Moment Closure Approximations in Epidemiology

Christopher Thomas Bauch

Submitted for the degree of Doctor of Philosophy

Mathematics Institute
University of Warwick
Coventry, England

August 2000

for my mother,
Birgit Rösler Bauch

Contents

List of Tables

Acknowledgements

Through the past four years a number of people have helped me in diverse ways and at various stages in my degree course, from the first days of my arrival at Warwick to the final months of hectic writing up. These acknowledgements would fill a chapter if I mentioned each of them and their unique contributions, however I would like to mention a number of them individually.

On the departmental side I would like to thank my supervisor Professor David Rand for several years of advice, support and encouragement. I am also grateful to Markus Kirkilionis, Keith Yates, Clemens Utzny, Hugo van den Berg, Jack Cohen, Nigel Burroughs, Rob Reid, Rolf Mantel, Andrew Morris, Natasha Komarova, Chris Bowman and many others in the Institute who have shared their knowledge and experience with me, and to Elaine Greaves Coelho, Alison Solman, Shayda Hussein, Georgina Copland, Ann Hemmings, the COMPOST team and all those who keep the department administration and computer network running.

Outside of the department friends and family have contributed to enriching my Ph.D. years. My family have always supported my academic goals. I am thankful to my friends in Cambridge, Coventry, and the Coventry Baha'i community as well for reminding me that there is life outside room 101. The likes of Bob Marley and W.A. Mozart have also made the job of writing up easier.

Finally I would like to acknowledge the National Science Foundation for funding my PhD research.

Declaration

The work of this thesis is entirely the work of the author unless otherwise noted. Where others have contributed, the extent of their contribution has been made clear. This thesis has not been submitted for a degree at another university.

Material from chapter 4 has been published as: Bauch, C.T. and Rand, D.A., 2000, *A Moment Closure Model for Sexually Transmitted Disease Transmission Through a Concurrent Partnership Network*. Proceedings of the Royal Society, Series B., vol. 267 (1465): 2019–2027.

Material from chapter 3 is in preparation as: Bauch, C.T., 2000, *Moment Closure Models for Infectious Diseases on Dynamic Sexual Partnership Networks*.

Material from chapters 2 and 5 is in preparation as: Bauch, C.T., and Kirkilionius, M., 2000, *Cluster Approximations for Infection Processes on a Perturbed Lattice*.

Summary

Moment closure approximation (MCA) is a method of obtaining dynamic deterministic approximations to models where spatiality is important. Such approximations track the time evolution of low-order correlations, for instance the correlation of disease status of nearest-neighbours in a square lattice. Thus they are able to capture aspects of population dynamics which traditional mean-field approximations are unable to.

This thesis extends the techniques of moment closure approximation and develops novel applications for MCA in epidemiology. Most existing moment closures were intended as deterministic approximations to static regular lattices. However we develop deterministic approximations for dynamic network models and continuous space models. The purpose of applying MCA to a different set of models is not only to demonstrate their flexibility; we also explore the dynamical properties of such models with the moment closure tools we derive and with simulation data. Comparisons are then made between processes on regular lattices and processes in dynamic networks and in continuous space. Additionally, we answer questions relating to the epidemiology of sexually transmitted diseases and epidemics in populations embedded in two-dimensional continuous space. Some of the new techniques we develop can be applied to other models in ecology and epidemiology. We conclude that moment closure approximations continue to provide fertile ground for research, and that application of MCA to models other than static regular lattices can be worthwhile.

Chapter 1 consists of background material and an introduction to moment closure approximations. In chapter 2 we look at the properties of moment closure approximations near critical points and during transient phases and consider their accuracy in such cases. Chapters 3 and 4 cover the application of pair approximations to sexually transmitted disease models, and chapter 5 is a preliminary study of a pair approximation for a continuous space model.

Chapter 1

Introduction and Background

If a man's wit be wandering, let him study the mathematics.

– Francis Bacon

1.1 Preliminaries

This thesis is about the mathematical modelling of epidemics in structured populations. We are interested in the effect of factors such as spatiality and the network structure of social contacts, and we are particularly interested in building analytical models incorporating these factors. The first chapter will present some relevant history and background from the beginnings of mathematical biology to recent developments in *moment closure approximations* (MCA). But first we will make a few comments about mathematical modelling in general.

1.2 About Modelling

We group epidemic models into two categories. The first consists of *empirical models*, which use statistical and/or extrapolation methods to predict the short- to middle-term progress of real diseases based on real data. The second category consists of *dynamical models* which seek to incorporate assumptions about the underlying biology into models which can be evolved through time to gain insight into the resulting dynamics. Many scientists restrict use of the term *model* to the second category or even to the subset of analytical models of the second category. Dynamical models focus on the more general aspects and basic principles of epidemics, while empirical models are more useful for prediction of the course of a disease in a particular population. In this thesis we deal exclusively with dynamical models.

1.2.1 Three Viewpoints on Modelling

There are three approaches to biomathematical modelling. Models should not be simply put in one of these categories or another, but this division into three approaches can help us clarify our understanding of what we are actually doing when we model, or what we should be doing. We present the approaches here in order of increasing faithfulness to what mathematical biologists actually attempt to do in most cases.

Modelling as Description

In this view, mathematical models simply express mathematically what is already understood from verbal descriptions and empirical knowledge. They are therefore not really that useful although they might provide a slightly different angle. This view is rather disappointing because such models do not really provide any insight

into biological systems. Rather they merely festoon established scientific truths. However this view about modelling is understandably held by many because many models (including some models we have developed) do not really go far beyond describing what is obvious to intuition or what is already known.

Modelling as Prediction

In this view, models are used to predict the course of epidemics in real populations. However there are many factors in real populations that are not taken into account in the relatively simple dynamical models which have been developed thus far. So this kind of prediction is perhaps best left to empirical models.

Modelling as Insight

Finally, there is the view that mathematical models are for gaining insight into biological systems. The modeller expresses real-world systems in terms of a small set of rules which are incorporated into the model, hopefully excluding unnecessary detail and retaining the essentials. If the results are similar to real-world systems, the model may have illumined the essential elements of real-world biological systems. If the model predicts novel phenomena, this can be sought out in empirical data to test the model and gain further insight. However different models might give the same behaviour, so care must be used. In this view of modelling, simplicity is a strength, not a weakness. Of course the real world might be too complicated to find a small and simple set of general principles and we may have to settle for rules of thumb, but only investigation will tell us whether or not these principles exist.

The study of nonlinearities is also an example of how mathematical modelling can contribute to insight about real biologies, in that nonlinearities can often produce seemingly random behaviour which is actually caused by simple deterministic equations. In fact, deterministic chaos, long thought to exist in real ecologies, was found in the population dynamics of flour beetles in 1997 in a laboratory setting [20].

1.2.2 Why Mathematical Modelling is a Good Thing

As mentioned in subsection 1.2.1 modelling first and foremost can provide insight into epidemic dynamics. Real data for ecological and epidemiological systems are often scant, or difficult to interpret, or complicated and involving much detail. Modelling makes the process of inference from data and interpretation of data easier. Inference from data is harder in biology than in physics or chemistry, but this fact can be seen

as a stronger reason to seek models for biological systems instead of something to discourage us.

There are practical advantages to modelling as well. The time and cost involved in carrying out mathematical analysis and simulations are a small fraction of the time and cost of full-scale experiments on real biological systems. Furthermore, experiments on biological systems often cannot be carried out because of ethical or environmental reasons. One example is the obvious difficulty of assessing the structure of sexual partnership networks for the purpose of studying sexually transmitted diseases.

1.3 Early Developments in Mathematical Biology

Mathematical biology is already a massive subject with a long history, although most progress has occurred in the twentieth century, and new possibilities have been opened up by increasing computer power in the last few decades. In this section we outline some of the major historical developments of its early history.

Daniel Bernoulli is credited with being the first to apply mathematical methods to epidemiology in 1760 by analyzing the effectiveness of the technique of variolation against smallpox [10]. A century later work continued with empirical models by Farr [37] and Brownlee [14], who fitted smallpox epidemic data to statistical distributions. A significant milestone was the work of Hamer [48] who in 1906 postulated the *mass-action principle*. This states that the course of an epidemic depends on the rate of contact between susceptibles and infecteds, which is modelled as the product of the density of susceptibles and infecteds in most cases.

Volterra was the first to write down a dynamical, mechanistic model in ecology, applied to understanding the oscillations in fish stock population in the Adriatic Sea in terms of a predator-prey relationship [91]. Let N be the number of prey and P the number of predators. Then the dynamics are described by the ODEs:

$$\begin{aligned} \frac{d}{dt}N &= N(a - bP) \\ \frac{d}{dt}P &= P(cN - d) \end{aligned} \tag{1.1}$$

In the absence of predators, the prey population N grows exponentially at rate aN. Prey are eaten at rate bNP (an application of Hamer's mass-action principle) and the predators thus grow at rate cNP. Finally, predators die at a constant rate dP. This model is known as the Lotka-Volterra model because Lotka studied the same equations in the context of chemical reactions a few years earlier; much of the history

of mathematical biology, including recent history, involves the cross-application of methods and ideas in physics and chemistry to biological systems.

This simple model is capable of exhibiting interesting dynamics such as oscillations and threshold behaviour. This is an encouraging step but the Lotka-Volterra model has its drawbacks. For one, many second-order ODEs exhibit oscillatory behaviour and so the fact that the Lotka-Volterra also exhibits oscillations may not indicate it has captured the cause of the oscillations seen in nature. Additionally the oscillatory solution curves are structurally unstable which is biologically unrealistic. Murray also points out an amusing example of real world predator-prey oscillations where, if one assumes the Lotka-Volterra mechanism is governing the population dynamics, the prey appear to be eating the predator [80]. The oscillations in this case must be caused by some other complicating factors like external forcing. Notwithstanding such complications, the application of ODEs by Volterra to a dynamical population biology model was a watershed event.

We discuss an extension of the Lotka-Volterra model to epidemiological systems in the following section.

1.4 An Epidemic Model by Kermack and McKendrick

An epidemiological system can be thought of as a kind of predator-prey system, in the sense that the infectious agent 'eats' the host by causing mortality, or in the more general sense that the dynamics of the infectious agent are dependent on the host density. The infectious agent can be roughly classified as either a *microparasite* (e.g. bacteria, viruses) or a *macroparasite* (e.g. worms) This has a bearing on how the disease transmission is modelled. We consider only microparasite-caused infectious diseases in this thesis. Microparasites are characterized by small size, short generation times, acquired immunity of the host and in some cases a short time span of infectiousness relative to the lifespan of the host. Also, the long-term macroparasite burden of an infected individual can be a function of the size of the initial inoculation, and individuals with different levels of macroparasite burden have different degrees of infectiousness and type of behaviour. Microparasites, on the other hand, quickly multiply in the host and so the final parasite burden is not as dependent on the size of the inoculation. The characteristics of microparasites, and in particular their behaviour in hosts, allows us to make the simplification of dividing individuals in the population into *compartments* according to their status: either susceptible, or infected, or recovered, etc. Models which divide the population into compartments are called (unsurprisingly) *compartmental models*. They are denoted

by a sequence of letters, one for each possible compartment. For instance in an SEIR model, individuals move from the susceptible S state to the exposed E state, where they are infected but not yet infectious. From the E state they move to the infectious I state and from the I state they move to the removed R state (immunity or death). The E state is introduced because many diseases such as measles exhibit a latency period. Often the infectivity of an individual can vary throughout the course of an infection, as for instance with HIV where peaks in infectiousness occur shortly after infection and once again before the onset of AIDS. In such cases further compartments might be necessary.

The first compartmental epidemic models for microparasites using integro-differential equations were developed in the ground-breaking work of Kermack and McKendrick in 1927 [62] (reprinted in 1991 in [63]). They looked at a model where individuals can become infected and are thereafter removed to an immune state or die. Using a generalized *force of infection* kernel which determines the rate at which susceptibles become infected, and applying the mass-action principle, they formulate the following equation of motion for the spatial density $S(t)$ of susceptibles:

$$\dot{S}(t) = S(t) \int_0^\infty \overline{A}(\tau)\dot{S}(t-\tau)d\tau \tag{1.2}$$

where $\overline{A}(\tau)$ is the expected infectivity of an individual that became infected a time τ ago. This equation is the model Kermack and McKendrick first proposed in 1927 although variants thereof for particular choices of the infection kernel are in practice analyzed most often. Klaus Dietz [24] notes that Ross and Hudson constructed a similar model in 1917 which is thought to have had some influence on the work of Kermack and McKendrick [86], but Ross and Hudson did not go into the depth of analysis that Kermack and McKendrick did.

If we choose the infectivity kernel $\overline{A}(\tau)$ to have the form of an exponentially decaying function in τ:

$$\overline{A} = \lambda e^{-\nu\tau} \tag{1.3}$$

and if we define the number of infecteds $I(t)$ at time t as $I(t) = -\frac{1}{\lambda}\int_0^\infty \overline{A}(\tau)\dot{S}(t-\tau)d\tau$ and differentiate, equation (1.2) becomes:

$$\frac{d}{dt}S(t) = -\lambda S(t)I(t) \tag{1.4}$$

$$\frac{d}{dt}I(t) = \lambda S(t)I(t) - \nu I(t) \tag{1.5}$$

$$\frac{d}{dt}R(t) = \nu I(t) \tag{1.6}$$

where $I(t)$ denotes the population density of infected individuals and $R(t)$ denotes the population density of recovered (immune) individuals. λ is the rate at which a

susceptible makes effective contact with an infected and catches the disease, and ν is the rate at which an infected is removed, corresponding to the recovery of infected individuals (but it could also be interpreted as mortality). Also note that the sum $S+I+R$ is constant since we are dealing with a fixed population density. This special case of the Kermack-McKendrick model is more widely recognizable. We can also formulate the model in terms of the numbers of susceptibles infecteds and recovereds instead of their densities–the mathematics is the same but their interpretation may differ according to population size.

By making other choices for the infectivity kernel $\overline{A}(\tau)$ we can generate any number of possible states to describe what happens to the individual after it is removed from the infected class, thus creating any number of compartments. Individuals in the population are lumped into compartments according to which state they are in at time t. Usually when researchers refer to models such as 'the SIR model' or 'the SIR equations' they mean compartmental models of this type, derived from the Kermack and McKendrick model. These kinds of compartmental models are one of the most widely studied models for epidemiological applications, in both their deterministic form and their stochastic counterparts. However there are times when we will want to retain the idea of being able to classify individuals into clearly defined disease status categories while abandoning the compartmental approach of defining transition rates between lumped population subgroups, in which case we will refer, for instance, to an 'SIS type' of infection, for a model where individuals move between infected and susceptible states but where we are not necessarily using a Kermack-McKendrick style deterministic model.

With the SIR equations (1.4) - (1.6) it is easy to analyze aspects of the epidemics such as threshold behaviour, time series, and final size. The original 1927 Kermack-McKendrick paper was followed up by further analyses in papers in 1932 and 1933.

One of the main conclusions of their analysis is that there exists a threshold susceptible density below which an epidemic dies out. This phenomenon is seen in real epidemics although the observed density at which the disease dies out is different from the critical density predicted by the mean-field compartmental models.

The existence of a threshold had been predicted two decades early by Ross [87] in a simpler mathematical model of malaria transmission formulated in 1909. Ross correctly hypothesized that malaria epidemics cannot occur if the mosquito density is below a certain value. Previously it was thought that as long as a few mosquitos were present, an outbreak could occur. This is a good example of how mathematical models can lead to new and surprising insights.

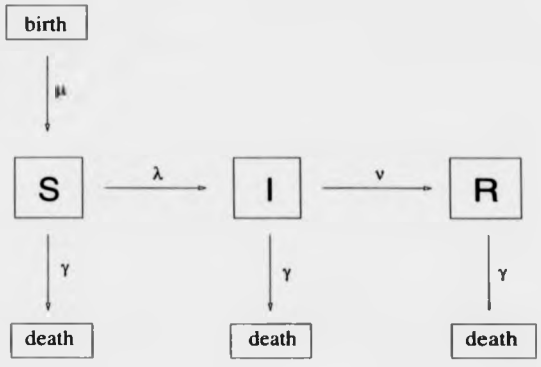


Figure 1.1: An SIR model with birth and death.

Incorporating Demographic Changes

Let us demonstrate one of the extensions of the SIR model by assuming that there is birth of fresh susceptibles into the population at rate μ, and that a removal rate γ (e.g. from death) is associated with each individual regardless of disease status. This model is depicted in figure 1.1. Introducing these changes to equations (1.4) - (1.6) produces:

$$\frac{d}{dt}S(t) = -\gamma S(t) + \mu - \lambda S(t)I(t) \tag{1.7}$$

$$\frac{d}{dt}I(t) = -\gamma I(t) + \lambda S(t)I(t) - \nu I(t) \tag{1.8}$$

$$\frac{d}{dt}R(t) = -\gamma R(t) + \nu I(t) \tag{1.9}$$

We note that at the demographic equilibrium the total population size N is equal to μ/γ.

This variant model exhibits what is known as *recurrent behaviour*; for the correct parameters there is a recurring pattern of epidemic outbreaks alternating with quiescent inter-epidemic periods (figure 1.2). After an epidemic outbreak the susceptible population is depleted and must be slowly built up again by the birth process. During this time the number of infecteds remains at extremely low numbers. Once the susceptible density is sufficiently high, a new epidemic becomes possible and there is another outbreak which persists until there are not enough susceptibles to sustain it. These repeated oscillations are damped and the system gradually settles down to an equilibrium state with a non-zero density of infecteds. This oscillatory behaviour is not possible in the system of equations (1.4) - (1.6) without birth and death.

These results are a useful illustration of the difference between the terms *epidemic*

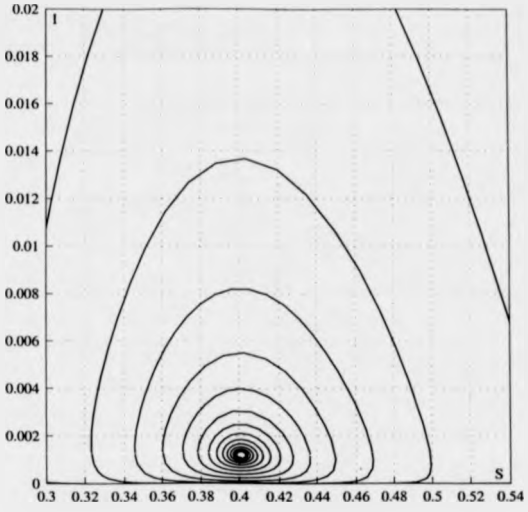


Figure 1.2: A phase portrait for the SIR equations with birth and death (1.7) - (1.9) in the I-S plane. Epidemic breakouts occur after susceptible density S reaches a sufficiently high value through immigration. Parameters: $\lambda = 25$, $\nu = 10$, $\mu = 0.02$, $\gamma = 0.02$. Initial conditions: $S = 0.99$, $I = 0.01$, $R = 0$.

and *endemic*. We say a disease is endemic when it persists in a population and the time scale of demographic change is of the same order as the time scale of infection dynamics. An epidemic situation occurs when the time scale of infection dynamics is faster than the time scale of demographic changes [6]. The transient orbits of figure 1.2 correspond to an alternation of epidemic phases where the number of infecteds explodes suddenly, and endemic phases where the number of infecteds remains unchanged while the birth process slowly builds up new susceptibles. Eventually the system reaches a steady-state nontrivial equilibrium which we can also describe as endemic. However, we note the term *epidemic* can also be used to denote a more gradual invasion of infecteds into a susceptible population as well, and some define a disease as *endemic* when the fraction of infecteds is bounded away from zero [52]. The terms are used in different ways by different researchers.

Although the SIR model with birth and death reproduces an important aspect of real epidemics, namely that of oscillations, it has drawbacks. The principal drawback is the unrealistically low numbers of infecteds which exist between epidemic outbreaks; $I(t)$ often takes values in inter-epidemic troughs such that apparently a fraction of one person is carrying the infectious agent. The disease can persist and

cause another outbreak no matter how few infected individuals there are between outbreaks, but in reality the disease is often wiped out by stochastic effects during the inter-epidemic troughs. Another unrealistic aspect of the SIR model is that the ratio of the peak infected density to the trough is larger than the ratio found in real time series for diseases such as measles [57].

Having explored one of the basic models used in mathematical epidemiology, we now turn to a more general discussion of some of the aspects of epidemic modelling and some of the tools which are used in that endeavour.

1.5 Some Aspects of Epidemiological Modelling

1.5.1 R_0

One of the fundamental topics of study in epidemiology is that of invasion. If we inoculate a few individuals with an infection in an otherwise susceptible population, under what conditions does the epidemic occur and under what conditions does it die out before infecting a significant proportion of the population?

To answer this question, mathematical epidemiologists predominantly use the *basic reproductive ratio* R_0. This quantity is defined as the expected number of secondary infections produced by an infected individual, in a wholly susceptible population, during that individual's period of infectiousness. Therefore if $R_0 > 1$ an epidemic is possible whereas if $R_0 < 1$ the epidemic necessarily dies out. The fact that the density of infecteds is initially small compared to the susceptible density is valuable for two reasons. Firstly, it corresponds to many real world situations where an infection is initially acquired by a few individuals in an otherwise susceptible population, and secondly it makes the infection process linear, since an infected individual is unlikely to have an effective contact with another infected individual in these early stages of the epidemic; therefore R_0 is easier to calculate than the solution to the final equilibrium.

The fundamental insight behind the basic reproductive ratio was discovered by Ross in 1909 [87] (cited from [50]). He developed a simple mathematical model for malaria transmission, showing the existence of a quantity which, when less than unity, implies the disappearance of malaria from the population. This quantity depended on the ratio of the density of the mosquito population to the density of the human population. Previously it had been thought that there was no such threshold behaviour, and that malaria could break out as long as at least a few mosquitos were present. This would make eradication and control of malaria virtually impossible.

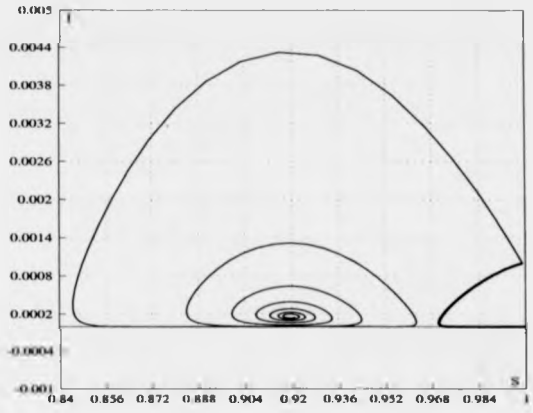


Figure 1.3: Phase portrait the SIR equations with birth and death (1.7)- (1.9) for two values of R_0. Thick line is orbit for $R_0 < 1$ and thin line is orbit for $R_0 > 1$. Parameters: $\nu = 10$, $\mu = 0.02$, $\gamma = 0.02$, $\lambda = 10.9$ for $R_0 > 1$ and 9.9 for $R_0 < 1$. Initial conditions: $S = 0.999$, $I = 0.001$ and $R = 0$ for both cases.

The hypothesis of the existence of a threshold was later confirmed in empirical studies. Ross did not originally use the symbol R_0 to denote his threshold equation, rather the symbol seems to have evolved over the course of the twentieth century until it reached its present form.

Figure 1.3 shows an S-I phase portrait for the SIR equations with birth and death (1.7) - (1.9) for $R_0 > 1$ and $R_0 < 1$.

We illustrate a derivation of R_0 for the SIR model described by equations (1.4), (1.5) and (1.6), i.e. without demographic processes. An infected individual remains infectious for a period of time $1/\nu$ on average, and in that time has an effective contact with each susceptible in the population at rate λ. Since the number of susceptibles is S, the rate at which secondary infections is produced is λS. Thus:

$$R_0 = \frac{\lambda S}{\nu} \tag{1.10}$$

Therefore there is a critical threshold susceptible density S_0 below which an epidemic cannot occur.

$R_0 > 1$ is a necessary but not sufficient condition for the disease to invade. Since the number of infecteds is often low, it is possible that the disease may be wiped out in its early stages due to stochasticity, even when $R_0 > 1$. If, on the other hand, the *percentage* of infecteds is small while the *number* of infecteds is large (so that the assumptions in the definition of R_0 still hold), then $R_0 > 1$ is a sufficient condition

for invasion since the disease is unlikely to be wiped out by stochasticity. There are other factors which may mitigate the usefulness of R_0 in ascertaining the probability of invasion, particularly the accuracy of the model on which the calculation is based. This is discussed in more depth at a later point in the thesis.

1.5.2 Variable Population Size Models

Most models assume that the population size is constant. This is valid in epidemic cases where the disease spreads quickly so that birth and death can be ignored, or endemic cases where births and deaths are balanced so that the population size remains constant over long periods. When these conditions do not hold one can take into account the changes in population size over the course of disease spread, either because one is interested in the effect of the epidemic on population size or because the population dynamics are coupled to the infection dynamics.

Models with variable population size often exhibit interesting behaviour not found in fixed-population models, such as multiple thresholds and multiple endemic equilibria [4][5][51][52] . Disease-induced mortality can stabilize or destabilize the equilibria of population size as well.

Most of the work on variable population size models have focussed on compartmental models and not on the kinds of models discussed in this thesis. We assume fixed population sizes in this thesis although some of the types of diseases we study, such as sexually transmitted diseases, can have a significant impact on both population size and other interesting aspects of demography such as patterns of social and sexual behaviour.

1.5.3 Stochasticity

The term *stochasticity* refers to randomness in dynamical processes. According to the dynamical systems perspective, stochasticity is the result of the projection of very high-dimensional phase spaces (which constitute, in principle, a full description of the real world) onto the low-dimensional phase spaces one uses in models. Stochastic effects are widespread in epidemiological systems, as in many physical systems. Sometimes these effects have only a limited impact at the population level. For instance when the numbers of infected individuals are large the fluctuations from equilibrium are small and so deterministic models serve quite well. In other cases stochastic effects can have a major effect on global dynamics, for example when numbers are small or when the structure of the phase space is complicated with many interacting basins of attraction. In these cases, of course, it would be ideal to

understand and model deterministically the microscopic dynamics which are responsible for the fluctuations. However because this can not usually be done we must use stochastic analysis in these situations.

One must consider stochasticity in deciding whether to apply deterministic models or stochastic models. Ultimately every deterministic model is an approximation to some stochastic model, because in writing down deterministic models we are averaging over microscopic states. Some information is thus excluded from the analysis.

The classic example of a fully stochastic epidemic model is the chain-binomial model of Reed and Frost formulated in a class lecture in 1952 [1] (anticipated by En'ko in 1889 [32]), although Reed did not think the model significant enough a contribution to publish [94]. In the chain-binomial model, we define S_t (resp. I_t) as the number of susceptibles (resp. infecteds) at discrete times $t = 0, 1, 2, 3 \ldots$. The population is well-mixed, and at each discrete time step the probability that a susceptible does *not* make an effective contact with an infected is p, thus the probability that a susceptible becomes infected, at time step t, is $1 - p^{I_t}$. Individuals infected at time t become infectious at time $t + 1$, and infectious individuals at time t are removed at time $t + 1$. This process continues until there are no more infecteds in the population, at which point there may or may not be any susceptibles left. Much analysis of this model and its variants has been carried out in the past century with results being obtained for distributions of the time to extinction of infecteds, number of susceptibles remaining, the existence of thresholds, etc. (see Daley [22] for an in-depth treatment). Most stochastic epidemic models in use today also rely on similar discrete or continuous time Markov chain formulations.

For spatial models the stochasticity is localized. In this thesis our approach is to define a stochastic network model of an individual-based process, and then to derive a deterministic approximation to the stochastic model. In moment closure approximations some of this localized stochasticity impacts dynamics at the population level and we incorporate it into our deterministic model, and some of the stochasticity can be modelled as random noise and is discarded. This will be explored in section 1.7.

