bullet^{(\Delta)}}(p_0 + p_1 z_2 + p_2 z_2^2)^{\Delta_\bullet^{(\Delta)}})$$

The probability of ultimate extinction of the backward process is given by

$$b^*(\bar{q}_b) = E((E(T)q_1^b + 1 - E(T))^S(p_0 + p_1 q_2^b + p_2 (q_2^b)^2)^\Delta).$$

We conclude that the expected final size of a major outbreak is given by

$$1 - b^*(\bar{q}_b).$$

5 An SIR model with general infectious periods

In section 4.1, we provided expressions for the rank based geometric growth rate. In the case with general infectious periods where infected individuals makes contact at Poisson rate 1, the expected number of cases infected in the true generation n can be computed by classifying the infected individuals according to the following three types.

- Type I: A node infected along a triangle edge whose brother is not susceptible at the time point of infection
- Type II: A node infected along a triangle edge whose brother is susceptible at the time point of infection
- Type III: A node infected along a single edge

As before, let the infectious periods of the individuals of the process have distribution F, and assume that an individual make contact with each of its neighbours at rate $\beta = 1$. Let the random variable t_{ij} be the time elapsed between the time point $t_I^{(v_i)}$ at which the node v_i contracts the disease and the first time point after $t_I^{(v_i)}$ at which v_i makes contact with v_j. Then $\{t_{ij}\}_{i,j}$ is an sequence of independent exponentially distributed random variables with expected value 1. Let further the infectious periods $\{\tau_i\}_{i=1}^N$ be a sequence of independent and identically distributed random variables with distribution F, independent of the t_{ij}.

The approximating process with the above described types does not constitute a branching process in general, with the exception of fixed infectious periods. Indeed, assume that the infectious period has strictly positive variance, and that v_1 is the first to be infected in the triangle v_1, v_2, v_3, and that v_1 makes (infectious) contact with v_2 before making contact with v_3. The offspring distribution of v_2 depends on the infectious period of v_1, since v_1 and v_2 compete to transmit the disease to v_3. That is, the offspring distribution of v_2 depends on the number of children of v_2.

As described in section 3.1 (Figure 3.1), in the limit as the population size $N \to \infty$, we may use a branching process (with transmission weight $T = 1$ almost surely) to approximate the local structure of the graph G_N. The graph so obtained is locally treelike, except for the triangles formed by triangle edges. For this reason, we approximate the number of susceptible neighbours of a case by the downshifted triangle (types I and II) or single (type III) size biased degree distribution, depending on the type of the edge along which the disease was transmitted to the case in question.

Suppose that the node v_1 is the first node in the triangle v_1, v_2, v_3 to be infected. In the early stage of the epidemic, we may ignore the possibility of transmission of the disease to v_2 or v_3 from nodes that are not members of the triangle v_1, v_2, v_3. The length of the path $v_1 \to v_3$ is given by

$$L(v_1 \to v_3) = \infty \cdot \mathbb{1}(t_{1,3} > \tau_1) + t_{1,3}$$

and the length of the path $v_1 \to v_2 \to v_3$ is given by

$$L(v_1 \to v_2 \to v_3) = \infty \cdot \mathbb{1}(t_{1,2} > \tau_1) + \infty \cdot \mathbb{1}(t_{2,3} > \tau_2) + t_{1,2} + t_{2,3}.$$

Let $D^{\tau'}$ be the cumulative distribution function of the transmission time $L(w_1 \to w_2)$ from a node w_1 to its neighbour w_2, conditioned on the infectious period τ' of w_1. That is, for $0 \leq x < \infty$,

$$D^{\tau'}(x) = 1 - e^{-\min(x, \tau')},$$

and let

$$d^{\tau'}(x) = \mathbb{1}(0 \leq x \leq \tau')e^{-x}$$

be the (improper) density of $D^{\tau'}(x)$.

Now, conditioning on the infectious period τ_i of v_i, $i = 1, 2$ and the length t of the path $v_1 \to v_3$ yields

$$\begin{aligned}\vartheta :=& P(L(v_1 \to v_2 \to v_3) < L(v_1 \to v_3) < \infty) \\ =& \iint_{\mathbb{R}_+^2} \left(\int_{[0,\tau_1]} \left(\int_{[0,t)} d^{\tau_1} * d^{\tau_2}(s) ds \right) (dD^{\tau_1}(t)) \right) d(F \times F)(\tau_1, \tau_2)\end{aligned}$$

where $*$ denotes convolution. Standard but tedious calculations give

$$\vartheta = \iint_{\mathbb{R}_+^2} H(\tau_1, \tau_2) d(F \times F)(\tau_1, \tau_2)$$

where

$$H(\tau_1, \tau_2) = \begin{cases} \frac{1}{4}\left(e^{\tau_1} - 1\right) e^{-3\tau_1 - 2\tau_2} \left(4e^{\tau_1+\tau_2} - 4e^{\tau_1+2\tau_2} + e^{2(\tau_1+\tau_2)} - e^{2\tau_1} + 2e^{2\tau_2}\tau_2 \right) & \text{if } \tau_1 > \tau_2 > 0 \\ \frac{1}{4} e^{-2\tau_2} \left(2\tau_2 - 4e^{\tau_2} + e^{2\tau_2} + 3 \right) & \text{if } \tau_2 \geq \tau_1 > 0. \end{cases}$$

The probability that v_1 transmits the disease to both v_2 and v_3 (that is, that the paths $v_1 \to v_2$ and $v_1 \to v_3$ are both of finite length and that

$$\max(L(v_1 \to v_2), L(v_1 \to v_3)) < \min(L(v_1 \to v_2 \to v_3), L(v_1 \to v_3 \to v_2))$$

holds) is given by

$$P(L(v_1 \to v_2) < \infty, L(v_1 \to v_3) < \infty) - 2\vartheta = (1 - 2\mathcal{L}(1) + \mathcal{L}(2) - 2\vartheta). \tag{5.1}$$

Hence, the expected number of type I individuals infected by v_1 in the triangle v_1, v_2, v_3 is given by the expression in (5.1).

Similarly, the probability that v_2 transmits the disease to v_3 if v_2 is the first node in the triangle to which v_1 transmits the disease is given by

$$\begin{aligned}\Upsilon :=& P(L(v_1 \to v_2 \to v_3) < L(v_1 \to v_3) | L(v_1 \to v_2) < L(v_1 \to v_3)) \\ =& \frac{P(L(v_1 \to v_2 \to v_3) < L(v_1 \to v_3))}{P(L(v_1 \to v_2) < L(v_1 \to v_3))} \\ =& \frac{\vartheta + P(L(v_1 \to v_2 \to v_3) < L(v_1 \to v_3) = \infty)}{P(L(v_1 \to v_2) < L(v_1 \to v_3))} \\ =& \frac{\vartheta + E(e^{-\tau_1}(1 - e^{-\tau_1})(1 - e^{-\tau_2}))}{\frac{1}{2}(1 - \mathcal{L}(2))} \\ =& 2\frac{\vartheta + (1 - \mathcal{L}(2))(\mathcal{L}(1) - \mathcal{L}(2))}{(1 - \mathcal{L}(2))}.\end{aligned} \tag{5.2}$$

Using similar notation and terminology as for branching processes, let the elements of the vector $\bar{V}_n^{\mathsf{T}} = (V_{n,1}, V_{n,2}, V_{n,3})$ be the number of type I, II and III individuals of

generation n. By linearity of expectation and the fact that the infectious period τ_u and contact processes $\{C^{(u,v)}\}_v$ of an infectious individual u are independent of the path of transmission to the individual u, the expected number of type j individuals of generation n is given by

$$\sum_{i=1}^{3} E(V_{n-1,i})\tilde{m}_{i,j}, \tag{5.3}$$

In view of (5.1) and (5.2)

$$\tilde{m}_{1,1} = (1 - 2\mathcal{L}(1) + \mathcal{L}(2) - 2\vartheta)E(\Delta_\bullet^{(\Delta)})$$

and

$$\tilde{m}_{2,1} = \tilde{m}_{1,1} + \Upsilon.$$

Similarly, the expected number of type I individuals infected by a type III individual is given by

$$\tilde{m}_{3,1} = (1 - 2\mathcal{L}(1) + \mathcal{L}(2) - 2\vartheta)E(\Delta_\bullet^{(s)})$$

Proceeding in this fashion yields the mean matrix $\tilde{M} = (\tilde{m}_{i,j})_{i,j=1}^3$

$$\tilde{M} = \begin{pmatrix} A_1 E(\Delta_\bullet^{(\Delta)}) & A_2 E(\Delta_\bullet^{(\Delta)}) & A_3 E(S_\bullet^{(\Delta)}) \\ A_1 E(\Delta_\bullet^{(\Delta)}) + \Upsilon & A_2 E(\Delta_\bullet^{(\Delta)}) & A_3 E(S_\bullet^{(\Delta)}) \\ A_1 E(\Delta_\bullet^{(S)}) & A_2 E(\Delta_\bullet^{(S)}) & A_3 E(S_\bullet^{(S)}). \end{pmatrix} \tag{5.4}$$

where $A_1 = (1 - 2\mathcal{L}(1) + \mathcal{L}(2) - 2\vartheta)$, $A_2 = (1 - \mathcal{L}(2))$ and $A_3 = (1 - \mathcal{L}(1))$.

Analogously to the rank based case, by (5.3) the expected number of individuals of generation n is given by

$$E(\bar{V}_n^\intercal) = E(\bar{V}_1)^\intercal \tilde{M}^{n-1}. \tag{5.5}$$

The expected value in (5.5) is similar to the expression for the expected number of individuals of generation n of a branching process given in (2.6). We stress that the approximating process is not a branching process in general. The Perron root r of M_f is strictly smaller than 1 if and only if he Perron root $\tilde{r}$ of $\tilde{M}$ is strictly smaller than 1, where M_f is the mean matrix of the approximating branching process Z of Section 4. This follows from

$$E\left(\sum_{i=0}^{n} \bar{V}_i^\intercal \bar{1}\right) \leq E\left(\sum_{i=0}^{n} \bar{Z}_i^\intercal \bar{1}\right) \leq E\left(\sum_{i=0}^{2n} \bar{V}_i^\intercal \bar{1}\right) \tag{5.6}$$

and the fact that for a matrix M satisfying the Frobenius-Perron theorem 2.1 (note that by assumption A3, $\tilde{M}$ is positively regular if the infectios periods are not concentrated at 0)

$$\lim_{n\to\infty} \sum_{i=0}^{n} M^i \bar{1} < \infty$$

holds if and only if the corresponding Perron root is strictly less than 1. Note that inequalities are to be interpreted element-wise. Similarly, by sending n to infinity in (5.6) we see that the Perron root r of M_f is equal to 1 if and only if he Perron root $\tilde{r}$ of $\tilde{M}$ is is equal to 1. Thus, the threshold property of the Perron root of the mean matrix is retained, even though the process is not a branching process in general.

5.1 Infectious periods with finite support

If the infectious period has finite support $\mathcal{H}$, then the process may be analysed as a multi-type branching process $Z^{(\mathcal{H})}$. We may, for instance, let the type space be given by $\mathcal{H} \times \{1, 2, 3\}$. The first component of the type of an individual v is then given by the infectious period τ_v of v. The second component of the type of v indicates which one of the following three cases holds.

1. v is infected along a single edge
2. v is infected along a triangle edge, and v transmits the disease to its brother
3. v is infected along a triangle edge, and v does not transmits the disease to its brother.

Denote the mean matrix of this branching process by $M_{\mathcal{H}}$, and the corresponding Perron root by $r_{\mathcal{H}}$. We have that $\tilde{r} = r_{\mathcal{H}}$, where $\tilde{r}$ is the Perron root of the mean matrix in (5.4).

Indeed, this follows from the fact that the expected number of individuals of generation n is

$$E(Z_1^{(\mathcal{H})})^{\mathsf{T}} M_{\mathcal{H}}^{n-1} \bar{1} = E(\bar{V}_1)^{\mathsf{T}} \tilde{M}^{n-1} \bar{1},$$

hence

$$(\tilde{r})^n \asymp_n (r_{\mathcal{H}})^n.$$

We conclude that the Perron root $r_{\mathcal{H}}$ of $M_{\mathcal{H}}$ is given by $\tilde{r}$.

6 Vaccination

We now turn our attention to vaccination. To incorporate immunity stemming from vaccination, we extend the model investigated in section 4.

6.1 Random vaccination with a perfect vaccine

Assume that a fraction $f_{\mathrm{v}} < 1$ of the population is vaccinated, and that the vaccinated individuals are chosen uniformly at random (without replacement) from the population. The vaccine is perfect, in the sense that a vaccinated individual gains full and lasting immunity to the disease. If the population size N is large, we may use a slightly different model, where each individual is vaccinated with probability f_{v}, independently of the vaccination status of other individuals. By the Strong Law of Large Numbers, for our purposes the models are equivalent in the limit as the population size $N \to \infty$.

A before, we may approximate the early phase of the epidemic by a multi-type branching process. Assume that, in the early phase of the epidemic, v is the initial infected in the triangle v, v_1, v_2. If v attempts to transmit the disease both to v_1 and v_2 and succeeds (that is, none of v_1 and v_2 are vaccinated) then both v_1 and v_2 are represented by type I individuals in the approximating branching process. This happens with probability

$$E(T^2)(1-f_{\mathrm{v}})^2. \tag{6.1}$$

If v attempts to transmit the disease both to v_1 and v_2, but only succeeds to transmit the disease to v_2 (that is, v_1 is vaccinated and v_2 is not vaccinated) then in the approximating branching process, the individual representing v gives birth to one type I individual (representing v_2) within the triangle v, v_1, v_2. This happens with probability

$$E(T^2)f_{\mathrm{v}}(1-f_{\mathrm{v}}). \tag{6.2}$$

If v attempts to transmit the disease only to v_1 and succeeds (that is, v_1 is not vaccinated) then in the approximating branching process, the individual representing v gives birth to one type II individual (representing v_1) within the triangle v, v_1, v_2. This happens with probability

$$E(T(1-T))(1-f_{\mathrm{v}}). \tag{6.3}$$

The above described events are illustrated in Figure 6.1.

In summary, the individuals of the approximating branching process are of the following three types.

- Type I: Infected along triangle edge and has a brother that is guaranteed not to be susceptible
- Type II: Infected along triangle edge and has a brother that might be susceptible
- Type III: Infected along single edge

Denote the mean matrix of the approximating branching process by $M_f^{(\mathrm{v})} = (m_{i,j}^{(\mathrm{v})})_{i,j=1}^3$. Using the expressions in (6.1) and (6.2) gives the expected number of type I individuals produced by a type I individual

$$\begin{aligned} m_{1,1}^{(\mathrm{v})} &= (2(1-f_{\mathrm{v}})^2 E(T^2) + 2(1-f_{\mathrm{v}})f_{\mathrm{v}}E(T^2))E(\Delta_\bullet^{(\Delta)}) \\ &= (1-f_{\mathrm{v}})2E(T^2)E(\Delta_\bullet^{(\Delta)}) \\ &= (1-f_{\mathrm{v}})m_{1,1} \end{aligned}$$

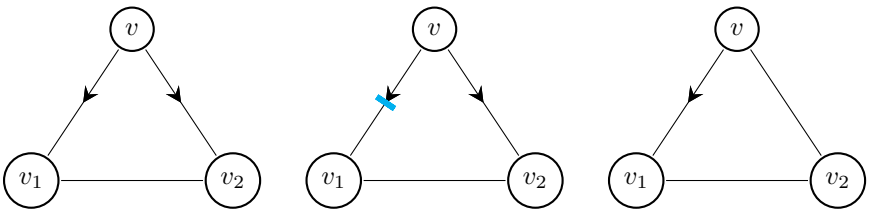

Figure 6.1: Three examples of transmission dynamics within a triangle v, v_1, v_2. An attempted transmission of the disease is represented by an arrow, an attempted transmission to a vaccinated individual is represented by an arrow and a blue bar. Left: v attempts to transmit the disease both to v_1 and v_2, and succeeds. Both v_1 and v_2 are represented by type I individuals in the approximating branching process. Center: v attempts to transmit the disease both to v_1 and v_2, the transmission to v_1 is blocked since v_1 is vaccinated. Then v_2 is represented by a type I individual. Right: v succeeds to transmit the disease to v_1, but does not attempt to infect v_2. Then v_1 is represented by a type II individual.

where $m_{1,1}$ is an element of he mean matrix M_f of the forward process presented in section 4.

Similarly, the expected number of type II offspring produced by a type I individual is given by

$$\begin{aligned} m_{1,2}^{(\mathrm{v})} &= 2(1-f_{\mathrm{v}})E(T(1-T))E(\Delta_{\bullet}^{(\Delta)}) \\ &= (1-f_{\mathrm{v}})m_{1,2}. \end{aligned}$$

Proceeding in the same fashion, we obtain the elements of the mean matrix $M_f^{(\mathrm{v})} = (m_{i,j}^{(\mathrm{v})})_{i,j=1}^3$ of the branching process with random vaccination. It turns out that

$$M_f^{(\mathrm{v})} = (1-f_{\mathrm{v}})M_f.$$

It is readily verified that the Perron root of $M_f^{(\mathrm{v})}$ is

$$r_f^{(\mathrm{v})} = (1-f_{\mathrm{v}})r_f, \tag{6.4}$$

where r_f is the Perron root of M_f. Setting $r_f^{(\mathrm{v})}$ to 1 in (6.4) and solving for f_{v} yields the *critical vaccination coverage*

$$f_{\mathrm{v}}^{(c)} = 1 - \frac{1}{r_f}.$$

The critical vaccination coverage $f_{\mathrm{v}}^{(c)}$ is defined as the fraction of the population necessary to vaccinate in order to be guaranteed to prevent a major outbreak (Britton 2010). That is, if a fraction $f_{\mathrm{v}}^{(c)}$ is vaccinated then the probability of a major outbreak is zero. We conclude that, for this particular graph model, equality holds between the basic reproduction number R_0 and the prefect vaccine-associated reproduction number R_V.

6.1.1 Probability of a major outbreak

Let $h : [0,1]^3 \to \mathbb{R}^3$ be the probability generating function of the offspring distribution of the three types. As in section 4.2, we use the probability generating function to approximate the probability of extinction of the epidemic. To this end, let $(\zeta_{i,1}, \zeta_{i,2}, \zeta_{i,3})$ be distributed as the offspring of a type i individual with transmission weight T, $i = 1, 2, 3$, and let $(S^{(i)}, \Delta^{(i)})$ be distributed as the joint degree of this individual. That is,

$$(S^{(1)}, \Delta^{(1)}) \stackrel{d}{=} (S^{(2)}, \Delta^{(2)}) \stackrel{d}{=} (S_\circ^{(\Delta)}, \Delta_\circ^{(\Delta)})$$

and

$$(S^{(3)}, \Delta^{(3)}) \stackrel{d}{=} (S_\circ^{(s)}, \Delta_\circ^{(s)}).$$

We assume that $(S^{(i)}, \Delta^{(i)})$ and T are independent.

