Technische Universität München

Fakultät für Mathematik

Contact tracing on stochastic graphs

Augustine Okebunor Okolie

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:	Prof. Dr. Christina Kuttler
Prüfer der Dissertation:	Prof. Dr. Johannes Müller Prof. Dr. Mirjam Kretzschmar (Utrecht University)

Die Dissertation wurde am 22.06.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 16.08.2022 angenommen.

With my signature below, I assert that the work in this thesis has been composed by myself independently and no source materials or aids other than those mentioned in the thesis have been used.

München, November 16, 2022 _____

Place, Date Signature

München, November 16, 2022 _____

Place, Date Signature

Kurzfassung

Wir betrachten ein stochastisches (Susceptible-Infected-Recovered) SIR Modell mit Kontaktverfolgung auf Zufallsbäumen und dem Konfigurationsmodell. Auf einem verwurzelten Baum, bei dem anfangs alle Individuen bis auf die Wurzel, die infiziert ist, anfällig sind, können wir exakte Formeln für die Verteilung der infektiösen Periode finden. Dazu zeigen wir, wie man die bestehende Theorie zur Kontaktverfolgung in homogen gemischten Populationen auf Bäume ausdehnen kann. Basierend auf diesen Formeln diskutieren wir den Einfluss von Zufälligkeiten im Baum auf die Basisreproduktionszahl. Wir finden die bekannten Ergebnisse für den Fall der homogenen Durchmischung als Grenzfall des vorliegenden Modells (baumförmiger Kontaktgraph). Darüber hinaus entwickeln wir approximative Meanfield Gleichungen für die Dynamik auf Bäumen und - unter Verwendung der Message Passing Methode - auch für das Konfigurationsmodell. Die Interpretation und Implikationen der Ergebnisse werden diskutiert.

Weitere Untersuchungen zur Verwendung dieses Baummodells zur Schätzung von Parametern werden ebenfalls betrachtet. Wir verwenden einen Maximum Likelihood Schätzer, der auf einem Multi-Typ-Geburts-Todes Verzweigungsmodell für eine heterogene Population basiert, in der jeder Wirt durch seine Übertragungsgruppe charakterisiert ist. Wir erweitern dieses Modell zu einem baumbasierten SIR Modell, bei dem ein Individuum durch seinen Grad charakterisiert wird. Wir beginnen mit einem einzelnen Individuum als Ausgangspunkt und gehen dann von der Betrachtung eines Individuums zu einer umfassenderen Betrachtung von mehreren Individuen über. Wir sind in der Lage, einen Maximum-Likelihood-Schätzer für die Parameter des Basismodells zu konstruieren. Anschließend kombinieren wir die Ergebnisse der Maximum Likelihood Schätzer des Basismodells mit den Ergebnissen der Kontaktverfolgung. Die Rückverfolgung von Kontakten enthält Informationen über die zugrundeliegende Kontaktstruktur, während die phylogenetischen Methoden besser geeignet sind den zeitlichen Ablauf des Infektionsprozesses abzuschätzen. Die kombinierte Methode verbessert die Schätzung der Parameter der epidemischen Prozesse sowie der zugrunde liegenden Kontaktstruktur.

Abstract

We consider a stochastic susceptible-infected-recovered (SIR) model with contact tracing on random trees and the configuration model. On a rooted tree, where initially all individuals are susceptible apart from the root which is infected, we are able to find exact formulas for the distribution of the infectious period. Thereto, we show how to extend the existing theory for contact tracing in homogeneously mixing populations to trees. Based on these formulas, we discuss the influence of randomness in the tree and the basic reproduction number. We find the well known results for the homogeneously mixing case as a limit of the present model (tree-shaped contact graph). Furthermore, we develop approximate mean field equations for the dynamics on trees, and – using the message passing method – also for the configuration model. The interpretation and implications of the results are discussed.

Further exploration of using this tree model to estimate parameters in infectious models is also considered. We adopt a maximum-likelihood estimator based on a multi-type birth-death branching model for a heterogeneous population where each host is characterized by its transmission group. We extend this model to a tree-based SIR mode where an individual is characterized by its degree distribution. We start with a single individual clade as an initial point about a subepidemic started by an individual is detected. Then we proceed from considering one individual to a broader view of multiple individuals. We are able to formulate theories of maximum likelihood estimators for parameter estimation. Then, we combine the results of maximum-likelihood estimators of the basic model with the results for contact tracing. The tracing events incorporate information about the underlying contact structure, while the phylogenetic methods are better to estimate the timing of the infectious process. The combined method will improve the estimation of parameters of epidemic processes, as well as the underlying contact structure.

Acknowledgement

First, I want to thank God for seeing me through this journey. I must express my sincere gratitude to my supervisor, Johannes Müller for his constant encouragement and guidance during my research. His contant follow-up on my work progress was highly instrumental in this journey. It was really a great honour for me to learn under his tutelage. I would also like to acknowledge and thank my advisor Carl-Friedrich Kreiner for his support and follow up. I thank Aurélien Tellier and Sona John for their collaborative research support and advice. I extend my heartfeit gratitude to my sponsors, the German Academic Exchange Service (DAAD) for awarding me this prestigious grant to be able to carry out my doctoral studies in Germany.

I would like to thank the TopMath Graduate School, it was really a great experience studying in the elite program within the Elite Network of Bavaria. I specifically want to appreciate Lydia Weber, Isabella Wiegand and Katja krösss for their kind and friendly advice on administrative matters. A big shout out to all my families and friends for your true love and care and for standing by me at all times. Special thanks to my siblings Andrew, Victor, Susan, Linda and to my fiancée Blessing. To those I had special and funny moments with, Victor, Chioma, Gerald, Ugo, CJ, Courad, Stephen, Victor, Faith, Fatimata and to mention a few, thank you all. To my wonderful doctoral colleagues Bright, Adebimpe, Vella, Beryl, Usman in this great citadel of learning, you have been wonderful. A special aknowledgement to Eva Stadler, Chris and Faithful, thank you for being a wonderful friends.

A special thanks goes to my late parents, Mr. and Mrs. Okolie for their support in my education which have brought me this far. This aknowledgement cannot be complete without thanking the German government for creating a conducive and facilitating environment for learning and conducting research.

Contents

Chapter 1

Introduction

Infectious diseases burden several communities and have been one of the major causes of morbidity and mortality all over the world [67]. Any population can suffer from an infectious disease if an infectious individual is present within the population and can spread the disease to other people. The most common form of disease transmission includes physical contact (direct or indirect), airborne, droplets, etc. Once the incidence or prevalence of infection starts to increase, several control measures or techniques are sort to control the outbreak. Depending on some factors, e.g the disease of interest, contact structure, etc, control measures can be easy to implement or not.

In epidemiology, the main interest is to investigate the quantitative and qualitative behaviour of an epidemic over a long period of time. These behaviours arise from two conceptions: to have a better understanding of morbidity and mortality patterns, and use these patterns to find solutions [72]. The development of control strategies, e.g contact tracing, vaccination, screening and antibiotics have been instrumental in reducing and eradicating infection, however, with the recent increase in emerging and re-emerging diseases, most control measures require additional or new strategies.

In this thesis, we will use mathematical tools to describe and provide meaningful insights into infectious disease dynamics and also the analysis of contact tracing as a control mechanism on networks.

1.1 Background

Infectious diseases are diseases that can be spread from one individual (carrier of the disease) to another individual (non-carrier of the disease but can be infected). Mathematical epidemiology has a long history and the early mathematical models of infectious disease dynamics were based on homogenous interaction without a contact structure [69, 70, 42]. However, in reality, most human interactions are not homogeneous as families, friends, colleagues, etc. tend to form a network of fixed contacts to whom they can infect [39].

This contact network usually represents individuals as nodes and the possible interactions for disease transmission as edges [21]. The study of diseases on networks has gained much attention in the last few decades and has helped to improve our understanding of epidemiological processes from individual-level interaction to population-level dynamics [39]. This individual-level mechanism allows for incorporation of high degree of complexity and heterogeneity. Different forms of real world and simulated networks have been studied. Real-world network data requires complete information about every individual and their relationships in a population, some of which are difficult to get in practice due to time consumption, lack of interest or incomplete information.

A network from an epidemiological approach should be rather subtle if we only consider individuals and interactions that play a major role in an epidemic process. Then we can focus on a sub-network of key players in the infection process. Networks can vary depending on the disease of interest. E.g a network of HIV transmission would be different from an airborne disease such as COVID-19 due to differences in the intensity of how both diseases spread and their mode of transmission. Also, since a mixing network is an abstract of different individual behavioural patterns, disease transmission is strongly influenced by a contact network. Newman [64] in his publication investigated the effect of network topology on the rate and pattern of disease spreads. Networks can vary from simple to complex structures and several studies also analyse different spatial properties of networks [22, 39]. E.g in Poisson networks (Fig 1.1), there are no spatial properties and connections between nodes are formed with independent probability.

The lattice network (Figs 1.1, 1.2) is a type of regular network where each node connects to all of its neighbouring nodes. A common example is a square lattice which is based on a spacially structured grid where individuals are positioned at different points on the grid forming different sites. Individual interactions within the lattice network are only possible locally due to highly cluster of localized contacts. Therefore, there tends to be a longer path length between two distant nodes which also slows down infection. Another well-known network is the "Small-world" (Fig 1.1) which is based on adding a few random connections to a lattice network. Shorter path lengths are formed between two distant nodes from these random connections and thus, infection tends to spread faster in the network. A lattice model is a form of a non-random regular network that generates random networks when different level of randomness is added to it. Other types of networks include spatial networks, scale-free networks, tree-like networks (Fig. 1.2), etc.

The study of epidemics on different types of networks (known as network epidemiology) is growing rapidly where disease dynamics is modelled as a stochastic process on a contact network [15, 78, 64]. Every node with individual attributes or behaviours can vary with time. Generally, network epidemiology has helped to improve our understanding of contact pattern formation and its influence on epidemic processes [48]. Several mathematical models and techniques have been put forward. E.g, the pair approximation consisting of ordinary differential equations (ODEs) [40, 27, 38, 23] considers all nodes and edges in all possible states and extends the mean field model to incorporate pair correlations.

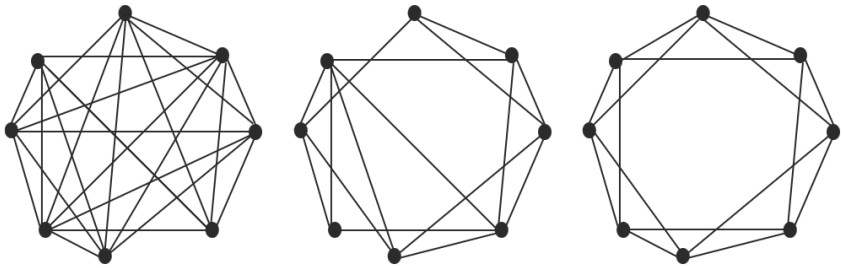


Figure 1.1: From right to left: Random network, small world network, circular (ring) lattice.

The message passing approach by Karrer and Newman [37] considers a probabilistic effects of all possible ways each individual can become infected. This message passing formulation is a generalized form of the susceptible-infected-recovered (SIR) model on both tree-like and non tree-like networks that allows for arbitrary distribution of infection and recovery times, and extends the pair approximation. Wilkinson et al. [79] also showed that the stochastic version of the homogeneous message passing model reduces to the deterministic Kermack McKendrick equations. Also, when the contact and recovery times is assumed to follow a Poisson distribution, the pair approximation can be derived from the message passing model.

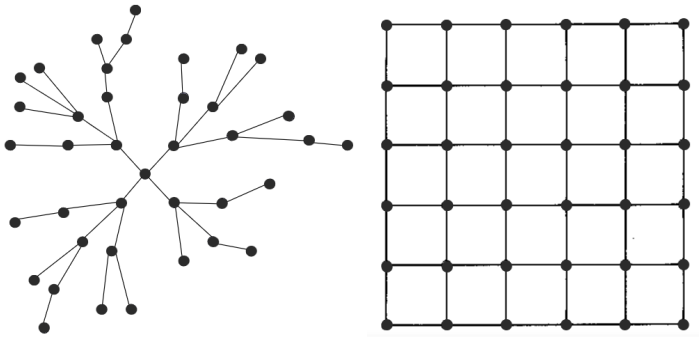


Figure 1.2: left: The tree-like network is a special type of random network with no cycles. Right: Square regular lattice in form of a grid.

The Edge-based compartmental model (EBCM) by Miller et al. [54] is an individual-based model that estimates the probability for a randomly chosen node to be susceptible. The EBCM is modelled as a stochastic process and gives the proportion of susceptible individuals in the population at a particular time. Sherborne et al. [71] extended the EBCM and showed that, for independent contact and recovery process, the extended version is equivalent to the message-passing models. Müller et al. [62] also proposed a stochastic SIRS model for contact tracing at the onset of an epidemic in a homogenous population. Just like the EBCM, the proposed model investigates the probability for a randomly chosen node to be infectious at a given age of infection.

All in all, these models have helped our understanding of epidemics on networks. While some of the models have exact solutions, numerical solutions have been proposed for some.

1.2 Contact tracing

Contact tracing is a scheme that helps to identify potential contacts of diagnosed infectives and if properly implemented, it attempts to reduce or eradicate infection [35]. Contact tracing aims to identify risky interactions within a population and it relies heavily on information about a person's contact network [47, 24]. Due to similarities between contact tracing and infection transmission patterns, there is a long history of using networks to analyse contact tracing [80, 30, 24, 47]. From an ecological perspective, contact tracing is dubbed "hyperparasitism" because it stops infection chains by spreading through the network in the same way as the disease but targets only infected individuals [55, 75].

For example, a gym instructor who showed symptoms of COVID-19 is diagnosed with the disease. The gym instructor called an "index case" is immediately interrogated and asked to name all his contacts. Realistically, he barely remembers everywhere he has been but can only remember a few places. First, he thinks about his immediate family members, his colleagues at his workplace, the gym and the restaurant he eats often. He mentions all possible contacts in these places and they are contacted by a health worker. Contacts are then advised to test themselves for possible infection. If the test is negative, the contact is confirmed not to have contracted the disease from this gym instructor. A positive result confirms an infection and the contact tracing process is repeated on this contact. In the case where there are no curative treatments, identified persons may be quarantined or isolated to prevent further infection. Also, if at the time of detection, requires an expensive diagnostic test or no quick and efficient test measure, then all contacts may be isolated to avoid further contact with individuals involved in this chain of transmission. The health workers monitor this scheme to ensure that all contacts with the disease are identified.

Most control measures, such as mass screening or mass vaccination, involve testing and treating random samples of individuals in a population [8, 27]. In contrast, contact tracing in almost the same way involves screening only subpopulations (contacts of index cases) with potentially high incidence. Also contact tracing involves identifying core groups in a population of interest, e.g in the case of gonorrhoea, the group of sexually active individuals are targetted. Several studies have empirically investigated factors improving contact tracing [47, 56, 18, 26, 13] yet our understanding of contact tracing effectiveness is still unclear based on different strategies to successfully implement.

Model analysis showed that contact tracing reduces disease prevalence by breaking chains associated with routes of transmission [24], a point that we will discuss in this thesis from a different perspective. Several intensive simulation studies have investigated the influence of contact graphs on contact tracing [44, 45]. The analysis of models for contact tracing is more involving and mathematically challenging than models e.g. targeting random

screening, as it is necessary to address local correlations. Several approaches are present in the literature. One major approach that focuses on correlations is pair approximation [38, 36, 23]. Originally, pair approximation has been developed to better analyse the effect of correlations in stochastic processes on graphs, e.g., infected individuals tend to cluster, and therefore the spread of an epidemic slows down. As contact tracing targets links between infected individuals, pair approximation is useful also in that context. Frazer et al. [29] did focus on the age of infection and analysed an (deterministic) age-structured model.

In a homogeneously mixing population, an approach often used is the analysis of a branching process with dependencies [62, 47, 58, 60, 5, 6]. The advantage of this method is an exact analysis of the corresponding models, at least during the onset of an epidemic. It is possible to handle an SIS and SIR model, as long as the onset of the infection is considered. Up to now, this method was restricted to a complete contact graph and a large (infinite) population. The analysis of epidemics on more realistic graphs is more challenging [7]. However, the underlying idea of the branching process approach to contact tracing resembles that of the message passing method for a stochastic SIR dynamics on a tree [37, 78, 46].

Starting from this observation, it became clear that it is possible to extend the branching theory for contact tracing to a stochastic SIR model on graphs which laid the foundation of this thesis.

1.3 Scope

We present in this thesis, contact tracing on stochastic graphs. We use a stochastic SIR model on a tree for the analysis of contact tracing from an individual level to population-level dynamics and then estimated epidemiology parameters in infectious models. Some parts of this thesis have been published while some are in still preparation for submission. The thesis is structured as follows:

In chapter 1, we introduce and give a background of infectious diseases and notable literature on the modelling of disease outbreaks. We start from the fundamentals of the study of epidemics on networks. We discuss a few types of networks and give an intuitive explanation of some of their underlying structure. Several modelling approach literature is also discussed, in particular how some of these models are just special cases of others. We introduce contact tracing in the last part of this chapter together with some contact tracing modelling techniques. The motivation and scope of this thesis is highlighted which we present in the later chapters.

In chapter 2, we start from the basic model assumptions and introduce the mathematical framework behind compartmental modelling. We estimate the reproduction number of this basic model and visualize the trajectories of the solution using a phase plane. Then we introduce the stochastic model for epidemics on graphs. Thereto, we introduce the basic graph theory definitions and notations and then we show how random graph models gives a

good representation of contact structure in the modelling of disease spread. We formulate the basic Erdős Rényi model and also the theorem about giant components. Furthermore, we gradually show techniques of modelling contact tracing by a simple contact process which is defined by a contact graph. We derive the mean-field approximation from this contact process by breaking correlations, then we also derive the pair approximation model as a special case of this contact process that incorporates correlations. This approximation is particularly interesting to better analyse the effect of correlations in stochastic processes on graphs and in the theory of contact tracing.

In chapter 3, we present the model for contact tracing on a random tree. The central ingredient for this analysis is the probability for a randomly chosen individual to still be infectious at a given age of infection. We obtain integrodifferential equations for the probability to be infectious at a given age of infection. Central information, such as the reproduction number, can be analysed, and a suited limit showed that it was possible to obtain results resembling that of a homogeneous population (complete graph) in Müller et al. [62], Müller and Koopmann [60] under suited limit conditions. We also present the influence of the graph structure on the probability to be infectious. By considering different degree distributions, we show the effect of the variance on the reproduction number and probability to be infectious. The theory of superspreaders is also highlighted and also show how the variance increases the speed of convergence for this probability. From this micro-level analysis, we derive through heuristic analysis, a mean-field equation on trees. Based on even more heuristic arguments, we apply the theory to a non-tree contact graph: the configuration model.

In chapter 4, we present a phylogenetic method for infectious models. The idea is based on estimating parameters from inferred phylogenies. On a root phylogenetic tree, we calculate probabilities from the tip along the branches down to the root for a complete estimate of the whole tree. We start by adopting a maximum-likelihood estimator which is based on a multi-type branching process for a heterogeneous population where each host is characterized by transmission groups. We extended this multi-type birth-death branching model to a tree-based SIR model which also we later incorporate contact tracing. Each individual is characterized by its degree distribution and we are able to derive likelihood estimators for estimating parameters. The theory of contact tracing introduced in this model is able to allow us to estimate the degree distribution of further detected cases by contact tracing and also estimate tracing probability as well as the underlying contact structure.

In chapter 5, we discuss and conclude the findings of this thesis.

Chapter 2

Modelling of outbreaks and contact tracing

2.1 Basic model assumptions and notation

Some of the pioneering works on modelling epidemics using mathematical equations are dated back to the early 20th century [69, 70, 42]. These basic models consisting of a system of ordinary differential equations are able to predict the dynamics of disease spread in a closed population. Over the decades, several models have evolved and are still evolving from it. The general idea is to divide a population of size N in consideration into 3 different compartments. The compartmental model group every individual in the population into 3 different disease state groups SIR which represents the total number of susceptible, infected and recovered individuals in the population at time t respectively. The susceptible individuals are a fraction of the population that are healthy but exposed to infection. For transmission to occur, there must be an infectious contact between a susceptible and an infected individual. After contraction of a disease, this susceptible individual becomes infected and transitions to the infected compartment. Upon recovery, an individual gains long lasting immunity. This SIR model is particularly relevant for a human to human transmission and assumes that a recovered individual becomes immune forever and can not be reinfected.

There are other forms of compartmental models, e.g, the SI compartmental model which simply assumes that an infected individual stays infected forever. Also, the SIS model assumes a recovered individual immediately transitions back into the susceptible class and can be reinfected. A similar form of the SIS is the SIRS model where there is a short time resistance for a recovered individual before he or she becomes susceptible again. A summary of various forms of compartmental models with description is given in the table 2.1.

Generally, these compartmental models based on differential equations described help to

Model	Meaning
SI	Susceptible – Infected
SIS	Susceptible – Infected – Susceptible
SIR	Susceptible – Infected – Recovered
SIRS	Susceptible – Infected – Recovered – Susceptible
SEIS	Susceptible – Exposed – Infected – Susceptible
SEIR	Susceptible – Exposed – Infected – Recovered
SIRD	Susceptible – Infected – Recovered – Deceased/Death
SIRV	Susceptible – Infected – Recovered – Vaccinated
MSIR	Maternally-derived Immunity – Susceptible – Infected – Recovered
MSEIR	Maternally-derived Immunity – Susceptible – Exposed – Infected – Recovered

Table 2.1: A summary of common compartmental models

Figure 2.1: Dynamics of the SIR model showing the transitions into the different states.

quantify and describe the evolution of the different compartments with respect to time. These models are called deterministic models and have been very useful in explaining spatial dynamics with respect to large populations. These models may fail to capture randomness and applied problems in epidemic events e.g local interactions, behaviour patterns among individuals and in particular, consequence of small population size.

The stochastic epidemic models which are more realistic are an adaptation of the deterministic models but formulated as Markov process [3]. These models are able to incorporate a high degree of complexity and heterogeneity among interacting individuals which gives better insights into individual-based modelling. Before introducing some of the stochastic approaches for modelling outbreaks, we will briefly explain the deterministic theory.

2.2 Deterministic SIR models

Given a total host population size $N(t) = S(t) + I(t) + R(t)$, the deterministic SIR model [42] is given by the following system of ordinary differential equations

$$\begin{aligned} S'(t) &= -\beta \frac{S(t)I(t)}{N} \\ I'(t) &= \beta \frac{S(t)I(t)}{N} - \gamma I \\ R'(t) &= \gamma I. \end{aligned} \tag{2.2.1}$$

Where β is the transmission rate constant and γ is the recovery rate (so the mean time spent in the infectious class is $1/\gamma$). The standard SIR model is basically a special simple case of the Kermack-McKendrick model which structures the infected population to include age since infection.

The only difference between the SIS and the SIR models is that, in the SIS model, recovered individuals transitions back to the group of susceptible. This results in the following differential equations

$$\begin{aligned} S'(t) &= -\beta \frac{S(t)I(t)}{N} + \gamma I \\ I'(t) &= \beta \frac{S(t)I(t)}{N} - \gamma I. \end{aligned} \tag{2.2.2}$$

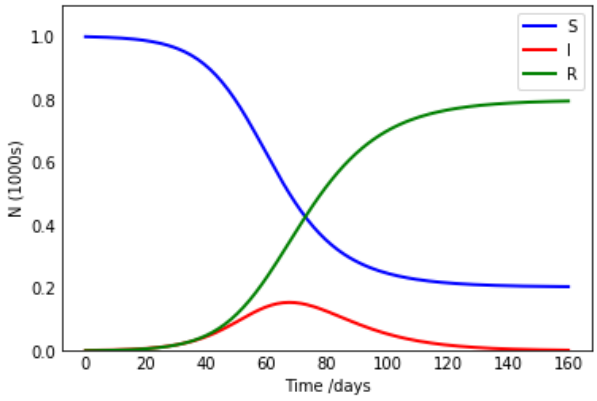


Figure 2.2: Plot showing the time rate of change in each of the SIR compartments. Parameters: $\beta = 0.2, \gamma = 0.1, N = 1000, S(0) = 999, I(0) = 1, R(0) = 0$.

Fig. 2.2 shows the behaviour of the model.

2.2.1 Basic reproduction number

Mathematical analysis can be used to analyse the quantitative and qualitative behaviour of an epidemic process. E.g to understand a certain threshold for a major outbreak or not.

If the infected compartment increases $I' > 0$, there will be a major disease outbreak. From Eqn. 2.2.1

$$I' > 0 \Longleftrightarrow \left(\beta \frac{S}{N} - \gamma\right) I > 0$$

At the onset of an outbreak, nearly every individual is susceptible so that $S \approx N$. We define

$$\mathcal{R}_o := \frac{\beta}{\gamma} > 1 \tag{2.2.3}$$

to be the basic reproduction number of the disease which is the average number of infectious contacts made by an average infectious individual in a completely susceptible population [4]. Now, we just define the reproduction number but later (in section Random graphs and the final size of epidemics), we will explain in details why this definition has this interpretation.

The reproduction number $\mathcal{R}_o$ plays an important role when modelling the dynamics of disease spread. It serves as a threshold value for an epidemic to occur or not. If $R_o \leq 1$, then the number of infecteds continually decreases to zero and dies out (disease-free equilibrium). If $R_o > 1$, the infecteds increase and there is a major outbreak (endemic equilibrium).

2.2.2 The phase plane (SI -plane)

The phase phane is a way to visualize the trajectories of the SIR model. The $R(t)$ compartment does not contribute to the $S(t)$ and $I(t)$ equations and thus can be disregarded when studying the dynamics of the SIR model. We can derive analytic result by dividing equation $\dot{I}$ by $\dot{S}$,

$$\frac{dI}{dS} = \frac{\beta SI/N - \gamma I}{-\beta SI/N}$$

Integrating both sides yields

$$I(t) = \frac{\gamma N}{\beta} \ln S(t) - S(t) - \frac{\gamma N}{\beta} \ln S(0) + S(0) + I(0)$$

Hence,

$$I(t) = I(0) + S(0) - S(t) + \rho N \ln\left(\frac{S(t)}{S(0)}\right), \tag{2.2.4}$$

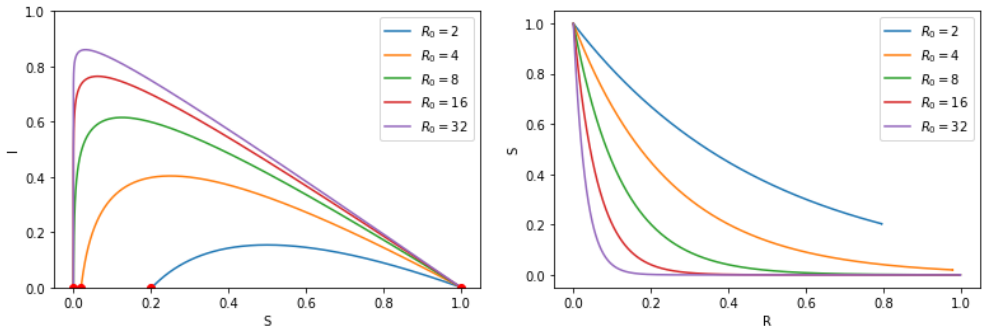


Figure 2.3: The SI and RS phase plane trajectory for the SIR model.

where $\rho = \frac{\gamma}{\beta} = \frac{1}{\mathcal{R}_o}$.

2.3 Epidemics on graphs

The nature of a contact network is of relevance for an outbreak to occur or not. E.g for a sparse network with many disconnected components, contacts are unlikely and disease is likely to die out. As soon as we have a contact network with pre-defined nodes and interacting links, we have a contact graph. We can use graph-based models to determine the distribution of an outbreak and epidemic sizes.

This section aims to understand how contact graphs are mathematically constructed and how it relates to modelling contact tracing.

2.3.1 Graph theory fundamentals

Graphs are mathematical structures that model the interactions between objects. To understand how graphs are used to model epidemics on networks, we simply introduce some of the most basic graph definitions and notations relevant to this thesis.

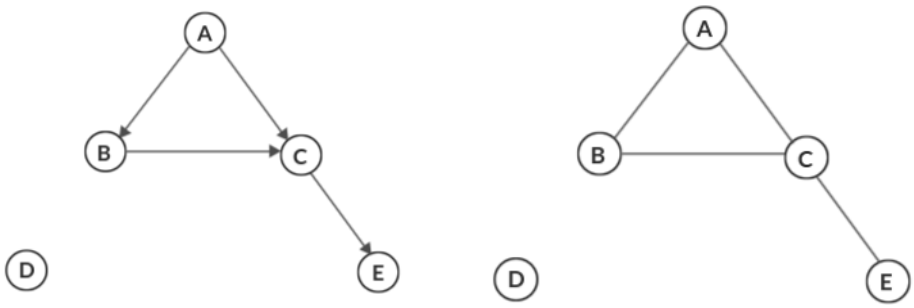


Figure 2.4: Left: Directed graph Right: Undirected graph.

Definition 2.3.1. *A directed graph is an ordered pair $G = (V, E)$ where V is a finite, non-empty set of objects called nodes, and E is a (possibly empty) 2-tuple of V called edges. If $e = (u, v) \in E(G)$, we say that nodes u and v are adjacent in G, and that e connects u and v. The edge e is said to be incident with u and v.*

Example 2.3.1. *In Fig 2.4, the pairs are ordered, so the graph G is directed with:*

$$V = \{A, B, C, D, E\}$$
$$E = \{(A, B), (A, C), (B, C), (C, E)\}$$

Remark 2.3.1. *If the graph is undirected, then the pair is unordered such that both directions are edges and we say edge e is incident with u and v, and vice versa (Fig 2.4):*

$$V = \{A, B, C, D, E\}$$
$$E = \{(A, B), (B, A), (A, C), (C, A), (B, C), (C, B), (C, E), (E, C)\}.$$

Example 2.3.2. *In epidemic context, a directed graph means I can only infect you with a disease but not vice-versa while an undirected graph case means infection can happen in both direction.*

Definition 2.3.2. *A node $u \in V$ is called a direct neighbour (an adjascent node) of node $w \in V$ if $(u, w) \in E$. The number of edges that a node contain is called the degree of that node.*

Example 2.3.3. *In Fig. 2.4, the nodes A and B have degrees 2, C is of degree 3, E is of degree 1 and D is of degree 0.*

Definition 2.3.3. *A graph $H(W, F)$ is a subgraph of $G(V, E)$ if $W \subset V$ and $F \subset E$.*

Definition 2.3.4. *A path $\rho(G)$ of graph G is an ordered sequence of vertices such that 2 subsequent vertices is always contained in the set of edges. If $\rho(G) = (v_1, v_2, .., v_{k-1}, v_k)$, then $\rho(G)$ is called a path in G if $(v_i, v_{i+1}) \in F$ for $i = 1, ..., k - 1$. The number of edges of a path is called the length of the path.*

Definition 2.3.5. *A path $\rho(G)$ in graph G is called a cycle if it starts from a given vertex u and ends at the same vertex u.*

Definition 2.3.6. *A graph $P(V, E)$ is said to be a complete graph if every 2-tupled distincts vertices $(u, v) \in V \times V$ forms an edge $(u, v) \in E$.*

Definition 2.3.7. *A graph G is called connected if for any pair of nodes (u, v), there exist a path in G starting from u and ending in v; otherwise G is disconnected.*

Definition 2.3.8. *A tree is a connected graph that contains no cycles.*

Event	Transition	Rates	Probabilities (up to higher order terms)
Infection	$S \to S-1, I \to I+1$	βSI	$\beta SI\Delta t$
Recovery	$I \to I-1, R \to R+1$	γI	$\gamma I\Delta t$

Table 2.2: Possible events in a standard stochastic SIR model, rates and probabilities of occurrence at a time interval. β and γ denotes the (constant) rate of infection and removal respectively.

2.3.2 Stochastic epidemic models

In previous sections, we introduced the basic compartmental model that classifies a population into 3 groups namely susceptible, infected and recovered. Some basic mathematical exploration of deterministic models shows the long term behaviour of the system and the conditions for a disease to persist or die out as well as the final epidemic size. These deterministic model results are well suited in particular for a very large well mixed population.

On the other hand, stochastic models have been very useful in addressing more complex situations. E.g high degree of complexity and heterogeneity where individual roles may differ according to behavioural pattern, spatial location, contact size, etc. Here, we present a stochastic version of the deterministic SIR model (eqn. 2.2.1) and explore the long term behaviour of the model. Some of the stochastic modelling frameworks we will discuss include discrete time Markov Chain model (DTMC), continuous time Markov Chain (CTMC) model, and stochastic differential equation (SDE) models. These modelling approaches differ based on the underlying time and state assumptions.

2.3.2.1 Transmission probabilities of state

Now, we turn to several individuals in a birth-death process where birth and death is possible. For an SIR model, we define the state of a random variable $X_t \in \{S, I, R\}$ the possible events of infection and recovery. In an infection event, the susceptibles class decreases by 1 and the infection class increases by 1. In a recovery event, the infected class decreases by 1 while the recovered class increase by 1. As we take a very small time step, we don't expect more than 1 event to occur at a time, the probability for two or more events to happen is of order $\mathcal{O}(\Delta t^2)$.

According to the standard stochastic SIR model, we write the transition probabilities (see Table. 2.2 for summary)

$$\begin{aligned} P(S_{t+\Delta t}, I_{t+\Delta t}) &= (S(t)-1, I(t)+1 | S_t = S(t), I_t = I(t)) \\ P(I_{t+\Delta t}, R_{t+\Delta t}) &= (I(t)-1, R(t)+1 | I_t = I(t), R_t = R(t)) \,. \end{aligned}$$

2.3.2.2 Random graphs and the final size of epidemics

The compartmental model assumes contacts are randomly mixed but ignores the contact structure. Random graphs gives a good representation of this contact structure with the same underlying assumptions. Here, we will focus on constructing the epidemic process as a stochastic process on random graph models from Müller and Kuttler [61]

In real-world applications, random graphs are used to describe huge networks not known in detail. E.g the internet, citation network, the network of disease or rumour spreading, etc. We know these networks are large and have typical properties but they are not known in detail. We can construct models defined on some similar principles to gain better insight into these network properties and overall structure and functioning. The study of random graphs was first introduced in the seminar paper of Erdös and Rényi [25]. Here, we will only sketch the idea of the model. A detailed proof of this model can found in the book of Durrett [19]. The model is defined as follows:

Model:

We start with an empty graph with N nodes, there will be at most $\binom{N}{2}$ possible edges. A randomly chosen node has at most $N-1$ edges, if we perform $\binom{N}{2}$ Bernoulli experiments inserting edges independently with probability p, then the random variable K that counts the number of realised edges is follows a binomial distribution.

$$K \sim Binom(N-1, p).$$

This Binomial random graph was introduced by Gilbert [31]. Usually, one is interested in large graphs. To describe these networks, if we do not scale p and allow $N \to \infty$ and $p \to 0$ while $\mathbb{E}[K]$ stays constant, the Binomial distribution is well approximated by the Poisson distribution with expectation $\mathbb{E}[K]$. This is why the Erdős Rényi (random) graphs are also referred to as the Possion random graphs.

Definition 2.3.9. *(Probability generating function) The probability generating function of a random variable K is*

$$G : [0,1] \to [0,1], \quad s \to \sum_{i=0}^{\infty} s^i \mathbb{P}(K=i) = \mathbb{E}\left[s^K\right],$$

where K is an $\mathbb{N}_0$-valued random variable and $\mathbb{P}$ denotes the probability mass function of K.

Definition 2.3.10. *(Properties of probability generating function) Let G be a probability generating function of a random variable K. Then the following properties hold:*

1. $G(i) = \mathbb{P}(K=i) = \left(\frac{1}{i!}\right) G^{(i)}(0) = \left(\frac{1}{i!}\right) \frac{d^i}{ds^i}\left(G(s)\right)\Big|_{s=0}$
2. $G'(1) = \sum_{i=0}^{\infty} 1^{K-1} k\, \mathbb{P}(K=i) = \mathbb{E}[K]$

3. $G''(1) = \mathrm{Var}[K] - \mathbb{E}[K](\mathbb{E}[K] - 1)$.

Giant component: Under certain conditions, the presence of giant component is central in the theory about random graphs. Many real world networks are connected and a close look at the local level properties of these connectedness are not clear.

Theorem 2.3.1. *Consider a Poisson graph with parameter* $\mathbb{E}[K]$. *Let* S_N *denote the average size of the largest connected component for a Poisson graph with* N *nodes and parameter* $\mathbb{E}[K]$. *If* $\mathbb{E}[K] < 1$, *then* $S_N/N \to 0$ *for* $N \to \infty$, *and for* $\mathbb{E}[K] > 1$ *we have* $S_N/N \to c > 0$ *almost surely.*

This theorem tells us that if the average realised edges $\mathbb{E}[K] < 1$, the graph is split into connected components that are small relative to the size of the graph. If $\mathbb{E}[K] > 1$, a positive fraction of the population is located in one single component, and also for large and huge graphs. This component is referred to as the giant component.

Proof. Let s denote the probability for a node not to belong to the largest component. Since every node is $i.i.d$, then the probability that a randomly chosen node do not belong to the largest component is the same for all neighbours that do not belong to this component. If the degree of the node under consideration is i, this probability reads s^i. Taking into account the degree distribution and the fact that our primary selected node has been randomly chosen, we state some important definitions:

Thus, we obtain the probability generating function

$$s = \sum_{i=0}^{\infty} s^i \mathbb{P}(k = i) = \sum_{i=0}^{\infty} \frac{1}{i!} s^i \mathbb{E}[K]^i e^{-\mathbb{E}[K]} = e^{-\mathbb{E}[K]} e^{s\mathbb{E}[K]} = e^{\mathbb{E}[K](s-1)}$$

The fraction u of nodes that belongs to the largest component satisfies $u = 1 - s$, thus

$$u = 1 - e^{-\mathbb{E}[K]u}$$

If $\mathbb{E}[K] < 1$, there is only the trivial fixed point $u = 0$. If $\mathbb{E}[K] > 1$, there is a second fixed point $u \in (0, 1)$. This is, for $\mathbb{E}[K] > 1$, a considerable fraction of nodes are located in the largest component. □

This model idea only takes into account that the number of neighbouring nodes of a randomly chosen node is infinite. In a branching process, specifically, one usually compares neighbouring nodes as a mother-daughter relationship. The appropriate arguments for the branching process approximation would also be discussed in the next chapters.

Final size of an epidemic on a Poissonian network: The basic idea of using random graph models to describe an epidemic is to first construct the contact graph. Then we assign independently on each edge, a probability of transmitting the disease if one of the connected nodes is infected. We remove all the edges that do not transmit the disease, and we focus

only on the infection graph. Last, we select one node at random, the estimated size of the connected component to which the node belongs is the final size of the epidemic.

We address a simple SIS model for this epidemic process. We will use the same parameter of the deterministic model described in previous sections to connect the theory about giant component. Here, we now have to think about the contact structure. The contact rate (β) in the deterministic model is defined per person, i.e, the rate a person contacts any other person, and is able to transmit the disease if infected and the other uninfected. In the random graph model, the contact is not defined per person but per edge. The average contact per person is given by $\mathbb{E}[K]$. Assume the time points of contacts on a given edge follow a Poisson process with intensity τ, the average total contact rates (contacts via any edge) per individual reads $\beta = \tau\mathbb{E}[K]$. The infected period is assumed to be exponentially distributed with an average length of $1/\gamma$.

Assuming that one of the two persons connected by an edge becomes infected at time zero, how likely will there be a contact before the person recovers? We define two random variables for the time of the next contact T_c, and for the time the person recovers again T_r. Thus the probability

$$q := P\left(T_c < T_r\right) = \frac{\tau}{\tau + \gamma}.$$

Therefore, a fraction of the population that recovers before making a contact $(1 - q)$ is eliminated, and we obtain a new Poissonian graph (the epidemic graph, which is a subgraph of the contact graph) with the expected number of nodes per individual $Npq = \tau\mathbb{E}[K]/(\tau+\gamma)$. This is the reproduction number R_0 which gives average number of persons infected per individual within a completely uninfected population. Therefore

$$R_0 = \frac{\tau\mathbb{E}[K]}{\tau + \gamma} = \frac{\beta}{\gamma + \beta/\mathbb{E}[K]}.$$

This is, we almost recover the R_0 for the deterministic model. The extra term $\beta/\mathbb{E}[K]$ is a correction due to the finite number of edges. If $\mathbb{E}[K]$ tends to infinity, this term vanishes. We say that a major outbreak can happen if the epidemic graph has a giant component ($R_0 > 1$). We can also recover this reproduction number by focusing on the infectious age of a single individual. We find the following result. Let $e^{-\gamma a}$ denote the probability to still be infectious at age a of infection, then the average number of contacts this individual makes in an infectious period $[0, a]$ is given as

$$\begin{aligned} \mathcal{R}_o &:= \int_0^a \tau\mathbb{E}[K]e^{-\gamma a}da \\ &= \frac{\tau\mathbb{E}[K]}{\gamma}(1 - e^{-\gamma a}) = \frac{\beta}{\gamma}(1 - e^{-\gamma a}). \end{aligned}$$

We find this $\mathcal{R}_o$ is similar for the deterministic model but with an additional term $1 - e^{-\gamma a}$ which incorporates age since infection a. This term vanishes if a tends to infinity. Other form of individual-level mechanisms, e.g adaptive behavior, variation in exposure, etc can also be incorporated in individual-based models which has advantage over deterministic epidemic models [16].

2.4 Modelling contact tracing

Mathematically, we can use modelling techniques to analyse contact tracing processes on different contact networks and also, find formulas to estimate tracing probabilities. Simple susceptible-infectious-recovered (SIR) models have shown to capture basic features of disease epidemic [4, 33] and control measures, e.g screening or vaccination have also been implemented [4, 22]. These control measures have proven to be effective at population level but may fail to capture individual local interactions and correlations [62].