1.5.4 Heterogeneity

Heterogeneity is an important aspect of virtually all infectious diseases. The fact that individuals have different ages, gender, personal habits, geographical locations etc. can all influence disease spread. Incorporating heterogeneous mixing of one

type or another continues to be one of the biggest challenges in epidemic modelling.

How much inaccuracy is actually introduced because of the assumption of homogeneous mixing? Anderson and May review measles data from England to answer this complex question [6]. They observe that while some evidence indicates consistency with the homogeneous mixing hypothesis for certain aspects of the epidemic, the strongest evidence is against homogeneous mixing, as shown in data indicating a variation of the force of infection with age–if individuals are mixing homogeneously they should all experience the same force of infection.

Because of the variety of factors involved in heterogeneously-mixing populations, one must choose which factors to average over and which to retain in the model. In the context of epidemiology, one way heterogeneity has its effect by altering the patterns of social mixing among individuals. Compartmental models deal with this by lumping individuals into compartments according to age, gender, etc. A matrix is defined where the (i, j) entry of this matrix tells us the rate at which an infected in group i will infect a susceptible in group j. This matrix is known as the WAIFW matrix (Who Acquires Infection From Whom).

Yet compartmental models with a WAIFW matrix are practically limited in how much heterogeneity they can incorporate. The assumption of homogeneous mixing works well for school classrooms or the workplace but for larger groups such as universities, towns, cities and countries, that assumption is less convincing. For example, each individual in a human population has on the order of 10-100 different individuals with whom they have repeated, regular social contacts (talking, handshakes), and there are even fewer established sexual partnerships per person. This situation, if modelled completely with a WAIFW compartmental model, would require hundreds of categories, making compartmental modelling impractical.

Additionally there are some aspects of social mixing which compartmental models cannot cope with, such as the repeated contacts within social and sexual networks. In such cases one must seek other modelling approaches. This is one of the challenges of this thesis. Examples of alternative approaches are given in section 1.6.

1.5.5 Waiting Times and the Principle of Universality

In most cases the probability distribution of the time to recovery is assumed to be exponential, as in equation (1.3). The same is assumed for other processes such as removal. This assumption produces ODEs instead of the integro-differential equations obtained by other distributions. Assuming an exponential distribution is usually unrealistic. For example normal distributions better describe the time to

recovery. Even so exponential distributions often work in practice, as pointed out by Hoppenseadt [53] (cited in [6]) and Grossman [47]. Another example of this, from the area of spatial modelling, is the work of Durrett et al. [29] who have proven that certain models which place individuals on a square lattice can reproduce features of dynamics in real populations which are not themselves restricted to a lattice. These observations are behind the *principle of universality*, formulated originally for statistical mechanics, which states that qualitative results at the population level are not generally affected by simplifying assumptions made about the rules for microscopic interactions.

There are nonetheless some interesting real-world situations where the assumption of exponential waiting times for disease parameters can lead to inaccuracies. For instance Keeling and Grenfell [57] study a stochastic model which incorporates normally distributed infectious and incubation periods. Their model predicts the threshold population density of measles more accurately, and also captures high-frequency components in the Fourier transform of the time series data which are not captured with a model based on an exponential distribution. So qualitative as well as quantitative inaccuracies may result when an exponential distribution is used. Thus, although the principle of universality is referred to as a 'principle', it will not always be immediately clear whether or not it can be applied to a particular problem.

Very often, being able to apply this principle can make a huge difference in how tractable models are. For infectious diseases there is a great diversity of processes occurring at microscopic scales (much more complexity than for physical and chemical systems) which may or may not be relevant for the qualitative behaviour of the system, and so the principle of universality is an important tool for mathematical epidemiology.

This issue of the relationship of microscopic dynamics to macroscopic dynamics leads to an interesting related topic, that of individual-based versus phenomenological modelling, which we now describe.

1.5.6 Individual-Based vs. Phenomenological Models

Models can be classified according to the fundamental unit of interaction. On one hand a model can be defined in terms of interactions between individuals, so-called *individual-based modelling*. In this case the time evolution of each individual is tracked. Individual-based models are usually implemented as simulations. On the other hand we can define the the basic unit of interaction on the population level, in terms of population densities and the rates of transition from one compartment of

the population to another. This is often referred to as *phenomenological modelling*, because they define phenomenological interaction terms between groups and not between individuals where the interaction actually occurs. Examples of individual-based models are given in section 1.6, and the SIR model of section 1.4 is a good example of a phenomenological model.

Individual-based models have the advantage that the individual is in nature the fundamental unit of epidemiological interaction [71][83]. Also, defining dynamics at the individual level provides a satisfyingly mechanistic foundation for a model. However these models lack analytical tractability and it is therefore much more difficult to glean insight or explore large regions of the parameter space. On the other hand, phenomenological models, although tractable, may not be able to incorporate interesting phenomena such as localized stochastic effects, some kinds of heterogeneity, etc. on account of their procedure of aggregation into subgroups. Because they define interactions between groups and not individuals (where most effective contacts actually take place) they are aptly described as phenomenological as opposed of mechanistic.

Some researchers have already sounded the death knell for analytical and phenomenological models in ecology and the dominance of the individual-based model [54][55]. They rightly point out that individual-based models can incorporate a level of detail that analytical models are not able to. It is true that analytical models are more generalized and less detailed than their individual-based counterparts. The traditional counterargument is that individual-based models are hard to interpret because of the large amount of information. But there is also the point that the capacity of analytical models to capture realism has not been fully utilized. We can make analytical models as complicated and realistic as we like, and that for each increase in computer power which allows us to create more complicated simulations, we are also able to improve the complexity of numerical solutions of equations of motion; it will still be more efficient to run a fourth-order Runge-Kutta integration on an arbitrarily large system of equations than to run the equivalent stochastic simulation and we can explore much more easily issues of stability, robustness and ubiquity. In limiting regimes one gets simple expressions for measures like R_0. The fact that analytical models are usually more generalized and less complicated than they could be reflects, to some extent, the opinions of those individuals who use them about what is worth modelling.

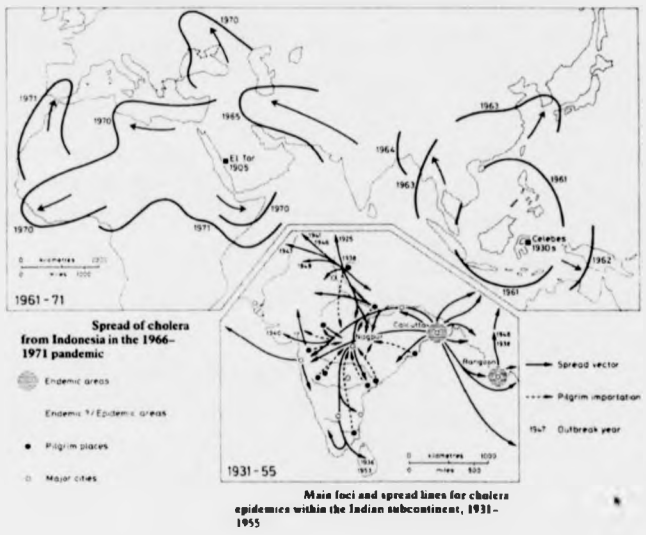


Figure 1.4: Cholera epidemic of 1966 – 1971 showing spread across Eurasia and Africa. Taken from Cliff and Haggett p. 170 [17].

1.6 Spatial Models

The aspect of epidemiological modelling which is most relevant to this thesis is modelling spatial structure. Space is a salient feature of epidemics; figure 1.4 shows the spread across Eurasia and Africa of the cholera epidemic of 1966 to 1971. Being able to predict how quickly infectious diseases spread across continents is clearly useful. Spatial models can help us devise control methods such as ring vaccination (where neighbours of an individual who is discovered to be infected are vaccinated) and to assess their potential impact on real epidemics. Moreover spatiality can influence disease dynamics in other ways, and learning about the spatial aspect of epidemics can offer rewarding insights. Finally, thinking about events occurring in space makes the problem more concrete, and the opportunity to visualize data and present them graphically is satisfying to the modeller.

Because conventional compartmental models are unable to capture spatial effects we must turn to other methods. Perhaps the most widely-known analytical spatial models are reaction-diffusion equations (RDE), which represent continuous spatial structure explicitly.

1.6.1 Reaction-Diffusion Equations

A reaction-diffusion system is a partial differential equation (PDE) with terms representing, in the continuous limit, diffusion of particles and reaction between particles of different species. RDEs take the form:

$$\frac{\partial}{\partial t}\mathbf{u} = \mathbf{f} + \nabla \cdot (\mathrm{D}\nabla\mathbf{u}) \tag{1.11}$$

where $\mathbf{u}$ is the dynamical variable of interest (for instance a density vector $\mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, \dots)$ of species $1, 2, \dots$), where $\mathbf{f}$ is the reaction term between interacting species, and where D is a matrix describing diffusion of the various species. In epidemiology RDEs are used for instance to model the spread of epidemics across continents; see Murray [80] for analysis of a model of the spread of rabies across Europe and a discussion of control methods. For infectious diseases the reaction term corresponds to the infection being transmitted from infecteds to susceptibles at a point $\underline{x}$ in a continuous space (usually one-dimensional or two-dimensional), and the diffusion term corresponds to migration of infected individuals. Thus in the RDE description a wave front of infection is caused by diffusion of infected individuals into new areas and infection of the susceptibles in those areas. For modelling rabies on the spatial scale of continents this is realistic, since rabid foxes are known to wander far afield from their home territory.

However RDE models are restricted in that they do not easily incorporate localized spatial correlations. The processes occurring at a given site are functions of the densities of organisms at that site, and the neighbourhood of position $\underline{x}$ does not affect the dynamics occurring at $\underline{x}$. Although the dynamics of an infection at location $\underline{x}$ may be influenced by migration from nearby locations in continuous space, the concept that each individual sees a different environment does not have meaning in an RDE description. For example, instead of defining the transmission of disease in terms of interactions between two individuals, it must be expressed as a product of two species densities, and the variance of the number of neighbours per person cannot be expressed as a density at a point in space either.

With discrete models it is often possible to take the limit of some discreteness parameter (for instance the size of an incremental box) to zero to obtain an RDE. The advantage of doing this is that the mathematical theory of PDEs is well-developed. However one is sometimes interested in retaining the discrete, individual-based description for reasons already given and for reasons which we discuss in future sections of this thesis. Models which retain this kind of information about individual neighbourhoods are reviewed in subsection 1.6.4 on network models.

1.6.2 Patch Models

On the way from mean-field compartmental models to explicit spatial models we find patch models. In these one considers a number of distinct patches, where each patch contains a population whose individuals mix with one another homogeneously, so that the within-patch dynamics are modelled by mean-field equations. The patches are linked with one another according to some interaction rule (for instance there is a defined rate of disease transmission, or migration is allowed). This inter-patch interaction is explicitly included in the system of ODEs which describe the situation, so that for k patches with m species each we have an order km system of equations. This is a rather neat model for archipelagoes [46] or towns and cities distributed throughout a countryside [72]. They are less useful when the distribution of the population in space is not so lumpy, or when one wants to ask what happens on the level of the individual, or when the within-patch mixing is not homogeneous at all.

We now turn from population-level phenomenological models to review the individual-based models.

1.6.3 Interacting Particle Systems

Interacting particle systems (IPS), also know as cellular automata (CA), are mostly a simulation-based approach. First developed by mathematician John von Neumann, they have since been applied to many biological, chemical and physical systems. Both IPS and their cousins the network models (discussed in subsection 1.6.4) have become popular because the impressive increase in computer power over the past few decades has made possible the simulation of complicated systems evolving over long periods of time. Also, they are easily understood and implemented, and are thus accessible to experts in a wide variety of fields.

In an IPS the population is thought of as a two-dimensional square grid isomorphic to $\mathbf{Z}^2$, with an individual at each grid point. The grid maps toroidally at the edges if we consider a finite-size population; an infinite population is usually assumed for mathematical analyses of IPS. Individuals can have a finite number of possible states. The transitions between states can be defined deterministically or probabilistically, as long as the transition probabilities for a given individual are a function of that individual's neighbourhood, i.e. the set of the 4 or 8 grid points closest to the individual. Methods of updating can vary, with systems being discrete asynchronous-, discrete synchronous- or continuously-updated in time.

IPS show a rich variety of behaviour and are particularly useful for studying pattern formation and emergent structures. Their accuracy in modelling two-dimensional

biologies in continuous space and time can be good notwithstanding their unnaturally regular structure. For instance Durrett [29] observes that an infection starting from a single grid point on a square grid has an approximately circular wavefront, as would be expected for epidemics in continuous space. Additionally, investigating IPS can provide intuition and can stimulate the formulation of hypotheses about real-world systems or help the modeller to make better deterministic approximations.

The biomathematical literature abounds with IPS models. Some examples of IPS models applied to epidemiology are those of Mollison [74] and Cox and Durrett [21].

The Contact Process

A widely used IPS is the *contact process* (CP) model developed by Harris [49]. Its popularity stems from its simplicity, and also from the fact that the CP can be interpreted as an ecological as well as an epidemiological model.

In the contact process we consider an infinite square grid. Each node can be empty or contain one individual. Individuals can give birth to any of the four nearest neighbour sites. Individuals give birth to an empty neighbouring site at rate λ, and individuals die at rate ν. For implementation in simulations we must consider a finite grid, usually with some rule for mapping the edges to one another. In the epidemiological interpretation, birth corresponds to infection and death corresponds to recovery, so that we have an SIS epidemiological process occurring on a square grid.

In this thesis we study the contact process in chapter 2 to illustrate certain aspects of moment closure approximations.

1.6.4 Network Models

One response to the challenge of developing individual-based models has been *network models*. Network models are closely related to IPS and are in some ways a generalization of them. We use the term 'network models' in this thesis to denote dynamic, irregular networks as opposed to the regular, static lattices used in IPA and the contact process. An excellent review of network models for epidemiology is provided by Andersson [7].

The population is modelled as a network where nodes represent individuals and edges represent possibilities for interaction, for instance through infection (figure 1.5). The edges of a given node are often referred to as *neighbours*. Again, individuals may be in one of a finite number of possible states and again transitions are defined according to their neighbourhood, in this case which nodes they are

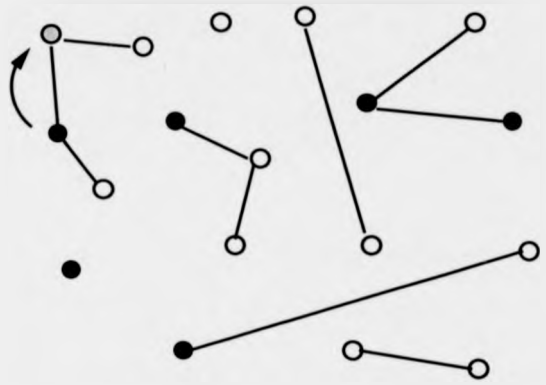


Figure 1.5: STD epidemic in a sexual partnership network model.

connected with through an edge. However with network models the network does not have to be a regular lattice embedded in one- or two-dimensional space. We can consider irregular networks, dynamic networks, networks where the strength of interaction depends on length and/or number of links, etc. The advantage of network models is that one can incorporate any of the wide variety of possible network structures and observe how it affects the spread of the epidemic. As with IPS, network models are primarily analyzed through simulation, although some limited mathematical analysis can be carried out.

Researchers have studied and attempted to apply network models to systems ranging from mobile phone networks and the internet [18][93] to sexually transmitted diseases [77]. Others have carried out questionnaire surveys to get an idea of how real world social contact networks look [30] [36], since very little data is available on this important question.

There are aspects of networks which have a significant impact on real and model epidemics, and which cannot be captured by a mean-field or IPS approach. Diekmann and Heesterbeek [25] note a few of the most obvious structural elements in networks which are relevant to epidemiology:

1. *Clusters* – members of a cluster are defined by the fact that they are more interconnected with one another than with those outside the cluster;
2. *Loops* – the size and number of closed loops of various lengths may be different from one network to the next;
3. *Degree distribution* – the distribution of number of edges per node can vary.

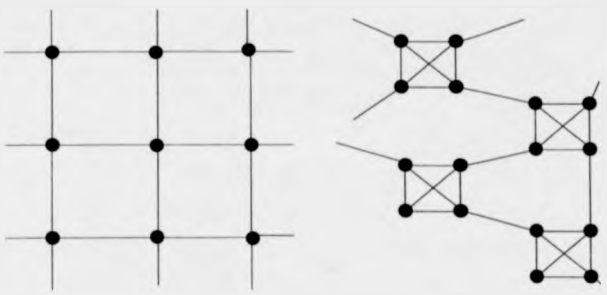


Figure 1.6: A comparison of two networks with identical degree distributions but with different amounts of clustering. The left- hand network shows the case of no clustering whereas a large amount of clustering is present in the right-hand network; in this case the clustering seems to reduce the dimensionality of the network from two-dimensions to a network which is almost one-dimensional.

These structures can radically alter disease dynamics. For instance consider the characteristic path length L of the network, which tells us the expected number of nodes we must go through, on average, in tracing a path via edges from one node to another node in the network. L reflects the distribution of the number of loops in the population and the clustering pattern, and is an important measure for epidemiology because it tells us roughly the number of individuals a disease must pass through to get from an index case to any other part of the network. A few random connections between nodes in a network which has an otherwise long characteristic path length (for instance a square grid) can greatly reduce L. A network with a few of these random connections in an otherwise regular network is known as a *small-world network*. Watts and Strogatz find that the neural network of *C. Elegans*, the power grid of the United States and the network of acquaintances amongst film actors are small-world networks, and they argue that small-world networks are widespread in natural and man-made systems [93].

We also give an example of how clustering can change disease spread. Consider a network where each individual has the same number of edges, but where there is a large amount of clustering. In other words, an individual's neighbours are likely neighbours to one another as well. This clustering can slow down the spread and decrease the final size of an epidemic [75], since the disease becomes 'trapped' in clusters for some length of time and potentially infectious contacts are expended on other individuals in the cluster who are already infected.

Clearly, understanding how network structures such as loops and clusters affect

epidemic spread is an important problem. It is also a difficult problem as far as finding an analytical approximation is concerned. One way to overcome this difficulty is to consider certain aspects of networks apart from the full network treatment (see Diekmann and Heesterbeek [25] for an example of this approach). Another way is moment closure approximations. This thesis makes some progress in incorporating various types of network structures into an analytical form via such approximations. Moment closure approximations are discussed in greater depth in section 1.7.

Networks and Geographical Space

A short discussion of the relationship between geographic space and networks is in order. Sometimes network models are invoked to model populations which are well-described by a non-geographical network, as with sexual partnership networks. Morris and Kretzschmar apply a network model for studying STDs [78]. This is generally more accurate for diseases which spread through social connections that are not necessarily geographically distributed. Another example of this is the network of acquaintances among workers in an office building. On the other hand, most IPS models intend that the square lattice description be an approximation to some continuous two-dimensional geographical space such as trees in a field or towns across a countryside. This distinction between the two types of networks is important, and is reinforced by the observation that increasing the population density of a city will increase the prevalence of density-dependent diseases such as flu or measles, but will not generally change the prevalence of STDs [6], since the partnership network is not much affected.

However there is a fundamental unifying feature of network descriptions and explicit spatial descriptions, namely the idea that different individuals experience, or 'see', different environments according to their location in the network. On the other hand, according to mean-field compartmental models, everyone experiences the same environment averaged over the entire population. Therefore we apply the terms *space* and *spatial* to both the pure network description and to geographical spatial models.

1.6.5 Pair-Formation Models

After development of network and IPS models started, researchers sought analytical treatments to gain more insight into these models. One such group of dynamical models are known as *pair models* or *pair formation models* and are used to incorporate the repeated contacts within relationships which occur in social and sexual

networks. This approach is particularly useful for studying STDs. It was developed for STDs by Dietz and Hadeler in 1988 [27] and will be discussed in more depth in section 1.8 on modelling sexually transmitted diseases.

Pair formation models can be seen as a special case of moment closure approximations, which are explained in the next section.

1.7 Moment Closure Approximations: To Second Order and Beyond

In the moment closure approximation we attempt to capture spatiality by employing pairs, triples and other high-order correlations as state variables and seeking the corresponding equations of motion. When we derive the equations of motion for correlations of a given level, we will find that it is necessary to know higher-order correlations to close the equations of motion. In fact what we have is an infinite hierarchy of equations of motion for correlations at every order. Thus, in practice, high-order correlations must be approximated as functions of lower-order correlations to close the equations of motion. We accomplish this closure by a suitable expression of the higher-order correlations in terms of lower-order correlations. We may also have to approximate combinations of higher-order correlations as random fluctuations around some mean value. The mean can hopefully be expressed in terms of lower-order correlations and incorporated into the equations of motion, and the remainder can be discarded or incorporated as noise into a stochastic differential equation version of the approximation. It turns out that higher-order correlations are in most cases wiped out by stochasticity, and so very often it is sufficient to use lower-order MCA to get an accurate approximation.

In the case where only singletons (e.g. the number of infecteds) are considered, we recover the mean-field equations, whereas increasing the order of the approximation to pairs, triples, etc. brings us closer to a full network model. So moment closure approximations can be thought of as occupying a middle ground between the extremes of individual-based and phenomenological models, since they consider structures which are larger than individuals and smaller than the entire population.

1.7.1 The Pair Approximation

The most widely used moment closure approximation is the *pair approximation* (PA), where pairs (numbers of edges of a given type) and singletons (numbers of individuals of a given state) are the state variables.

The first application of pair approximations was in solid-state physics by researchers such as Kikuchi [64], Bethe [11] and Burley [15]. More recently Nord and Evans [81] have applied them to dimer adsorption problems in chemical physics. Moment closure methods are also employed in the study of fluid dynamics, in particular in the modelling of turbulence where the correlations between different points in the velocity field are tracked [82].

Researchers at Kyushu University were the first to apply pair approximations to ecological systems [56][73][88], initially to a spatial version of the Lotka-Volterra model. Pair approximation models have since been used to study a wide variety of problems in ecology, such as altruism [90], measles [58], invasion [70][85] and spatial games such as the hawk-dove game and the prisoner's dilemma [75]. In this section we give a basic outline of the moment closure technique in a biological context. We illustrate it in greater depth with particular examples in the other chapters of this thesis. A survey of the pair approximation technique can be found in Rand [84].

Consider a network where nodes represent individuals and edges represent possibilities for interaction between nodes, for instance through competition or infection. Let f be a real-valued function of the state of the network at time t which can be approximated as continuous. The equation of motion for f is derived by summing over all events in the population which affect f, and the amount of change produced by those events:

$$\frac{d}{dt}f = \sum_{e \in Events} r(e)\Delta f_e \tag{1.12}$$

where $r(e)$ is the rate of event e, and Δf_e is the change induced in f by event e.

In evaluation of equation (1.12) we sum over each node in the network. At each node we express the rates $r(e)$ and the change Δf_e at that node in terms of population-averaged expectation values of $r(e)$ and Δf_e plus the deviation of those values from the expectation values at that node. We must perform the summation and choose our substitution so that any effect of stochasticity which is nonzero when summed over the population is incorporated into the expectation values of $r(e)$ and Δf_e. Then, any left over stochasticity can be treated as random noise.

We denote the states of nodes by letters i, j, k, etc. and the number of such nodes in the population by $[i]$, $[j]$, $[k]$, etc. We want to derive the equations of motion for quantities such as $[i]$. When deriving the equation of motion for $[i]$, we end up with terms with quantities such as $[ij]$ the number of edges connecting an i node to a j node. For instance in an epidemic model the rate at which susceptibles are infected is proportional the number of susceptible-infected edges. So we end up

with an equation of motion for $[i]$ which has the form:

$$\frac{d}{dt}[i] = \frac{d}{dt}[i]([i], [ij], \dots)$$

Clearly we must know what the $[ij]$ quantities are in order to obtain an ODE for $[i]$. Under a mean-field assumption, we use the approximation $[ij] \propto [i][j]$ thus eliminating any information about the impact of the network structure on the disease spread. However with the PA we accept a marginal increase in complexity by retaining $[ij]$ as a state variable, and we write down its equation of motion. Since in an ij pair the j individual can be affected not just by the i neighbour but also by others, we are led to incorporation of triples $[ijk]$ when we write down the equations of motion for $[ij]$:

$$\frac{d}{dt}[ij] = \frac{d}{dt}[ij]([i], [ij], [ijk], \dots)$$

Continuing in this fashion produces an infinite hierarchy of equations, but for pair approximations we close the equations of motion at the pair level, finding ways to approximate triples in terms of pair and singleton quantities. Note that as we move up in the hierarchy, the number of required equations increases drastically: for two possible disease states the pair, (resp. triple, quadruple) approximations require $2^2 - 2^0 = 3$, (resp. $2^3 - 2^1 = 7$, $2^4 - 2^2 = 12$) equations of motion before symmetries. So there are practical reasons not to go far beyond the pair level in addition to the fact that at some point (depending on the system under study) higher-order correlations are swamped by stochastic noise. However it is sometimes possible to boil the model down to a lower-dimensional set of equations by retaining the most important higher-order correlations and approximating the others (see Keeling et al. [58] for an example) with a minimal decrease in accuracy.

As important as the decision about where to close the hierarchy is deciding how to carry out the closure. The most widely used method is to assume conditional independence of the various neighbours of an individual from one another. Katori et al. [56] formulate this assumption in the following way. Let $q_{\sigma/\sigma'\sigma''}$ be the conditional probability that under a condition that a neighbour of a σ' site is a σ'' site, another randomly chosen neighbour of a σ' site is a σ site. Under the assumption of conditional independence we have

$$q_{\sigma/\sigma'\sigma''} = q_{\sigma/\sigma'} \tag{1.13}$$

This is known as the *ordinary pair approximation* (OPA) because we truncate the hierarchy at the level of pairs. Levin and Durrett [70] also make use of the OPA in their study of epidemics. Originally derived for the contact process, we will use it

in this thesis to denote any pair approximation where this assumption of conditional independence is made. Note, however, that this assumption does not free us from taking the local underlying network structure into account. In other words, knowing that a σ' site has a σ'' neighbour *does* tell us that there are at most $Q-1$ other σ neighbours, where Q is the number of neighbours per node, since at least one neighbour is a σ'' site.

The implications of this for our moment closure derivations is made more clear by using the notation of Morris, Keeling and Rand which we use throughout the rest of this thesis (see [75][84] for examples). Here we explain their notation and the form taken by the OPA under their notation for two widely used types of network structures, regular lattices and random networks. The moment closures are derived rigorously by Morris [75] but we only give a heuristic justification here.

Let $Q_x(i)$ denote the number of state i neighbours of node x. Similarly let $Q_{xy}(i)$ denote the number of state i neighbours of a node x which has node y as a neighbour. Also, let us denote the state of node x by ς_x and the state of an edge involving x and y by ς_{xy}. As mentioned in the introduction to this section, in order to close the equations, for certain quantities we must substitute the values of those quantities at a given node by their population-averaged value plus the deviation from the average at that node. So let $Q(i \mid j)$ be the population-averaged value of $Q_x(i)$ when $\varsigma_x = j$, and let $Q(i \mid jk)$ be the population-averaged value of $Q_{xy}(i)$ when $\varsigma_{xy} = jk$, in other words it expresses the expected number of state i neighbours of a state j individual who has at least one state k neighbour. Since we are seeking a pair approximation we need to approximate third-order correlations such as $Q(i \mid jk)$ in terms of second-order correlations. We note at this point that in our derivations in this thesis we will sometimes express third-order correlations in terms of numbers of triples $[ijk]$ and sometimes in terms of expectation values such as $Q(i \mid jk)$. The following identities relating the two apply for all network types:

$$\begin{aligned} Q(i \mid jk) &= \frac{[ijk]}{[jk]} \qquad\qquad i \neq k \\ Q(i \mid ji) &= 1 + \frac{[iji]}{[ij]} \end{aligned} \tag{1.14}$$

and for lower-order quantities:

$$\begin{aligned} Q(i \mid j) &= \frac{[ij]}{[j]} \\ Q(i \mid i) &= \frac{[ii]}{[i]} \end{aligned} \tag{1.15}$$

where the edge numbers are defined by:

$[ij]$ the number of edges connecting a state i individual to a state j individual where $i \neq j$.