By conditional independence

$$E\left(z_1^{\zeta_{1,1}} z_2^{\zeta_{1,2}} z_3^{\zeta_{1,3}}\right) = E\left(E\left(z_3^{\zeta_{1,3}} | S^{(1)}, \Delta^{(1)}, T\right) E\left(z_1^{\zeta_{1,1}} z_2^{\zeta_{1,2}} | S^{(1)}, \Delta^{(1)}, T\right)\right)$$

for $\bar{z} = (z_1, z_2, z_3)^\mathsf{T} \in [0,1]^3$.

Conditioned on the transmission weight T and the joint degree $(S^{(1)}, \Delta^{(1)})$, the number of attempted transmissions from a type I individual along single edges has a binomial distribution with parameters $S^{(1)}$ and T, and each attempted transmission succeeds with probability $(1 - f_\mathrm{v})$. Thus

$$\begin{aligned} E\left(z_3^{\zeta_{1,3}} | S^{(1)}, \Delta^{(1)}, T\right) &= \sum_{k_0+k_1=S^{(1)}} \binom{S^{(1)}}{k_0, k_1} z_3^{k_1} \big(T(1-f_\mathrm{v})\big)^{k_1} \big((1-T) + T f_\mathrm{v}\big)^{k_0} \\ &= \big(T(1-f_\mathrm{v}) z_3 + 1 - T + T f_\mathrm{v}\big)^{S^{(1)}}. \end{aligned} \tag{6.5}$$

Similarly, for a type I individual v with triangle degree $\Delta^{(1)}$, by conditioning on the number of attempted transmissions (in k_i of the $\Delta^{(1)} - 1$ triangles that is not yet affected by the disease, v attempts to transmit the disease to i individuals, $i = 0, 1, 2$) and the vaccination status of the individuals contacted by v we obtain

$$
\begin{aligned}
&E(z_1^{\zeta_{1,1}} z_2^{\zeta_{1,2}} | S^{(1)}, \Delta^{(1)}, T) \\
&= \sum_{k_0+k_1+k_2=\Delta^{(1)}-1} \binom{\Delta^{(1)}-1}{k_0, k_1, k_2} (1-T)^{2k_0} \big(2T(1-T)\big)^{k_1} T^{2k_2} \\
&\qquad \left(\sum_{\tilde{k}_0+\tilde{k}_1+\tilde{k}_2=k_2} \binom{k_2}{\tilde{k}_0, \tilde{k}_1, \tilde{k}_2} \big((1-f_{\mathrm{v}})z_1\big)^{2\tilde{k}_2} \big(2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1\big)^{\tilde{k}_1} f_{\mathrm{v}}^{2\tilde{k}_0} \right) \\
&\qquad \left(\sum_{k_0'+k_1'=k_1} \binom{k_1}{k_0', k_1'} (1-f_{\mathrm{v}})^{k_1'} z_2^{k_1'} f_{\mathrm{v}}^{k_0'} \right) \\
&= \sum_{k_0+k_1+k_2=\Delta^{(1)}-1} \binom{\Delta^{(1)}-1}{k_0, k_1, k_2} (1-T)^{2k_0} \big(2T(1-T)\big)^{k_1} T^{2k_2} \\
&\qquad \big(\big((1-f_{\mathrm{v}})z_1\big)^2 + 2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1 + f_{\mathrm{v}}^2\big)^{k_2} \\
&\qquad \big((1-f_{\mathrm{v}})z_2 + f_{\mathrm{v}}\big)^{k_1} \\
&= \Big((1-T)^2 + 2T(1-T)\big((1-f_{\mathrm{v}})z_2 + f_{\mathrm{v}}\big) \\
&\qquad + T^2\big(\big((1-f_{\mathrm{v}})z_1\big)^2 + 2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1 + f_{\mathrm{v}}^2\big)\Big)^{\Delta^{(1)}-1}.
\end{aligned} \tag{6.6}
$$

Combining (6.5) and (6.6) yields

$$
\begin{aligned}
E\left(z_1^{\zeta_{1,1}} z_2^{\zeta_{1,2}} z_3^{\zeta_{1,3}}\right) = E\Big[&\Big(T(1-f_{\mathrm{v}})z_3 + 1 - T + Tf_{\mathrm{v}}\Big)^{S_\bullet^{(\Delta)}} \\
&\Big((1-T)^2 + 2T(1-T)\big((1-f_{\mathrm{v}})z_2 + f_{\mathrm{v}}\big) \\
&\quad + T^2\big(\big((1-f_{\mathrm{v}})z_1\big)^2 + 2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1 + f_{\mathrm{v}}^2\big)\Big)^{\Delta_\bullet^{(\Delta)}}\Big].
\end{aligned} \tag{6.7}
$$

By noting that the offspring distribution of a type II individual is identical to the offspring distribution of a type I individual, except that a type II may give birth to one additional type I individual with probability $T(1-f_{\mathrm{v}})$

$$
\begin{aligned}
E\left(z_1^{\zeta_{2,1}} z_2^{\zeta_{2,2}} z_3^{\zeta_{2,3}}\right) = E\Big[&\Big(T(1-f_{\mathrm{v}})z_3 + 1 - T + Tf_{\mathrm{v}}\Big)^{S_\bullet^{(\Delta)}} \\
&\Big((1-T)^2 + 2T(1-T)\big((1-f_{\mathrm{v}})z_2 + f_{\mathrm{v}}\big) \\
&\quad + T^2\big(\big((1-f_{\mathrm{v}})z_1\big)^2 + 2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1 + f_{\mathrm{v}}^2\big)\Big)^{\Delta_\bullet^{(\Delta)}} \\
&\Big(z_1 T(1-f_{\mathrm{v}}) + 1 - T(1-f_{\mathrm{v}})\Big)\Big].
\end{aligned} \tag{6.8}
$$

Similarly

$$E\left(z_1^{\zeta_{3,1}} z_2^{\zeta_{3,2}} z_3^{\zeta_{3,3}}\right) = E\Big[\Big(T(1-f_{\mathrm{v}})z_3 + 1 - T + Tf_{\mathrm{v}}\Big)^{S_\bullet^{(s)}} \Big((1-T)^2 + 2T(1-T)\big((1-f_{\mathrm{v}})z_2 + f_{\mathrm{v}}\big) + T^2\big(\big((1-f_{\mathrm{v}})z_1\big)^2 + 2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1 + f_{\mathrm{v}}^2\big)\Big)^{\Delta_\bullet^{(s)}}\Big]. \tag{6.9}$$

Combining these results yields the probability generating function h of the offspring distribution of a type I, II, III individual respectively. That is, $h(\bar{z})_1$ is given by (6.7), , $h(\bar{z})_2$ is given by (6.8) and $h(\bar{z})_3$ is given by (6.9).

The probability generating function h^* of the initial case is given by

$$\begin{aligned} h^*(\bar{z}) =& E(z_1^{\zeta_{*,1}} z_2^{\zeta_{*,2}} z_3^{\zeta_{*,3}}) \\ =& E\Big[\Big(T(1-f_{\mathrm{v}})z_3 + 1 - T + Tf_{\mathrm{v}}\Big)^{S} \\ & \Big((1-T)^2 + 2T(1-T)\big((1-f_{\mathrm{v}})z_2 + f_{\mathrm{v}}\big) \\ & + T^2\big(\big((1-f_{\mathrm{v}})z_1\big)^2 + 2f_{\mathrm{v}}(1-f_{\mathrm{v}})z_1 + f_{\mathrm{v}}^2\big)\Big)^{\Delta}\Big]. \end{aligned} \tag{6.10}$$

for $\bar{z} = (z_1, z_2, z_3)^{\mathsf{T}} \in [0,1]^3$, where (S, Δ) is distributed as the joint degree of the initial case and independent of T. The probability of extinction of the approximating branching process is given by

$$h^*(\bar{q}^{\,(\mathrm{v})}),$$

where $\bar{q}^{\,(\mathrm{v})}$ is given by the point in $[0,1]^3$ closest to the origin that satisfies

$$\bar{q}^{\,(\mathrm{v})} = h(\bar{q}^{\,(\mathrm{v})}).$$

Thus, the probability of a major outbreak is

$$1 - h^*(\bar{q}^{\,(\mathrm{v})}).$$

6.1.2 The backward process

We now turn our attention to the backward process and final size of an epidemic in a population where a fraction f_{v} is vaccinated with a perfect vaccine. To this end, we introduce the following three types, where individuals are classified by their vaccination status and the type of the edge along which they would transmit the disease, if infected.

- Type I: Transmits along single edge, no information on vaccination status is available
- Type II: Transmits along triangle edge and is guaranteed not to be vaccinated
- Type III: Transmits along triangle edge, no information on vaccination status is available

Let v, v_1, v_2 be a given triangle. At least one of v_1 and v_2 belongs to the susceptibility set of v by virtue of potential transmissions within the triangle if and only if some the following events, illustrated in Figure 6.2, happens. Note that all cases infected by virtue of transmission within the triangle v, v_1, v_2 are attributed to v.

E_1) v_1 attempts to infect v and v_2 attempts to infect v_1, both succeed, and v_2 does not attempt to infect v. Or the same thing might happen, with v_1 and v_2 interchanged. This results in one type II and one type III individual in the approximating branching process. If v is represented by a type I or III individual this happens with probability

$$2\big(1-f_{\mathrm{v}}\big)^2 E(T)E\big(T(1-T)\big),$$

if v is represented by an individual of type II this happens with probability

$$2(1-f_{\mathrm{v}})E(T)E\big(T(1-T)\big).$$

E_2) Only one of v_1 and v_2 attempts to infect v, and succeeds. The other node does not attempt to infect any node within the triangle. This results in one type III offspring. If v is represented by an individual of type I or III this happens with probability

$$2(1-f_{\mathrm{v}})E(T)E\big(T(1-T)\big),$$

if v is represented by an individual of type II this happens with probability

$$2E(T)E\big(T(1-T)\big).$$

E_3) v_1 and v_2 both attempt to infect v and succeeds. This results in two type III individuals born in the approximating branching process. If v is represented by an individual of type I or III this happens with probability

$$(1-f_{\mathrm{v}})E(T^2),$$

if v is represented by an individual of type II this happens with probability

$$E(T^2).$$

E_4) v_1 attempts to infect v and succeeds. The other node, v_2, attempts to infect v_1, but fails due to v_1 being vaccinated. The individual v_2 does not attempt to infect v. In this scenario, v_1 belongs to the susceptibility set of v. However, we choose not to include v_1 is the approximating branching process. This does not have any impact on the result of our analysis, since we are only interested in the probability of extinction of the backward process.

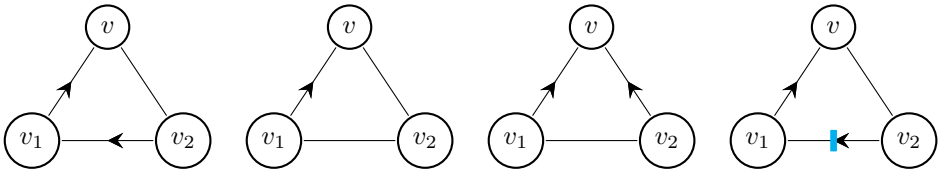

Figure 6.2: At least one of v_1 and v_2 will belong to the susceptibility set of v by virtue of potential transmissions within the triangle if and only if some of the following types of scenarios (left to right in the picture) occur: E_1, E_2, E_3, E_4. An attempted transmission of the disease is represented by an arrow, an attempted transmission to a vaccinated individual is represented by an arrow and a blue bar.

6.1.3 Expected final size

Let $b^{(\mathrm{v})}$ be the probability generating function of the offspring distribution of the three types of the approximating backward branching process. That is

$$b^{(\mathrm{v})}(\bar{z})_i = E(z_1^{\zeta^b_{i,1}} z_2^{\zeta^b_{i,2}} z_3^{\zeta^b_{i,3}})$$

where $\bar{\zeta}_i = (\zeta^b_{i,1}, \zeta^b_{i,2}, \zeta^b_{i,3})$ is distributed as the offspring of a type $i, i = 1, 2, 3$, individual. For $i = 1, 3$, denote by E_s expectation conditioned on that the parent of $(\zeta^b_{i,1}, \zeta^b_{i,2}, \zeta^b_{i,3})$ is not vaccinated. Let further $b^{(\mathrm{v})}_*$ be the probability generating function of the offspring distribution of the initial case, that is

$$b^{(\mathrm{v})}_*(\bar{z}) = E\left(z_1^{\zeta^b_{*,1}} z_2^{\zeta^b_{*,2}} z_3^{\zeta^b_{*,3}}\right)$$

where $(\zeta^b_{*,1}, \zeta^b_{*,2}, \zeta^b_{*,3})$ is distributed as the offspring of the ancestor.

Analogously to the forward process, the probability that the bloodline started by a type i, $i = 1, 2, 3$, individual will go extinct is given by q^b_i, where $\bar{q}_b = (q^b_1, q^b_2, q^b_3)^\mathsf{T}$ is the solution of

$$\bar{q}_b = b^{(\mathrm{v})}(\bar{q}_b)$$

in $[0, 1]^3$ closest to the origin. The probability of ultimate extinction of the backward process is then given by

$$b^{(\mathrm{v})}_*(\bar{q}_b). \tag{6.11}$$

To find an expression for $(b^{(\mathrm{v})})_1$, we note that for $\bar{z} = (z_1, z_2, z_3)^\mathsf{T}$

$$E\left(\bar{z}^{\bar{\zeta}_1}\right) = f_\mathrm{v} + (1 - f_\mathrm{v}) E_s\left(E_s\left(z_1^{\zeta^b_{1,1}} | S^{(1)}, \Delta^{(1)}\right) E_s\left(z_3^{\zeta^b_{1,3}} z_2^{\zeta^b_{1,2}} | S^{(1)}, \Delta^{(1)}\right)\right) \tag{6.12}$$

where, as before, $(S^{(i)}, \Delta^{(i)})$, is distributed as the joint degree of a type i individual, $i = 1, 2, 3$.

Now

$$\begin{aligned} E_s\left(z_1^{\zeta_{1,1}} | S^{(1)}, \Delta^{(1)}\right) &= \sum_{k_0 + k_1 = S^{(1)} - 1} \binom{S^{(1)} - 1}{k_0, k_1} z_1^{k_1} E(T)^{k_1} E(1 - T)^{k_0} \\ &= \big(E(T) z_1 + 1 - E(T)\big)^{S^{(1)} - 1}. \end{aligned} \tag{6.13}$$

By conditioning on the number of triangles k_2 in which an event of type E_3 occurs, the number of triangles k^a_1 in which an event of type E_1 occurs, the number of triangles k^b_1 in which an event of type E_4 occurs and the number of triangles k^c_1 in which an event of

type E_2 occurs we obtain

$$
\begin{aligned}
E_s(z_2^{\zeta_{1,2}} z_3^{\zeta_{1,3}} | S^{(1)}, \Delta^{(1)}) = \sum_{k_0+k_1^a+k_1^b+k_1^c+k_2=\Delta^{(1)}} & \binom{\Delta^{(1)}}{k_0, k_1^a, k_1^b, k_1^c, k_2} E(1-T)^{2k_0} \\
& \Big(2E(T)E\big(T(1-T)\big)(1-f_{\mathrm{v}})\Big)^{k_1^a} \\
& \Big(2E(T)E\big(T(1-T)\big)f_{\mathrm{v}}\Big)^{k_1^b} \Big(2E(T)E\big((1-T)^2\big)\Big)^{k_1^c} \\
& E(T)^{2k_2} z_2^{k_1^a} z_3^{k_1^a+k_1^c+2k_2} \\
= \Big(\big(E(1-T)\big)^2 + 2E\big(T\big)E\big(T(1-T)\big)(1-f_{\mathrm{v}}) z_2 z_3 & + 2E(T)E\big(T(1-T)\big)f_{\mathrm{v}} \\
& + 2E(T)E\big((1-T)^2\big) z_3 + E(T)^2 z_3^2\Big)^{\Delta^{(1)}}.
\end{aligned}
\tag{6.14}
$$

Inserting the right hand sides of (6.13) and (6.14) in (6.12) gives

$$
\begin{aligned}
E(z_1^{\zeta_{1,1}} z_2^{\zeta_{1,2}} z_3^{\zeta_{1,3}}) = f_{\mathrm{v}} + (1-f_{\mathrm{v}}) E\Big[& \Big(E(T)z_1 + 1 - E(T)\Big)^{S_\bullet^{(s)}} \\
& \Big(\big(E(1-T)\big)^2 + 2E(T)E\big(T(1-T)\big)(1-f_{\mathrm{v}}) z_2 z_3 \\
& + 2E(T)E\big(T(1-T)\big)f_{\mathrm{v}} \\
& + 2E(T)E\big((1-T)^2\big) z_3 + E(T)^2 z_3^2\Big)^{\Delta_\bullet^{(s)}}\Big].
\end{aligned}
\tag{6.15}
$$

Similarly

$$
\begin{aligned}
E(z_1^{\zeta_{2,1}} z_2^{\zeta_{2,2}} z_3^{\zeta_{2,3}}) = E\Big[& \Big(E(T)z_1 + 1 - E(T)\Big)^{S_\bullet^{(\Delta)}} \\
& \Big(\big(E(1-T)\big)^2 + 2E(T)E\big(T(1-T)\big)(1-f_{\mathrm{v}}) z_2 z_3 \\
& + 2E(T)E\big(T(1-T)\big)f_{\mathrm{v}} \\
& + 2E(T)E\big((1-T)^2\big) z_3 + E(T)^2 z_3^2\Big)^{\Delta_\bullet^{(\Delta)}}\Big].
\end{aligned}
\tag{6.16}
$$

and

$$
E(z_1^{\zeta_{3,1}} z_2^{\zeta_{3,2}} z_3^{\zeta_{3,3}}) = f_{\mathrm{v}} + (1-f_{\mathrm{v}}) E(z_1^{\zeta_{2,1}} z_2^{\zeta_{2,2}} z_3^{\zeta_{2,3}}). \tag{6.17}
$$

Combining these results yields the probability generating function of the offspring distribution of the three types; $(b^{(\mathrm{v})}(\bar{z}))_1$ is given by (6.15) and $(b^{(\mathrm{v})}(\bar{z}))_2$ is given by (6.16). By replacing $(S_\bullet^{(s)}, \Delta_\bullet^{(s)})$ in the right hand side of (6.15) by $(S_\bullet^{(\Delta)}, \Delta_\bullet^{(\Delta)})$ we obtain $(b^{(\mathrm{v})}(\bar{z}))_3$.