Contact tracing control strategy, unlike the mean-field approach, requires information about individual contact links in the transmission network. Modelling contact tracing using the mean-field approach has proven to be unrealistic due to the loss of information about connected pairs in the network. With information about these pairs, contact tracing has been shown to reduce prevalence of infection. We will look at 2 common approches for modelling contact tracing. In this section, we will derive the first approach: the pair approximation by Eames and Keeling [22] and in the next chapter, we will present the second approach: the branching process approximation which is an extension of Müller et al. [62], Müller and Koopmann [60].

2.4.1 Contact Process

Here, we introduce the contact process of a prototypical spatially structured model in interacting populations [61]. We start by introducing a model for the spread of a plant specie on a square grid [49, 20] where an empty site means susceptible and presence of a plant means infected. We now define a contact graph by this grid which looks like a spatial structure but we have no space in mind but only the contact graph (See Fig. 1.1). In the following subsections, we show in steps the contact process of a prototypical contact structured model to help gain insights on the basic properties of approximations techniques by deterministic mean field and pairwise models. We derive the mean-field approximation and the pair approximation from the basic SIS model. In between, contact tracing has been introduced as one of the events that may lead to the recovery of an infected individual.

2.4.1.1 SIS model

An individual may be susceptible (we define as "0") or infected (we define as "1"). This notation and model assumption follows the SIS framework where an infectious event switches a 0 to 1 when an individual is infected. Also, a recovery event switches a 1 to 0.

Definition 2.4.1. *Contact graph of the contact process: Let* $\Gamma = \mathbb{Z}^d$ *or* $\Gamma = \mathbb{Z}^d/\mathbb{Z}_m^d$ *be a d-dimensional torus. We define for a susceptible individual,* $x = 0 \in \Gamma$ *the neighbourhood* $U(0) = \{x \in \Gamma : \|x\|_\infty \leq 1\}$ *(Moore neighbourhood) or* $U(0) = \{x \in \Gamma : \|x\|_1 \leq 1\}$ *(Neumann neighbourhood. Let* $x + U(0)$ *be the neighbourhood of* $x \in \Gamma$*.*

Definition 2.4.2. *State of the contact process: Let* $E = \{0, 1\}$ *be the set of all possible states of a site. We define* Γ *such that the map*

$$\varphi : \Gamma \to E$$

yields the state of the system. Let $(\Omega, \mathscr{F}, P)$ *be a probability space, we define a random funtion,* $\varphi_t^{(\omega)}(x)$ *that yields the state of the site* x *at a time* t *in a certain realisation* $\omega \in \Omega$*.*

Definition 2.4.3. *Dyanmics of the contact process: Let*

$$\varphi : \Omega \times \mathbb{R}_+ \times \Gamma \to E, \qquad \varphi(\omega, t, x) = \varphi_t^{(\omega)}(x)$$

a random variable that evolves according to $(g > 0)$

$$P\left(\varphi_{t+\Delta t}^{(\cdot)}(x) = 0 | \varphi_t^{(\cdot)}(x) = 1\right) = g\,\Delta t + \mathcal{O}(\Delta t)$$

and $(\beta > 0)$

$$P\left(\varphi_{t+\Delta t}^{(\cdot)}(x) = 1 | \varphi_t^{(\cdot)}(x) = 0\right) = \tfrac{\beta}{K} \# \left\{y \in U(x) | \varphi_t^{(\cdot)}(x) = 1\right\} \Delta t + \mathcal{O}(\Delta t)$$

where $K = |U(0)| - 1$*.*

2.4.1.2 Counting configurations

In the following, we need some definitions.

Definition 2.4.4. *Let* $G(v, v')$ *be an incidence matrix for the graph implied by* $(\Gamma, U(0))$*, i.e*

$$G(v, v') = \begin{cases} 1, & \text{for } v' \in U(v) \backslash \{v\} \\ 0, & \text{otherwise,} \end{cases} \tag{2.4.1}$$

and $E = \{0, 1\}$*. Then, we define for one realisation* $\omega \in \Omega$ *and* $A, B, C \in E$

$$
\begin{aligned}
[A](t,\omega) &= \sum_{v\in\Gamma} \chi^{(\omega)}_{\varphi_t(v),A} \\
[AB](t,\omega) &= \sum_{v_1,v_2\in\Gamma} \chi^{(\omega)}_{\varphi_t(v_1),A}\chi^{(\omega)}_{\varphi_t(v_2),B}G(v_1,v_2) \\
[ABC](t,\omega) &= \sum_{\substack{v_1,v_2,v_3\in\Gamma \\ v_1\neq v_2\neq v_3\neq v_1}} \chi^{(\omega)}_{\varphi_t(v_1),A}\chi^{(\omega)}_{\varphi_t(v_2),B}\chi^{(\omega)}_{\varphi_t(v_3),C}G(v_1,v_2)G(v_2,v_3)
\end{aligned}
\tag{2.4.2}
$$

and the expectations

$$
\begin{aligned}
|[A]|(t) &= E([A](t,\cdot)) \\
|[AB]|(t) &= E([A,B](t,\cdot)) \\
|[ABC]|(t) &= E([A,B,C](t,\cdot))
\end{aligned}
\tag{2.4.3}
$$

We note that due to symmetry, $[A,B](t,\omega) = [B,A](t,\omega)$ and $[A,B,C](t,\omega) = [C,B,A](t,\omega)$. We exclude edges connecting a node to itself on the graph and additionally note that a node in Γ cannot be its own neighbour. This implies that $G(v,v) = 0$. We also note that $G(v,v')$ in eqn. 2.4.1 is based on the chosen neighbourhood $U(0)$ not on the graph structure. The number of configurations with symmetry is counted twice since the orientation of contacts between partners is not taken into account. On the contrary, the number of configurations without symmetry is counted once. For example, in Fig. 3.2, we find for singletons

$$
\begin{aligned}
[0] &= |\{v_1, v_2\}| = 2 \\
[1] &= |\{v_3, v_4\}| = 2
\end{aligned}
$$

and for pairs denoted as the ordered pair of the nodes v_i and v_j

$$
\begin{aligned}
[00] &= |\{(v_1,v_2),(v_2,v_1)\}| = 2 \\
[01] &= |\{(v_1,v_3),(v_2,v_3)\}| = 2 \\
[11] &= |\{(v_3,v_4),(v_4,v_3)\}| = 2
\end{aligned}
$$

Definition 2.4.5. *In one dimension, for $d = 1$, i.e $\Gamma = \mathbb{Z}^1$, we find for any state for*

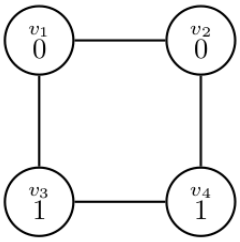


Figure 2.5: Examples for counting of configurations

$$\begin{aligned}
[0] &= \tfrac{1}{2}\left([0,0]+[0,1]\right) = \tfrac{1}{2}\,[0,0,0]+[0,0,1]+\tfrac{1}{2}\,[1,0,1] \\
[1] &= \tfrac{1}{2}\left([0,1]+[1,1]\right) = \tfrac{1}{2}\,[0,1,0]+[0,1,1]+\tfrac{1}{2}\,[1,1,1] \\
[0,0] &= [0,0,0]+[1,0,0] \\
[0,1] &= [0,0,1]+[1,0,1] \\
[1,1] &= [0,1,1]+[1,1,1]\,.
\end{aligned} \tag{2.4.4}$$

Proof. In this proof, the notation $\varphi(v)$ is used in place of $\varphi_t^{(\omega)}(x)$.

step 1: Sum over G. For any $v_1 \in \Gamma$, we find

$$\sum_{v_2 \in \Gamma} G(v_1, v_2) = K = 2$$

since the sum takes into the account the numbers of neighbours of v_1. Similarly

$$\sum_{v_1 \in \Gamma} G(v_1, v_2) = K = 2$$

Furthermore, if $v_3 \in U(v_1)\backslash\{v_1\}$, we have

$$\sum_{\substack{v_2 \in \Gamma \\ v_2 \neq v_3}} G(v_1,v_2) = K - 1 = 1$$

since we again count the size of neighbours of v_1, but exclude one specific neighbour (i.e v_3).

Step 2: From pairs to singletons.

Let $A \in E$, then

$$\begin{aligned}[0,A]+[1,A] &= \sum_{v_1,v_2\in\Gamma}\left(\chi_{\varphi(v_1),0}+\chi_{\varphi(v_1),1}\right)\chi_{\varphi(v_2),A}G(v_1,v_2)\\ &= \sum_{v_2\in\Gamma}\chi_{\varphi(v_2),A}\sum_{v_1\in\Gamma}G(v_1,v_2)=\sum_{v_2\in\Gamma}\chi_{\varphi(v_2),A}K\\ &= 2\,[A]\end{aligned}$$

Step 3: From triplets to pairs (recall that $K=2$ for $\Gamma=\mathbb{Z}^1$)

For every pair $[A,B]\,, A, B\in E$, we obtain

$$\begin{aligned}[0,A,B]+[1,A,B] &= \sum_{\substack{v_1,v_2,v_3\in\Gamma\\ v_1\neq v_2\neq v_3\neq v_1}}\left(\chi_{\varphi(v_1),0}+\chi_{\varphi(v_1),1}\right)\chi_{\varphi(v_2),A}\chi_{\varphi(v_3),B}G(v_1,v_2)G(v_2,v_3)\\ &= \sum_{\substack{v_1,v_2,v_3\in\Gamma\\ v_1\neq v_2\neq v_3\neq v_1}}\chi_{\varphi(v_2),A}\chi_{\varphi(v_3),B}G(v_1,v_2)G(v_2,v_3)\\ &= \sum_{\substack{v_2,v_3\in\Gamma\\ v_1\neq v_2\neq v_3\neq v_1}}\chi_{\varphi(v_2),A}\chi_{\varphi(v_3),B}G(v_2,v_3)\sum_{v_1\in\Gamma\setminus\{v_3\}}G(v_1,v_2)\\ &= \sum_{\substack{v_2,v_3\in\Gamma\\ v_2\neq v_3}}\chi_{\varphi(v_2),A}\chi_{\varphi(v_3),B}G(v_2,v_3)(K-1)\\ &= [A,B]\end{aligned}$$

Step 4: From triplets to singletons.

$$\begin{aligned}&\tfrac{1}{2}\,[0,A,0]+[0,A,1]+\tfrac{1}{2}\,[1,A,1]\\ &=\tfrac{1}{2}\left(\{[0,A,0]+[1,A,0]\}\right)+\tfrac{1}{2}\left(\{[0,A,1]+[1,A,1]\}\right)\\ &=\tfrac{1}{2}\left([A,0]+[A,0]\right)=\tfrac{1}{2}\left(2\,[A]\right)\\ &=[A]\end{aligned}$$

□

Corollary 2.4.1. *The expectations of the number of pairs and singletons in a certain configuration satisfy*

$$
\begin{aligned}
\frac{d}{dt}|[0]| &= -\tfrac{\beta}{K}|[01]| + g|[1]| \\
\frac{d}{dt}|[1]| &= \tfrac{\beta}{K}|[01]| - g|[1]| \\
\frac{d}{dt}|[00]| &= -\tfrac{2\beta}{K}|[001]| + 2g|[01]| \\
\frac{d}{dt}|[01]| &= \tfrac{\beta}{K}(|[001]| - |[101]| - |[01]|) - g(|[01]| - |[11]|) \\
\frac{d}{dt}|[11]| &= \tfrac{2\beta}{K}(|[01]| + |[101]|) - 2g|[11]|
\end{aligned}
\tag{2.4.5}
$$

where β and g are the infection and recovery rates respectively.

Proof. See Appendix A.1 for case $f = 0$. □

2.4.2 Mean field equation

First order dynamics: The idea of the mean field equation is to break any form of correlations with the assumption no correlation exist. From Cor. 2.4.1, we wish to derive the number of pairs $|[01]|$ given that we know only singletons $|[0]|$ and $|[1]|$. If we assume Γ to be large, we find that a cell is in state $|[1]|$ with probability $|[1]|/|\Gamma|$. A site in state "0" has K neighbours, each of them in state "1" with probability $|[1]|/|\Gamma|$. Thus the expected number of pairs in state $(0, 1)$ is

$$|[01]| \approx K|[0]|\frac{|[1]|}{|\Gamma|}.$$

Furthermore, with the coination $|\Gamma| = |[0]| + |[1]|$, we find the mean field approximation.

Corollary 2.4.2. *The mean field approximation for the contact process is given as*

$$
\begin{aligned}
\frac{d}{dt}|[0]|(t) &= -\beta|[0]|\frac{|[1]|(t)}{|\Gamma|} + g|[1]|(t) \\
\frac{d}{dt}|[1]|(t) &= \beta|[0]|\frac{|[1]|(t)}{|\Gamma|} - g|[1]|(t).
\end{aligned}
$$

2.4.3 Pair approximation without contact tracing

Second order dynamics: The pair approximation assumes only pair correlations exists and we only have to use similar approach in the mean field to break correlations over three sites

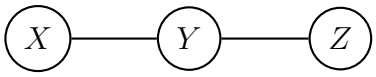


Figure 2.6: Example of a connected triplet

(triplets). In this case, we wish to derive the number of triplets of a certain configuration in terms of the number of pairs and singletons. If we assume Γ to be large, then the number of configurations $(0,0,1)$ can be written as the product of the number of pairs $(0,0)$ and the probability to find a "1" next to "0", i.e., the number of pairs $[0,1]$ divided by the number of zeros, $[0]$. Hence, we get the approximation

$$[0,0,1] \approx ([0,0]/2)\frac{[0,1]}{[0]}$$

The pair $[0,0]$ is symmetrical, so we have to divide by 2 to get the actual number of pairs of type $(0,0)$. Also, the number of configuration $(1,0,1)$ can be written as a product of the number of pairs $(1,0)$ and the probability to find a "1" next to a "0". Similar computation yields $[1,0][0,1]/[0]$. The procedure is also symmetrical as the number of triplets of type $(1,0,1)$ is twice the number of this type. To get the exact number of this type, we again divide $[1,0,1]$ by 2 so that we have the approximation as

$$[1,0,1]/2 \approx \frac{[0,1]^2}{[0]}$$

Hence, we have the closed form equation to be

$$[XYZ] \approx \frac{[XY][YZ]}{[Y]}. \tag{2.4.6}$$

Closing the equations with these terms results in the pair approximation.

Corollary 2.4.3. *The pair approximation of the contact process without contact tracing reads*

$$\begin{aligned}
\frac{d}{dt}|[0]| &= -\tfrac{\beta}{K}|[01]| + g|[1]| \\
\frac{d}{dt}|[1]| &= \tfrac{\beta}{K}|[01]| - g|[1]| \\
\frac{d}{dt}|[00]| &= -\tfrac{2\beta}{K}|[001]| + 2g|[01]| \\
\frac{d}{dt}|[01]| &= \tfrac{\beta}{K}(|[001]| - |[101]| - |[01]|) - g(|[01]| - |[11]|) \\
\frac{d}{dt}|[11]| &= \tfrac{2\beta}{K}(|[101]| + |[01]|) - 2g|[11]|.
\end{aligned}$$

Proof. See Appendix A.1 for case $f = 0$. □

Information about the time evolution of the connected pairs $|[SI]|$ has to be modelled differently which leads to information about connected triplets. However, if we continue this process, computations gets cumbersome and unrealistic to solve due to information about quadruple, quintuple, etc. We therefore use the moment closure approximation in Def. 2.4.6 for the triplets and we result to a closed form equation.

2.4.4 Pair approximation with contact tracing

We have derived above the pair approximation without contact tracing. That is, an individual upon recovery, does not trigger a tracing event. We will now introduce contact tracing into the pair approximation where a recovery event triggers a tracing event. Contact tracing can be achieved via a one-step or recursive scheme which we will discuss.

One step tracing: One-step tracing involves tracing only immediate neighbours of an index case. This means that tracing stops after successfully identifying an immediate contact. For sexually transmitted diseases, one-step tracing is mostly applicable and very useful in this context because both infected partners may seek medical attention and are also treated at once. In this case, the contact tracing process is referred to as partner notification.

We now introduce one-step tracing in the pair approximation model where a treated fraction f of the contacts of an index case receives treatment simultaneously with the index case thereby leading to the recovery of more than one individual in a single recovery event.

Proposition 2.4.1. *The pair approximation of the contact process with one-step tracing is defined as*

$$\begin{aligned}
\frac{d}{dt}|[0]| &= -\tfrac{\beta}{K}|[01]| + g|[1]| + fg|[11]| \\
\frac{d}{dt}|[1]| &= \tfrac{\beta}{K}|[01]| - g|[1]| - fg|[11]| \\
\frac{d}{dt}|[00]| &= -\tfrac{2\beta}{K}|[001]| + 2g|[01]| + 2fg|[11]| + 2fg|[011]| \\
\frac{d}{dt}|[01]| &= \tfrac{\beta}{K}(|[001]| - |[101]| - |[01]|) - g|[01]| + g(1-f)|[11]| + fg(|[111]| - |[011]|) \\
\frac{d}{dt}|[11]| &= \tfrac{2\beta}{K}(|[101]| + |[01]|) - 2g|[11]| - 2fg|[111]|.
\end{aligned}$$

Proof. See Appendix A.1. □

Recursive tracing: In the recursive tracing scheme, traced individuals become new index cases and tracing is triggered again. That is, neighbours of traced persons are further targetted until all expected contacts involved in the chain of infection are traced. Contact

tracing has shown to be very effective, in particular for uncovering high-risk individuals or groups in a population. Recursive tracing may be expensive and time consuming, however, it is also very applicable in scenarios where extra efforts and measures are required to completely eradicate infection.

To include recursive tracing in the pair approximation model, a new tracing class T is introduced where recovered individuals (via tracing or not) become index cases, and triggers contact tracing recursively. In this case, these index cases do not immediately transition into the susceptible class but spend some time in the tracing class before moving into the susceptible class. This is usually the case when index patients are isolated after testing positive for a disease and placed under medical surveillance or monitoring while their potential contacts are traced and informed of their exposure, and advised to seek treatment. After the tracing exercise, these individuals may become exposed again. Thus, we have a $SITS$ model which only differs slightly from the SIS and $SIRS$ model.

We note that the contact process in this section was derived only for the SIS model. We will only state the model for the recursive tracing but the proof can be adapted as derived in Appendix A.1. We now have to include the tracing class T in the configurations and modify the calculations for recovery and tracing events.

Proposition 2.4.2. *The pair approximation of the contact process with recursive tracing is defined as*

$$\begin{aligned}
\frac{d}{dt}|[S]| &= -\tfrac{\beta}{K}|[SI]| + a|[T]| \\
\frac{d}{dt}|[I]| &= \tfrac{\beta}{K}|[SI]| - c|[IT]| - g|[I]| \\
\frac{d}{dt}[T] &= c|[IT]| + g|[I]| - a|[T]| \\
\frac{d}{dt}[SS] &= -\tfrac{2\beta}{K}|[SSI]| + 2a|[ST]| \\
\frac{d}{dt}|[SI]| &= \tfrac{\beta}{K}(|[SSI]| - |[ISI]|) - \tau|[SI]| - g|[SI]| + a|[IT]| - c|[SIT]| \\
\frac{d}{dt}|[ST]| &= -\tfrac{\beta}{K}|[IST]| + g|[SI]| + a(|[TT]| - |[ST]|) + c|[SIT]| \\
\frac{d}{dt}|[II]| &= \tfrac{2\beta}{K}|[ISI]| + 2\tau|[SI]| - 2g|[II]| - 2c|[IIT]| \\
\frac{d}{dt}|[IT]| &= \tfrac{\beta}{K}|[IST]| + g(|[II]| - |[IT]|) - a|[IT]| + c(|[IIT]| - |[TIT]|) - c|[IT]| \\
\frac{d}{dt}|[TT]| &= 2g|[IT]| - 2a|[TT]| + 2c|[IT]| + 2c|[TIT]|,
\end{aligned}$$

where tracing happens at rate c and the average time spent in the tracing class is $1/a$. The evolution of pairs is also modelled and applying the moment closure, we result in a closed

form. The effect of the additional recovery via contact tracing shows reduced prevalence [24].

Chapter 3

Contact tracing on random trees

We follow and extend in some parts of this thesis the publication of Okolie and Müller [66].

The present work starts with a stochastic SIR model on a contact graph that is a rooted tree (See fig. 3.1), an extension of the stochastic SIR model in a homogeneous population by by Müller and Koopmann [60]. That is, a contact graph that is a complete graph. Initially, only the root is infected while all other nodes are susceptible. We follow a similar approach as in the thoery about giant components on random graphs (see section Epidemics on graphs). It turns out that the basis for this analysis is the probability for a randomly choosen node to be still be infectious at a given age of infection.

An individual-based model is used to analyse the contact process and contact tracing on the tree. At the onset of the epidemics, an infected individual who makes an infectious contact with another susceptible individual is called an infector while the individual that is infected is called an infectee. This analogy is the same when this infectee infects another individual. Following the contact structure and transmission routes, the infection spreads along this path forming new generations of newly infected individuals. Recall, an index case is called a primary case while the individual infected by the primary case is called the secondary case.

3.1 Model and analysis

In this section, we state our model assumption based on a stochastic SIR model on a rooted tree. By focusing on the onset of an oubreak, our model is analysed using a branching process approximation [58, 62]. On the tree-network (See fig. 3.1), we start by denoting B as a "downstreams" of A (also A as an "upstream" of B). Given the incidence of transmission, A is the infector of B likewise B is an infectee of A. Let K be the random variable for the number of downstream edges. We note that the root node C is different and has no infector, therefore, there is no upstream edge. It turns that K applies only to

Figure 3.1: Concept of contact tracing. The focal node A can be traced from one of its infectees (B or D) via a backward tracing scheme and from its infector (C) via a forward tracing scheme with probability p. Full tracing is the combination of both forward and backward tracing. The black arrows shows the direction of infection while the green arrow show the direction of contact tracng.

the root individual such that $K + 1$ (infectees plus infector) will be the total number of edges for non-root individuals.

At the onset of the epidemic, i.e time $t = 0$, it is assumed that only the root individual is infected and while other individuals are susceptible. Any infectious contact between a susceptible and an infected individual will transmit the infection. Contacts happens on each edge at an exponentially distributed waiting time with rate β. When an index case is identified, an infector or infectee of the index case, if still infectious at the time of detection, is successfully traced with probability p. Let γ be the rate at which infected individuals recover from a disease without contact tracing. As tracing can only be triggered by directly observing an individual, we can split γ such that the recovery term is the combination of unobserved and observed recovery, i.e, $\gamma := \alpha + \sigma$. First, we assume that an individual recovers spontaneously (unobserved) at rate $\alpha := (1 - p_{obs})\gamma$ or by a direct observation (not via contact tracing) at rate $\sigma = p_{obs}\gamma$ where p_{obs} is the probability for a recovered individual to be observed.

The number of traced individuals within a certain period of time is measured by the rate of tracing. The variables α and σ are generally not constant, but may during the infectious period (within or between individuals). Thus, they are to be considered as functions depending on age since infection a, but in the following, they are assumed to be constant

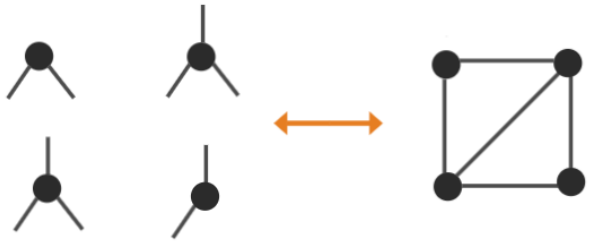


Figure 3.2: The configuration model generates a random network in which stubs are assigned to nodes and are randomly joined together.

for simplicity. The same applies to the contact rate β, which describes the number of contacts between an infected individual and a susceptible neighbour over a certain period of time. Note that we focus on the onset of the epidemic, so that, all contacts of an infected individual connect to susceptible individuals.

Below we investigate the probability for an infected individual to still be infectious at age of infection a (often, we will just write "age", as we do not use chronological age but only age of infection). This probability depends in general on the location of the individual within the contact tree because of contact tracing. Let the generation i of a node be the distance between the node and the root (w.r.t. the standard graph metric). We define

$$\kappa_i(a) = P(\text{a randomly chosen infected node of generation } i \text{ is infectious at age of infection } a) \quad (3.1.1)$$

which turns out to be central for our analysis of the process. In order to obtain this probability, we divide the tracing process in "backward" and "forward" tracing: We call a tracing event "backward tracing" if an individual is detected via a downstream index case, and "forward tracing", if the index case is upstream. In backward tracing we artificially switch off all forward tracing events, while in forward tracing, we only allow for forward tracing events. Of course, this is done for mathematical convenience only. In reality, we always find the full tracing process [35]. This full process can be easily understood as a combination of forward- and backward tracing [62]. For the full tracing, we take into account that an infector can be traced by both his or her infectee or infector. Within the backward or forward tracing, a distinction is also made between the recursive and one step tracing. For the recursive tracing, recovered traced individuals becomes new index cases again and triggers another tracing. In this case, the tracing does not stop once a traced individual becomes an index case. In one step tracing, the tracing stops after one step which means we truncate the tracing at the immediate secondary cases of the primary cases (index cases).

Notation: The symbol κ will appear with several sub- and superscripts. In order to avoid confusion, we summarize the different roles of κ here: Apart from forward- backward- and full tracing, we will consider one step and recursive tracing. Therefore, we use $\kappa^+_{*,i}(a)$ for forward tracing, $\kappa^-_{*,i}(a)$ for backward tracing, and $\kappa_{*,i}(a)$ for full tracing. The asterisk $*$ is either "r" (for recursive) or "o" (for one-step tracing). That is, $\kappa^+_{r,i}(a)$ refers to the

probability to be infectious at age (of infection) a for an individual of generation i, subject to recursive forward tracing. We furthermore denote by

$$\widehat{\kappa}(a) = e^{-(\alpha+\sigma)a}$$

the probability to be infectious at age a if no tracing takes place ($p = 0$). Last, in backward tracing, we note that only the downstream individuals of an (infected) focal individual would be "actors" in the tracing process. Hence, the generation of the focal individual plays no role such that the contact graph becomes an infinite tree. So $\kappa^-_*(a)$ is the (generation-independent) probability in case of backward tracing. Importantly, as we don't know the exact location of this randomly selected focal individual with k downstream edges on the contact graph, we define $\kappa^-_{*,i+1}(\cdot)$ to be the expected value of $\kappa^-_{*,k,i+1}(\cdot)$ (the probabilities for each of the k downstream individuals of a randomly selected infected individual to still be infectious at some age $\cdot$) over all possible k, where the random variable K takes on k.

3.2 Backward tracing

In the following subsections, we will discuss the recursive and one-step mode of backward tracing. In this analysis, we follow the publication of Okolie and Müller [66].

3.2.1 Recursive mode

Theorem 3.2.1. *Let K be a random variable that determines the probability $\kappa^-_{r,k,i}(a)$ that a focal individual of generation i with k downstream edges is still infectious at age a of infection and the probability generating function*

$$\kappa^-_{r,i}(a) = \sum_{k=0}^{\infty} P(K = k)\kappa^-_{r,k,i}(a) \tag{3.2.1}$$

be the unconditional probability to be infectious. With $G(s) = E(s^K)$, we find $\kappa^-_{r,i}(a) = \kappa^-_r(a)$ where

$$\kappa^-_r(a) = \quad e^{-(\alpha+\sigma)a} G\left(1 - p\int_0^a (1-e^{-\beta(\alpha-\tilde{a})})(-\tfrac{d}{d\tilde{a}}\kappa^-_r(\tilde{a}) - \alpha\kappa^-_r(\tilde{a}))d\tilde{a}\right). \tag{3.2.2}$$

Proof. An infected individual can loose his or her infectivity in 3 possible ways;

1 An unobserved recovery at the rate α

2 A direct observation and diagnosed at rate σ

3 A successful tracing event with probability p

Because of these considerations and the differential equations for the SIR model, $\kappa_r^-(a)$ regardless of the type of tracing, satisfies the following differential equation:

$$\frac{d}{da}\kappa(a) = -\kappa(a)\left(\alpha + \sigma + \text{tracing(a)}\right), \tag{3.2.3}$$

where $\kappa(0) = 1$.

If there is no contact tracing, i.e $p = 0$, an infected individual can loose his or her infectivity either by an unobserved recovery at rate α or by direct observation at rate σ. In the absence of tracing, the probability that an infected individual reaches the age of infection a is increased, then

$$\hat{\kappa}(a) := e^{-(\alpha+\sigma)a}.$$

If we include contact tracing, this probability $\kappa_r^-(a)$ is decreased. For the individual to reach the age of infection a, we use a parallel approach by simply multiplying $\hat{\kappa}(a)$ by the probability that no tracing event occurred during the infectious period [0,a] [62]. This, in turn is 1 minus the probability that a tracing event did occur in the interval $[0, a]$. Thus;

$$\kappa_r^-(a) = \hat{\kappa}(a)[1 - \mathbb{P}(\text{ tracing event in the interval } [0, \text{a}))]. \tag{3.2.4}$$

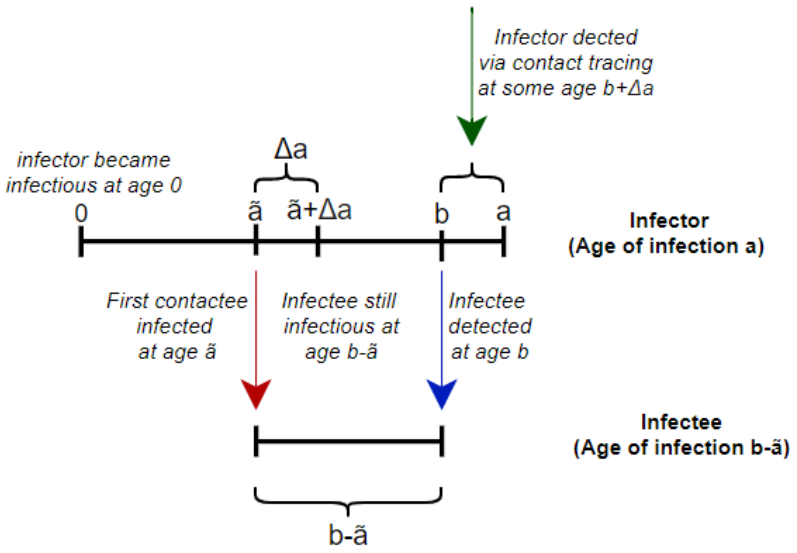


Figure 3.3: A summary of backward tracing events.

For notation clarity in backward tracing, we define which assumptions plays a role or not. As only the root node is known to be infectious at the start of the epidemic process, it is important to note that only forward tracing moves in the same direction as the incidence. That is, a focal individual who is traced via his or her infector is more or less in the same scenario as contracting the disease from his or her infector. In backward tracing, the tracing flows in the opposite direction as the incidence because the individual is traced by his or her infectee (see fig. 3.1). As the tree is considered to be infinite and only the infectees are "actors" in the tracing process, the generation of this focal individual plays no role. So in backward tracing, $\kappa^-_{r,i}(a) = \kappa^-_r(a)$ becomes the (generation-independent) probability to still be infectious.

Following a single edge downstream, the infected individual A is the only person who can infect B. As such, A can only be traced via a backward tracing event triggered by B. Recall that β is the constant rate on the edge between the an infected individual A and a given downstream neighbour of A. The time at B is infected by A is exponentially distributed, so the probability that B is not infected by A until time $\tilde{a}$ is, according to SIR model $e^{-\beta\tilde{a}}$ (See fig. 3.3). Thus, the probability that B is infected by A until time $\tilde{a}$ is $1 - e^{-\beta\tilde{a}}$.

The incidence (probability density) of that downstream node (at age of infection of our focal individual $\tilde{a}$) reads

$$\beta e^{-\beta\tilde{a}}.$$

It is now assumed that individual B was infected at age since infection $\tilde{a}$ of A. In addition, a tracing event is assumed to have been triggered by B at age b. Individual A can only be discovered by this tracing event, though he is still infected until age b. The probability that A is still infected at time b is $\kappa^-_r(b - \tilde{a})$.

This is the case when individual B infection is either diagnosed with rate α or via contact tracing. Equation (3.2.3) gives the hazard

$$\alpha + \sigma + \text{tracing}(\tilde{a}) = -\frac{\kappa^{-\prime}_r(b - \tilde{a})}{\kappa^-_r(b - \tilde{a})}.$$

We subtract the spontaneous recovery rate α from the hazard to obtain the rate of direct or indirect detection of an infected individual since age of infection $\tilde{a}$. Thus

$$\text{hazard}(\tilde{a}) = -\frac{\kappa^{-\prime}_r(b - \tilde{a})}{\kappa^-_r(b - \tilde{a})} - \alpha.$$

Hence to get the rate of tracing at age $b \in [0, a]$ via a given downstream edge, we integrate over b and thus

$$\begin{aligned} \text{rate of tracing}(b) &= p\int_0^b \beta e^{-\beta\tilde{a}}\kappa_r^-(b-\tilde{a})\left(\frac{-\kappa_r^{-\prime}(b-\tilde{a})}{\kappa_r^-(b-\tilde{a})} - \alpha\right)d\tilde{a} \\ &= p\int_0^b \beta e^{-\beta\tilde{a}}\left(-\kappa_r^{-\prime}(b-\tilde{a}) - \alpha\kappa_r^-(b-\tilde{a})\right)d\tilde{a} \end{aligned}$$

and the probability that a successful tracing event by that downstream edge in interval [0,a] is given by p(tracing event by one edge during the infectious interval [0,a]). We have the

$$\begin{aligned} \text{probability of tracing in [0,a]} &= p\int_0^a\int_0^b \beta e^{-\beta\tilde{a}}\left(-\kappa_r^{-\prime}(b-\tilde{a}) - \alpha\kappa_r^-(b-\tilde{a})\right)d\tilde{a}db \\ &= p\int_0^a\int_0^b \beta e^{-\beta(b-\tilde{a})}\left(-\kappa_r^{-\prime}(\tilde{a}) - \alpha\kappa_r^-(\tilde{a})\right)d\tilde{a}db \\ &= p\int_0^a\int_{\tilde{a}}^a -\frac{d}{d\tilde{a}}e^{-\beta(b-\tilde{a})}db\left(-\kappa_r^{-\prime}(\tilde{a}) - \alpha\kappa_r^-(\tilde{a})\right)d\tilde{a} \\ &= p\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\left(-\kappa_r^{-\prime}(\tilde{a}) - \alpha\kappa_r^-(\tilde{a})\right)d\tilde{a}. \end{aligned}$$

We recall that $\kappa_{r,k,i}^-(a)$ denotes the probability to still be infectious under the condition that the individual is in generation i and has k downstream neighbours. That is

$$\kappa_{r,k,i}^-(a) = e^{-(\alpha+\sigma)a}\left(1 - p\int_0^a (1-e^{-\beta(a-\tilde{a})})(-\kappa_r^{-\prime}(\tilde{a}) - \alpha\kappa_r^-(\tilde{a}))d\tilde{a}\right)^k.$$

and

$$G(s) = E(s^K) = \sum_{i=0}^{\infty} P(K=k)s^k$$

We remove the generation i condition from $\kappa_{r,k,i}^-(a)$ because the generation of the "actors" in backward tracing plays no role. Last, for independent k number of downstream edges, we remove the condition k by

$$\kappa_r^-(a) = \sum_{i=0}^{\infty} P(K=k)\kappa_{r,k}^-(a)$$

such that

$$\kappa_r^-(a) = e^{-(\alpha+\sigma)a} G\left(1 - p\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\left(-\kappa_r^{-\prime}(\tilde{a}) - \alpha\kappa_r^-(\tilde{a})\right)d\tilde{a}\right).$$

□

Proposition 3.2.1. *Let $\beta \neq \alpha + \sigma$. For all moments of K finite, the first order approximation of $\kappa_r^-(a)$ in p reads*

$$\kappa_r^-(a) = \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1 - \hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2). \tag{3.2.5}$$

Proof. The first order approximation of $\kappa_r^-(\tilde{a})$ are only the tracing events of interest which are initiated by the focal individual or neighbours of the focal individual at rate σ. From Theorem 3.2.1, we can write

$$\begin{aligned}
\kappa_r^-(a) &= e^{-(\alpha+\sigma)a}\, G\left(1 - p\int_0^a (1-e^{-\beta(a-\tilde{a})})\sigma\kappa_r^-(\tilde{a})d\tilde{a}\right) + \mathcal{O}(p^2), \\
&= e^{-(\alpha+\sigma)a}\, G\left(1 - p\sigma\int_0^a (1-e^{-\beta(a-\tilde{a})})\kappa_r^-(\tilde{a})d\tilde{a}\right) + \mathcal{O}(p^2).
\end{aligned}$$

Now consider the expression;

$$\int_0^a (1-e^{-\beta(a-\tilde{a})})\kappa_r^-(\tilde{a})d\tilde{a}.$$

To obtain a first order approximation, we may replace $\kappa_r^-(\tilde{a})$ in the argument of $G(.)$ by its zeroth order approximation. The zeroth order approximation of $\kappa_r^-(\tilde{a})$ is given by $\hat{\kappa}(a)$

$$\kappa_r^-(a) = e^{-(\alpha+\sigma)a} + \mathcal{O}(p) = \hat{\kappa}(a) + \mathcal{O}(p).$$

Then

$$\begin{aligned}
\int_0^a (1-e^{-\beta(a-\tilde{a})})\kappa_r^-(\tilde{a})d\tilde{a} &= \int_0^a (1-e^{-\beta(a-\tilde{a})})e^{-(\alpha+\sigma)\tilde{a}}d\tilde{a} + \mathcal{O}(p) \\
&= \frac{-1}{\alpha+\sigma}\left(e^{-(\alpha+\sigma)\tilde{a}}\right)\Big|_0^a - \frac{e^{-\beta a}}{\alpha+\sigma-\beta}\left(-e^{-(\alpha+\sigma-\beta)\tilde{a}}\right)\Big|_0^a \\
&= \frac{1}{\alpha+\sigma}\left(1 - e^{-(\alpha+\sigma)}\right) - \frac{1}{\alpha+\sigma-\beta}\left(e^{-\beta a} - e^{-(\alpha+\sigma)a}\right) \\
&= \frac{1}{\alpha+\sigma}\left(1 - \hat{\kappa}(a)\right) - \frac{1}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) + \mathcal{O}(p).
\end{aligned}$$

Furthermore, for any $x \in \mathbb{R}$, we have

$$\begin{aligned} G(1-px) &= \sum_{k=0}^{\infty} P(K=k)(1-px)^k = \sum_{k=0}^{\infty} P(K=k)(1-pxk) + \mathcal{O}(p^2) \\ &= G(1) - pxG'(1) + \mathcal{O}(p^2). \end{aligned}$$

Therewith,

$$\begin{aligned} \kappa_r^-(a) &= e^{-(\alpha+\sigma)a}\left(G(1) + \frac{p\sigma G'(1)}{\alpha+\sigma-\beta}\left(e^{-\beta a} - e^{-(\alpha+\sigma)a}\right) - \frac{p\sigma G'(1)}{\alpha+\sigma}\left(1 - e^{-(\alpha+\sigma)a}\right)\right) + \mathcal{O}(p^2) \\ &= \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2). \end{aligned}$$

□

As shown in Fig. 3.4, the probability to be infectious is reduced as a result of the tracing term. The numerical solution of eqn. 3.2.2 and the first order approximation given in eqn. 3.2.5 is compared with numerical simulations. For small p, which is given in most applications, we find a good agreement (Fig. 3.4).

Find some notes related to the simulation algorithm in the appendix.

3.2.2 One step mode

The only difference between the recursive and the one-step method is that at the one-step method the tracing can be aborted after one step. Illustratively, this means that person A can only be discovered through tracing events, which are triggered by direct neighbors of A with rate σ, but not by tracing events from which neighbors of A are discovered. Therefore, the one-step tracing as infected individual is the rate of direct detection σ. From equation 3.2.2, we replace the removal term with σ.