$[ii]$ twice the number of edges connecting a state i individual to another state i individual.

$[ijk]$ the number of triples connecting a state j individual to individuals of states i and k, $i \neq k$.

$[iji]$ twice the number of triples connecting a state j individual to two state i individuals.

$\vdots$

Consider first the situation where we have a regular network, i.e. where the number of neighbours Q per node is the same for all nodes. To find $Q(i \mid jk)$ we must know how $Q_{xy}(i)$ is distributed across the population. Given that we assume conditional independence and given that Q is constant, a binomial distribution is appropriate; determining the distribution of neighbours is like a problem of sampling without replacement with probability of success $Q(i \mid j)/Q$. We can derive the expression for $Q(i \mid jk)$ rigorously under this assumption, but from the assumption of conditional independence it is clear on heuristic grounds that:

$$\begin{aligned} Q(i \mid jk) &= \frac{Q-1}{Q} Q(i \mid j) \\ Q(i \mid ji) &= 1 + \frac{Q-1}{Q} Q(i \mid j) \end{aligned} \tag{1.16}$$

where $i \neq k$. The first equation can be understood by observing that since the state j individual has already at least one $k \neq i$ neighbour, he can have at most $Q-1$ state i neighbours. Hence the fraction $Q-1/Q$. We are also making the implicit assumption that the state of the k individual does not influence the states of the other neighbours of the j individual. The reasoning for the second equation is similar.

The expression is slightly different for a random network where $Q_x(i)$ is assumed to obey a Poisson truncation which also allows for conditional independence in neighbour status. In this case the higher order correlations take the form:

$$\begin{aligned} Q(i \mid jk) &= Q(i \mid j) \\ Q(i \mid ji) &= 1 + Q(i \mid j) \end{aligned} \tag{1.17}$$

Also, we note this expression is more suitable for an OPA applied to a dynamic network. Although we refer to both equations (1.16) and equation (1.17) as OPA,

sometimes one expression will be suitable and sometimes another. Therefore we also refer to binomial OPA and Poisson OPA at times.

Other closures which are more sophisticated than OPA have since been developed, and developing and analyzing better closures is one of the themes of this thesis. Once we have obtained a closure such as binomial OPA or Poisson OPA we have a set of relatively simple ODEs we can proceed to analyze. A full derivation of an MCA is presented in chapter 2.

Because of the errors which may be introduced by moment closure, and especially because we cannot get an error estimate directly from the deterministic approximations (except for some kinds of moment closures, see Bolker and Pacala [12] for an example), the MCA is best used in conjunction with simulations.

Often there are constraints on the state variables and these constraints can influence how the moment closure is made. Constraints arise if the network is static or if a dynamic network has reached an equilibrium, allowing us to impose a constraint. For a static network model where individuals can be in one of two possible states i and j, the following constraint on pair numbers holds:

$$\overline{Q}N = [ii] + [jj] + 2[ij] \tag{1.18}$$

where $\overline{Q}$ is the average number of edges per node and N is the population size. Note that the network does not have to be regular as long as $\overline{Q}$ is known. For the equilibrium of a dynamic network we usually calculate the equilibrium value of the number of edges L, so that the above constraint takes on the form::

$$2L = [ii] + [jj] + 2[ij] \tag{1.19}$$

Triples $[ijk]$ and higher order correlations can be similarly constrained. These identities can be derived straightforwardly by thinking about the structure of networks. If necessary, the constraints can be extended to models where individuals have more than two possible states, but since this thesis focuses on SIS models we do not derive a general form of the constraint for an arbitrary number of possible states. The total population size N is also constrained for all models we consider in this thesis:

$$N = [i] + [j] \tag{1.20}$$

and we note that the following identities apply which relate singleton to pair numbers:

$$\begin{aligned} [i] &= ([ij] + [ii])/\overline{Q} \\ [j] &= ([ij] + [jj])/\overline{Q} \end{aligned} \tag{1.21}$$

PA for Epidemiological Systems

Tracking pairs is a natural thing to do for epidemiological systems since the infection is transmitted through pairwise contacts; our main parameter of interest then is the probability of transmission per unit time in a susceptible-infected partnership. This description is closer to the fundamental event of real epidemics–transmission of infection between two individuals–than the phenomenological transmission rates defined between compartments in compartmental models or the density differences in RDE models.

An example of PA applied to epidemics is the study of measles by Keeling et al. [58][61] where a pair approximation to an SEIR network model is compared to the behaviour produced by a classical mean-field SEIR model. It is found that the pair approximation model captures aspects of real-world measles dynamics which are lacking in the SEIR model, for instance regular periodicity and less violent oscillations. Additionally the SEIR model predicts more frequent extinctions of the disease than is realistic.

However large-scale network structures such as loops and clusters are clearly more difficult to capture with pair approximations. These large-scale structures can be important for epidemic spread as noted in the subsection on network models, and this issue will be discussed in more depth in section 1.8 on STDs.

1.7.2 Higher-Order Clusters

A number of modellers have in fact gone one step further than the pair approximation and studied models with triples, quadruples or other higher-order correlations [15][64][75]. The motivation is that higher-order correlations are relevant in many systems, especially near critical points, so using higher-order approximations becomes necessary. Bolker and Pacala [12], for instance, note that higher-order approximations might be necessary to capture spatial patterns with sharp boundaries such as wavefronts. An epidemic model based on higher-order approximations is presented in chapter 2, where we see that such approximations can provide some advantages and have qualitatively different behaviour from the PA.

Before returning to the topic of MCA in chapter 2, we discuss in our final section of the introduction issues related particularly to the modelling of sexually transmitted diseases.

1.8 Modelling Sexually Transmitted Diseases

According to the June 2000 *UNAids Report on the Global HIV/AIDS Epidemic* [33], 5.4 million people contracted the HIV virus in 1999 and 2.8 million died from AIDS that same year. The report notes that for some African countries

> there is now compelling evidence ... that the trend in HIV infection will have a profound impact on future rates of infant, child and maternal mortality, life expectancy and economic growth ... AIDS is unique in its devastating impact on the social economic and demographic underpinnings of development.

Interestingly, the report notes that the 1991 forecast of the disease prevalence for 2000 was underestimated by a factor of three. Comparisons with the Black Death of the 14th century are sounding less alarmist.

Even less fatal STDs which have been endemic for a long time in human populations can still cause serious health problems. *Chlamydia trachomatis* and *Neisseria gonorrhoeae* are causative factors in pelvic inflammatory disease, sterility and ectopic pregnancies in women, who often are asymptomatic carriers of the bacterium [44]. There are compelling reasons to better understand STD epidemiology.

1.8.1 What is Special About Modelling STDs

STDs differ from other infectious diseases in ways which affect how they are modelled:

1. Carriers (especially female carriers) are often asymptomatic until much later in the course of infection.
2. There is little or no acquired immunity.
3. STDs are restricted to the sexually active population.
4. The number of individuals from which one can catch the disease is usually lower than for many other diseases.

The last point suggests that the sexual contact structure of the population will be particularly important and that the traditional mean-field assumption will be less viable for STD epidemiology. Indeed, there has recently been great interest in the role played by the pattern of sexual contacts in the spread of sexually transmitted diseases (see [28][43][78][92]) and our discussion of network models in subsection 1.6.4 gave some theoretical reasons as to why the contact structure is significant.

Even when considering compartmental models for STDs the choice of the WAIFW matrix is particularly important and may reflect age, race, geographical and socioeconomic factors. Mixing can be either assortative, where individuals in a group mix with one another, or disassortative, where individuals mix with people outside their own group. Anderson and May [6] point out that

> ...most sexually transmitted diseases (STDs) cannot be understood without acknowledging the marked heterogeneity in degrees of sexual activity within the overall population.

The importance of the sexual partnership network structure is highlighted by the example of the AIDS pandemic. In Africa HIV is widespread among the heterosexual population, whereas in other countries it is predominantly found among homosexuals and IV drug users, both groups which engage in high-risk behaviour (although recently inroads have been made into the heterosexual population in India [35]). Some authors feel that the different demography of HIV in these two cases reflects the difference in the sexual partnership networks. In Uganda for instance data from high HIV seroprevalence areas indicate a low number of lifetime sex partners for both males and females, yet very high rates of *concurrency*, where individuals have more than one partner simultaneously [76] (cited in [77]). This is from the widespread practice of married men taking mistresses. On the other hand in most countries serial monogamy is the median behaviour, *although* the total lifetime number of partners is comparable to that of Uganda. The enormous difference in HIV demography in these two cases could be due to the significant differences in the distribution of sexual partnerships in the populations.

The importance of the structure of sexual partnership networks, and the role of mathematical models in stimulating and reinforcing this observation, has also been noted by those individuals who are close to HIV control programmes [34], who have devised questionnaires in order to get an idea of how real sexual networks look.

Sexual partnership network structure also has implications for the infection risk of individuals independent of the risk introduced by their own sexual behaviour. The rise in HIV among heterosexuals in India noted by Rodrigues [35] is among monogamous married women who claim no history of sex working. They have contracted the disease from their husbands who are part of high-risk core groups of prostitutes and their clients. Figure 1.7 illustrates how the placement of an individual can impact their risk of getting an STD.

Collecting data on sexual networks in the general population is difficult because one must know the identities of each individual's sexual partners. *Contact tracing*,

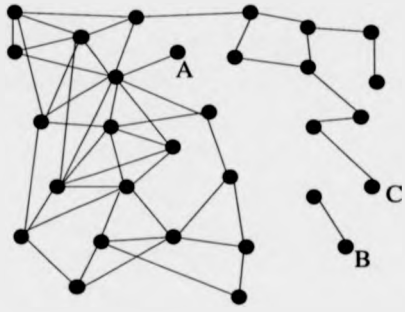


Figure 1.7: An illustration of the risk of acquiring an STD relative to one's location in a sexual partnership network. Individual A is at much greater risk than either individual B or C, although all three individuals have only one partner.

where the partners of infected individuals are themselves screened for infections, gives us limited information about sexual networks for certain subgroups of the population (those who come to STD clinics for treatment) [79] but the network structure of high-risk groups is not representative of the general population, and contact tracing usually only goes through a few nodes before stopping. So modelling is particularly important when it comes to STDs and sexual networks because of the lack of information about network structure on the one hand and the importance of network structure in STD transmission on the other hand. The best which can be achieved is to embody some assumptions about the global structure and to see whether the resulting dynamics can supply an explanation for real-world epidemics.

The natural history of STDs mixes two dynamical processes. Firstly, there are the partnership dynamics in which partnerships are formed and broken. This can be regarded as producing a dynamical network in which the nodes represent individuals and where edges connect individuals in a partnership. On top of this we have an infection process which is constrained by the network in that only an individual's partners can be infected. It is the combination of these two dynamical processes that makes the mathematical analysis of such systems more difficult than for disease models that assume homogeneous mixing [6] or a fixed network [84].

Bearing in mind these facts about the distinguishing features of STDs, we now describe the main classes of STD models and their historical development.

1.8.2 Prologue: Compartmental Models

The first attempt at constructing a dynamical model for STDs was by Cooke and Yorke [19], who developed a compartmental model of gonorrhea transmission. This

significant work was aimed at public health workers in the hope that it would have an impact on epidemic control. Since then the advent of the HIV epidemic has stimulated further research into all areas of STD epidemiology, including the use of mathematical models to understand their transmission dynamics.

However there are some aspects of STDs that cannot modelled with compartmental models. One of their weaknesses is their assumption that each individual in a compartment makes an instantaneous one-time sexual contact with one of many individuals in the other compartments. Yet in reality, individuals make repeated contacts with a small set of other individuals they are in partnership with. This repetition of contacts is present in non-sexual social networks as well but the node degree and characteristic path length L for those are shorter and hence the mean-field assumption is not as unrealistic.

Dietz and Hadeler attempted to remove this deficiency with a new class of models called pair models (not be be confused with the pair approximation). We mentioned these in passing in subsection 1.6.5 and describe them in more depth now.

1.8.3 Pair Models Come Onto the Scene

Dietz and Hadeler introduced pair models into STD epidemiology in 1988 [27]. These models allow for monogamous partnerships between individuals through which the STD can be transmitted during the duration of the partnership. Either a transmission rate per unit time is defined, or there is a per-partnership transmission probability. Thus one can study the impact of partnerships on STD epidemiology, something not possible with a compartmental model. It was thought that the existence of partnership could slow down the spread of disease since an infection is 'locked' into the partnership for its duration and once both partners are infected, effective contacts are wasted on individuals who are already infected. Indeed, they found it is possible for partnerships to slow the spread of STDs as in comparison with compartmental models. Hans Heesterbeek, using the model of Dietz and Hadeler, extends a calculation of R_0 for compartmental STD models to pair STD models [50].

Pair models are not without drawbacks, the primary one being ambiguity in the definition of a partnership. If two individuals have two sexual contacts separated by one year, does that count as two partnerships separated by some length of time or is it counted as one long partnership? How quickly does the probability of transmission within a partnership approach unity as contacts are repeated throughout the duration of the partnership? Should we define a per partnership or per unit time transmission

probability? In the formulation of Dietz and Hadeler a partnership starts with a sexual contact, however it would also be possible to define a partnership according to a certain rate of contact occurring over time for the duration of the partnership, as in [66] or even just a rate of disease transmission for its duration. Defining an exponentially distributed waiting time for disease transmission in a partnership is standard, and has the advantage that the equations of motion become ODEs instead of integro-differential equations. Debate continues over the relationship between the number of effective contacts per partnership and the probability of transmission per partnership [6].

Although compensating for a major weakness of compartmental STD models, pair models still do not allow for the highly structured sexual partnership networks of real populations, because they cannot incorporate the existence of simultaneous partnerships. We want to know about the impact of network structures such as closed loops, clusters and *core groups* of high-activity assortatively-mixing individuals. Can we create IPS models, or better still network models for sexually transmitted diseases which take these interesting features into account?

1.8.4 Enter Network STD Models and Moment Closure Approximations

Thinking of a sexually active population as a network is natural for STDs. Modelling STD transmission with a network model gives us freedom to define structures corresponding to various possible sexual behaviours defined on the individual level and we can assess their impact on network structure and STD epidemiology. The structure of a real-life sexual partnership network is much better represented by a network model than a compartmental or pair model. And the choice of the individual as the unit of interaction is a natural choice for STDs as for all infectious diseases. One can control the distribution of partnerships per person, their duration, one can distinguish between different types of partnerships, etc. This mechanistic foundation gives the network models a satisfying concreteness not possible with compartmental models.

Kretzschmar and others [69] have used simulation models to predict the effect of various prevention strategies on the spread of gonorrhea and chlamydia. They have also looked at the impact of various network structures on STD spread. However network models have the drawback of being analytically intractable. Although Kretzschmar [67] made some inroads into combining insights from deterministic pair models and stochastic network models, and one can also use results from percolation

theory [89], what is still needed is an approach to constructing deterministic ODE approximations to network models, in order to extend their usefulness in gaining insights into STD epidemiology. Moment closure approximations seem to be the way to proceed at this point.

Heterosexual versus Homosexual Networks

Before continuing in this endeavour we make an aside about the use of homosexual versus heterosexual networks. This is clearly important for cases where sexual behaviour varies between the two groups or where parameters such as recovery and transmission rates vary according to gender, or when age-structure and demography must be accounted for. But consider the hypothetical situation where the two sexes are identical in all respects, so that the only difference between a homosexual and a heterosexual network is the fact that in a heterosexual network a person can only form partnerships with members of the opposite sex. How does this affect epidemiology? We carry out here an impromptu analysis based on the moment hierarchies. We do not derive the moment hierarchies here but a derivation for the contact process is presented in chapter 2.

Consider the following STD network model:

1. There is a 1:1 sex ratio in the population.
2. All epidemiological parameters are the same for males and females–in particular, the recovery rate to the susceptible class is ν for both and the transmission rate is λ for both.
3. Only heterosexual partnerships are possible.
4. Partnerships break up at rate σ and any two male and female individuals form a partnership at rate $2\rho/N$, where N is the population size.

We compare the moment hierarchy which describes this network model with the hierarchy which describes an equivalent homosexual network. By *equivalent* we mean the epidemiological parameters are the same, and the partnership formation parameters are such that the total rate of partnership dissolution and the equilibrium number of partnerships are also the same. Thus for an equivalent homosexual STD network model we require:

1. The infection is transmitted between partners at rate λ and recovery to the susceptible class occurs at rate ν.

2. Partnerships break up at rate σ and any two individuals form a partnership at rate ρ/N.

We adopt the standard notation to describe the network variables with subscripts to denote male or female. For instance $[I_m S_f]$ denotes the number of partnerships with an infected male and a susceptible female. We expect for a heterosexual network at equilibrium that:

$$\begin{aligned} [I_f] &= [I_m] \\ [I_m S_f] &= [I_f S_m] \\ [I I_m S] &= [I I_f S] \\ &\vdots \end{aligned} \tag{1.22}$$

since the partnership and transmission dynamics are the same for the two sexes; they are essentially mirror images of one other. We will use this assumption in the following analysis.

The equation of motion for the number of infecteds $[I]^{hom}$ in the homosexual network is:

$$\frac{d}{dt}[I]^{hom} = -\nu[I]^{hom} + \lambda[SI]^{hom} \tag{1.23}$$

The equations of motion for $[I_m]$ and $[I_f]$ for the heterosexual network are:

$$\begin{aligned} \frac{d}{dt}[I_f] &= -\nu[I_f] + \lambda[S_f I_m] \\ \frac{d}{dt}[I_m] &= -\nu[I_m] + \lambda[S_m I_f] \end{aligned} \tag{1.24}$$

Using the identities $[I]^{het} = [I_m] + [I_f]$ and $[SI]^{het} = [S_m I_f] + [S_f I_m]$ the equation of motion for the total number of infecteds in the heterosexual network is:

$$\frac{d}{dt}[I]^{het} = -\nu[I]^{het} + \lambda[SI]^{het} \tag{1.25}$$

which is identical to equation (1.23) for the homosexual network. At the level of pairs the partnership formation dynamics come into play. We do not present here an in-depth derivation of the equations of motion for pairs, but the reader can refer to section 2.2 to better understand the derivation. The result of this derivation is the equation of motion for $[SI]^{hom}$:

$$\begin{aligned} \frac{d}{dt}[SI]^{hom} &= -\sigma[SI]^{hom} + \frac{\rho}{N}[S]^{hom}[I]^{hom} + \nu[II]^{hom} - \nu[SI]^{hom} \\ &\quad +\lambda[ISS]^{hom} - \lambda[ISI]^{hom} - \lambda[SI]^{hom} \end{aligned} \tag{1.26}$$

For the heterosexual network, the equations of motion for $[S_f I_m]$ and $[S_m I_f]$ are:

$$\begin{aligned}
\frac{d}{dt}[S_f I_m] &= -\sigma[S_f I_m] + \frac{2\rho}{N}[S_f][I_m] + \nu[I_f I_m] - \nu[S_f I_m] \\
&\quad +\lambda[S_f S_m I] - \lambda[I S_f I_m] - \lambda[S_f I_m] \\
\frac{d}{dt}[S_m I_f] &= -\sigma[S_m I_f] + \frac{2\rho}{N}[S_m][I_f] + \nu[I_m I_f] - \nu[S_m I_f] \\
&\quad +\lambda[S_m S_f I] - \lambda[I S_m I_f] - \lambda[S_m I_f] \qquad (1.27)
\end{aligned}$$

Combining them with the identities $[SI]^{het} = [S_f I_m] + [S_m I_f]$ and $[II]^{het} = [I_m I_f] + [I_f I_m]$ gives, after some simplification:

$$\begin{aligned}
\frac{d}{dt}[SI]^{het} &= -\sigma[SI]^{het} + \frac{\rho}{N}[S]^{het}[I]^{het} + \frac{\rho}{N}([I_m] - [I_f])^2 + \nu[II]^{het} - \nu[SI]^{het} \\
&\quad +\lambda[ISS]^{het} - \lambda[ISI]^{het} - \lambda[SI]^{het} \qquad (1.28)
\end{aligned}$$

which is identical to equation (1.26) except for the $\frac{\rho}{N}([I_m] - [I_f])^2$ term. If we are considering an equilibrium situation where we can assume $[I_f] = [I_m]$ (equation (1.22)), equations (1.26) and (1.28) are identical. Whenever $[I_f] \neq [I_m]$ the $[SI]^{het}$ dynamics are different from the $[SI]^{hom}$ dynamics, even if $[I]^{het} = [I]^{hom}$. Thus the fact that the network is heterosexual instead of homosexual can have a non-trivial effect on the transient dynamics of $[SI]$ and $[I]$. To settle definitively the question of the impact on transient behaviour one would have to investigate the stability of the $[I_m] = [I_f]$ manifold.

However this still does not answer the question of the differences in equilibria, since we do not know generally if $[ISI]^{hom} = [ISI]^{het}$ or if $[ISS]^{hom} = [ISS]^{het}$. However if we use a Poisson OPA, we get:

$$\begin{aligned}
[ISI]^{het} &= \frac{[SI]^{het}[SI]^{het}}{[S]^{het}} + [SI]^{het} \\
[ISS]^{het} &= \frac{[SS]^{het}[SI]^{het}}{[S]^{het}} \\
[ISI]^{hom} &= \frac{[SI]^{hom}[SI]^{hom}}{[S]^{hom}} + [SI]^{hom} \\
[ISS]^{hom} &= \frac{[SS]^{hom}[SI]^{hom}}{[S]^{hom}}
\end{aligned}$$

and then we find that the equilibrium behaviour of the solutions $[I]^{het}$ and $[I]^{hom}$ are the same (with constraints in accordance to the definition of the two models), since equations (1.26) and (1.28) with the above closures are identical. A more rigorous argument would require induction on the k^{th} level of the hierarchy of equations of motion.

Therefore the existence of heterosexuality in a network can change the transient behaviour of the total number of infecteds but not the equilibrium, a conclusion

which does not disagree with what our intuition tells us. A general theory based on moment hierarchies about the effect on epidemiology of the partition of the population into k disassortative-mixing subpopulations in network models would prove a useful tool and should be feasible. In this thesis we use homosexual networks exclusively but we keep in mind the differences with heterosexual networks which the models are actually intended to study.

1.9 Summary

In this introductory chapter we have reviewed some of the history and background to epidemic modelling, the basic ideas behind moment closure approximations, and some of the particularities of modelling sexually transmitted diseases.

In the next chapter we will explore in greater depth the use of moment closure approximations in modelling invasion and the behaviour of MCA near the threshold values, i.e. near *critical points*. We will use the contact process as an example to illumine these features of moment closure approximations.

Chapter 2

Critical Points, Invasion and Moment Closure Approximations

For every complex problem, there is a solution that is simple, neat, and wrong.

–Henry L. Mencken

2.1 Introduction

We now study two situations where moment closure approximations sometimes do not work very well. Qualitative and quantitative inaccuracies can occur near critical points (e.g. low endemicity) and in the case of transient dynamics (particularly invasion). *Critical points* are points in parameter space where there is a qualitative change in the behaviour of the stable equilibrium solutions, i.e. a stability-exchanging bifurcation. In solid state physics they are associated with phase transitions, and in epidemiology with the threshold phenomenon discussed in chapter 1. The two problems are related to one another because they both deal with situations where the infected individuals occupy a relatively small part of the lattice and, more importantly, they both are caused by phenomena such as clustering which involve high-order correlations (by *clustering* we mean, in this chapter and others, not only a clustering of the underlying network structure but also clustering of infecteds with one another, superimposed on the network structure). There are two approaches to overcoming the inaccuracy of MCA in such cases; in part I of this chapter we attempt to improve the MCA by going from pairs to higher-order clusters, and in part II we consider alternative pair approximations, one of which is novel. We will study the contact process, where each node of the regular lattice has Q neighbours, where the rate of infection transmission in infected-susceptible edges is λ and where the recovery rate is ν and the population size is N.

To help us visualize invasion and the equilibrium situation near to and far away from the critical point, we present stochastic data from the contact process. Figure 2.1 shows snapshots of the time evolution of the contact process on a square lattice after a small number of infecteds have been introduced, and figure 2.2 shows populations at equilibrium for low and high endemicity for the contact process. Clustering is obvious in both cases but seems to be reduced in the last frame of figure 2.1 on account of the large density of infecteds.

The inaccuracy of moment closure approximations near critical points has been noted by researchers from their first applications in solid-state physics to recent applications in ecology [70]. Near critical points, long-range correlations develop which are poorly modelled by most low-order approximations [89]. The effect of such long-range correlations can be seen in figure 2.2. Researchers found that it is possible to increase accuracy by using higher-order approximations until the size of the basic element of the MCA reaches the typical length scale of the relevant correlations [15][75][81]. We will use the term *cluster approximations* to refer generally to higher-order moment closure approximations as opposed to pair approximations, but the

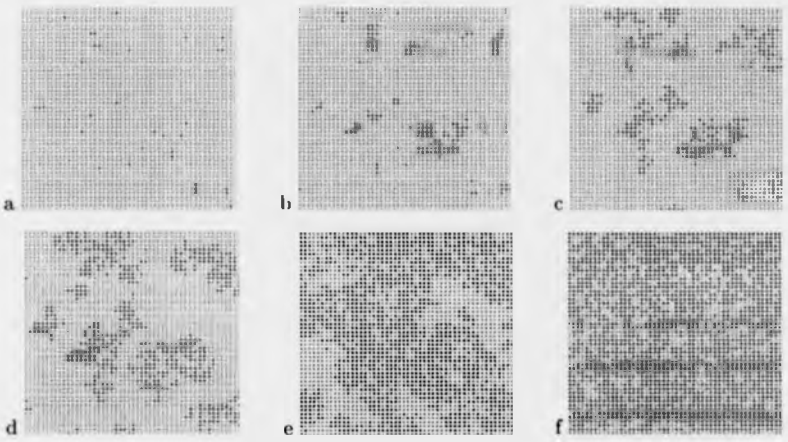


Figure 2.1: Time evolution of contact process on a square grid. Parameters: $\lambda = 0.1$, $\nu = 0.1$, $N = 2500$. Number of iterations = 10, 200, 1000, 2000, 8500, 40000 respectively for snapshots a,b,c,d,e,f. Dark gray sites are infected individuals and light gray sites are susceptible individuals.

term has been used previously to denote MCA generally.

Triple approximations for instance may well capture features of spatial spread with sharp boundaries [12] [13] such as wavefronts. However in ecology and epidemiology (unlike in solid-state physics) there has been little effort to extend moment closures beyond the pair level. In the case of epidemiology this has often been motivated by the observation that pairs are the fundamental unit of interaction in epidemiological systems, and hence it is inferred that low-order correlations dominate dynamics. Although this is true, there are many examples in epidemiological models where pair approximations are highly inaccurate. The accuracy of the PA seems to have more to do with a high average number of neighbours Q per individual than with the biological systems under consideration [13]. However, for the CP at least it will turn out that cluster approximations do not provide much improvement over OPA in predicting the final size of the epidemic, whereas alternative pair approximations fare much better.