By replacing $(S_\bullet^{(s)}, \Delta_\bullet^{(s)})$ in the right hand side of (6.15) by (S, Δ) (recall that (S, Δ) has the size biased degree distribuion) we obtain the probability generating function $b_*^{(\mathrm{v})}(\bar{z})$

of the offspring of the initial case. The expected final size of the epidemic, conditioned on that a major outbreak occurs, is given by

$$1 - b_*^{(\mathrm{v})}(\bar{q}_b),$$

where $b_*^{(\mathrm{v})}(\bar{q}_b)$ is the probability of extinction in (6.11).

7 Summary and discussion

In this thesis we have considered SIR epidemics in populations where the social structure is represented by graphs with clustering. The main goal and contribution of this thesis was to extend previous results obtained by Miller (2009). Miller analysed the spread of SIR epidemics on graphs of this model under the assumption of homogenous infectivity and susceptibility, by attributing all secondary and tertiary cases in a triangle to the initial infective in that triangle, regardless of the true path of transmission within the triangle. We have used a branching process approach to provide expressions for the probability that a major outbreak occurs, and the expected final size of a major outbreak. We have extended the Miller's results by allowing for heterogeneity in individual infectivity. Furthermore, we have calculated the rank based basic reproduction number, and shown that vaccinating a fraction $1 - \frac{1}{R_0}$ of the population with a perfect vaccine is sufficient to prevent a major outbreak. In Appendix D, we prove that the branching process approximations are exact in the limit as the population size $N \to \infty$.

In most modelling efforts, a balance must be struck between accuracy, interpretability and tractability. Including too little detail in the model puts one at the risk of missing important mechanisms of the modelled process, which may result in a too simplistic model that does not faithfully replicate reality. Including too much detail in the model may result in an intractable model, whose key features are difficult to interpret. An important feature of the contact patterns in a population is the concept of groups. People typically belong to a wide range of groups (workplaces, families, sport teams, groups of friends, to name a few) whose members interact frequently with each other. Many network models, including the standard configuration model, do not capture this feature at all. By considering the configuration model with clustering, we have included an element of social groups in the model. Although the configuration model with clustering does incorporate group structure, the group size is restricted to three (or two). A graph of this model is locally treelike, except for the triangles formed by triangle edges. In other words, except for small groups of two or three individuals, there are virtually no groups. A real world social network is, however, typically characterised by layers of groups of varying size, some tightly and some more loosely formed. The branching process approximation used in this thesis relies heavily on the fact that, except for the triangles formed by triangle edges, short cycles are rare, provided that the number of nodes N of the graph is large. For this reason, it is not straightforward to directly extend the techniques employed in this thesis to more general graph models with clustering, where the structure is not locally treelike at some level.

The focus of this thesis was epidemics in discrete time, since we were mainly interested in the probability of a major outbreak and the expected final size of a major outbreak. A possible extension of this thesis would be epidemics in continuous time on graphs generated by the configuration model with clustering. A possible issue to address would be the real time growth rate of the epidemic. An SIR epidemic on a graph generated by the standard configuration model grows exponentially during the early phase (Pellis et al. 2015). We expect similar results to hold for SIR epidemics on a graphs generated by the configuration model with clustering. The challenge would lie in describing the real time epidemic process as a branching process, since dependencies arise from the presence of triangles. To be more specific, the offsprings of the first two cases in a triangle are not independent, since these cases compete to transmit the disease to the remaining susceptible individual in the triangle in question. One possible approach would be to describe the real-time spread of the disease as an infinite type branching process, see for instance Bollobás et al. (2007).

A List of symbols, notation conventions and assumptions

Table 1: Frequently used notation.

Notation	Usual meaning	Page
$\sim X$	The distribution of the random variable X	
$f(x) \asymp_x g(x)$	$\frac{f(x)}{g(x)} \to 1$ as $x \to \infty$	
$\mathbb{N}$	$\{1, 2, 3, \ldots\}$	
$\mathbb{N}_0$	$\{0, 1, 2, 3, \ldots\}$	
$\mathbb{N}_\infty$	$\{1, 2, 3, \ldots, \infty\}$	
$\mathbb{R}_+$	$[0, \infty)$	
$\bar{e}_1, \ldots, \bar{e}_s$	Standard basis of $\mathbb{R}^s$	
$\bar{1}$	$(1, \ldots, 1)^\mathsf{T}$	
$\bar{0}$	$(0, \ldots, 0)^\mathsf{T}$	
$\bar{a}^{\bar{b}}$	$a_1^{b_1} \cdot \ldots \cdot a_s^{b_s}$ where $\bar{a} = (a_1, \ldots, a_s)^\mathsf{T}$ and $\bar{b} = (b_1, \ldots, b_s)^\mathsf{T}$	
$A \overset{P}{=} A$	The events A and B are identical up to null sets, i.e. $P(A \setminus B) = P(B \setminus A) = 0$	
$C^{(u,v)}$	A point process whose points are the time points at which the individual u attempts to make contact with the individual v	4
$t_I^{(v)}$	The time point at which the individual v contracts the disease	4
τ_{v_i} or τ_i	The infectious period of v_i	4
R_0	Basic reproduction number	5
R_V	Perfect vaccine-associated reproduction number	6
$f_\mathrm{v}^{(c)}$	Critical vaccination coverage	6
N	The population size	
G_N	Graph consisting of N nodes, generated by the configuration model with clustering	9
$D_S^{(N)}$	The total single degree of G_N	9
$D_\Delta^{(N)}$	The total triangle degree of G_N	9
$v_1 \to \ldots \to v_n$	The path $\{(v_i, v_{i+1})\}_{i=1}^{n-1}$	7
$L(v_1 \to \ldots \to v_n)$	The length of the path $v_1 \to \ldots \to v_n$	7
$d(u, v)$	The distance between the nodes u and v	7
p	The degree distribution $\{p(k_\Delta, k_S)\}_{k_\Delta, k_S \in \mathbb{N}}$	8
(S, Δ)	Random vector with distribution p	8
$\mathfrak{S}(v)$	The susceptibility set of the node v	19
$\mathcal{L}(z)$	Laplace transform $\int e^{-zx} dF(x)$ of the infectious period	24

Table 1: Frequently used notation.

Notation	Usual meaning	Page
$p_\circ^{(s)}$	Single size biased degree distribution	18
$p_\circ^{(\Delta)}$	Triangle size biased degree distribution	18
$p_\bullet^{(s)}$	Downshifted single size biased degree distribution	18
$p_\bullet^{(\Delta)}$	Downshifted triangle size biased degree distribution	18
$p^{(N)}$	Empirical degree distribution of G_N	18
$p_\circ^{(N,s)}$	Single size biased empirical degree distribution of G_N	18
$p_\circ^{(N,\Delta)}$	Triangle size biased empirical degree distribution of G_N	18
$p_\bullet^{(N,s)}$	Downshifted single size biased empirical degree distribution of G_N	18
$p_\bullet^{(N,\Delta)}$	Downshifted triangle size biased empirical degree distribution of G_N	18
$\{Y \to 0\}$	The event that the branching process Y eventually becomes extinct	
$\{Y \not\to 0\}$	The event that the branching process Y avoids extinction	

Table 2: Assumptions.

Assumption	Page
A1	10
A2	10
A3	10
A4	53
A5	53
A6	57
A7	57

B Tools to compare epidemic models

In this section, we provide tools to compare the final size and the probability of non-extinction of an epidemic on G_N for different transmission mechanisms. The results described in the present section can, for instance, be used to compare the probability and final size of a major outbreak for different distributions of the transmission weights in the epidemic model presented in section 4. We illustrate this in Example B.1 and Example B.2 by comparing the probability of a major outbreak for six different models.

First, we will need some notation. Let $\{(S_i, T_i)\}_i$ be a sequence of i.i.d. random vectors taking values in $\mathbb{R}^2$ and let $\theta : \mathbb{R}^2 \to \mathbb{R}$ be a non-decreasing function (that is $\theta(x_1, y_2)$ is non-decreasing in x_i for fixed x_j, $i \neq j$) such that $0 \leq \theta(T_1, S_1) \leq 1$ holds almost surely. We often assume that $\theta(x, y) = \theta(xy)$. This covers the models considered in this project. We refer to T_i as the transmission weight of v_i and S_i as the susceptibility weight of v_i. We denote by $M := (\theta, \{(S_i, T_i)\}_i)$ the model for an epidemic governed by the weights $\{(S_i, T_i)\}_i$ and the transmission probability function θ. Let (v_i, v_j) be an edge of G_N. For the model M, the conditional probability of transmission along (v_i, v_j) if the tail v_i contracts the disease is $\theta(T_i, S_j)$. Given the transmission and susceptibility weights, (potential) transmissions along the edges of G_N are independent for different edges. We assume that the weights of different nodes are independent. However, the transmission weight T_i and the susceptibility weight S_i need not be independent, $i = 1, \ldots, N$.

Example B.1. Let $\theta : \mathbb{R} \to \mathbb{R}$ be the identity operator and let T be some real valued random variable with support in $[0, 1]$. Let further T_{Be} be Bernoulli distributed with expected value $E(T)$.

Consider the following three models:

M_1) The probability of transmission is distributed as $\theta(T) = T$

M_2) The probability of transmission is distributed as $\theta(E(T)) = E(T)$

M_3) The probability of transmission is distributed as $\theta(T_{Be})) = T_{Be}$

Note that the distribution governing the susceptibility of the individuals is degenerate and concentrated at 1, that is $S_i = 1$ almost surely for the models M_1-M_3. The marginal probability of transmission along an edge (v_i, v_j) is the same for the three models. However, as we will se in the present section, the dynamics of the spread of the disease for these three cases are not the same. This model was investigated further in section 4. ■

Example B.2. Let τ be a nonnegative extended real valued random variable (that is, τ takes values in $\mathbb{R}_+ \cup \{\infty\}$). Let the infectious period T_i of an individual v_i have distribution $\sim \tau$ and assume that each individual makes contact independently with each of its neighbours at a Poission rate with intensity 1.

Assume that a fraction f_v of the population is vaccinated with a perfect vaccine, so that an attempted transmission to a vaccinated individual succeeds with probability 0 and an attempted transmission to an unvaccinated individual succeeds with probability 1. To incorporate immunity from vaccination, let the susceptibility weights $\{S_i\}_i$ be Bernoulli distributed with success probability $1 - f_\mathrm{v}$. We assume that the susceptibility weights $\{S_i\}_i$ are independent of the transmission weights $\{T_i\}_i$.

The probability of transmission from v_i to its neighbour v_j is then distributed as

$$\tilde{\theta}(T_i, S_j) = 1 - e^{-S_j T_i}.$$

We refer to this model as M_4. It was investigated further in section 6.

In a slightly different model, which we call M_5, the entire population is vaccinated with a so called *leaky vaccine* with *vaccine efficacy* f_v. A leaky vaccine with efficay f_v reduces

the per-contact infection probability by a (multiplicative) factor $1-f_{\mathrm{v}}$, so that, for model M_5, the probability of transmission per infectious contact is $1-f_{\mathrm{v}}$. The probability of transmission from v_i to its neighbour v_j is distributed as

$$\tilde{\theta}(T_i, 1-f_{\mathrm{v}}) = 1 - e^{(1-f_{\mathrm{v}})T_i}.$$

Now assume that the entire population is vaccinated with a non-perfect vaccine, as in model M_5. The difference between M_5 and the present model, which we call M_6, is that the vaccine in M_6 provides a random multiplicative reduction in susceptibility distributed as the random variable $\tilde{S}$, $E(\tilde{S}) = 1-f_{\mathrm{v}}$, with support in $[0,1]$. Let $\{\tilde{S}_i\}_i$ be the susceptibility weights of M_6, where $\tilde{S}_i \sim \tilde{S}$. We assume that the susceptibility weights $\{\tilde{S}_i\}_i$ are independent of the transmission weights $\{T_i\}_i$. The (conditional) probability of transmission per infectious contact with the individual v_j is $\tilde{S}_j$. That is, the probability of transmission from v_i to its neighbour v_j is distributed as

$$\tilde{\theta}(T_i, \tilde{S}_j) = 1 - e^{-\tilde{S}_j T_i}.$$

To summarize, we consider the following three models:

M_4) A fraction f_{v} of the population is vaccinated with a perfect vaccine.

M_5) The entire population is vaccinated with a leaky vaccine with efficacy f_{v}. That is, the vaccine reduces the susceptibility of every vaccinated individual by a multiplicative factor $1-f_{\mathrm{v}}$.

M_6) The entire population is vaccinated with a non-perfect vaccine, and the individual immune response to vaccination is random and independent between individuals. The expected multiplicative reduction of the susceptibility is by a factor $1-f_{\mathrm{v}}$.

■

Let u be any node of G_N and let $\mathcal{N}_{\rightarrow}$ be a subset of the nodes from which an edge with head u emanate. Similarly, let $\mathcal{N}_{\leftarrow}$ be a subset of the heads of the edges with tail u. We assume that $\mathcal{N}_{\rightarrow}$ and $\mathcal{N}_{\leftarrow}$ are disjoint. Let further $\bar{t}$ and $\bar{s}$ denote the realizations of the transmission weights of the nodes of $\mathcal{N}_{\rightarrow}$ and the susceptibility weights of the nodes of $\mathcal{N}_{\leftarrow}$, respectively.

Following Meester and Trapman (2011) we define the *zero function*

$$P_0(u, \bar{t}, \mathcal{N}_{\rightarrow}, \bar{s}, \mathcal{N}_{\leftarrow}) \tag{B.1}$$

as the probability that the disease is not transmitted from any node of $\mathcal{N}_{\rightarrow}$ to any node of $\mathcal{N}_{\leftarrow}$ via u, conditioned on the transmission weights $\bar{t}$ of the nodes of $\mathcal{N}_{\rightarrow}$ and the susceptibility weights $\bar{s}$ of the nodes of $\mathcal{N}_{\leftarrow}$.

If $\mathcal{N}_{\rightarrow}$ is the empty set $\varnothing$ then we define (B.1) as the probability that u would not infect any member of $\mathcal{N}_{\leftarrow}$, if u is the initial case. Similarly, if $\mathcal{N}_{\leftarrow}$ is the empty set then (B.1) is the probability that u escapes infection from the nodes in $\mathcal{N}_{\rightarrow}$. That is

$$P_0(u, \bar{t}, \mathcal{N}_{\rightarrow}, \bar{s}, \mathcal{N}_{\leftarrow}) = E\left(1 - \left(1 - \prod_{v_i \in \mathcal{N}_{\rightarrow}} (1 - \theta(t_i, S_u))\right)\left(1 - \prod_{v_j \in \mathcal{N}_{\leftarrow}} (1 - \theta(T_u, s_j))\right)\right)$$

$$P_0(u, \bar{t}, \mathcal{N}_{\rightarrow}, \varnothing, \varnothing) = E\left(\prod_{v_i \in \mathcal{N}_{\rightarrow}} (1 - \theta(t_i, S_u))\right) \tag{B.2}$$

$$P_0(u, \varnothing, \varnothing, \bar{s}, \mathcal{N}_{\leftarrow}) = E\left(\prod_{v_j \in \mathcal{N}_{\leftarrow}} (1 - \theta(T_u, s_j))\right)$$

$$P_0(u, \varnothing, \varnothing, \varnothing, \varnothing) = 1$$

where S_u denotes the susceptibility weight of u and T_u is the transmission weight of u. We sometimes write $P_0(u, \bar{t}, \mathcal{N}_\rightarrow, \bar{s}, \mathcal{N}_\leftarrow, M)$ to make the underlying model M explicit.

If for two models $M_{(1)}$ and $M_{(2)}$

$$P_0(u, \bar{t}, \mathcal{N}_\rightarrow, \bar{s}, \mathcal{N}_\leftarrow, M_{(1)}) \leq P_0(u, \bar{t}, \mathcal{N}_\rightarrow, \bar{s}, \mathcal{N}_\leftarrow, M_{(2)})$$

holds for every $\mathcal{N}_\rightarrow$ and $\mathcal{N}_\leftarrow$ and corresponding weights $\bar{s}$ and $\bar{t}$ we write

$$P_0(M_{(1)}) \leq P_0(M_{(2)}).$$

For a node u of G_N and a positive integer a, let $E_a(u)$ be the set of nodes of G_N that can only be reached from u following a path consisting of at least a edges. That is, $v \in E_a(u)$ if and only if any path with starting point u and end point v consists of at least a edges, and at least one path from u to v exists. Let further

$$\boldsymbol{\Xi}_a(u)$$

be the collection of paths starting in u and ending in some node of $E_a(u)$.

We denote the event that at least one path of $\boldsymbol{\Xi}_a(u)$ is of finite length (that is, that the disease is transmitted along some path of $\boldsymbol{\Xi}_a(u)$) by $\mathcal{D}^{\boldsymbol{\Xi}_a(u)}$. For a given model M, let

$$P_M(\mathcal{D}^{\boldsymbol{\Xi}_a(u)})$$

be the probability that $\mathcal{D}^{\boldsymbol{\Xi}_a(u)}$ occurs.

If, for two models $M_{(1)} = (\theta_1, \{(S_i, T_i^{(1)})\}_i)$ and $M_{(2)} = (\theta_2, \{(S_i, T_i^{(2)})\}_i)$ with independent transmission and susceptibility weights it holds that

$$E(\theta_k(T_i^{(k)}, s_j)) = s_j$$

for $k = 1, 2$, then we say that the susceptibility distributions of $M_{(1)}$ and $M_{(2)}$ are *consistent*.

The following theorem was presented by Miller (2008, Lemma 2), see also Kuulasmaa (1982) for related results.

Theorem B.1. *Consider the two models $M_{(1)} = (\theta_1, \{(S_i, T_i^{(1)})\}_i)$ and $M_{(2)} = (\theta_2, \{(S_i, T_i^{(2)})\}_i)$, where the S_i are independent of the $T_i^{(1)}$ and the $T_i^{(2)}$, and the susceptibility distributions of $M_{(1)}$ and $M_{(2)}$ are consistent.*

If, for each node u of G and all sets $\mathcal{N}_\leftarrow$ of neighbours of u with susceptibility weights $\bar{s}$

$$P_0(u, \varnothing, \varnothing, \bar{s}, \mathcal{N}_\leftarrow, M_{(1)}) \leq P_0(u, \varnothing, \varnothing, \bar{s}, \mathcal{N}_\leftarrow, M_{(2)}) \tag{B.3}$$

then

$$P_{M_{(2)}}\left(\mathcal{D}^{\boldsymbol{\Xi}_a(u)}\right) \leq P_{M_{(1)}}\left(\mathcal{D}^{\boldsymbol{\Xi}_a(u)}\right). \tag{B.4}$$

The proof given below is due to Miller (2008). We present it here since it is instructive.