Theorem 3.2.2. *The probability that an infected individual under one step backward tracing is still be infectious at age a reads*

$$\kappa_o^-(a) = e^{-(\alpha+\sigma)a}\, G\left(1 - p\int_0^a (1-e^{-\beta(\alpha-\tilde{a})})\sigma\kappa_o^-(\tilde{a})d\tilde{a}\right). \tag{3.2.6}$$

Proof. Following the same arguments in the recursive mode, the prove is similar to Theorem 3.2.1. As before, we obtain the equation stated above. □

Proposition 3.2.2. *The first order approximation of κ_o^- reads*

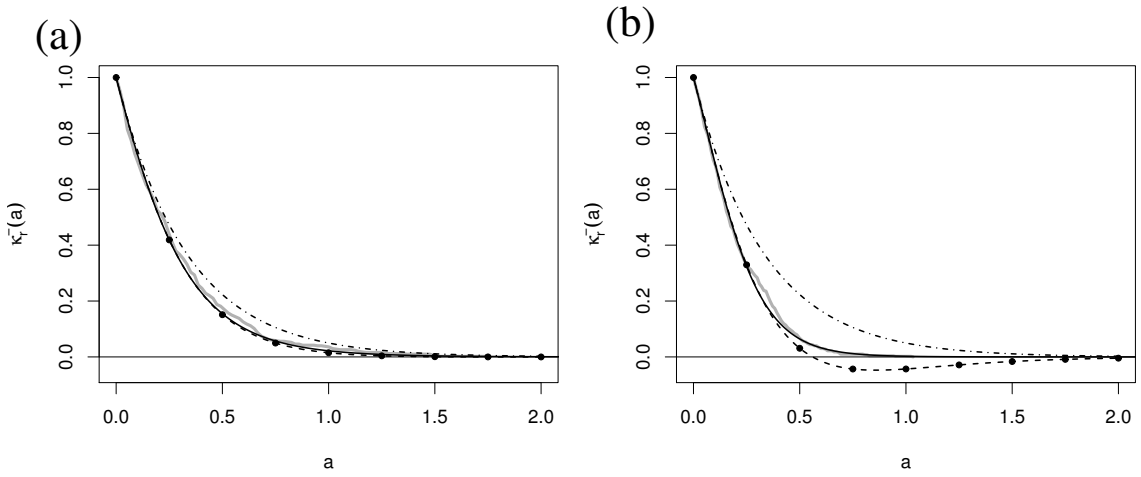


Figure 3.4: Recursive backward tracing. The probability to be infectious at age a after infection ($\kappa_r^-(a)$) for $p = 0.3$ (panel a), and $p = 0.8$ (panel b). Dashed-dotted line: without tracing ($\hat{\kappa}(a) = e^{-(\sigma+\alpha)a}$); solid line: numerical solution of $\kappa_r^-(a)$; dashed-lines with bullets: first order approximation of $\kappa_r^-(a)$; grey line: simulation. Parameters: $\beta = 1.5$, $\alpha = 0.1$, $\sigma = 2.9$, $\mathbb{E}[K] = 4$, fixed degree.

$$\kappa_o^-(a) = \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2). \tag{3.2.7}$$

Proof. The proof parallels Proposition 3.2.1 and applies also to $\kappa_o^-(a)$. Hence, we can proceed to determine the approximation exactly as with the recursive method. □

Remark 3.2.1. *In the above, we have examined the backward tracing event, taking into consideration the recursive mode κ_r^- and one step tracing κ_o^-. Results of the first order approximations of κ_r^- and κ_o^- shows to coincides. Tracing longer chains with length k contributes with probability p^k. Since the tracing events with length 2 or more steps do not play a role in the first order approximation, they disappear in the remainder term $\mathcal{O}(p^2)$. The results seem plausible since both approximations in both cases seem to follow a trend. The first order approximation of the recursive mode takes into account only its immediate neighbours which is also similar to one step tracing approach.*

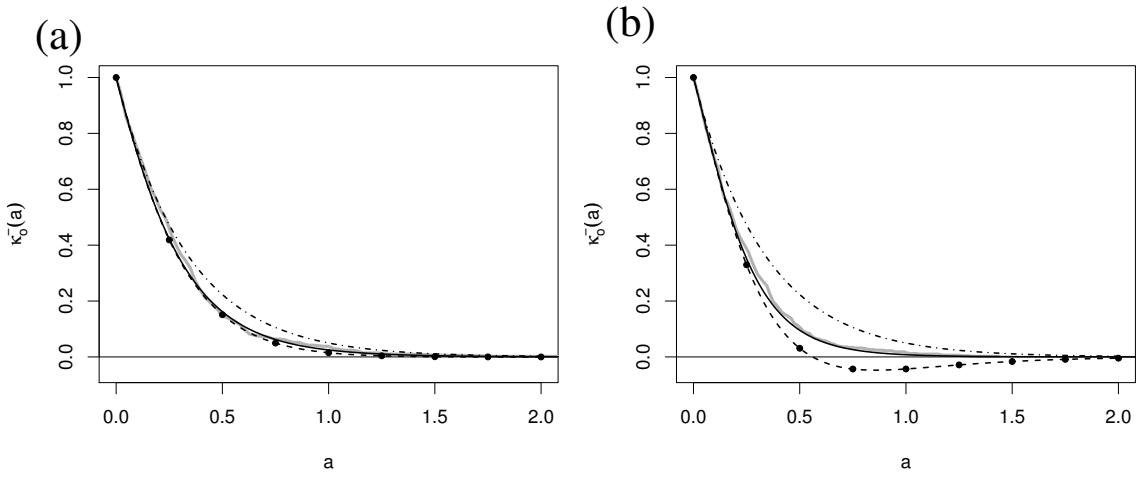


Figure 3.5: One-step backward tracing. The probability to be infectious at age a after infection ($\kappa_o^-(a)$) for $p = 0.3$ (panel a), and $p = 0.8$ (panel b). Dashed-dotted line: without tracing ($\hat{\kappa}(a) = e^{-(\sigma+\alpha)a}$); solid line: numerical solution of $\kappa_o^-(a)$; dashed-lines with bullets: first order approximation of $\kappa_o^-(a)$; grey line: simulation. Parameters: $\beta = 1.5$, $\alpha = 0.1$, $\sigma = 2.9$, $\mathbb{E}[K] = 4$, fixed degree.

3.3 Connection to results of homogeneously mixing populations

3.3.1 First approximation argument

The approximation results of $\kappa_*^-(a)$ calculated in this present work is based on a population whose contact graph is a (stochastic) tree. Interestingly enough, we can relate the result for random trees to the corresponding results for randomly mixing population, derived in [60, 62]. In a randomly mixing population, the contact rate is not defined per edge (β), but for an individual (β_{ind}). The model and the meaning of all other parameters parallel that of the present work, with the exception that the contact graph is not a tree but a complete graph. The corresponding probability $\kappa_h^-(a)$ to be infectious at age a of infection for an homogeneous population (considered in the onset of an infection, and one step backward tracing only) is given by the equation (Proposition 2.1 and Proposition 2.2 in Müller and Koopmann [60])

$$\frac{d}{da}\kappa_h^-(a) = -\kappa_h^-(a)\left[\alpha + \sigma + p\sigma\beta_{ind}\int_0^a \kappa_h^-(\tilde{a})d\tilde{a}\right], \quad \kappa_h^-(0) = 1 \tag{3.3.1}$$

with the first order approximation given as

$$\kappa_h^-(a) = \hat{\kappa}(a) - p\frac{\sigma}{\alpha+\sigma}\beta_{ind}\hat{\kappa}(a)\left(a - \frac{1-\hat{\kappa}(a)}{\alpha+\sigma}\right) + \mathscr{O}\left(p^2\right) \tag{3.3.2}$$

One central difference between this present equations and the homogeneous models is the appearance of β_{ind}. If we select randomly a non root node in the present model, then the total contact rate of that node is the sum of all contacts on its edges. For all individuals apart from the root, the relation between β and β_{ind} is given by

$$\beta_{ind} = \beta(1 + \mathbb{E}[K]).$$

Here we obtain $1+\mathbb{E}[K]$, as contacts happen on upstream and on downstream nodes. It is near at hand to consider the limit $\mathbb{E}[K] \to \infty$, $\beta \to 0$, while $\beta_{ind} := \beta(1+\mathbb{E}[K])$ is constant to approximate a randomly mixing population (full graph). We start from the first order approximation results which will be shown below.

Proposition 3.3.1. *The following applies:*

$$\kappa_h^-(a) = \lim_{\beta\to 0,\, \mathbb{E}[K]\to\infty} \kappa_o^-(a) + \mathcal{O}(p^2), \tag{3.3.3}$$

under the condition that the constant $\beta_{ind} := \beta(1 + \mathbb{E}[K])$ is satisfied.

Proof. According to Proposition 3.2.1, the following applies to $\kappa_o^-(a)$;

$$\kappa_o^-(a) = \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2).$$

Now, consider the following expression;

$$\begin{aligned}
&\lim_{\beta\to 0} \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - e^{-(\alpha+\sigma)a}\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1 - e^{-(\alpha+\sigma)a}\right)\Big|_{\beta_{ind}=\beta(1+\mathbb{E}[K])} \\
&= \lim_{\beta\to 0} \frac{p\sigma\beta_{ind}}{\alpha+\sigma}\frac{1}{\beta}\left(\frac{\alpha+\sigma}{\alpha+\sigma-\beta}\left(e^{-\beta a} - e^{-(\alpha+\sigma)a}\right)\right) - \frac{\alpha+\sigma}{\alpha+\sigma}\left(1 - e^{-(\alpha+\sigma)a}\right) \\
&= \frac{p\sigma\beta_{ind}}{\alpha+\sigma}\frac{d}{d\beta}\left(\frac{\alpha+\sigma}{\alpha+\sigma-\beta}\left(e^{-\beta a} - e^{-(\alpha+\sigma)a}\right)\right)\Big|_{\beta=0} \\
&= \frac{p\sigma\beta_{ind}}{\alpha+\sigma}\left(\frac{1}{\alpha+\sigma}\left(1 - e^{-(\alpha+\sigma)a}\right) - a\right).
\end{aligned}$$

Hence for $\beta_{ind} := \beta(1 + \mathbb{E}[K])$ is constant and $\beta \to 0$, while $\mathbb{E}[K] \to \infty$, we have the first order approximation of $\kappa_o^-(\tilde{a})$ in p tends to

$$\kappa_o^-(a) = e^{-(\alpha+\sigma)a} - \frac{p\sigma\beta_{ind}}{\alpha+\sigma} e^{-(\alpha+\sigma)a} \left(a - \frac{1}{\alpha+\sigma} \left(1 - e^{-(\alpha+\sigma)a} \right) \right) + \mathcal{O}(p^2).$$

This corresponds to the first order approximation in the homogeneous case $\kappa_h(a)$ □

3.3.2 Second approximation argument

We have shown above that under the limit for any degree distribution, the results of our tree model and random mixing model of [60] shows to agree. Let us now assume a natural degree distribution K that follows the Poisson process $K \sim Pois\,(\mathbb{E}[K])$. The random mixing model [60] assumes that contacts happens at randomly distributed times, i.e homogeneous Poisson process, and that the number of contactees over a given period of time follows a Poisson distribution. Thus, we wish to examine our tree model with a Poisson degree distribution.

Proposition 3.3.2. *Let* $E(s^K) = \sum_{i=0}^{\infty} \frac{1}{i!}(s\mathbb{E}[K])^i e^{-\mathbb{E}[K]}$ *be the probability generating function for a poissionic distribution, then the following applies*

$$\frac{d}{da}\kappa_o^-(a) = -\kappa_o^-(a) \left(\alpha + \sigma + p\sigma\beta_{ind} \int_0^a \kappa_0^-(\tilde{a}) d\tilde{a} \right). \tag{3.3.4}$$

Proof. The PGF reads

$$E(s^K) = \sum_{i=0}^{\infty} \frac{1}{i!}(s\mathbb{E}[K])^i e^{-\mathbb{E}[K]} = e^{(s-1)\mathbb{E}[K]} = e^{-(1-s)\mathbb{E}[K]}.$$

We recall from Theorem 3.2.2 that

$$\kappa_o^-(a) = e^{-(\alpha+\sigma)a}\, G\left(1 - p \int_0^a (1 - e^{-\beta(\alpha-\tilde{a})})\sigma\kappa_o^-(\tilde{a}) d\tilde{a} \right).$$

We have

$$\kappa_o^-(a) = e^{-(\alpha+\sigma)a} \exp\left\{ - \left(1 - p\sigma \int_0^a (1 - e^{-\beta(a-\tilde{a})}\kappa_o^-(\tilde{a}) d\tilde{a} \right) \mathbb{E}[K] \right\}.$$

The time derivative gives

$$\frac{d}{da}\kappa_o^-(a) = -\kappa_o^-(a)\left(\alpha + \sigma + \mathbb{E}[K]p\sigma\beta\int_0^a e^{-\beta(a-\tilde{a})}\kappa_o^-(\tilde{a})d\tilde{a}\right).$$

If we again take for $\beta_{ind} := \beta(1 + \mathbb{E}[K])$ remains constant, $\beta \to 0$ while $\mathbb{E}[K] \to \infty$, we obtain

$$\frac{d}{da}\kappa_o^-(a) = -\kappa_o^-(a)\left(\alpha + \sigma + p\sigma\beta_{ind}\int_0^a \kappa_o^-(\tilde{a})d\tilde{a}\right).$$

□

Remark 3.3.1. *Results from the tree model [66] shows to agree with the results obtained in a randomly mixing homogeneous population [60] under suited limit conditions. For a small tracing probability, we have shown that not only that both models agree at the first order approximation in p, but also that in this limit (and that if K follows a well behaved degree distribution), the models themselves coincide.*

3.4 Forward tracing

In contrast to backward tracing, there is no difference in forward tracing, whether one assumes a homogeneous population or a network of contacts with a tree structure, because in forward tracing, one moves in the direction in which the infection spreads and in both models, it becomes assumed that all neighbours of an infected person except the infector are susceptible and each infected can be unambiguously assigned to one infector. So the following formulas developed the same procedure as Müller and Koopmann [60].

3.4.1 Recursive mode

Definition 3.4.1. *Let $\kappa_{r,i}^+(a|b)$ be the probability that an individual of generation i is still infected at age a of infection given that the infector has age $a + b$ of infection.*

Proposition 3.4.1. *For the probability that an ith-generation individual with age a of infection is still infected, the following applies for $\kappa_{r,i}^+(a|b)$, the recursion formula*

$$\kappa_{r,i-1}^+(a) = \hat{\kappa}(a)\left(1 - p\frac{\int_0^\infty\int_0^a -\kappa_{r,i-1}^+{}'(b+c) - \alpha\kappa_{r,i-1}^+(b+c)dcdb}{\int_0^\infty \kappa_{r,i-1}^+(b)db}\right) \tag{3.4.1}$$

with

$$\kappa_{r,0}^+(a) = e^{-(\alpha+\sigma)a} = \hat{\kappa}(a).$$

Proof. If an individual with age a of infection has not yet been traced, he or she is still infectious with a probability $\hat{\kappa}(a)$. This corresponds exactly to the probability that an individual of the 0th-generation with infection age a is still infected, since they can not be traced by forward tracing.

However, if the individual is not from the 0th-generation, this probability is decreased by tracing via the infector. Hence, to compute $\kappa^+_{r,i}(a|b)$, we multiply $\hat{\kappa}(a)$ by the probability not to be traced through the infector. This is exactly one minus the probability to be traced via the infector. The probability of being traced by the infector is just p times the probability that the infector is part of a tracing event.

For the probability that the infector is still infected at time $a+b$, the following applies

$$\kappa^+_{r,i-1}(a+b) = \frac{\kappa^+_{r,i-1}(a+b)}{\kappa^+_{r,i-1}(b)}.$$

because at age a, an infectious event did take place. The infectious individual infected the infected individual at a time b and thus had to be infectious at time b. The rate at which the infector is observed at age $b+c$ with $c \in [0,a)$ reads

$$\frac{\kappa^+_{r,i-1}(b+c)}{\kappa^+_{r,i-1}(b)}\left(\frac{-{\kappa^+_{r,i-1}}'(b+c)}{\kappa^+_{r,i-1}(b+c)} - \alpha\right) = \left(\frac{-{\kappa^+_{r,i-1}}'(b+c)}{\kappa^+_{r,i-1}(b)} - \frac{\alpha\kappa^+_{r,i-1}(b+c)}{\kappa^+_{r,i-1}(b)}\right).$$

Therefore, the different times in $c \in [0,a)$ at which the infector triggers a tracing event sums up to

$$\int_0^a \left(\frac{-{\kappa^+_{r,i-1}}'(b+c)}{\kappa^+_{r,i-1}(b)} - \frac{\alpha\kappa^+_{r,i-1}(b+c)}{\kappa^+_{r,i-1}(b)}\right) dc.$$

This term indicates the probability of a tracing event at the infector until age a of infection of the infected individual. From this, it is now possible as described above to calculate $\kappa^+_{r,i}(a|b)$. Thus

$$\kappa^+_{r,i}(a|b) = \hat{\kappa}(a)\left\{1 - p\int_0^a \left(\frac{-{\kappa^+_{r,i-1}}'(b+c)}{\kappa^+_{r,i-1}(b)} - \frac{\alpha\kappa^+_{r,i-1}(b+c)}{\kappa^+_{r,i-1}(b)}\right) dc\right\}$$

The probability that an infected individual is still infectious at age a of infection if the infector has age $a+b$ of infection is now multiplied by $\kappa^+_{r,i-1}(b)$, i.e the probability that the infector was infectious at the time of infection. We have

$$\kappa^+_{r,i}(a|b)\kappa^+_{r,i-1}(b) = \hat{\kappa}(a)\left\{\kappa^+_{r,i-1}(b) - p\int_0^a \left(-\kappa^+_{r,i-1}{}'(b+c) - \alpha\kappa^+_{r,i-1}(b+c)\right)dc\right\}.$$

The probability density of infection age of the infector in an infection is given by

$$\frac{\beta\kappa^+_{r,i-1}(b)}{\int_0^\infty \beta\kappa^+_{r,i-1}(b)db} = \frac{\kappa^+_{r,i-1}(b)}{\int_0^\infty \kappa^+_{r,i-1}(b)db}.$$

This gives the probability that an infected individual with age a of infection was infected by an infector with infection age $a+b$ multiplied by the probability that the infector had infection age b at the time of infection:

$$\kappa^+_{r,i}(a|b)\frac{\kappa^+_{r,i-1}(b)}{\int_0^\infty \kappa^+_{r,i-1}(b)db}.$$

The summation over the different times b at which there is an infection between the infector and the infectee then gives a recursive formula for the probability that the infected individual has age a of infection. This reads

$$\begin{aligned}
\kappa^+_{r,i}(a) &= \int_0^\infty \kappa^+_{r,i}(a|b)\frac{\kappa^+_{r,i-1}(b)}{\int_0^\infty \kappa^+_{r,i-1}(b)db}db \\
&= \frac{\int_0^\infty \kappa^+_{r,i}(a|b)\kappa^+_{r,i-1}(b)db}{\int_0^\infty \kappa^+_{r,i-1}(b)db} \\
&= \frac{\int_0^\infty \hat{\kappa}(a)\left(\kappa^+_{r,i-1}(b) - p\int_0^a \left(-\kappa^+_{r,i-1}{}'(b+c) - \alpha\kappa^+_{r,i-1}(b+c)\right)dc\right)db}{\int_0^\infty \kappa^+_{r,i-1}(b)db} \\
&= \hat{\kappa}(a)\left(1 - p\frac{\int_0^\infty\int_0^a \left(-\kappa^+_{r,i-1}{}'(b+c) - \alpha\kappa^+_{r,i-1}(b+c)\right)dcdb}{\int_0^\infty \kappa^+_{r,i-1}(b)db}\right)
\end{aligned}$$

Since

$$\begin{aligned}
\int_0^\infty\int_0^a f(b+c)dcdb = \int_0^\infty\int_b^{a+b} f(c)dcdb &= \int_0^a\int_0^c f(c)dbdc + \int_a^\infty\int_{c-a}^c f(c)dbdc \\
&= \int_0^\infty \min\{a,c\}f(c)dc
\end{aligned}$$

we find

$$\kappa^+_{r,i}(a) = \hat{\kappa}(a)\left(1 - p\frac{\int_0^\infty \min\{a,b\}\left(-\kappa^{+\prime}_{r,i-1}(b) - \alpha\kappa^+_{r,i-1}(b)\right)db}{\int_0^\infty \kappa^+_{r,i-1}(b)db}\right)$$

and it follows the claim. □

The given formula for $\kappa^+_{r,i}(a)$ is exact but for calculations rather cumbersome, we determine in the following, the first order approximation of $\kappa^+_{r,i}(a)$.

Proposition 3.4.2. *The first order approximation of $\kappa^+_{r,i}(a)$ reads (recall $p_{obs} = \sigma/\alpha + \sigma$)*

$$\kappa^+_{r,i}(a) = \hat{\kappa}(a)\left(1 - pp_{obs}\left(1 - \hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2) \tag{3.4.2}$$

Proof. For the recursive formula for $\kappa^+_{r,i}(a)$, it is easy to see that a first order approximation of $\kappa^+_{r,i}(a)$ is required. We not that $\kappa^+_{r,i-1}(a) = e^{-(\alpha+\sigma)a} + \mathcal{O}(p) = \hat{\kappa}(a) + \mathcal{O}(p)$.

This applies to the approximation of the first order taking into account the fact that the tracing event that traced the infector is not important for the first-order approximation but only the tracing events triggered by the infector itself is included in the approximation. Thus

$$\begin{aligned}
\kappa^+_{r,i}(a) &= \hat{\kappa}(a)\left(1 - p\frac{\int_0^\infty\int_0^a \sigma\kappa^+_{r,i-1}(b+c)dcdb}{\int_0^\infty \kappa^+_{r,i-1}(b)db}\right) + \mathcal{O}(p^2)\\
&= \hat{\kappa}(a)\left(1 - p\frac{\sigma\int_0^\infty e^{-(\alpha+\sigma)b}\int_0^a e^{-(\alpha+\sigma)c}dcdb}{\int_0^\infty e^{-(\alpha+\sigma)b}db}\right) + \mathcal{O}(p^2)\\
&= \hat{\kappa}(a)\left(1 - p\frac{\sigma\int_0^\infty e^{-(\alpha+\sigma)b}\frac{1}{\alpha+\sigma}\left(1 - e^{-(\alpha+\sigma)a}\right)db}{\frac{1}{\alpha+\sigma}}\right) + \mathcal{O}(p^2)\\
&= \hat{\kappa}(a)\left(1 - pp_{obs}\left(1 - \hat{\kappa}(a)\right)\frac{\sigma\int_0^\infty e^{-(\alpha+\sigma)b}db}{\frac{1}{\alpha+\sigma}}\right) + \mathcal{O}(p^2)\\
&= \hat{\kappa}(a)\left(1 - pp_{obs}(1 - \hat{\kappa}(a))\right) + \mathcal{O}(p^2).
\end{aligned}$$

□

3.4.2 One step mode

Similar to backward tracing, the difference between the recursive method and the one-step method is that in the one-step method, an infected individual discovered via tracing cannot trigger another tracing event.

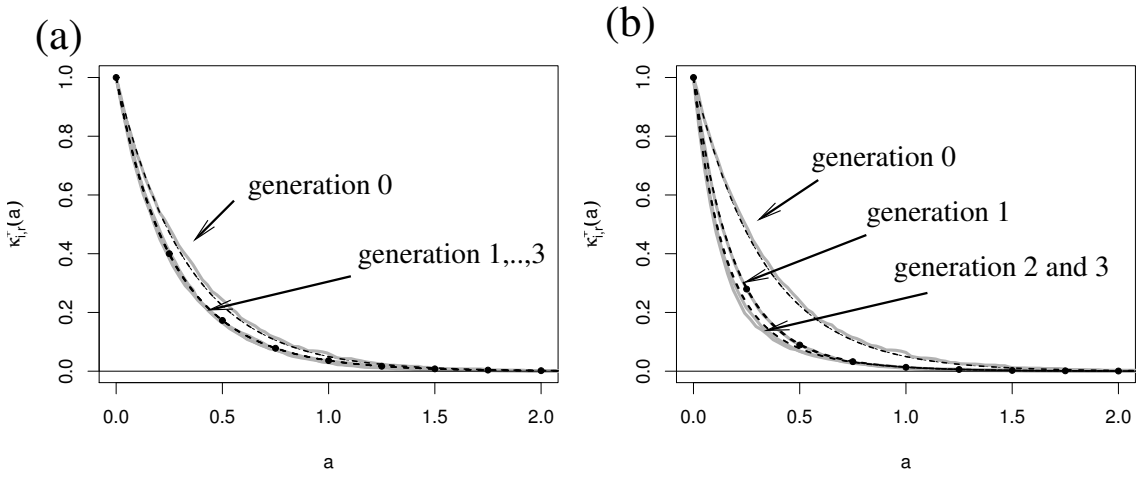


Figure 3.6: Recursive forward tracing. The probability to be infectious at age a after infection ($\kappa^+_{r,i}(a)$) for $p = 0.3$ (panel a), and $p = 0.8$ (panel b). Dashed lines: numerical solution of $\kappa^+_{r,i}(a)$; dashed-dotted line: without tracing ($\hat{\kappa}(a) = e^{-(\sigma+\alpha)a}$) coinciding with $\kappa^+_{r,0}(a)$; dashed-lines with bullets: first order approximation of $\kappa^+_{r,i}(a)$; grey line: simulation. Parameters: $\beta = 1.5$, $\alpha = 0.1$, $\sigma = 2.9$, $\mathbb{E}[K] = 4$, K is constant.

Proposition 3.4.3. *For the probability that an individual of generation i reaches age a of infection in forward tracing with the one-step method, the following applies:*

$$\kappa^+_{o,i} = \hat{\kappa}(a)\left(1 - p\frac{\int_0^\infty \int_0^a \sigma\kappa^+_{o,i-1}(b+c)dcdb}{\int_0^\infty \kappa^+_{o,i-1}(b)db}\right) \tag{3.4.3}$$

with

$$\kappa^+_{o,0}(a) = e^{-(\alpha+\sigma)a} = \hat{\kappa}(a).$$

Proof. The proof follows Proposition 3.4.1 by replacing $-\kappa^{+\prime}_{r,i-1}(a) - \alpha\kappa^+_{r,i-1}(a)$ with $-\sigma\kappa^+_{o,i-1}(a)$ for the reason stated above. □

Proposition 3.4.4. *The first order approximation for $\kappa^+_{o,i}(a)$ reads*

$$\kappa^+_{o,i}(a) = \hat{\kappa}(a)\left(1 - pp_{obs}\left(1 - \hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2). \tag{3.4.4}$$

Proof. See proof of Proposition 3.2.2 □

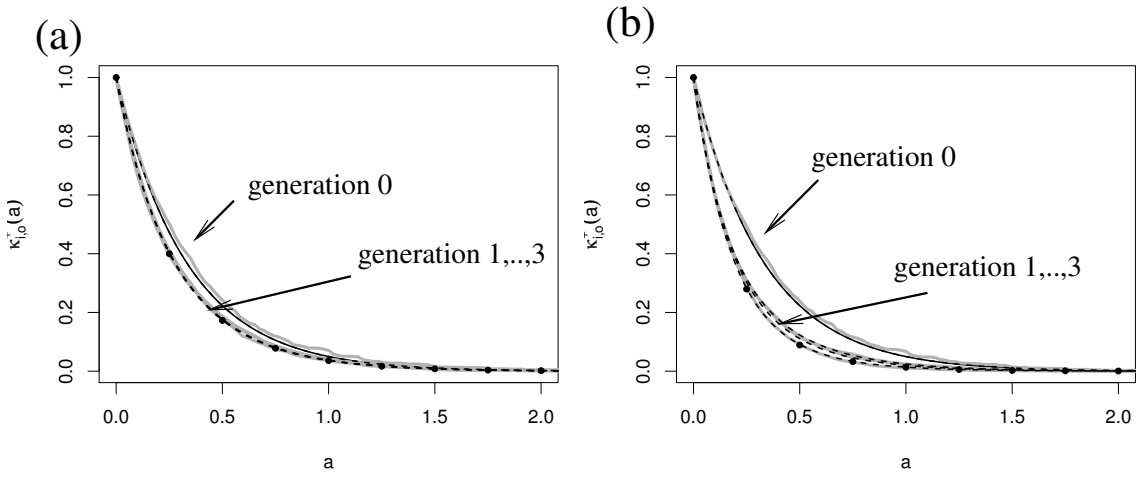


Figure 3.7: One-step forward tracing. The probability to be infectious at age a after infection ($\kappa^+_{o,i}(a)$) for $p = 0.3$ (panel a), and $p = 0.8$ (panel b). Dashed lines: numerical solution of $\kappa^+_{o,i}(a)$; solid line: without tracing ($\hat{\kappa}(a) = e^{-(\sigma+\alpha)a}$) coinciding with $\kappa^+_{o,0}(a)$; dashed-lines with bullets: first order approximation of $\kappa^+_{o,i}(a)$; grey line: simulation. Parameters: $\beta = 1.5$, $\alpha = 0.1$, $\sigma = 2.9$, $\mathbb{E}[K] = 4$, K is constant.

Remark 3.4.1. *As expected, the first order approximations of $\kappa^+_{r,i}$ and $\kappa^+_{o,i}$ coincide. Because the difference between the recursive and the one-step method results only from the second tracing step. This comes with a probability of $\mathcal{O}(p^2)$. Thus, the difference between the two methods in the approximation is expressed only in the remainder of the term. The approximation for the first order of $\kappa^+_{r,i}$ and $\kappa^+_{o,i}$ are independent of the generation of the individual in consideration and also agree with the given first order approximation of Müller and Koopmann [60]. Again, this is not surprising given that the formulas for $\kappa^+_{r,i}$ and $\kappa^+_{o,i}$ have been developed using the same approach.*

3.5 Full tracing

In the following, a formula for full tracing will be developed. This is precisely the probability that an infected person reaches the age of infection a when full tracing is applied. For this, the formulas for forward and backward tracing set up in the previous sections (recursive or one step mode) must be suitably combined.

3.5.1 Recursive mode

Proposition 3.5.1. *The following applies*

$$\kappa_{r,i}(a) = \frac{\int_0^\infty \kappa_r(a)\left(\kappa_{r,i-1}(b) - p\int_0^a -\kappa_{r,i-1}{}'(b+c) - \alpha\kappa_{r,i-1}(b+c)dc\right)db}{\int_0^\infty \kappa_{i-1}(b)db} \tag{3.5.1}$$

with

$$\kappa_r(a) = e^{-(\alpha+\sigma)a}\, G\left(1 - p\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\left(-\kappa_r^-(\tilde{a}) - \alpha\kappa_r^-(\tilde{a})\right)d\tilde{a}\right).$$

Proof. (Proposition 3.5.1) The concept of full tracing is to do the same things as the forward tracing only with the difference that an individual of the zeroth generation can be traced by recursive backward tracing. We use the notation and thus $\kappa_r(a) = \kappa_r^-(a)$. Thus, the formulas of Theorem 3.2.1 is well suited and the claim follows. □

Proposition 3.5.2. *The first order approximation of $\kappa_{r,i}(a)$ reads*

$$\kappa_{r,i}(a) = \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right) - \frac{p\sigma}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2). \tag{3.5.2}$$

Proof. We follow the claim that $\kappa_r(a) = \kappa_r^-(a)$ and also by proposition 3.2.1, the following applies

$$\kappa_r(a) = \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2)$$

Taking the approximation into the formula for $\kappa_{r,i}(a)$ in eqn. 3.5.2 and extracting it from the integral yields

$$\begin{aligned}\kappa_{r,i}(a) = \hat{\kappa}(a)&\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(\hat{\kappa}(a) - e^{-(\alpha+\sigma)a}\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right)\\ &\times \frac{\int_0^\infty \left(\kappa_{r,i-1}(b) - p\int_0^a -\kappa'_{r,i-1}(b+c) - \alpha\kappa_{r,i-1}(b+c)dc\right)db}{\int_0^\infty \kappa_{r,i-1}(b)db} + \mathcal{O}(p^2).\end{aligned}$$

The second part of this expression can be approximated exactly as in Proposition 3.4.2 and this results in

$$\kappa_{r,i}(a) = \hat{\kappa}(a)\left(1+\frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a}-\hat{\kappa}(a)\right)-\frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)-\frac{p\sigma}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right)+\mathcal{O}(p^2)$$

and it follows the claim. □

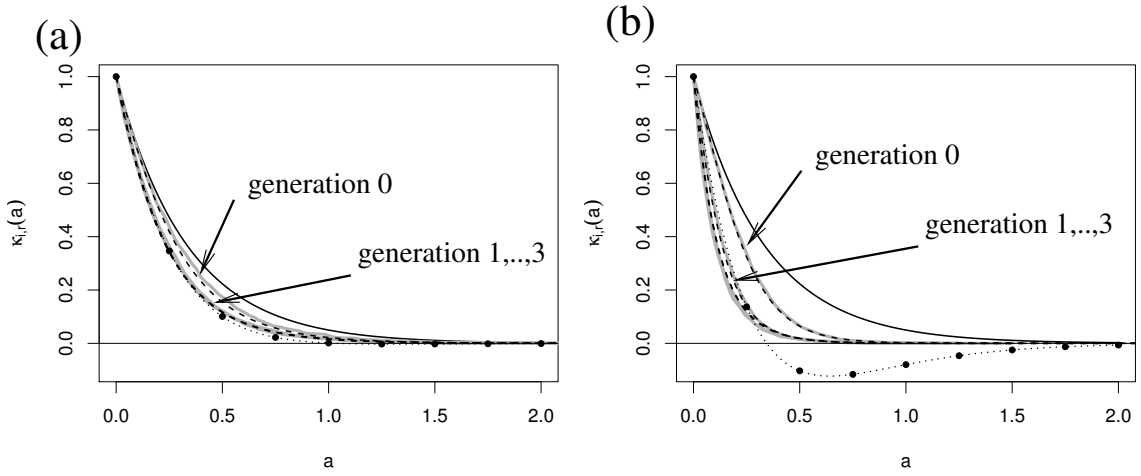


Figure 3.8: Recursive full tracing. The probability to be infectious at age a after infection ($\kappa_{r,i}(a)$) for $p = 0.3$ (panel a), and $p = 0.8$ (panel b). Solid line: without tracing ($\hat{\kappa}(a) = e^{-(\sigma+\alpha)a}$); dashed lines: numerical solution of $\kappa_{r,i}(a)$; dashed-dotted lines with bullets: first order approximation of $\kappa_{r,i}(a)$; grey line: simulation. Parameters: $\beta = 1.5$, $\alpha = 0.1$, $\sigma = 2.9$, $\mathbb{E}[K] = 4$, K is constant.

3.5.2 One step mode

Proposition 3.5.3. *Under one step full tracing, we have* $\kappa_{o,0}(a) = \kappa_o^-(a)$, *and for* $i > 0$

$$\kappa_{o,i}(a) = \frac{\int_0^\infty \kappa_o^-(a)\left(\kappa_{o,i-1}(b) - p\sigma\int_0^a \kappa_{o,i-1}(b+c)dc\right)db}{\int_0^\infty \kappa_{o,i-1}(b)db} \tag{3.5.3}$$

where $\kappa_o^-(a)$ *is given by eqn. 3.2.6*

Proof. The proof parallels Proposition 3.5.1 if we replace the removal term with σ. For the first order approximation, we have the proof identical to Proposition 3.5.2

□

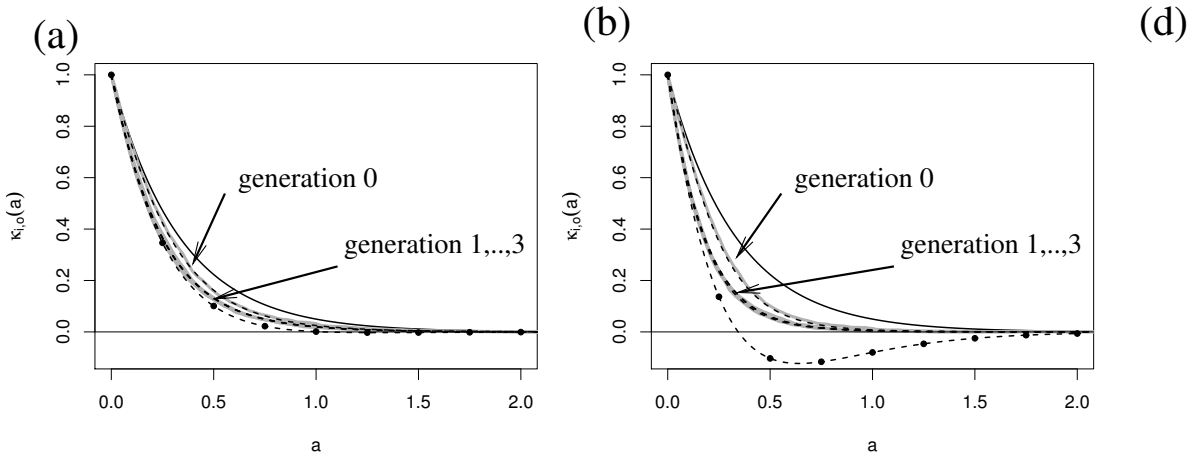


Figure 3.9: One-step full tracing. The probability to be infectious at age a after infection ($\kappa_{o,i}(a)$) for $p = 0.3$ (panel a), and $p = 0.8$ (panel b). Solid line: without tracing ($\hat{\kappa}(a) = e^{-(\sigma+\alpha)a}$); dashed lines: numerical solution of $\kappa_{o,i}(a)$; dashed-lines with bullets: first order approximation of $\kappa_{o,i}(a)$; grey line: simulation. Parameters: $\beta = 1.5$, $\alpha = 0.1$, $\sigma = 2.9$, $\mathbb{E}[K] = 4$, K is constant.

3.6 Reproduction number

In the following, an approximation of the reproduction number R_0 will be determined. We have here only a marginal reproduction number based on heuristics in that no threshold theorem is available right now.

Without contact tracing ($p = 0$), the basic reproduction number is given by

$$R_0 = \int_0^\infty \beta \left(\sum_{k=0}^{\infty} kP(K = k) \right) \hat{\kappa}(a)da = \int_0^\infty \beta \mathbb{E}[K]\hat{\kappa}(a)da = \frac{\beta \mathbb{E}[K]}{\alpha + \sigma}. \quad (3.6.1)$$

This reproduction number is decreased by contact tracing; we denote the reproduction number with contact tracing by R_{ct}.

Theorem 3.6.1. *Under full tracing (either recursive or one-step), the reproduction number of the ith generation reads*

$$R_{ct} = R_0 \left(1 - \frac{p\sigma\beta\mathbb{E}[K]}{2(\alpha + \sigma)(\alpha + \sigma + \beta)} - \frac{p\sigma}{2(\alpha + \sigma)} \right) + \mathscr{O}\left(p^2\right). \quad (3.6.2)$$

Proof. For $p > 0$, we use the first order approximation of full tracing $\kappa_{*,i}(a) = \kappa_i(a)$ (for * recursive or one step tracing) and find

$$\begin{aligned} R_{ct} &= \int_0^\infty \beta \left(\sum_{k=0}^{\infty} P(K=k)\kappa_i(a) \right) da, \\ &= \int_0^\infty \beta \mathbb{E}[K]\hat{\kappa}(a) \left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} \left(e^{-\beta a} - \hat{\kappa}(a) \right) - pp_{obs}\mathbb{E}[K](1-\hat{\kappa}(a)) - pp_{obs}(1-\hat{\kappa}(a)) \right) da \\ &\qquad + \mathcal{O}(p^2). \end{aligned} \tag{3.6.3}$$

Consider the following expressions in eqn. 3.6.3

$$\int_0^\infty \hat{\kappa}(a) \left(e^{-\beta a} - \hat{\kappa}(a) \right) da = \int_0^\infty e^{-(\alpha+\sigma+\beta)a} - e^{-2(\alpha+\sigma)a} = \frac{\alpha+\sigma-\beta}{2(\alpha+\sigma)(\alpha+\sigma+\beta)}$$

and

$$\int_0^\infty \hat{\kappa}(a)(1-\hat{\kappa}(a))da = \int_0^\infty e^{-(\alpha+\sigma)a} - e^{-2(\alpha+\sigma)a}da = \frac{1}{2(\alpha+\sigma)}.$$

By substituting the results of the above expressions in eqn. 3.6.3, the proof follows.

□

Remark 3.6.1. *We again consider the limit $\beta \to 0$, $\mathbb{E}[K] \to \infty$, while $\beta(1+\mathbb{E}[K]) \to \beta_{ind}$ of the reproduction number R_{ct} (Theorem 3.6.1).*

$$\begin{aligned} &\lim_{\beta\to 0} \frac{p\sigma\beta\mathbb{E}[K]}{2(\alpha+\sigma)(\alpha+\sigma+\beta)}\bigg|_{\beta_{ind}=\beta\mathbb{E}[K]} = \lim_{\beta\to 0} \frac{p\sigma\beta_{ind}}{2(\alpha+\sigma)} \left(\frac{1}{\alpha+\sigma+\beta} \right) \\ &= \frac{p\sigma\beta_{ind}}{2(\alpha+\sigma)} \frac{d}{d\beta} \left(\frac{1}{\alpha+\sigma+\beta} \right)\bigg|_{\beta=0} = \frac{-p\sigma\beta_{ind}}{2(\alpha+\sigma)}. \end{aligned}$$

With $R_0 = \beta_{ind}/(\sigma+\alpha)$ and $p_{obs} = \sigma/(\alpha+\sigma)$, we obtain

$$R_{ct} \to R_0 \left(1 - \frac{1}{2}\, p\, p_{obs}\, (R_0+1) \right) + \mathcal{O}(p^2), \tag{3.6.4}$$

which is identical with the first order approximation of the reproduciton number obtained in the randomly mixing model [62, 60].

3.7 Influence of the Graph structure

The structure of the random tree given by the generating function of the downstream degree distribution $G(s)$, only appeared in the consideration of backward tracing (see Theorem 3.2.1 and Theorem 3.2.2). As full tracing is the combination of forward- and on backward tracing, also there, the graph structure comes in. However, all first order approximations depend only on $\mathbb{E}[K]$. If we expand $\kappa_*^-(a)$ w.r.t. p, a closer look shows that the i'th order correction term (proportional to p^i) depends on the first i moments of K. As in applications, p is rather small, we expect indeed the special choice of the random tree only has a minor influence, if all moments of K are finite.

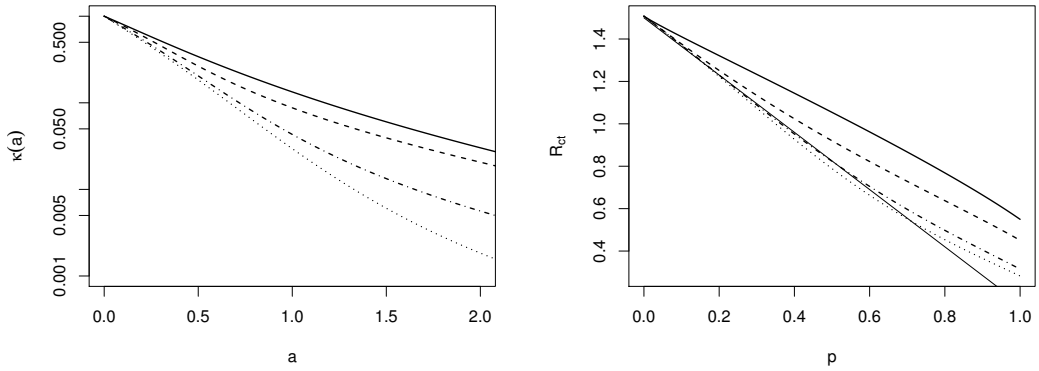


Figure 3.10: Influence of the distribution of K. Left panel: $\kappa_{r,5}(a)$ (in semi-logarithmic representation), right panel: R_{ct} over p. Curves, from the top down: solid, K with large second moment (see text, variance 4064); dashes, geometric distribution (variance 23.11); dashed-dotted, Poisson distribution (variance 4.3); dotted, K = constant (variance 0). Additional (thin, solid) line in the right panel: First order approximation of R_0. Parameters: $\mathbb{E}[K] = 4.3$, $\beta = 1.5$, $\alpha = \sigma = 0.5$.