Pair approximations were basically developed for equilibrium analysis, and therefore the behaviour of PA near critical points has been studied. However the inaccurate behaviour of PA transients has received less attention. Figure 2.3 illustrates the phenomenon we are talking about. The left-hand side of figure 2.3 shows the predicted time series of the contact process according to a pair approximation, and the right-hand side shows the equivalent simulation time series for corresponding

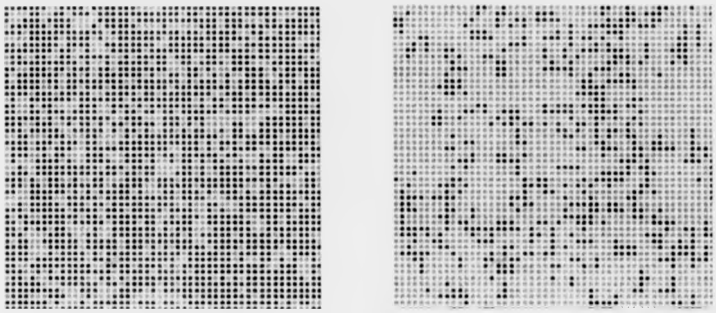


Figure 2.2: The contact process in two dimensions at equilibrium. $\lambda = 0.1$, $N = 2500$, $\nu = 0.1$ for left-hand lattice and $\nu = 0.2$ for right-hand lattice. Dark gray sites are infected individuals and light gray sites are susceptible individuals.

initial conditions (we go into detail about the derivation of the pair approximation later in this chapter). Although the time scale for the simulation has not been adjusted to fit the time scale of the PA time series, it is clear that there is qualitative disagreement. The moment closure approximation predicts logistic growth, i.e. it is initially exponential in time, but the simulation exhibits linear growth for most of the transient phase until the epidemic starts to approach steady-state. The linearity of the stochastic time series stems from clustering in a spatially-distributed population where the disease can only spread through neighbours. Figure 2.1 shows the process of invasion whereby fairly well-defined patches expand and coalesce. As we can see, infected individuals within patches will often 'waste' effective contacts on other infecteds. The problem with moment closure approximations with respect to this inaccuracy is that they are not explicit spatial models. They, too, are mean-field models in a sense; while they do not assume that each individual sees the same environment, they do assume that each type of pair or cluster sees the same environment. Therefore, as with a mean-field model, each pair or cluster can be in contact with any other pair or cluster, and so the transient appears to grow exponentially for early times. The assumption of conditional independence under OPA ignores clustering, and so the only way clustering can come into the picture to produce more realistic time series is through choosing a PA which incorporates clustering or by going to sufficiently high-order clusters. Filipe et al. devote some attention to this interesting question of transients and pair approximations and employ more rigorous analytical arguments, although their explanation and the spatial epidemic models they study differ from ours [39][40][41].

This inaccuracy however does not preclude all use of moment closure approxi-

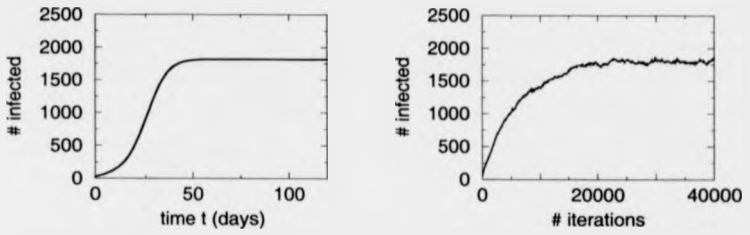


Figure 2.3: A comparison of stochastic model and cluster approximation time series for number of infecteds $[I]$. The right-hand graph is the stochastic data, the left-hand graph is the data from the cluster approximation. (For these parameters the pair approximation gives almost an identical curve to the cluster approximation). Parameters: $\lambda = 0.1$, $\nu = 0.1$, $N = 2500$. Initial Conditions: $\phi_0^I = 0.01$, $\phi_1^S = 0.04$, $\phi_0^S = 0.95$, corresponding to inoculation of 25 randomly chosen individuals in the population.

mations for transients and the study of invasion, and there are examples of good agreement between stochastic time series and pair approximations for certain types of models. For instance Filipe and Gibson [41] consider the contact process on a square lattice with the possibility of infection from nearest neighbours as well as some background probability from long-distance sources (this is the same model analyzed by Filipe in [38]). In the presence of background infection risk, the time series takes on a linear shape in both stochastic data and the pair approximation. However as the contribution from the background force of infection falls the pair approximation time series takes on a logistic shape, unlike the stochastic model which retains its linear shape. Another example is the work by Ellner et al. [31] which combines the equilibrium predictions of the OPA with random walk analysis to obtain good estimates of the speed of infection wavefronts in lattice populations models. However their derivation rests on assumptions about the shape of the wave front, and it is more desirable to obtain expressions for wave speed directly from the lattice transition rules, if possible.

There are other examples where pair approximations have constituted useful improvements on the time series from mean-field models. For instance Keeling [58], Morris [75] and others have studied time series of measles dynamics and compared them to real-world data for England and Wales. They find that the pair approximation can capture the annual and biannual cycles of measles dynamics more realistically than mean-field SEIR models, particularly by preventing the number of infecteds in inter-epidemic troughs from falling to unrealistically low levels. Finally, we will see in this chapter and others that moment closure approximations can in

limited ways capture qualitative features of invasion such as the time evolution of local network structure; this is not possible with conventional mean-field models and it makes valuable insights possible and can give us improved calculations for quantities such as R_0.

In the next section, part I, we derive two moment closure approximations to the contact process: the ordinary pair approximation, and a higher-order *cluster approximation*, and we analyze their behaviour near critical points and during invasion.

2.2 Part I–Deriving the Cluster Approximation and the OPA for the Contact Process

This section presents a full derivation of two moment closure approximations for the contact process. The basic procedure was outlined in chapter 1. We start with the derivation of the pair approximation.

2.2.1 The Ordinary Pair Approximation

Our state variables are $[SI]$, $[II]$ and $[SS]$, the number of susceptible-infected, infected-infected and susceptible-susceptible edges respectively. We will derive the equation of motion for $[SI]$ as an example, using the notation introduced in chapter 1.

The first step is to list all possible network events affecting $[SI]$:

$SI \rightarrow SS$	recovery
$SI \rightarrow II$	infection
$SS \rightarrow SI$	infection
$II \rightarrow SI$	recovery

Infection events will generally involve triples since individuals can be infected by more than one neighbor, while recovery only involves a single individual. Next we want to sum over all nodes in the population where such events can occur, using equation 1.12 and our knowledge of the effect of each event on the $[SI]$ pair numbers. Thus the equation of motion for $[SI]$ before simplification is:

$$\begin{aligned}\frac{d}{dt}[SI] &= \sum_{\varsigma_x=S} \nu(-Q_x(I)) + \sum_{\varsigma_x=I} \nu(+Q_x(I)) \\ &\quad + \sum_{\varsigma_{xy}=SI} \lambda Q_{xy}(I)(-1) + \sum_{\varsigma_{xy}=SS} \lambda Q_{xy}(I)(+1)\end{aligned}$$

Next we substitute for the $Q_x(I)$ and $Q_{xy}(I)$ quantities their population-averaged means plus the stochastic fluctuations of those quantities from the means at the individual nodes x and node pairs xy. For instance we substitute $Q_{xy}(I) = Q(I \mid SI) + \eta_{xy}(I \mid SI)$ where $Q(I \mid SI)$ is the mean and $\eta_{xy}(I \mid SI)$ is the stochastic fluctuation from the mean:

$$\begin{aligned}\frac{d}{dt}[SI] &= -\sum_{\varsigma_x=S} \nu(Q(I \mid S) + \eta_x(I \mid S)) + \sum_{\varsigma_x=I} \nu(Q(I \mid I) + \eta_x(I \mid I)) \\ &\quad - \sum_{\varsigma_{xy}=SI} \lambda(Q(I \mid SI) + \eta_{xy}(I \mid SI)) \\ &\quad + \sum_{\varsigma_{xy}=SS} \lambda(Q(I \mid SS) + \eta_{xy}(I \mid SS))\end{aligned}$$

Next we can take the constants such as $Q(I \mid SI)$ and the parameters out of the sums. Also we note that terms such as $\sum_{\varsigma_x=S} \eta_x(I \mid S)$ which represent fluctuations are zero by definition. Evaluating the sums produces:

$$\frac{d}{dt}[SI] = -\nu[SI] + \nu[II] - \lambda[SI]Q(I \mid SI) + \lambda[SS]Q(I \mid SS) \tag{2.1}$$

Now we apply the binomial OPA since we are on a regular network (see equation (1.16)):

$$\begin{aligned}Q(I \mid SI) &= 1 + \frac{Q-1}{Q}Q(I \mid S) \\ Q(I \mid SS) &= \frac{Q-1}{Q}Q(I \mid S)\end{aligned} \tag{2.2}$$

Inserting these into equation (2.1) produces an equation of motion for $[SI]$ in terms of pair numbers only. Similarly for $[II]$ and $[SS]$ we get:

$$\frac{d}{dt}[II] = -2\nu[II] + 2\lambda[SI]Q(I \mid SI) \tag{2.3}$$

$$\frac{d}{dt}[SS] = 2\nu[SI] - 2\lambda[SS]Q(I \mid SS) \tag{2.4}$$

the factor of two comes from the counting convention for SS and II pairs. Finally, applying the constraint

$$QN = [II] + 2[SI] + [SS]$$

which is true for all regular networks reduces the dimensionality to two, and we are ready to do some analysis with our closed equations of motion.

2.2.2 A Cluster Approximation

The second moment closure approximation we use in this chapter is a cluster approximation. Our clusters consist of an individual and his Q neighbours. The individual

in the centre of the cluster is called the *central individual* while the neighbours are sometimes called *peripheral individuals*. Let c_n^ξ denote a cluster where the central individual is in state $\xi \in \{I, S\}$ and where n of his neighbours are infected, $0 \leq n \leq Q$. Similarly we define ϕ_n^ξ as the density of type c_n^ξ clusters in the lattice. Therefore $\sum_{k=0..Q}(\phi_k^S + \phi_k^I) = 1$. Also $[I] = N\sum_{k=0..Q}\phi_k^I$ where N is the population size, although the equations of motion will be scaled out so that we consider only cluster densities. With this formulation, we know how many infected neighbours the central individual has but not which ones are which.

The hope is that this model has advantages because the number of infected neighbours of a central individual is known exactly, and so there is no ambiguity in the *force of infection* experienced by a susceptible from his infected neighbours (the term *force of infection* was originally used in mean-field models to denote the rate at which susceptibles become infected; the meaning is not so different here). Hence the calculation of the final size of the epidemic should be more accurate. Also, larger clusters will be able to capture more of the higher-order correlations. Finally, the local structure evolution of the cluster approximation will be different from that of the pair approximation and we will be able to see how the neighbourhoods of individuals evolve over the course of the epidemic.

These clusters can be created or destroyed either by events which affect the central individual or by events which affect the neighbours. For events relating to recovery or to infection of the central individual, no approximation is necessary. However in order to close the equations of motion we will have to introduce an approximation when it comes to considering infection of peripheral individuals. A list of all possible events affecting cluster numbers is:

$c_n^I \rightarrow$	c_n^S	recovery
$c_n^I \rightarrow$	c_{n-1}^I	recovery
$c_n^I \rightarrow$	c_{n+1}^I	infection
$c_n^S \rightarrow$	c_n^I	infection
$c_n^S \rightarrow$	c_{n-1}^S	recovery
$c_n^S \rightarrow$	c_{n+1}^S	infection

We derive the equations of motion again using equation (1.12). Let $\mathcal{N}_x(\xi)$ denote the set of state ξ neighbours of the individual at node x, where $\xi \in S, I$. if some node y is in the set $\mathcal{N}_x(\xi)$ of x, we write $y \in \mathcal{N}_x(\xi)$. Consider first the equation of

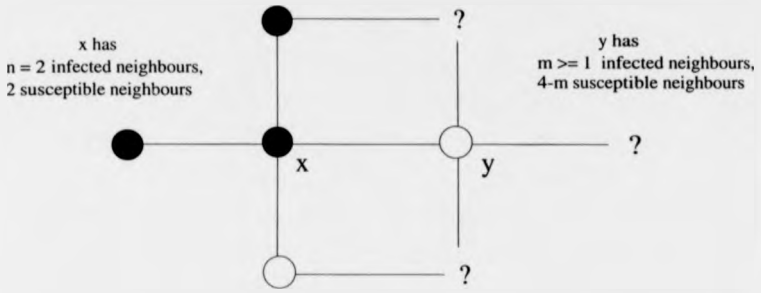


Figure 2.4: Illustration for derivation of $\frac{d}{dt}\phi_n^I$. Black circles represent infected individuals and white circles represent susceptible individuals.

motion for ϕ_n^I, $0 < n < Q$:

$$\begin{aligned} N\frac{d}{dt}\phi_n^I &= \sum_{\varsigma_x = c_{n-1}^I} \sum_{y \in \mathcal{N}_x(S)} \lambda Q_y(I)(+1) + \sum_{\varsigma_x = c_n^I} \sum_{y \in \mathcal{N}_x(S)} \lambda Q_y(I)(-1) \\ &+ \sum_{\varsigma_x = c_n^S} \lambda Q_x(I)(+1) + \sum_{\varsigma_x = c_n^I} \sum_{y \in \mathcal{N}_x(I)} \nu(-1) \\ &+ \sum_{\varsigma_x = c_n^I} \nu(-1) + \sum_{\varsigma_x = c_{n+1}^I} \sum_{y \in \mathcal{N}_x(I)} \nu(+1) \end{aligned} \tag{2.5}$$

We know that in the third term $Q_x(I) = n$. The other $Q_y(I)$ terms, where y is a peripheral individual of some known cluster type, will require approximation. We substitute for the $Q_y(I)$ terms as follows:

$$\begin{aligned} Q_y(I) &= Q(I \mid SI_n) + \eta_y(I \mid SI_n) \qquad y \in \mathcal{N}_x(S), x \in c_n^I \\ Q_y(I) &= Q(I \mid SI_{n-1}) + \eta_y(I \mid SI_{n-1}) \qquad y \in \mathcal{N}_x(S), x \in c_{n-1}^I \end{aligned} \tag{2.6}$$

where $Q(I \mid SI_n)$ is the number of infected neighbours of a susceptible with at least one infected neighbour who is the central individual of a c_n^I cluster, averaged over that entire subpopulation. $Q(I \mid SI_{n-1})$ is similarly defined. Figure 2.4 depicts a typical situation. Using the substitutions, taking the constant parameters and averages from equation (2.5) out of the sums and noting that terms for fluctuations sum to zero produces:

$$\begin{aligned} N\frac{d}{dt}\phi_n^I &= \lambda Q(I \mid SI_{n-1}) \sum_{\varsigma_x \in c_{n-1}^I} \sum_{\varsigma_y \in \mathcal{N}_S(x)} -\lambda Q(I \mid SI_n) \sum_{\varsigma_x \in c_n^I} \sum_{\varsigma_y \in \mathcal{N}_S(x)} \\ &+ n\lambda \sum_{\varsigma_x \in c_n^S} \quad -\nu \sum_{\varsigma_x \in c_n^I} \sum_{\varsigma_y \in \mathcal{N}_I(x)} \\ &- \nu \sum_{\varsigma_x \in c_n^I} \quad +\nu \sum_{\varsigma_x \in c_{n+1}^I} \sum_{\varsigma_y \in \mathcal{N}_I(x)} \end{aligned}$$

Finally evaluating all sums and dividing through by N results in:

$$\begin{aligned}\frac{d}{dt}\phi_n^I &= \lambda(n-1)Q(I \mid SI_{n-1})\phi_{n-1}^I - \lambda nQ(I \mid SI_n)\phi_n^I \\ &\quad +n\lambda\phi_n^S - n\nu\phi_n^I - \nu\phi_n^I + (n+1)\nu\phi_{n+1}^I \end{aligned} \tag{2.7}$$

Next we must approximate $Q(I \mid SI_n)$ and $Q(I \mid SI_{n-1})$. We start by ignoring the information about the peripheral individuals, an assumption which follows from conditional independence:

$$Q(I \mid SI_n) = Q(I \mid SI_{n-1}) = Q(I \mid SI) \tag{2.8}$$

So the only information we are keeping so far is the fact that the susceptible has at least one infected neighbour. Because we are working on a regular lattice, we can use the binomial OPA closure from chapter 1:

$$Q(I \mid SI) = 1 + \frac{Q-1}{Q}Q(I \mid S) \tag{2.9}$$

Substituting this into equation (2.7) produces a coupled ordinary differential equation of motion for ϕ_n^I for $0 < n < Q$. It seems wasteful to throw away so much information by using the approximations (2.8) and (2.9). However in our research we tried a number of simple ad-hoc closures, and some other not-so-simple closures as well, in order to retain more cluster information, and we found that none of them worked better than this simple closure, and many were even worse.

We can follow a similar procedure for ϕ_0^I to get:

$$\frac{d}{dt}\phi_0^I = -\nu\phi_0^I + \nu\phi_1^I - Q\lambda Q(I \mid SI)\phi_0^I \tag{2.10}$$

and when $n = Q$ we have:

$$\frac{d}{dt}\phi_Q^I = -Q\nu\phi_Q^I - \nu\phi_Q^I + Q\lambda\phi_Q^S + \lambda Q(I \mid SI)\phi_{Q-1}^I \tag{2.11}$$

A similar derivation produces the following for the equations of motion for susceptible clusters $\frac{d}{dt}\phi_n^S$. In these equations the higher-order correlation $Q(I \mid SS)$ appears instead of $Q(I \mid SI)$. According to binomial OPA, $Q(I \mid SS) = \frac{Q-1}{Q}Q(I \mid S)$. When $0 < n < Q$:

$$\begin{aligned}\frac{d}{dt}\phi_n^S &= -n\nu\phi_n^S + (n+1)\nu\phi_{n+1}^S + \nu\phi_n^I - n\lambda\phi_n^S \\ &\quad -(Q-n)\lambda Q(I \mid SS)\phi_n^S + (Q+1-n)\lambda Q(I \mid SS)\phi_{n-1}^S \end{aligned} \tag{2.12}$$

Similarly when $n = 0$ we have:

$$\frac{d}{dt}\phi_0^S = \nu\phi_1^S + \nu\phi_0^I - Q\lambda Q(I \mid SS)\phi_0^S \tag{2.13}$$

and when $n = Q$ we have:

$$\frac{d}{dt}\phi_Q^S = -Q\nu\phi_Q^S + \nu\phi_Q^I - Q\lambda\phi_Q^S + \lambda Q(I \mid SS)\phi_{Q-1}^S \tag{2.14}$$

After applying the total population density constraint $\sum_{k=0..Q} \phi_k^I + \phi_k^S = 1$ we are left with a nine-dimensional system of equations, difficult to analyze with pencil and paper but amenable to numerical analysis.

Although this cluster approximation is somewhat more complicated than the pair approximation, the formulation and motivation are clear, which is most important. In the next section we look at what we gain for this loss of simplicity.

2.2.3 Comparison of the Contact Process and the Pair and Cluster Approximations

Comparison of Equilibria

We use the notation $[I]_\infty$ throughout this thesis to denote the final size of the epidemic, either derived from numerical methods, direct calculation or measurement from simulations. To compare $[I]_\infty$ in the stochastic model and its deterministic approximations we looked at a case where $Q = 4$ and we implemented the contact process as a stochastic Monte Carlo model. We measured $[I]_\infty$ as a function of ν. Results are shown in figure 2.5. The cluster approximation is only marginally more accurate than the pair approximation, and both are particularly poor closer to the critical point. So it seems, for this range of parameters at least, that the cluster model does not have much better accuracy than the pair approximation. A triple approximation which distinguishes between collinear triples (i.e. where the three nodes are collinear on the lattice) and triples with a ninety degree angle (i.e. the three nodes form a right triangle) might do better than the cluster approximation since the cluster approximation does not distinguish between the various cluster configurations possible for each value n of infected individuals per cluster, and so there is some averaging over configurations. Morris finds that the triple approximation has significantly improved accuracy for a similar ecological IPS [75], the spatialized hawk-dove model.

Comparison of Time Evolution of Local Network Structure

Here we look at the time evolution of the stochastic model and its deterministic approximations, focusing on invasion and the evolution of local network structure. It is in the time series of the state variables that differences between PA and cluster approximations will most likely emerge.

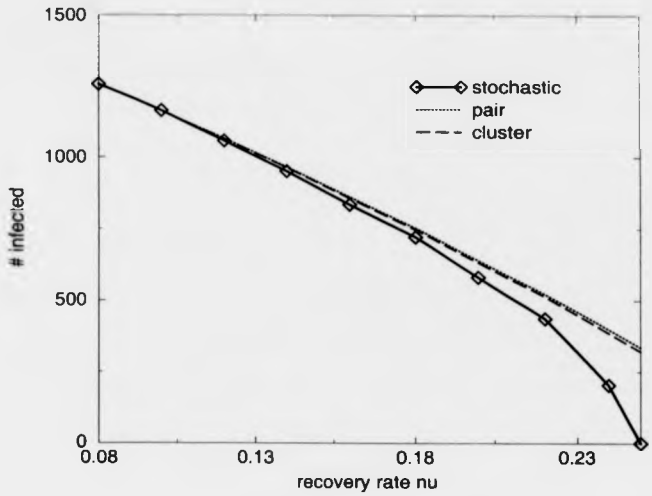


Figure 2.5: A comparison of the equilibrium predictions of $[I]$ for the stochastic model, pair approximation and the cluster approximation. Parameters: $\lambda = 0.1$, $N = 1600$. Each diamond is a long-time average of one stochastic run at a given value of ν

Figures 2.6 through 2.9 show time series of the cluster densities ϕ_n^ξ for the stochastic model and the pair and cluster approximation, along with the corresponding time series for $[I]$. In the case of the pair approximation, the cluster densities are reconstructed using the assumption of a binomial distribution for $Q_x(I)$.

The stochastic time series of figure 2.6 exhibit three basic types:

1. Exponential decay – e.g. ϕ_0^S
2. Growth followed by decay – e.g. ϕ_0^I, ϕ_1^I,ϕ_1^S, ϕ_2^S.
3. Monotonic growth – e.g. ϕ_2^I, ϕ_3^I, ϕ_4^I, ϕ_3^S, ϕ_4^S.

The type of time series of a particular cluster density will, of course, vary according to the parameters chosen. This particular set of time series captures all three types of behaviour. Here we have chosen $\nu = 0.1$, producing a large final size and rapid growth.

Notably, for the ϕ_0^I time series there is a rapid initial decline to $\phi_0^I \approx 0.002$ before the cluster density recovers and starts its upward climb. This local minimum is present in the data of figure 2.6 but is not visible because the time scale of this process is small compared to the time scale of the process of reaching equilibrium. However, the data for the ϕ_0^I time series for this short initial phase are shown in

figure 2.7. This minimum is only seen in the time series data for ϕ_0^I and not the other cluster densities, at least for this choice of parameters.

We interpret this local minimum as the establishment of a characteristic local network structure after the initial inoculation, followed by a spreading through the rest of the network. The characteristic local structure is in quasi-equilibrium and changes only slowly over the course of the epidemic, relative to the time scale of the first phase. Such behaviour during invasion has been observed by others in IPS models [90] and we can describe such a stochastic invasion as consisting of three parts [9]:

1. *Inoculation* phase. The first stage is largely stochastic and relatively quick: a number of infective individuals (with $[I] \ll N$) are randomly placed in the population and start infecting their neighbours.

2. *Establishment* phase. If the disease does not die out, then there is a second stage which is characterized by the fact that at its beginning we still have $[I] \ll N$ but now the individual infectives have grown into small local populations with a well-defined local correlation structure that can be calculated.

3. *Development* phase. Here, the disease may either die out or grow until it reaches some equilibrium. If the disease grows, the characteristic local correlation structure undergoes changes as the various patches of infection start to come into contact with one another.

Regarding the other time series, it is apparent that the number of clusters with low n increases rapidly early in the epidemic, and then decreases as more individuals in the population become infected. However the cluster densities with large numbers of infected neighbours increase monotonically in time. ϕ_0^S decreases monotonically, which is unsurprising. Thus the relative densities of clusters change somewhat as the disease invades and the network becomes more crowded with infecteds.

We now turn to the moment closure approximations to compare their predictions of the cluster density time series. Figure 2.8 shows the time series for the pair approximation constructed from the binomial assumption. The qualitative features of time series of the pair approximation match the data of the stochastic model, however there are some problems. The decay of ϕ_0^S is monotone but not exponential. Also the curves tend to have a logistic shape, but the stochastic data indicate linear growth in the early phase of the epidemic. Yet we do find that as n increases, the time series adhere more to type 3, as we would like them to. Also, the time series for ϕ_0^I shows the local minimum for early times which we interpret as the establishment

of the pseudo-equilibrium, given those initial conditions. As the disease spreads through the network, ϕ_0^I increases and then declines as the density of infecteds increases dramatically. This behaviour is also what is expected for growth from a small initial number of infected individuals to a large final size..

Given the similar accuracies of the cluster and pair approximation in predicting $[I]_\infty$, it is not surprising that the cluster approximation produces very similar results for the time evolution of the local epidemic structure, including the local minimum in the time series for ϕ_0^I (see figure 2.6).

We also compared the time evolution of the measures $Q(I \mid S)$, $Q(I \mid SI)$ and $Q(I \mid SS)$ in a similar way. These data are shown in figures 2.10 through 2.12. Figure 2.10 shows the stochastic data from a single simulation run. The data for $\nu = 0.08$ show a linear increase in $Q(I \mid S)$, but the growth is concave in time t for $Q(I \mid SI)$ and convex in time for $Q(I \mid SS)$. It is less clear what is happening in the cases $\nu = 0.20$ and $\nu = 0.24$.

Figures 2.12 and 2.11 show the data from the pair and cluster approximations. These data are uninteresting except for the fact that they both predict logistic growth curves for $Q(I \mid S)$, $Q(I \mid SI)$ and $Q(I \mid SS)$ for all values of ν, which is clearly at odds with the stochastic data.

So, in summary, the cluster approximation for these parameter values does not offer a significantly better prediction for the time series of the cluster densities than the pair approximation does. Going to higher-order does not always increase accuracy. This is surprising because higher-order correlations become more important near the critical point, so we would expect a higher-order moment closure to better capture the dynamics near the critical point. Yet both deterministic approximations are deficient; for higher values of n, they predict exponential growth in early times instead of linear growth. Also, their predicted decay for ϕ_0^S is not exponential as it is in the simulation data. However they do give meaningful time series for ϕ_0^I and are a significant improvement on the mean-field approximation which cannot capture spatial structure at all.

We now attempt to improve our moment closure approximations by looking more closely at the data from the stochastic model to gain insight into the epidemic processes.

2.2.4 Scrutinizing the Stochastic Model for More Clues

To derive better moment closures we analyzed data from the stochastic model. Data on measurements of $Q(I \mid S)$, $Q(I \mid SI)$ and $Q(I \mid SS)$ in the stochastic model are

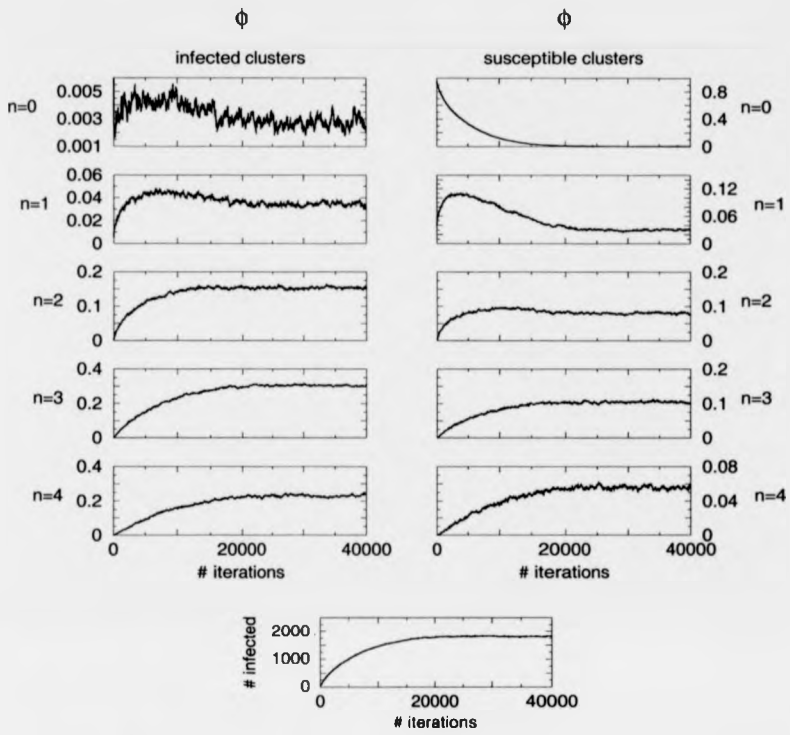


Figure 2.6: Time evolution of ϕ_n^I and ϕ_n^S in the stochastic model. Parameters: $N = 2500$, $\lambda = 0.1$, $\nu = 0.1$. Initial Conditions: 25 randomly chosen individuals are inoculated. Each datum at time t is the average of values from eight different simulation runs at time t for the same parameters and initial conditions.