Proof. Assume without loss of generality that $\{T_i^{(1)}\}_i$ and $\{T_i^{(2)}\}_i$ are independent. To each node v_i of G_N assign the three numbers S_i, $T_i^{(1)}$ and $T_i^{(2)}$. Partition the nodes of G_N into two subsets $\mathcal{U}_1$ and $\mathcal{U}_2$. A node u_i transmits the disease to a neighbour u_j with probability $\theta_1(T_i^{(1)}, S_j)$ if $u_i \in \mathcal{U}_1$ and with probability $\theta_2(T_i^{(2)}, S_j)$ if $u_i \in \mathcal{U}_2$, conditioned on the weights.

Choose some $w \in \mathcal{U}_1$, and consider the impact of moving w from $\mathcal{U}_1$ to $\mathcal{U}_2$. There are two possibilities:

1. The occurrence of the event $\mathcal{D}^{\Xi_a(u)}$ is not determined by which of its neighbours w infects.
2. Alternative 1 does not hold

If alternative 1 holds then moving w from $\mathcal{U}_1$ to $\mathcal{U}_2$ have no impact on the occurrence of the event $\mathcal{D}^{\Xi_a(u)}$.

If alternative 2 holds then there is a set $\mathcal{N}_{\leftarrow}$ of neighbours of w such that if w transmits the disease to any of the members of $\mathcal{N}_{\leftarrow}$ then $\mathcal{D}^{\Xi_a(u)}$ occurs. Let $\bar{s}$ denote the susceptibility weights of the nodes in $\mathcal{N}_{\leftarrow}$. By (B.3) $P\left(\mathcal{D}^{\Xi_a(u)}\right)$ cannot increase by moving w from $\mathcal{U}_1$ to $\mathcal{U}_2$. Note that there exists a finite subset $\mathbf{\Pi}_a(u)$ of $\Xi_a(u)$ of finite paths such that $\mathcal{D}^{\Xi_a(u)}$ occurs if and only if $\mathcal{D}^{\mathbf{\Pi}_a(u)}$ occurs. The assertion (B.4) now follows from induction. □

Miller (2008) showed that increased homogeneity of the infectivity of the nodes increase the probability $P\left(\mathcal{D}_a^{\Xi}(u)\right)$ that some path of $\Xi_a(u)$ is of finite length and vice versa.

Theorem B.2 (Maximal homogeneity maximizes $P\left(\mathcal{D}^{\Xi_a(u)}\right)$)**.** *Consider the model $M = (\theta, \{(S_i, T_i)\}_i)$, where the S_i and the T_i are independent, and let $\theta_{hom}(t_i, s_j) = E(\theta(T_i, s_j))$. Let $M_{hom} = (\theta_{hom}, \{(S_i, T_i)\}_i)$. Then*

$$P_M\left(\mathcal{D}_a^{\Xi}(u)\right) \leq P_{M_{hom}}\left(\mathcal{D}^{\Xi_a(u)}\right)$$

Theorem B.3 (Maximal heterogeneity minimizes $P\left(\mathcal{D}^{\Xi_a(u)}\right)$)**.** *Consider the model $M = (\theta, \{(S_i, T_i)\}_i)$, where the S_i and the T_i are independent, and let $\{U_i\}_i$ be a sequence of independent uniform $(0, 1)$ random variables, independent of the S_i. Let the function θ_{het} be such that*

$$\theta_{het}(u, s) = \mathbb{1}(u < E(\theta(T_1, s)))$$

and let $M_{het} = (\theta_{het}, \{(S_i, U_i)\}_i)$. Then

$$P_{M_{het}}\left(\mathcal{D}^{\Xi_a(u)}\right) \leq P_M\left(\mathcal{D}^{\Xi_a(u)}\right)$$

Furthermore, Miller (2008) showed that, for a given marginal probability of transmission, $P\left(\mathcal{D}^{\Xi_a(u)}\right)$ is maximized for homogeneous susceptibility and infectitivity of the nodes, provided that the transmission weights and susceptibility weights are independent.

Recall that a real function φ defined on the interval (a, b), where $-\infty \leq a < b \leq \infty$ is said to be *concave* if

$$\varphi((1-t)x + ty) \geq (1-t)\varphi(x) + t\varphi(y)$$

whenever $a < x, y < b$ and $0 \leq t \leq 1$. If φ is concave then $-\varphi$ is said to be *convex*.

The following theorem is due to Meester and Trapman (2011). Note that no assumptions on independence between the transmission weights and susceptibility weights are made.

Theorem B.4. *Consider the two models $M_{(1)} = (\theta, \{(S_i^{(1)}, T_i^{(1)})\}_i)$ and $M_{(2)} = (\theta, \{(S_i^{(2)}, T_i^{(2)})\}_i)$, where $\theta(x, y) = \theta(xy)$ is concave.*
If, for each node u of G

$$P_0(M_{(1)}) \leq P_0(M_{(2)})$$

then

$$P_{M_{(2)}}\left(\mathcal{D}^{\Xi_a(u)}\right) \leq P_{M_{(1)}}\left(\mathcal{D}^{\Xi_a(u)}\right).$$

We now have the tools to compare the models of Example B.1.

Example B.1 (Continued)**.** By Theorem B.2 and Theorem B.3 we have

$$P_{M_3}\left(\mathcal{D}^{\Xi_a(u)}\right) \leq P_{M_1}\left(\mathcal{D}^{\Xi_a(u)}\right) \leq P_{M_2}\left(\mathcal{D}^{\Xi_a(u)}\right)$$

for each node u of G_N and each positive integer a. ■

To compare the models of Example B.2, we use Jensen's inequality (Rudin 1987, Theorem 3.3, page 62).

Theorem B.5 (Jensen's inequality)**.** *Let μ be a positive measure on a σ-algebra $\mathcal{A}$ on a set Ω, so that $\mu(\Omega) = 1$. If f is a real function in $L^1(\mu)$, $a < f(\omega) < b$ for all $\omega \in \Omega$, and if φ is convex on (a, b) then*

$$\varphi\left(\int_\Omega f d\mu\right) \leq \int_\Omega (\varphi \circ f) d\mu.$$

Example B.2 (Continued)**.** To compare the models M_4, M_5 and M_6 we only need to compare zero functions on the form (B.2), since the three models have common transmission function $\tilde{\theta}$ and transmission weights $\{T_i\}_i$, which are independent of the susceptibility weights. Since $\varphi(s) = e^{-sc}$ is a convex function on $[0, 1]$ for every $c > 0$, we have

$$e^{-cs} = e^{-c(s\cdot 1 + (1-s)\cdot 0)} \leq se^{-c} + (1 - s).$$

Thus (with $\bar{T} = \{T_{i_j}\}_{j=1}^k$)

$$\begin{aligned} E(e^{-\tilde{S}_l(T_{i_1}+\ldots+T_{i_k})}|\bar{T}) &\leq E(\tilde{S}_l e^{-(T_{i_1}+\ldots+T_{i_k})} + (1 - \tilde{S}_l)|\bar{T}) \\ &= (1 - f_{\mathrm{v}})e^{-(T_{i_1}+\ldots+T_{i_k})} + f_{\mathrm{v}} \\ &= E(e^{-S_l(T_{i_1}+\ldots+T_{i_k})}|\bar{T}) \end{aligned}$$

for each set of distinct indices $\{i_j\}_{j=1}^k$ such that $l \notin \{i_j\}_{j=1}^k$. Furthermore, by Jensen's inequality (Theorem B.5)

$$e^{-(1-f_{\mathrm{v}})(T_{i_1}+\ldots+T_{i_k})} \leq E(e^{-\tilde{S}_l(T_{i_1}+\ldots+T_{i_k})}|\bar{T}).$$

By Theorem B.4 it follows that

$$P_{M_4}\left(\mathcal{D}^{\Xi_a(u)}\right) \leq P_{M_6}\left(\mathcal{D}^{\Xi_a(u)}\right) \leq P_{M_5}\left(\mathcal{D}^{\Xi_a(u)}\right)$$

for each node u of G_N and each positive integer a. ■

We now discuss briefly how the above described results relate to the branching process approximations employed in this thesis (see section 3 and Appendix D). To this end, we introduce some additional notation. Let $G_\infty(u^*)$ be a graph obtained in the approximating branching process where the transmission probability is $T = 1$ almost surely. That is, $G_\infty(u^*)$ is obtained as follows. Let the joint degree of u^* be that of the initial case of the epidemic. If the initial case is chosen uniformly at random, then we draw the degree of u^* from p. We then construct the graph by exploring the neighbourhood of the part of the graph already explored, as described in section 2.5. The degrees of the nodes are sampled from the size biased degree distributions $p_\circ^{(s)}$ and $p_\circ^{(\Delta)}$. Note that, apart from the triangles formed by triangle edges, the infinite graph $G_\infty(u^*)$ is a tree.

We assume that the order in which half-edges are paired is such that for each positive integer a, eventually each half-edge attached to a node that can be reached from u^* by traversing at most a (undirected) edges is paired.

Let $B^{(a)}(u^*)$ be the ball centered at the node u^* with radius a. That is, $B^{(a)}(u^*)$ is the part of $G_\infty(u^*)$ that can be reached from u^* by traversing at most a edges, starting in u^*. Let further $\Xi_a^\infty(u^*)$ be the collection of paths from u^* to some node in $B^{(a)}(u^*) \setminus B^{(a-1)}(u^*)$. In other words, $\Xi_a^\infty(u^*)$ is the collection of paths to nodes that can be reached from u^* by traversing a half-edges but not by traversing less than a half-edges.

As we will see in Appendix D, if the spread of the disease is well approximated by a branching process in the early phase for a model M, the probability of a large outbreak equals (in the limit as $N \to \infty$)

$$\lim_{a\to\infty} P_M(\mathcal{D}^{\Xi_a^\infty(u^*)}). \tag{B.5}$$

Thus, for models M_1-M_5, the results presented in this section can be used to compare the probability of and final size of a major outbreak. As pointed out in section 3, if the direction of the edges of the graph representing the epidemic is reversed, the final size and the probability of an outbreak are interchanged (see for instance Miller (2008)). Thus, the results presented in the present section can be used to compare the expected final size for different models, provided that the backward epidemic process can be described as a branching process.

Because of the dependencies that arise from the presence of triangles, it is not as straightforward to describe the spread of the disease of model M_6 as a branching process as for models M_1-M_5. We conjecture that (B.5) is the limiting probability of a major outbreak also for this model. One possible approach to show this is to describe the spread of the disease of model M_6 as an infinite-type branching process, where the type of an individual is determined by its susceptibility and transmission weights, see for instance Bollobás et al. (2007). This is, however, beyond the scope of this project, but would be a nice future work.

C Coupling

We now introduce the concept of coupling. In Appendix D we use these results to prove the limiting results described in section 3. All results presented in this section can be found in the book by Thorisson (2000, Chapter 3).

Thorisson (2000) describes coupling as "the joint construction of two or more random elements (variables, processes), usually in order to deduce properties of the individual elements". Before giving a formal treatment of the concept of coupling, we remind the reader of the following two well-known definitions.

Definition C.1. *A* measurable space *is a pair* $(A, \mathscr{A})$*, where* A *is a set and* $\mathscr{A}$ *is a* σ*-algebra of subsets of* A*.*

Let $(A_1, \mathscr{A}_1)$ *and* $(A_2, \mathscr{A}_2)$ *be two measurable spaces. A function* $f : A_1 \to A_2$ *is said to be* measurable *if for every* $B \in \mathscr{A}_2$

$$f^{-1}(B) \in \mathscr{A}_1,$$

where $f^{-1}(B)$ *denotes the pre-image of* B *under* f*.*

Definition C.2 (Random elements)**.** *A* random element Y *in a measurable space* $(A, \mathscr{A})$ *defined on the probability space* $(\Omega, \mathscr{F}, P)$ *is a* measurable mapping *from* $(\Omega, \mathscr{F}, P)$ *to* $(A, \mathscr{A})$*. That is,* $Y^{-1}B \in \mathscr{F}$ *for every* $B \in \mathscr{A}$*.*

Let $\{Y_i\}_{i \in I}$ be a collection of random elements (Y_i may be a random variable, a process etc.), where I is an index set. A *coupling* of the Y_i is a family $\{\hat{Y}_i\}_{i \in I}$ of random elements, defined on the same probability space, that satisfies

$$Y_i \overset{D}{=} \hat{Y}_i$$

for every $i \in I$. That is, a coupling of $\{Y_i\}_{i \in I}$ is any family of random elements $\{\hat{Y}_i\}_{i \in I}$ such that the marginal distribution of the random element Y_i coincides with the marginal distribution of the random element $\hat{Y}_i$ for each $i \in I$. Note that a coupling of the Y_i is not unique in general.

Definition C.3 (Coupling)**.** *For each index* $i \in I$*, let* Y_i *be a random element in the measurable space* $(A_i, \mathscr{A}_i)$ *defined on a probability space* $(\Omega_i, \mathscr{F}_i, P_i)$*. A family of random elements* $\{\hat{Y}_i\}_{i \in I}$ *defined on a common probability space* $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$ *is a coupling of* $\{Y_i\}_{i \in I}$ *if*

$$\hat{Y}_i \overset{D}{=} Y_i$$

for each index $i \in I$*.*

Consider a measurable space $(A, \mathscr{A})$, and two measures μ and ν on $(A, \mathscr{A})$. The measure ν is a *component* of μ if

$$\nu(E) \leq \mu(E)$$

for each $E \in \mathscr{A}$, and we write $\nu \leq \mu$. Let $\{\mu_i\}_{i \in I}$ be a collection of measures on $(A, \mathscr{A})$. We say that ν is a *common component* of the μ_i if ν is a component of μ_i for each $i \in I$. If, in addition, every common component of the μ_i is a component of ν, then ν is said to be the *greatest common component* of the μ_i.

Theorem C.1 (cf. Thorisson (2000), Theorem 7.1, Chapter 3)**.** *Let* $(A, \mathscr{A})$ *be a measurable space, and* $\{\mu_i\}_{i \in I}$ *a collection of measures on* $(A, \mathscr{A})$*. Then there exists a unique greatest component* $\bigwedge_{i \in I} \mu_i$ *of the* μ_i*, given by*

$$\left(\bigwedge_{i \in I} \mu_i \right)(E) = \sup\{\nu(E) : \ \nu \leq \mu_i \text{ for all } i \in I\}$$

where the supremum runs over all common components ν of the μ_i.

Now consider the collection $\{Y_i\}_{i\in\mathbb{N}_\infty}$ of random elements in the measurable space $(A, \mathscr{A})$. Note that the index set is given by $\mathbb{N}_\infty = \mathbb{N} \cup \{\infty\}$. A *coupling index* K of a coupling $\{\hat{Y}_i\}_{i\in\mathbb{N}_\infty}$ is any random variable in $\mathbb{N}_\infty$ defined on the same probability space $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$ as the $\hat{Y}_i$ that satisfies

$$\hat{Y}_n = \hat{Y}_\infty \text{ whenever } n \geq K. \tag{C.1}$$

For any measure μ on a measurable space $(A, \mathscr{A})$ we denote the *total mass* $\mu(A)$ of μ by $\|\mu\|$.

Theorem C.2 (cf. Thorisson (2000), Theorem 9.1 Chapter 3)**.** *Let K be the coupling index in (C.1). Then*

$$\left\| \bigwedge_{n\leq k\leq\infty} P(Y_k \in \cdot) \right\| \geq \hat{P}(K \leq n) \tag{C.2}$$

for any $n \in \mathbb{N}_\infty$.

That is, it is not possible to find a coupling $\{\hat{Y}_i\}_{i\in\mathbb{N}_\infty}$ of $\{Y_i\}_{i\in\mathbb{N}_\infty}$ such that the probability that $\hat{Y}_i$ coincide for $i = n, \ldots, \infty$ exceeds the total mass of $\bigwedge_{i\geq n} P(Y_i \in \cdot)$. The coupling $\{\hat{Y}_i\}_i$ of Theorem C.2 with the coupling index K is said to be *maximal at each index* if equality holds in (C.2) for each $n \in \mathbb{N}_\infty$. It can be shown that there exists a coupling that is maximal at each index for any sequence $\{Y_i\}_{i\in\mathbb{N}_\infty}$ of random elements in $(A, \mathscr{A})$ (Thorisson 2000, Theorem 9.3 Chapter 3).

D Proof of limiting results

In this appendix, we prove that the branching process approximation is exact in the limit as the population size N tends to infinity. The proof draws heavily on ideas used by Ball et al. (2009, 2014), who proved similar results for two related models, the households model and epidemics on random intersection graphs. The main difference between our proof and the proofs presented by Ball et al. (2009, 2014) is that we use maximal coupling, which enables us to obtain almost sure convergence for the approximating forward branching process and convergence in probability for the approximating backward process as the population size $N \to \infty$.

For ease of presentation, we prove the results for the standard configuration model with no triangle edges, under the assumption of heterogeneous infectivity and homogeneous susceptibility. We use a framework that is straightforward to extend to the models considered in this report (with the exception of model M_6, Appendix B) and the configuration model with clustering.

We will need some notation. Let p be a probability measure with support in $\mathbb{N}_0$ and let K be a random variable with distribution p. We assume that $E(K^2 \log K < \infty)$. Let $p_\circ$ be the size biased degree distribution, that is

$$p_\circ(k) = \frac{kp(k)}{E(K)},$$

and let $G_\circ$ be the cumulative distribution function of the size biased degree distribution

$$G_\circ(k) = \sum_{j \leq k} p_\circ(j)$$

Let further the ordered sequence $\bar{d} = (d_1, d_2, \ldots)$ be a given degree sequence in $\mathbb{N}_0$ and let $\bar{d}_N$ be the sequence $\bar{d}$ truncated after the Nth element, that is

$$\bar{d}_N = (d_1, \ldots, d_N)$$

and let G_N be a standard configuration model graph based on $\bar{d}_N$. Let further $p^{(N)}$ be the distribution of $\bar{d}_N$,

$$p^{(N)}(k) = \tfrac{1}{N} \sum_{i=1}^{N} \mathbb{1}(d_i = k)$$

and let $p_\circ^{(N)}$ be the size biased distribution corresponding to $p^{(N)}$

$$p_\circ^{(N)}(k) = \tfrac{k}{D^{(N)}} \sum_{i=1}^{N} \mathbb{1}(d_i = k)$$

where

$$D^{(N)} = \sum_{i=1}^{N} d_i$$

is the total degree of G_N. We make the following assumptions on $\bar{d}$.

A4) $p^{(N)}(k) \to p(k)$ as $N \to \infty$ for each $k \in \mathbb{N}_0$.

A5) $p_\circ^{(N)}(k) \to p_\circ(k)$ as $N \to \infty$ for each $k \in \mathbb{N}_0$.

We will need two classical results (Grimmett and Stirzaker 1992, sections 7.3, 7.5 and 7.9) later, which we include for completeness.