A large or even infinite second moment, however, will lead to different results. To exemplify this conjecture, we consider four different degree distributions for the random variable K: (a) K is constant, (b) K follows a Poisson distribution, (c) K follows a geometric distribution, and (d) K follows a truncated power law, $P(K = i) = c\, i^{-\gamma}$ for $i \in \{1, \ldots, 15000\}$, and $P(K = i) = 0$ else, where we choose $\gamma = 2.1$. Clearly if we do not restrict ourselves to a $K \leq 15000$, but extend the distribution for all positive integers, the expectation is finite, while the variance is not. For numerical reasons, we exclude in this last case the large numbers of K and truncate the power law.

In the simulations, we adapt the parameters such that the expectations of the four distributions are identical, while the variance is increasing from (a) to (d) – see caption of Fig. 3.10 for the exact numbers. We find indeed that the effect of contact tracing decreases with the variance of K (Fig. 3.10). Moreover, the first order approximation for R_0 is only

appropriate if the variance is not too large; even for the geometric distribution, there is some reasonable deviation for $p > 0.3$, while for the distribution (d), even small values of p show a different behaviour. Nevertheless, at $p = 0$, approximation and exact result are tangential to each other, as predicted by the theory. Our finding bears some similarity with the observation that the reproduction number monotonously depends on the variance of the inter-generation time [77]. In that case, the larger the variance the smaller the reproduction number.

3.7.1 Analysis of superspreaders

Using the theory developed for $\kappa_{*,i}(a)$, the analysis of the effect of superspreaders on contact tracing is also interesting in the theory of random trees [59]. A superspreader is an infected individual who infects a very high number of secondary cases at an exceedingly rate more than normal. In large networks, superspreaders are dubbed "hubs" because they form a major component of the network. Superspreaders plays an important role on the reproduction number as the graph structure is influenced by the downstream degree distribution K.

An individual who is not a superspreader infects on average R_0 uninfected individuals. In an infection events where superspreaders are present, this R_0 is increased. There, the second moment (variance) in the second order approximation of $\kappa_{*,i}(a)$ also comes into play. Before investigating the effect of the superspreaders on $\kappa_{*,i}(a)$, the long term behaviour of the probability should be clear.

Proposition 3.7.1. *$\kappa_{*,i}(a)$ at the age of infection a converges to 0 for $a \to \infty$.*

Remark 3.7.1. *The existence of superspreaders increase the individual variation as shown in the publication of Lloyd-Smith et al. [51], thus* $\mathrm{Var}[K]$ *for the random variable K which denotes the number of downstream edges.*

In order to understand the effect of superspreaders on the probability to be infectious, we observe the coefficient of the variance in the second order approximation of $\kappa_{r,i}(a)$. Since the probability to be infectious converges to 0, the coefficient of $\mathrm{Var}[K]$ in the second order approximation may only have an influence on the speed of convergence.

Proposition 3.7.2. *The second order approximation under recursive step full tracing*

$\kappa_{r,i}(a)$ *reads*

$$\begin{aligned}
\kappa_{r,i}(a) \quad = \quad & \hat{\kappa}(a)\left(1+\frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma-\beta}\left(e^{-\beta a}-\hat{\kappa}(a)\right)-\frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma}(1-\hat{\kappa}(a))-\frac{p\sigma}{(\alpha+\sigma)}(1-\hat{\kappa}(a))\right.\\
& \frac{p^2}{2}\left(p_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))\left(\frac{-p_{obs}\mathbb{E}[K](\alpha+\sigma-\beta)-\sigma\mathbb{E}[K]-p_{obs}(\alpha+\sigma-\beta)}{2(\alpha+\sigma-\beta)(\alpha+\sigma)}+\frac{\sigma\mathbb{E}[K]}{(\alpha+\sigma-\beta)(\alpha+\sigma+\beta)}\right)\right.\\
& +p_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a)\left(\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\sigma+\alpha-\beta))}\right)\right.\\
& -\hat{\kappa}(a)(\alpha+\sigma)\left(\left(\frac{(1-\hat{\kappa}(a)e^{\beta a})}{(\alpha+\sigma-\beta)^2}\right)\left(\mathbb{E}[K]\sigma-\frac{\sigma\alpha\mathbb{E}[K]}{\alpha+\sigma-\beta}\right)\right.\\
& +\left(\frac{(1-\hat{\kappa}(a)^2)}{4(\alpha+\sigma)^2}\right)\left(\frac{\alpha\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}+\mathbb{E}[K]\sigma\left(1+p_{obs}\right)+p_{obs}(2\alpha+2\sigma-1)\right)\\
& \left.+\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2}\right)(-p_{obs}(\sigma\mathbb{E}[K]+\alpha+\sigma+1))\right)\\
& \left.+\frac{\hat{\kappa}(a)}{2}(\mathrm{Var}[K]+\mathbb{E}[K](\mathbb{E}(K)-1))\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2+\mathcal{O}\left(p^3\right)\right).
\end{aligned}$$

Proof. See Appendix A.2.5 □

For the second order approximation of $\kappa_{r,i}(a)$. Let $v(a), a \geq 0$ be the coefficient of the variance:

$$v(a)=\hat{\kappa}(a)\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2$$

The function is continuous on the domain $a \in [0,\infty)$ for $\alpha+\sigma \neq \beta$. Since $\hat{\kappa}(a) \in (0,1]$ and the result in the brackect is squared, the function is always positive or equal zero. This means that the variance increases the probability $\kappa_{r,i}(a)$, if $\mathbb{E}[K] > 1$, which is the case for superspreaders. Superspreaders increase the variance term of the degree distribution and the coefficient of the variance is positive in the second order approximation of $\kappa_{r,i}(a)$.

From this reason, we expect that the convergence in the second order approximation is rather slower than the convergence in first order approximation and this fact may decrease the efficiency of contact tracing as already shown by simulation in Fig. 3.10. Other than the variance, there are also other variables that influence the probability $\kappa_{r,i}(a)$.

3.8 Mean field

So far, we have mainly focused on single individuals. The fate of an individual is well characterized by the infectious interval, described by $\kappa(a)$. Now we turn to the population.

Preliminary studies ([62, 60]) did show that in homogeneous populations, the local correlations caused by contact tracing (removal of several neighbouring infected individuals at the same time) are less important than the reduction of the infectious period. We also test this observation in the present setting. The situation we did analyse is a tree, where initially the root is infected and infectious.

We develop an idea about the dynamics of the prevalence in the next section. In the section thereafter, we extend our ideas by means of heuristic arguments to the configuration model.

3.8.1 Mean field without contact tracing on a tree

The dynamics of an SIR model without contact tracing on a tree is well known [17, 39]. We briefly sketch the idea of the analysis. For the initial state, only the root is infected. β is the contact rate per edge, γ recovery rate and K is random variable for the nnumber of downstream edges (for all individuals, including the root). That is, the root has K_0 edges (all are downstream), all other nodes have $K_i + 1$ edges (K_i downstream, one upstream). K_0, K_i are i.i.d.

Let I_t and R_t denote the number of infected and recovered nodes at time t respectively.

Clearly,

$$\frac{d}{dt}E(R_t) = \gamma E(I_t).$$

It remains to focus on I.

Therefore we focus on an infected node. Let $z_k(a)$ denote the expected number of secondary infecteds if the age of infection of the focal individual is a (where we do assume that the focal individual is still infectious). Clearly, $z_k(0) = 0$, $z_k(a) \leq k$. We have

$$\frac{d}{da}z_k(a) = \beta\,(k - z_k(a)), \qquad z_k(0) = 0.$$

Hence,

$$\frac{d}{da}\ln(k - z_k(a)) = \beta$$

and (using $z_k(0) = 0$)

$$z_k(a) = k\left(1 - e^{-\beta a}\right).$$

If we want to know the number of children of a randomly selected individuals that became infected a time units ago, we need to remove the information about K,

$$z(a) = \sum_{k=0}^{\infty} k\left(1 - e^{-\beta a}\right)\, P(K = k) = \mathbb{E}[K]\,\left(1 - e^{-\beta a}\right).$$

Hence the rate $\theta(a)$ at which a randomly chosen infected individual produces downstream infecteds is given by

$$\theta(a) = \frac{d}{da} z(a) = \mathbb{E}[K]\, \beta\, e^{-\beta a}.$$

The last term, $\beta e^{-\beta a}$ refers to the time distribution of the first infectious event, while $\mathbb{E}[K]$ is the average number of nodes that can be infected.

Using these considerations, we are able to set up an age-structured model for the density of infected individuals,

$$\begin{aligned} (\partial_t + \partial_a) i(t,a) &= -\gamma i(t,a) && (3.8.1) \\ i(t,0) &= \int_0^\infty \theta(a)\, i(t,a)\, da. && (3.8.2) \end{aligned}$$

As usual, this age structured model will tend to an exponential growing solution with a stable age structure,

$$i(t,a) = I_0\, e^{\lambda t}\, i(a), \qquad i(0) = 1.$$

Therewith,

$$i'(a) = -(\gamma - \lambda) i(a) \quad \Rightarrow \quad i(a) = e^{-(\gamma+\lambda)\, a}.$$

We have exponent λ, where

$$1 = \int_0^\infty \theta(a)\, e^{-(\lambda+\gamma)a}\, da = \mathbb{E}[K]\, \beta \int_0^\infty e^{-(\lambda+\gamma+\beta)a}\, da = \frac{\beta \mathbb{E}[K]}{\lambda+\gamma+\beta}.$$

Solving this expression for λ, we have

$$\lambda = \beta(\mathbb{E}[K] - 1) - \gamma.$$

Hence, the mean field reads

$$\begin{aligned} \frac{d}{dt} I &= \beta(\mathbb{E}[K] - 1)\, I - \gamma I && (3.8.3) \\ \frac{d}{dt} R &= \gamma I && (3.8.4) \end{aligned}$$

and the limiting age stucture reads

$$i(t,a) = I(t)\, (\gamma + \lambda)\, e^{-(\lambda+\gamma)\, a}.$$

Interpretation of $\beta(\mathbb{E}[K] - 1)$: We do not have $\lambda = \beta\mathbb{E}[K] - \gamma$ (which would be the case if any infected individual would have only susceptible downstream nodes) but $\lambda = \beta(\mathbb{E}[K]-1)-\gamma$ to take into account that some downstream nodes are already infected.

3.8.1.1 The number of missed nodes

A further interesting subpopulation is the number of missed downstream neighbours. If the upstream neighbour of a susceptible is recovered, that individual escaped the infection. Let $\Sigma(t)$ be the expected number of susceptible downstream neighbours of recovered individuals (nodes that the infected did miss).
We find

$$\begin{aligned}
(\partial_t + \partial_a) i(t,a) &= -\gamma i(t,a) && (3.8.5)\\
i(t,0) &= \int_0^\infty \theta(a)\, i(t,a)\, da && (3.8.6)\\
\frac{d}{dt} R &= \gamma \int_0^\infty i(t,a)\, da && (3.8.7)\\
\frac{d}{dt}\Sigma &= \gamma \int_0^\infty i(t,a)\,(\mathbb{E}[K] - z(a))\, da = \mathbb{E}[K]\,\gamma \int_0^\infty i(t,a)\, e^{-\beta a}\, da && (3.8.8)
\end{aligned}$$

In the long run,

$$i(t,a) = I(t)\,(\gamma+\lambda)\, e^{-(\lambda+\gamma)\,a}$$

and hence we have the following result

$$\begin{aligned}
\frac{d}{dt}\Sigma &= \mathbb{E}[K]\,\gamma I \int_0^\infty (\gamma+\lambda)\, e^{-(\lambda+\gamma+\beta)a}\, da\\
&= \mathbb{E}[K]\,\frac{\gamma+\lambda}{\lambda+\gamma+\beta}\,\gamma\, I = \mathbb{E}[K]\,\frac{\beta(\mathbb{E}[K]-1)}{\beta\mathbb{E}[K]}\,\gamma\, I = (\mathbb{E}[K]-1)\,\gamma\, I
\end{aligned}$$

Theorem 3.8.1.

$$\begin{aligned}
\frac{d}{dt} I &= \beta(\mathbb{E}[K]-1)\, I - \gamma I && (3.8.9)\\
\frac{d}{dt} R &= \gamma I && (3.8.10)\\
\frac{d}{dt}\Sigma &= (\mathbb{E}[K]-1)\,\gamma\, I && (3.8.11)
\end{aligned}$$

Proposition 3.8.1. *In the long run, the expected number of susceptible downstream nodes is given by*

$$\mathbb{E}[K] - 1. \tag{3.8.12}$$

Proof. The age structure of the infected individuals (in the long run) is given by

$$i(a) = (\gamma+\lambda)\, e^{-(\lambda+\gamma)a}.$$

The average number of susceptible downstream nodes of an individual of age a reads

$$z(a) = \mathbb{E}[K](1 - e^{-\beta a}).$$

Hence, the average no of infected downstream nodes is given by

$$\begin{aligned}\int_0^\infty z(a)\, i(a)\, da &= \mathbb{E}[K](\lambda+\gamma)\int_0^\infty (1-e^{-\beta a})e^{-(\lambda+\gamma)a}\, da = \mathbb{E}[K](\lambda+\gamma)\left(\frac{1}{\lambda+\gamma} - \frac{1}{\lambda+\gamma+\beta}\right) \\ &= \mathbb{E}[K]\frac{\beta(\lambda+\gamma)}{(\lambda+\gamma)(\lambda+\beta+\gamma)} = \mathbb{E}[K]\frac{\beta(\lambda+\gamma)}{(\lambda+\gamma)(\beta\mathbb{E}[K])} = 1.\end{aligned}$$

Hence, the average no of susceptible nodes is $\mathbb{E}[K] - 1$, independently on β and γ (only assuming that $E(I) \to \infty$). □

□

3.8.1.2 Asymptotic ratios

Determine

$$\lim_{t\to\infty} \frac{\Sigma(t)}{I(t)}, \qquad \lim_{t\to\infty} \frac{R(t)}{I(t)}.$$

With $\zeta(t) = \Sigma(t)/I(t)$ and $\eta(t) = R(t)/I(t)$, we find

$$\begin{aligned}\zeta'(t) &= \frac{\Sigma'(t)}{I(t)} - \frac{\Sigma(t)}{I(t)}\frac{I'(t)}{I(t)} = (\mathbb{E}[K]-1)\gamma - \lambda\,\zeta(t) \\ \eta'(t) &= \frac{R'(t)}{I(t)} - \frac{R(t)}{I(t)}\frac{I'(t)}{I(t)} = \gamma - \lambda\,\eta(t)\end{aligned}$$

and hence

Corollary 3.8.1.

$$\lim_{t\to\infty} \frac{\Sigma(t)}{I(t)} = (\mathbb{E}[K]-1)\frac{\gamma}{\lambda} = (\mathbb{E}[K]-1)\frac{\gamma}{\beta(\mathbb{E}[K]-1)-\gamma}, \qquad \lim_{t\to\infty} \frac{R(t)}{I(t)} = \frac{\gamma}{\beta(\mathbb{E}[K]-1)-\gamma}.$$

3.8.2 Mean field with contact tracing on a tree

In order to incorporate contact tracing, we replace in eqn. (3.8.5)-(3.8.6), the constant removal rate γ by the age structured non-constant hazard rate we did determine. That is, we replace γ by

$$\frac{-\kappa'(a)}{\kappa(a)} = -\log(\kappa(a))'$$

where $\kappa(a)$ is the probability to be infectious at age of infection a (in case of forward/full contact tracing, the asymptotic probability, if we take the number of generations to ∞). Using the first order approximation of $\kappa(a)$ in eqn. 3.5.1, that is

$$\kappa_{r,i}(a) = \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right) - \frac{p\sigma}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right) + \mathcal{O}(p^2), \tag{3.8.13}$$

the age-structured model now becomes

$$\begin{aligned}(\partial_t + \partial_a) i(t,a) &= \log(\kappa(a))' i(t,a) && (3.8.14)\\ i(t,0) &= \int_0^\infty \theta(a)\, i(t,a)\, da. && (3.8.15)\end{aligned}$$

Asymptotically, $i(t,a) = I_0\, e^{\lambda t}\, i(a)$, where

$$i'(a) = (\log(\kappa(a))' + \lambda)\, i(a).$$

Thus,

$$i(a) = i(0)\, \kappa(a)\, e^{-\lambda a}$$

and the equation that determines λ becomes

$$1 = \int_0^\infty \theta(a)\, \kappa(a)\, e^{-\lambda a}\, da. \tag{3.8.16}$$

Recall that $\theta(a) = \mathbb{E}[K]\, \beta\, e^{-\beta a}$. If we plug in the first rder approximation for κ, we are able to solve the integral

$$\begin{aligned}
1 &= \int_0^\infty \theta(a)\, \kappa(a)\, e^{-\lambda a}\, da \\
&= \mathbb{E}[K]\, \beta \int_0^\infty \left[\hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right.\right. \\
&\qquad \left.\left. - \frac{p\sigma}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\right)\right], e^{-(\beta+\lambda)a}\, da + \mathcal{O}(p^2) \\
&= \mathbb{E}[K]\, \beta \int_0^\infty \left[\hat{\kappa}(a)\left(\left(1 - \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma} - \frac{p\sigma}{\alpha+\sigma}\right) + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\, e^{-\beta a}\right.\right. \\
&\qquad \left.\left. + \left(-\frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma} + \frac{p\sigma}{\alpha+\sigma}\right)\hat{\kappa}(a)\right)\right], e^{-(\beta+\lambda)a}\, da + \mathcal{O}(p^2) \\
&= \mathbb{E}[K]\, \beta \int_0^\infty \left[\left(1 - \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right) + \frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\, e^{-\beta a}\right. \\
&\qquad \left. + \left(-\frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} + \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right)\hat{\kappa}(a)\right] e^{-(\alpha+\sigma+\beta+\lambda)a}\, da + \mathcal{O}(p^2) \\
&= \frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+\beta+\lambda}\left(1 - \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right) + \frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+2\,\beta+\lambda}\left(\frac{p\sigma E\mathbb{E}K]}{\alpha+\sigma-\beta}\right) \\
&\quad + \frac{\mathbb{E}[K]\,\beta}{2\alpha+2\sigma+\beta+\lambda}\left(-\frac{p\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} + \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right).
\end{aligned}$$

We expand λ in terms of p,

$$\lambda = \lambda(p) = \lambda_0 + p\,\lambda_1 + \mathcal{O}(p^2). \tag{3.8.17}$$

Hence, as before, (note that for $p = 0$ we have $\gamma = \alpha + \sigma$)

$$\lambda_0 = \beta(\mathbb{E}[K] - 1) - \alpha - \sigma.$$

It is

$$\frac{1}{A + p\lambda_1} = \frac{1}{A} - p\,\lambda_1\,\frac{1}{A^2} + \mathcal{O}(p^2).$$

Therewith, the first order perturbation λ_1 satisfies

$$\begin{aligned} 0 \;=\; & -\frac{\mathbb{E}[K]\,\beta}{(\alpha+\sigma+\beta+\lambda_0)^2}\,\lambda_1 - \frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+\beta+\lambda_0}\left(\frac{\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right) + \frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+2\,\beta+\lambda_0}\left(\frac{\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\right) \\ & + \frac{\mathbb{E}[K]\,\beta}{2\alpha+2\sigma+\beta+\lambda_0}\left(-\frac{\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} + \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right). \end{aligned}$$

That is,

$$\begin{aligned} \lambda_1 \;=\; & \frac{(\alpha+\sigma+\beta+\lambda_0)^2}{\mathbb{E}[K]\,\beta}\;\left\{ -\frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+\beta+\lambda_0}\left(\frac{\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right) + \frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+2\,\beta+\lambda_0}\left(\frac{\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\right)\right. \\ & \left. + \frac{\mathbb{E}[K]\,\beta}{2\alpha+2\sigma+\beta+\lambda_0}\left(-\frac{\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} - \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right)\right\} \\ =\; & \frac{(\alpha+\sigma+\beta+\lambda_0)^2}{\mathbb{E}[K]\,\beta}\;\left\{ -\frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+\beta+\lambda_0}\left(\frac{\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right) + \frac{\mathbb{E}[K]\,\beta}{\alpha+\sigma+2\,\beta+\lambda_0}\left(\frac{\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}\right)\right. \\ & \left. + \frac{\mathbb{E}[K]\,\beta}{2\alpha+2\sigma+\beta+\lambda_0}\left(-\frac{\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} - \frac{p\sigma(\mathbb{E}[K]+1)}{\alpha+\sigma}\right)\right\}. \end{aligned}$$

By substituting λ_0 and λ_1 in the equation 3.8.17, we solve the integral expression for λ.

Furthermore, a short computation of the asymptotic ratios which incorporates contact tracing yields

Corollary 3.8.2.

$$\lim_{t\to\infty}\frac{\Sigma(t)}{I(t)} = \frac{\mathbb{E}[K]\left(1-(\lambda+\beta)\int_0^\infty \kappa(a)e^{-(\lambda+\beta)a}\,da\right)}{\lambda\int_0^\infty \kappa(a)e^{-\lambda a}\,da}, \qquad \lim_{t\to\infty}\frac{R(t)}{I(t)} = \frac{1}{\lambda\int_0^\infty \kappa(a)e^{-\lambda a}\,da} - 1.$$

For different tracing probabilities, the effect of contact tracing on the exponent and asymptotic ratios is established, and with the age structured non-constant hazard rate, we find a satisfying agreement with simulation results (Fig. 3.11). It is not surprising that for increased p, the prediction for the non-age structured removal rate ($\lambda = \beta(\mathbb{E}[K]-1) - 1/\int_0^\infty \kappa(a)da$) breaks down, this is because we have used a first-order approximation of $\kappa(a)$ to compute λ.

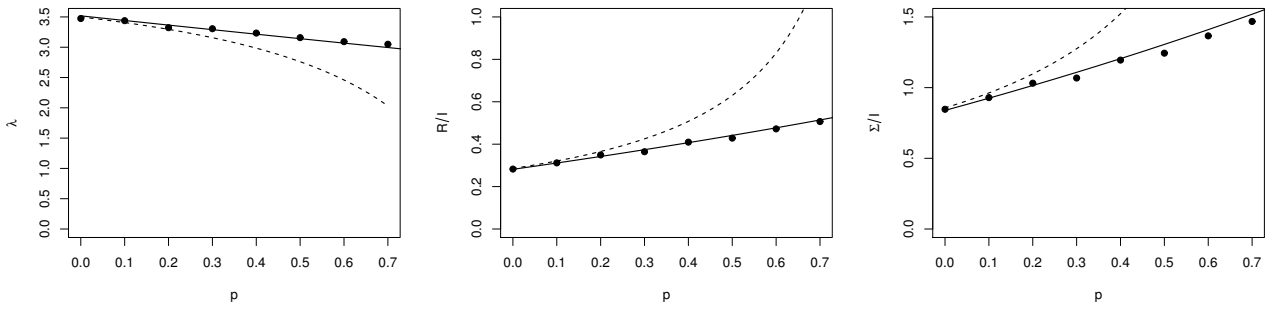


Figure 3.11: λ (left panel), $\lim_{t\to\infty} \frac{R(t)}{I(t)}$ (middle panel) and $\lim_{t\to\infty} \frac{\Sigma(t)}{I(t)}$ (right panel) with different values of p. The bullets are the average simulating results in 50 runs, solid line is the predictions with the age structured non-constant hazard rate ($log(\kappa(a))'$) while the dashed line is the prediction with the non-age structured removal rate ($1/\int_0^\infty \kappa(a)da$).

3.8.3 Mean field on the configuration model

Obviously, trees are rather mathematically convenient but not appropriate models for natural contact graphs. The configuration model (CM) is better suited [53]. In the CM, the number of nodes N and the degree distribution of nodes are given. The CM is constructed in generating for each node a random number of stubs, according to the degree distribution. Afterwards, subs are paired randomly to form edges (See Fig. 3.2).

Newman and co-workers developed a macroscopic, approximate description of an SIR dynamics on a large configuration model, the message passing approach [37, 78]. Interestingly, this approach resembles the idea used in the present work to obtain $\kappa(a)$, and is exact on trees. In the centre of that theory is the probability $H(t)$ that the infection has not been passed to a focal individual by a given edge at time t. Let us assume that no node is recovered at time $t = 0$, and that nodes are independently and randomly assigned to be either susceptible (probability z) or infected (probability $1 - z$).

Furthermore, β is the rate of contacts on a given edge, and $r(a)\,da$ is the probability for removal of an infected individual at age a. As we lose the graph structure, we cannot speak of upstream- and downstream edges. Let $\hat{K}$ denote the degree distribution (all edges) of a node. Furthermore, $G(s)$ denotes the generating function of that degree distribution, and $G_1(s) = G'(s)/G'(1)$ is the degree distribution of a randomly chosen neighbour of a randomly chosen individual [65]. Then [37]

$$H(t) = 1 - \int_0^t \beta e^{-\beta a} \int_a^\infty r(\tau)\,d\tau\,(1 - zG_1(H(t-a)))\,da. \tag{3.8.18}$$

The probabilities $P_S(t)$, ($P_I(t)$, and $P_R(t)$) for a randomly chosen individual to be suscep-

tible (infected, recovered) at time t are given by [37]

$$\begin{aligned} P_S(t) &= zG(H(t)) && (3.8.19)\\ \frac{d}{dt}P_I(t) &= -\frac{d}{dt}P_S(t) - (1-z)r(t) + \int_0^t r(t-t')\frac{d}{dt'}P_S(t')\,dt',\ P_I(0) = 1-z && (3.8.20)\\ P_R(t) &= 1 - P_S(t) - P_I(t). && (3.8.21) \end{aligned}$$

For $p = 0$, we are in the setting of the message passing method, where

$$r(a) = (\alpha + \sigma)\, e^{-(\alpha+\sigma)a}.$$

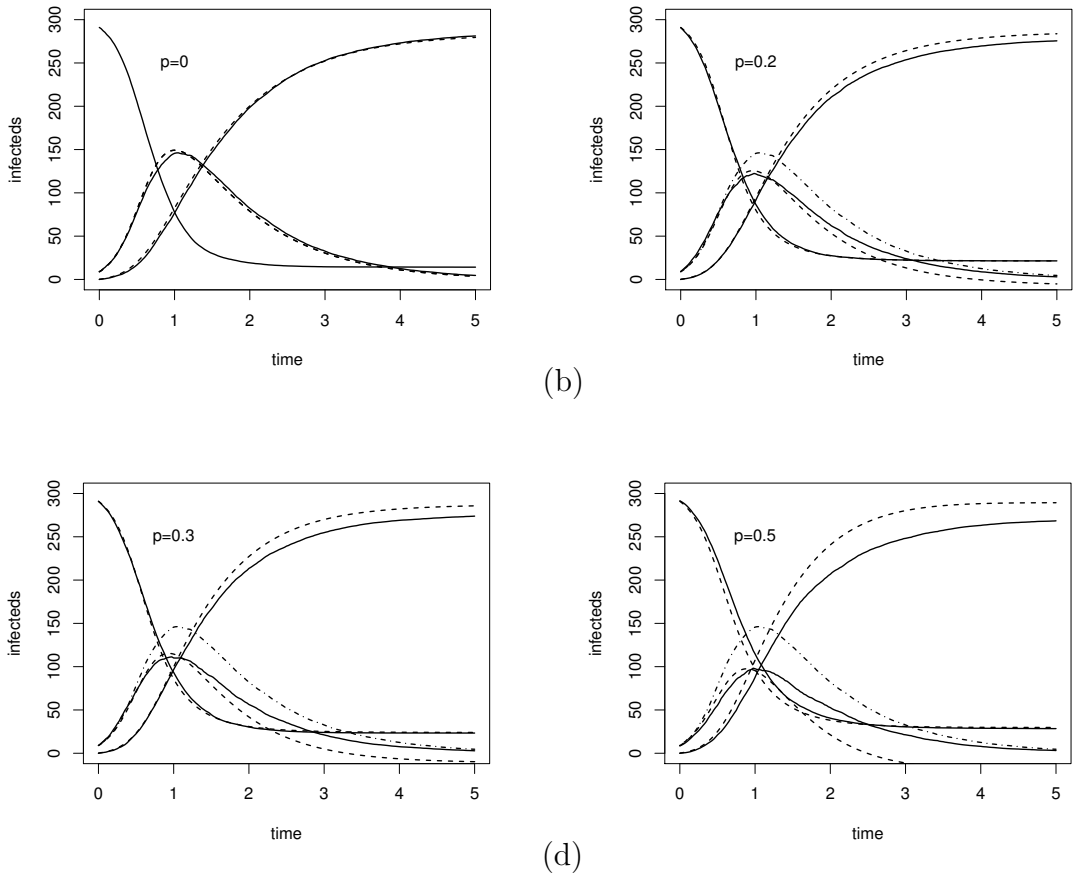
And indeed, Fig. 3.12 (a) shows an excellent agreement of the prediction given by that method and simulations.

We incorporate contact tracing into that model in replacing $r(a)$ by $-\kappa'(a)$, where we choose the first order approximation as developed in eqn. (3.5.1) for $\kappa(a)$. We focus here on one-step tracing (see discussion below). However, there is a fundamental difference between a tree where only the root is infected and the CM: It is possible that neighbouring nodes are infected and trigger tracing events, though the focal node is neither an infectee nor infector (see also [62]). We can estimate the rate to be traced by such a contact at age of infection a by multiplying the number of edges minus the edge of the infector $(E(\hat{K}) - 1)$ times the probability that no contact happened till the present age $e^{-\beta a}$ times the probability that the contactee is infected $(P_I(t))$, times the rate of direct detection σ times the tracing probability p. We obtain

$$\begin{aligned} \kappa(a, P_I) := \hat{\kappa}(a)\Bigg(1 &+ \frac{p\sigma(E(\hat{K})-1)}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma(E(\hat{K})-1)}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right)\\ &- \frac{p\sigma}{\alpha+\sigma}\left(1-\hat{\kappa}(a)\right) - p\left(E(\hat{K})-1\right)\sigma P_I(t)\int_0^a e^{-\beta a'}\,da'\Bigg). \end{aligned} \tag{3.8.22}$$

We assume that the infectious period of an individual is shorter than the time scale at which P_I does change. Therewith, we replace in eqn. (3.8.18) $r(\tau)$ by $-\frac{\partial}{\partial\tau}\kappa(\tau, P_I(t))$.

We find for small values of p a reasonable approximation (Fig. 3.12 (b), (c)); however for $p = 0.5$ and larger the residuals become notable (Fig. 3.12 (d)). As we only use a first-order approximation as the basis for our definition of $\kappa(a)$, we cannot expect a better result.
For recursive tracing, a reasonable approximation is more involving, as tracing can be triggered by individual in a distance of several steps (in the graph metric). The correlations between individuals, as well as the dependency on the background prevalence becomes more involving to control.



(a) (b)

(c) (d)

Figure 3.12: Mean field approximation on the CM model for full tracing and different values of p (as indicated in the graphs). The decreasing functions represent S, the increasing function R, and the unimodal function I. The solid lines are the average of 50 simulations, the dashed line the predictions of the message passing method for contact tracing (see text). The dashed-dotted line in panels (b)-(d) is the simulating result for I and $p = 0$. Parameters: $N = 300$, $\beta = 1.5$, $\sigma = \alpha = 0.5$, Poissonian edge-distribution with $\mathbb{E}[K] = 4$.

Chapter 4

Phylogenetic methods for infectious models

Infectious models have been instrumental in the study of many infectious diseases. Usually these models are dependent of several epidemiological parameters. However, most of these parameters are unknown and may cause these models to lack robustness [43] if not chosen appropriately. Missing data poses a major quantification challenge in epidemiology due to unobserved or partially observed events [50]. This makes parameter estimation very essential in modelling disease spread. Parameter estimation is the process of experimentally determining model parameters from data. Often, parameters are maximised following the model predictions on sets of parameter values. In other to achieve parameter estimation, the model system property must be identifiable [11, 12, 14]. Several estimation methods, e.g statistically based techniques such Approximate Bayesian computation (ABS) [10], Markov chain Monte Carlo (MCMC) integration [68], optimal control theory approach [32], classical least-squares method [2], etc have been instrumental in increasing the confidence level of parameter estimation and making inferences in epidemic models.

In this proposed paper, we aim at using phylogenetic tools to estimate parameters in infectious models. A phylogeny is a branching diagram or tree showing the evolutionary history of (groups of) species. Several models [74, 52, 28, 63] have proposed proobability models for studying phylogenies which is based on a reconstruction technique of the phylogenetic tree topology that can be completely or partially resolved. Nee et al. [63] in their publication, derived models for reconstructed phylogenies that do not capture extinct lineages. Maddison et al. [52] proposed a binary-state speciation and extinction (BISSE) model for estimating the effect of character state change on speciation and extinction. In this proposed model, species can be in one of a binary state (0 for e.g herbivory or 1 for e.g carnivory). A similar approach by Stadler and Bonhoeffer [74] relies on a multi-type branching process (MTBP) for heterogeneous population where each taxa (individual) is characterised by a transmission group.

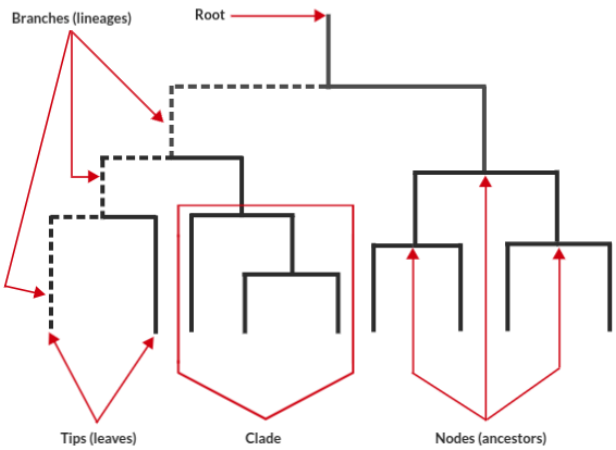


Figure 4.1: A phylogeny: parts of a phylogenetic tree.

These models assume a complete and fully resolved phylogenetic tree where all extant taxa are known. However, unresolved phylogenetic trees will not be completely represented by these likelihoods. FitzJohn et al. [28] proposed a generalised likelihood methods for the BISSE where it is possible to derive valid likelihoods for both resolved and unresolved phylogeny. These likelihoods methods generally assume a rooted phylogenetic tree where the probability calculations are estimated by moving backwards in time from the tips down to the root. Here, we adopt similar maximum-likelihood framework based on phylogenetic trees by [74]. We can extend or translate this multi-type birth-death branching model to a tree-based SIR model where each taxa (individual) is assigned to a degree distribution. We incorporate our tree model theory of contact tracing derived in previous chapter into the likelihood. As we expect, the phylogenetic methods is better suited for estimating the timing of the infectious process while the incorporation of contact tracing is better suited for understanding and estimating the underlying contact structure.

4.1 Model and analysis

In this section, we will forumulate the branching process for a stochastic SIR model based on phylogenetic trees. In this section, the model analysis do not include contact tracing. Later in this chapter, we will incorporate contact tracing. We start by stating the model assumption and parameters.

On a tree-shaped contact graph, each individual is characterised by its degree distribution. On the rooted phylogenetic tree, an individual can be in one of SIR disease state. We augment the state of this individual by assigning, at a given time point, a 2-tuple (i, k) character type which is of interest in the model analysis. k denotes the total number of downstream nodes (infected + uninfected) and i the number of already infected downstream nodes such that $k - i$ is the number of uninfected downstream nodes.

Note that the type of an individual will be $(0, k)$ at the time point of infection and i is increasing (non-decreasing) during the infection period. When the individual recovers, its state will be (i^*, k) with $0 \leq i^* \leq k$. On the phylogenetic tree, we assume the state of an observed individual is known and also each tip is assigned a known state.

The effective infection rate of an infected individual in state i is given as $(k - i)\,\beta$. Under this process, we assume constant infection and recovery rate for every individual,

$\mu + \sigma$: Recovery rate (direct observation or unobserved),

$p_{obs} = \frac{\sigma}{\mu+\sigma}$: Probability for observed recovery,

$1 - p_{obs}$: Probability for unobserved recovery.

Notation: In order to avoid confusion, we reduce the notation of the state of an individual from a 2-tuple (i, k) to a 1-tuple i. As the central analysis of the fate of a lineage focus on the number of already infected downstream edges, we use this notation i always with the assumption that it is coupled with k. In a more general term, the state of an individual can be arbitrary, that is an individual represented by $*$ has always state $(*, k)$. The asterisk $*$ can take any character.

Below, we will define some important terminologies associated with the phylogenetic methods. We will keep the definitions as simple as possible. See Fig. 4.1 for references of the definitions below.

Phylogenetic tree: A phylogenetic tree (also known as a phylogeny) is a branching diagram or a tree that gives a visual representation of the relationship between different organisms, showing the path through evolutionary time from a common ancestor to different descendants.

Lineage: (*backward in time*) A lineage is the sequence of ancestors of a given individual.

Clade: (*forward in time*) A clade is a group of biological taxa (such as species) that includes all descendants of one common ancestor.

4.1.1 Formulating probabilities along a branch

On a rooted phylogenetic tree, we investigate the probability density of a sampled tree given some epidemiological parameters. We start with one individual of type i somewhere in the process and the descendants of this individual form a clade. For a randomly selected type i individual of an observed clade, we define $D^i_j(t)$ to be the probability density for observing a type j individual in the clade at time t. That is, the likelihood that if we start with a type i individual in the past, we will observe a type j individual in the present for the clade to survive. To be able to calculate $D^i_j(t)$, we also have to calculate independently the probability $E_i(t)$ for the lineage i and all its descendants to go undetected before the present time. We define the present time to be $t = 0$ and with time $t + \Delta t$ increasing into the past (i.e. we measure time backwards), we derive probabilities for all possible events.

In the next session, we calculate the probability $E_i(t)$, then also derive $D_j^i(t)$ along the branch represented by individual i.

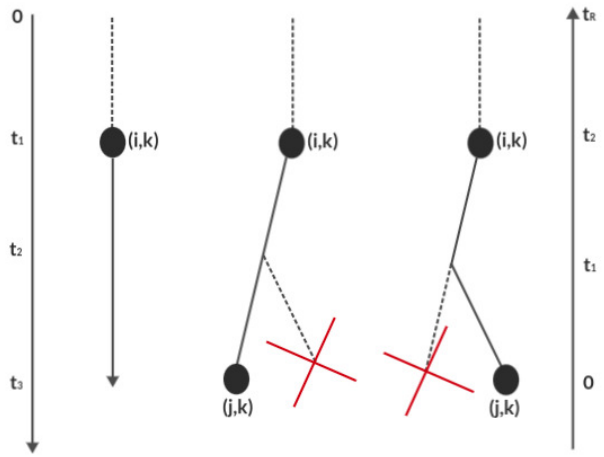


Figure 4.2: Possible events for $D_j^i(t)$. Left: undetected recovery. Middle: No birth, no death and therefore birth of a new lineage (right) which goes undetected before the present time. Right: Birth of a new lineage (left) which goes undetected before the present time

4.1.1.1 Probability to stay undetected, $E_i(t)$

To be able to solve the differential $D_j^i(t)$, we must determine the probability $E_i(t)$ for a type i individual and all descendants altogether stays undetected for a time interval t. An individual at time of infection $t = 0$ is not sampled, i,e. we have

$$E_i(0) = 1 \tag{4.1.1}$$

For one small time step (Δt) further down the branch, we calculate the corresponding probabilities for $E_i(t + \Delta t)$.