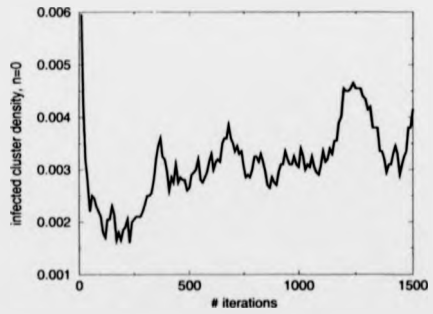


Figure 2.7: Early time evolution (up to 1500 iterations) of ϕ_0^I in the stochastic model. Parameters and initial conditions are as in figure 2.6

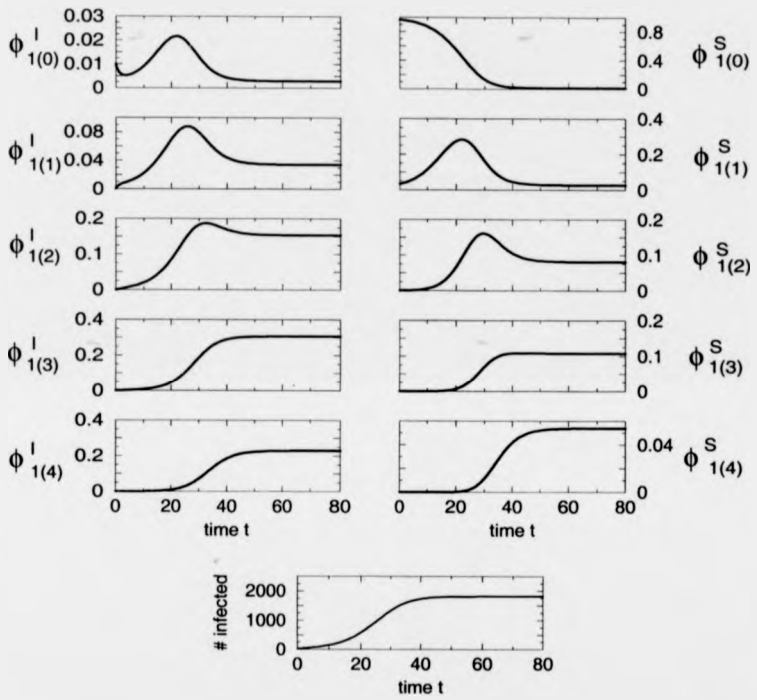


Figure 2.8: Time evolution of ϕ_n^I and ϕ_n^S in the pair approximation. Parameters: $N = 2500$, $\lambda = 0.1$, $\nu = 0.1$. Initial conditions: $\phi_0^I = 0.01$, $\phi_1^S = 0.04$, $\phi_0^S = 0.95$ (corresponds to inoculation of 25 randomly selected individuals). From top left to bottom left are time series for ϕ_0^I to ϕ_4^I, and from top right to bottom right are time series for ϕ_0^S to ϕ_4^S. Also shown is a time series for the number of infecteds, at bottom.

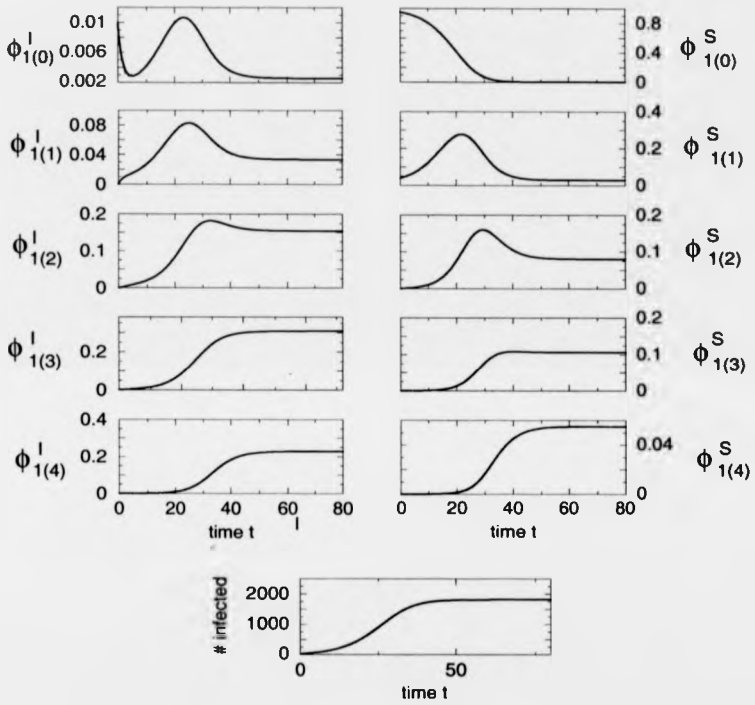


Figure 2.9: Time evolution of ϕ_n^I and ϕ_n^S in the cluster approximation. Parameters: $\lambda = 0.1$, $\nu = 0.1$, $N = 2500$. Initial conditions: $\phi_0^I = 0.01$, $\phi_1^S = 0.04$, $\phi_0^S = 0.95$ (corresponds to inoculation of 25 randomly selected individuals). From top left to bottom left are time series for ϕ_0^I to ϕ_4^I, and from top right to bottom right are time series for ϕ_0^S to ϕ_4^S. Also shown is a time series for the number of infecteds, at bottom.

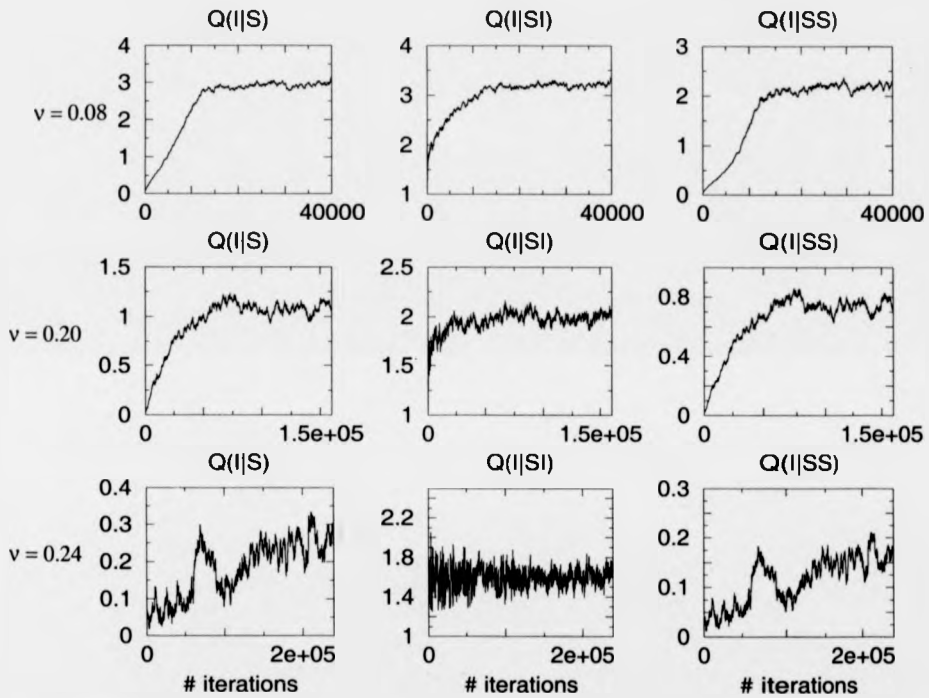


Figure 2.10: Time evolution of $Q(I \mid S)$, $Q(I \mid SI)$, and $Q(I \mid SS)$ in the stochastic model. Parameters: $N = 2500$, $\lambda = 0.1$, $\nu = 0.08$ (resp. 0.20, 0.24) for the top (resp. middle, bottom) row. Initial conditions: 25 randomly chosen individuals are inoculated.

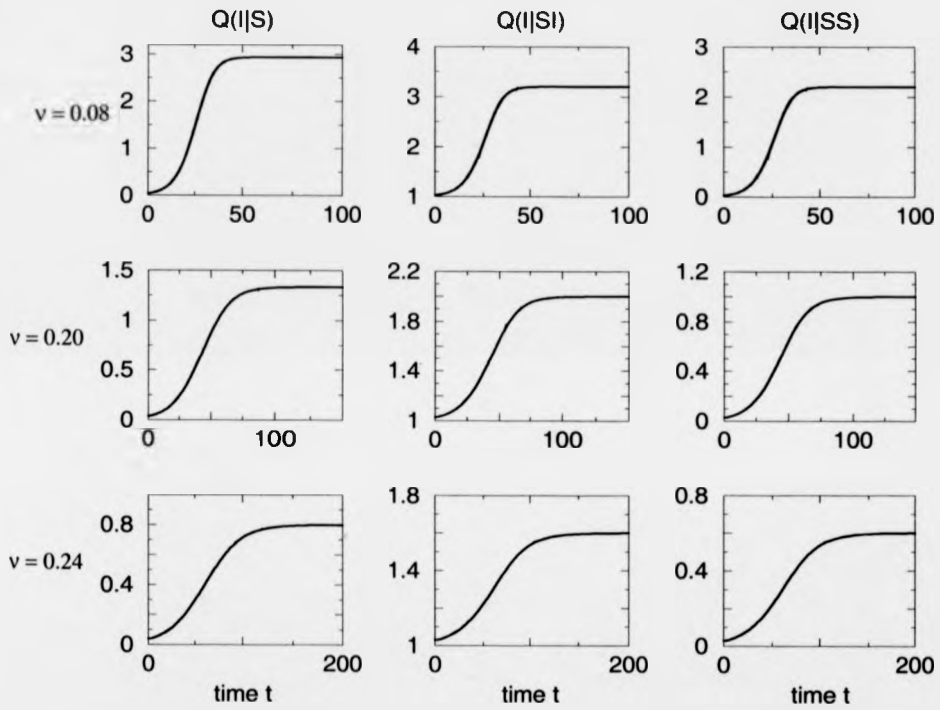


Figure 2.11: Time evolution of $Q(I \mid S)$, $Q(I \mid SI)$ and $Q(I \mid SS)$ in the pair approximation. Parameters: $N = 2500$, $\lambda = 0.1$, $\nu = 0.08$ (resp. 0.20, 0.24) for the top (resp. middle, bottom) row. Initial conditions: $[SI] = 100$, $[II] = 0$

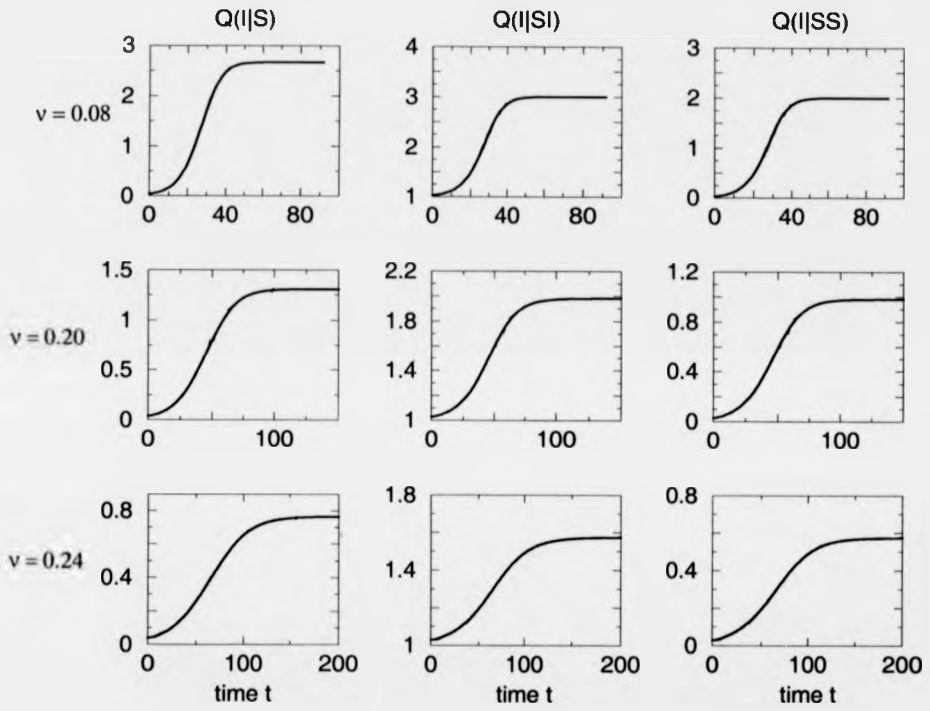


Figure 2.12: Time evolution of $Q(I \mid S)$, $Q(I \mid SI)$ and $Q(I \mid SS)$ in the cluster approximation. Parameters: $N = 2500$, $\lambda = 0.1$, $\nu = 0.08$ (resp. 0.20, 0.24) for the top (resp. middle, bottom) row. Initial conditions: $\phi_0^I = 0.01$, $\phi_1^S = 0.04$, $\phi_0^S = 0.95$

1	2	3	4	5	6		
t	$[I]$	$Q(I \mid SI)$	$Q(I \mid SS)$	$1 + 0.75\times$ $Q(I \mid S)$	$0.75\times$ $Q(I \mid S)$	% error col. 5 wrt 3	% error col. 6 wrt 4
510	234.3	1.823	0.106	1.139	0.139	31.3 %	37.5 %
1870	601.7	2.178	0.233	1.339	0.339	45.5 %	38.5 %
3030	863.6	2.285	0.323	1.482	0.482	49.2 %	35.1 %
5520	1291.6	2.626	0.554	1.862	0.862	55.6 %	29.1 %
8130	1571.1	2.845	0.882	2.304	1.304	47.8 %	19.0 %
10240	1795.6	2.968	1.462	2.766	1.766	20.8 %	6.8 %
21560	1973.2	3.175	2.162	3.172	2.172	0.46 %	0.09 %

Table 2.1: A comparison of local structure evolution for the stochastic model and the binomial assumption. Snapshots of time series at various values of time t for parameters $\lambda = 0.08$, $\nu = 0.1$, $N = 2500$. Initial conditions: 25 randomly chosen individuals are inoculated. Times are chosen to match with values of $[I]$ found in table 2.2. Each datum at time t is the average of nine samples from points at times $t + 10k$, $k = -4 \ldots 4$.

presented tables 2.1 and 2.2. In table 2.1 we look at equilibrium values for various final sizes as controlled by ν, and in table 2.2 we look at how these measures evolve over the course of an invasion. In both tables we compare the stochastic values for the third-order correlations $Q(I \mid SI)$ and $Q(I \mid SS)$ to the predicted values for those measures assuming an OPA for $Q(I \mid SI)$ and $Q(I \mid SS)$ based on the stochastic data for $Q(I \mid S)$. Such a comparison allows us to determine how much of a deviation there is from the OPA and will give us some idea of what corrections we need to make. Also, note that the times t at which the snapshots are taken in table 2.1 are chosen so that the number of infecteds $[I]$ at those times correspond approximately to the equilibrium number of infecteds for the values of ν in table 2.2. This will help us to compare and contrast the situation of low endemicity to the situation of the early stages of invasion when $[I]$ is low. We can also consider figures 2.1 and 2.2 in our comparison.

The data on epidemic growth in table 2.1 show that the error of the binomial truncation peaks in the middle phase of the epidemic and goes to almost zero at equilibrium. Even close to the equilibrium at $t = 10240$ the error is still high. These results mean that the network structure of the epidemic does not converge significantly to the equilibrium structure until very close to the equilibrium. On the other hand for the corresponding data in table 2.2 we see that the error increases with increasing recovery rate ν. As expected, in both cases $Q(I \mid SI) > 1 + 0.75 \times Q(I \mid S)$ and $Q(I \mid SS) < 0.75 \times Q(I \mid S)$ always because of clustering. Comparing figure 2.1.b to the low endemicity case of figure 2.2 also describes the difference between

1	2	3	4	5	6		
ν	$[I]$	$Q(I \mid SI)$	$Q(I \mid SS)$	$1 + 0.75\times$ $Q(I \mid S)$	$0.75\times$ $Q(I \mid S)$	% error col. 5 wrt 3	% error col. 6 wrt 4
0.24	231.9	1.602	0.157	1.185	0.185	26.030 %	17.834 %
0.22	603.0	1.797	0.469	1.527	0.527	15.044 %	12.367 %
0.20	866.4	1.983	0.737	1.803	0.803	9.077 %	8.955 %
0.16	1293.0	2.382	1.232	2.297	1.297	3.568 %	5.278 %
0.13	1571.3	2.702	1.571	2.642	1.642	2.221 %	4.519 %
0.10	1795.9	2.975	1.921	2.956	1.956	0.639 %	1.822 %
0.08	1976.7	3.210	2.194	3.206	2.206	0.125 %	0.547 %

Table 2.2: A comparison of local structure evolution for the stochastic model and the OPA assumption; Equilibrium data for several values of ν for the parameters $\lambda = 0.1$, $N = 2500$.

equilibrium behaviour near a critical point and the behaviour in early stages of an invasion.

We can understand the results of table 2.1 by looking at figure 2.1, where we clearly see the qualitative differences between a growing epidemic and the equilibrium state. The growing epidemic is divided into well-distinguished patches. The fact that the error is high even close to the equilibrium in the time series data implies that the local network structure is relatively constant and only changes very close to the final size.

The results mostly confirm our intuition about invasion and the behaviour of the stochastic model near the critical point. In the next section we use these insights to improve pair approximations.

2.3 Part II–Better Pair Approximations

We have seen in the first part of this chapter that the cluster approximation does not improve significantly on the pair approximation, at least for the contact process in the parameter regime we considered. In this section we try the other approach to improving the MCA, i.e. seeking better pair approximations. Improving the pair approximations instead of going to higher-order has the advantage that we do not end up with unwieldy and complicated systems of equations. However finding good pair approximations which capture higher-order correlations can be tricky.

One recently developed approximation which works well for the contact process is the *hybrid pair approximation* (HPA) of Filipe [38][40]. The hybrid pair approximation works by creating a hybrid of two other pair approximations, one which overpredicts higher-order correlations and one which underpredicts them. When

these are balanced we have a hybrid pair approximation where the higher-order correlations are well-approximated. The HPA is discussed in more depth in derivation of the steady/casual model of chapter 3. We do not go into depth here about the HPA for the contact process because Filipe has already considered this application [38][40]. The HPA works much better than the OPA near critical points. For instance, under HPA the location of the critical value ν_c of ν for the contact process is given by:

$$\nu_c = \lambda \frac{Q^2 \sqrt{1 + 4(Q-1)/Q^2} - 1}{2} \tag{2.15}$$

For the parameters of figure 2.5, the critical point according to HPA is located at $\nu = 0.258$, which is significantly closer to the stochastic value of approximately $\nu = 0.24$ than the predictions of other pair approximations and the cluster approximation ($\nu = 0.30$). Since HPA has been studied for the contact process we do not investigate it further. Also, there are some limitations to HPA so that it cannot be used for certain types of models. However we do attempt to apply it to the dynamic network models of chapters 3 and 4.

In the next two subsections, we consider two alternative pair approximations designed to improve upon the ordinary pair approximation.

2.3.1 The Improved Pair Approximation

We have noted from the data of tables 2.1 and 2.2 that the OPA overestimates $Q(I \mid SI)$ and underestimates $Q(I \mid SS)$. This is because of the clustering of infecteds; if we know that a susceptible already has one infected neighbour, it is more likely to have other infected neighbours as well. It is difficult to find a mechanistic foundation for these results and to incorporate them into a pair approximation, but one can try half-measures.

For instance, based on the observation that the the deviation in table 2.2 generally increases as the ratio $[I]/N$ decreases, we could introduce a multiplicative term of the form $[I]/(Nk+[I])$ (the Michaelis-Menten functional form) into equation (2.2). However we found that neither this function nor any other simple functional forms we tried could fit the curve well. Another more effective correction is to introduce a parameter which contains the missing information about clustering. Such an analysis will at least help us get a better feel for how correlations affect the infection dynamics, and perhaps suggest a mechanistic treatment for clustering. The new parameter ϵ is introduced into the equation for $Q(I \mid SI)$:

$$Q(I \mid SI) = 1 + \epsilon \frac{Q-1}{Q} Q(I \mid S) \tag{2.16}$$

For $Q(I \mid SS)$ we use the following identity which applies to regular lattices:

$$\begin{aligned} Q &= Q(I \mid SI) + Q(S \mid SI) \\ &= Q(I \mid SI) + \frac{[SSI]}{[SI]} \\ &= Q(I \mid SI) + \frac{[SSI]}{[SS]}\frac{[SS]}{[SI]} \\ &= Q(I \mid SI) + Q(I \mid SS)\frac{[SS]}{[SI]} \\ \Rightarrow Q(I \mid SS) &= (Q - Q(I \mid SI))\frac{[SI]}{[SS]} \\ &= \left(Q - 1 - \epsilon\frac{Q-1}{Q}Q(I \mid S)\right)\frac{[SI]}{[SS]} \end{aligned} \tag{2.17}$$

This technique is the *improved pair approximation* (IPA) and was developed by Sato et al. [88] for a similar model with host mortality. We implement the IPA for the contact process described by equations (2.1) and (2.4). We use ϵ as a free parameter, although one could estimate ϵ from the simulation data as Sato et al. do. Also, we could incorporate the $(Q-1)/Q$ factor into the ϵ parameter, but we choose to leave it separate because ϵ is used as a measure of clustering.

The introduction of a new parameter is perhaps not as much a setback as one might think. ϵ very neatly encapsulates much information about clustering and we can carry out our analysis with this in mind. The drawback is that it incorporates other information besides clustering, and in fact all error in the moment closure approximation comes in through ϵ.

Figure 2.13 shows $[I]_\infty$ as a function of the recovery rate ν for various values of ϵ. Also, bifurcation diagrams for the stable nontrivial branches of $[SI]$ and $[II]$ are shown in figure 2.14. The results are interesting. The agreement with the simulation data is best for moderate to high endemicity for $\epsilon = 1$, as expected. This means conditional independence applies when the density of infecteds is moderate to high. When $\epsilon > 1$ the IPA underpredicts $[I]_\infty$ for moderate to high endemicity but again overpredicts $[I]_\infty$ for low endemicity. This is not surprising since the higher-order correlations are much stronger at low endemicity than high endemicity and so different amounts of correction to the OPA are necessary. Since the deviation of table 2.2 of the binomial OPA approximation for $Q(I \mid SI)$ from its actual value depends on $[I]/N$, we would not expect any one value of ϵ to give a good fit for all values of the recovery rate ν. However for a given value of ν it is clear that increasing clustering decreases $[I]_\infty$.

What is notable about figure 2.13 is that the stable solution branches converge at $\nu = 0.3$, $[I] = 0$ for all values of ϵ. Thus we can infer the existence of a critical

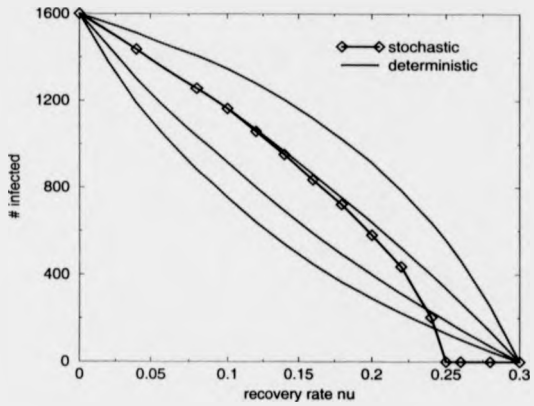


Figure 2.13: A comparison of $[I]_\infty$ as a function of ν for the stochastic model and the IPA. Values of ϵ are, from top curve to bottom, 0.5, 1, 2, 3. Parameters: $N = 1600$, $\lambda = 0.1$.

value ν_c at $\nu = 0.3$ which, interestingly, has no ϵ dependence. ϵ, which is supposed to determine clustering and hence influence the final size of the disease, might also be expected to have an impact on the value of the invasion threshold. We can also examine this by considering the nontrivial equilibrium solution $[I]_\infty$ derived from applying the IPA to equations 2.1 and 2.4:

$$[I]_\infty = NQ\frac{-\lambda + \lambda Q - \nu}{\lambda Q^2 - Q\nu + \nu\epsilon Q - \lambda Q - \nu\epsilon} \tag{2.18}$$

Solving for $[I]_\infty = 0$ in terms of ν we get:

$$\nu_c = \lambda(Q - 1) \tag{2.19}$$

which has no dependence on ϵ. On the other hand, the relative size of $[I]_\infty$ for various values of ϵ in the limit $\nu \to \nu_c$ does vary according ϵ. To see this consider two values ϵ' and ϵ'' of ϵ, and consider the limit:

$$\lim_{\nu\to\nu_c}\frac{[I]_\infty(\epsilon')}{[I]_\infty(\epsilon'')} = \lim_{\nu\to\nu_c}\frac{\lambda Q^2 - \nu Q + \nu\epsilon'' Q - \lambda Q - \nu\epsilon''}{\lambda Q^2 - \nu Q + \nu\epsilon' Q - \lambda Q - \nu\epsilon'} = \frac{\epsilon''}{\epsilon'}$$

Notwithstanding this result, the behaviour of the IPA near ν_c must be treated with caution because of the lack of agreement with the stochastic model in this regime. More fundamentally, this lack of agreement near ν_c has to do with the fact that in spatial models, what is happening locally at the level of individuals is important and determines whether or not the disease can invade. For invasion, the description of local structure evolution given in subsection 2.2.3 applies. Because of the significant changes in the local structure in the early phase of invasion, no single value of ϵ

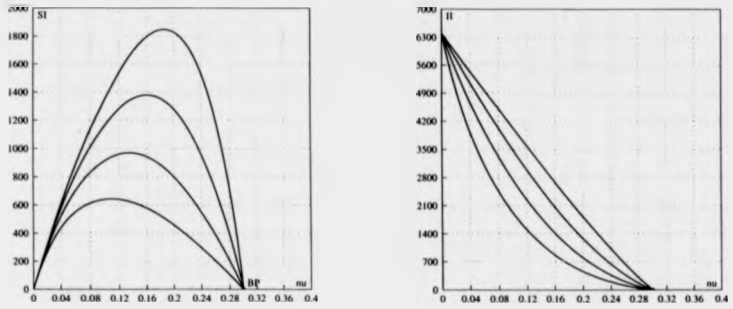


Figure 2.14: Bifurcation diagrams of $[SI]$ and $[II]$ versus ν. Parameters: $N = 1600$, $\lambda = 0.1$. Values of ϵ are, from top curve to bottom, 0.5,1,2,4.

will be able to capture what happens for different values of ν near ν_c. To capture the effect of clustering on the ability to invade, we need a pair approximation which models more explicitly the long-range correlations which develop near the critical point. Such an explicit, mechanistically-based PA will be able to model how long-range correlations are altered by different parameter choices and how they change throughout the course of the early phases of invasion. HPA is an example of a pair approximation which manages this, and predicts ν_c more accurately than IPA. In contrast to spatial models, this dichotomous relationship between the invasion process and the final size does not exist in mean-field models. For instance equation (1.10) for R_0 in the mean-field SIR model has the same form of the $[I]_\infty = 0$ threshold solution of the SIR model.