Proposition D.1 (Strong Law of Large Numbers). *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of iid random variables. Then*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \overset{a.s.}{\to} \mu \tag{D.1}$$

for some constant μ as $n \to \infty$ if and only if $E(|X|) < \infty$. In this case $E(X_1) = \mu$. If, in addition, $E(X^2) < \infty$, then the convergence (D.1) holds in mean square.

Proposition D.2 (Borel Cantelli Lemma). *Let $\{A_m\}_m$ be a sequence of events and let*

$$A = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

be the event that infinitely many of the A_m occur. Then

$$P(A) = 0 \text{ if } \sum_{m=1}^{\infty} P(A_m) < \infty.$$

If, in addition, the A_m are independent then

$$P(A) = 1 \text{ if } \sum_{m=1}^{\infty} P(A_m) = \infty.$$

Note that if $\bar{d}$ is a sequence of independent copies of K then by the Strong Law of Large Numbers, the assumptions A4 and A5 are almost surely satisfied.

The following theorem (Meyer 2000, page 666) will be useful in later sections.

Theorem D.3 (Collatz-Wielandt formula). *Let the $s \times s$ matrix M be positively regular. The Perron root r of M is then given by*

$$r = \max_{\bar{x} \in \mathcal{N}} \left(\min_{\substack{1 \le i \le d \\ x_i \neq 0}} \frac{(M\bar{x})_i}{x_i} \right),$$

where $\mathcal{N} = \{\bar{x} \in \mathbb{R}^d : \bar{x} \geq \bar{0},\ \bar{x} \neq \bar{0}\}$.

Here and in what follows, inequalities of vectors and matrices are to be interpreted element-wise.

Remark D.1. Meyer (2000) states Theorem D.3 for matrices with strictly positive entries. Following the proof presented by Meyer (2000), it is straightforward to verify that Theorem D.3 holds for positively regular matrices.

D.1 Branching process framework

Since we perform the proof for the standard configuration model under the assumption of heterogeneous infectivity and homogeneous susceptibility, the approximating forward and backward branching processes are single-type processes. We present a more extensive multi-type branching process framework, although we perform the proof for the single type case $s = 1$, since this enables straightforward extension of the proof to the configuration model with clustering. The framework presented in this section is based on results described in the book by Jagers (1975, Chapter 4).

We consider branching processes with the common finite type space $\{1, 2, \ldots, s\}$. Denote the space of all individuals (including the *ancestor*) by $\mathcal{X}$. We denote the ancestor by (a, τ_a), where τ_a denotes the type of the ancestor. We typically consider branching processes where the type of the ancestor is unique. That is, the ancestor is the only individual of its type. If this is the case we write $\tau_a = 0$, and call the branching process an s-type branching process although there are $s + 1$ types; $0, 1, \ldots, s$.

Individuals $x \in \mathcal{X}$ are vectors of the form

$$x = (a, \tau_a; j_1, \tau_1; j_2, \tau_2; \ldots; \tau_n, j_n) \tag{D.2}$$

where the lineage of x is fully specified by $\{j_i\}_{i=1}^n \subset \mathbb{N}$, $\{\tau_i\}_{i=1}^n \subset \mathbb{N}$ and (a, τ_a); the individual x is the j_nth child of type τ_n of the individual

$$x_p = (a, \tau_a; j_1, \tau_1; j_2, \tau_2; \ldots; \tau_{n-1}, j_{n-1}) \in \mathcal{X}. \tag{D.3}$$

We refer to the individual x_p in (D.3) as the *parent* of x, and we write $(x_p; \tau_n, j_j)$ for x. Any individual with lineage

$$(a, \tau_a; j_1, \tau_1; j_2, \tau_2; \ldots; \tau_n, j_n, \ldots, \tau_m, j_m) \tag{D.4}$$

where $m \geq n$ is said to be a *descendant* of x. If y is a descendant of x then we call x a *ascendant* of y. Note that $\mathcal{X}$ is countable and that "most" individuals of $\mathcal{X}$ are not *realized* (defined below). In the sequel, we consider several branching process, and each individual of $\mathcal{X}$ may be realized in none, some, or several of the branching processes under consideration. Given a branching process B, we sometimes write $x(B)$ to make the underlying branching process B explicit. We refer to $x(B)$ as the individual x of B.

For a given branching process B with type space $\{1, \ldots, s\}$ (or $\{0, \ldots, s\}$ if the type of the ancestor is unique), the offspring vector of an individual $x(B)$

$$\bar{\xi}_{x(B)} = (\xi_1^{x(B)}, \ldots, \xi_s^{x(B)})^\mathsf{T}$$

is a random variable in $\mathbb{N}_0^s$. An individual $(x; j, \tau)$ of B is said to be *realized* if its parent $x(B)$ is realized and if $j \leq \xi_\tau^{x(B)}$. The ancestor (a, τ_a) is always realized.

Given the branching process B, we define

$$R_{x(B)} = \begin{cases} 1 & \text{if } x \text{ of } B \text{ is realized} \\ 0 & \text{otherwise} \end{cases}$$

for each individual $x \in \mathcal{X}$. We define the *generation* $|x|$ of the individual

$$x = (a, \tau_a; j_1, \tau_1; j_2, \tau_2; \ldots; \tau_n, j_n) \in \mathcal{X}$$

as

$$|x| := n.$$

If $|x| = n$ then x is said to belong to the nth generation. Given a branching process B, define

$$B^{(n)} := \{x \in \mathcal{X} : \ |x| \leq n, R_{x(B)} = 1\}.$$

That is, $B^{(n)}$ is the set of individuals that are realized in generations $1, \ldots, n$ of the branching process B.

D.2 Limit of the approximating branching process

To analyse the spread of the disease on G_N, we employ a branching process approximation. To this end, for each population size N we construct an epidemic process $\hat{E}_{(N)}$. The epidemic process $\hat{E}_{(N)}$ can be coupled with a branching process $\hat{Z}_{(N)}$ in such a way that they coincide in the early phase of the epidemic. In this section, we show that in the limit as the population size N tends to infinity, the branching process $\hat{Z}_{(N)}$ coincides almost surely with a limiting branching process $\hat{Z} = \hat{Z}_{(\infty)}$.

Let

$$\mathcal{Z}_N := \bigcup_{N \leq N' \leq \infty} \{\hat{Z}_{(N')}\}$$

be the collection of branching processes $\hat{Z}_{(N')}$ corresponding to populations of size at least N (including $\hat{Z} = \hat{Z}_{(\infty)}$), and let

$$\mathcal{Z} := \mathcal{Z}_1.$$

The offspring distribution of $\hat{Z}_{(N)}$ is governed by the mechanisms of transmission of $\hat{E}_{(N)}$ and by the size biased empirical degree distribution $p_\circ^{(N)}$ of G_N. Similarly, the offspring distribution of $\hat{Z}$ is governed by the mechanisms of transmission and by the size biased degree distribution $p_\circ$. In other words, for each branching process $\hat{Z}_{(N)} \in \mathcal{Z}$, the offspring distribution of $\hat{Z}_{(N)}$ is identical to the distribution of the number of cases caused by an infected node, if the degree of the node is drawn from the size biased degree distribution $p_\circ^{(N)}$ (with $p_\circ^{(\infty)} = p_\circ$). We construct the branching process B of $\mathcal{Z}$ by assigning degrees to the individuals of B. If $B = \hat{Z}_{(N)}$, then the degree $D_{x(B)}$ of $x(B)$ is drawn from $p_\circ^{(N)}$. Given the degree $D_{x(B)}$, the offspring distribution of the individual x of B is governed by the probability law of transmission, specified in the epidemic model.

More specifically, we construct the branching processes of $\mathcal{Z}$ as follows. Let $\{U_x\}_{x \in \mathcal{X}}$ be a sequence of independent random variables, uniformly distributed on the interval $[0, 1]$. We use U_x to draw the degree of the individual $x(B)$, $B \in \mathcal{Z}$. Unless the individual x is the ancestor (a, τ_a), the degree $D_{x(B)}$ of $x\hat{(}Z)$ is drawn as follows.

If

$$G_\circ(k-1) \leq U_x < G_\circ(k) \tag{D.5}$$

then the individual $x(\hat{Z})$ is assigned the degree k.

If

$$G_\circ(k-1) \leq U_x < \min(p_\circ^{(N)}(k), p_\circ(k)) + G_\circ(k-1) \tag{D.6}$$

then the individual $x(\hat{Z}_{(N)})$ is assigned the degree k. If this does not hold for any k then the degree of the individual $x(\hat{Z}_{(N)})$ is drawn according to some rule (which we do not specify yet since the details are not important for this stage of the proof) so that the marginal degree distribution of this individual agrees with the empirical size biased degree distribution. The procedure of assigning degrees is illustrated in Figure D.1.

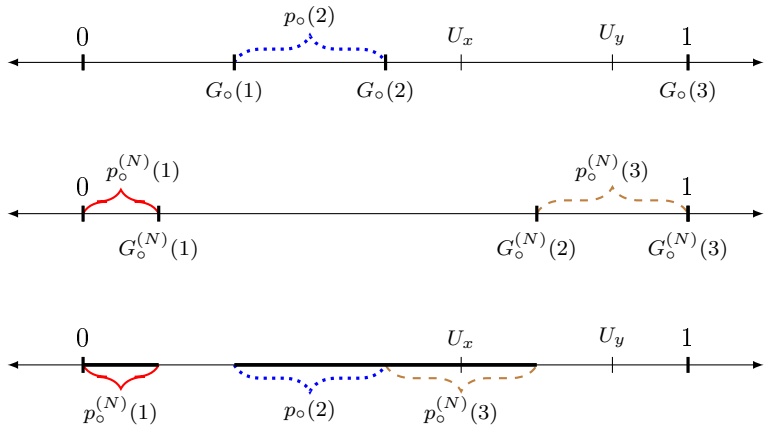

Figure D.1: Schematic illustration of the procedure of assigning degrees. Top: Size biased degree distribution. The individuals $x(\hat{Z})$ and $y(\hat{Z})$ both have degree 3. Center: Empirical size biased degree distribution of G_N. Bottom: The degrees of $x(\hat{Z})$ and $x(\hat{Z}_{(N)})$ coincide since U_x belongs to the intervals $[G_\circ(2), G_\circ(3))$ and $[G_\circ(2), G_\circ(2) + p_\circ^{(N)}(3))$. The degrees of $y(\hat{Z})$ and $y(\hat{Z}_{(N)})$ do not coincide since $y(\hat{Z}_{(N)})$ is assigned the degree 2.

The degrees of the ancestor are drawn analogously, with the size biased degree distribution $p_\circ$ replaced by the underlying degree distribution p, and the empirical size biased degree distribution $p_\circ^{(N)}$ replaced by the empirical degree distribution $p^{(N)}$, since the initial case is assumed to be chosen uniformly at random.

Not making details explicit, the mechanism of transmission of the disease is assumed to satisfy:

A6) The mechanism of transmission is such that if we have $D_{x(B_1)} = D_{x(B_2)}$ and $R_{B_1(x)} = R_{B_2(x)}$ for two branching processes $B_1, B_2 \in \mathcal{Z}$, then almost surely $R_{B_1(y)} = R_{B_2(y)}$ for each child y of x. That is, if the individuals $x(B_1)$ and $x(B_2)$ are both realized (or both not realized) and if their degrees coincide, then almost surely $y(B_1)$ is realized if and only $y(B_2)$ is realized for each child y of x.[3]

A7) The transmission mechanism is assumed to be independent of $\bar{d}$ and $\{U_x\}_{x \in \mathcal{X}}$.

We say that two branching processes B_1, B_2 of $\mathcal{Z}$ *coincide up to generation* n if $R_{x(B_1)} = R_{x(B_2)}$ and $D_{x(B_1)} = D_{x(B_2)}$ for each individual x of generation $|x| \leq n$. We also say that $B_1^{(n)}$ and $B_1^{(n)}$ coincide. That is, $B_1^{(n)}$ and $B_1^{(n)}$ coincide if the realized individuals of generation $1, \ldots, n$ and their degrees coincide for the two branching processes.

We now show that there exists some event H_1 and some *generation function* $\kappa : \mathbb{N} \to \mathbb{N}$ that satisfies

1. $\hat{P}(H_1) = 1$
2. For each $\omega \in H_1$ there exists $N_1(\omega)$ such that whenever $N \geq N_1$, it holds that $B^{(\kappa(N))}$ coincide for all $B \in \mathcal{Z}_N$
3. κ is non-decreasing and $\kappa(N) \to \infty$ as $N \to \infty$

[3] We may, for instance, construct the mechanism of transmission as follows. Suppose that, on the current probability space there exists a supply of independent random variables $\{U_x^{(i)}\}_{i,x}$, uniformly distributed on $[0, 1]$, and a collection $\{T_x\}_x$ of independent transmission weights, distributed as T (T plays the same role as in section 4) and independent of the U_x. For a branching process $B \in \mathcal{Z}$, the number of children of $x(B)$ (provided that $x(B)$ is realized) is the number of indices $i, 1 \leq i \leq D_{x(B)} - 1$ for which $U_x^{(i)} \leq T_x$. This construction satisfies the assumption A6.

That is, we show that we may choose the generation function κ so that with probability one, the branching processes of $\mathcal{Z}_N$ coincide up to generation $\kappa(N)$ for all but finitely many values of the population size N.

To this end, define

$$A_N^{(n)} := \{B^{(n)} \text{ coincides for every } B \in \mathcal{Z}_N\}.$$

In words, $A_N^{(n)}$ is the event that $\hat{Z}$ and $\hat{Z}_{(N')}$ coincide up to generation n whenever the population size N' is at least N. Note that the event $A_N^{(n)}$ is non-decreasing in N.

For each fixed generation n,

$$\hat{P}\left(\bigcup_N A_N^{(n)}\right) = 1 \tag{D.7}$$

Indeed, for each individual $x \in \mathcal{X}$ the event that $D_{x(Z_{(N')})} = D_{x(Z)}$ for all $N' \geq N$ is the event that

$$G_\circ(k-1) \leq U_x < \inf_{N' \geq N} \min\left(p_\circ^{(N')}(k), p_\circ(k)\right) + G_\circ(k-1) \tag{D.8}$$

for some $k \in \mathbb{N}$.

By assumption A5

$$\inf_{N' \geq N} \min(p_\circ^{(N')}(k), p_\circ(k)) \nearrow p_\circ(k)$$

as $N \to \infty$ for each $k \in \mathbb{N}$. Since (we may ignore the null event $\{U_x = 1\}$)

$$G_\circ(D_{x(\hat{Z})} - 1) \leq U_x < G_\circ(D_{x(Z)})$$

the event in (D.8) happens for some population size $N \in \mathbb{N}$ almost surely. In other words, for each individual $x \in \mathcal{X}$ there exists with probability one some $N \in \mathbb{N}$ such that $D_{x(\hat{Z}_{(N')})} = D_{x(\hat{Z})}$ for all $N' \geq N$.

Thus, it is readily checked that the assertion in (D.7) holds for $n = 0$ (recall that generation 0 consists of the ancestor). The assertion in (D.7) now follows by induction over n and assumption A6.

Since

$$\hat{P}\left(A_N^{(n)}\right) \to 1$$

as $N \to \infty$ for each $n \in \mathbb{N}$ there exists a strictly increasing (non-random) sequence $\{\kappa^{-1}(n)\}_{n \in \mathbb{N}}$ in $\mathbb{N}$ such that

$$\sum_{n \in \mathbb{N}} \left(1 - \hat{P}\left(A_{\kappa^{-1}(n)}^{(n)}\right)\right) < \infty.$$

By the Borel Cantelli Lemma (Theorem D.2) the event

$$\left(A_{\kappa^{-1}(n)}^{(n)}\right)^c \tag{D.9}$$

happens only for finitely many $n \in \mathbb{N}$ almost surely. That is, with probability one $\hat{Z}$ and $\hat{Z}_{(N)}$ coincide up to generation $\kappa(N)$ for all but finitely many values of the population size $N \in \mathbb{N}$, where $\kappa : \mathbb{N} \to \mathbb{N}$ is the function that satisfies

$$\kappa(N) = \sup\{n \in \mathbb{N}; \kappa^{-1}(n) \leq N\}$$

for all $n \in \mathbb{N}$.

Thus, there exists some event H_1, $\hat{P}(H_1) = 1$, such that for each $\omega \in H_1$ there exists $N_1(\omega)$ such that $\hat{Z}$ and $\hat{Z}_{(N)}$ coincide up to generation $\kappa(N)$ for each population of size at least N_1. Note that $\kappa(N) \to \infty$ as $N \to \infty$.

Remark D.2. If $\bar{d}$ is a sequence of independent copies of $K \sim p$ then the results presented in this section holds for almost every realization of $\bar{d}$ by the Strong Law of Large Numbers. By following the steps performed above, it is straightforward to show that we may choose a common non-random generation function κ for almost every realization of $\bar{d}$.

D.2.1 Approximation of the epidemic process

In this section, we construct the epidemic process $\hat{E}_{(N)}$ and show that in the early phase of the epidemic $\hat{E}_{(N)}$ coincides with the branching process $\hat{Z}$ in the limit as the population size N tends to infinity. The cornerstone of the proof is the result described in section 2.5; we may construct the graph G_N as the epidemic propagates, by pairing the half-edges of infected nodes. In other words, we explore G_N along with the spread of the disease, pairing the infectious half-edges along which the disease is transmitted. Since we pair each infectious half-edge with a half-edge chosen uniformly at random among the *free* (not yet paired) half-edges, the probability that the node v_i of G_N is chosen is proportional to the number of free half-edges attached to v_i.

We say that a node v is *involved* at a certain stage of this pairing procedure if some half-edge attached to v is already paired. The branching process coupling breaks down if a *collision* occurs, that is if a half-edge attached to an already involved node is chosen in the pairing procedure. We show that the generation function κ may be chosen so that almost surely the branching coupling breaks down before generation $\kappa(N)$ only for finitely many values N of the population size.