(i) The probability $E_k(t)$ is straight to forward compute. As there is no possibility for further infection before the present time, we focus on whether this individual recovers detected or stays forever undetected in $[0, t]$.

let $q(t)$ denote the probability to be alive, undetected and $r(t)$, the probability to be recovered, undetected:

$$\begin{aligned} \frac{d}{dt}q(t) &= -(\mu + \sigma)\, q(t) \\ \frac{d}{dt}r(t) &= \mu\, q(t) \end{aligned}$$

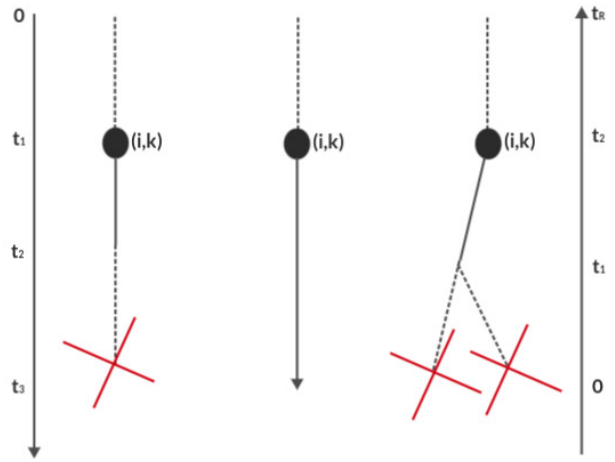


Figure 4.3: Possible events for $E_i(t)$. Left: No birth, lineage recovers undetected. Middle: No birth, no death and lineage remains undetected. Right: Birth of a new lineage and both lineages goes undetected

With $q(0) = 1$, $r(0) = 0$, we have

$$q(t) = e^{-(\mu+\sigma)t}, \qquad r(t) = \int_0^t \mu\, e^{-(\mu+\sigma)t}\, dt = \frac{\mu}{\mu+\sigma}(1 - e^{-(\mu+\sigma)t})$$

that is

$$E_k(t) \;=\; e^{-(\mu+\sigma)(t)} + \frac{\mu}{\mu+\sigma}(1 - e^{-(\mu+\sigma)t}) = \frac{\mu}{\mu+\sigma} + \frac{\mu}{\mu+\sigma} e^{-(\mu+\sigma)(t)}$$

Hence we have

$$\frac{d}{dt}E_k(t) \;=\; -\sigma\, e^{-(\mu+\sigma)t} = \mu - (\sigma+\mu)\, E_k(t), \qquad E_k(0) = 1 \tag{4.1.2}$$

(ii) Now, we turn to $E_{k-1}(t)$, i.e, with case $k-1$ already infected downstream nodes. We have three possible events for this individual and all its descendants in a small time interval, (a) direct observation, (b) undetected recovery, (c) infection of one downstream node. We use the same argument as in the first case only that we include an infection event at rate β.

$$\begin{aligned} \frac{d}{dt}q(t) &= -(\mu+\sigma+\beta)\, q(t) \\ \frac{d}{dt}r(t) &= \mu\, q(t) \end{aligned}$$

$$q(t) = e^{-(\mu+\sigma+\beta)t}, \quad r(t) = \int_0^t \mu e^{-(\mu+\sigma+\beta)t} dt = \frac{\mu}{\mu+\sigma+\beta}\left(1 - e^{-(\mu+\sigma+\beta)t}\right)$$

We need to incorporate the infection. The probability density to produce one downstream infected individual of state is

$$\beta q(t) = \beta e^{-(\mu+\sigma+\beta)t}$$

We note that, at infection, a newly infected individual would have 0 infected downstream, so that the state of every new born would always be 0. If a birth event did occur at time $a < t$, the state of this parent lineage changes from $k-1$ to k. Thus, the probability for both lineages (newly infected and parent) to stay undetected becomes $E_0(t-a)E_k(t-a)$. We need to integrate over a to obtain the probability of all new infections as well as the probability to stay undetected in $[0, t]$,

$$\int_0^t \beta e^{-(\mu+\sigma+\beta)a} E_0(t-a)E_k(t-a)da.$$

Thus, the sum of these probabilities gives the overall probability for both lineages to stay undetected

$$\begin{aligned} E_{k-1}(t) &= e^{-(\mu+\sigma+\beta)t} + \frac{\mu}{\mu+\sigma+\beta}\left(1 - e^{-(\mu+\sigma+\beta)t}\right) + \int_0^t \beta e^{-(\mu+\sigma+\beta)a} E_0(t-a)E_k(t-a)da \\ &= \frac{\mu}{\mu+\sigma+\beta} + \frac{\sigma+\beta}{\mu+\sigma+\beta} e^{-(\mu+\sigma+\beta)t} + \int_0^t \beta e^{-(\mu+\sigma+\beta)(t-a)} E_0(a)E_k(a)da \end{aligned}$$

Hence,

$$\begin{aligned} \frac{d}{dt}E_{k-1}(t) =& -(\mu+\sigma+\beta)\left(\frac{\sigma+\beta}{\mu+\sigma+\beta} e^{-(\mu+\sigma+\beta)t} + \int_0^t \beta e^{-(\mu+\sigma+\beta)(t-\tau)} E_0(a)E_k(a)da\right) \\ &+ \beta E_0(t)E_k(t) \\ =& -(\mu+\sigma+\beta)\left(E_{k-1}(t) - \frac{\mu}{\mu+\sigma+\beta}\right) + \beta E_0(t)E_k(t) \\ =& \mu - (\mu+\sigma+\beta)E_{k-1}(t) + \beta E_0(t)E_k(t) \end{aligned}$$

For $n = 0, .., k$, we have

$$\frac{d}{dt}E_{k-n}(t) = \mu - (\mu+\sigma+n\beta)E_{k-n}(t) + n\beta E_0(t)E_{k-n+1}(t),$$

where $E_{k+1}(t) = 0$. If we replace $i = k - n$, we obtain differential equations for the rate of change in the probability $E_i(t)$ as we descend along the branch ($i = 0, .., k$ and $E_{k+1}(t) = 0$)

$$\frac{d}{dt}E_i(t) = \mu - \big(\mu + \sigma + (k-i)\,\beta\big)E_i(t) + (k-i)\,\beta E_0(t)E_{i+1}(t) \qquad E_i(0) = 1. \quad (4.1.3)$$

In this case, equation 4.1.3 is a non-linear coupled system of differential equations which can be solved numerically. As illustrated in Fig. 4.4, the number of already infected downstreams greatly influences the probability for a clade to stay undetected. The clade of a randomly chosen node with no infected downstream ($k = 0$) has the least probability go undetected because of possible further infections. An infectee or one of the descendants of the infectee of this focal individual can trigger a direct recovery event which decreases the probability for the clade to go undetected. This probability $E_i(t)$ increases as the number of already infected downstreams increases. For all downstream edges infected ($i = k$), there would be no possible further infection, hence the most probability to go undetected. For an infinite tree and a fixed degree distribution, we find in Fig. 4.4, for appropriate parameter chosen, a satisfying agreement with simulation results.

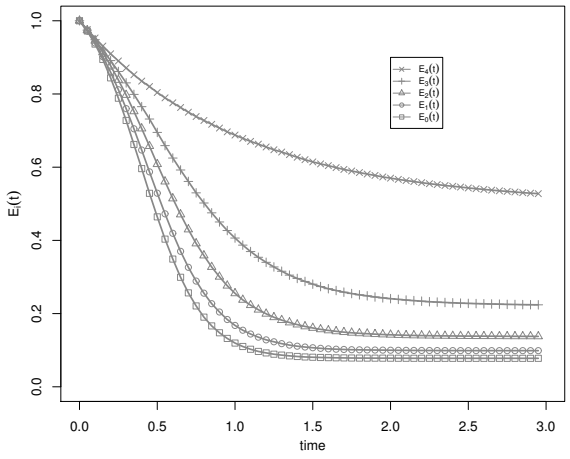


Figure 4.4: The probability $E_i(t)$ for a lineage and all of its descendants to go unobserved after time t. Gray thick lines: 100000 simulated results for the different types, other points symbols for ODE results respectively - square: $E_0(t)$, circle: $E_1(t)$, triangle point up: $E_2(t)$, plus: $E_3(t)$, cross: $E_k(t)$. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

4.1.1.2 Calculating $D_j^i(t)$

Now, we switch back to the probability $D_j^i(t)$ which depends not only on time but also on the probability $E_i(t)$. Recall that $D_j^i(t)$ is probability density for a clade starting with a type i individual to have evolved to a type j individual as inferred. We calculate $D_j^i(t)$ backward in time starting with the leaf of the sampled tree. We assume our sampled tree is

known as well as other properties of the tree. Consider a leaf in state j who had recovered observed in the past at time τ with rate σ.

Thus, we have the initial condition for this probability to be

$$D_j^i(\tau) \;=\; \sigma\,\chi_i(j), \tag{4.1.4}$$

where χ_i is an indicator function defined as

$$\chi_i(j) := \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{if } j \neq i. \end{cases}$$

At time τ, this individual in state i who produces the sampled tree together with the sampled leaf in state j is sampled if $j = i$. For $t > \tau$, we calculate $D_j^i(t + \Delta t)$ backward in time, starting with a j individual along the branch represented by lineage i. We investigate the different probabilities by considering all possible events in a small time interval.

(i) The probability $D_j^k(t)$ is also straight forward as in $E_k(t)$. The only possibility is for the lineage k to have evolved from itself between now and the past, and so must recover observed for the clade to survive. Let $\phi(t)$ denote the probability the clade survives in $[0, t - \tau]$.

$$\frac{d}{dt}\phi(t) \;=\; -(\mu + \sigma)\,\phi(t), \qquad \phi(\tau) = \sigma. \tag{4.1.5}$$

Recall that the initial condition is such that the rate of recovery occurs with rate $\mu + \sigma$ and probability to be observed $\frac{\sigma}{\mu+\sigma}$.

Solving for ϕ yields

$$\phi(t) = \sigma\, e^{-(\mu+\sigma)(t-\tau)}.$$

Thus, the probability that the clade represented by node k survives in $[0, t - \tau]$ reads

$$D_j^k(t) \;=\; \sigma\, e^{-(\mu+\sigma)(t-\tau)}\chi_i(j).$$

Hence, we have

$$\frac{d}{dt}D_j^k(t) \;=\; -(\mu + \sigma)\, D_j^k(t) \tag{4.1.6}$$

(ii) For $D_j^{k-1}(t)$, there is a possibility for one further infection before the present time. Considering every possible event during in a small time interval which is the combination of

(a) no birth, no death from lineage $k-1$,

(b) birth of a new lineage 0 such that this new lineage and its descendants must stay undetected before the present time with probability $E_0(t)$ and thus the clade survives by a detection from lineage $k-1$,

(c) birth of a new lineage in state 0 such that lineage $k-1$ must stay undetected with probability $E_{k-1}(t)$ before the present time and thus the clade survives by a detection from or one of the descendants from lineage 0.

For event (a), we use the assumption in the first case (eqn. 4.1.5) and include infection term β. We have

$$\begin{aligned} \frac{d}{dt}\phi(t) &= -(\mu+\sigma+\beta)\,\phi(t) \\ \phi(t) &= \sigma\, e^{-(\mu+\sigma+\beta)(t-\tau)}. \end{aligned}$$

For event (b), we incorporate infection. The probability density for this individual to produce a downstream infected individual is

$$\beta\phi(t) = \beta\,\sigma\, e^{-(\mu+\sigma+\beta)(t-\tau)}$$

and this newly infected plus all its descendants must stay undetected before the present time. If infection did occur at time $c < t-\tau$, the probability for the new born lineage (always in state 0) and the parent lineage (now in state k) to stay detected and undetected respectively reads $E_0(t-\tau-c)\,D_j^k(t-\tau-c)$.

For event (c), we use the reverse argument as event (b). Lineage k stays undetected while lineage 0 survives with the probability $E_k(t-\tau-c)\,D_j^0(t-\tau-c)$. We note that the subclade, $D_j^0(t)$ represented by this newly infected must produce an observed leaf j in the present time for the whole clade $D_j^{k-1}(t)$, represented by $k-1$ individual to survive.

We have to integrate over c to obtain the corresponding probabilities and thus, the combination of these events gives the overall probability for $D_j^{k-1}(t-\tau-c))$,

$$\begin{aligned} D_j^{k-1}(t) &= \sigma\, e^{-(\mu+\sigma+\beta)(t-\tau)} + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)c} E_0(t-\tau-c) D_j^k(t-\tau-c)dc \\ &\quad + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)c} E_k(t-\tau-c) D_j^0(t)(t-\tau-c)dc, \\ &= \sigma\, e^{-(\mu+\sigma+\beta)(t-\tau)} + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)(t-\tau-c)} E_0(c) D_j^k(c)dc \\ &\quad + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)(t-\tau-c)} E_k(c) D_j^0(c)dc. \end{aligned}$$

$$\frac{d}{dt}D_j^{k-1}(t) = -\frac{\sigma}{\sigma+\mu+\beta}\, e^{-(\mu+\sigma+\beta)(t-\tau)} - (\sigma+\mu+\beta)\Bigg(\int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)(t-\tau-c)} E_0(c) D_j^k(c) dc + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)(t-\tau-c)} E_k(c) D_j^0(c) dc\Bigg) + \beta\, E_0(t)\, D_j^k(t) + \beta\, E_k(t)\, D_j^0(t),$$

$$= -(\sigma+\mu+\beta)\Bigg(\sigma\, e^{-(\mu+\sigma+\beta)\tau} + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)(t-\tau-c)} E_0(c) D_j^k(c) dc + \int_0^{t-\tau} \beta\sigma e^{-(\mu+\sigma+\beta)(t-\tau-c)} E_k(c) D_j^0(c) dc\Bigg) + \beta\, E_0(t)\, D_j^k(t) + \beta\, E_k(t)\, D_j^0(t).$$

Thus

$$\frac{d}{dt}D_j^{k-1}(t) = -(\sigma+\mu+\beta)\, D_j^{k-1}(t)) + \beta\, E_0(t)\, D_j^k(t) + \beta\, E_k(t)\, D_j^0(t)).$$

Corollary 4.1.1. *For $n = 0, .., k$, we have*

$$\frac{d}{dt}D_j^{k-n}(t) = -(\sigma+\mu+n\,\beta)\, D_j^{k-n}(t) + n\,\beta\, E_0(t)\, D_j^{k-n+1}(t) + n\,\beta\, E_{k-n+1}(t)\, D_j^0(t).$$

Similarly, if we replace $i = k - n$, we obtain differential equations

$$\frac{d}{dt}D_j^{i}(t) = -\big(\sigma+\mu+(k-i)\,\beta\big)\, D_j^{i}(t) + (k-i)\,\beta\, E_0(t)\, D_j^{i+1}(t)) + (k-i)\,\beta\, E_{i+1}(t)\, D_j^0(t) \quad (4.1.7)$$

where $E_{k+1}(t) = 0$ and $D_j^{k+1}(t) = 0$, with initial condition

$$D_j^i(\tau) = \begin{cases} \sigma, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$$

As shown in Figures. 4.5, 4.6, 4.7, 4.8 and 4.9, we find the different probability densities $D_j^i(t)$ for an individual represented by node i and all its descendants to have evolved between the past and the present. For the different state of the sampled leaf j at time τ, we find a satisfying agreement with simulated results. It is not surprising in Fig. 4.5 we have zero densities $D_j^i(t)$ for $i = k, j \neq k$, this is because a focal node k (all downstreams infected) cannot produce any further downstreams and hence its clade could have evolved only from itself following an observed recovery iff $j = k$ (first graph at the top panel in Fig. 4.5). This probability decreases as the number of already infected downstreams decreases. The probability densities $D_j^i(t)$ for other focal nodes with infected downstreams $i = 0, 1, ..., k-1$ will all be non-zero because of possible infection events and different state observations.

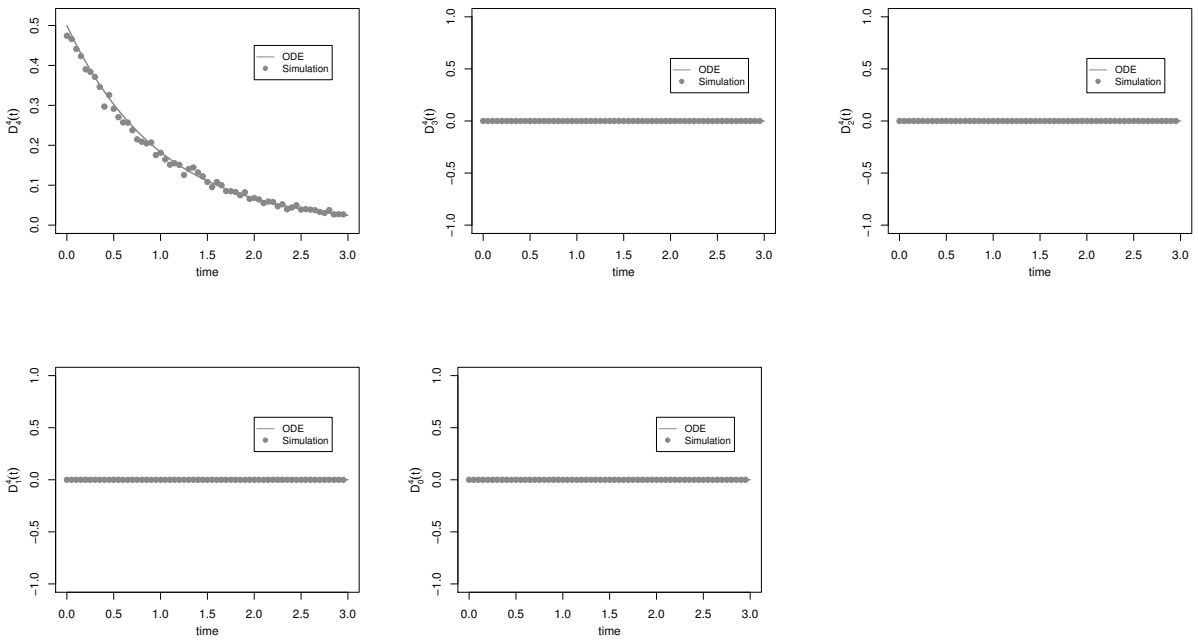


Figure 4.5: The probability density $D_j^4(t)$ for the clade and all its descendants to have evolved as the observed sampled tree given that the lineage started in state 4. $D_j^4(t)$ for the different state of each observed leaf j in the sampled tree: 100000 simulated results (grey dots), ODE results (grey lines). Top row from left to right: state 4, 3, 2, Second row: state 1 and 0. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

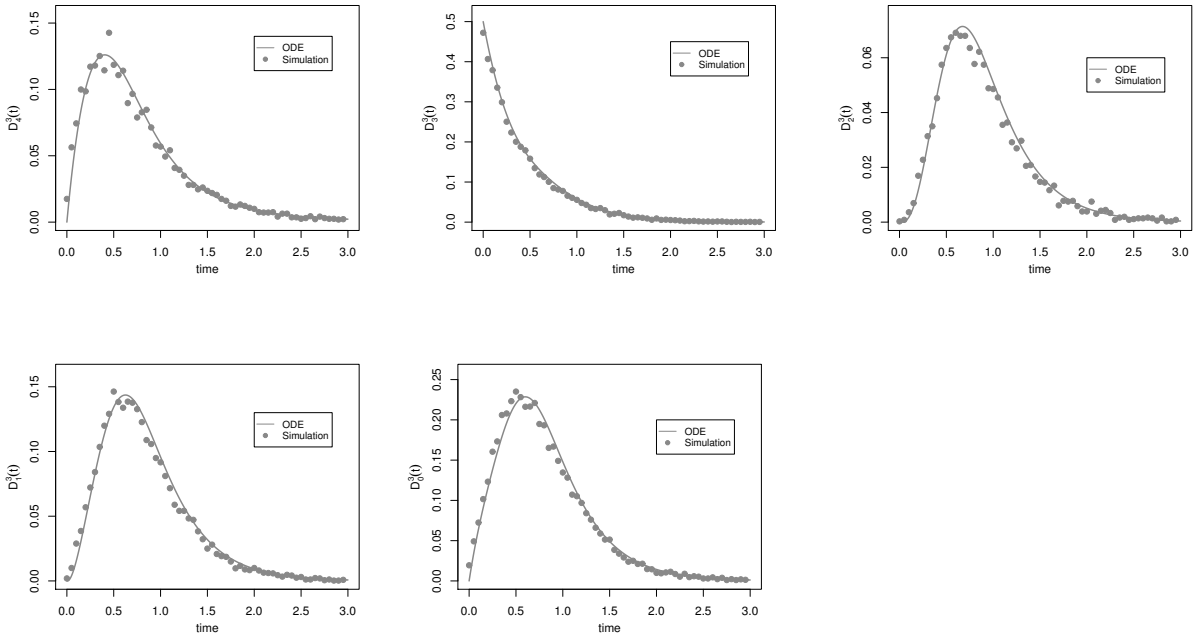


Figure 4.6: The probability density $D_j^3(t)$ for the clade and all its descendants to have evolved as the observed sampled tree given that the lineage started in state 3. $D_j^3(t)$ for the different state of each observed leaf j in the sampled tree: 100000 simulated results (grey dots), ODE results (grey lines). Top row from left to right: state 4, 3, 2, Second row: state 1 and 0. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

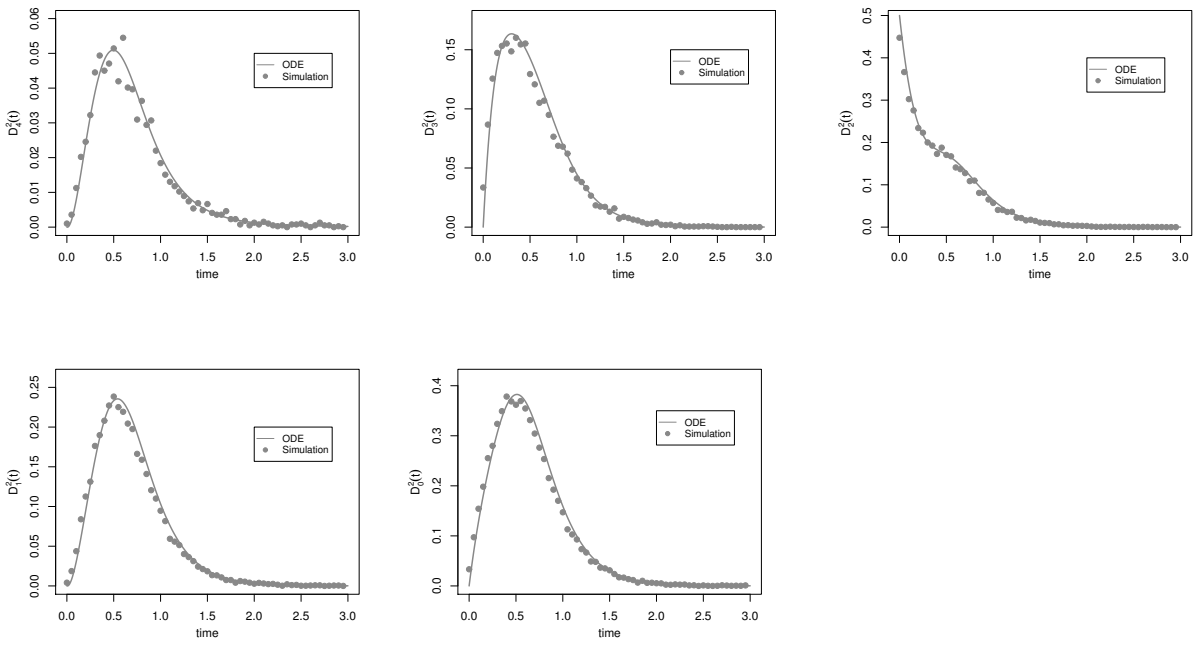


Figure 4.7: The probability density $D_j^2(t)$ for the clade and all its descendants to have evolved as the observed sampled tree given that the lineage started in state 2. $D_j^2(t)$ for the different state of each observed leaf j in the sampled tree: 100000 simulated results (grey dots), ODE results (grey lines). Top row from left to right: state 4, 3, 2, Second row: state 1 and 0. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

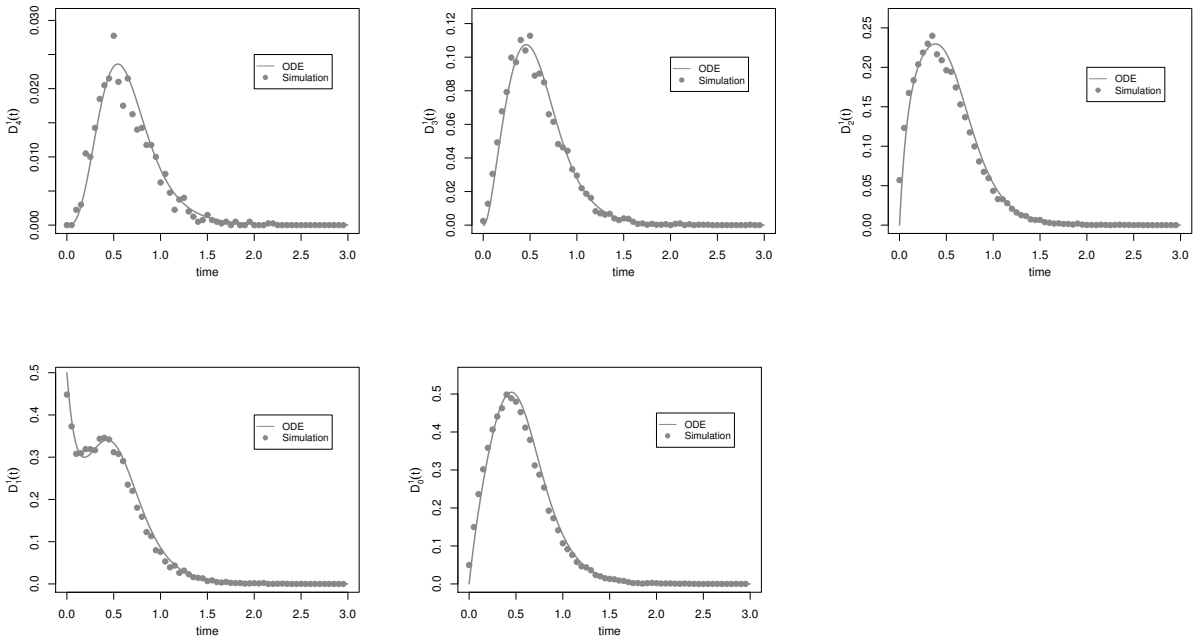


Figure 4.8: The probability density $D_j^1(t)$ for the clade and all its descendants to have evolved as the observed sampled tree given that the lineage started in state 1. $D_j^1(t)$ for the different state of each observed leaf j in the sampled tree: 100000 simulated results (grey dots), ODE results (grey lines). Top row from left to right: state 4, 3, 2, Second row: state 1 and 0. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

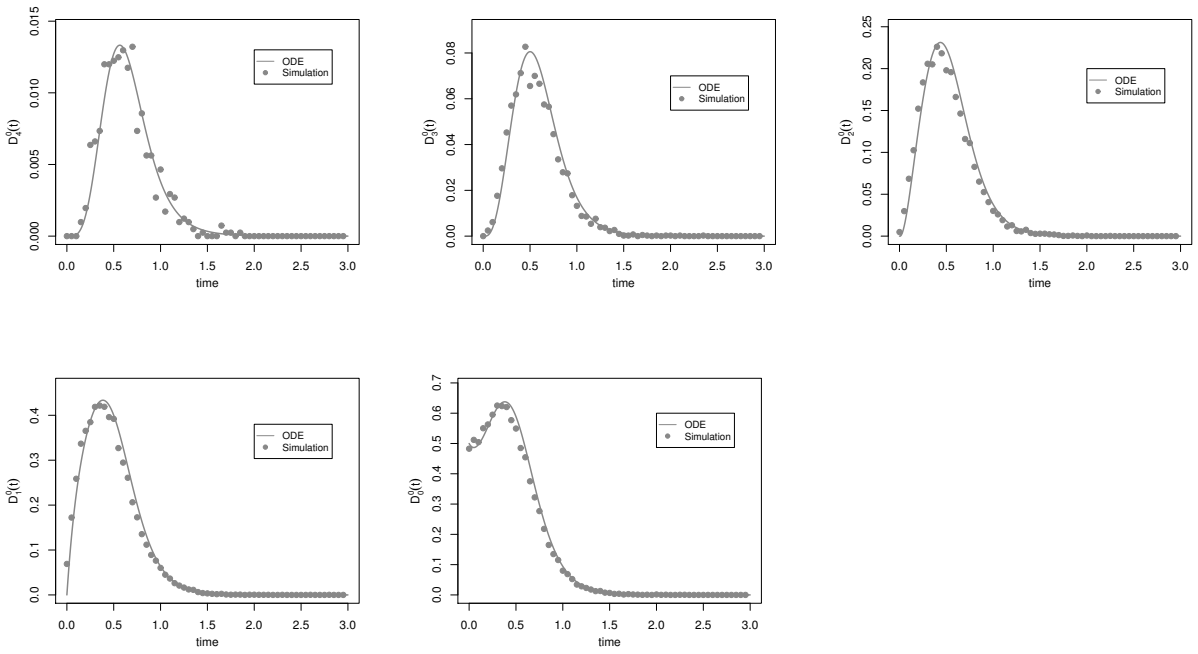


Figure 4.9: The probability density $D_j^0(t)$ for the clade and all its descendants to have evolved as the observed sampled tree given that the lineage started in state 0. $D_j^0(t)$ for the different state of each observed leaf j in the sampled tree: 100000 simulated results (grey dots), ODE results (grey lines). Top row from left to right: state 4, 3, 2, Second row: state 1 and 0. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

Remark 4.1.1. *So far, we started with how a single individual, as an initial starting point, is detected from a subepidemic started by an individual. The clade started by the individual is observed following the observation of the first individual in the clade.*

In the next section, we will proceed from considering one individual to a broader view of multiple individuals. That is, we do not explicitly take into account the first observed individual but all expected observed individuals. This would allow for the analysis of contact tracing in the epidemic process.

4.2 Estimations using expectations

We are able to formulate theories of maximum likelihood estimators from the expected number of observed individuals in different character state.

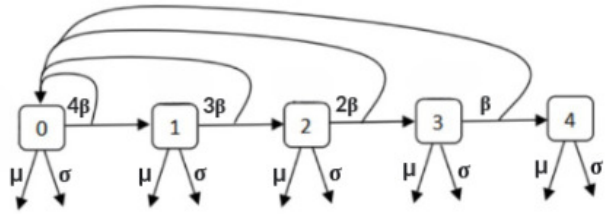


Figure 4.10: A flowchart and summary of the modifided MTBD process. The different compartments shows the expectated number of observed individuals in the complete phylogeny.

4.2.1 Type expectations

We start by defining this framework as follow;

Let $N_i(t)$ be the expected number of active individuals of type i at time t, where $i \in \mathbb{N}_0$. A node of type i has $k-i$ susceptible downstream individuals, and thus will infect one of them at rate $(k-i)\beta$ and immediately transition to type $i+1$. Newly infecteds are only possible from states 0 to $k-1$ (See Fig. 4.10 for a summary of the model).

We obtain the differential equation

$$\begin{aligned} \frac{d}{dt}N_0(t) &= -(\mu+\sigma+4\beta)N_0(t) + \beta\sum_{j=0}^{k-1}(k-j)N_j(t), \\ \frac{d}{dt}N_i(t) &= -(\mu+\sigma+(k-i)\beta)N_i(t) + (k-i+1)\beta N_{i-1}(t), \quad i \geq 1. \end{aligned} \tag{4.2.1}$$

With $N_0(0) = 1$ and $N_i(0) = 0, i \neq 0$, we are able to set up the probability density $\psi_i(t)$ for the rate at which we observe an individual of type i at time t

$$\psi_i(t) = \sigma\, N_i(t). \tag{4.2.2}$$

As shown in Fig. 4.11, the results for the rate of observation for the different states is expected to grow exponentially fast. As the chain of sequence is triggered by a state change, the expected number of lineages in each state decreases as we move from type 0 to k. As we have more influx in state 0 due to births from all possible states, it is expected to have the most individuals in state 0. These numbers decreases as move along these chain due to possible recovery events. For appropriate parameter chosen, we find in Fig. 4.11 a perfect agreement with simulation results.

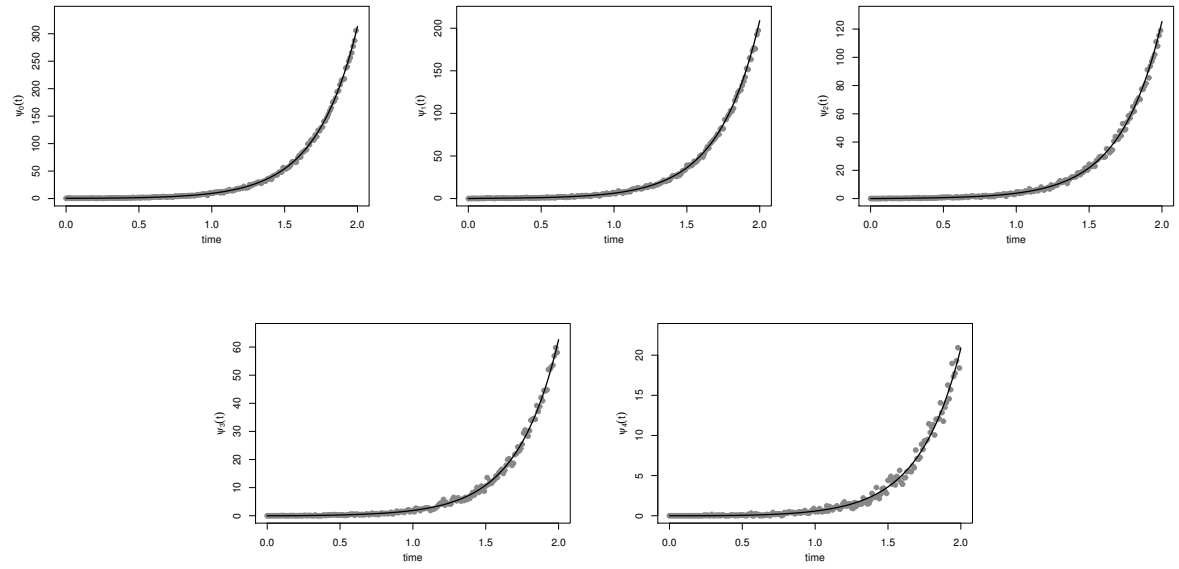


Figure 4.11: Density $\phi_i(t)$ for a lineage to be in i state at time t. gray dots: 2000 simulated results for the different types, black thick lines for ODE results respectively. $\psi_0(t)$, $\psi_1(t)$, $\psi_2(t)$, $\psi_3(t)$, $\psi_4(t)$. Choice of parameters: $\mu, \sigma = 0.5$, $\beta = 1.5$, $k = 4$ (constant).

4.2.2 Maximum likelihood estimator

In the long run, we have solutions $N_i(t)$ of the ODE, where $i \in \mathbb{N}_0$ is the type of an individual and $\psi_i(t)$ is the rate of observing an individual of type i. We can interpret this setting as a non-stationary multitype Poisson process, where the intensity of the process for type i is given by $\psi_i(t)$. We slowly think about an appropriate estimator for the parameters.

Data:

We have n observations, where $t_\ell \in \mathbb{R}_+, \ell = 1, ., n$ are the time point of observations, and $i_\ell \in \{0, .., K\}$ are the types of the observations.

Model:

$\sigma N_i(t) = \sigma N_i(t; \vec{\mu})$ is the intensity of observations of type i. Here, $\vec{\mu}$ is the parameter vector that we wish to estimate. These parameters influence the solution of the ODE $N_i(t)$.

Single type Poisson process, constant intensity:

We start with thinking about the standard Poisson process [9] with a constant intensity and only one type. That is, we assume that we only observe type 0 , and that $N_0(t; \vec{\mu}) = N_0(\vec{\mu})$ is constant in time. Then,

$$\Delta_\ell := t_{\ell+1} - t_\ell \sim \text{Exp}\left(\sigma\, N_0(\vec{\mu})\right)$$

and hence, the likelihood for one data point Δ_ℓ is given by

$$\sigma N_0(\vec{\mu}) e^{-\sigma N_0(\vec{\mu})\Delta_\ell}$$

The likelihood for the data thus reads

$$\mathcal{L}\left(\Delta_\ell, \ell = 1, .., n-1 \mid \vec{\mu}\right) = \prod_{\ell=1}^{n-1} \sigma N_0(\vec{\mu}) e^{-\sigma\, N_0(\vec{\mu})\Delta_\ell}$$

resp. the log-likelihood is given by

$$\begin{aligned} \mathcal{LL}\left(\Delta_\ell, \ell = 1, .., n-1 \mid \vec{\mu}\right) &= (n-1)\ln\left(\sigma N_0(\vec{\mu})\right) - \sigma N_0(\vec{\mu}) \sum_{\ell=1}^{n-1} \Delta_\ell \\ &= (n-1)\ln\left(\sigma N_0(\vec{\mu})\right) - \sigma N_0(\vec{\mu})\left(t_n - t_1\right) \end{aligned} \quad (4.2.3)$$

Single type Poisson process, variable intensity:

We proceed to a Poisson process where we only observe one type, but the intensity is time dependent. That is, we assume that we only observe type 0, and that $N_0(t; \vec{\mu})$ is in general not constant in time. How likely is it to find the observation time $t_{\ell+1}$ after the observation at time t_ℓ? We might reformulate the situation: We start in $t = t_\ell$ with a state "no next observation", and with rate $\sigma N_0(t)$, we jump into the state "next observation did happen". Thus,

$$P(\text{ next observation after } t) = \sigma N_0(t_\ell; \vec{u}) e^{-\int_{t_\ell}^{t} \sigma N_0(s;\vec{u}) ds}$$

Hence, the likelihood to have the next observation at $t = t_{\ell+1}$ is the negative derivative,

$$\mathcal{L}\left(t_\ell, \ell = 1, .., n-1 \mid \vec{\mu}\right) = \prod_{\ell=1}^{n-1} \sigma N_0\left(t_\ell; \vec{\mu}\right) e^{-\int_{t_\ell}^{t_{\ell+1}} \sigma N_0(s;\vec{u}) ds}$$

resp. the log-likelihood is given by

$$\begin{aligned}\mathcal{LL}\left(t_\ell, \ell = 1, .., n-1 \mid \vec{\mu}\right) &= \sum_{\ell=1}^{n-1} \ln\left(\sigma N_0\left(t_\ell; \vec{\mu}\right)\right) - \sum_{\ell=1}^{n-1} \int_{t_\ell}^{t_{\ell+1}} \sigma N_0(s; \vec{u}) ds \\ &= \sum_{\ell=1}^{n-1} \ln\left(\sigma N_0\left(t_\ell; \vec{\mu}\right)\right) - \int_{t_1}^{t_\ell} \sigma N_0(s; \vec{u}) ds\end{aligned}$$

Let T denotes the duration of the epidemic process. The overall log-likelihood considering all time intervals t_ℓ, $\ell \in (0, T]$ where $0 < t_1 \leq ... \leq t_n < T$ is

$$\mathcal{LL}\left(t_\ell, \ell = 1, .., n \mid \vec{\mu}\right) = n \ln(\sigma) + \sum_{\ell=1}^{n} \ln\left(N_0\left(t_\ell; \vec{\mu}\right)\right) - \sigma \int_0^T N_0(s; \vec{u}) ds \tag{4.2.4}$$

Multi-type Poisson process, variable intensity:

Now, we proceed to a Poisson process where we observe different types with time-dependent intensities. That is, we assume that we observe type i, and that $N_i(t; \vec{\mu})$ is not constant in time for $i = 0, .., k$. Considering all type i individuals are $i.i.d$, the sum of k independent Poisson process respectively gives a Possion process. Using the same formulation as above, we are able to set up the likelihood framework.

We define

$$N(t; \vec{u}) = N(t) = \sum_{i=1}^{k} N_i(t)$$

to be the overall intensity of all type i individuals at time t and

$$P_i(t) = \frac{N_i(t)}{N(t)},$$

the relative frequency that an observation is of type i. Thus $N_i(t) = N(t)P_i(t)$ is the intensity of observation of type i. Let's consider k consecutive time intervals for observing a type i individual,

$$P(\text{ next type } i \text{ observation after } t) = \sigma N(t_\ell) P_{i_\ell}(t_\ell) \exp\left\{ - \int_{t_\ell}^{t} \sigma N(s) P_{i_\ell}(s) ds \right\}$$

Hence, the likelihood to have the next observation at $t = t_{\ell+1}$ is

$$\mathcal{L}\left(t_\ell, \ell = 1, .., n \mid \vec{\mu}\right) = \prod_{\ell=1}^{n} \sigma N(t_\ell) P_{i_\ell}(t_\ell) \exp\left\{ - \int_{t_\ell}^{t_{\ell+1}} \sigma N(s) P_{i_\ell}(s) ds \right\}$$

resp. the log-likelihood over all type i is given by

$$\mathcal{LL}\left((t_\ell, i_\ell), \ell = 1, .., n \mid \vec{\mu}\right) \quad = \quad n\ \ln(\sigma) + \sum_{\ell=1}^{n} \left(\ln\left(N(t_\ell)\right) + \ln\left(P_{i_\ell}(t_\ell)\right) \right) - \sigma \int_0^T N(s)ds,$$

where $\sum_0^k P_{i_\ell}(t_\ell) = 1$.

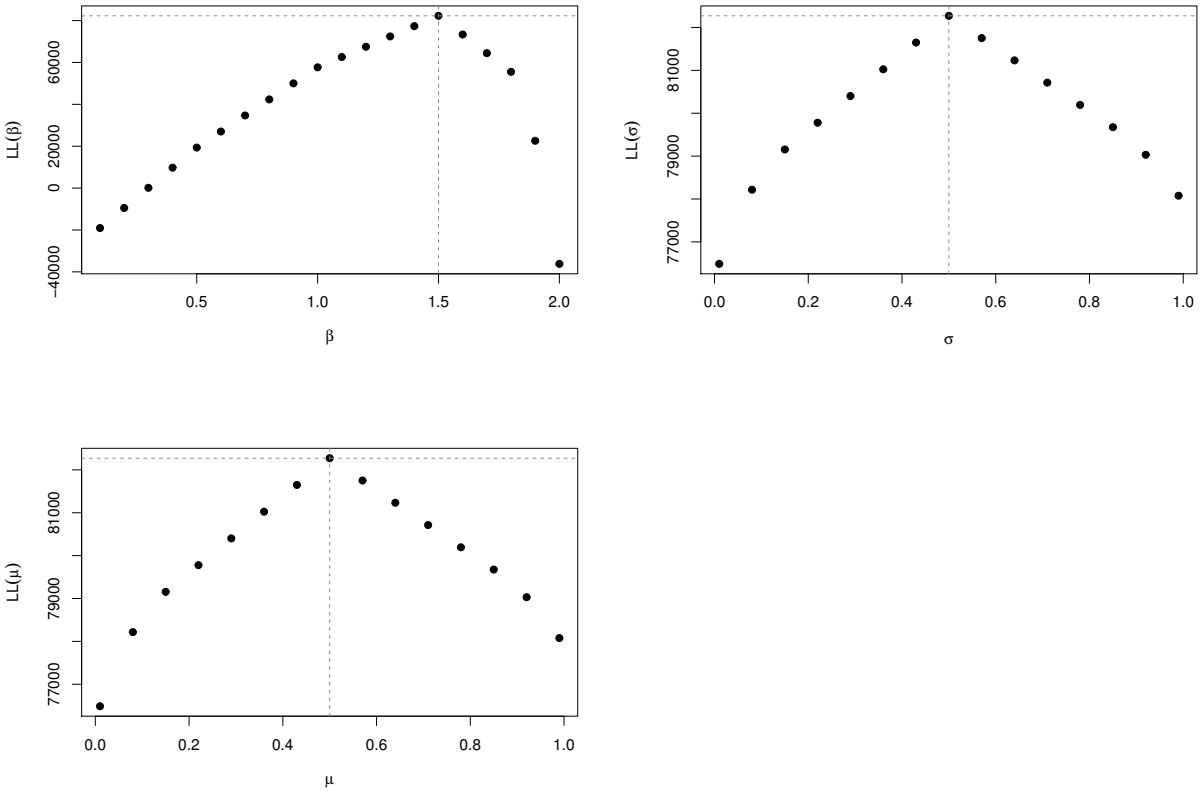


Figure 4.12: The Log likelihoods for estimated parameters β, μ, σ.