We illustrate this point further by considering the calculation of R_0 on a square lattice. For a spatial model, the number of secondary infections produced by an infected individual depends on the status of the neighbours of the infected individual. In the case of a square lattice, the rate at which an infected individual produces secondary infections in a susceptible population is $\lambda Q(S \mid I)_{inv}$, where $Q(S \mid I)_{inv}$ denotes the value of $Q(S \mid I)$ for the invading patches of infecteds. The expected length of time an individual remains infected is $1/\nu$. Therefore:

$$R_0 = \frac{\lambda Q(S \mid I)_{inv}}{\nu}$$

In an initially susceptible population, the first individual to be infected has $Q = 4$ susceptible neighbours. However, thereafter each infected has at most $Q - 1 = 3$ susceptible neighbours since they have each received the infection from one of their

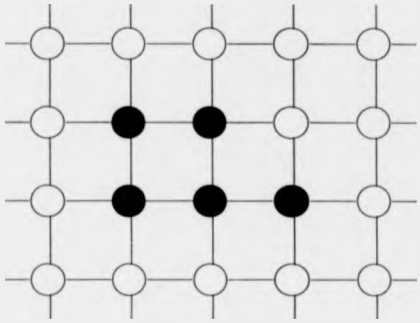


Figure 2.15: Illustration of how clustering reduces $Q(S \mid I)_{inv}$ in early stages of an epidemic on a square lattice. Black circles denote infected individuals and white circles are susceptibles.

neighbours. Thus, in the absence of clustering, $Q(S \mid I) = Q - 1$, so:

$$R_0 = \frac{3\lambda}{\nu} \tag{2.20}$$

which is consistent with figure 2.13. However if we consider the actual situation where clustering occurs early in the development of the epidemic, the derivation is wrong. In a real square lattice, infecteds will start to interfere with one another early on, so that $Q(S \mid I) < Q - 1$. Recovery will tend to weaken this interference but the interference is zero only when the recovery rate is so high that the disease dies out anyway. Figure 2.15 shows an example of this phenomenon: most of the infected individuals in this invading cluster have less than three susceptible neighbours. Because of this effect, in order to get a realistic expression for R_0 for a square lattice, our moment closure must incorporate clustering and our derivation for R_0 must take this clustering into account. For such a derivation it is not necessary to discard IPA or even OPA. In fact we can use them along with these insights about stochastic invasions in order to get a simple analytical expression for R_0 which captures invasions in spatial populations more realistically. An example of this is discussed in chapter 4.

We now turn our attention to figure 2.16, a bifurcation diagram of $[I]$ versus ϵ. Apparently, $[I] \to N$ as $\epsilon \to 0$ for all values of ν. When $\epsilon = 0$ we have $Q(I \mid SI) = 1$ and $Q(I \mid SS) = 0$ always. This means that if there is an SI or SS pair, they do not have any infected neighbours. This is a contradiction unless all individuals in the population are infected (in which case there are no SI or SS pairs); hence $[I] = N$ necessarily. However the case of very small ϵ is not physically realistic so we do not have to be too concerned with this limit. The important region is $\epsilon > 1$, and we

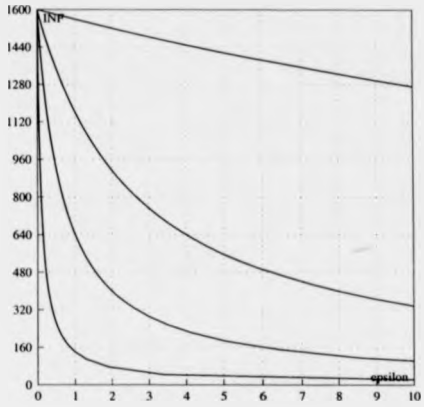


Figure 2.16: Bifurcation diagram of $[I]$ versus ϵ. Parameters: $N = 1600$, $\lambda = 0.1$ Values of ν are, from top curve to bottom, 0.01, 0.1, 0.2, 0.28.

note that $[I]$ is monotone decreasing in ϵ here. This means that the clustering of infecteds decreases the equilibrium size of the epidemic. By considering the slope of $[I]_\infty$ as a function of ϵ in figure 2.16, it seems that $[I]_\infty$ is most sensitive to changes in ϵ when $o(\nu) = o(\lambda)$. For large or small values of ν, changing ϵ has less effect on the equilibrium size of $[I]$. The fact that this holds for small ν is not surprising since when most people are infected, higher-order correlations do not matter as much. However when ν is large this is more interesting because it implies that even when higher-order correlations are present, changes in clustering can have smaller effects on the final size of the epidemic. This is not obvious; one might be led to conclude the opposite on the basis of table 2.2 since as ν decreases the divergence between the actual value of $Q(I \mid SI)$ and the value of $Q(I \mid SI)$ according to the assumption of a binomial OPA increases.

2.3.2 The Invasory Pair Approximation

Having reviewed some of the existing improvements to the OPA we now present a novel pair approximation which is based on knowledge of the invasion process gleaned from results of the stochastic model and analysis of IPA. It is a quite straightforward restructuring of the OPA.

One of the drawbacks of OPA has to do with its inability to incorporate clustering. For instance we expect in a real network that $Q(I \mid SI) > 1$, yet for the binomial

OPA, $Q(I \mid SI)$ is approximated by:

$$\begin{aligned} Q(I \mid SI) &= 1 + \frac{Q-1}{Q}\frac{[SI]}{[S]} \\ &= 1 + \frac{Q-1}{Q}Q(I \mid S) \end{aligned}$$

To see why OPA does poorly, consider the situation where the number of infecteds is small relative to the number of susceptibles. Then, in the limit $[I] \to 0$ we have $[SI] \to 0$, but also $[S] \to N$, and so $Q(I \mid SI) \to 1$. This contradicts what we would expect, since a local structure measure like $Q(I \mid SI)$ should not depend on N, and it also contradicts simulation data showing independence of $Q(I \mid SI)$ from N in the early stages of the epidemic.

A more appropriate moment closure which corrects for this particular flaw of OPA is:

$$\begin{aligned} Q(I \mid SI) &= 1 + \frac{Q-1}{Q}\frac{[SI]}{[S]_{inv}} \\ &= 1 + \frac{Q-1}{Q}Q(I \mid S)_{inv} \end{aligned} \tag{2.21}$$

where $[S]_{inv}$ is defined as the number of susceptible individuals within or bordering the active infected patches. Once we obtain an expression $[S]_{inv}$ we need only $Q(I \mid SS)$ to have a closed system of equations. Applying the constraint $Q = Q(I \mid SI) + Q(S \mid SI)$ and going through a similar calculation to that which resulted in identity (2.17) produces an expression for $Q(I \mid SS)$:

$$Q(I \mid SS) = \frac{Q-1}{Q}Q(I \mid S)_{inv}$$

However this is inaccurate because the heuristic calculation for $Q(I \mid SI)$ is intended to apply only within the invading cluster whereas $Q(I \mid SS)$ applies to the whole population. In our numerical analysis, when we implemented this expression for $Q(I \mid SS)$, the approximation unsurprisingly became highly inaccurate. In fact what we should do is to divide $Q(I \mid SS)$ into contributions to $Q(I \mid SS)$ from within the invading clusters ($Q(I \mid SS)_{inv}$) and contributions external to the invading clusters ($Q(I \mid SS)_{ext}$). Within the invading cluster we have

$$Q(I \mid SS)_{inv} = \frac{Q-1}{Q}Q(I \mid S)_{inv} \tag{2.22}$$

and outside

$$Q(I \mid SS)_{ext} = 0$$

In the master equations (2.1) and (2.4) we could equally well have used the term $\lambda[SS]_{inv}Q(I \mid SS)_{inv}$ instead of $\lambda[SS]Q(I \mid SS)$, since infection events by definition

can only happen at the boundaries of or within the invading clusters:

$$[SS]Q(I \mid SS) = [ISS] = [SS]_{inv}Q(I \mid SS)_{inv}$$

and so we do not actually have to use the expression (2.22) for $Q(I \mid SS)inv$ or derive a new one for $[SS]_{inv}$; we can simply use the usual binomial OPA expression for $Q(I \mid SS)$, at least for the equations of motion for this model:

$$Q(I \mid SS) = \frac{Q-1}{Q}Q(I \mid S)$$

This convenience of being able to use the old OPA expression for $Q(I \mid SS)$ applies to all the models we analyze in this thesis.

We call this pair approximation the *invasory pair approximation* (IVPA), although it could also be thought of as a modified IPA approach where ϵ is estimated from empirical knowledge of spatial dynamics. IVPA basically corrects for the existence of patches of purely susceptible individuals in the population; with the OPA these pure patches of susceptibles 'weigh down' the approximation and cause it to be inaccurate. At all times, both before and after equilibrium is reached, there will be at least some pockets of susceptibles which are distinct from the active infected patches, and so we expect $[S]_{inv} < [S]$ always if $[S]_{inv}$ is well-chosen. Therefore with IVPA we should see improved accuracies for all values of $[I]_\infty$ and for the transient phase as well. However one problem is that the IVPA does not correct for all of the long-range correlations which determine $Q(I \mid SI)$. And a more fundamental problem with IVPA is robustness, since $[S]_{inv}$ is calculated differently for different models and the derivation is often heuristic; we often do not know how good the approximation will be until it is implemented and investigated numerically.

Here we give an example of deriving an IVPA for the contact process. From figure 2.1 and similar data in the literature [29][40], and from observations which have been made about invasion in the CP [31], we know that behind the advancing wavefront, the system has reached a local pseudo-equilibrium which resembles the final equilibrium when the infection has spread throughout the network. In fact the process of invasion in some sense continues even at final equilibrium as active patches of infecteds migrate across the lattice. In this case we will define $[S]_{inv}$ as the number of susceptibles inside an active infected patch, excluding those susceptibles at or near the patch border. We can calculate the final equilibrium values under OPA and use these to calculate

1. the number of susceptibles within active infected patches, and
2. how large the pockets of susceptibles are in the final equilibrium.

Both factors can be incorporated into the derivation of $[S]_{inv}$.

The equations of motion for the contact process under OPA are:

$$\begin{aligned}
\frac{d}{dt}[SI] &= -\nu[SI] + \nu[II] - \lambda[SI]\left(1 + \frac{Q-1}{Q}\frac{[SI]}{[S]}\right) + \lambda[SS]\frac{Q-1}{Q}\frac{[SI]}{[S]} \\
\frac{d}{dt}[II] &= -2\nu[II] + 2\lambda[SI]\left(1 + \frac{Q-1}{Q}\frac{[SI]}{[S]}\right) \\
\frac{d}{dt}[SS] &= +2\nu[SI] - 2\lambda[SS]\frac{Q-1}{Q}\frac{[SI]}{[S]}
\end{aligned} \tag{2.23}$$

and the equation of motion for $[I]$ is:

$$\frac{d}{dt}[I] = -\nu[I] + \lambda[SI] \tag{2.24}$$

At equilibrium in a population of size N on a lattice we have:

$$\begin{aligned}
[SI]_\infty &= \nu QN\frac{\lambda Q - \nu - \lambda}{\lambda(\lambda Q^2 - \lambda Q - \nu)} \\
[II]_\infty &= \frac{\lambda^2Q^2 - \lambda^2 Q - 2\nu\lambda Q + \nu^2 + \nu\lambda)QN}{\lambda(\lambda Q^2 - \lambda Q - \nu)} \\
[SS]_\infty &= QN - 2[SI]_\infty - [II]_\infty \\
[I]_\infty &= \frac{\lambda}{\nu}[SI]_\infty
\end{aligned} \tag{2.25}$$

where the notation $[\cdot]_\infty$ denotes the nontrivial equilibrium value of quantity $[\cdot]$. The local network structure of the epidemic is approximately constant for large values of ν as the disease spreads through the network (see figure 2.10). For smaller values of ν the local structure changes more significantly over time. Thus we say $Q(I \mid S)^{inv}$ at equilibrium has the same value $Q(I \mid S)^{inv}$ as it does during the transient phase, *as long as* we restrict ourselves to the infected patches:

$$\begin{aligned}
Q(I \mid S)^{inv} &= Q(I \mid S)^{inv}_\infty \\
\Rightarrow \frac{[SI]}{[S]^{inv}} &= \frac{[SI]_\infty}{[S]^{inv}_\infty}
\end{aligned} \tag{2.26}$$

We need to approximate $[S]^{inv}_\infty$ in order to simplify further. We know that $[S]^{inv}_\infty < [S]_\infty$ since there exist pockets of susceptibles even at equilibrium. Therefore to get $[S]^{inv}_\infty$ we must subtract those individuals. The probability that a susceptible does not have any infected partners, i.e. is a member of such a pocket of susceptibles, we take to be $(Q(S \mid S)_\infty/Q)^Q$. Therefore:

$$[S]^{inv}_\infty = [S]_\infty\left(1 - \left[\frac{Q(S \mid S)_\infty}{Q}\right]^Q\right) \tag{2.27}$$

Substituting this expression into equation (2.28) gives an expression for $[S]^{inv}$:

$$\Rightarrow [S]^{inv} = [SI]\frac{[S]_\infty(1 - (Q(S \mid S)_\infty/Q)^Q)}{[SI]_\infty} \tag{2.28}$$

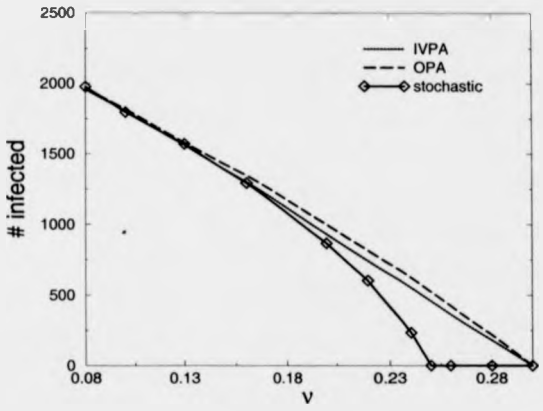


Figure 2.17: A comparison of the accuracy of IVPA and OPA for the contact process. $[I]_\infty$ is plotted against the recovery rate ν. Each black diamonds represents a long-time average of a single stochastic run, Parameters: $N = 2500$, $\lambda = 0.1$.

where $[SI]_\infty$ and $[S]_\infty$ are obtained from equations (2.25). We retain the binomial OPA expression for $Q(I \mid SS)$.

This derivation of $[S]_\infty$ is imperfect in a number of ways. One drawback is the time independence of $Q(I \mid S)_{inv}$. One expects some variation in $Q(I \mid S)_{inv}$ with time, especially for cases of smaller recovery rate ν, as suggested by the time series for $Q(I \mid SI)$ in figure 2.10. However this derivation for $[S]_{inv}$ does not incorporate any time dependence. Also, the only difference between $Q(I \mid S)_{eq}$ and $Q(I \mid S)^{inv}_{eq}$ comes in through equation (2.28), which is a rough and conservative estimate and will seriously underestimate the size of the invading clusters when $[I]_\infty$ is low. However in chapter 3 we will see a derivation of $[S]_\infty$ for an STD model which works much better than this derivation for the contact process. We presented the IVPA for the contact process first because the contact process, in all other ways, is an easier introduction to moment closure approximations

We compare the predictions for $[I]_\infty$ under OPA and IVPA in figure 2.3.2. IVPA has slightly increased accuracy for all values of ν but, as expected, does not provide a perfect fit to the simulation data. Also it predicts a wrong threshold value of $\nu = 0.3$, the same value predicted by the IPA and the OPA.

The IVPA also changes the speed with which the epidemic spreads through the network. Figure 2.18 shows time series for $[I]$ for OPA and IVPA. The IVPA time series shows slower epidemic growth for the same initial conditions and parameters,

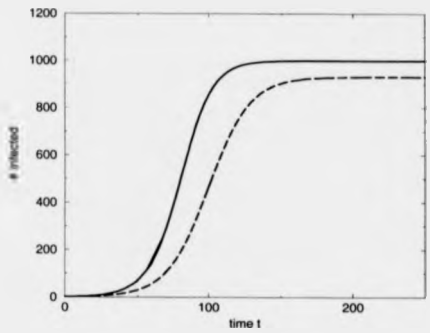


Figure 2.18: Time Series of $[I]$ under IVPA and OPA. The solid line is the OPA time series and the dashed line is the IVPA time series. Parameters: $N = 2500$, $\lambda = 0.1$, $\nu = 0.2$. Initial conditions: $SI = 4$, $II = 0$.

reproducing the slowing effect of clustering seen in stochastic simulations. From our distinction between $Q(I \mid S)$ and $Q(I \mid S)_{inv}$ we would expect IVPA to have a slowing effect on the transient phase of the epidemic. Additionally, we note that if we set $[S]^{eq}_{inv} = [S]^{eq}$ in equation 2.27, then the final size of the epidemic is the same for IVPA and OPA for all parameter choices, but the IVPA time series still approaches equilibrium more slowly than the OPA time series. The slowing which is observed is entirely because the IVPA partially accounts for the largest-scale patches in the network but not for the intermediate-scale correlations.

The trickiest part of IVPA approximations is defining and calculating $[S]_{inv}$, since numerous possibilities exist and some empirical knowledge of the infection correlation structure may come into the derivation (for instance if we want to include susceptibles near the patch border). Approaches may be very ad-hoc or more rigorous. We will see other examples of IVPA in chapters 3 and 4 for STD models.

why this is a bad example....

2.4 Discussion

This chapter investigated the applicability of moment closure approximations for modelling invasion and the equilibrium behaviour near critical points, using the contact process on a square lattice as a subject of study.

In part I we considered improving the moment closure approximation by using a higher-order cluster approximation, an approach not often used in MCA work in ecology and epidemiology. We found that a cluster approximation in this particular

application does not improve on the ordinary pair approximation (OPA). This illustrates the fact that going to higher-order is not always desirable. Also, both cluster and pair approximations share some deficiencies. For instance both are unable to reproduce the linearity of the time series in the transient invading phase of the contact process, and both are equally bad in predicting $[I]_\infty$ accurately near the critical point.

In part II we looked at alternative pair approximations. We implemented the improved pair approximation of Sato et al. [88] in order to study the problem of invasion and behaviour near critical points. We found that not all pair approximations are able to characterize well the epidemic near critical points and in the invading phase of the epidemic. This is on account of the localized, individual-based nature of invasion in spatial populations and the importance of higher-order correlations near the critical point. We also introduced a new pair approximation called the invasory pair approximation, which relies for its derivation on certain phenomenological knowledge of spatialized invasions. We found that IVPA is more accurate than OPA in predicting $[I]_\infty$, at least for the parameter choices we studied. More importantly, IVPA slows down epidemic spread because it incorporates clustering, and it predicts values of $[I]_\infty$ which are marginally closer to the simulation data. Also, IVPA is based on a simple insight into the spatial dynamics, is intuitive, and can be applied to all moment closure approximations in ecology. However one of the drawbacks is that for each new model we must derive a unique expression for $[S]_{inv}$ (the number of susceptibles within the active infected patches of the network). Also, and partially because of this first weakness, there are questions about the robustness of the IVPA approach. For some choices of $[S]_{inv}$ it can underpredict the final size significantly, in the case of the contact process.

Most attempts to improve the pair approximation have used approaches similar to the ones reviewed and/or introduced here. That is, they have resorted to higher order approximations or applied heuristic arguments to get better pair approximations. As we can see, such approaches can prove difficult. Perhaps what is needed it a more exhaustive survey of the physics literature of the twentieth century, where the issues of criticality are investigated. Certainly some powerful tools could be cross-applied from solid-state physics to ecological PA models. However, to our knowledge, this has not been attempted yet by ecology PA theorists.

This takes us to the next chapter, where we shift our focus from static regular lattices to dynamic network models for sexually transmitted diseases. Although we will be studying a different kind of network, much of what we have discussed in the first two chapters is relevant, and the differences between the epidemiology of the

contact process and dynamic network models are interesting.

Chapter 3

STD Dynamics in a Steady/Casual Partnership Network

They certainly give very strange names to diseases.

– Socrates

3.1 Introduction

As discussed in chapter 1, population heterogeneity, and particularly the structure of the sexual partnership network, are especially important when it comes to understanding the epidemiology of sexually transmitted diseases. One of the most salient forms of heterogeneity in sexual partnership networks is the existence of two types of partnerships: long-term monogamous partnerships and short-term casual partnerships which are often concurrent to the long-term partnerships [67] (*concurrency* is the term used in STD epidemiology to describe the situation where individuals have more than one partner simultaneously). Because this appears to be a common feature of real sexual networks, it is worthwhile to study the impact of this dichotomy on STD epidemiology. Evidence from sociological surveys indicates that sexual behaviour in the two types of partnerships differs, for instance with individuals in casual partnerships being more likely to practice safe sex with their casual partners than with their steady partners [66]. Such differences can be incorporated into our steady/casual model and their effect studied.

Dynamic and irregular networks are more well-suited to modelling STDs than regular static networks are. Some of the qualities of such networks have been discussed in chapter 1. The characteristics that make these networks different from regular static lattices are that:

1. edges can be created and destroyed;
2. the network does not necessarily correspond well to an n-dimensional geographical space as regular lattices do;
3. there can be variation in the number of edges per node.

Also, for the steady/casual partnership model in particular (and for many STD network models) there are two further differences:

1. the average number of edges per node is lower;
2. there are two types of edges.

The special characteristics of dynamic and irregular networks must be taken into account in deriving a moment closure approximation. For instance if the network is very disconnected and the average node degree is low, then the neat picture of spatial invasion with spreading patches of infecteds for regular lattices discussed in chapter 2 must be altered. For one thing, the nature of clustering is different; triangles are rare in the networks we use and so clustering will not have the slowing

effect on invasion that it does in regular lattices where one finds triangles. This is because individuals can be partners with anyone else in the population, and so the probability that an individual's two partners are also partners with one another is negligible for large population sizes. Of course one could also define dynamic and irregular networks where this is not the case.

Also, as noted in chapter 1, we have the two dynamical processes of partnership network formation and disease transmission through the partnership network. The network of the infected-susceptible contact structure is superimposed on the underlying partnership network structure; so effectively we have to concern ourselves with two networks instead of just one. This makes the analysis more complicated. For instance, the relative time scales of partnership dynamics and infection dynamics can be important. As the time scale for edge creation and dissolution decreases, the model approaches a mean-field situation. Another difference of the MCA for the steady/casual model is that there are two types of edges, which is not a common assumption.

Previous researchers have studied models with steady and casual partnerships using simulations and pair formation models. Kretzschmar et al. [66][67] have analyzed a stochastic model with steady and concurrent casual partnerships, concluding that a small number of concurrent casual partnerships can have a significant impact on epidemic spread and final size. They reach similar conclusions with a deterministic pair formation model where individuals can have either steady or casual partners but not both simultaneously. Our model is the same as the model in [67], the steady/casual partnership model with concurrency. Although we do not necessarily expect any new epidemiological insights, we are providing more flexible and realistic deterministic approximations for sexual partnership networks, and this is a valuable step forward.

This work is also an opportunity to begin exploring moment closure approximations for dynamic and irregular networks. To our knowledge little research has been done in this area, and certainly not for any biological applications. Most models analyzed have been regular lattices or irregular static networks where the number of neighbours Q per person is constant (in the latter case, a novel pair approximation must be used to incorporate the clustering which characterizes irregular static networks where all nodes have the same degree [60][75][90]). The fact that most models considered are square lattices stems partially from the fact that they are simpler to analyze, and partially from the fact that the original area of application of pair approximations was in solid-state physics, where one finds regular crystalline structures.

It is also possible to modify the pair approximation for regular lattices in other ways to incorporate some aspects of dynamic, irregular networks. For example, allowing for empty sites in an SIR model on a regular lattice can result in a binomial distribution for the number of occupied neighbouring sites per site, which is closer to the reality of social networks than assuming a fixed number of neighbouring sites per site [59]. A lattice model which allows for the existence of empty sites could thus give similar results to a dynamic, irregular network which does not have empty sites. However even with this approach there are limitations and it would not be the ideal solution to apply such an model to sexual and social networks.

In this chapter we will describe the stochastic model, derive a number of moment closure approximations and compare them, and analyze the model to answer epidemiological questions. We derive several moment closure approximations to illustrate the unique and significant challenges of applying MCA to dynamic, irregular networks. Once we have derived our approximations we will analyze them in order to answer questions of epidemiological interest, although some of these questions have been explored by other research, as already noted.

3.2 Description of the Stochastic Model

We start with a population of N individuals. They can form steady partnerships (which are monogamous and long-lived) with one other individual, and can also form casual partnerships (which are short-lived and may be concurrent to one other and to steady partnerships). Thus, an epidemic of an STD does not change people's sexual habits in this model. Rates are defined for the formation and dissolution of both types of partnerships, for transmission of the disease from an infected to his susceptible partner, and for recovery to a susceptible state. A more detailed description of parameters is in table 3.2. This model is not written with any particular STD in mind, however many common diseases fit the SIS description we use, such as gonorrhea and syphilis. On the other hand this model is unsuitable for infections such as HIV which require a more complex description.

As we have defined the model, the infection dynamics are decoupled from the network dynamics; individuals do not die from the disease (i.e. there is no disease-induced mortality) and having the disease does not affect behaviour (i.e. there is no disease-induced morality). Yet this simplifying assumption does not mean that correlations of disease states between sites are not present, and we still expect to find clustering of infecteds in areas of the network where node degrees are higher. There are other ways in which this model is unrealistic. In real sexual networks,

X	number of single individuals (i.e. lacking a steady partner)
N	population size
$\frac{\rho f}{X}$	rate at which any two single individuals form a steady partnership
$\frac{\rho(1-f)}{N}$	rate at which any two individuals form a casual partnership
σ_m	steady partnership separation rate
σ_c	casual partnership separation rate
f	fraction of newly-formed partnerships which are steady
λ_m	disease transmission rate in a steady SI partnership
λ_c	disease transmission rate in a casual SI partnership
ν	recovery rate

Table 3.1: Parameters for steady/casual partnership model.

we find *core groups*, which are high-activity assortatively-mixing subgroups of the population. They are very important for STD epidemiology because the STD is often able to persist in the general population only because of the presence of core groups. The large average node degree and within-group mixing makes it easier for the disease to persist within the core group. Our model, however, does not incorporate them.

The question posed in the introduction about the impact of casual partnerships can be investigated by considering the behaviour of the model under changes in f, a parameter which neatly encapsulates information on the relative numbers of steady and casual partnerships.

In practice, when running a simulation we will first allow the partnership dynamics to come to an equilibrium. Then we will introduce the disease into the wholly susceptible population by inoculating a few individuals, and we will study the resulting epidemic.

Figure 3.1 shows the time series for $[I]$ for the stochastic model. As can be seen, the invasion of the disease is linear in time until it approaches equilibrium. In this respect the invasion on a dynamic irregular network is the same as invasion in regular lattice models such as the contact process. Data on the local structure evolution would also be valuable, but for this we need to average over many realizations, and our computational speed was too limited to accomplish this. However we expect that some of what applies to the contact process also applies here, with the disease establishing some initial pseudo-equilibrium (inoculated individuals quickly infect their steady and casual partners) and then spreading through the rest of the network as the partnership network reforms itself gradually.