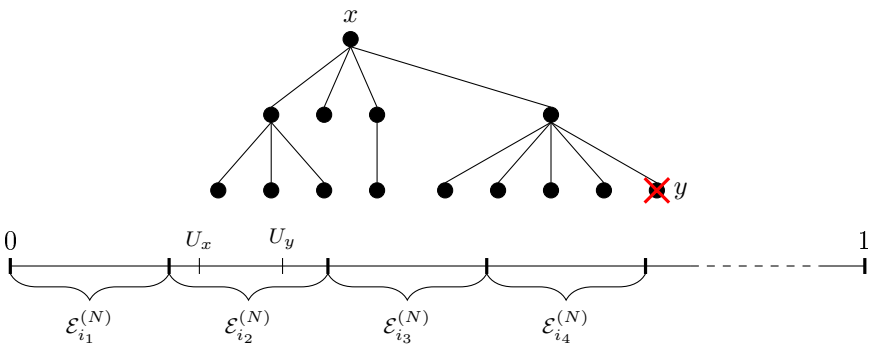

Figure D.2: The individuals $x(\hat{Z}_{(N)})$ and $y(\hat{Z}_{(N)})$ are both assigned the degree d_{i_2} of the node v_{i_2}. The coupling of the epidemic process $\hat{E}_{(N)}$ with the approximating branching process $\hat{Z}_{(N)}$ breaks down when the individual y of $\hat{Z}_{(N)}$ is realized, since $x(\hat{Z}_{(N)})$ and $y(\hat{Z}_{(N)})$ collide. That is, v_{i_2} is chosen for the second time in the pairing procedure when y of $\hat{Z}_{(N)}$ is assigned the degree d_{i_2}.

We now describe how the epidemic processes $\hat{E}_{(N)}$, $N \in \mathbb{N}$, are constructed. To this end, we give some additional details on the procedure of assigning degrees to the individuals of the branching processes of $\mathcal{Z}$. See Figure D.2 for an illustration of this procedure. We partition the interval $[0,1]$ into $D^{(N)}$ disjoint Borel sets of the form

$$\mathcal{E}^{(N)}_{i,j}, \tag{D.10}$$

$1 \leq i \leq N, 1 \leq j \leq d_i$, each of measure $\frac{1}{D^{(N)}}$, independently of $\{U_x\}_{x \in \mathcal{X}}$. For some given enumeration of the d_i half-edges attached to the node v_i of G_N, we assign the set $\mathcal{E}^{(N)}_{i,j}$ to the jth half-edge of v_i. If

$$U_x \in \mathcal{E}^{(N)}_{i,j} \tag{D.11}$$

then the individual x of $\hat{Z}_{(N)}$ is assigned the degree d_i of the node v_i. The construction given in (D.6) translates to

$$\left[G_\circ(k-1),\ G_\circ(k-1)+\min(p_\circ^{(N)}(k), p_\circ(k))\right) \subset \bigcup_{v_i \in \mathcal{V}_k^{(N)}} \mathcal{E}_i^{(N)} \tag{D.12}$$

for each $k \in \mathbb{N}$, where $\mathcal{V}_k^{(N)}$ is the set of nodes, v_i say, of G_N such that the degree d_i of v_i equals k, and

$$\mathcal{E}_i^{(N)} := \left(\bigcup_{1 \le j \le d_i} \mathcal{E}_{i,j}^{(N)} \right). \tag{D.13}$$

To construct the epidemic process $\hat{E}_{(N)}$, we let $\hat{Z}_{(N)}$ and $\hat{E}_{(N)}$ coincide. If the individual x of $\hat{Z}_{(N)}$ is realized, then the event in (D.11) corresponds to chosing the jth half-edge of the node v_i in the pairing procedure. The coupling breaks down if a node is chosen for the second time in the pairing procedure. This happens if

$$U_x, U_y \in \mathcal{E}_i^{(N)}$$

for some node v_i and two distinct realized individuals x and y of $\hat{Z}_{(N)}$, and we say that $x(\hat{Z}_{(N)})$ and $y(\hat{Z}_{(N)})$ *collide*. In the example of Figure D.2, x and y of $\hat{Z}_{(N)}$ collide. This construction ensures that $\hat{Z}$, $\hat{Z}_{(N)}$ and $\hat{E}_{(N)}$ coincide until a collision occurs or until the degrees of a realized individual of $\hat{Z}$ and $\hat{Z}_{(N)}$ do not coincide, whichever occurs first.

The following requirement is a convenient way to ensure that two specific individuals can only collide for a finite number of values of the population size N. We include it since it is instructive, although it is not strictly needed for the proof. For fixed N and each $k \in \mathbb{N}$, let $j_k = j_k(N)$ be the largest integer that satisfies

$$j_k \frac{k}{D^{(N)}} \le \min\left(p_\circ(k), p_\circ^{(N)}(k)\right).$$

In addition to the constraint given in (D.12), we assume that for each degree k, there is a collection consisting of j_k sets $\mathcal{E}_i^{(N)}$ that are intervals and that form a partition of the interval

$$\left[G_\circ(k-1), G_\circ(k-1) + j_k \frac{k}{D^{(N)}}\right),$$

as shown in Figure D.3. A collision occurs if $U_x \in \mathcal{E}_i^{(N)}$ and $U_y \in \mathcal{E}_i^{(N)}$ for two distinct realized individuals $x, y \in \mathcal{X}$ of $Z_{(N)}$ and some node v_i of G_N. Loosely speaking, this requirement ensures that collisions happens if the distance between U_x and U_y is small for two distinct individuals x and y. Note that

$$j_k \frac{k}{D^{(N)}} \to p_\circ(k)$$

as $N \to \infty$.

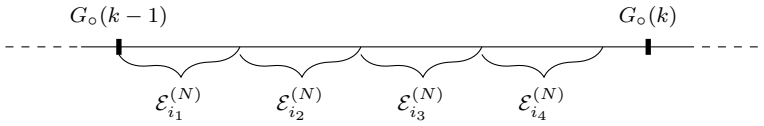

Figure D.3: The sets $\mathcal{E}_{i_1}^{(N)}, \ldots, \mathcal{E}_{i_4}^{(N)}$ are intervals. Here $j_k = 4$.

Let the random variable $\tau_{(N)}$ be the smallest value of n such that two distinct realized individuals $x(Z_{(N)})$ and $y(Z_{(N)})$ collide and $|x| = n$. That is, n is the first generation in which a collision occurs. To show that we may choose the generation function κ so that almost surely

$$\tau_{(N)} > \kappa(N)$$

for all but finitely many values of the population size N, we note that if

$$G_\circ(k-1) \leq U_x < G_\circ(k-1) + j_k \frac{k}{D^{(N)}}$$

for some k, then the individual x of $Z_{(N)}$ can only collide with the realized individual y of $Z_{(N)}$ if U_x is within Euclidean distance $\frac{k}{D^{(N)}}$ of U_y. It is now straightforward to employ the technique in section D.2 to show that we may choose the generation function κ such that almost surely $\tau_{(N)} > \kappa(N)$ for all but finitely many values of the population size N and $\kappa(N) \to \infty$ as $N \to \infty$.

Remark D.3. Remark D.2 applies here as well. That is to say, if the degree sequence $\bar{d}$ consists of independent copies of $K \sim p$ we may choose the generation function κ so that for almost every realization of $\bar{d}$ it holds almost surely that $\tau_{(N)} > \kappa(N)$ for all but finitely many values of the population size N

D.2.2 Asymptotic growth rate

In the previous sections we have seen that there exists a non-decreasing function κ such that $\kappa(N) \to \infty$ as $N \to \infty$, and there exists almost surely some (random) number N' such that for each population of size $N \geq N'$ the three processes $\hat{Z}$, $\hat{Z}_{(N)}$ and $\hat{E}_{(N)}$ coincide up to generation $\kappa(N)$. We construct the epidemic forward process $\hat{E}_{(N)}$ up to generation $\kappa(N)$, and refer to the epidemic process so obtained as a *stopped* epidemic process.

If $\hat{Z}$ is supercritical, by Theorem (2.3) the asymptotic growth rate of $\hat{Z}$ is almost surely exponential provided that $\hat{Z}$ does not go extinct, i.e

$$\frac{|\hat{Z}_n|}{r^n} \overset{a.s}{\to} W$$

as $n \to \infty$ where $\hat{Z}_n$ is the number of realized individuals of generation n and the random variable W is as in (2.12) up to multiplication by a positive constant.

Thus

$$\frac{\Psi_{(N)}}{r^{\kappa(N)}} \overset{a.s}{\to} W \tag{D.14}$$

as $N \to \infty$, where $\Psi_{(N)}$ is the number of individuals realized in generation $\kappa(N)$ of $\hat{Z}_{(N)}$. Note that we may choose $\kappa(N)$ such that

$$\frac{|\hat{Z}_{(N)}^{(\kappa(N))}|}{N} \overset{a.s.}{\to} 0 \tag{D.15}$$

as $N \to \infty$, where $|\hat{Z}_{(N)}^{(\kappa(N))}|$ is the number of realized individuals of $\hat{Z}_{(N)}$ belonging to a generation $\leq \kappa(N)$. Indeed, as pointed out in the last paragraph of section 3.1, the coupling breaks down when the epidemic reaches a size of order $\sqrt{N}$. Thus, (D.15) is a necessary condition on the generation function κ.

D.3 Backward processes and the size of a major outbreak

In this section, we show that a backward process can be used to approximate the expected final size of the epidemic, as described in section 3.2. In the limit as the population size $N \to \infty$, the expected final size of a major outbreak conditioned on that the approximating forward branching process $\hat{Z}$ avoids extinction is given by the probability of survival of a backward approximating process $\hat{Y}$.

Let v be some node of G_N. Denote the backward epidemic process (on some probability space) corresponding to v by $S^{(N)}(v)$. In the remainder of this section, we suppress the dependence on v and simply write $S^{(N)}$ for $S^{(N)}(v)$. Let the parameter space

$$\Theta = \mathbb{N}_\infty \times \mathbb{N}_\infty \times [0, 1)$$

have elements of the form

$$\theta = (N, m, \varepsilon).$$

Let further $\{\hat{Y}_\theta\}_{\theta \in \Theta}$ be a collection of single-type branching processes with common type space (the $\hat{Y}_\theta$ are actually two-type branching processes, since the ancestor a is of the unique type 0). The space of individuals of these branching processes is denoted by $\mathcal{X}_b$. We use the branching process terminology and notation of section D.1 also for the branching processes considered in this section.

The main idea is as follows. We assume that v is chosen uniformly at random. For each parameter vector $\theta = (N, m, \varepsilon) \in \Theta$ we construct a copy $\hat{S}_\theta$ of the epidemic process $S^{(N)}$. Note that the the population sizes are consistent. In other words, N is the first component of the parameter vector θ. For each θ, we couple the epidemic process with a backward branching process $\hat{Y}_\theta$. This results in a collection $\{\hat{S}_\theta\}_\theta$ of copies of the backward epidemic processes of $\{S_{(N)}\}_N$. Recall that v contracts the disease if and only if some initial case is a member of the susceptibility set $\mathfrak{S}(v)$ of v, as described in section 3.2. If a half-edge of some node of the coupled backward epidemic process $\hat{S}_\theta$ is paired with an infectious half-edge attached to a node of generation $\kappa(N)$ of the stopped forward epidemic process $\hat{E}_{(N)}$ then the initial case belongs to the susceptibility set of v for the specific coupling. We then find a sequence $\{\Theta_b\}_b$ of subsets of the parameter space Θ that satisfies

$$\inf_{\Theta_b}\{N : \ (N, m, \varepsilon) \in \Theta_b\} \to \infty$$

and

$$\begin{aligned}
&\{\hat{Z} \to 0, v_* \notin \mathfrak{S}_{\theta_b}(v)\} \xrightarrow{\hat{P}} \{\hat{Z} \to 0\} \\
&\{\hat{Z} \not\to 0, v_* \notin \mathfrak{S}_{\theta_b}(v)\} \xrightarrow{\hat{P}} \{\hat{Z} \not\to 0, \hat{Y} \to 0\} \\
&\{\hat{Z} \not\to 0, v_* \in \mathfrak{S}_{\theta_b}(v)\} \xrightarrow{\hat{P}} \{\hat{Z} \not\to 0, \hat{Y} \not\to 0\}
\end{aligned} \tag{D.16}$$

as $b \to \infty$ for any parameter sequence $\{\theta_b\}_b$, $\theta_b \in \Theta_b$, where $\mathfrak{S}_{\theta_b}(v)$ is the susceptibility set of v corresponding to the epidemic backward process $\hat{S}_{\theta_b}$, v_* is the initial case and $\hat{P}$ is the probability measure governing the couplings.

D.3.1 Construction of the coupling

We write $\hat{Y}$ for the branching process $\hat{Y}_{(\infty,\infty,0)}$ and $\hat{Y}_{(N)}$ for the branching process $\hat{Y}_{(N,\infty,0)}$. Note that $\hat{Y} = \hat{Y}_{(\infty)}$. The branching processes of the form $\hat{Y}_{(N)}$ plays a similar role for $S_{(N)}$ as $\hat{Z}_{(N)}$ for $\hat{E}_{(N)}$. More specificly, the branching process $\hat{Y}$ plays a similar role as $\hat{Z}$ in the sense that it is the limiting branching process of the approximating branching processes as the population size $N \to \infty$.

We construct the $\hat{Y}_{(N)}$ in a similar manner as the forward branching processes $\hat{Z}_{(N)}$, $N \in \mathbb{N}_\infty$. To this end, let $\{U^{(y)}\}_{y\in\mathcal{X}_b}$ be a collection of independent random variables, uniformly distributed on $[0,1]$ and independent of the forward epidemic process and the forward branching processes constructed in the previous section. Each individual $y \in \mathcal{X}_b$ is assigned the random variable $U^{(y)}$. The individual y of $\hat{Y}$ is then assigned a degree $D_{y(\hat{Y})}$ according to (D.5), with U_x replaced by $U^{(y)}$. Similarly, the individual y of $\hat{B} = \hat{Y}_{(N)}$, $N < \infty$, is assigned a degree $D_{y(\hat{B})}$ according to (D.6), with U_x replaced by $U^{(y)}$.

Analogously to the forward processes, the number of offspring of a realized individual $x \in \mathcal{X}_b$ of $\hat{Y}_{(N)} = \hat{B}$ with degree $D_{x(\hat{B})}$ is distributed as the number of neighbours that would attempt to infect a node of degree $D_{x(\hat{B})} - 1$. For the branching processes of $\{\hat{Y}_{(N)}\}_{N\in\mathbb{N}_\infty}$ the equivalents of assumptions A6-A7 are assumed to hold. That is, for any two branching processes $B_1, B_2 \in \{\hat{Y}_{(N)}\}_{\mathbb{N}_\infty}$, if the individuals $x(B_1)$ and $x(B_2)$ are both realized (or both not realized) and if their degrees coincide, then almost surely $y(B_1)$ is realized if and only $y(B_2)$ is realized for each child y of x.

Analogously to the forward processes, we couple the branching process $\hat{Y}_{(N)}$ with the epidemic processes $\hat{S}_{(N,\infty,0)}$ by letting them coincide. As before, the coupling breaks down if a realized individual of $\hat{Y}_{(N)}$ collides with another realized individual of $\hat{Y}_{(N)}$ or $\hat{Z}_{(N)}$. That is, the coupling breaks down if for some realized individual x of $\hat{Y}_{(N)}$ and node v_i of G_N

$$U^{(x)}, U^{(y_1)} \in \mathcal{E}_i^{(N)}$$

for some realized individual y_1 of $\hat{Y}_{(N)}$ or

$$U^{(x)}, U_{y_2} \in \mathcal{E}_i^{(N)}$$

for some realized individual $y_1(Z_{(N)})$ of generation $|z_2| \leq \kappa(N)$.

As pointed out in the last paragraph of section 3.1, by a birthday problem type of argument collision occurs when the epidemic reaches a size of order $\sqrt{N}$. For this reason, in the limit as $N \to \infty$, the coupling breaks down before the backward process gets connected with the forward process. To address the possibility that the coupling breaks down at a too early stage, we approximate the epidemic backward process by the branching process $\hat{Y}_{(N,m,\varepsilon)}$, rather than $\hat{Y}_{(N)}$, where $\varepsilon > 0$ is some small positive number and m is some large positive integer.

D.3.2 Trimmed branching processes

To address the possibility that the coupling breaks down at a too early stage, we approximate the epidemic backward process by *trimmed* branching processes of the form $\hat{Y}_{(N,m,\epsilon)}$ where $m < \infty$. To this end, we introduce the *trimmed* empirical degree distribution $p_\circ^{(N,m)}$

$$p_\circ^{(N,m)}(k) = \begin{cases} p_\circ^{(N)}(k) & \text{if } 2 \leq k < m \\ p_\circ^{(N)}(1) + \sum_{h\geq m} p_\circ^{(N)}(h) & \text{if } k = 1. \end{cases} \tag{D.17}$$

Note that $p_\circ^{(N,\infty)} = p_\circ^{(N)}$ and $p_\circ^{(N,m)} \to p_\circ^{(\infty,m)}$ as $N \to \infty$. The offspring distribution of the branching process $\hat{Y}_{(N,m,0)}$ is identical to the offspring distribution of $\hat{Y}_{(N)} = \hat{Y}_{(N,\infty,0)}$ except that the size biased empirical degree distribution $p_\circ^{(N)}$ is replaced by its trimmed counterpart $p_\circ^{(N,m)}$ in the construction of $\hat{Y}_{(N,m,0)}$.

More specifically, the branching process $\hat{Y}_{(N,m,0)}$ is constructed by altering the degrees assigned to the individuals of $\hat{Y}_{(N)} = \hat{Y}_{(N,\infty,0)}$; if an individual x of $\hat{Y}_{(N)}$ is assigned

the degree $D_{x(\hat{Y}_{(N)})} \geq m$ then x of $\hat{Y}_{(N,m,0)}$ is assigned the degree 1. Note that if this happens then the individual x of $\hat{Y}_{(N,m,0)}$ have 0 children. Apart from this, the realized individuals of the two coupled branching processes $\hat{Y}_{(N,m,0)}$ and $\hat{Y}_{(N)}$ and their degrees coincide. The construction of the trimmed branching processes is illustrated in Figure D.4.

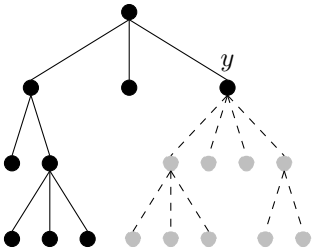

Figure D.4: Construction of the trimmed branching process $\hat{Y}_{(N,m,0)}$ for $m = 5$. The individual y of $\hat{Y}_{(N)}$ is realized and is assigned the degree $D_{x(\hat{Y}_{(N)})} = 5$. The descendants of the individual y of $\hat{Y}_{(N,5,0)}$ are not realized since $D_{x(\hat{Y}_{(N)})} \geq m$.

We now consider the construction of the copy $\hat{S}_{(N,m,0)}$ of the backward epidemic process $S_{(N)}$. As we saw in section 2.5, the order in which the half-edges are paired does not affect the distribution of the topology of G_N, as long as the pairing is uniform. Similar to the case $m = \infty$, we construct a coupling of the epidemic process $\hat{S}_{(N,m,0)}$ and the approximating branching process $\hat{Y}_{(N,m,0)}$ by letting them coincide. As before, the coupling breaks down if a collision occurs. This coupling has the following interpretation. If a node v of degree $k \geq m$ is chosen in the pairing procedure, then we postpone the pairing of the half-edges attached to v until the coupling breaks down.