As shown in Fig. 4.12, we find the estimation of the contact rate, direct and indirect recovery rate parameters. The parameters $\beta = 1.5, \mu = 0.5$ and $\sigma = 1.5$ of the model that maximises the likelihood is well estimated given the data points.

4.3 Estimations using contact tracing

While the phylogenetic methods are better suited in estimating the timing of the infectious process, we expect contact tracing events to incorporate information about the heterogene-

ity in the contact structure. We are able to find the distributions of ages since infection and further detected cases by contact tracing.

4.3.1 Distribution of ages since infection

As the contact graph which is tree is mainly influenced via the downstream degree distribution, we are able find approximate distributions of the age since infection. Starting from the root node, the phylogenetic tree is constructed in generating downstreams for an infected node on a fly. That is, only when an individual is infected, his/her downstreams (susceptible children) are generated. This process results in an exponential increase in the number of nodes.

Using these considerations, we recall the age-structured model (set up in section Mean field without contact tracing on a tree)

$$\begin{aligned} (\partial_t + \partial_a) i(t,a) &= -(\mu+\sigma) i(t,a) && (4.3.1)\\ i(t,0) &= \int_0^\infty \theta(a)\, i(t,a)\, da. && (4.3.2) \end{aligned}$$

where

$$\theta(a) = \mathbb{E}[K]\, \beta\, e^{-\beta a}$$

is the age dependent rate at which an infected individual produces downstream infecteds.

This age structured model will tend to an exponential growing solution with a stable age structure,

$$i(t,a) = I_0\, e^{\lambda t}\, e^{-\lambda\, a}\, \hat{\kappa}_\infty(a)$$

where

$$\hat{\kappa}_\infty(a) = e^{-(\mu+\sigma)a}$$

is the probability to be infectious at age a if no tracing takes place (i.e tracing probability $p = 0$). The exponent λ is the unique real root of

$$1 = \int_0^\infty \theta(a)\, e^{-(\lambda+\mu+\sigma)a}\, da = \mathbb{E}[K]\, \beta \int_0^\infty e^{-(\lambda+\mu+\sigma+\beta)a}\, da.$$

We have

$$\lambda = \beta(\mathbb{E}[K] - 1) - \mu - \sigma$$

and the limiting age stucture reads

$$i(t,a) = I(t)\, (\mu+\sigma+\lambda)\, e^{-(\lambda+\mu+\sigma)\, a} = I(t)\, \beta\, (\mathbb{E}[K]-1)\, e^{-\beta(\mathbb{E}[K]-1)a}.$$

The asymptotic age distribution tends to

$$\varphi(a) = \frac{\sigma\, i(t,a)}{\int_0^\infty \sigma\, i(t,a)da} = \beta\left(\mathbb{E}[K]-1\right)\, e^{-\beta(\mathbb{E}[K]-1)a}. \tag{4.3.3}$$

As shown in Figure. 4.13, we have a higher density of younger age group in the population. For any randomly chosen individual given by the age since infection, it is not surprising to have a younger dominating age class. Due to the exponentially fast growing population, this asymptotic age distribution is expected.

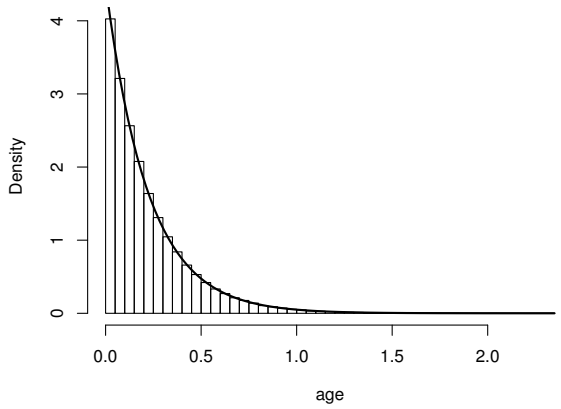


Figure 4.13: Theoretical and simulated age distributions since infection.

4.3.2 Distribution of detected cases

Furthermore, we estimate the distribution of detected cases through contact tracing by an index case. We start with forward tracing (only downstream infecteds can be traced). We combine this result with backward tracing to yield full tracing.

Proposition 4.3.1. *Let $\hat{p}$ be the probability that an infected downstream is successfully traced given that the focal individual is still infectious in the time interval $[0,a]$.*

$$\hat{p} = p\frac{\beta}{\mu+\sigma-\beta}\left(e^{-\beta a} - e^{-(\mu+\sigma)a}\right) \tag{4.3.4}$$

Proof. We condition on the focal individual to be infectious in the time interval $[0,a]$. Let s_1 represent the state for a downstream to still be susceptible at age 0, s_2 for this downstream to be infected by the focal individual at some age $c \in [0,a]$, s_3 the state for the downstream to be removed following a succesful tracing event triggered by the focal individual.

We have the following ODE's

$$\begin{aligned} \dot{s_1} &= -\beta\, s_1 & s_1(0) = 1 \\ \dot{s_2} &= \beta s_1 - (\mu+\sigma)\, s_2 & s_2(0) = 0 \\ \dot{s_3} &= (\mu+\sigma)\, s_2 & s_2(0) = 0. \end{aligned}$$

The probability for the downstream to still be infectious given by s_2 at age a is given as

$$s_2 = \frac{\beta}{\mu+\sigma-\beta}\big(e^{-\beta a} - e^{-(\mu+\sigma)a}\big).$$

The probability $\hat{p}$ for a succesful event is the probability p for an edge to be removed independently times the probability s_2 for this edge to still be active. □

Proposition 4.3.2. *Let T be the random variable for the total number of i succesfully traced individuals by one index case. The asymptotic probability distribution of T reads*

$$P(T=i) = \int_0^\infty \Big[pP(I_a=1)\,P(T=i-1\,|\,a) + \big(1-pP(I_a=1)\big)\,P(T=i\,|\,a)\Big]\varphi(a)\,da \tag{4.3.5}$$

Proof. T follows a Binomial distribution $T \sim Binom(k,\hat{p})$ conditioned on age a. Thus, the probability of i downstream detectees given age a and k total downstreams is given as

$$P(T=i\,|\,a) = \binom{k}{i}\hat{p}(a)^i\,(1-\hat{p}(a))^{k-i}$$

Last we remove the condition a by integrating over all possible age of index cases $\varphi(a)$ (eqn. 4.3.3). Thus we have

$$\begin{aligned} P(T=i) &= \int_0^\infty P(T=i\,|\,a)\,\varphi(a)\,da \\ &= \int_0^\infty \binom{k}{i}\hat{p}^i\;(1-\hat{p})^{k-i}\,\beta\,(\mathbb{E}[K]-1)\;e^{-\beta\left(\mathbb{E}[K]-1\right)a}\,da \end{aligned} \tag{4.3.6}$$

As contacts of the index case happens on both upstream and downstreams, the infector plays a role in the distribution of the total number of detectees. For backward tracing, an edge connects from the index case to the infector. This infector is either still infectious, $P(I_a=1)$ or no longer infectious, $P(I_a=0)$ since age a. The probability for this infector to still be infectious at age a of infection of the index case [58] is given as

$$P(I_a = 1) = e^{-(\mu+\sigma)a} + \mathcal{O}(p).$$

Considering full tracing, we observe two scenerios:

1 The infector is still infectious and detected, and contributes to total number of further detected cases.

2 The infector is no longer infectious and not detected, and do not contribute to total number of further detected cases.

By reformulating eqn. 4.3.6, we have the asymptotic probability distribution of T to be

$$P(T = i) = \int_0^\infty \Big[pP(I_a = 1)\, P(T = i - 1 \,|\, a) + \big(1 - pP(I_a = 1)\big)\, P(T = i \,|\, a) \Big] \varphi(a)\, da.$$

□

4.3.3 Maximum likelihood estimator with contact tracing

4.3.3.1 Fixed degree

Data:

We have n observations, where $i_\ell \in \mathbb{N}_0, \ell = 1, ., n$ are the observations of the total number of detectees by an index case.

Model: We are able to set an estimator $P(T = i) = P(T = i \mid \vec{\mu})$ for this data points where $\vec{\mu}$ are the parameters of the estimator and the random variable T for the total number of traced individuals. The likelihood for one data point is given by

$$\int_0^\infty \Big[pP(I_a = 1)\, P(T = i_l - 1) + \big(1 - pP(I_a = 1)\big)\, P(T = i_l) \Big] \varphi(a)\, da.$$

For the data points, the likelihood result reads

$$\mathcal{L}\,(i_\ell, \ell = 1, .., n \mid \vec{\mu}) = \prod_{\ell=1}^{n} \int_0^\infty \Big[pP(I_a = 1)\, P(T = i_l - 1 \,|\, a) + \big(1 - pP(I_a = 1)\big)\, P(T = i_l \,|\, a) \Big] \varphi(a)\, da.$$

The log-likelihood is given by

$$\mathcal{LL}\,(i_\ell, \ell = 1, , .., n \mid \vec{\mu}) = \sum_{\ell=1}^{n} \ln \Big(\int_0^\infty \Big[pP(I_a = 1)\, P(i_l - 1 \,|\, a) + \big(1 - pP(I_a = 1)\big)\, P(i_l \,|\, a) \Big] \varphi(a)\, da \Big).$$

4.3.3.2 Poisson degree

So far, the likelihood estimator is formulated for a fixed degree distribution such that we have only detected cases by contact as an unknown parameter. In reality, we do not always know k due to randomness in individual contact structure. Here, we assume k and p unknown parameters and assume k possion graph. The likelihood estimator derived in the fixed case is adapted only we have to sum over all possible k such that

$$P(T = i \,|\, a) \;=\; \sum_{k}^{\infty} \binom{k}{i} \hat{p}(a)^i \, (1 - \hat{p}(a))^{k-i} \, P(K = k). \tag{4.3.7}$$

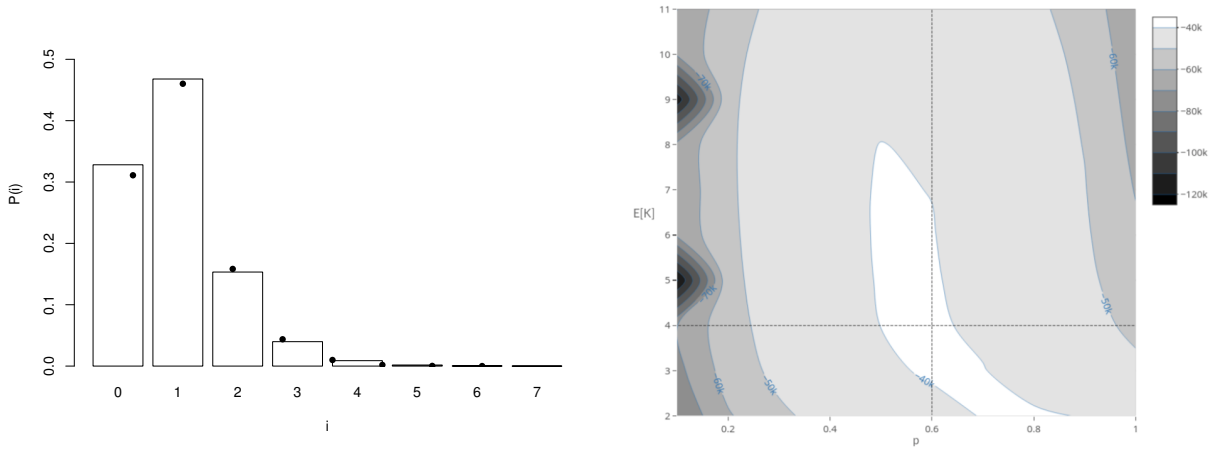


Figure 4.14: Left panel - bars: theory $P(T = i)$, black dots: 10000 simulated results. Right panel - true estimated tracing probability ($p = 0.6$) and expected degree distribution ($\mathbb{E}[K] = 4$).

As shown in Fig. 4.14, we find our estimated values of tracing probability p and the expected downstream degree distribution $\mathbb{E}[K]$ via a contour plot. The white region (the elbow spot) is able to show a global minimum for the true parameter values. For the estimated Poisson degree distribution, we find a satisfying result based on our theory assumption.

4.4 Application on real-world data

In the previous section, we have used the maximum likelihood estimator with contact tracing on simulated data to estimate the tracing probability and expected degree distribution. Clearly, in tracing data, we have information about the distribution of contacts, and also

about the fraction of infectious contacts we are able to identify via contact tracing. For simulated data, we show that the method yields nice results.

In this section, we would like to see how his methods works with empirical data. We sort to estimate these parameters from the COVID-19 pandemic data.

Data

The contact tracing we did analyse was from Karnataka, India where 956 confirmed index cases were reported between 9th of March 2020 and 20th of May, 2020 [34]. For each observation, the variables in the extracted dataset includes the ID's, primary chidren, secondary children and all chidren of index cases. The ID's corresponds to the chronological order of detection of index cases. The primary children is the number of detected cases which have been discovered by tracing the direct contacts of an index case (including those index cases where contact tracing did not lead to the detection of further cases) while secondary children is the number of direct and indirect contacts of an index case. A summary of the dataset is shown in Table 4.1.

We are not after a statistical model correlating co-factors of persons with some target variables, but what is intended is to consider the empirical number of detectees per index case. We only look at one-step tracing in the moment because only immediate neighbours of an index case is accounted for in the likelihood estimator. In the dataset, only immediate secondary cases are reported, thus we are able to analyse this scenerio with the estimator for forward tracing.

Results

For the relevant parameters needed for our estimator, we looked at literature that addressed the ancestrial variant that was most prevalent at that time in India. From Figure 4.15, we find the estimated tracing probability ($p = 0.7$) and expected degree distribution ($\mathbb{E}[K] = 37$). We should expect this high degree because the study was very careful in searching for secondary cases. We note that in the contact graph, the expected degree are not infectious contacts but only potential contacts. We have quite often that no case is detected and somehow one would expect if there would be an infectious contact, then basically we would expect that $p\,\mathbb{E}[K]$ is the expected number of detectees which is obviously not the case. In principle, we expect the number of secondary infectees to be much lesser than $\mathbb{E}[K]$ which should be in the magnitude of the reproduction number.

Number of detectees	0	1	2	3	4	5	6	7	8	10	11	12	13	15	16	19	22	28	29
Frequency	766	87	34	19	16	12	3	4	3	2	1	1	1	1	1	2	1	1	1

Table 4.1: Total number and frequencies of the total number of detected cases.

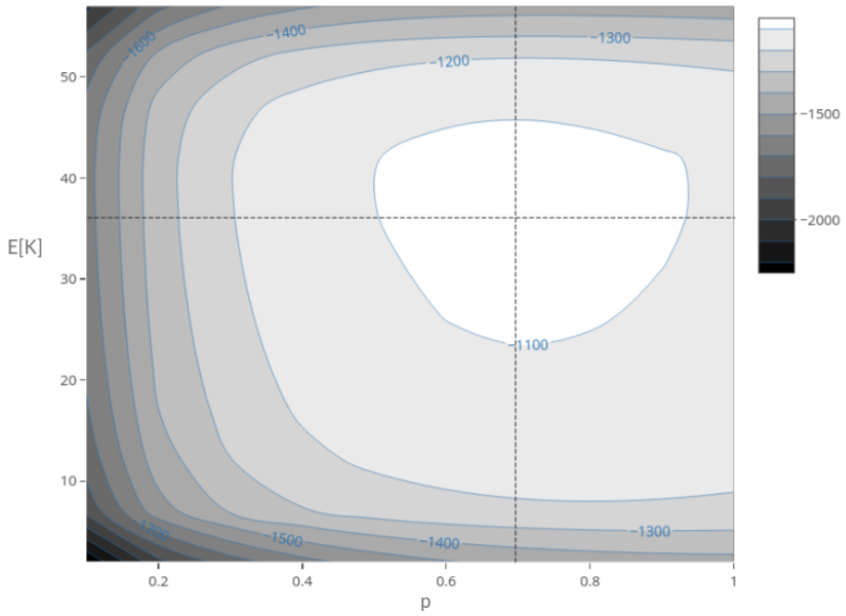


Figure 4.15: Estimated tracing probability and degree distribution from the reported data. Estimated parameters are $p = 0.7$ and $\mathbb{E}[K] = 37$

Chapter 5

DISCUSSION AND CONCLUSION

5.1 Discussion

In the first part of this thesis, we presented an approach for contact on random trees which extends the existing branching-process-theory for contact tracing in a randomly mixing population to a situation where contacts are only allowed along a prescribed, non-trivial contact graph. We restricted ourselves to an SIR model, while in homogeneous populations also SIS models can be handled. The central difference between the two situations is that in homogeneous models with large populations repeated contacts between the same individuals during the infectious period can be neglected, while they do occur most likely in graph models. The analytical machinery we did use cannot handle reinfection by an infectee, and therefore we did focus on the SIR model. Herein, the mathematical technique resembles that of the message passing method, which does also only apply to SIR models.

For the exact results in this thesis, the contact graph is chosen to be a (stochastic) tree. While the tree of infecteds is dynamically constructed in homogeneous models (at a given time point, the nodes are the infecteds, and a directed edge goes from infector to infectee), in the present case we focus on the tree of possible contacts, which is prescribed and static. In the present case, the tree of infecteds is a subtree of the contact tree. In order to better understand the relation between the different approaches, we clarify (some of) the time scales involved.

In particular, for real-world populations, we identify four relevant time scales: (a) The time scale at which the contact graph changes. A contact network is the abstraction of e.g. pair formation and sexual relations, family structure, or contacts at work. All these relations are likely to change on a slow time scale (children may grow up and leave the core family, a person changes his/her job position etc.). (b) The time scale at which a given edge within the contact graph is activated. Of course, there are different contacts with different intensities. The time scale for contacts within the family may be hours or

minutes, while the time scale of contacts in leisure activities as a chorus may be given by a week etc. (c) The time scale of the duration of an epidemic outbreak, and (d) the time scale of an individuals' infectious period.

Only if the time scale of an infectious outbreak (c) is much shorter than that time scale describing the change in social interactions (a), we can assume the contact graph to be static. That might not be the case for endemic diseases, or infections with a long infectious period as HIV. Also asymptomatic cases of gonorrhoea and chlamydia may have an infectious period around a year, such that a static contact graph may not be suited to cover all aspects of the disease dynamics. For the difference of the contact graph and the graph of infecteds, the time scale of the infectious period (d) and that of contacts on a single edge (b) becomes interesting. If the infectious period is long in comparison with the frequency of contacts on a given edge, this edge will most likely spread the infection; there is almost no difference between the contact graph and the graph of infecteds. In the contrary case, the graph of infecteds is a distinct subgraph of the contact graph. In that case, "contact tracing" will be rather "infectious contact tracing".

There is no difference for the effect of contact tracing if we focus on the (rather abstract) setting of a contact graph that is a tree, where initially only the root is infected. Any downstream contactee can be only infected by an upstream individual, and the effect of contact tracing and infectious contact tracing coincides. That becomes different in more complex and realistic models (e.g. the configuration model). Downstream individuals may get infected by another downstream individual. If we focus on contact tracing, that infected downstream individual may be detected, while in infectious contact tracing, that infected downstream infected may be missed. In the view of these considerations, it is an encouraging insight, that there is a limit that shows that contact tracing and infectious contact tracing lead to similar results for the appropriate scaling of contract tree and contact rates. There is no fundamental difference between homogeneous models and contact graph models, but only a gradual one.

A further interesting finding of the present work is the fact that the degree distribution of the contact graph influence the dynamics mainly via the expected number of edges (at least, if the tracing probability p is small, which is given in most applications). The higher moments play only a role if they become large (or even infinite). Also in the pair approximation approach, the first moment of the edge distribution enters the equations, but not the higher moments [41]. This result may possibly contradict the idea that the effect of contract tracing relies on the detection of super-spreaders: Super-spreader strongly influence the edge distribution of a randomly chosen neighbour of a randomly chosen node; the existence of super-spreaders, however, is expressed by the second moment of the edge distribution (the variance). As this second moment is only of minor importance for the description of the overall effect of contact tracing, it might be, that also the detection of super-spreaders is not that central.

This interpretation of our results is in line with the findings in [44, 45], where the authors showed that contact tracing works well in assortatively and associatively mixing contact graphs. Contact tracing seems to be rather robust w.r.t. the details of the contact graph structure. In our opinion, the present work indicates that we gain a better and better insight into contact tracing on a complete graph (homogeneous model) or on a static contact network. However, the interplay of the different time scales involved (infectious period, number of contacts per edge, reformation of the contact network) is still not well understood, and requires refined models with dynamic contact graphs, and analytical methods to handle these models.

Furthermore, in the final part of this thesis, we presented a phylogenetic method for estimating parameters in infectious models. Maddison et al. [52] had highlighted how the BiSSE model could be extended to a situation where branching occurs simultaneously with a character change. For the model studied here, our model assumtion is the same as an infection event causes a change of character state of an individual (i.e the infected downstream increases upon infection). Using the multi-type birth-death branching model idea of [74], we started with an approach which is based on calculating probability densities for observing a clade as inferred. However, we are only able to get results along the branch so far. For the probabilities at the node, the work is still in progress. For our estimation, we switched the approach from considering one observed individual to a situation of observing multiple individuals. There, we used a system of ODE's to capture the expected number of individuals in different character state at a given time. Thus, our model excludes information of lineages that have gone extinct in the pylogeny.

Other models, e.g Mossong et al. [57] carried out an empirical study to estimate contact rates directly by asking people who they have come in contact with daily. This study is particularly interesting for airborne diseases that can be spread through close contact. The different patterns of individual contact structures, e.g age groups, and social interactions allow for evaluating the risk of disease transmissions. Also Soetens et al. [73] used a maximum likelihood estimation to estimate the exposure type-specific attack rates and reproduction number from contact tracing data. Another important parameter estimated is the tracing probability. Müller and Hösel [58] did investigate parameter estimation from the distribution of further detected cases per index case in a homogeneous population. The expected number of traced persons per index case was estimated using a rule of thumb and the tracing probability was estimated by multiplying the expected total traced cases per an index case with a factor.

Here, we use a similar approach as [74, 52] by simulating phylogenetic trees with character states. We then developed a model for parameter estimation with and without contact tracing. Without contact tracing, we analyse different intensities of observation. We are able to estimate the infection and recovery rate parameters using maximum likelihood estimators. Then we proceed to contact tracing where we set up an age-structured model. Following closely Müller and Hösel [58], the distribution of ages since infection and further

detected cases per index are estimated. Using a fixed and Poisson degree distribution, we are able to find an estimate for the tracing probability and expected degree distribution which gives some information about the underlying contact structure. This result is particularly interesting for our analysis of contact tracing because in most applications, tracing probability is not easy to estimate but rather chosen to be small. In reality, as we do not always have much information about individual contact structure, this result is also useful in this context.

A further interesting part of this work was applying our theory in estimating tracing probability and the underlying contact structure from COVID-19 data reported in India from 9th of March 2020 to 20th of May 2020. On the one hand, we want to learn how our method will get on with empirical data, on the other hand, how the estimated parameters from the micro-level data might give us additional information about the performance of covid-19 tracing programs. We are able to estimate the tracing probability and expected degree of contact. For the high estimated tracing probability of 0.7, we are not surprised at the result because the performance of covid-19 tracing programs was at its peak during the time the data was collected. Also, we suspect that the lockdown measures imposed during this period may have contributed to a high adoption rate of contact tracing schemes leading to a high tracing probability. Van Egeren et al. [76] did estimate the reproduction number of the SARS-CoV-2 ancestral variant to be below 2 despite relaxation of lockdown and Gupta et al. [34] in their publication found significant heterogeneity in the individual-level effective contact rate of the SARS-CoV-2.

In our estimation study, we have used fixed contact and recovery rates, thus we suspect this may have also contributed to the high tracing probability due to the high heterogeneity in the individual-level contact network which is in line with the findings of Gupta et al. [34]. Furthermore, the number of cases in the reported data as described by Gupta et al. [34] were grouped in different clusters to analyse effective contact tracing guidelines for superspreading events targetting different age groups. They found 111 superspreader index cases surpass the threshold of ≥ 8 [1]. Our estimated degree of 37 seems to be plausible with the findings of Gupta et al. [34] given that this number is only the expected number of contacts and not the number of expected secondary infections. Also, from the number of analysed clusters from the reported data, the number of cases ranged between 0 and 40+. Our findings which were based on simulated and empirical data reveal that the results for the maximum-likelihood estimator is well suited for estimating tracing probability and the underlying contact structure. This finding reveals that the combination of results for the maximum-likelihood estimators of the basic model and contact tracing is well suited for better estimating the timing of infectious process and the underlying contact structure. In general, we expect this approach to improve estimating parameters in graph-based models.

5.2 Conclusion

Models are fundamentally all about capturing reality in the most simplified approximation, and the power of an epidemiological model choice lies within the nature of the epidemiological question. In this thesis, we have explored contact tracing on random graphs. Graphs are fundamentally mathematical representations of networks, and the study of epidemics using graph theory has increased our understanding of epidemiological processes from a microscopic to a macroscopic level.

The contact graph we did explore was a tree which is a special type of graph. In (infinite) homogeneous mixing populations, the graph is (somewhat) complete. There exist many forms of graph that deviate from a tree and a complete graph. From a microscopic level, trees are a simple decomposed form of complex graphs and an extensive study from this microscopic level can allow us to gain a better understanding of the overall mechanism and functioning of larger and more complex graphs. There is still so much ongoing research in developing robust models for constructing new emerging networks that deviates from well-known population structure with respect to spatial dynamics and time scales.

Appendix A

Appendix

A.1 Pair approximation

Proposition A.1.1. *The pair approximation of the contact process with one-step tracing is defined as*

$$\begin{aligned}
\frac{d}{dt}|[0]| &= -\tfrac{\beta}{K}|[01]| + g|[1]| + fg|[11]| \\
\frac{d}{dt}|[1]| &= \tfrac{\beta}{K}|[01]| - g|[1]| - fg|[11]| \\
\frac{d}{dt}|[00]| &= -\tfrac{2\beta}{K}|[001]| + 2g|[01]| + 2fg|[11]| + 2fg|[011]| \\
\frac{d}{dt}|[01]| &= \tfrac{\beta}{K}(|[001]| - |[101]| - |[01]|) - g|[01]| + g(1-f)|[11]| + fg(|[111]| - |[011]|) \\
\frac{d}{dt}|[11]| &= \tfrac{2\beta}{K}(|[101]| + |[01]|) - 2g|[11]| - 2fg|[111]|.
\end{aligned}$$

Proof. To derive ODE's for these configurations, we have to first derive the probability for any event to occur in a (small) time interval Δt. A switch of state from "0" to "1" and vice versa implies the occurence of an event. We also note that all configurations with a width that allows to read the changes in the number of pairs of various type has to be taking into account on the vector

$$([0], [1], \tfrac{1}{2}[0,0], [0,1], \tfrac{1}{2}[1,1])^T$$

From here, for a more precise disease state of an individual, we replace the notations "0" and "1" with "S" and "I" respectively. The strategy is that, we first determine the local configuration needed for an infection or a recovery event. Only afterwards, we can

determine the rate at which the event happens and the corresponding effect of this event on the state vector. In the end, we simplify the resulting equations.

For an infectious event, we require at least one susceptible and one infected individual in the neighbourhood. This event switches an "S" to "I". We find for this event, two possible neighbourhoods.

$$(S, S, I) \to (S, I, I)$$

$$(I, S, I) \to (I, I, I)$$

Since we require at least one infected individual in the neighbourhood, the event $(S, S, S) \to (S, I, S)$ is not possible for an infection event.

For a recovery event, we require at least one infected individual in the configuration. We also note that a susceptible individual stops the spread of detection of an infected individual via contact tracing. There is one focal individual (index case) that is detected and the following possibilities under recovery event is investigated.

1 The smallest configuration is given, if this focal individual has two susceptible neighbours. We find for this one configuration and one event

$$(S, I, S) \to (S, S, S)$$

With only two susceptible individuals in the neighbourhood of the infected individual, it means this individual recovers in the absence of contact tracing. That is, without contact tracing, the three neighbourhood is the only configuration to consider.

2 Now, we come to contact tracing, the next larger is when this focal individual has one susceptible neighbour on one side, and on the other side, has an infected neighbour, that in turn has a susceptible or an infected neighbour. We find for this two neighbourhood configurations: (S, I, I, S) and (I, I, I, S). For each configuration, we have the possibility that an infected individuals can recover without contact tracing or that, with contact tracing, some fraction, f of partners of the focal individual are treated simultaneously with the focal individual.

$$(S, I, \textcircled{I}, S) \begin{matrix} \to (S,I,S,S) \\ \to (S,S,S,S) \end{matrix}$$

$$(I, I, \textcircled{I}, S) \begin{matrix} \to (I,I,S,S) \\ \to (I,S,S,S) \end{matrix}$$

3 The next more complex configuration is the $I, \textcircled{I}, I$, where the central I is the recovering individual. This focal individual has two infected neighbours that in turn,

has another susceptible or infected neighbour. To count all pairs, we need to consider $(*I, \textcircled{I}, I*)$ where * is either S or I. We find for this, the neighbourhood below.

$$(S, I, \textcircled{I}, I, S) \begin{array}{l} \nearrow (S, I, S, I, S) \\ \rightarrow (S, I, S, S, S) \\ \rightarrow (S, S, S, I, S) \\ \searrow (S, S, S, S, S) \end{array}$$

$$(I, I, \textcircled{I}, I, S) \begin{array}{l} \nearrow (I, I, S, I, S) \\ \rightarrow (I, I, S, S, S) \\ \rightarrow (I, S, S, I, S) \\ \searrow (I, S, S, S, S) \end{array}$$

$$(I, I, \textcircled{I}, I, I) \begin{array}{l} \nearrow (I, I, S, I, I) \\ \rightarrow (I, I, S, S, I) \\ \rightarrow (I, S, S, I, I) \\ \searrow (I, S, S, S, I) \end{array}$$

We now discuss in detail every event of infection and recovery with one neighbourhood.

Probability for $(S, S, I) \rightarrow (S, I, I)$:

$$P\left((S, S, I) \rightarrow (S, I, I)\right) = [SSI] \, \tfrac{\beta}{K} \Delta t + \mathcal{O}(\Delta t)$$

The probability for the occurrence of this infectious event in a (small) time interval Δt is the number of configurations (S, S, I) multiplied by the transmission rate, β, divided by the number of sites neighbouring the central site in the triplet and multiplied by the length of the time interval. $1/K$ is the probability that an individual in a configuration is placed in the central site.

Also,

$$P\left((I, S, I) \rightarrow (I, I, I)\right) = \tfrac{1}{2} \, [ISI] \, \tfrac{2\beta}{K} \Delta t + \mathcal{O}(\Delta t).$$

We note that configurations with symmetry is divided by 2 to get the exact number of triplets in the configuration. Since the direction in which the disease can spread in this configuration is not specified, we have to take into account every possible way of infection. Consequently, the additional transmission rate in the $(I, S, I) \rightarrow (I, I, I)$ configuration stems from the fact that the central susceptible individual can be infected by any of its two infected neighbours.

For a recovery event in the configuration $(S, I, \textcircled{I}, S)$, there are two possible recovery events. First, the focal individual can be diagnosed without triggering a tracing event. Second, this focal individual triggers a tracing event that subsequently traces the other infected partner. Thus we have the probability for $(S, I, I, S) \to (S, I, S, S)$ and $(S, I, I, S) \to (S, I, S, S)$ to be

$$P\left((S, I, I, S) \to (S, I, S, S)\right) = \tfrac{1}{2}\left[SIIS\right] 2(1-f)g + \mathcal{O}(\Delta t)$$
$$P\left((S, I, I, S) \to (S, S, S, S)\right) = \tfrac{1}{2}\left[SIIS\right] 2fg + \mathcal{O}(\Delta t)$$

Similar to the infection event, the probability for the occurrence of this recovery event in a (small) time interval Δt is the number of configurations (S, I, I, S) multiplied by the recovery rate, g, multiplied by the length of the time interval. We take into account the additional symmetry $(S, I, \textcircled{I}, S) = (S, \textcircled{I}, I, S)$ and also note the focal individual in circle which will result in an additional recovery factor of two. As already discussed above, configurations without symmetry is counted once. Since other configuration has been derived in the same way, this argument applies to other configurations in this context. We show below the summary of probabilities of the whole process and its effects on singletons and pairs.

Event	Infection (a) $(S,S,I)\to(S,I,I)$	Infection (b) $(I,S,I)\to(I,I,I)$	Recovery (a) $(S,I,S)\to(S,S,S)$	Recovery (b) $(S,I,I,S)\to(S,I,S,S)$	Recovery (b) $(S,I,I,S)\to(S,S,S,S)$	Recovery (c) $(I,I,I,S)\to(I,I,S,S)$	Recovery (c) $(I,I,I,S)\to(I,S,S,S)$
Probabilities (up to higher order terms)	$[SSI]\,\frac{\beta}{K}\Delta t$	$\frac{1}{2}[ISI]\,\frac{2\beta}{K}\Delta t$	$\frac{1}{2}[SIS]\,g\Delta t$	$\frac{1}{2}[SIIS]\,2(1-f)g\Delta t$	$\frac{1}{2}[SIIS]\,2fg\Delta t$	$[IIIS]\,(1-f)g\Delta t$	$[IIIS]\,fg\Delta t$
Effect on							
$[S]$	-1	-1	$+1$	$+1$	$+2$	$+1$	$+2$
$[I]$	$+1$	$+1$	-1	-1	-1	-1	-2
$\frac{1}{2}[SS]$	-1	0	$+2$	$+1$	$+3$	$+1$	$+2$
$[SI]$	0	-2	-2	0	-2	0	0
$\frac{1}{2}[II]$	$+1$	$+2$	0	-1	-1	-1	-2

Event	Recovery (d) $(S,I,I,I,S)\to(S,I,S,I,S)$	Recovery (d) $(S,I,I,I,S)\to(S,I,S,S,S)$	Recovery (d) $(S,I,I,I,S)\to(S,S,S,S,S)$	Recovery (e) $(I,I,I,I,S)\to(I,I,S,I,S)$	Recovery (e) $(I,I,I,I,S)\to(I,I,S,S,S)$	Recovery (e) $(I,I,I,I,S)\to(I,S,S,I,S)$	Recovery (e) $(I,I,I,I,S)\to(I,S,S,S,S)$
Probabilities (up to higher order terms)	$\frac{1}{2}[SIIIS]\,(1-f)^2g\Delta t$	$\frac{1}{2}[SIIIS]\,2f(1-f)g\Delta t$	$\frac{1}{2}[SIIIS]\,f^2g\Delta t$	$[IIIIS]\,(1-f)^2g\Delta t$	$[IIIIS]\,f(1-f)g\Delta t$	$[IIIIS]\,f(1-f)g\Delta t$	$[IIIIS]\,f^2g\Delta t$
Effect on							
$[S]$	$+1$	$+2$	$+3$	$+1$	$+2$	$+2$	$+3$
$[I]$	-1	-2	-3	-1	-2	-2	-3
$\frac{1}{2}[SS]$	0	$+2$	$+4$	0	$+2$	$+1$	$+3$
$[SI]$	$+1$	0	-2	$+2$	0	$+2$	0
$\frac{1}{2}[II]$	-2	-2	-2	-2	-2	-3	-3

Event	Recovery (f) $(I,I,I,I,I)\to(I,I,S,I,I)$	Recovery (f) $(I,I,I,I,I)\to(I,I,S,S,I)$	Recovery (f) $(I,I,I,I,I)\to(I,S,S,S,I)$
Probabilities (up to higher order terms)	$\frac{1}{2}[IIIII]\,(1-f)^2g\Delta t$	$\frac{1}{2}[IIIII]\,2f(1-f)g\Delta t$	$\frac{1}{2}[IIIII]\,f^2g\Delta t$
Effect on			
$[S]$	$+1$	$+2$	$+3$
$[I]$	-1	-2	-3
$\frac{1}{2}[SS]$	0	$+1$	$+2$
$[SI]$	$+2$	$+2$	$+2$
$\frac{1}{2}[II]$	-2	-3	-4

Table A.1: Summary of the probabilities and effects of infection and recovery events. Contact tracing has been taking into account in the recovery process.

From table A.1, it is now straight forward to write out and derive the differential equations

for the effect of each probabilities on singletons and pairs in all events. The expected values of the number of pairs and singletons in a certain configuration satisfy.

$$
\begin{aligned}
|[\dot{S}]| = & -|[SSI]|\tfrac{\beta}{K} - \tfrac{1}{2}|[ISI]|\tfrac{2\beta}{K} + \tfrac{1}{2}|[SIS]|g + 2\tfrac{1}{2}|[SIIS]|(1-f)g + 2\tfrac{1}{2}|[SIIS]|fg \\
& + |[IIIS]|(1-f)fg + 2|[IIIS]|fg + \tfrac{1}{2}|[SIIIS]|(1-f)^2g + 2\tfrac{1}{2}|[SIIIS]|(1-f)2fg \\
& + 3\tfrac{1}{2}|[SIIIS]|f^2g + |[IIIIS]|(1-f)^2g + 2|[IIIIS]|(1-f)fg + 2|[IIIIS]|(1-f)fg \\
& + 3|[IIIIS]|f^2g + \tfrac{1}{2}|[IIIII]|(1-f)^2g + 2\tfrac{1}{2}|[IIIII]|(1-f)2fg + 3\tfrac{1}{2}|[IIIII]|2f^2g \\
|[\dot{I}]| = & |[SSI]|\tfrac{\beta}{K} + \tfrac{1}{2}|[ISI]|\tfrac{2\beta}{K} - \tfrac{1}{2}|[SIS]|g - 2\tfrac{1}{2}|[SIIS]|(1-f)g - 2\tfrac{1}{2}|[SIIS]|fg \\
& - |[IIIS]|(1-f)fg - 2|[IIIS]|fg - \tfrac{1}{2}|[SIIIS]|(1-f)^2g - 2\tfrac{1}{2}|[SIIIS]|(1-f)2fg \\
& - 3\tfrac{1}{2}|[SIIIS]|f^2g - |[IIIIS]|(1-f)^2g - 2|[IIIIS]|(1-f)fg - 2|[IIIIS]|(1-f)fg \\
& - 3|[IIIIS]|f^2g - \tfrac{1}{2}|[IIIII]|(1-f)^2g - 2\tfrac{1}{2}|[IIIII]|(1-f)2fg - 3\tfrac{1}{2}|[IIIII]|2f^2g \\
\tfrac{1}{2}|[\dot{SS}]| = & -|[SSI]|\tfrac{\beta}{K} + 2\tfrac{1}{2}|[SIS]|g + 2\tfrac{1}{2}|[SIIS]|(1-f)g + 2\tfrac{1}{2}|[SIIS]|fg \\
& + |[IIIS]|(1-f)fg + 2|[IIIS]|fg + 2\tfrac{1}{2}|[SIIIS]|(1-f)2fg \\
& + 4\tfrac{1}{2}|[SIIIS]|f^2g + 2|[IIIIS]|(1-f)fg + 2|[IIIIS]|(1-f)fg \\
& + 3|[IIIIS]|f^2g + 2\tfrac{1}{2}|[IIIII]|(1-f)fg + 2\tfrac{1}{2}|[IIIII]|f^2g \\
\tfrac{1}{2}|[\dot{II}]| = & -|[SSI]|\tfrac{\beta}{K} + 2\tfrac{1}{2}|[ISI]|\tfrac{2\beta}{K} - 2\tfrac{1}{2}|[SIIS]|(1-f)g - 2\tfrac{1}{2}|[SIIS]|fg \\
& - |[IIIS]|(1-f)fg - 2|[IIIS]|fg - 2\tfrac{1}{2}|[SIIIS]|(1-f)^2g - 2.2\tfrac{1}{2}|[SIIIS]|(1-f)fg \\
& - 2\tfrac{1}{2}|[SIIIS]|f^2g - 2|[IIIIS]|(1-f)^2g - 2|[IIIIS]|(1-f)fg - 3|[IIIIS]|(1-f)fg \\
& - 3|[IIIIS]|f^2g - 2\tfrac{1}{2}|[IIIII]|(1-f)^2g - 3.2\tfrac{1}{2}|[IIIII]|(1-f)fg \\
& - 4\tfrac{1}{2}|[IIIII]|f^2g \\
|[\dot{SI}]| = & \tfrac{1}{2}|[\dot{SS}]| - \tfrac{1}{2}|[\dot{II}]|
\end{aligned}
\tag{A.1.1}
$$

For the $|[\dot{SI}]|$, we claim that $\frac{1}{2}|[\dot{SS}]| + \frac{1}{2}|[\dot{II}]|$ is the total number of pairs on the graph, which is constant. By expanding and simplifying the system of differential equations above gives the expectations of the number of pairs and singletons and we resolve to

$$
\begin{aligned}
|[\dot{S}]| = & -\tfrac{\beta}{K}|[SI]| + g|[I]| + fg|[II]| \\
|[\dot{I}]| = & \tfrac{\beta}{K}|[SI]| - g|[I]| - fg|[II]| \\
|[\dot{SS}]| = & -\tfrac{2\beta}{K}|[SSI]| + 2g|[SI]| + 2fg|[II]| + 2fg|[SII]| \\
|[\dot{SI}]| = & -\tfrac{\beta}{K}(|[SSI]| - |[ISI]|) - \tfrac{\beta}{K}|[SI]| - g|[SI]| + g(1-f)|[II]| + fg(|[III]| - |[SII]|) \\
|[\dot{II}]| = & -\tfrac{2\beta}{K}|[SI]| + \tfrac{2\beta}{K}|[ISI]| - 2g|[II]| - 2fg|[III]|,
\end{aligned}
\tag{A.1.2}
$$

which follows the claim. We note that the triplet configurations has also been simplified in terms of singles and pairs following same rules in Def. 2.4.6.

for $f = 0$, we have the pair approximation without one step contact tracing

$$\begin{aligned}
\frac{d}{dt}|[0]| &= -\tfrac{\beta}{K}|[01]| + g|[1]| \\
\frac{d}{dt}|[1]| &= \tfrac{\beta}{K}|[01]| - g|[1]| \\
\frac{d}{dt}|[00]| &= -\tfrac{2\beta}{K}|[001]| + 2g|[01]| \\
\frac{d}{dt}|[01]| &= \tfrac{\beta}{K}(|[001]| - |[101]| - |[01]|) - g(|[01]| - |[11]|) \\
\frac{d}{dt}|[11]| &= \tfrac{2\beta}{K}(|[101]| + |[01]|) - 2g|[11]|.
\end{aligned}$$

□

A.2 Second order approximation

A.2.1 Backward tracing - recursive

Proposition A.2.1. *For all moments of K finite, the second order approximation of backward recursive tracing $\kappa_r^-(a)$ in p for $\beta \neq \alpha + \sigma$ is given as*

$$\begin{aligned}
\kappa_r^-(a) &= \hat{\kappa}(a) - p\hat{\kappa}(a)\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right)\right) \\
&- \frac{p^2}{2}\hat{\kappa}(a)\Bigg(\mathbb{E}[K]^2\Bigg(\frac{1 - e^{(\alpha+\sigma+\beta)a}(-\sigma - \frac{\alpha\sigma}{\alpha+\sigma-\beta}}{\alpha+\sigma+\beta} + \frac{\left(1-\hat{\kappa}(a)^2\right)\left(\frac{\alpha\sigma}{(\alpha+\sigma-\beta)} + \sigma\left(1+p_{obs}\right)\right)}{2(\alpha+\sigma)} \\
&+ \frac{(1-\hat{\kappa}(a))\left(-\sigma p_{obs} + \sigma e^{-\beta a} + \frac{\alpha\sigma}{\alpha+\sigma-\beta}e^{-\beta a}\right)}{\alpha+\sigma} + \frac{\left(e^{-\beta a} - \hat{\kappa}(a)\right)\left(-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)} + p_{obs}\sigma\right)}{\alpha+\sigma-\beta} \\
&+ \frac{\left(e^{-\beta a} - \hat{\kappa}(a)^2\right)\left(-\sigma\left(1+p_{obs}\right)\right)}{2\alpha+2\sigma-\beta}\Bigg)\Bigg) \\
&+ \frac{p^2}{2}\hat{\kappa}(a)(\mathrm{Var}(K) + \mathbb{E}[K]\big(\mathbb{E}[K]-1\big))\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma\left(e^{-\beta a} - \hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2 + \mathcal{O}\left(p^3\right).
\end{aligned}$$

Proof. We apply the taylor's expansion with respect to p up to the second degree

$$\kappa_r^-(a) = \sum_{k=0}^{\infty} \frac{p^k}{k!}\kappa_r^{(k)-}(a)$$

in Theorem. 3.2.1.

$$
\begin{aligned}
\kappa_r^-(a) &= \kappa_r^{(0)-}(a) + p\kappa_r^{(1)-}(a) + \frac{p^2}{2}\kappa_r^{(2)-}(a) + \mathcal{O}\left(p^3\right) \\
&= e^{-(\alpha+\sigma)a}\mathcal{G}\left(1 - p\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\right. \\
&\qquad \left.\left(-\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a}) + p\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a})\right) + \mathcal{O}\left(p^2\right)\right) d\tilde{a}\right) \\
&= e^{-(\alpha+\sigma)a}\mathbb{E}\left(\left(1 - p\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\right.\right. \\
&\qquad \left.\left.\left(-\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a}) + p\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a})\right) + \mathcal{O}\left(p^2\right)\right) d\tilde{a}\right)^K\right) \\
&= e^{-(\alpha+\sigma)a}\mathbb{E}\left(1^K - K\left(\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\right.\right. \\
&\qquad \left.\left(-p\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - p\alpha\kappa_r^{(0)-}(\tilde{a}) + p^2\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a})\right) + \mathcal{O}\left(p^3\right)\right) d\tilde{a}\right)\right) \\
&\qquad +K(K-1)\left(p^2\left(\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\right.\right. \\
&\qquad \left.\left.\left(-\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a}) + p\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a})\right)\right) d\tilde{a}\right)^2 + \mathcal{O}\left(p^3\right)\right) \\
&\qquad +\mathbb{E}[K(K-1)]\left(p^2\left(\int_0^a -\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a})\left(1 - e^{(-\beta(a-\tilde{a})}\right) d\tilde{a} + \mathcal{O}(p)\right)^2\right) \\
&= \hat{\kappa}(a)(\mathbb{E}[1] - \mathbb{E}[K]\left(\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right)\right. \\
&\qquad \left.\left(-p\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - p\alpha\kappa_r^{(0)-}(\tilde{a}) + p^2\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a})\right) + \mathcal{O}\left(p^3\right) d\tilde{a}\right)\right) \\
&\qquad +\mathbb{E}[K(K-1)]\left(p^2\left(\int_0^a -\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a})\left(1 - e^{(-\beta(a-\tilde{a})}\right) d\tilde{a} + \mathcal{O}(p)\right)^2\right).
\end{aligned}
$$

We apply the binomial theorem in the 4th equality and the linearity of the expectation in the 5th equality. We can use the binomial theorem in our case, because K random variable denotes the number of downstream edges.