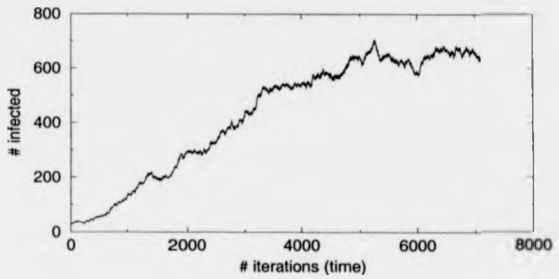


Figure 3.1: Stochastic time series for $[I]$. Parameters: $\rho = 0.001$, $\sigma_c = 0.005$, $\sigma_s = 0.00012$, $\lambda_s = \lambda_c = 0.14$, $\nu = 0.00125$, $f = 0.95$, $N = 1500$. Initial conditions: 15 randomly selected individuals were inoculated with the infection.

$[S]^x$	number of single susceptibles
$[I]^x$	number of single infecteds
$[SI]^c$	number of casual infected-susceptible partnerships
$[II]^c$	twice the number of casual infected-infected partnerships
$[SS]^c$	twice the number of casual susceptible partnerships
$[SI]^m$	number of steady infected-susceptible partnerships
$[II]^m$	twice the number of steady infected-infected partnerships
$[SS]^m$	twice the number of steady susceptible-susceptible partnerships

Table 3.2: State variables for steady/casual partnership model.

3.3 Derivation of Master Equations

In this section we derive the master equations up to the pair level. The state variables for the approximation are shown in table 3.3. The notation for the state variables could be simplified but we choose this notation for conformity with the notation in the rest of the thesis. Because there are two types of edges, superscripts m and c denote steady and casual partnerships respectively.

Because of our assumption that partnership dynamics are at equilibrium when the disease is introduced into the population, the following constraints on the state variables apply:

$$\begin{aligned} 2C &= 2[SI]^c + [II]^c + [SS]^c \\ 2M &= 2[SI]^m + [II]^m + [SS]^m \\ X &= [S]^x + [I]^x \end{aligned} \tag{3.1}$$

where C is the equilibrium number of casual partnerships, M is the equilibrium

number of steady partnerships, and X is the equilibrium number of single individuals ('single' means they do not have a steady partner). These constraints reduce the number of equations from eight to five. We also note that the number of infecteds $[I]$ is given by $[I] = [I]^x + [SI]^m + [II]^m$.

We extend the notation of previous chapters in straightforward fashion using superscripts and subscripts m and c to denote the two types of partnerships. For instance, $Q^c(i \mid j^m k)$ is the expected number of state i casual partners of a state j individual who has a state k steady partner. Sometimes it will be convenient to use 'Q' formulation (e.g. $Q^c(i \mid j^m k)$) and at other times it will be better to use bracket $[ijk]^{cm}$ formulation. We use both in this chapter and the following identities which relate the two conventions come in handy:

$$\begin{aligned}
Q^c(i \mid j) &= \tfrac{[ij]^c}{[j]} \\
Q^m(i \mid j) &= \tfrac{[ij]^m}{[j]} \\
Q^c(i \mid j^m k) &= \tfrac{[ijk]^{cm}}{[jk]^m} \\
Q^c(i \mid j^c k) &= \tfrac{[ijk]^{cc}}{[jk]^c} && i \neq k \\
Q^c(i \mid j^c i) &= \tfrac{[iji]^{cc}}{[ij]^c} + [ij]^c \\
Q^m(i \mid j^c k) &= \tfrac{[kji]^{cm}}{[jk]^c} \\
Q^m(i \mid j^m k) &= 0 && i \neq k \\
Q^m(i \mid j^m i) &= 1
\end{aligned} \tag{3.2}$$

We illustrate the pair approximation with a full derivation of the equation of motion for $[SI]^m$. We must extend slightly the notation used for the derivation in chapter 1 because the steady/casual model involves edge formation. Let $\varsigma_x =$ (state description) denote the state of individual x. For instance $\varsigma_x = (M)$ indicates that individual x is in a steady partnership, and may be infected or uninfected. Let $\varsigma_{xy} =$ (state description) denote the state of two individuals (x, y) who may or may not be in partnership. For instance $\varsigma_{xy} = (S, X; I, M)$ indicates that x is a susceptible and single, and y is infected and in a steady partnership, but not with x. On the other hand a statement such as $\varsigma_{xy} = (S^m I)$ indicates that x is susceptible, y is infected, and they are steady partners.

We write terms for each of the processes of partnership formation, partnership

dissolution, infection and recovery as in chapter 2:

$$\begin{aligned}\frac{d}{dt}[SI]^m &= \sum_{\varsigma_z=(S,X)} \frac{\rho f}{X}[I]^x(+1) - \sum_{\varsigma_{zy}=S^mI} \sigma(-1) + \sum_{\varsigma_{zy}=S^mI} \lambda_m(-1) \\ &+ \sum_{\varsigma_{zy}=S^mI} \lambda_c Q^c_z(I)(-1) + \sum_{\varsigma_{zy}=S^mS} \lambda_c Q^c_z(I)(+1) \\ &+ \sum_{\varsigma_{zy}=S^mI} \nu(-1) + \sum_{\varsigma_{zy}=I_mI} \nu(+1)\end{aligned}$$

We expand terms such as $Q^c_z(I)$ as $Q^c(I \mid S^mI) + \eta_z(I \mid S^mI)$ as usual; we take the averages out of the sums, and note that the sums of the η fluctuation terms are zero as before:

$$\begin{aligned}\frac{d}{dt}[SI]^m &= -\sigma_m[SI]^m + \frac{\rho f}{X}[S]^x[I]^x - \lambda_m[SI]^m - \nu[SI]^m + \nu[II]^m \\ &-\lambda_c \sum_{\varsigma_{zy}=S^mI} (Q^c(I \mid S^mI) + \eta^c_z(I \mid S^mI)) \\ &+\lambda_c \sum_{\varsigma_{zy}=S^mS} (Q^c(I \mid S^mS) + \eta^c_z(I \mid S^mS)) \\ &= -\sigma_m[SI]^m + \frac{\rho f}{X}[S]^x[I]^x - \lambda_m[SI]^m - \nu[SI]^m + \nu[II]^m \\ &-\lambda_c[SI]^mQ^c(I \mid S^mI) + \lambda_c[SS]^mQ^c(I \mid S^mS) \qquad (3.3)\end{aligned}$$

which leaves us with higher-order correlations $Q^c(I \mid S^mI)$ and $Q^c(I \mid S^mS)$ which must be approximated. Both terms arise because of the presence of concurrent casual partnerships.

The master equations for $[II]^m$ and $[SS]^m$ are derived similarly:

$$\begin{aligned}\frac{d}{dt}[II]^m &= -\sigma_m[II]^m + \frac{\rho f}{X}([I]^x)^2 - 2\nu[II]^m \\ &+2\lambda_m[SI]^m + 2\lambda_c[SI]^mQ^c(I \mid S^mI) \qquad (3.4) \\ \frac{d}{dt}[SS]^m &= -\sigma_m[SS]^m + \frac{\rho f}{X}([S]^x)^2 + 2\nu[SI]^m \\ &-2\lambda_c[SS]^mQ^c(I \mid S^mS) \qquad (3.5)\end{aligned}$$

For casual partnerships we follow the same procedure. For $[SI]^c$:

$$\begin{aligned}\frac{d}{dt}[SI]^c &= \sum_{\varsigma_{zy}=S^cI} \sigma_c(-1) + \sum_{\varsigma_z=S} \frac{\rho(1-f)}{N}[I](+1) \\ &+ \sum_{\varsigma_{zy}=S^cI} \nu(-1) + \sum_{\varsigma_{zy}=I^cI} \nu(+1) \\ &+ \sum_{\varsigma_{zy}=S^cI} \lambda_c(-Q^c_z(I)) + \sum_{\varsigma_{zy}=S^cI} \lambda_c(+Q^c_z(S)) \\ &+ \sum_{\varsigma_{zy}=S^mI} \lambda_m(-Q^c_z(I)) + \sum_{\varsigma_{zy}=S^mI} \lambda_m(+Q^c_z(S))\end{aligned}$$

$$
\begin{aligned}
&= -\sigma_c[SI]^c + \frac{\rho(1-f)}{N}[S][I] - \nu[SI]^c + \nu[II]^c \\
&\quad -\lambda_c \sum_{\varsigma_{zy}=S^cI} (Q^c(I \mid S^cI) + \eta_z^c(I \mid S^cI)) \\
&\quad +\lambda_c \sum_{\varsigma_{zy}=S^cI} (Q^c(S \mid S^cI) + \eta_z^c(S \mid S^cI)) \\
&\quad -\lambda_m \sum_{\varsigma_{zy}=S^mI} (Q^c(I \mid S^mI) + \eta_z^c(I \mid S^mI)) \\
&\quad +\lambda_m \sum_{\varsigma_{zy}=S^mI} (Q^c(S \mid S^mI) + \eta_z^c(S \mid S^mI))
\end{aligned}
$$

$$
\begin{aligned}
&= -\sigma_c[SI]^c + \frac{\rho(1-f)}{N}[S][I] - \nu[SI]^c + \nu[II]^c \\
&\quad -\lambda_c[SI]^cQ^c(I \mid S^cI) + \lambda_c[SI]^cQ^c(S \mid S^cI) \\
&\quad -\lambda_m[SI]^mQ^c(I \mid S^mI) + \lambda_m[SI]^mQ^c(S \mid S^mI) \qquad (3.6)
\end{aligned}
$$

And for $[SS]^c$ and $[II]^c$ the derivation is similar:

$$
\begin{aligned}
\frac{d}{dt}[II]^c &= -\sigma_c[II]^c + \frac{\rho(1-f)}{N}[I]^2 - 2\nu[II]^c \\
&\quad +2\lambda_c[SI]^cQ^c(I \mid S^cI) + 2\lambda_m[SI]^mQ^c(I \mid S^mI) \qquad (3.7) \\
\frac{d}{dt}[SS]^c &= -\sigma_c[SS]^c + \frac{\rho(1-f)}{N}[S]^2 - 2\nu[SI]^c \\
&\quad -2\lambda_c[SI]^cQ^c(S \mid S^cI) - 2\lambda_m[SI]^mQ^c(S \mid S^mI) \qquad (3.8)
\end{aligned}
$$

For the equation of motion of $[I]^x$, the number of infected individuals who do not have a steady partner, we do not have to calculate any third-order correlations:

$$
\begin{aligned}
\frac{d}{dt}[I]^x &= \sum_{\varsigma_z=S_x} \frac{\rho f}{X}[I]^x(-1) + \sum_{\varsigma_z=I_x} \frac{\rho f}{X}\frac{1}{2}[I]^x(-2) + \sum_{\varsigma_{zy}=S_mI} \sigma_m(+1) \\
&\quad + \sum_{\varsigma_{zy}=I_mI} \sigma_m(+1) + \sum_{\varsigma_z=I_x} \nu(-1) + \sum_{\varsigma_z=S_x} \lambda_c Q_z^c(I)(+1) \\
&= -\frac{\rho f}{x}[S]^x[I]^x - \frac{\rho f}{x}([I]^x)^2 + \sigma_m[SI]^m + \sigma_m[II]^m - \nu[I]^x \\
&\quad +\lambda_c \sum_{\varsigma_z=S_x} (Q^c(I \mid S^x) + \eta_z^c(I \mid S^x)) \\
&= -\rho f[I]^x + \sigma_m[SI]^m + \sigma_m[II]^m - \nu[I]^x + \lambda_c[IS]^{cx} \qquad (3.9)
\end{aligned}
$$

where $[IS]^{cx}$ is the number of casual susceptible-infected partnerships where the susceptible individual has no steady partner. $[II]^{cx}$ is similarly defined. These pairs are ordered, so we do not count II^{cx} partnerships twice (as we do with II^c partnerships for instance). Luckily, we can express $[IS]^{cx}$ and $[II]^{cx}$ exactly in terms of the state

variables:

$$\begin{aligned} [II]^{cx} &= [II]^c - [IIS]^{cm} - [III]^{cm} \\ [SI]^{cx} &= [SI]^c - [SIS]^{cm} - [SII]^{cm} \end{aligned} \tag{3.10}$$

which give the following identities:

$$\begin{aligned} Q^c(I \mid I^x) &= \frac{1}{[I]^x}([II]^c - [IIS]^{cm} - [III]^{cm}) \\ Q^c(I \mid S^x) &= \frac{1}{[S]^x}([SI]^c - [ISS]^{cm} - [ISI]^{cm}) \end{aligned} \tag{3.11}$$

Similarly for $[S]^x$ we have:

$$\frac{d}{dt}[S]^x = -\rho f[S]^x + \sigma_m[SI]^m + \sigma_m[SS]^m + \nu[I]^x - \lambda_c[IS]^{cx} \tag{3.12}$$

From the master equations (?? - (3.9) it is possible to determine the quantities M, C and X of equation (3.1):

$$\begin{aligned} M &= \frac{\rho f N}{2(\rho f + \sigma_m)} \\ C &= \frac{\rho(1-f)N}{2\sigma_c} \\ X &= \frac{\sigma_m N}{\rho f + \sigma_m} \end{aligned} \tag{3.13}$$

For the master equations (3.3) – (3.12) to be closed, the higher-order quantities $Q^c(I \mid S^m I)$, $Q^c(I \mid S^m S)$, $Q^c(S \mid S^c I)$, $Q^c(I \mid S^c I)$ and $Q^c(S \mid S^m I)$ must be approximated. Our first attempt at closure is with the OPA.

3.4 Closing the Master Equations with the OPA

The OPA for the steady/casual model is obtained in heuristic fashion by considering conditional independence and also taking network structure into account. From these considerations, we arrive at the expressions:

$$\begin{aligned} Q^c(I \mid S^c I) &= 1 + \frac{Q^c_{>1} - 1}{Q^c}\frac{[SI]^c}{[S]} \\ Q^c(S \mid S^c I) &= \frac{Q^c_{>1} - 1}{Q^c}\frac{[SS]^c}{[S]} \\ Q^c(I \mid S^m I) &= \frac{[SI]^c}{[S]} \\ Q^c(S \mid S^m I) &= \frac{[SS]^c}{[S]} \\ Q^c(I \mid S^m S) &= \frac{[SI]^c}{[S]} \end{aligned} \tag{3.14}$$

where $Q^c_{\geq 1}$ is the expected number of casual partners of an individual who has at least one casual partner. This quantity is calculated from knowledge of the distribution of casual partners per person, i.e. a Poisson distribution with mean $2C/N$:

$$Q^c_{\geq 1} = \frac{\rho(1-f)}{\sigma_c(1-e^{-\rho(1-f)/\sigma_c})} \tag{3.15}$$

The last three approximations hold because of the independence of the processes of casual and steady partner acquisition, while in the first two approximations we need the factor $(Q^c_{\geq 1} - 1)/Q^c$ to account for the fact that the susceptible individual already has at least one casual partner. This is similar to the multiplicative factor $(Q-1)/Q$ used in the binomial OPA for regular lattices.

3.4.1 Comparison of OPA with the Simulation Data

A stochastic Monte-Carlo simulation was implemented to test the stochastic model and compare it with the deterministic approximations. Once partnership dynamics have equilibrated, $N/5$ randomly selected individuals are inoculated and we start to take a running time-average for $[I]$ and other variables after the infection dynamics have apparently reached an equilibrium. Although $N/5$ is a large inoculation, we were interested only in comparing the predictions for $[I]_\infty$ of the stochastic model and the deterministic approximations.

Figure 3.2 shows poor agreement of the OPA and the simulation data. The OPA is much more inaccurate here than in the other biological models to which it has been applied since Katori's first paper [56]. The most likely reason is that the number of edges per node is typically lower, and thus higher-order correlations are much more important and the assumption of conditional independence breaks down. This inaccuracy for low Q has also been seen in regular lattice models. For other parameter choices perhaps agreement would be better, but the parameters we have chosen were not arbitrary but rather based on what what we would expect for real sexual networks and real STDs. The high values of transmissibility λ_c and λ_m also make the higher-order correlations more important, since the disease is spread quickly within steady-casual triples, and slower relative partnership dynamics move us away from the mean-field limit where conditional independence applies. The agreement worsens closer to the critical point for much the same reasons as in other moment closure approximations.

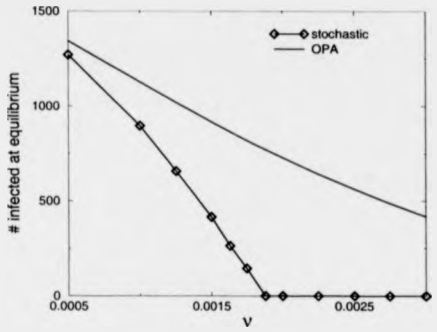


Figure 3.2: Comparison of results from stochastic model and OPA. The epidemic final size is plotted against the recovery rate ν. Other parameters are $\rho = 0.001$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $f = 0.95$, $\lambda_c = 0.14$, $\lambda_m = 0.14$, $N = 1500$.

3.5 Four Other Moment Closure Approximations

With these observations in mind we seek alternative moment closure approximations which are more accurate. We consider in this section an invasory pair approximation, a triple approximation, a *pseudo-equilibrium pair approximation* and a hybrid pair approximation. Although the work required to check and compare all four approximations is significant, we feel that an improvement on the OPA is definitely necessary, and that comparing various approaches will help us gain an intuition for when certain approximations work and with what types of models.

3.5.1 Invasory Pair Approximation

The IVPA for the contact process was described in chapter 2. We implement an IVPA for the steady/casual model by following a similar procedure:

$$\begin{aligned}
Q^c(I \mid S^c I) &= 1 + \frac{Q^c_{>1} - 1}{Q^c} \frac{[SI]^c}{[S]_{inv}} \\
Q^c(S \mid S^c I) &= \frac{Q^c_{>1} - 1}{Q^c} \frac{[SS]^c}{[S]_{inv}} \\
Q^c(I \mid S^m I) &= \frac{[SI]^c}{[S]_{inv}} \\
Q^c(S \mid S^m I) &= \frac{[SS]^c}{[S]_{inv}} \\
Q^c(I \mid S^m S) &= \frac{[SI]^c}{[S]}
\end{aligned} \tag{3.16}$$

We note that the issue of constraints does not come up in deriving the IVPA for the steady/casual model. This is because the variance in node degree guarantees that $Q^c[SI]^c \neq [ISS]^{cc} + [ISI]^{cc}$; an individual with large (resp. small) node degree is more likely to have proportionately more (resp. less) infected partners. We must take a different approach in defining and deriving $[S]_{inv}$ than we did for the contact process because the network structure is somewhat different. The fact that the network is dynamic and the average node degree is low means that we use a somewhat different picture of spatial invasion than the one developed in chapter 2. So for the steady/casual model let us derive $[S]_{inv}$ by saying a susceptible is part of an active infected patch if it is connected to an infected individual through a steady or a casual partnership. We can approximate the number of susceptibles who fit this description as:

$$[S]_{inv} \approx [SI]^m + [SI]^c - \frac{[SI]^m[SI]^c}{[S]}$$

The negative cross term accounts for those susceptible individuals who are members of both a casual and a steady partnership. This expression overestimates $[S]_{inv}$ somewhat because the subtracted term $\frac{[SI]^m[SI]^c}{[S]}$ is a conservative estimate for the number of overlapping partnerships.

As a side note, another possible choice for $[S]_{inv}$ which leads to less accurate values for I_∞ is $[S]_{inv} = [SI]^m + [SS]^m[I]/N + [S]^x$. This estimate is more conservative and only applies when most individuals in the population are in a steady partnership.

3.5.2 Ordinary Triple Approximation

In the *ordinary triple approximation* (OTA) we go to a third-order moment closure approximation. The qualifier 'ordinary' is used because the assumption of conditional independence of disease status is also made in the OTA as in the OPA. The motivation for OTA is as follows. In the steady/casual model we are usually interested in the effect of a small number of casual partnerships in a population where steady partnerships dominate. We look at areas of the parameter space where there are only a few casual partnerships around, and when $[I]_\infty$ is not too large we expect most of the epidemic activity to be happening where individuals have both a steady and a casual partner. This is a different situation from other network models where the number of edges per node is higher and where there is only one type of edge. Thus, the seeming importance of steady-casual triples suggests that for the steady/casual model we need a third-order approximation.

Our OTA is therefore based on concurrent steady-casual partnerships. Although the dimensionality and complication of the new system of equations is considerable, the extension from the pair case is mostly straightforward and the reader can skip the derivation without loss.

Consider the state variable $[ijk]^{\alpha\gamma}$ which is the number of triples where a state j individual is in a type γ partnership with a state k individual and a type α partnership with a state i individual. We will consider only steady-casual triples in this approximation, and not the casual-casual triples which also occur (these are rare when f is close to one). The total number of cm triples T^{cm} is constant when partnership dynamics have equilibrated, thus:

$$\begin{aligned} T^{cm} &= [III]^{cm} + [IIS]^{cm} + [ISI]^{cm} + [ISS]^{cm} \\ &\quad + [SII]^{cm} + [SIS]^{cm} + [SSI]^{cm} + [SSS]^{cm} \end{aligned} \tag{3.17}$$

Because the cm triples are ordered, we do not count any of them twice. On the other hand we count cc triples twice because of the counting convention used in moment closure approximations, which simplifies the final approximation. We can calculate the value of T^{cm} from the partnership network dynamics. Because the casual and steady partnership dynamics are independent from one another, we have $T^{cm} = MC/2N$, which gives us:

$$T^{cm} = \frac{\rho^2 f(1-f)}{(\rho f + \sigma_m)\sigma_c} N \tag{3.18}$$

We derive the equation of motion for $[SSI]^{cm}$ as an example of the OTA. The dynamical processes affecting $[SSI]^{cm}$ are:

$S^cS^mI \rightarrow$	S^cS^mS	recovery
$S^cS^mI \rightarrow$	S^cI^mI	infection
$S^cS^mI \rightarrow$	I^cS^mI	infection
$S^cS^mS \rightarrow$	S^cS^mI	infection
$S^cI^mI \rightarrow$	S^cS^mI	recovery
$I^cS^mI \rightarrow$	S^cS^mI	recovery
$S^cS^mI \rightarrow$	$S^cS + I$	steady partnership dissolution
$S^cS^mI \rightarrow$	$S + S^mI$	casual partnership dissolution
$S^cS + I \rightarrow$	S^cS^mI	steady partnership formation
$S + S^mI \rightarrow$	S^cS^mI	casual partnership formation

Therefore our equation of motion is:

$$\begin{aligned}\frac{d}{dt}[SSI]^{cm} &= \sum_{\varsigma_{xyz}=S^cS^mI}\nu(-1)+\sum_{\varsigma_{xyz}=S^cS^mS}\lambda_cQ_z^c(I)(+1)\\ &+\sum_{\varsigma_{xyz}=S^cS^mI}(\lambda_mQ_y^m(I)+\lambda_cQ_y^c(I))(-1)\\ &+\sum_{\varsigma_{xyz}=S^cS^mI}(\lambda_mQ_x^m(I)+\lambda_cQ_x^c(I))(-1)\\ &+\sum_{\varsigma_{xyz}=S^cI^mI}\nu(+1)+\sum_{\varsigma_{xyz}=I^cS^mI}\nu(+1)\\ &+\sum_{\varsigma_{xyz}=S^cS^mI}\sigma_m(-1)+\sum_{\varsigma_{xyz}=S^cS^mI}\sigma_c(-1)\\ &+\sum_{\varsigma_{xy}=S^cS^x}\frac{\rho f}{X}[I]^x(+1)+\sum_{\varsigma_{xy}=S^mI}\frac{\rho(1-f)}{N}[S](+1)\end{aligned}\tag{3.19}$$

We apply the substitutions for the higher-order correlations much as before:

$$\begin{aligned}Q_x^c(I) &= Q^c(I\mid S^cS^mI)+\eta_x^c(I\mid S^cS^mI) && \varsigma_{xyz}=S_cS_mI\\ Q_x^m(I) &= Q^m(I\mid S^cS^mI)+\eta_x^m(I\mid S^cS^mI) && \varsigma_{xyz}=S_cS_mI\\ Q_y^c(I) &= Q^c(I\mid S<_m^{cS}I)+\eta_y^c(I\mid S<_m^{cS}I) && \varsigma_{xyz}=S_cS_mI\\ Q_y^m(I) &= 1 && \varsigma_{xyz}=S_cS_mI\\ Q_z^c(I) &= Q^c(I\mid S^mS^cS)+\eta_z^c(I\mid S^mS^cS) && \varsigma_{xyz}=S_cS_mS\end{aligned}\tag{3.20}$$

where $Q^c(I \mid S <_m^{cS} I)$ is the expected number of infected casual partners of a susceptible with an infected steady partner and a susceptible casual partner. As usual the stochastic fluctuations represented by the η terms sum to zero by definition. Taking constants out of the summations and evaluating sums produces:

$$\begin{aligned}\frac{d}{dt}[SSI]^{cm} &= -\nu[SSI]^{cm}-\left(\lambda_m+\lambda_cQ^c(I\mid S<_m^{cS}I)\right)[SSI]^{cm}\\ &-(\lambda_mQ^m(I\mid S^cS^mI)+\lambda_cQ^c(I\mid S^cS^mI))[SSI]^{cm}\\ &+\lambda_cQ^c(I\mid S^mS^cS)[SSS]^{cm}+\nu[SII]^{cm}+\nu[ISI]^{cm}\\ &-(\sigma_m+\sigma_c)[SSI]^{cm}+\frac{\rho(1-f)}{N}[S][SI]^m\\ &+\frac{\rho f}{X}[I]^x([SS]^c-[SSI]^{cm}-[SSS]^{cm})\end{aligned}\tag{3.21}$$

Similarly for the other triple quantities:

$$\begin{aligned}
\frac{d}{dt}[III]^{cm} &= -3\nu[III]^{cm} + (\lambda_m + \lambda_c Q^c(I \mid S^m I^c I))[IIS]^{cm} \\
&\quad +(\lambda_m + \lambda_c Q^c(I \mid S <^{cI}_{mI}))[ISI]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S^c I^m I) + \lambda_m Q^m(I \mid S^c I^m I))[SII]^{cm} \\
&\quad -(\sigma_c + \sigma_m)[III]^{cm} + \frac{\rho(1-f)}{N}[I][II]^m \\
&\quad +\frac{\rho f}{X}[I]^x([II]^c - [IIS]^{cm} - [III]^{cm}) \qquad (3.22)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[IIS]^{cm} &= -2\nu[IIS]^{cm} - (\lambda_m + \lambda_c Q^c(I \mid S^m I^c I))[IIS]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S <^{cI}_{mS}) + 0)[ISS]^{cm} + \nu[III]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S^c I^m S) + \lambda_m Q^m(I \mid S^c I^m S))[SIS]^{cm} \\
&\quad -(\sigma_c + \sigma_m)[IIS]^{cm} + \frac{\rho(1-f)}{N}[I][SI]^m \\
&\quad +\frac{\rho f}{X}[S]^x([II]^c - [IIS]^{cm} - [III]^{cm}) \qquad (3.23)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[ISI]^{cm} &= -2\nu[ISI]^{cm} - (\lambda_c Q^c(I \mid S <^{mI}_{cI}) + \lambda_m)[ISI]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S^m S^c I) + 0)[ISS]^{cm} + \nu[III]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S^c S^m I) + \lambda_m Q^m(I \mid S^c S^m I))[SSI]^{cm} \\
&\quad -(\sigma_c + \sigma_m)[ISI]^{cm} + \frac{\rho(1-f)}{N}[I][SI]^m \\
&\quad +\frac{\rho f}{X}[I]^x([SI]^c - [SIS]^{cm} - [SII]^{cm}) \qquad (3.24)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[ISS]^{cm} &= -\nu[ISS]^{cm} - (\lambda_c Q^c(I \mid S^m S^c I) + 0)[ISS]^{cm} \\
&\quad -(\lambda_c Q^c(I \mid S <^{cI}_{mS}) + 0)[ISS]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S^c S^m S) + \lambda_m Q^m(I \mid S^c S^m S))[SSS]^{cm} \\
&\quad +\nu([IIS]^{cm} + [ISI]^{cm}) \\
&\quad -(\sigma_c + \sigma_m)[ISS]^{cm} + \frac{\rho(1-f)}{N}[I][SS]^m \\
&\quad +\frac{\rho f}{X}[S]^x([SI]^c - [SIS]^{cm} - [SII]^{cm}) \qquad (3.25)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SII]^{cm} &= -2\nu[SII]^{cm} + (\lambda_c Q^c(I \mid S <^{cS}_{mI}) + \lambda_m)[SSI]^{cm} \\
&\quad -(\lambda_c Q^c(I \mid S^c I^m I) + \lambda_m Q^m(I \mid S^c I^m I))[SII]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S^m I^c S) + \lambda_m)[SIS]^{cm} + \nu[III]^{cm} \\
&\quad -(\sigma_c + \sigma_m)[SII]^{cm} + \frac{\rho(1-f)}{N}[S][II]^m \\
&\quad +\frac{\rho f}{X}[I]^x([SI]^c - [SIS]^{cm} - [SII]^{cm}) \qquad (3.26)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SIS]^{cm} &= -\nu[SIS]^{cm} - (\lambda_c Q^c(I \mid S^m I^c S) + \lambda_m)[SIS]^{cm} \\
&\quad -(\lambda_c Q^c(I \mid S^c I^m S) + \lambda_m Q^m(I \mid S^c I^m S))[SIS]^{cm} \\
&\quad +(\lambda_c Q^c(I \mid S <^{cS}_{mS}) + 0)[SSS]^{cm} + \nu([IIS]^{cm} + [SII]^{cm}) \\
&\quad -(\sigma_c + \sigma_m)[SIS]^{cm} + \frac{\rho(1-f)}{N}[S][SI]^m \\
&\quad +\frac{\rho f}{X}[S]^x([S^{-1}]^c - [SIS]^{cm} - [SII]^{cm}) \qquad (3.27)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SSS]^{cm} &= -(\lambda_c Q^c(I \mid S^c S^m S) + \lambda_m Q^m(I \mid S^c S^m S))[SSS]^{cm} \\
&\quad -(\lambda_c Q^c(I \mid S <^{cS}_{mS}) + 0)[SSS]^{cm} \\
&\quad -(\lambda_c Q^c(I \mid S^m S^c S) + 0)[SSS]^{cm} \\
&\quad +\nu([SSI]^{cm} + [SIS]^{cm} + [ISS]^{cm}) \\
&\quad -(\sigma_c + \sigma_m)[SSS]^{cm} + \frac{\rho(1-f)}{N}[S][SS]^m \\
&\quad +\frac{\rho f}{X}[S]^x([SS]^c - [SSI]^{cm} - [SSS]^{cm}) \qquad (3.28)
\end{aligned}$$

These equations involve pair quantities such as $[SI]^m$. For these we can use the pair master equations (3.3) - (3.12) except that we do not have to approximate the third-order correlations. To reduce the dimensionality of the system of equations one could also express the pair quantities in terms of the steady-casual triples using constraints.