D.3.3 Erased branching processes

If $\varepsilon > 0$ we refer to $\hat{Y}_{(N,m,\varepsilon)}$ as an *erased* branching process. To address the possibility that the coupling with the backward epidemic process breaks down before the backward epidemic process and the forward epidemic process has become connected or gone extinct, we consider erased branching processes. The idea of erased backward processes was used by Ball et al. (2009, 2014).

Take some trimmed branching process $\hat{Y}_{\theta_0}$, where $\theta_0 = (N, m, 0)$ and let $\theta_\varepsilon = (N, m, \varepsilon)$. The number $\varepsilon \in (0, 1)$ is the sum of two strictly positive terms $\varepsilon = \varepsilon_1 + \varepsilon_2(m, \varepsilon_1)$ (to be specified later in this section). That is, $\varepsilon = \varepsilon(m, \varepsilon_1)$ is a function of m and ε_1. The coupling of $\hat{S}_{\theta_0}$ with $\hat{Y}_{\theta_0}$ breaks down if a node v_i is chosen for the second time in the pairing procedure. This might occur for three reasons. First, an already paired half-edge of the forward process might be chosen in the pairing procedure of the backward process. Second, an already paired half-edge of the backward process might be chosen in the pairing procedure of the backward process. Third, an unpaired half-edge attached to a node that is involved in the backward epidemic process might be chosen in the pairing procedure.

By *erasing* realized individuals of the approximating branching process $\hat{Y}_{\theta_0}$, we ensure that in the limit $N \to \infty$, a collision does not occur until a certain proportion of the nodes of G_N are involved. As we will see later in this section, this allows us to connect the forward and backward epidemic processes on the set $\{\hat{Z} \not\to 0, \hat{Y} \not\to 0\}$ in the limit as $N \to \infty$.

The offspring distribution ν_{θ_ε} of $\hat{Y}_{\theta_\varepsilon} = \hat{Y}_{(N,m,\varepsilon)}$ and the offspring distribution ν_{θ_0} of

$\hat{Y}_{\theta_0} = Y_{(N,m,0)}$ are related as follows

$$\nu_{\theta_\varepsilon}(k) = \sum_{j \geq k} \nu_{\theta_0}(j) \binom{j}{k} (1-\varepsilon)^k \varepsilon^{j-k}. \tag{D.18}$$

That is, erasing each realized individual of $\hat{Y}_{\theta_0}$ independently with probability ε results in a branching process with the same reproduction law as $\hat{Y}_{\theta_\varepsilon}$.

The coupling of the erased branching processes is constructed a follows. Pick some arbitrary (but small) $\varepsilon_1 > 0$. As mentioned above, $\varepsilon = \varepsilon_1 + \varepsilon_2(\varepsilon_1, m)$, where the function ε_2 is yet to be specified. The two branching processes $\hat{Y}_{\theta_0}$ and $\hat{Y}_{\theta_\varepsilon}$ coincide, except that some realized individuals are erased from $\hat{Y}_{\theta_0}$ with marginal probability ε, which results in $\hat{Y}_{\theta_\varepsilon}$. This happens independently of the degree and other characteristics of the individuals of $\hat{Y}_{\theta_0}$. More specifically, we construct $\hat{Y}_{\theta_\varepsilon}$ from $\hat{Y}_{\theta_0}$ by erasing the individuals of $\hat{Y}_{\theta_0}$ that collide. We erase some additional individuals (according to some rule which we do not specify since the details are not important here) so that that the marginal probability of being erased is ε, independently of the degree and other characteristics of the individual. The construction of the coupling of the erased processes is illustrated in Figure D.5.

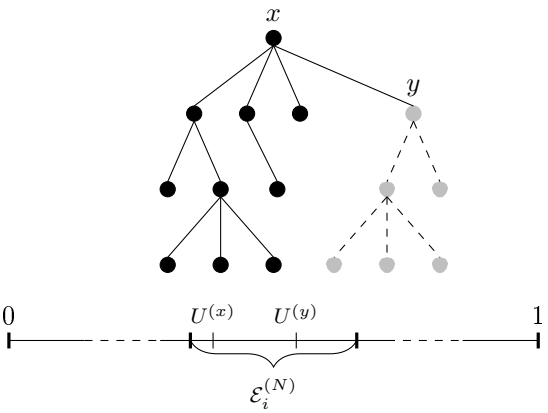

Figure D.5: Construction of the erased branching process $\hat{Y}_{\theta_\varepsilon}$. The individual $y(\hat{Y}_{\theta_\varepsilon})$ and its descendants are erased since $x(\hat{Y}_{\theta_0})$ and $y(\hat{Y}_{\theta_0})$ collide.

The construction of $\hat{Y}_{\theta_\varepsilon}$ breaks down if the marginal probability of erasing an individual is larger than ε, or if the event that an individual x of $\hat{Y}_{\theta_\varepsilon}$ is erased is not independent of its degree. The latter happens if the proportion of involved nodes of (trimmed) degree k is large enough for some degree k. There are two ways in which an node v can take part in a collision in the construction of the backward epidemic process $\hat{E}_{\theta_\varepsilon}$. First, a (paired or unpaired) half-edge attached to an already involved node u might be chosen when some of the half-edges attached to v is paired. If this happens, the individual corresponding to the half-edge attached to v is erased. Second, a (paired or unpaired) half-edge attached to v might be chosen in the pairing procedure after v has become involved. If the chosen half-edge is unpaired, then the corresponding individual of $\hat{Y}_{\theta_\varepsilon}$ is erased.

The idea is as follows. We run the backward processes as long as the proportion of involved nodes of trimmed degree k is at most ε_1 for each k, and the probability that some unpaired half-edge attached to a node involved in the backward process is chosen in the pairing procedure is at most $\varepsilon_2(m, \varepsilon_2)$ (the function $\varepsilon_2(m, \varepsilon_2)$ is specified below). This ensures that the event that an individual x of $\hat{Y}_{\theta_\varepsilon}$ is erased is independent of its degree. Since the degrees are assigned from the trimmed empirical degree distribution $p_\circ^{(N,m)}$ we are guaranteed that the proportion of involved nodes of trimmed degree k is at most ε_1 for each k until at least

$$\varepsilon_1 \inf\left\{\left|\mathcal{V}_k^{(N,m)}\right| :\ \mathcal{V}_k^{(N,m)} \neq \varnothing,\ 1 \leq k \leq m\right\} - \left|Z_{(N)}^{(\kappa(N))}\right| \tag{D.19}$$

nodes of G_N are involved. Here $|Z_{(N)}^{(\kappa(N))}|$ is the number of realized individuals x of the forward approximating branching process $Z_{(N)}$ stopped in generation $\kappa(N)$, and $\mathcal{V}_k^{(N,m)}$ is the set of the nodes of G_N that are assigned the trimmed degree k. That is

$$\mathcal{V}_k^{(N,m)} = \begin{cases} \mathcal{V}_1^{(N)} \cup \left(\bigcup_{j \geq k} \mathcal{V}_j^{(N)}\right) & \text{if } k = 1. \\ \mathcal{V}_k^{(N)} & \text{if } 2 \leq k \leq m-1 \\ \varnothing & \text{otherwise.} \end{cases}$$

Dividing by N in (D.19) and using $|Z^{(\kappa(N))}|/N \to 0$ as $N \to \infty$ yields that in the limit as $N \to \infty$, a lower bound on the proportion of the nodes of G_N that are involved when for some degree k an individual of degree k is realized and the proportion of involved nodes of degree k is larger than ε_1 is given by

$$\varepsilon_1 \delta_m > 0, \tag{D.20}$$

where

$$\delta_m := \min\left\{p_\circ^{(\infty,m)}(k) :\ p_\circ^{(\infty,m)}(k) > 0\right\}. \tag{D.21}$$

It can be shown (Ball et al. 2009) that for fixed m and $\varepsilon_1 > 0$ small enough, there exists a strictly positive function $\varepsilon_2(\varepsilon_1, m)$ such that in the limit as the population size $N \to \infty$, the (marginal) probability that an unpaired half-edge attached to a node that is involved in the backward epidemic process is chosen in the pairing procedure is bounded from above by $\varepsilon_2(\varepsilon_1, m)$ until the backward process reaches a size of at least $\lfloor N\varepsilon_1\delta_m \rfloor$. Moreover, $\varepsilon_2(\varepsilon_1, m) \to 0$ as $\varepsilon_1 \to 0$ for fixed m. That is, $\varepsilon \to 0$ as $\varepsilon_1 \to 0$ for fixed m. The proof, which is to be found in Ball et al. (2009), is omitted. That is, in the limit $N \to \infty$, we may pair at least a fraction $\varepsilon_1\delta_m$ of the half-edges, before the coupling breaks down. Note that the coupling might not break down, since $\hat{Y}_{\theta_\varepsilon}$ might become extinct.

The coupling of the erased processes has an important interpretation, similar to the interpretation for the trimmed processes. Recall that the order in which the half-edges are paired does not affect the distribution of the topology of the network, provided uniform pairing. The coupling of $\hat{S}_{\theta_\varepsilon}$ and the epidemic process $S_{(N)}$ corresponding to the erased backward process $\hat{Y}_{\theta_\varepsilon}$ has the following interpretation: In the pairing procedure we choose half-edges (paired or unpaired) uniformly at random among all half-edges of G_N. If we choose an already paired half-edge, then the pairing of the associated half-edge is postponed until after the coupling breaks down. If we choose an unpaired half-edge attached to an already involved node, then the half-edges are paired, and both of the corresponding individuals are erased from $\hat{Y}_{\theta_\varepsilon}$.

D.3.4 Connecting the processes

Let $\{\Theta_b\}_{b\in\mathbb{N}}$ be a sequence of subsets of the parameter space Θ such that

$$\begin{aligned} &\inf_{\Theta_b}\{N :\ (N,m,\varepsilon) \in \Theta_b\} \to \infty \\ &\inf_{\Theta_b}\{m :\ (N,m,\varepsilon) \in \Theta_b\} \to \infty \\ &\sup_{\Theta_b}\{\varepsilon :\ (N,m,\varepsilon) \in \Theta_b\} \to 0 \end{aligned} \tag{D.22}$$

as $b \to \infty$.

We write $\liminf_{\Theta_b}$ for $\lim_{b\to\infty} \inf_{\theta\in\Theta_b}$. Repeating the analysis performed in section D.2.1 gives

$$\{\hat{Y} \to 0\} \overset{\hat{P}}{\subset} \liminf_{N} \{\hat{Y}_{(N)} \to 0\} \subset \liminf_{\Theta_b} \{\hat{Y}_\theta \to 0\} \tag{D.23}$$

where $\overset{\hat{P}}{\subset}$ denotes inclusion up to a set of $\hat{P}$-measure zero. That is, if $A \overset{\hat{P}}{\subset} B$ then $\hat{P}(A \setminus B) = 0$. Note that (Friedman 1982, Theorem 1.2.1-1.2.2)

$$\begin{aligned}
\limsup_{\Theta_b} P(\{\hat{Y} \to 0\} \setminus \{\hat{Y}_\theta \to 0\}) &\leq \lim_{b} P(\{\hat{Y} \to 0\} \setminus \inf_{\Theta_b} \{\hat{Y}_\theta \to 0\}) \\
&= P(\{\hat{Y} \to 0\} \setminus \liminf_{\Theta_b} \{\hat{Y}_\theta \to 0\}) \\
&= 0
\end{aligned} \tag{D.24}$$

and

$$\begin{aligned}
&\limsup_{\Theta_b} P(\{\hat{Y}_\theta \to 0\} \setminus \{\hat{Y} \to 0\}) \\
&= \limsup_{\Theta_b} (P(\{\hat{Y}_\theta \to 0\}) - P(\{\hat{Y} \to 0\} \cap \{\hat{Y}_\theta \to 0\})) \\
&\leq \limsup_{\Theta_b} P(\{\hat{Y}_\theta \to 0\}) - \liminf_{\Theta_b} P(\{\hat{Y} \to 0\} \cap \{\hat{Y}_\theta \to \bar{0}\}) \\
&\leq \limsup_{\Theta_b} P(\{\hat{Y}_\theta \to 0\}) - P(\{\hat{Y} \to 0\} \cap \liminf_{\Theta_b} \{\hat{Y}_\theta \to 0\}) \\
&= \limsup_{\Theta_b} P(\{\hat{Y}_\theta \to 0\}) - P(\{\hat{Y} \to 0\}\}) \\
&= 0,
\end{aligned} \tag{D.25}$$

where the second last step follows from (D.23) and the last step follows from Theorem D.4 below. For proofs of the special case $s = 1$ of Theorem D.4, see Britton et al. (2007, Lemma 4.1) or Leskelä and Ngo (2017, Lemma 2.6).

In view of (D.24) and (D.25)

$$\begin{aligned}
\{\hat{Y}_{\theta_b} \to 0\} &\overset{\hat{P}}{\to} \{\hat{Y} \to 0\} \\
\{\hat{Y}_{\theta_b} \not\to 0\} &\overset{\hat{P}}{\to} \{\hat{Y} \not\to 0\}
\end{aligned} \tag{D.26}$$

for any sequence $\{\theta_b\}_b$ such that $\theta_b \in \Theta_b$.

Theorem D.4. *Let Z be a non-singular positively regular s-type Galton-Watson branching process with extinction probabilities $\bar{q}$ and offspring distribution ν. For each $n \in \mathbb{N}$, let $Z_{(n)}$ be an s-type Galton-Watson branching processes with extinction probabilities $\bar{q}_{(n)}$ and offspring distribution $\nu^{(n)}$. If*

$$\nu \overset{d}{\to} \nu^{(n)} \text{ as } n \to \infty,$$

where $\overset{d}{\to}$ denotes convergence in distribution, then

$$\bar{q}_{(n)} \to \bar{q} \text{ as } n \to \infty$$

Proof. By Fatous lemma (Friedman 1982, Theorem 2.10.5)

$$\liminf_{n} \lim_{k} P(Z_{(n)} \text{ has gone extinct before generation } k) \geq \bar{q},$$

where the inequality holds element-wise, and lim inf is taken element-wise. If $\bar{q} = \bar{1}$, the assertion follows.

Assume that $|\bar{q}| < 1$. We proceed by contradiction. Assume that $\bar{q}_{(n)} \not\to \bar{q}$ as $n \to \infty$. That is, there exists a subsequence $\{\bar{q}_{(n_k)}\}_k$ of $\{\bar{q}_{(n)}\}_n$ such that $\bar{q}$ is not an accumulation point of $\{\bar{q}_{(n_k)}\}_k$. Assume without loss of generality that $\{\bar{q}_{(n_k)}\}_k$ is the whole sequence $\{\bar{q}_{(n)}\}_n$.

We have that $\{\bar{q}_{(n)}\}_n$ is a sequence in $[0,1]^s$, and since $[0,1]^s$ is compact, closed and bounded it is sequentially compact (Rudin 1976, Theorem 2.40-2.41). That is, there exists some subsequence of $\{\bar{q}_{(n)}\}_n$ that converge to some point $\bar{q}_*$ of $[0,1]^s$. Assume without loss of generality that

$$\bar{q}_{(n)} \to \bar{q}_*. \tag{D.27}$$

Let f_Z and $f_{Z_{(n)}}$ be the probability generating functions of the offspring distributions of Z and $Z_{(n)}$, respectively. For two vectors $\bar{z} = (z_1, \ldots, z_s)$ and $\bar{k} = (k_1, \ldots, k_s)$, we define $\bar{z}^{\,\bar{k}} = z_1^{k_1} \cdot \ldots \cdot z_s^{k_s}$. By the triangle inequality, for any $\bar{z} \in [0,1]^s$ and $\bar{y} \in [0,1]^s$

$$\begin{aligned}
|(f_Z(\bar{z}))_i - (f_{Z_{(n)}}(\bar{x}))_i| &\leq \sum_{k \in \mathbb{N}_0^s} |\nu_i(k)\bar{z}^k - \nu_i^{(n)}(k)\bar{x}^k| \\
&\leq \sum_{k \in \mathbb{N}_0^s} |\nu_i(k)\bar{z}^k - \nu_i(k)\bar{x}^k| + \sum_{k \in \mathbb{N}_0^s} |\nu_i(k)\bar{x}^k - \nu_i^{(n)}(k)\bar{x}^k| \\
&\leq \sum_{k \in \mathbb{N}_0^s} \nu_i(k)|\bar{z}^k - \bar{x}^k| + \sum_{k \in \mathbb{N}_0^s} |\nu_i(k) - \nu_i^{(n)}(k)|
\end{aligned} \tag{D.28}$$

for $1 \leq i \leq s$, where $\nu = (\nu_1, \ldots, \nu_s)$ and $\nu^{(n)} = (\nu_1^{(n)}, \ldots, \nu_s^{(n)})$.

Note that convergence in distribution implies convergence in total variation norm for discrete random variables, that is

$$\sum_k |\nu_i(k) - \nu_i^{(n)}(k)| \to 0 \tag{D.29}$$

as $n \to \infty$.

Indeed, let $\mathcal{K}$ be some subset of $\mathbb{N}_0^s$ of finite cardinality. Then

$$\begin{aligned}
\sum_{k \in \mathbb{N}_0^s} |\nu_i(k) - \nu_i^{(n)}(k)| &\leq \sum_{k \in \mathcal{K}} |\nu_i(k) - \nu_i^{(n)}(k)| + \sum_{k \in K^c} \nu_i(k) + \sum_{k \in \mathcal{K}^c} \nu_i^{(n)}(k) \\
&= \sum_{k \in \mathcal{K}} |\nu_i(k) - \nu_i^{(n)}(k)| + 2 - \sum_{k \in K} \nu_i(k) - \sum_{k \in \mathcal{K}} \nu_i^{(n)}(k).
\end{aligned} \tag{D.30}$$

We see that for any $\epsilon > 0$, we may choose $\mathcal{K}$ such that for n large enough, the right hand side of (D.30) is smaller than ϵ. Hence the assertion (D.29) holds.

Combining (D.27), (D.28) and (D.29) gives $\bar{q}_{(n)} = f_{Z_{(n)}}(\bar{q}_{(n)}) \to f_Z(\bar{q}_*)$. Hence $f(\bar{q}_*) = \bar{q}_*$. Since Z is assumed to be positively regular and non-singular, the only fixed points of f_Z are $\bar{1}$ and $\bar{q}$. Thus $\bar{q}_* = \bar{1}$ or $\bar{q}_* = \bar{q}$. By assumption, $\bar{q}_* = \bar{1}$.

We now derive a contradiction by showing that $\bar{q}_{(n)}$ is bounded away from $\bar{1}$ for large n.