Equating the powers of p from the last equation gives

$$\begin{aligned}
p^0 : \kappa_r^{(0)-}(a) &= \hat{\kappa}(a)\mathbb{E}(1) = \hat{\kappa}(a) \\
p^1 : \kappa_r^{(1)-}(a) &= \hat{\kappa}(a)\left(-\mathbb{E}[K]\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\left(-\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a})\right)d\tilde{a}\right) \\
p^2 : \frac{\kappa_r^{(2)-}(a)}{2} &= \hat{\kappa}(a)\left(-\mathbb{E}[K]\left(\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a})\right)d\tilde{a}\right)\right) \\
&\quad +\mathbb{E}[K(K-1)]\left(\left(\int_0^a \left(-\frac{d}{da}\kappa_r^{(0)-}(\tilde{a}) - \alpha\kappa_r^{(0)-}(\tilde{a})\right)\left(1-e^{(-\beta(a-\tilde{a})}\right)d\tilde{a}\right)^2\right).
\end{aligned} \tag{A.2.1}$$

Since we already calculated the p^1-term as

$$\kappa_r^{(1)-}(a) = -\hat{\kappa}(a)\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma}{\alpha+\sigma-\beta}\left(e^{-\beta a}-\hat{\kappa}(a)\right)\right).$$

We can move to the p^2-term and simplify the terms that appear in the integral. We start with the derivative of $\kappa_r^{(1)-}(a)$ as follows:

$$\begin{aligned}
\frac{d}{da}\kappa_r^{(1)-}(a) &= (\alpha+\sigma)\hat{\kappa}(a)\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma}{(\alpha+\sigma-\beta)}\left(e^{-\beta a}-\hat{\kappa}(a)\right)\right) \\
&\quad -\hat{\kappa}(a)\mathbb{E}[K]\left(\sigma\hat{\kappa}(a) + \frac{\sigma\beta e^{-\beta a} - (\alpha+\sigma)\sigma\hat{\kappa}(a)}{(\alpha+\sigma-\beta)}\right) \\
&= \hat{\kappa}(a)\mathbb{E}[K]\left(\sigma e^{-\beta a} + p_{obs}(\alpha - \alpha\hat{\kappa}(a)+\sigma) - \sigma\hat{\kappa}(a)\left(1+p_{obs}\right)\right)
\end{aligned}$$

and the subtraction inside the integral as

$$\begin{aligned}
& -\frac{d}{da}\kappa_r^{(1)-}(\tilde{a}) - \alpha\kappa_r^{(1)-}(\tilde{a}) \\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]\left(\sigma e^{-\beta a} + p_{obs}(\alpha-\alpha\hat{\kappa}(a)+\sigma) - \sigma\hat{\kappa}(a)\left(1+p_{obs}\right)\right) + \alpha\mathbb{E}[K]\hat{\kappa}(a)\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{\sigma-\beta+\alpha}\right) \\
= \; & -\mathbb{E}[K]\sigma\hat{\kappa}(a)e^{-a\beta} - \mathbb{E}[K]\hat{\kappa}(a)\left(p_{obs}(-\alpha\hat{\kappa}(a)+\sigma+\alpha) - \left(p_{obs}+1\right)\sigma\hat{\kappa}(a)\right) - \frac{\mathbb{E}[K]\alpha\sigma\left(e^{-a\beta}-\hat{\kappa}(a)\right)\hat{\kappa}(a)}{\sigma-\beta+\alpha} \\
& \quad +\mathbb{E}[K]\alpha p_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a) \\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]\sigma e^{-\beta a} - \frac{\alpha\hat{\kappa}(a)\mathbb{E}[K]\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)} - \mathbb{E}[K]\hat{\kappa}(a)p_{obs}\sigma + \hat{\kappa}(a)^2\mathbb{E}[K]\sigma\left(1+p_{obs}\right).
\end{aligned}$$

From the first term of the p^2-term in Eqn. A.2.1

$$
\begin{aligned}
& -\hat{\kappa}(a)\mathbb{E}[K]\left(\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a})-\alpha\kappa_r^{(1)-}(\tilde{a})\right)d\tilde{a}\right)\\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]\left(\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\right.\\
& \left.\left(-\hat{\kappa}(\tilde{a})\mathbb{E}[K]\sigma e^{-\beta\tilde{a}}-\frac{\alpha\hat{\kappa}(\tilde{a})\mathbb{E}[K]\sigma\left(e^{-\beta\tilde{a}}-\hat{\kappa}(\tilde{a})\right)}{(\alpha+\sigma-\beta)}-\mathbb{E}[K]\hat{\kappa}(\tilde{a})p_{obs}\sigma+\hat{\kappa}(\tilde{a})^2\mathbb{E}[K]\sigma\left(1+p_{obs}\right)\right)d\tilde{a}\right)\\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]\left(\int_0^a\left(-\hat{\kappa}(\tilde{a})\mathbb{E}[K]\sigma e^{-\beta\tilde{a}}-\frac{\alpha\hat{\kappa}(\tilde{a})\mathbb{E}[K]\sigma\left(e^{-\beta\tilde{a}}-\hat{\kappa}(\tilde{a})\right)}{(\alpha+\sigma-\beta)}-\mathbb{E}[K]\hat{\kappa}(\tilde{a})p_{obs}\sigma+\hat{\kappa}(\tilde{a})^2\mathbb{E}[K]\sigma\left(1+p_{obs}\right)\right)d\tilde{a}\right.\\
& -\int_0^a\left(-e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})\mathbb{E}[K]\sigma e^{-\beta\tilde{a}}-\frac{e^{-\beta(a-\tilde{a})}\alpha\hat{\kappa}(\tilde{a})\mathbb{E}[K]\sigma\left(e^{-\beta\tilde{a}}-\hat{\kappa}(\tilde{a})\right)}{(\alpha+\sigma-\beta)}-e^{-\beta(a-\tilde{a})}\mathbb{E}[K]\hat{\kappa}(\tilde{a})p_{obs}\sigma\right.\\
& \left.\left.+e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})^2\mathbb{E}[K]\sigma\left(1+p_{obs}\right)\right)d\tilde{a}\right)\\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]\left(-\mathbb{E}[K]\sigma\int_0^a\hat{\kappa}(\tilde{a})e^{-\beta\tilde{a}}d\tilde{a}-\frac{\alpha\mathbb{E}[K]\sigma}{(\alpha+\sigma-\beta)}\int_0^a\hat{\kappa}(\tilde{a})\left(e^{-\beta\tilde{a}}-\hat{\kappa}(\tilde{a})\right)d\tilde{a}-\mathbb{E}[K]p_{obs}\sigma\int_0^a\hat{\kappa}(\tilde{a})d\tilde{a}\right.\\
& +\mathbb{E}[K]\sigma\left(1+p_{obs}\right)\int_0^a\hat{\kappa}(\tilde{a})^2d\tilde{a}+\mathbb{E}[K]\sigma\int_0^a e^{-\beta a}\hat{\kappa}(\tilde{a})d\tilde{a}+\frac{\alpha\mathbb{E}[K]\sigma}{(\alpha+\sigma-\beta)}\int_0^a\hat{\kappa}(\tilde{a})e^{-\beta a}-e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})d\tilde{a}\\
& \left.+\mathbb{E}[K]p_{obs}\sigma\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})d\tilde{a}-\mathbb{E}[K]\sigma\left(1+p_{obs}\right)\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})^2d\tilde{a}\right)\\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]^2\left(\int_0^a\hat{\kappa}(\tilde{a})e^{-\beta\tilde{a}}d\tilde{a}\left(-\sigma-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}\right)+\int_0^a\hat{\kappa}(\tilde{a})^2d\tilde{a}\left(\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}+\sigma\left(1+p_{obs}\right)\right)\right.\\
& +\int_0^a\hat{\kappa}(\tilde{a})d\tilde{a}\left(-p_{obs}\sigma+\sigma e^{-\beta a}+\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}e^{-\beta a}\right)+\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})d\tilde{a}\left(-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}+p_{obs}\sigma\right)\\
& \left.+\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})^2d\tilde{a}\left(-\sigma\left(1+p_{obs}\right)\right)\right).
\end{aligned}
$$

Consider the following integrals;

$$
\begin{aligned}
\int_0^a\hat{\kappa}(\tilde{a})e^{-\beta\tilde{a}}d\tilde{a} &= \frac{1-e^{-a(\alpha+\sigma+\beta)}}{\alpha+\sigma+\beta}\\
\int_0^a\hat{\kappa}(\tilde{a})d\tilde{a} &= \frac{1-\hat{\kappa}(a)}{\alpha+\sigma}\\
\int_0^a\hat{\kappa}(\tilde{a})^2d\tilde{a} &= \frac{1-\hat{\kappa}(a)^2}{2(\alpha+\sigma)}\\
\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})d\tilde{a} &= \frac{e^{-\beta a}-\hat{\kappa}(a)}{\alpha+\sigma-\beta}\\
\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})^2d\tilde{a} &= \frac{e^{-\beta a}-\hat{\kappa}(a)^2}{2\alpha+2\sigma-\beta}.
\end{aligned}
$$

we get the following result for the first term of the p^2- term in Eqn. A.2.1

$$
\begin{aligned}
& -\hat{\kappa}(a)\mathbb{E}[K]\left(\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\left(-\frac{d}{da}\kappa_r^{(1)-}(\tilde{a})-\alpha\kappa_r^{(1)-}(\tilde{a})\right)d\tilde{a}\right) \\
= \; & -\hat{\kappa}(a)\mathbb{E}[K]^2\left(\frac{\left(1-e^{-a(\alpha+\sigma+\beta)}\right)\left(-\sigma-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}\right)}{\alpha+\sigma+\beta}+\frac{(1-\hat{\kappa}(a)^2)\left(\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}+\sigma\left(1+p_{obs}\right)\right)}{2(\alpha+\sigma)}\right. \\
& +\frac{(1-\hat{\kappa}(a))\left(-p_{obs}\sigma+\sigma e^{-\beta a}+\frac{\alpha\sigma}{\alpha+\sigma}-\beta a\right)}{(\alpha+\sigma-\beta)}e^{-\beta a}+\frac{\left(e^{-\beta a}-\hat{\kappa}(a)\right)\left(-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}+p_{obs}\sigma\right)}{\alpha+\sigma-\beta} \\
& \left.+\frac{\left(e^{-\beta a}-\hat{\kappa}(a)^2\right)\left(-\sigma\left(1+p_{obs}\right)\right)}{2\alpha+2\sigma-\beta}\right).
\end{aligned}
$$

From the second term of the p^2- term in Eqn. A.2.1

$$
\begin{aligned}
& \hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(\int_0^a -\frac{d}{da}\kappa_r^{(0)-}(\tilde{a})-\alpha\kappa_r^{(0)-}(\tilde{a})\left(1-e^{(-\beta(a-\tilde{a})}\right)d\tilde{a}+\mathcal{O}(p)\right)^2 \\
= \; & \hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(\sigma\int_0^a \hat{\kappa}(\tilde{a})\left(1-e^{-\beta(a-\tilde{a})}\right)d\tilde{a}+\mathcal{O}(p)\right)^2 \\
= \; & \hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(\sigma\int_0^a \hat{\kappa}(\tilde{a})d\tilde{a}-\sigma\int_0^a \hat{\kappa}(\tilde{a})e^{-\beta(a-\tilde{a})}\right)d\tilde{a}+\mathcal{O}(p)\Bigg)^2 \\
= \; & \hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2.
\end{aligned}
$$

We obtain the p^2- term as

$$
\begin{aligned}
& \frac{\kappa_{r,i}^{(2)-}(a)}{2} \\
= & -\hat{\kappa}(a)\Bigg(\quad \mathbb{E}[K]^2\left(\frac{\left(1-e^{-a(\alpha+\sigma+\beta)}\right)\left(-\sigma-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}\right)}{\alpha+\sigma+\beta}+\frac{\left(1-\hat{\kappa}(a)^2\right)\left(\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}+\sigma(1+p_{obs})\right)}{2(\alpha+\sigma)}\right. \\
& +\frac{(1-\hat{\kappa}(a))\left(-p_{obs}\sigma+\sigma e^{-\beta a}+\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}e^{-\beta a}\right)}{\alpha+\sigma}+\frac{\left(e^{-\beta a}-\hat{\kappa}(a)\right)\left(-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)}+p_{obs}\sigma\right)}{\alpha+\sigma-\beta} \\
& \left.\left.+\frac{\left(e^{-\beta a}-\hat{\kappa}(a)^2\right)(-\sigma(1+p_{obs}))}{2\alpha+2\sigma-\beta}\right)\right) \\
& +\hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2.
\end{aligned}
$$

By using the root idea from the beginning and the properties of the probability generation function,

$$
\begin{aligned}
\kappa_{r,i}^{-}(a) &= \kappa_{r,i}^{(0)-}(a) + p\kappa_{r,i}^{(1)-}(a) + \frac{p^2}{2}\kappa_{r,i}^{(2)-}(a) + \mathcal{O}\left(p^3\right) \\
&= \hat{\kappa}(a) - p\hat{\kappa}(a)\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right)\right) \\
&\quad + \frac{p^2}{2}\left(-\hat{\kappa}(a)\left(\mathbb{E}[K]^2\left(\frac{\left(1-e^{-a(\alpha+\sigma+\beta)}\right)\left(-\sigma - \frac{\alpha\sigma}{(\alpha+\sigma-\beta)}\right)}{\alpha+\sigma+\beta} + \frac{\left(1-\hat{\kappa}(a)^2\right)\left(\frac{\alpha\sigma}{(\alpha+\sigma-\beta)} + \sigma\left(1+p_{obs}\right)\right)}{2(\alpha+\sigma)}\right.\right.\right. \\
&\qquad + \frac{(1-\hat{\kappa}(a))\left(-p_{obs}\sigma + \sigma e^{-\beta a} + \frac{\alpha\sigma}{(\alpha+\sigma-\beta)}e^{-\beta a}\right)}{\alpha+\sigma} + \frac{\left(e^{-\beta a} - \hat{\kappa}(a)\right)\left(-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)} + p_{obs}\sigma\right)}{\alpha+\sigma-\beta} \\
&\qquad \left.\left. + \frac{\left(e^{-\beta a} - \hat{\kappa}(a)^2\right)\left(-\sigma\left(1+p_{obs}\right)\right)}{2\alpha+2\sigma-\beta}\right)\right) \\
&\qquad + \hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma\left(e^{-\beta a} - \hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2\Bigg) + \mathcal{O}\left(p^3\right) \\
&= \hat{\kappa}(a) - p\hat{\kappa(a)}\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right)\right) \\
&\quad - \frac{p^2}{2}\hat{\kappa}(a)\left(\mathbb{E}[K]^2\left(\frac{\left(1-e^{-a(\alpha+\sigma+\beta)}\right)\left(-\sigma - \frac{\alpha\sigma}{(\alpha+\sigma-\beta)}\right)}{\alpha+\sigma+\beta} + \frac{\left(1-\hat{\kappa}(a)^2\right)\left(\frac{\alpha\sigma}{(\alpha+\sigma-\beta)} + \sigma\left(1+p_{obs}\right)\right)}{2(\alpha+\sigma)}\right.\right. \\
&\qquad + \frac{(1-\hat{\kappa}(a))\left(-p_{obs}\sigma + \sigma e^{-\beta a} + \frac{\alpha\sigma}{(\alpha+\sigma-\beta)}e^{-\beta a}\right)}{\alpha+\sigma} + \frac{\left(e^{-\beta a} - \hat{\kappa}(a)\right)\left(-\frac{\alpha\sigma}{(\alpha+\sigma-\beta)} + p_{obs}\sigma\right)}{\alpha+\sigma-\beta} \\
&\qquad \left.\left. + \frac{\left(e^{-\beta a} - \hat{\kappa}(a)^2\right)\left(-\sigma\left(1+p_{obs}\right)\right)}{2\alpha+2\sigma-\beta}\right)\right) \\
&\qquad + \frac{p^2}{2}\hat{\kappa}(a)(\mathrm{Var}[K] + \mathbb{E}[K](\mathbb{E}(K)-1))\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma\left(e^{-\beta a} - \hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2 + \mathcal{O}\left(p^3\right).
\end{aligned}
$$

□

A.2.2 Backward tracing - one step

Proposition A.2.2. *The second order approximation of κ_o^- is*

$$\begin{aligned}
\kappa_o^-(a) &= \hat{\kappa}(a) - p\hat{\kappa}(a)\mathbb{E}[K]\sigma\left(\frac{\kappa(a)-1}{-(\alpha+\sigma)} + \frac{\hat{\kappa}(a)-e^{-\beta a}}{\alpha+\sigma-\beta}\right) \\
&+ \frac{p^2}{2}(\hat{\kappa}(a)\left(\mathbb{E}[K]^2\sigma^2\left(-(\alpha+\sigma)\left(\frac{(1-\hat{\kappa}(a)^2)}{2(\alpha+\sigma)} - \frac{1-\hat{\kappa}(a)}{(\alpha+\sigma)}\right)\right.\right. \\
&+ (\alpha+\sigma-\beta)\left(\frac{1-\hat{\kappa}(a)^2}{2(\alpha+\sigma)} - \frac{1-\hat{\kappa}(a)e^{-\beta a}}{(\alpha+\sigma+\beta)}\right) \\
&+ (\alpha+\sigma)\left(\frac{e^{-a\beta}-\hat{\kappa}(a)^2}{2\sigma-\beta+2\alpha} - \frac{e^{-\beta a}-\hat{\kappa}(a)}{(\alpha+\sigma-\beta)}\right) \\
&- \left.(\alpha+\sigma-\beta)\left(\frac{e^{a\beta}-\hat{\kappa}(a)^2}{2\sigma-\beta+2\alpha} - \frac{e^{-\beta a}-\hat{\kappa}(a)e^{-\beta a}}{(\alpha+\sigma)}\right)\right) \\
&+ \left.\hat{\kappa}(a)\sigma^2(\mathrm{Var}[K]+\mathbb{E}[K(K-1)])\left(\frac{1-\hat{\kappa}(a)^2}{2(\alpha+\sigma)} - 2\frac{\left(e^{-a\beta}-\hat{\kappa}(a)^2\right)}{(2\sigma-\beta+2\alpha)} + \frac{e^{-2a\beta}-\hat{\kappa}(a)^2}{2(\alpha-\beta+\sigma)}\right)\right)
\end{aligned}$$

Proof. The main idea is again to obtain the second order approximation by using the Taylor expansion as follows

$$\begin{aligned}
&\sum_{k=0}^{\infty}\frac{p^k}{k!}\kappa_0^{(k)-}(a) \\
&= \hat{\kappa}(a)\mathcal{G}\left(1-p\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\sum_{k=0}^{\infty}\frac{p^k}{k!}\kappa_0^{(k)-}(\tilde{a})d\tilde{a}\right) \\
&= \hat{\kappa}(a)\mathbb{E}\left[\left(1-p\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\sum_{k=0}^{\infty}\frac{p^k}{k!}\kappa_0^{(k)-}(\tilde{a})d\tilde{a}\right)^K\right] \\
&= \hat{\kappa}(a)\mathbb{E}\left[1^K - Kp\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\sum_{k=0}^{\infty}\frac{p^k}{k!}\kappa_0^{(k)-}(\tilde{a})d\tilde{a}\right. \\
&\qquad \left.+K(K-1)p^2\left(\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\left(\sum_{k=0}^{2}\frac{p^k}{k!}\kappa_0^{(k)-}(\tilde{a})+\mathcal{O}\left(p^3\right)\right)d\tilde{a}\right)^2+\mathcal{O}\left(p^3\right)\right] \\
&= \hat{\kappa}(a)\mathbb{E}\left[1^K\right] - \hat{\kappa}(a)\mathbb{E}[K]\left(p\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\left(\kappa_0^{(0)-}(\tilde{a})+p\kappa_0^{(1)-}(\tilde{a})+\mathcal{O}\left(p^2\right)\right)d\tilde{a}\right) \\
&\qquad +\hat{\kappa}(a)\mathbb{E}[K(K-1)]p^2\left(\int_0^a\left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\left(\kappa_0^{(0)-}(\tilde{a})+\mathcal{O}\left(p^2\right)\right)d\tilde{a}\right)^2+\mathcal{O}\left(p^3\right)
\end{aligned}$$

As before we can again read the p-terms

$$
\begin{aligned}
p^0\text{-term: } \kappa_o^{(0)-}(\tilde{a}) &= \hat{\kappa}(a)\mathbb{E}[1] = \hat{\kappa}(a) \\
p^1\text{-term: } \kappa_o^{(1)-}(\tilde{a}) &= \hat{\kappa}(a)\Big(-\mathbb{E}[K] \int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right) \sigma \left(\kappa_o^{(0)-}(\tilde{a}) d\tilde{a}\right)\Big) \\
p^2\text{-term: } \kappa_o^{(2)-}(\tilde{a}) &= \hat{\kappa}(a)\Big(-\mathbb{E}[K] \int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right) \sigma \kappa_0^{(1)-}(\tilde{a}) d\tilde{a} \\
&\qquad + \mathbb{E}[K(K-1)] \left(\int_0^a \left(1 - e^{-\beta(a-\tilde{a})}\right) \sigma \kappa_o^{(0)-}(\tilde{a}) d\tilde{a} \right)^2 \Big)
\end{aligned}
$$

We calculate the p^1-term by inserting the $\kappa_0^{(0)-}(a)$ into the definition of $\kappa_0^{(1)-}(a)$ and obtain:

$$
\begin{aligned}
\kappa_o^{(1)-}(a) &= -\hat{\kappa}(a)\mathbb{E}(K)\sigma \left(\frac{\hat{\kappa}(a) - 1}{-(\alpha + \sigma)} + \frac{\hat{\kappa}(a) - e^{-\beta a}}{\alpha + \sigma - \beta} \right) \\
&= -\mathbb{E}[K]\sigma \frac{\left(\hat{\kappa}(a)^2 - \hat{\kappa}(a)\right)(\alpha + \sigma - \beta) + \left(\hat{\kappa}(a)^2 - \hat{\kappa}(a)e^{-\beta a}\right)(-(\alpha + \sigma))}{-(\alpha + \sigma)(\alpha - \beta + \sigma)} \\
&= \mathbb{E}[K]\sigma \frac{-\hat{\kappa}(a)^2\beta - \hat{\kappa}(a)(\alpha + \sigma - \beta) + e^{-\beta a}\hat{\kappa}(a)(\alpha + \sigma)}{(\alpha + \sigma)(\alpha - \beta + \sigma)}
\end{aligned}
$$

and some pre-calculations to insert in the p^2- term later:

$$
\begin{aligned}
&\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\kappa_0^{(1)-}(\tilde{a})d\tilde{a}\\
&=\sigma\int_0^a \kappa_0^{(1)-}(\tilde{a})d\tilde{a}-\sigma\int_0^a e^{-\beta(a-\tilde{a})}\kappa_0^{(1)-}(\tilde{a})d\tilde{a}\\
&=\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\int_0^a -\hat{\kappa}(\tilde{a})^2\beta-\hat{\kappa}(\tilde{a})(\alpha+\sigma-\beta)+e^{-\beta\tilde{a}}\hat{\kappa}(\tilde{a})(\alpha+\sigma)d\tilde{a}\\
&-\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\int_0^a e^{-\beta(a-\tilde{a})}\left(-\hat{\kappa}(\tilde{a})^2\beta-\hat{\kappa}(\tilde{a})(\alpha+\sigma-\beta)+e^{-\beta\tilde{a}}\hat{\kappa}(\tilde{a})(\alpha+\sigma)\right)d\tilde{a}\\
&=\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(-\beta\int_0^a \hat{\kappa}(\tilde{a})^2d\tilde{a}-(\alpha+\sigma-\beta)\int_0^a \hat{\kappa}(\tilde{a})d\tilde{a}+(\alpha+\sigma)\int_0^a e^{-\beta\tilde{a}}\hat{\kappa}(\tilde{a})d\tilde{a}\right)\\
&-\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(-\beta\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})^2d\tilde{a}-(\alpha+\sigma-\beta)\int_0^a e^{-\beta(a-\tilde{a})}\hat{\kappa}(\tilde{a})d\tilde{a}+(\alpha+\sigma)e^{-\beta a}\int_0^a \hat{\kappa}(\tilde{a})d\tilde{a}\right)\right)\\
&=\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(-\beta\frac{1-\hat{\kappa}(a)^2}{2(\alpha+\sigma)}-(\alpha+\sigma-\beta)\frac{1-\hat{\kappa}(a)}{\alpha+\sigma}+(\alpha+\sigma)\frac{1-e^{-a(\alpha+\sigma+\beta)}}{\alpha+\sigma+\beta}\right)\\
&-\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(-\beta\frac{e^{-\beta a}-\hat{\kappa}(a)^2}{2\alpha+2\sigma-\beta}-(\alpha+\sigma-\beta)\frac{e^{-\beta a}-\hat{\kappa}(a)}{\alpha+\sigma-\beta}+(\alpha+\sigma)e^{-\beta a}\frac{1-\hat{\kappa}(a)}{\alpha+\sigma}\right)\\
&=\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(\frac{-\beta\left(1-\hat{\kappa}(a)^2\right)-2(\alpha+\sigma-\beta)(1-\hat{\kappa}(a))}{2(\alpha+\sigma)}\right.\\
&\left.+(\alpha+\sigma)\frac{1-e^{-a(\alpha+\sigma+\beta)}}{\alpha+\sigma+\beta}+\beta\frac{e^{-\beta a}-\hat{\kappa}(a)^2}{2\alpha+2\sigma-\beta}+\left(e^{-\beta a}-\hat{\kappa}(a)\right)-\left(e^{-\beta a}(1-\hat{\kappa}(a))\right)\right)\\
&=\frac{\sigma^2\mathbb{E}[K]}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(\frac{-\beta\left(1-\hat{\kappa}(a)^2\right)-2(\alpha+\sigma-\beta)(1-\hat{\kappa}(a))}{2(\alpha+\sigma)}+(\alpha+\sigma)\frac{1-e^{-a(\alpha+\sigma+\beta)}}{\alpha+\sigma+\beta}+\beta\frac{e^{-\beta a}-\hat{\kappa}(a)^2}{2\alpha+2\sigma-\beta}\right.\\
&\left.+\left(\hat{\kappa}(a)\left(e^{-\beta a}-1\right)\right)\right)
\end{aligned}
\tag{A.2.2}
$$

and

$$
\left(\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\kappa_0^{(0)-}(\tilde{a})d\tilde{a}\right)^2=\left(\frac{\sigma(1-\hat{\kappa}(a))}{\alpha+\sigma}+\frac{\sigma\left(\hat{\kappa}(a)-e^{-\beta a}\right)}{\alpha+\sigma-\beta}\right)^2
$$

For the p^2- term we insert $\kappa_0^{(1)-}$ and $\kappa_0^{(0)-}$

$$
\begin{aligned}
&\kappa_0^{(2)-}(a)\\
&=\hat{\kappa}(a)\left(-\mathbb{E}(K)\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)\sigma\kappa_0^{(1)-}(\tilde{a})d\tilde{a}+\mathbb{E}(K(K-1))\int_0^a \left(1-e^{-\beta(a-\tilde{a})}\right)^2\sigma^2\left(\kappa_0^{(0)-}(\tilde{a})^2+\mathcal{O}(p)d\tilde{a}\right)\right.\\
&=\frac{-\hat{\kappa}(a)\sigma^2\mathbb{E}[K]^2}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\left(\frac{-\beta\left(1-\hat{\kappa}(a)^2\right)-2(\alpha+\sigma-\beta)(1-\hat{\kappa}(a))}{2(\alpha+\sigma)}+(\alpha+\sigma)\frac{1-e^{-a(\alpha+\sigma+\beta)}}{\alpha+\sigma+\beta}+\beta\frac{e^{-\beta a}-\hat{\kappa}(a)^2}{2\alpha+2\sigma-\beta}\right.\\
&\quad\left.+\left(\hat{\kappa}(a)\left(e^{-\beta a}-1\right)\right)\right)\\
&\quad+\hat{\kappa}(a)(\mathrm{Var}[K]+\mathbb{E}(K(K-1)))\left(\frac{\sigma(1-\hat{\kappa}(a))}{\alpha+\sigma}+\frac{\sigma\left(\hat{\kappa}(a)-e^{-\beta a}\right)}{\alpha+\sigma-\beta}\right)^2
\end{aligned}
$$

By inserting the foundings to $\kappa^-_{o,i} = \kappa^{(0)-}_{o,i} + p\kappa^{(1)-}_{o,i}(a) + \frac{p^2}{2}\kappa^{(2)-}_{o,i}(a)$ we obtain the second order approximation for $\kappa^-_o(a)$. □

A.2.3 Forward tracing - recursive

Proposition A.2.3. *The second order approximation of $\kappa^+_{r,i}(a)$ reads*

$$\begin{aligned}&\kappa^+_{r,i}(a)\\ &= \hat\kappa(a) - pp_{obs}(1-\hat\kappa(a))\hat\kappa(a)\\ &\quad - \frac{p^2}{2}\hat\kappa(a)\left(\frac{3\alpha p_{obs} - 2\sigma + 2p^2_{obs}(\alpha+\sigma) - \hat\kappa(a)\left(4\alpha p_{obs} - 4\sigma + 2p^2_{obs}(\alpha+\sigma)\right) - \hat\kappa(a)^2\left(2\sigma - \alpha p_{obs}\right)}{4(\alpha+\sigma)}\right) + \mathcal{O}\left(p^3\right)\end{aligned}$$

Proof. We have $\kappa^+_{r,i}(a) = \hat\kappa(a)\left(1 - p\frac{\int_0^\infty\int_0^a\left(-\kappa^{+'}_{r,i-1}(b+c) - \alpha\kappa^+_{r,i-1}(b+c)\right)dcdb}{\int_0^\infty \kappa^+_{r,i-1}(\tau)d\tau}\right)$ with $\hat\kappa(a) = e^{-(\alpha+\sigma)a}$.

For the second order Taylor approximation we have

$$\sum_{k=0}^{k=2}\kappa^{+(k)}_{r,i}(a)\frac{p^{(k)}}{k!} + \mathcal{O}\left(p^3\right) = \hat\kappa(a)\left(1 - p\frac{\int_0^\infty\int_0^a\sum_{k=0}^{k=2}\frac{p^k}{k!}\left(-\kappa^{(k)+'}_{r,i-1}(b+c) - \alpha\kappa^{(k)+}_{r,i-1}(b+c)\right)dcdb + \mathcal{O}\left(p^3\right)}{\int_0^\infty\sum_{k=0}^{k=2}p^k_{k!}{}^{(k)+}_{r,i-1}(\tau)d\tau + \mathcal{O}(p^3)}\right)$$

As we have nasty terms to consider, we slightly abuse notation and write sums stated in a large bracket

$$\begin{aligned}&\kappa^+_{r,i}(a)\\ =\ &\kappa^{+(0)}_{r,i}(a) + \kappa^{+(1)}_{r,i}(a)p + \kappa^{+(2)}_{r,i}(a)\frac{p^2}{2!} + \mathcal{O}\left(p^3\right)\\ =\ &\hat\kappa(a)\left(\begin{array}{c}\int_0^\infty\int_0^a\left(-\kappa^{+'(0)}_{r,i-1}(b+c) - \alpha\kappa^{+(0)}_{r,i-1}(b+c)\right)dcdb\\ +p\int_0^\infty\int_0^a\left(-\kappa^{+'(1)}_{r,i-1}(b+c) - \alpha\kappa^{+(1)}_{r,i-1}(b+c)\right)dcdb\\ +1 - p\frac{+\frac{p^2}{2!}\int_0^\infty\int_0^a\left(-\kappa^+_{r,i-1}(2)(b+c) - \alpha\kappa^{+(2)}_{r,i-1}(b+c)\right)dcdb + \mathcal{O}\left(p^3\right)}{\int_0^\infty\kappa^{+(0)}_{r,i-1}(\tau)d\tau + p\int_0^\infty\kappa^{+(1)}_{r,i-1}(\tau)d\tau + \frac{p^2}{2!}\int_0^\infty\kappa^{+(2)}_{r,i-1}(\tau)d\tau + \mathcal{O}(p^3)}\end{array}\right) \qquad \text{(A.2.3)}\\ =\ &\hat\kappa(a)\left(1 - p\left(\frac{1}{\int_0^\infty\kappa^{(0)+}_{r,i-1}(b)db} - p\frac{\int_0^\infty\kappa^{(1)+}_{r,i-1}(b)db}{\left(\int_0^\infty\kappa^{(0)+}_{r,i-1}(b)db\right)^2} + \mathcal{O}\left(p^2\right)\right)\right.\\ &\left.\left(\begin{array}{c}\int_0^\infty\int_0^a\left(-\kappa^{+'(0)}_{r,i-1}(b+c) - \alpha\kappa^{+(0)}_{r,i-1}(b+c)\right)dcdb\\ +p\int_0^\infty\int_0^a\left(-\kappa^{+'(1)}_{r,i-1}(b+c) - \alpha\kappa^{+(1)}_{r,i-1}(b+c)\right)dcdb\\ +\frac{p^2}{2!}\int_0^\infty\int_0^a\left(-\kappa^{+'(2)}_{r,i-1}(b+c) - \alpha\kappa^{+(2)}_{r,i-1}(b+c)\right)dcdb + \mathcal{O}\left(p^3\right)\end{array}\right)\right) \qquad \text{(A.2.4)}\end{aligned}$$

From Eqn. A.2.3 to A.2.4, we use the trick:

$$
\begin{aligned}
& \frac{1}{\int_0^\infty \kappa_{r,i-1}^{+(0)}(b)db + p\int_0^\infty \kappa_{r,i-1}^{+(1)}(b)db + \frac{p^2}{2!}\int_0^\infty \kappa_{r,i-1}^{+(2)}(b)db + \mathcal{O}\left(p^3\right)} \\
= & \frac{1}{\int_0^\infty \kappa_{r,i-1}^{+(0)}(b)db + p\int_0^\infty \kappa_{r,i-1}^{+(1)}(b)db + \mathcal{O}\left(p^2\right)} \\
= & \frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db}\frac{1}{1+p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)+}(b)db}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db}} + \mathcal{O}\left(p^2\right) \\
= & \frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db}\left(1 - p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)+}(b)db}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db} + \mathcal{O}\left(p^2\right)\right) \\
= & \left(\frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db} - p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)+}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db\right)^2} + \mathcal{O}\left(p^2\right)\right),
\end{aligned}
$$

where we in the 2. equality used the fact that the probability is smaller than 1, thus have interpreted as geometric series. Extracting the p-terms:

$$
\begin{aligned}
p^0 - \text{ Term } : \kappa_i^{(0)+}(a) &= e^{(\alpha+\sigma)a} = \hat{\kappa}(a) \\
p^1 - \text{ Term } : \kappa_i^{(1)+}(a) &= -\hat{\kappa}(a)\frac{\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(0)+'}(b+c) - \alpha\kappa_{r,i-1}^{(0)+}(b+c)\right)dcdb}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(\tau)d\tau} \\
p^2 - \text{ Term } : \frac{\kappa_i^{(2)+}(a)}{2} &= -\hat{\kappa}(a)\left(\frac{\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(1)++'}(b+c) - \alpha\kappa_{r,i-1}^{(1)+}(b+c)\right)dcdb}{\kappa_{r,i-1}^{(0)+}(b)db}\right) \\
&\quad +\hat{\kappa}(a)\left(\frac{\int_0^\infty \kappa_{r,i-1}^{(1)+}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db\right)^2}\left(\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(0)+'}(b+c) - \alpha\kappa_{r,i-1}^{(0)+}(b+c)\right)dcdb\right)\right)
\end{aligned}
$$

Since we already have calculated the p^0- term and the p^1- term in the first order approximation, we can now iteratively calculate $\kappa_i^{(2)+}(a)$ by using the $\kappa_i^{(1)+}(a)$.

$$
\begin{aligned}
\frac{\kappa_i^{(2)+}(a)}{2} &= -\hat{\kappa}(a)\frac{\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(1)+'}(b+c) - \alpha\kappa_{r,i-1}^{(1)+}(b+c)\right)dcdb}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db} \\
&\quad +\hat{\kappa}(a)\left(\frac{\int_0^\infty \kappa_{r,i-1}^{(1)+}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db\right)^2}\left(\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(0)+'}(b+c) - \alpha\kappa_{r,i-1}^{(0)+}(b+c)\right)dcdb\right)\right)
\end{aligned} \tag{A.2.5}
$$

We do the following calculations to calculate the first term in Eqn. A.2.5

$$\begin{aligned}\kappa_i^{(1)+'}(a) &= (-\hat{\kappa}(a)p_{obs}(1-\hat{\kappa}(a)))' = (\alpha+\sigma)p_{obs}\hat{\kappa}(a) - 2\hat{\kappa}^2(a)p_{obs}(\alpha+\sigma) \\ &= \sigma\hat{\kappa}(a)(1-2\hat{\kappa}(a))\end{aligned}$$

$$\begin{aligned}\implies &-\kappa_i^{(1)+'}(b+c) - \alpha\kappa_{r,i-1}^{(1)+}(b+c) \\ &= -\sigma\hat{\kappa}(b+c) + 2\sigma\hat{\kappa}(b+c)^2 + \alpha p_{obs}\hat{\kappa}(b+c) - \alpha p_{obs}\hat{\kappa}(b+c)^2 \\ &= \hat{\kappa}(b+c)\left(\alpha p_{obs} - \sigma\right) + \hat{\kappa}(b+c)^2\left(2\sigma - \alpha p_{obs}\right)\end{aligned}$$

Now integrating this term:

$$\begin{aligned}\implies &\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(1)+'}(b+c) - \alpha\kappa_{r,i-1}^{(1)+}(b+c)\right) dcdb \\ &= \int_0^\infty \int_0^a \hat{\kappa}(b+c)\left(\alpha p_{obs} - \sigma\right) + \hat{\kappa}(b+c)^2\left(2\sigma - \alpha p_{obs}\right) dcdb \\ &= \int_0^\infty \left(\left(\alpha p_{obs} - \sigma\right)\int_0^a \hat{\kappa}(b+c)dc + \left(2\sigma - \alpha p_{obs}\right)\int_0^a \hat{\kappa}(b+c)^2 dc\right) db \\ &= \left(\alpha p_{obs} - \sigma\right)\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2}\right) + \left(2\sigma - \alpha p_{obs}\right)\left(\frac{(1-\hat{\kappa}(a)^2)}{4(\alpha+\sigma)^2}\right) \\ &= \frac{3\alpha p_{obs} - 2\sigma - 4\left(\alpha p_{obs} - \sigma\right)\hat{\kappa}(a) - \left(2\sigma - \alpha p_{obs}\right)\hat{\kappa}(a)^2}{4(\alpha+\sigma)^2}\end{aligned}$$

and

$$\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db = \int_0^\infty e^{-(\alpha+\sigma)b}db = \frac{1}{(\alpha+\sigma)}$$

we get

$$\begin{aligned}(14) &= -\hat{\kappa}(a)\frac{\frac{3\alpha p_{obs} - 2\sigma - 4(\alpha p_{obs} - \sigma)\hat{\kappa}(a) - (2\sigma - \alpha p_{obs})\hat{\kappa}(a)^2}{4(\alpha+\sigma)^2}}{1/\frac{1}{(\alpha+\sigma)}} \\ &= -\hat{\kappa}(a)\left(\frac{3\alpha p_{obs} - 2\sigma - 4\left(\alpha p_{obs} - \sigma\right)\hat{\kappa}(a) - \left(2\sigma - \alpha p_{obs}\right)\hat{\kappa}(a)^2}{4(\alpha+\sigma)}\right)\end{aligned}$$

For the second term of expression in Eqn. A.2.5, we have

$$
\begin{aligned}
& \hat{\kappa}(a)\left(\frac{\int_0^\infty \kappa_{r,i-1}^{(1)+}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db\right)^2}\left(\int_0^\infty\int_0^a\left(-\kappa_{r,i-1}^{(0)+'}(b+c)-\alpha\kappa_{r,i-1}^{(0)+}(b+c)\right)dcdb\right)\right)\\
= \;& \hat{\kappa}(a)\left(\frac{\frac{-p_{obs}}{2(\alpha+\sigma)}}{\frac{1}{(\alpha+\sigma)^2}}\right)\left(\frac{p_{obs}(1-\hat{\kappa}(a))}{(\alpha+\sigma)}\right)\\
= \;& \hat{\kappa}(a)\left(\frac{-p_{obs}}{2(\alpha+\sigma)}(\alpha+\sigma)^2\right)\left(p_{obs}\left(\frac{1-\hat{\kappa}(a)}{(\alpha+\sigma)}\right)\right)\\
= \;& \frac{-p_{obs}^2\hat{\kappa}(a)(1-\hat{\kappa}(a))}{2}
\end{aligned}
$$

By putting the foundings for $\kappa_i^{(1)+}$ and $\kappa_i^{(0)+}$:

$$
\begin{aligned}
\kappa_i^+(a) = \;& p^0\kappa^{(0)+}(a)+p\kappa_i^{(1)+}(a)+\frac{p^2}{2}\kappa_i^{(2)+}(a)+\mathcal{O}\left(p^3\right)\\
= \;& \hat{\kappa}(a)-pp_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a)\\
& -\frac{p^2}{2}\left(\hat{\kappa}(a)\left(\frac{3\alpha p_{obs}-2\sigma-4\left(\alpha p_{obs}-\sigma\right)\hat{\kappa}(a)-\left(2\sigma-\alpha p_{obs}\right)\hat{\kappa}(a)^2}{4(\alpha+\sigma)}\right)-\frac{p_{obs}^2\hat{\kappa}(a)(1-\hat{\kappa}(a))}{2}\right)+\mathcal{O}\left(p^3\right)\\
= \;& \hat{\kappa}(a)-pp_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a)\\
& -\frac{p^2}{2}\hat{\kappa}(a)\left(\frac{3\alpha p_{obs}-2\sigma+2p_{obs}^2(\alpha+\sigma)-\hat{\kappa}(a)\left(4\alpha p_{obs}-4\sigma+2p_{obs}^2(\alpha+\sigma)\right)-\hat{\kappa}(a)^2\left(2\sigma-\alpha p_{obs}\right)}{4(\alpha+\sigma)}\right)+\mathcal{O}\left(p^3\right)
\end{aligned}
$$

one obtains the second order approximation for forward tracing.

□

A.2.4 Forward tracing - one step

Proposition A.2.4. *The second order approximation for $\kappa_{o,i}^+(a)$ is*

$$
\kappa_{o,i}^+(a)=\hat{\kappa}(a)-pp_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))+\frac{p^2}{2}\hat{\kappa}(a)\left(\frac{p_{obs}((\hat{\kappa}(a)-1)}{2}\right)^2
$$

Proof. Similarly, we use the Taylor approximation and the Neumann series as in forward tracing with recursive method and get:

$$
\begin{aligned}
& \kappa_{o,i}^{+}(a) \\
= \; & \kappa_{o,i}^{(0)+}(a) + \kappa_{o,i}^{(1)+}(a)p + \kappa_{o,i}^{(2)+}(a)\frac{p^2}{2!} + \mathcal{O}\left(p^3\right) \\
= \; & \hat{\kappa}(a) \begin{pmatrix} \int_0^\infty \int_0^a \sigma \kappa_{o,i-1}^{(0)+}(b+c)dcdb \\ +p\int_0^\infty \int_0^a \sigma \kappa_{o,i-1}^{(1)+}(b+c)dcdb \\ +\frac{p^2}{2!}\int_0^\infty \int_0^a \kappa_{o,i-1}^{(2)+}(b+c)dcdb + \mathcal{O}\left(p^3\right) \\ 1 - p\frac{\infty}{\int_0^\infty \kappa_{0,i-1}^{(0)+}(\tau)d\tau + p\int_{0,i-1}^{\infty}(\tau)d\tau + \frac{p^2}{2!}\int_0^\infty \sigma\kappa_{o,i-1}^{(2)+}(\tau)d\tau + \mathcal{O}(p^3)} \end{pmatrix} \\
= \; & \hat{\kappa}(a) \left(1 - p \left(\frac{1}{\int_0^\infty \kappa_{o,i-1}^{(0)+}(b)db} - p\frac{\int_0^\infty \kappa_{o,i-1}^{(1)+}(b)db}{\left(\int_0^\infty \kappa_{o,i-1}^{(0)+}(b)db\right)^2} + \mathcal{O}\left(p^2\right) \right) \right. \\
& \left. \begin{pmatrix} \int_0^\infty \int_0^a \sigma \kappa_{0,i-1}^{(0)+}(b+c)dcdb \\ +p\int_0^\infty \int_0^a \sigma \kappa_{0,i-1}^{(1)+}(b+c)dcdb \\ +\frac{p^2}{2!}\int_0^\infty \int_0^a \sigma \kappa_{o,i-1}^{(2)+}(b+c)dcdb + \mathcal{O}\left(p^3\right) \end{pmatrix} \right)
\end{aligned}
$$

Since we already have the p^0- and p^1- term from the first order approximation as:

$$
\begin{aligned}
& p^0 - \text{term} : \kappa_{o,i}^{(0)+}(a) = \hat{\kappa}(a) \\
& p^1 - \text{ term } : \kappa_{o,i}^{(1)+}(a) = \frac{-\sigma \int_0^\infty \int_0^a \kappa_{o,i-1}^{(0)+}(b+c)dcdb}{\int_0^\infty \kappa_{0,i-1}^{(0)+}(b)db} = -p_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))
\end{aligned}
$$

For the p^2- term:

$$
\begin{aligned}
& \kappa_{o,i}^{(2)+}(a) \\
& = -\hat{\kappa}(a)\frac{\sigma \int_0^\infty \int_0^a \kappa_{o,i-1}^{(1)+}(b+c)dcdb}{\left(\int_0^\infty \kappa_{o,i-1}^{(0)+}(b)db\right)} + \hat{\kappa}(a)\frac{\int_0^\infty \kappa_{o,i-1}^{(1)+}(b)db}{\left(\int_0^\infty \kappa_{o,i-1}^{(0)+}(b)db\right)^2}\left(\sigma \int_0^\infty \int_0^a \kappa_{o,i-1}^{(0)+}(b+c)dcdb\right) \\
& = \hat{\kappa}(a)\frac{-\sigma \int_0^\infty \int_0^a -p_{obs}\hat{\kappa}(b+c)(1-\hat{\kappa}(b+c))dcdb}{\int_0^\infty \hat{\kappa}(b)db} + \hat{\kappa}(a)\frac{\int_0^\infty (-p_{obs}\hat{\kappa}(b)(1-\hat{\kappa}(b))db)}{\left(\int_0^\infty \hat{\kappa}(b)db\right)^2}\sigma \int_0^\infty \int_0^a \hat{\kappa}(b+c)dcdb,
\end{aligned}
$$

where

$$
\begin{aligned}
& \hat{\kappa}(a)\frac{-\sigma \int_0^\infty \int_0^a -p_{obs}\hat{\kappa}(b+c)(1-\hat{\kappa}(b+c))dcdb}{\int_0^\infty \hat{\kappa}(b)db} \\
= & \frac{\sigma\hat{\kappa}(a)p_{obs}\left(\int_0^\infty \int_0^a \hat{\kappa}(b+c)dcdb - \int_0^\infty \int_0^a \hat{\kappa}(b+c)^2 dcdb\right)}{\int_0^\infty \hat{\kappa}(b)} \\
= & \frac{\sigma\hat{\kappa}(a)p_{obs}\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2} - \frac{1-\hat{\kappa}(a)^2}{4(\alpha+\sigma)^2}\right)}{\frac{1}{(\alpha+\sigma)}} \\
= & \frac{\hat{\kappa}(a)(1-\hat{\kappa}(a))(3-\hat{\kappa}(a))p_{obs}^2}{4}
\end{aligned}
$$

and

$$\begin{aligned}
&\hat{\kappa}(a)\frac{\int_0^\infty(-p_{obs}\hat{\kappa}(b)(1-\hat{\kappa}(b))db)}{\left(\int_0^\infty\hat{\kappa}(b)db\right)^2}\sigma\int_0^\infty\int_0^a\hat{\kappa}(b+c)dcdb\\
&=\sigma\hat{\kappa}(a)\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2}\right)\left(\frac{-p_{obs}\left(\frac{1}{(\alpha+\sigma)}-\frac{1}{2(\alpha+\sigma)}\right)}{\frac{1}{(\alpha+\sigma)^2}}\right)\\
&=\frac{-p_{obs}^2\hat{\kappa}(a)(1-\hat{\kappa}(a))}{2}
\end{aligned}$$

Altogether

$$\begin{aligned}
\kappa_{o,i}^{+}(a)&=\kappa_{o,i}^{+(0)}(a)+p\kappa_{o,i}^{+(1)}(a)+\frac{p^2}{2}\kappa_{o,i}^{+(2)}(a)\\
&=\hat{\kappa}(a)-pp_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))+\frac{p^2}{2}\left(\frac{-p_{obs}^2\hat{\kappa}(a)(1-\hat{\kappa}(a))}{2}+\frac{\hat{\kappa}(a)(1-\hat{\kappa}(a))(3-\hat{\kappa}(a))p_{obs}^2}{4}\right)\\
&=\hat{\kappa}(a)-pp_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))+\frac{p^2}{2}\hat{\kappa}(a)\left(\frac{p_{obs}(1-\hat{\kappa}(a))}{2}\right)^2
\end{aligned}$$

□

A.2.5 Full tracing - recursive

Proposition A.2.5. *The second order approximation for the full recursive tracing*

$$\begin{aligned}
\kappa_{r,i}(a)=\hat{\kappa}(a)\Bigg(&1+\frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma-\beta}\left(e^{-\beta a}-\hat{\kappa}(a)\right)-\frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma}(1-\hat{\kappa}(a))-\frac{p\sigma}{(\alpha+\sigma)}(1-\hat{\kappa}(a))\\
&\frac{p^2}{2}\Bigg(p_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))\left(\frac{-p_{obs}\mathbb{E}[K](\alpha+\sigma-\beta)-\sigma\mathbb{E}[K]-p_{obs}(\alpha+\sigma-\beta)}{2(\alpha+\sigma-\beta)(\alpha+\sigma)}+\frac{\sigma\mathbb{E}[K]}{(\alpha+\sigma-\beta)(\alpha+\sigma+\beta)}\right)\\
&+p_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a)\left(\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\sigma+\alpha-\beta))}\right)\right.\\
&-\hat{\kappa}(a)(\alpha+\sigma)\left(\left(\frac{\left(1-\hat{\kappa}(a)e^{\beta a}\right)}{(\alpha+\sigma-\beta)^2}\right)\left(\mathbb{E}[K]\sigma-\frac{\sigma\alpha\mathbb{E}[K]}{\alpha+\sigma-\beta}\right)\right.\\
&+\left(\frac{\left(1-\hat{\kappa}(a)^2\right)}{4(\alpha+\sigma)^2}\right)\left(\frac{\alpha\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}+\mathbb{E}[K]\sigma\left(1+p_{obs}\right)+p_{obs}(2\alpha+2\sigma-1)\right)\\
&\left.\left.+\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2}\right)\left(-p_{obs}(\sigma\mathbb{E}[K]+\alpha+\sigma+1)\right)\right)\right.\\
&\left.+\frac{\hat{\kappa}(a)}{2}(\mathrm{Var}[K]+\mathbb{E}[K](\mathbb{E}(K)-1))\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2+\mathcal{O}\left(p^3\right)\right)\Bigg)
\end{aligned}$$

Proof. We look for the second order approximation. We expand $\kappa_{r,i}(a)$ as a power series in p.

$$
\begin{aligned}
& \kappa_r(a) \\
= & \sum_{k=0}^{2} \frac{p^{(k)}}{k!} \kappa_{r,i}^{(k)}(a) + \mathcal{O}\left(p^3\right) = \kappa_{r,i}^{(0)}(a) + p\kappa_{r,i}^{(1)}(a) + \frac{p^2}{2}\kappa_{r,i}^{(2)}(a) + \mathcal{O}\left(p^3\right) \\
= & \left(\sum_{k=0}^{2} \frac{p^{(k)}}{k!} \kappa_r^{(k)-}(a) + \mathcal{O}\left(p^3\right)\right)\left(1 - p\frac{\sum_{k=0}^{2} \frac{p^k}{k!} \int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(k)'}(b+c) - \alpha\kappa_{r,i-1}^{(k)}(b+c)\right) dcdb + \mathcal{O}\left(p^3\right)}{\sum_{k=0}^{2} \frac{p^k}{k!} \int_0^\infty \kappa_{r,i-1}^{(k)}(b)db + \mathcal{O}\left(p^3\right)}\right) \\
= & \left(\kappa_r^{(0)-}(a) + p\kappa_r^{(1)-}(a) + \frac{p^2}{2}\kappa_r^{(2)-}(a) + \mathcal{O}\left(p^3\right)\right)\left(1 - p\frac{\sum_{k=0}^{2} \frac{p^k}{k!} \int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(k)'}(b+c) - \alpha\kappa_{r,i-1}^{(k)}(b+c)\right) dcdb + \mathcal{O}\left(p^3\right)}{\sum_{k=0}^{2} \frac{p^k}{k!} \int_0^\infty \kappa_{r,i-1}^{(k)}(b)db + \mathcal{O}\left(p^3\right)}\right) \\
= & \left(\kappa_r^{(0)-}(a) + p\kappa_r^{(1)-}(a) + \frac{p^2}{2}\kappa_r^{(2)-}(a) + \mathcal{O}\left(p^3\right)\right)\Bigg(1 - p \\
& \left(\frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db} - p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db\right)^2} + \mathcal{O}\left(p^3\right)\right)\begin{pmatrix} \int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{+'(0)}(b+c) - \alpha\kappa_{r,i-1}^{+(0)}(b+c)\right) dcdb \\ +p\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(1)(b+c) - \alpha\kappa_{r,i-1}^{(1)}(b+c)\right) dcdb \\ +\frac{p^2}{2!}\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(2)(b+c) - \alpha\kappa_{r,i-1}^{(2)}(b+c)\right) dcdb + \mathcal{O}\left(p^3\right) \end{pmatrix}\Bigg) \\
= & \kappa_r^{(0)-}(a) - p\kappa_r^{(0)-}(a) \\
& \left(\frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db} - p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db\right)^2} + \mathcal{O}\left(p^3\right)\right)\begin{pmatrix} \int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(b)\right. \\ +p\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(1)}(b+c) - \alpha\kappa_{r,i-1}^{(1)}(b+c)\right) dcdb \\ +\frac{p^2}{2!}\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'^{(2)}(b+c) - \alpha\kappa_{r,i-1}^{(2)}(b+c)\right) dcdb + \mathcal{O}\left(p^3\right) \end{pmatrix} \\
& +p\kappa_r^{(1)-}(a) - p^2\kappa_r^{(1)-}(a) \\
& \left(\frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)+}(b)db} - p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db\right)^2} + \mathcal{O}\left(p^3\right)\right)\begin{pmatrix} \int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(0)\right. \\ +p\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{(1)}(b+c) - \alpha\kappa_{r,i-1}^{(1)}(b+c)\right) dcdb \\ \left.+\kappa_{r,i-1}^{(b)}(b+c)\right) dcdb \\ +\frac{p^2}{2!}\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(2)(b+c) - \alpha\kappa_{r,i-1}^{(2)}(b+c)\right) dcdb + \mathcal{O}\left(p^3\right) \end{pmatrix} \\
& +\frac{p^2}{2}\kappa_r^{(2)-}(a) - \frac{p^3}{2}\kappa_r^{(2)-}(a) \\
& \left(\frac{1}{\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db} - p\frac{\int_0^\infty \kappa_{r,i-1}^{(1)}(b)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db\right)^2} + \mathcal{O}\left(p^3\right)\right)\begin{pmatrix} \int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(b) + c\right) - \alpha\kappa_{r,i-1}^{(0)}(b+c)\right) dcdb \\ +p\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}'(1)(b+c) - \alpha\kappa_{r,i-1}^{(1)}(b+c)\right) dcdb \\ +\frac{p^2}{2!}\int_0^\infty \int_0^a \left(-\kappa_{r,i-1}^{+'(2)}(b+c) - \alpha\kappa_{r,i-1}^{(2)}(b+c)\right) dcdb + \mathcal{O}\left(p^3\right) \end{pmatrix}.
\end{aligned}
$$

We combine the p^0 and p^1 term from the first order approximation, we only read the p^2 term

$$\frac{\kappa_{r,i}^{(2)}(a)}{2} = \kappa_r^{(0)-}(a)\left(\frac{\int_0^\infty \kappa_{r,i-1}^{(1)}(b)db\int_0^\infty\int_0^a\left(-\kappa_{r,i-1}^{(0)'}(b+c)-\alpha\kappa_{r,i-1}^{(0)}(b+c)dc\right)db}{\left(\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db\right)^2}\right) \quad \text{(A.2.6)}$$

$$-\kappa_r^{(1)-}(a)\left(\frac{\int_0^\infty\int_0^a\left(-\kappa_{r,i-1}^{(0)'}(b+c)-\alpha\kappa_{r,i-1}^{(0)}(b+c)dc\right)db}{\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db}\right) \quad \text{(A.2.7)}$$

$$-\kappa_r^{(0)-}(a)\left(\frac{\int_0^\infty\int_0^a\left(-\kappa_{r,i-1}^{(1)'}(b+c)-\alpha\kappa_{r,i-1}^{(1)}(b+c)dc\right)db}{\int_0^\infty \kappa_{r,i-1}^{(0)}(b)db}\right) \quad \text{(A.2.8)}$$

$$+\frac{\kappa_2^{(2)-(a)}}{2} \quad \text{(A.2.9)}$$

Divide the calculation process from Eqn. A.2.6, A.2.7, A.2.8, A.2.9

$$A.2.6 = p_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))\left(\frac{-p_{obs}\mathbb{E}[K](\alpha+\sigma-\beta)-\sigma\mathbb{E}[K]-p_{obs}(\alpha+\sigma-\beta)}{2(\alpha+\sigma-\beta)(\alpha+\sigma)}+\frac{\sigma\mathbb{E}[K]}{(\alpha+\sigma-\beta)(\alpha+\sigma+\beta)}\right)$$

$$A.2.7 = p_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a)\left(\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\sigma+\alpha-\beta)}\right)\right)$$

$$\begin{aligned}A.2.8 = &-\hat{\kappa}(a)(\alpha+\sigma)\left(\left(\frac{\left(1-\hat{\kappa}(a)e^{\beta a}\right)}{(\alpha+\sigma-\beta)^2}\right)\left(\mathbb{E}[K]\sigma-\frac{\sigma\alpha\mathbb{E}[K]}{\alpha+\sigma-\beta}\right)\right.\\&+\left(\frac{\left(1-\hat{\kappa}(a)^2\right)}{4(\alpha+\sigma)^2}\right)\left(\frac{\alpha\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta}+\mathbb{E}[K]\sigma\left(1+p_{obs}\right)+p_{obs}(2\alpha+2\sigma-1)\right)\\&\left.+\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2}\right)\left(-p_{obs}(\sigma\mathbb{E}[K]+\alpha+\sigma+1)\right)\right)\end{aligned}$$

$$A.2.9 = \hat{\kappa}(a)\mathbb{E}[K(K-1)]\left(p_{obs}(1-\hat{\kappa}(a))-\frac{\sigma\left(e^{-\beta a}-\hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2\frac{1}{2}$$

By putting the foundings together

$$
\begin{aligned}
\kappa_r(a) &= \kappa_{r,i}^{(0)}(a) + p\kappa_{r,i}^{(1)}(a) + \frac{p^2}{2}\kappa_{r,i}^{(2)}(a) + \mathcal{O}\left(p^3\right) \\
&= \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma}(1-\hat{\kappa}(a)) - \frac{p\sigma}{(\alpha+\sigma)}(1-\hat{\kappa}(a))\right. \\
&\quad \frac{p^2}{2}\left(p_{obs}\hat{\kappa}(a)(1-\hat{\kappa}(a))\left(\frac{-p_{obs}\mathbb{E}[K](\alpha+\sigma-\beta) - \sigma\mathbb{E}[K] - p_{obs}(\alpha+\sigma-\beta)}{2(\alpha+\sigma-\beta)(\alpha+\sigma)} + \frac{\sigma\mathbb{E}[K]}{(\alpha+\sigma-\beta)(\alpha+\sigma+\beta)}\right)\right. \\
&\quad +p_{obs}(1-\hat{\kappa}(a))\hat{\kappa}(a)\left(\mathbb{E}[K]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma\left(e^{-\beta a} - \hat{\kappa}(a)\right)}{(\sigma+\alpha-\beta)}\right)\right) \\
&\quad -\hat{\kappa}(a)(\alpha+\sigma)\left(\left(\frac{\left(1-\hat{\kappa}(a)e^{\beta a}\right)}{(\alpha+\sigma-\beta)^2}\right)\left(\mathbb{E}[K]\sigma - \frac{\sigma\alpha\mathbb{E}[K]}{\alpha+\sigma-\beta}\right)\right. \\
&\quad +\left(\frac{\left(1-\hat{\kappa}(a)^2\right)}{4(\alpha+\sigma)^2}\right)\left(\frac{\alpha\sigma\mathbb{E}[K]}{\alpha+\sigma-\beta} + \mathbb{E}[K]\sigma\left(1+p_{obs}\right) + p_{obs}(2\alpha+2\sigma-1)\right) \\
&\quad \left.+\left(\frac{(1-\hat{\kappa}(a))}{(\alpha+\sigma)^2}\right)\left(-p_{obs}(\sigma\mathbb{E}[K]+\alpha+\sigma+1)\right)\right) \\
&\quad \left.+\frac{\hat{\kappa}(a)}{2}\mathbb{E}[K(K-1)]\left(p_{obs}(1-\hat{\kappa}(a)) - \frac{\sigma\left(e^{-\beta a} - \hat{\kappa}(a)\right)}{(\alpha+\sigma-\beta)}\right)^2 + \mathcal{O}\left(p^3\right)\right).
\end{aligned}
$$

□

A.2.6 Full tracing - one step

For completeness, we also state the second order approximation for one-step full tracing. The proof is analogous to recursive full tracing by replacing $-\kappa_{r,i}^{'} - \alpha\kappa_{r,i-1}$ with $\sigma\kappa_{o,i-1}$.

Proposition A.2.6. *The second order approximation for $\kappa_{o,i}(a)$ is*

$$
\begin{aligned}
&\kappa_{o,i}(a) \\
&= \hat{\kappa}(a)\left(1 + \frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma-\beta}\left(e^{-\beta a} - \hat{\kappa}(a)\right) - \frac{p\sigma\mathbb{E}(K)}{\alpha+\sigma}(1-\hat{\kappa}(a)) - \frac{p\sigma}{\alpha+\sigma}(1-\hat{\kappa}(a))\right) \\
&\quad + \frac{p^2}{2}\left(\hat{\kappa}(a)\left(\frac{-p_{obs}(1+2\mathbb{E}[K])}{2(\alpha+\sigma)} + \frac{\mathbb{E}[K]\sigma}{(\alpha+\sigma-\beta)^2} - \frac{p_{obs}\mathbb{E}[K]\beta}{2(\alpha+\sigma)(\alpha+\sigma-\beta)}\right)(\sigma(1-\hat{\kappa}(a)))\right. \\
&\quad - \left(\mathbb{E}[K]\sigma\frac{-\hat{\kappa}(a)^2\beta - \hat{\kappa}(a)(\alpha+\sigma-\beta) + e^{-\beta a}\hat{\kappa}(a)(\alpha+\sigma)}{(\alpha+\sigma)(\alpha-\beta+\sigma)}\right)(\sigma p_{obs}(1-\hat{\kappa}(a))) \\
&\quad - \hat{\kappa}(a)\left(\frac{-\mathbb{E}[K]p_{obs}\beta}{\alpha+\sigma-\beta}\left(\frac{1-\hat{\kappa}(a)^2}{4(\alpha+\sigma)^2}\right) - \mathbb{E}[K]p_{obs}\left(\frac{1-\hat{\kappa}(a)}{(\alpha+\sigma)^2}\right) + \frac{\mathbb{E}[K]\sigma\left(1-\hat{\kappa}(a)e^{\beta a}\right)}{(\alpha+\sigma-\beta)^3}\right)\sigma(\alpha+\sigma) \\
&\quad \left.+ \frac{\hat{\kappa}(a)\mathbb{E}(K(K-1))}{2}\left(\frac{\sigma(1-\hat{\kappa}(a))}{\alpha+\sigma} + \frac{\sigma\left(\hat{\kappa}(a) - e^{-\beta a}\right)\right)^2}{\alpha+\sigma-\beta}\right)^2
\end{aligned}
$$

List of Figures

List of Tables

Bibliography

[1] Dillon C Adam, Peng Wu, Jessica Y Wong, Eric HY Lau, Tim K Tsang, Simon Cauchemez, Gabriel M Leung, and Benjamin J Cowling. Clustering and superspreading potential of sars-cov-2 infections in hong kong. *Nature Medicine*, 26(11):1714–1719, 2020.

[2] FB Agusto and MA Khan. Optimal control strategies for dengue transmission in pakistan. *Mathematical biosciences*, 305:102–121, 2018.

[3] Linda JS Allen and Amy M Burgin. Comparison of deterministic and stochastic sis and sir models. *Dept. Math. Stat. Tech. Rep. Ser*, pages 98–103, 1998.

[4] Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[5] Frank G. Ball, Edward S. Knock, and Philip D. O'Neill. Threshold behaviour of emerging epidemics featuring contact tracing. *Adv. Appl. Probab.*, 43:1048–1065, 2011.

[6] Frank G. Ball, Edward S. Knock, and Philip D. O'Neill. Stochastic epidemic models featuring contact tracing with delays. *Math. Biosci.*, 266:23 – 35, 2015. ISSN 0025-5564.

[7] Andrew Barbour and Denis Mollison. Epidemics and random graphs. In *Stochastic processes in epidemic theory*, pages 86–89. Springer, 1990.

[8] C Bauch and DA Rand. A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. *Proceedings of the Royal Society of London B: Biological Sciences*, 267(1456):2019–2027, 2000.

[9] Niels G Becker. *Analysis of infectious disease data*. Routledge, 2017.

[10] Michael GB Blum and Viet Chi Tran. Hiv with contact tracing: a case study in approximate bayesian computation. *Biostatistics*, 11(4):644–660, 2010.

[11] Begoña Cantó, Carmen Coll, and Elena Sánchez. Structural identifiability of a model of dialysis. *Mathematical and computer modelling*, 50(5-6):733–737, 2009.

[12] Begoña Cantó, Carmen Coll, and Elena Sánchez. Identifiability for a class of discretized linear partial differential algebraic equations. *Mathematical Problems in Engineering*, 2011, 2011.

[13] SM Cavany, T Sumner, E Vynnycky, et al. An evaluation of tuberculosis contact investigations against national standards. thorax [electronic article]. 2017; 72 (8): 736–745, 2017.

[14] Gheorghe Craciun and Casian Pantea. Identifiability of chemical reaction networks. *Journal of Mathematical Chemistry*, 44(1):244–259, 2008.

[15] Leon Danon, Ashley P Ford, Thomas House, Chris P Jewell, Matt J Keeling, Gareth O Roberts, Joshua V Ross, and Matthew C Vernon. Networks and the epidemiology of infectious disease. *Interdisciplinary perspectives on infectious diseases*, 2011, 2011.

[16] Donald L DeAngelis and Volker Grimm. Individual-based models in ecology after four decades. *F1000prime reports*, 6, 2014.

[17] Odo Diekmann, MCM De Jong, and Johan Anton Jacob Metz. A deterministic epidemic model taking account of repeated contacts between the same individuals. *Journal of Applied Probability*, 35(2):448–462, 1998.

[18] R Duarte, M Neto, A Carvalho, and H Barros. Improving tuberculosis contact tracing: the role of evaluations in the home and workplace. *The International journal of tuberculosis and lung disease*, 16(1):55–59, 2012.

[19] Richard Durrett. *Random graph dynamics*, volume 200. Citeseer, 2007.

[20] Rick Durrett. Ten lectures on particle systems. In *Lectures on Probability Theory*, pages 97–201. Springer, 1995.

[21] K. Eames, S. Bansal, S. Frost, and S. Riley. Six challenges in measuring contact networks for use in modelling. *Epidemics*, 10:72–77, 2015. ISSN 1755-4365. doi: https://doi.org/10.1016/j.epidem.2014.08.006. URL https://www.sciencedirect.com/science/article/pii/S1755436514000413. Challenges in Modelling Infectious DIsease Dynamics.

[22] Ken TD Eames and Matt J Keeling. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Sciences*, 99(20):13330–13335, 2002.

[23] Ken TD Eames and Matt J Keeling. Contact tracing and disease control. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1533):2565–2571, 2003.

[24] KTD Eames. Contact tracing strategies in heterogeneous populations. *Epidemiology & Infection*, 135(3):443–454, 2007.

[25] P Erdös and A Rényi. On random graphs i. publ math debrecen. 1959.

[26] Razia Fatima, Ejaz Qadeer, Aashifa Yaqoob, Mahboob ul Haq, Suman S Majumdar, Hemant D Shewade, Robert Stevens, Jacob Creswell, Nasir Mahmood, and Ajay MV Kumar. Extending 'contact tracing' into the community within a 50-metre radius of an index tuberculosis patient using xpert mtb/rif in urban, pakistan: did it increase case detection? *PLoS One*, 11(11):e0165813, 2016.

[27] Neil M Ferguson and Geoffrey P Garnett. More realistic models of sexually transmitted disease transmission dynamics: sexual partnership networks, pair models, and moment closure. *Sexually transmitted diseases*, 27(10):600–609, 2000.

[28] Richard G FitzJohn, Wayne P Maddison, and Sarah P Otto. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic biology*, 58(6):595–611, 2009.

[29] Ch. Fraser, S. Riley, R.M. Anderson, and N.M. Ferguson. Factors that make an infectious disease outbreak controllable. *PNAS*, 101:6146 – 6151, 2004.

[30] Christophe Fraser, Steven Riley, Roy M Anderson, and Neil M Ferguson. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences*, 101(16):6146–6151, 2004.

[31] Edgar N Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4): 1141–1144, 1959.

[32] Thomas Götz, Nicole Altmeier, Wolfgang Bock, Robert Rockenfeller, Karunia Putra Wijaya, et al. Modeling dengue data from semarang, indonesia. *Ecological complexity*, 30:57–62, 2017.

[33] Bryan T Grenfell, Andrew P Dobson, HK Moffatt, et al. *Ecology of infectious diseases in natural populations*, volume 7. Cambridge University Press, 1995.

[34] Mohak Gupta, Giridara G Parameswaran, Manraj S Sra, Rishika Mohanta, Devarsh Patel, Amulya Gupta, Bhavik Bansal, Vardhmaan Jain, Archisman Mazumder, Mehak Arora, et al. Contact tracing of covid-19 in karnataka, india: Superspreading and determinants of infectiousness and symptomatic infection. *Plos one*, 17(7):e0270789, 2022.

[35] Herbert W Hethcote and James A Yorke. *Gonorrhea transmission dynamics and control*, volume 56. Springer, 2014.

[36] Thomas House and Matt J. Keeling. The impact of contact tracing in clustered populations. *PLoS Comput. Biol.*, 6:e1000721, 2010. doi: 10.1371/journal.pcbi.1000721.

[37] Brian Karrer and Mark EJ Newman. Message passing approach for general epidemic models. *Physical Review E*, 82(1):016101, 2010.

[38] M. Keeling. Correlation equations for endemic diseases. *Proceedings of the Royal Society B: Biological Sciences.*, 266:953–961, 1999.

[39] Matt J Keeling and Ken TD Eames. Networks and epidemic models. *Journal of the royal society interface*, 2(4):295–307, 2005.

[40] Matthew J Keeling. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1421): 859–867, 1999.

[41] MJ Keeling, DA Rand, and AJ Morris. Correlation models for childhood epidemics. *Proceedings of the Royal Society of London B: Biological Sciences*, 264(1385):1149–1156, 1997.

[42] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721, 1927.

[43] Muhammad Altaf Khan et al. Parameter estimation and fractional derivatives of dengue transmission model. *AIMS Mathematics*, 5(3):2758–2779, 2020.

[44] Istvan Z Kiss, Darren M Green, and Rowland R Kao. Infectious disease control using contact tracing in random and scale-free networks. *Journal of The Royal Society Interface*, 3:55–62, 2005. doi: 10.1098/rsif.2005.0079.

[45] Istvan Z Kiss, Darren M Green, and Rowland R Kao. The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *J. R. Soc. Interface*, pages 791–799, 2007. doi: 10.1098/rsif.2007.1272.

[46] Istvan Z Kiss, Gergely Röst, and Zsolt Vizi. Generalization of pairwise models to non-Markovian epidemics on networks. *Phys. Rev. E*, 115:078701, 2015.

[47] Don Klinkenberg, Christophe Fraser, and Hans Heesterbeek. The effectiveness of contact tracing in emerging epidemics. *PloS one*, 1(1):e12, 2006.

[48] Tae Jin Lee, Masayuki Kakehashi, and Arni S.R. Srinivasa Rao. Chapter 8 - network models in epidemiology. In Arni S.R. Srinivasa Rao and C.R. Rao, editors, *Data Science: Theory and Applications*, volume 44 of *Handbook of Statistics*, pages 235–256. Elsevier, 2021. doi: https://doi.org/10.1016/bs.host.2020.09.002. URL https://www.sciencedirect.com/science/article/pii/S0169716120300444.

[49] Thomas Milton Liggett and Thomas M Liggett. *Interacting particle systems*, volume 2. Springer, 1985.

[50] Roderick J Little, Ralph D'Agostino, Michael L Cohen, Kay Dickersin, Scott S Emerson, John T Farrar, Constantine Frangakis, Joseph W Hogan, Geert Molenberghs, Susan A Murphy, et al. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*, 367(14):1355–1360, 2012.

[51] James O Lloyd-Smith, Sebastian J Schreiber, P Ekkehard Kopp, and Wayne M Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359, 2005.

[52] Wayne P Maddison, Peter E Midford, and Sarah P Otto. Estimating a binary character's effect on speciation and extinction. *Systematic biology*, 56(5):701–710, 2007.

[53] Joel C Miller, Anja C Slim, and Erik M Volz. Edge-based compartmental modelling for infectious disease spread. *J. R. Soc. Interface*, 9:890–906, 2011.

[54] Joel C Miller, Anja C Slim, and Erik M Volz. Edge-based compartmental modelling for infectious disease spread. *Journal of the Royal Society Interface*, 9(70):890–906, 2012.

[55] NJ Mills and AP Gutierrez. Prospective modelling in biological control: an analysis of the dynamics of heteronomous hyperparasitism in a cotton-whitefly-parasitoid system. *Journal of Applied Ecology*, pages 1379–1394, 1996.

[56] Regina Monteiro, Aurora Carvalho, and Raquel Duarte. New opportunities in tuberculosis control. *European Respiratory Journal*, 39(5):1271–1273, 2012.

[57] Joël Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, 5(3):e74, 2008.

[58] Johannes Müller and Volker Hösel. Estimating the tracing probability from contact history at the onset of an epidemic. *Mathematical Population Studies*, 14(4):211–236, 2007.

[59] Johannes Müller and Volker Hösel. Contact tracing & super-spreaders in the branching-process model. *arXiv preprint arXiv:2010.04942*, 2020.

[60] Johannes Müller and Bendix Koopmann. The effect of delay on contact tracing. *Mathematical biosciences*, 282:204–214, 2016.

[61] Johannes Müller and Christina Kuttler. *Methods and models in mathematical biology*. Springer, 2015.

[62] Johannes Müller, Mirjam Kretzschmar, and Klaus Dietz. Contact tracing in stochastic and deterministic epidemic models. *Mathematical biosciences*, 164(1):39–64, 2000.

[63] Sean Nee, Robert Mccredie May, and Paul H Harvey. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 344(1309):305–311, 1994.

[64] Mark EJ Newman. Spread of epidemic disease on networks. *Physical review E*, 66(1): 016128, 2002.

[65] Mark EJ Newman, Steven H Strogatz, and Duncan J Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, 2001.

[66] Augustine Okolie and Johannes Müller. Exact and approximate formulas for contact tracing on random trees. *Mathematical biosciences*, 321:108320, 2020.

[67] Madan K Oli, Meenakshi Venkataraman, Paul A Klein, Lori D Wendland, and Mary B Brown. Population dynamics of infectious diseases: a discrete time model. *Ecological Modelling*, 198(1-2):183–194, 2006.

[68] PD O'Neill, GO Roberts, and Bradford Univ.(United Kingdom). Dept. of Mathematics;. *Bayesian inference for partially observed stochastic epidemics*. University of Bradford. School of Mathematical Sciences, 1997.

[69] Ronald Ross and Hilda P Hudson. An application of the theory of probabilities to the study of a priori pathometry. part i. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 92(638):204–230, 1916.

[70] Ronald Ross and Hilda P Hudson. An application of the theory of probabilities to the study of a priori pathometry. part ii. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 93(650):212–225, 1917.

[71] Neil Sherborne, Joel C Miller, Konstantin B Blyuss, and Istvan Z Kiss. Mean-field models for non-markovian epidemics on networks. *Journal of mathematical biology*, 76(3):755–778, 2018.

[72] John Snow. Cholera, and the water supply in the south districts of london. *British Medical Journal*, 1(42):864, 1857.

[73] Loes Soetens, Don Klinkenberg, Corien Swaan, Susan Hahné, and Jacco Wallinga. Real-time estimation of epidemiologic parameters from contact tracing data during an emerging infectious disease outbreak. *Epidemiology*, 29(2):230–236, 2018.

[74] Tanja Stadler and Sebastian Bonhoeffer. Uncovering epidemiological dynamics in heterogeneous host populations using phylogenetic methods. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1614):20120198, 2013.

[75] Daniel J Sullivan and Wolfgang Völkl. Hyperparasitism: multitrophic ecology and behavior. *Annual review of entomology*, 44(1):291–315, 1999.

[76] Debra Van Egeren, Madison Stoddard, Alexander Novokhodko, Michael S Rogers, Diane Joseph-McCarthy, Bruce Zetter, and Arijit Chakravarty. Rapid relaxation of pandemic restrictions after vaccine rollout favors growth of sars-cov-2 variants: A model-based analysis. *PloS one*, 16(11):e0258997, 2021.

[77] Jacco Wallinga and Marc Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007.

[78] Robert R Wilkinson and Kieran J Sharkey. Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks. *Physical Review E*, 89(2): 022808, 2014.

[79] Robert R Wilkinson, Frank G Ball, and Kieran J Sharkey. The relationships between message passing, pairwise, kermack–mckendrick and stochastic sir epidemic models. *Journal of mathematical biology*, 75(6):1563–1590, 2017.

[80] Donald E Woodhouse, Richard B Rothenberg, John J Potterat, William W Darrow, Stephen Q Muth, Alden S Klovdahl, Helen P Zimmerman, Helen L Rogers, Tammy S Maldonado, John B Muth, et al. Mapping a social network of heterosexuals at high risk for hiv infection. *Aids*, 8(9):1331–1336, 1994.