By Theorem 2.2, the Perron root $\rho(M)$ of the mean matrix M of Z satisfies $\rho(M) > 1$. Let $\underline{\mathcal{P}}$ be the collection of offspring distributions of s-type Galton-Watson processes that

have finite support and are dominated by ν. That is, for each $\eta = (\eta_1, \ldots, \eta_s) \in \underline{\mathcal{P}}$ we have that η_i has finite support and for each $k \in \mathbb{N}_0^s$, $k \neq \bar{0}$

$$\begin{aligned} &\eta_i(k) < \nu_i(k) \text{ if } \nu_i(k) > 0 \\ &\eta_i(k) = \nu_i(k) \text{ if } \nu_i(k) = 0 \end{aligned} \tag{D.31}$$

for $1 \leq i \leq s$. Note that η_i puts more probability mass on $\bar{0}$ than ν_i. For any $\eta \in \underline{\mathcal{P}}$, let Z_η be an s-type branching process with offspring distribution η, and denote the corresponding mean matrix by M_η.

By the Collatz-Wielandt formula (Theorem D.3) and monotone convergence (Friedman 1982, Theorem 2.10.4), we have

$$\sup_{\eta \in \underline{\mathcal{P}}} \rho(M_\eta) = \rho(M).$$

Take some $\eta \in \underline{\mathcal{P}}$ such that $\rho(M_\eta) > 1$ and Z_η is positively regular. It follows that the extinction probabilities $\bar{q}_\eta$ of Z_η satisfies $|\bar{q}_\eta| < 1$. Since the offspring distribution η of Z_η is dominated by the offspring distribution ν of Z and η has finite support, there exists some n' such that $\nu^{(n)}$ dominates η for all $n \geq n'$. By a simple coupling argument, $\bar{q}_{(n)} \leq \bar{q}_\eta$ for all $n \geq n'$, where the inequality holds element-wise. Thus, by contradiction, $\bar{q}_{(n)} \to \bar{q}$ as $n \to \infty$. □

Now consider the coupling $\{\hat{S}_\theta, \hat{Y}_\theta\}$ of the epidemic backward process $S_{(N)}$ and the branching process $\hat{Y}_\theta$. There are three possibilities.

*P*1) The coupling breaks down before $\hat{Y}_\theta$ goes extinct and before the forward and backward processes become connected.

*P*2) The forward and backward process become connected before the coupling breaks down and before $\hat{Y}_\theta$ goes extinct.

*P*3) $\hat{Y}_\theta$ goes extinct before the forward and backward process become connected and before the coupling breaks down.

Take some fixed m' and $\varepsilon' = \varepsilon_1' + \varepsilon_2'(m', \varepsilon_1')$, where $\varepsilon_1' > 0$ is "small enough". For any population size N, denote the number of infectious unpaired half-edges attached to an individual belonging to generation $\kappa(N)$ of the limiting forward branching process $\hat{Z}$ by $\Psi_{(N)}$. For each $\theta_N = (N, m', \varepsilon')$, we introduce the following three stopping times in $\mathbb{N}_\infty$, each corresponding to one of the possibilities *P*1-*P*3.

1. $T_{\theta_N}^{(b)}$ is the number of nodes involved in the backward process when the coupling of $\hat{Y}_{\theta_N}$ and $S_{(N)}$ **b**reaks down.
2. $T_{\theta_N}^{(e)}$ is the total number of realilized individuals of $\hat{Y}_{\theta_N}$ when $\hat{Y}_{\theta_N}$ becomes **e**xtinct.
3. $T_{\theta_N}^{(c)}$ is the number of half-edges drawn when one of the $\Psi_{(N)}$ infectious half-edges is drawn and the corresponding individual is realized.

That is, the forward and backward epidemic processes become **c**onnected if $T_{\theta_N}^{(c)} < \min(T_{\theta_N}^{(b)}, T_{\theta_N}^{(e)})$.

We first note that on the set $\{\hat{Z} \to 0\}$, in the limit as $N \to \infty$ we have that the initial case v_* can only (up to a set of $\hat{P}$-measure zero) be a member of the susceptibility set $\mathfrak{S}_{\theta_N}(v)$ of v if v is one of the $\lim_N \Psi_{(N)} < \infty$ nodes involved in the forward process $\hat{Z}$. This follows from the fact that the forward epidemic process $\hat{E}_{(N)}$ coincides with the limiting forward process $\hat{Z}$ for all but finitely many values of N almost surely. If v is chosen uniformly at random then the event $\{\hat{Y} \to 0, v_* \in \mathfrak{S}_{\theta_N}(v)\}$ converges to zero in $\hat{P}$-measure as $N \to \infty$.

Recall that in the limit $N \to \infty$ the coupling does not break down until the proportion given in (D.20) is involved. Thus

$$\liminf_N \frac{T^{(b)}_{\theta_N}}{N} \geq \varepsilon_1' \delta_{m'} > 0 \tag{D.32}$$

almost surely. Furthermore,

$$T^{(c)}_{\theta_N} \xrightarrow{\hat{P}} \infty \tag{D.33}$$

as $N \to \infty$. Indeed, for fixed θ_N, we may view the pairing procedure as choosing elements (half-edges) uniformly at random with replacement from a set consisting of $D^{(N)}$ elements, where $D^{(N)}$ is the total degree of G_N. Some of these half-edges/elements correspond to realized individuals, some correspond to erased individuals. Let the random variable $T^{(N)}$ have a geometric distribution with parameter $\frac{\Psi_{(N)}}{N(1-\varepsilon')} \wedge 1$, where $\wedge$ denotes the minimum function. Note that the probability that a specific half-edge is one of the $\Psi_{(N)}$ infectious half edges attached to a node belonging to generation $\kappa(N)$ of the epidemic forward process $E_{(N)}$ is bounded from above by $\frac{\Psi_{(N)}}{N(1-\varepsilon')}$, provided that the corresponding individual is not erased. Hence, conditioned on $\Psi_{(N)}$, $T^{(c)}_{\theta_N}$ is stochastically larger than $T^{(N)}$. Now

$$\left(1 - 1 \wedge \frac{\Psi_{(N)}}{N(1-\varepsilon')}\right)^t \to 1 \tag{D.34}$$

almost surely as $N \to \infty$ for each fixed $t \in \mathbb{N}$. Since the left hand side in (D.34) is bounded by 1, the convergence holds also in mean (Grimmett and Stirzaker 1992, p. 277) and the assertion in (D.33) follows.

Since the offspring distribution of $\hat{Y}_{\theta_N}$ converge to the offspring distribution of $\hat{Y}_{(\infty,m',\varepsilon')}$ as $N \to \infty$, by (D.32) and (D.33)

$$\hat{P}\left(\hat{Y}_{\theta_N} \to 0, \max(\, T^{(b)}_{\theta_N}, T^{(c)}_{\theta_N}) < T^{(e)}_{\theta_N}\right) \to 0 \tag{D.35}$$

as $N \to \infty$.

Let the random variable $T^{(N)}_{\varepsilon'}$ have a geometric distribution with parameter $\frac{\Psi_{(N)}}{N} \wedge 1$. Note that, conditioned on $\Psi_{(N)}$, it holds that $T^{(c)}_{\theta_N} \wedge N\varepsilon'\delta_{m'}$ is stochastically smaller than $T^{(N)}_{\varepsilon'}$.

Now, by the standard limit $\lim_{x\to\infty}(1-1/x)^x = e^{-1}$ we have for any constant $c > 0$

$$\left(1 - \frac{\Psi_{(N)}}{N} \wedge 1\right)^{\lfloor cN \rfloor} \xrightarrow{a.s} 0 \tag{D.36}$$

as $N \to \infty$ on $\{Z \not\to 0\}$.

Thus

$$\hat{P}(T^{(c)}_{\theta_N} \geq cN, Z \not\to 0) \to 0 \tag{D.37}$$

as $N \to \infty$ for any constant $c > 0$.

In view of (D.32) and (D.37)

$$\hat{P}(\hat{Y}_{\theta_N} \not\to 0, \max(T^{(b)}_{\theta_N}, T^{(e)}_{\theta_N}) < T^{(c)}_{\theta_N}, \hat{Z} \not\to 0) \to 0 \tag{D.38}$$

as $N \to \infty$.

Take some sequence $\{(m_b, \varepsilon^{(b)})\}_{b\in\mathbb{N}}$ in $\mathbb{N}\times(0,1)$ such that $m_b \to \infty$ and $\varepsilon^{(b)} = \varepsilon_1^{(b)} + \varepsilon_2^{(b)}(m_b, \varepsilon_1^{(b)}) \to 0$ as $b \to \infty$. For each b, choose some N_b such that for any parameter vector

$$\theta_b \in \Theta_b := \{(N, m_b, \varepsilon^{(b)}) : \ N \geq N_b\}$$

it holds that

$$\hat{P}(\hat{Z} \not\to 0, \hat{Y}_{\theta_b} \to 0, \max(\ T^{(b)}_{\theta_b}, T^{(c)}_{\theta_b}) < T^{(e)}_{\theta_b}) \leq \frac{1}{b}$$

$$\hat{P}(\hat{Z} \not\to 0, \hat{Y}_{\theta_b} \not\to 0, \max(T^{(b)}_{\theta_b}, T^{(e)}_{\theta_b}) < T^{(c)}_{\theta_b}) \leq \frac{1}{b}.$$

This implies

$$\begin{aligned}
&\hat{P}(\hat{Z} \not\to 0, \hat{Y}_{\theta_b} \to 0, v_* \in \mathfrak{S}_{\theta_b}(v)) \leq \frac{1}{b} \\
&\hat{P}(\hat{Z} \not\to 0, \hat{Y}_{\theta_b} \not\to 0, v_* \notin \mathfrak{S}_{\theta_b}(v))) \leq \frac{1}{b}.
\end{aligned} \tag{D.39}$$

for every $\theta_b \in \Theta_b$, where v_* is the initial case and $\mathfrak{S}_\theta(v)$ is the susceptibility set of v corresponding to $\hat{S}_{\theta_b}$. Combining (D.39) with (D.26) gives (D.16). Since $\hat{Z}$ and $\hat{Y}$ are independent, it holds that

$$\begin{aligned}
&\hat{P}(v_* \in \mathfrak{S}_{\theta_b}(v) | \hat{Z} \to 0) \to 0 \\
&\hat{P}(v_* \in \mathfrak{S}_{\theta_b}(v) | \hat{Z} \not\to 0) \to \hat{P}(\hat{Y} \not\to 0) \\
&\hat{P}(v_* \notin \mathfrak{S}_{\theta_b}(v) | \hat{Z} \not\to 0) \to \hat{P}(\hat{Y} \to 0)
\end{aligned} \tag{D.40}$$

as $b \to \infty$.

Thus, in the limit as the population size $N \to \infty$, the expected fraction of the population ultimately infected is given by the probability $P(\hat{Y} \not\to 0)$ that the limiting backward branching process $\hat{Y}$ avoids extinction, conditioned on that a major outbreak occurs.

D.4 Maximal coupling of branching processes

In this section, we show that the coupling of the approximating branching processes of $\mathcal{Z} = \{\hat{Z}_N\}_{N\in\mathbb{N}_\infty}$ is maximal at every index.

Recall that, for a fixed generation n, two branching processes B_1 and B_2 are said to coincide up to generation n if $R_{x(B_1)} = R_{x(B_2)}$ and $D_{x(B_1)} = D_{x(B_2)}$ whenever $R_{x(B_1)} = 1$ for each individual such that $|x| \leq n$. That is, B_1 and B_2 coincide up to generation n if the realized individuals and their degrees coincide up to generation n.

We may characterize the fate of a branching process B up to generation n by the (random) set

$$\Sigma_B^{(n)} := \{(x, D_{x(B)}) : |x| \leq n, R_{x(B)} = 1\}$$

Note that the random element $\Sigma_B^{(n)}$ has a countable support, $S_\Sigma^{(n)}$ say, for each fixed generation n, and that two processes B_1 and B_2 coincide up to generation n if and only if

$$\Sigma_{B_1}^{(n)} = \Sigma_{B_2}^{(n)}.$$

We now consider any coupling $\hat{\hat{\mathcal{Z}}} = \{\hat{\hat{Z}}_{(N)}\}_{N\in\mathbb{N}_\infty}$ of $\mathcal{Z}$ and some coupling index $K^{(n)}$ of $\{\hat{\hat{Z}}^{(n)}_{(N)}\}_{N\in\mathbb{N}_\infty}$. Let $\hat{\hat{\mathcal{Z}}}_N = \{\hat{\hat{Z}}^{(n)}_{(N)}\}_{N\leq N'\leq\infty}$ and let $\hat{\hat{P}}$ be the probability measure governing the coupling $\hat{\hat{\mathcal{Z}}}$.

We have for each population size $N' \in \mathbb{N}$

$$\begin{aligned}\hat{\hat{P}}(K^{(n)} \leq N') &= \hat{\hat{P}}\left(\Sigma_B^{(n)} \text{ coincide for each } B \in \hat{\hat{\mathcal{Z}}}_{N'}\}\right) \\ &= \sum_{\sigma \in S_\Sigma^{(n)}} \hat{\hat{P}}\left(\Sigma_B^{(n)} = \sigma \text{ for each } B \in \hat{\hat{\mathcal{Z}}}_{N'}\}\right) \\ &\leq \sum_{\sigma \in S_\Sigma^{(n)}} \inf_{B \in \hat{\hat{\mathcal{Z}}}_{N'}} \hat{\hat{P}}\left(\Sigma_B^{(n)} = \sigma\right).\end{aligned} \tag{D.41}$$

By construction, equality holds in (D.41) for the coupling constructed previously in this section. We conclude that for each generation n, the coupling of the branching processes of $\mathcal{Z}$ up to generation n constructed in this section is maximal at each index.

References

Athreya, Krishna and Peter Ney (1972). *Branching processes*. Springer-Verlag.

Ball, Frank, David Sirl, and Pieter Trapman (2009). "Threshold behaviour and final outcome of an epidemic on a random network with household structure". In: *Advances in Applied Probability* 41.3, pp. 765–796.

– (2010). "Analysis of a stochastic SIR epidemic on a random network incorporating household structure". In: *Mathematical Biosciences* 224.2, pp. 53 –73.

– (2014). "Epidemics on random intersection graphs". In: *The Annals of Applied Probability* 24.3, pp. 1081–1128.

Ball, Frank, Lorenzo Pellis, and Pieter Trapman (2016). "Reproduction numbers for epidemic models with households and other social structures II: Comparisons and implications for vaccination". In: *Mathematical Biosciences* 274, pp. 108 –139.

Bellman, Richard and Robert Kalaba (1967). *Modern Analytic and Computational Methods in Science and Mathematics*. American Elsevier Publishing Co.

Bollobás, Béla, Svante Janson, and Oliver Riordan (2007). "The Phase Transition in Inhomogeneous Random Graphs". In: *Random Struct. Algorithms* 31.1, pp. 3–122.

Britton, Tom (2010). "Stochastic epidemic models: A survey". In: *Mathematical Biosciences* 225.1, pp. 24 –35.

Britton, Tom, Svante Janson, and Anders Martin-Löf (2007). "Graphs with Specified Degree Distributions, Simple Epidemics, and Local Vaccination Strategies". In: *Advances in Applied Probability* 39.4, pp. 922–948.

Deijfen, Maria and Willemien Kets (2009). "Random intersection graphs with tunable degree distribution and clustering". Version 1. In: *Probability in the Engineering and Informational Sciences* 23, pp. 661–674. arXiv: 1509.07019v1.

Friedman, Avner (1982). *Foundations of modern analysis*. Dover Publications Inc.

Grimmett, Geoffrey and David Stirzaker (1992). *Probability and random processes*. Oxford: Oxford University Press.

Haccou, Patsy, Peter Jagers, and Vladimir Vatutin (2005). *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press.

Harbeck, Michaela et al. (2013). "Yersinia pestis DNA from Skeletal Remains from the 6th Century AD Reveals Insights into Justinianic Plague". In: *PLOS Pathogens* 9.5, pp. 1–8.

Heesterbeek, Hans (2002). "A Brief History of R_0 and a Recipe for its Calculation". In: *Acta Biotheoretica* 50.3, pp. 189–204.

van der Hofstad, Remco (2016). Random Graphs and Complex Networks Volume I. https://www.win.tue.nl/~rhofstad/NotesRGCN.pdf.

Jagers, Peter (1975). *Branching processes with Biological Applications*. John Wiley & Sons.

Kuulasmaa, Kari (1982). "The Spatial General Epidemic and Locally Dependent Random Graphs". In: *Journal of Applied Probability* 19.4, pp. 745–758.

Leskelä, Lasse and Hoa Ngo (2017). "The impact of degree variability on connectivity properties of large networks". Version 3. In: *Internet Mathematics* 13, pp. 1–24. arXiv: 1508.03379v3.

Liggett, Thomas Milton (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Berlin: Springer.

Meester, Ronald and Pieter Trapman (2011). "Bounding the Size and Probability of Epidemics on Networks". In: *Advances in Applied Probability* 43.2, pp. 335–347.

Meyer, Carl (2000). *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics.

Miller, Joel (2007). "Epidemic size and probability in populations with heterogeneous infectivity and susceptibility". In: *Physical Review E* 76, 010101(R).

– (2008). "Bounding the Size and Probability of Epidemics on Networks". In: *Journal of Applied Probability* 45.2, pp. 498–512.

Miller, Joel (2009). "Percolation and epidemics in random clustered networks". In: *Phys. Rev. E* 80 (2), p. 020901.

Mosteller, Frederick (1962). "Understanding the birthday problem". In: *The Mathematics Teacher* 55.5, pp. 322–325.

Newman, Mark (2009). "Random Graphs with Clustering". In: *Phys. Rev. Lett.* 103 (5), p. 058701.

Pellis, Lorenzo, Frank Ball, and Pieter Trapman (2012). "Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R_0". In: *Mathematical Biosciences* 235.1, pp. 85–97.

Pellis, Lorenzo, Simon Spencer, and Thomas House (2015). "Real-time growth rate for general stochastic SIR epidemics on unclustered networks". In: *Mathematical Biosciences* 265, pp. 65 –81.

Rudin, Walter (1976). *Principles of mathematical analysis*. 3d ed. McGraw-Hill New York.

– (1987). *Real and Complex Analysis, 3rd Ed.* New York, NY, USA: McGraw-Hill, Inc.

Taubenberger, Jeffery K and David M Morens (2006). "1918 influenza: the mother of all pandemics". In: *Rev Biomed* 17, pp. 69–79.

Thorisson, Hermann (2000). *Coupling, stationarity, and regeneration*. New York: Springer-Verlag Inc.

Tibayrenc, Michel. (2007). *Encyclopedia of Infectious Diseases: Modern Methodologies*. Hoboken, N.J.: Wiley-Liss.

Trapman, Pieter (2007). "On analytical approaches to epidemics on networks". In: *Theoretical Population Biology* 71.2, pp. 160 –173.