To close the equations we must express the fourth-order correlations in terms of lower-order correlations. Let us use conditional independence again, making adjust-

ments for the partnership network structure:

$$\begin{array}{lcll}
Q^c(I \mid S^cS^mj) & \approx & Q^c(I \mid S^cS) & \approx & \frac{Q^c_{\geq 1}-1}{Q^c}Q^c(I \mid S) \\
Q^c(I \mid S^cI^mj) & \approx & Q^c(I \mid S^cI) & \approx & 1+\frac{Q^c_{\geq 1}-1}{Q^c}Q^c(I \mid S) \\
Q^m(I \mid S^cj^mk) & \approx & Q^m(I \mid S^cj) & = & \frac{[jSI]^{cm}}{[Sj]^c} \\
Q^c(I \mid S^mj^ck) & \approx & Q^c(I \mid S^mj) & = & \frac{[ISj]^{cm}}{[Sj]^m} \\
Q^c(I \mid S <^{cS}_{mj}) & \approx & \frac{Q^c_{\geq 1}-1}{Q_c}Q^c(I \mid S^mj) & = & \frac{Q^c_{\geq 1}-1}{Q_c}\frac{[ISj]^{cm}}{[Sj]^m} \\
Q^c(I \mid S <^{cI}_{mj}) & \approx & 1+\frac{Q^c_{\geq 1}-1}{Q_c}Q^c(I \mid S^mj) & = & \frac{Q^c_{\geq 1}-1}{Q_c}\frac{[ISj]^{cm}}{[Sj]^m}
\end{array} \tag{3.29}$$

where $Q^m(I \mid S) = [SI]^m/[S]$ and $Q^c(I \mid S) = [SI]^c/[S]$. Most of these approximations are straightforward application of conditional independence to higher-order cases. The last two for $Q^c(I \mid S <^{cS}_{mj})$ and $Q^c(I \mid S <^{cI}_{mj})$ are special because of the T-shaped element they involve. We lose more information in approximating these than we do with the other correlations because we carry out a pair approximation when we neglect the disease status of the attached casual partner. However this set of approximations should be sufficient and we do not want to increase the complexity of the OTA more than necessary.

3.5.3 Pseudo-Equilibrium Pair Approximation

In the pseudo-equilibrium pair approximation we employ heuristic arguments and apply certain assumptions to boil down the system to only the most important higher-order correlations. We exploit the fact that steady partnerships are long-lived. We also assume concurrent casual partnership are rare. Then we calculate the pseudo-equilibrium of the local structure and can thus approximate the higher order terms in equations (3.3) - (3.12), incorporating them into a pair approximation. Although this approach is complicated, the resulting equations are not worse than the OTA, and it is hoped that we can simplify things by capturing the most important correlations. We calculate the pseudo-equilibrium values for each of the higher order correlations $Q^c(I \mid S^mI)$, $Q^c(S \mid S^mI)$, $Q^c(I \mid S^cI)$ and $Q^c(S \mid S^cI)$ individually according to this method. Again, reading the first derivation for $Q^c(I \mid S^mI)$ and $Q^c(S \mid S^mI)$ is enough for the reader to understand the method; the succeeding four derivations can be skipped over.

Calculation of $Q^c(I \mid S^mI)$ and $Q^c(S \mid S^mI)$

Consider two individuals in a steady partnership and assume that the partnership does not break up over the time scales we are considering (the time scale of infection,

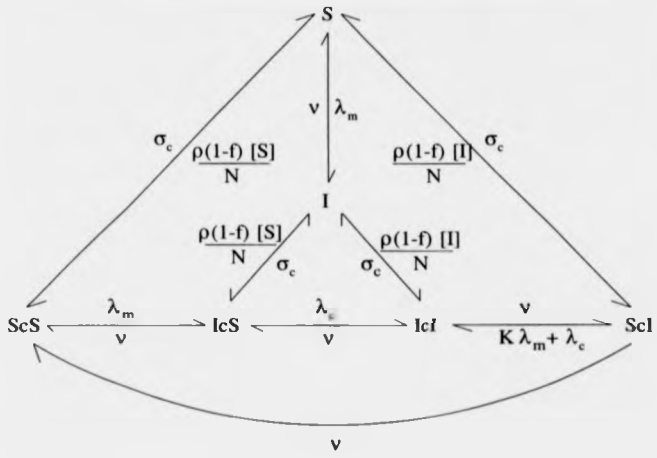


Figure 3.3: Detailed balance diagram for calculation of $Q^c(I \mid S^m I)$ and $Q^c(S \mid S^m I)$.

recovery and casual partnership dynamics). Furthermore, assume one of the partners is infected and remains infected for all time; we call him the *fixed individual*. His partner, the *variable individual* can switch between infected and susceptible states and can have up to one casual partner. This assumption of a fixed individual is hard to justify but it will be seen that the method works well notwithstanding this gross simplification.

Let g_I (resp. g_S) be the expected proportion of time the variable individual is infected (resp. susceptible) and has no casual partners; let g_{S_cS} (resp. g_{I_cS}) be the expected proportion of time the variable individual is susceptible (resp. infected) and has one casual susceptible partner; let g_{S_cI} (resp. g_{I_cI}) be the expected proportion of time the variable individual is susceptible (resp. infected) and has one infected casual partner. We have $g_I+g_S+g_{S_cS}+g_{I_cS}+g_{S_cI}+g_{I_cI} = 1$. The rates at which the steady pair moves from one state to another are determined from the model parameters; the detailed balance diagram from which we calculate the pseudo-equilibrium for this arrangement is shown in figure 3.3.

The extra factor K we take to be $\frac{Q^m(I|S^cI)}{Q^m}$ where Q^m is the average number of steady partners per person. This constant is included as a partial correction for the unrealistic assumption that the fixed individual is infected for all time. The equilibrium of this configuration must be calculated to give us the quantities g_S, g_I, g_{S_cS}, g_{S_cI}, g_{I_cS} and g_{I_cI} in terms of pairs and singletons. We use these equilibrium

values to obtain the expressions for higher order correlations:

$$\begin{aligned} Q^c(I \mid S^m I) &= \frac{g_{S_c I}}{g_{S_c I} + g_{S_c S} + g_S} \\ Q^c(S \mid S^m I) &= \frac{g_{S_c S}}{g_{S_c I} + g_{S_c S} + g_S} \end{aligned} \tag{3.30}$$

The explicit solutions for $Q^c(I \mid S^m I)$ and $Q^c(S \mid S^m I)$ are not written down because they are too large and cannot be simplified. The same is unfortunately true for the other higher-order correlations which are calculated in the remainder of this subsection. Thus pencil-and-paper analysis is not feasible under the pseudo-equilibrium approximation and we must stick to numerical methods. However this complexity at least partially reflects the complexity of the model itself since we have so many processes occurring at once.

Calculation of $Q^c(I \mid S^c I)$ and $Q^c(S \mid S^c I)$

The derivation here is very similar to the derivation for $Q^c(I \mid S^m I)$ and $Q^c(S \mid S^m I)$. We again assume the existence of a fixed infected individual and a variable individual, connected this time by a casual partnership. The partnership is also assumed to last for a sufficiently long time, even though in this case the assumption is much less accurate. Let g_I (resp. g_S) be the expected proportion of time the variable individual is infected (resp. susceptible) and has no extra casual partner, let $g_{S_c S}$ (resp. $g_{I_c S}$) be the expected proportion of time the variable individual is susceptible (resp. infected) and has one casual susceptible partner, and let $g_{S_c I}$ (resp. $g_{I_c I}$) be the expected proportion of time the variable individual is susceptible (resp. infected) and has one infected casual partner. The detailed balance diagram for this arrangement is shown in figure 3.4 and we calculate the equilibrium solution to get $Q^c(I \mid S^c I)$ and $Q^c(S \mid S^c I)$ as before.

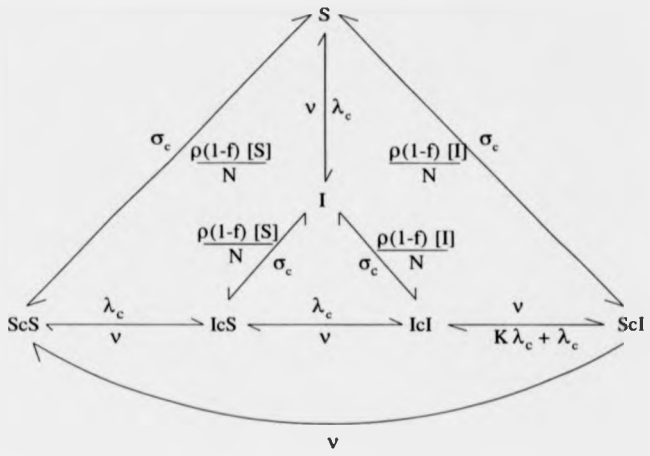


Figure 3.4: Detailed balance diagram for calculation of $Q^c(I \mid S^cI)$ and $Q^c(S \mid S^cI)$

Calculation of $Q^c(I \mid S^mS)$ and $Q^c(I \mid S^z)$

$Q^c(I \mid S^mS)$ can be calculated from the other state variables:

$$\begin{aligned}
Q^c(I \mid S^mS) &= \frac{[I^cS^mS]}{[SS]^m} \\
&= \frac{[I^cS^mS]}{[SI]^c}\frac{[SI]^c}{[SS]^m} \\
&= Q^m(S \mid S^cI)\frac{[SI]^c}{[SS]^m} \\
&= (Q^m - Q^m(I \mid S^cI))\frac{[SI]^c}{[SS]^m} \\
&= \left(Q^m - \frac{[I^mS^cI]}{[SI]^c}\right)\frac{[SI]^c}{[SS]^m} \\
&= \left(Q^m - \frac{[I^mS^cI]}{[SI]^m}\frac{[SI]^m}{[SI]^c}\right)\frac{[SI]^c}{[SS]^m} \\
&= \left(Q^m - Q^c(I \mid S^mI)\frac{[SI]^m}{[SI]^c}\right)\frac{[SI]^c}{[SI]^m}
\end{aligned}$$

3.5.4 Hybrid Pair Approximation

The final approximation we apply to the steady/casual model is the hybrid pair approximation [38][40], described briefly in chapters 1 and 2. For the HPA it is necessary for the stochastic model to incorporate events which have the opposite effect of nearest-neighbour infection, i.e. events which hamper the clustering caused

by disease transmission through edges. We will see how this is necessary to form a pair approximation that overestimates the local correlations, so that we can combine it with the underestimating OPA to obtain a hybrid pair approximation in which the effects are balanced. Therefore we must incorporate *nearest-neighbour recovery* in our steady/casual model. In nearest-neighbour recovery, there is a certain probability that the infected neighbours of a recovering individual also recover. We do not have to consider this effect in the eventual analysis of the model, and we are able to take the limit where the probability of nearest-neighbour recovery is zero while still using HPA. Nearest-neighbour recovery has a real-world analogue in STD epidemiology and corresponds to the practice of contact tracing.

The extent of nearest-neighbour recovery in the population is controlled by the parameter α, where $0 \leq \alpha \leq 1$. For $\alpha = 1$, all infected partners of an infected individual recover when he recovers, and for $\alpha = 0$, none of them do. For other values of α there is a probability α that a given infected partner recovers. Here are the altered equations of motion for the steady/casual model with nearest-neighbour recovery:

$$\begin{aligned}
\frac{d}{dt}[SI]^m &= -\sigma_m[SI]^m + \frac{\rho f}{X}[S]^x[I]^x - \lambda_m[SI]^m - \nu[SI]^m + \nu[II]^m \\
&\quad -\lambda_c[ISI]^{cm} + \lambda_c[ISS]^{cm} + \alpha\nu[III]^{cm} - \alpha\nu[IIS]^{cm} \\
\frac{d}{dt}[II]^m &= -\sigma_m[II]^m + \frac{\rho f}{X}([I]^x)^2 - 2\nu[II]^m \\
&\quad +2\lambda_m[SI]^m + 2\lambda_c[ISI]^{cm} - 2\alpha\nu[III]^{cm}
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SS]^m &= -\sigma_m[SS]^m + \frac{\rho f}{X}([S]^x)^2 + 2\nu[SI]^m \\
&\quad -2\lambda_c[ISS]^{cm} + 2\alpha\nu[IIS]^{cm} \\
\frac{d}{dt}[SI]^c &= -\sigma_c[SI]^c + \frac{\rho(1-f)}{N}[S][I] - \nu[SI]^c + \nu[II]^c \\
&\quad -\lambda_c([ISI]^{cc} + [SI]^c) + \lambda_c[SSI]^{cc} - \lambda_m[ISI]^{cm} + \lambda_m[SSI]^{cm} \\
&\quad +\alpha\nu[III]^{cm} + \alpha\nu[III]^{cc} - \alpha\nu[SII]^{cm} - \alpha\nu[SII]^{cc}
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[II]^c &= -\sigma_c[II]^c + \frac{\rho(1-f)}{N}[I]^2 - 2\nu[II]^c + 2\lambda_c([ISI]^{cc} + [IS]^c) \\
&\quad +2\lambda_m[ISI]^{cm} - 2\alpha\nu[III]^{cm} - 2\alpha\nu[III]^{cc} \\
\frac{d}{dt}[SS]^c &= -\sigma_c[SS]^c + \frac{\rho(1-f)}{N}[S]^2 - 2\nu[SI]^c - 2\lambda_c[SSI]^{cc} \\
&\quad -2\lambda_m[SSI]^{cm} + 2\alpha\nu[SII]^{cm} + 2\alpha\nu[SII]^{cc}
\end{aligned}$$

$$
\begin{aligned}
\frac{d}{dt}[I]^x &= -\rho f[I]^x + \sigma_m[SI]^m + \sigma_m[II]^m - \nu[I]^x \\
&\quad +\lambda_c([SI]^c - [ISS]^{cm} - [ISI]^{cm}) \\
&\quad -\alpha\nu([II]^c - [IIS]^{cm} - [III]^{cm}) \\
\frac{d}{dt}[S]^x &= -\rho f[S]^x + \sigma_m[SI]^m + \sigma_m[SS]^m + \nu[I]^x \\
&\quad -\lambda_c([SI]^c - [ISS]^{cm} - [ISI]^{cm}) \\
&\quad +\alpha\nu([II]^c - [IIS]^{cm} - [III]^{cm})
\end{aligned} \tag{3.31}
$$

We illustrate the process of deriving HPA for this fearsome system of equations, analyzing first the equation of motion for $[SI]^c$. Under the OPA we have the following approximation for the triples appearing in $\frac{d}{dt}[SI]^c$ which represent nearest-neighbour recovery:

$$
\begin{aligned}
&\alpha\nu([III]^{cm} + [III]^{cc} - [SII]^{cm} - [SII]^{cc}) \\
&= \alpha\nu\left(\frac{[II]^c[II]^m}{[I]} + \kappa\frac{([II]^c)^2}{[I]} - \frac{[SI]^c[II]^m}{[I]} - \kappa\frac{[SI]^c[II]^c}{[I]}\right)
\end{aligned} \tag{3.32}
$$

where $\kappa = (Q^c_{\geq 1} - 1)/Q^c$. We know that the OPA overestimates the final size of the epidemic. The triples in equation (3.32) represent the process of nearest-neighbour recovery, which counteracts the clustering caused by nearest-neighbour infection. We can construct a pair approximation that underestimates the final size of the epidemic by approximating the triples for nearest-neighbour recovery in the equation of motion (3.31) for $\frac{d}{dt}[SI]^c$ differently from the triples for infection; i.e. to overestimate clustering we want to make those triples smaller than they really are:

$$
\begin{aligned}
&\alpha\nu(-[SII]^{cm} - [SII]^{cc} + [III]^{cm} + [III]^{cc}) \\
&= -\alpha\nu([S*I]^{cm} - [SSI]^{cm} + [S*I]^{cc} - [SSI]^{cc}) \\
&\quad \alpha\nu([I*I]^{cm} - [ISI]^{cm} + [I*I]^{cc} - [ISI]^{cc}) \\
&= -\alpha\nu\left(\frac{2T^{cm}[S][I]}{N^2} - \frac{[SS]^c[SI]^m}{[S]} + \frac{2T^{cc}[S][I]}{N^2} - \kappa\frac{[SS]^c[SI]^c}{[S]}\right) \\
&\quad \alpha\nu\left(\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]} + \frac{2T^{cc}[I]^2}{N^2} - \kappa\frac{([SI]^c)^2}{[S]}\right)
\end{aligned} \tag{3.33}
$$

where the $*$ denotes an individual who can be either susceptible or infected. The approximations of terms with $*$ in them leave out even more of the local correlation structure than the OPA, and therefore the $[I*I]^{cm} - [ISI]^{cm}$ and $[I*I]^{cc} - [ISI]^{cc}$ terms are more positive than under the OPA, contributing to overestimation of the triple density (i.e. clustering).

What remains is to find a way to balance these under- and over-estimating approximations. We combine them into the equation of motion for $[SI]^c$ as follows,

using a mixing parameter δ:

$$\begin{aligned}
\frac{d}{dt}[SI]^c &= -\sigma_c[SI]^c + \frac{\rho(1-f)}{N}[S][I] - \nu[SI]^c + \nu[II]^c \\
&\quad -\lambda_c\left(\frac{\kappa([SI]^c)^2}{[S]} + [SI]^c\right) + \lambda_c\kappa\frac{[SS]^c[SI]^c}{[S]} - \lambda_m\frac{[SI]^c[SI]^m}{[S]} \\
&\quad +\lambda_m\frac{[SS]^c[SI]^m}{[S]} - (1-\delta)\left[\alpha\nu\left(\frac{[SI]^c[II]^m}{[I]} + \kappa\frac{[SI]^c[II]^c}{[I]}\right)\right] \\
&\quad -\delta\left[\alpha\nu\left(\frac{2T^{cm}[I][S]}{N^2} - \frac{[SS]^c[SI]^m}{[S]} + \frac{2T^{cc}[S][I]}{N^2}\right.\right. \\
&\quad \left.\left. -\kappa\frac{[SS]^c[SI]^c}{[S]}\right)\right] + (1-\delta)\left[\alpha\nu\left(\frac{[II]^c[II]^m}{[I]} + \kappa\frac{([II]^c)^2}{[I]}\right)\right] \\
&\quad +\delta\left[\alpha\nu\left(\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]} + \frac{2T^{cc}[I]^2}{N^2} - \kappa\frac{([SI]^c)^2}{[S]}\right)\right]
\end{aligned} \tag{3.34}$$

and the other equations of motion are similarly modified. Because we are interested in a network without nearest-neighbour recovery we need to consider the case $\alpha = 0$. In the limit $\alpha \to 0$ we want to know what happens to the product $\Delta \equiv \alpha\delta$. For the square lattice model studied Filipe et al. [40] it is possible to define and calculate the value of δ exactly by solving the equations of motion, and to use this value of δ to take the limit. For this model it is not easy to fix a value for δ because the solution to the equations of motion is too large. Also the same criterion used by Filipe to fix δ (he takes the least value of δ for which the transition from a stable trivial solution to a stable nontrivial solution is continuous) cannot be used because there is no such discontinuity in the steady/casual model.

Therefore we must fix δ empirically. This might not be a problem for a fixed grid because once the value of δ is set it should apply for all possible parameter choices (for the model of Filipe, δ depends only on Q which is usually fixed). But for a random dynamic network, the value of δ can only be fixed for a given choice of parameters since different partnership dynamics are expected to give different networks and hence different values for δ. Short of any other way to approximate δ we cannot get around this limitation with a model as complex as the steady/casual model. This is a general drawback in applying HPA to dynamic networks. All we can do is choose a value for δ which provides the best agreement in one case, and then compare the agreement produced in other cases.

So we must simply assume that Δ is nonzero in the limit $\alpha \to 0$. Therefore

equation (3.34) becomes:

$$\begin{aligned}
\frac{d}{dt}[SI]^c &= -\sigma_c[SI]^c + \frac{\rho(1-f)}{N}[S][I] - \nu[SI]^c + \nu[II]^c \\
&\quad -\lambda_c\left(\kappa\frac{([SI]^c)^2}{[S]} + [SI]^c\right) + \lambda_c\kappa\frac{[SS]^c[SI]^c}{[S]} - \lambda_m\frac{[SI]^c[SI]^m}{[S]} \\
&\quad +\lambda_m\frac{[SS]^c[SI]^m}{[S]} + \Delta\nu\left[\frac{[SI]^c[II]^m}{[I]} + \kappa\frac{[SI]^c[II]^c}{[I]}\right] \\
&\quad -\Delta\nu\left[\frac{2T^{cm}[I][S]}{N^2} - \frac{[SS]^c[SI]^m}{[S]} + \frac{2T^{cc}[S][I]}{N^2} - \kappa\frac{[SS]^c[SI]^c}{[S]}\right] \\
&\quad +\Delta\nu\left[\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]} + \frac{2T^{cc}[I]^2}{N^2} - \kappa\frac{([SI]^c)^2}{[S]}\right] \\
&\quad -\Delta\nu\left[\frac{[II]^c[II]^m}{[I]} + \kappa\frac{([II]^c)^2}{[I]}\right] \qquad (3.35)
\end{aligned}$$

The other equations are:

$$\begin{aligned}
\frac{d}{dt}[II]^c &= -\sigma_c[II]^c + \frac{\rho(1-f)}{N}[I]^2 - 2\nu[II]^c + 2\lambda_m\frac{[SI]^c[SI]^m}{[S]} \\
&\quad +2\lambda_c\left(\kappa\frac{([SI]^c)^2}{[S]} + [IS]^c\right) + 2\Delta\nu\left[\frac{[II]^c[II]^m}{[I]} + \kappa\frac{([II]^c)^2}{[I]}\right] \\
&\quad -2\Delta\nu\left[\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]} + \frac{2T^{cc}[I]^2}{N^2} - \kappa\frac{([SI]^c)^2}{[S]}\right]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SS]^c &= -\sigma_c[SS]^c + \frac{\rho(1-f)}{N}[S]^2 - 2\nu[SI]^c - 2\lambda_c\kappa\frac{[SS]^c[SI]^c}{[S]} \\
&\quad -2\lambda_m\frac{[SS]^c[SI]^m}{[S]} - 2\Delta\nu\left[\frac{[SI]^c[II]^m}{[I]} + \kappa\frac{[SI]^c[II]^c}{[I]}\right] \\
&\quad +2\Delta\nu\left[\frac{2T^{cm}[I][S]}{N^2} - \frac{[SS]^c[SI]^m}{[S]} + \frac{2T^{cc}[S][I]}{N^2} - \kappa\frac{[SS]^c[SI]^c}{[S]}\right]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SI]^m &= -\sigma_m[SI]^m + \frac{\rho f}{X}[S]^x[I]^x - \lambda_m[SI]^m - \nu[SI]^m + \nu[II]^m \\
&\quad -\lambda_c\frac{[SI]^c[SI]^m}{[S]} + \lambda_c\frac{[SI]^c[SS]^m}{[S]} + \Delta\nu\left[\frac{[II]^c[SI]^m}{[I]}\right] \\
&\quad -\Delta\nu\left[\frac{2T^{cm}[I][S]}{N^2} - \frac{[SI]^c[SS]^m}{[S]}\right] - \Delta\nu\left[\frac{[II]^c[II]^m}{[I]}\right] \\
&\quad +\Delta\nu\left[\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]}\right]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[SS]^m &= -\sigma_m[SS]^m + \frac{\rho f}{X}([S]^x)^2 + 2\nu[SI]^m - 2\lambda_c\frac{[SI]^c[SS]^m}{[S]} \\
&\quad -2\Delta\nu\left[\frac{[II]^c[SI]^m}{[I]}\right] + 2\Delta\nu\left[\frac{2T^{cm}[I][S]}{N^2} - \frac{[SI]^c[SS]^m}{[S]}\right]
\end{aligned}$$

$$
\begin{aligned}
\frac{d}{dt}[II]^m &= -\sigma_m[II]^m + \frac{\rho f}{X}([I]^x)^2 - 2\nu[II]^m + 2\lambda_m[SI]^m \\
&\quad +2\lambda_c\frac{[SI]^c[SI]^m}{[S]} + 2\Delta\nu\left[\frac{[II]^c[II]^m}{[I]}\right] \\
&\quad -2\Delta\nu\left[\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]}\right] \\
\\
\frac{d}{dt}[I]^x &= -\rho f[I]^x + \sigma_m[SI]^m + \sigma_m[II]^m - \nu[I]^x \\
&\quad +\lambda_c\left([SI]^c - \frac{[SI]^c[SS]^m}{[S]} - \frac{[SI]^c[SI]^m}{[S]}\right) \\
&\quad +\Delta\nu\left[\frac{2T^{cm}[I][S]}{N^2} - \frac{[SI]^c[SS]^m}{[S]}\right] - \Delta\nu\left[\frac{[II]^c[SI]^m}{[I]}\right] \\
&\quad +\Delta\nu\left[\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]}\right] - \Delta\nu\left[\frac{[II]^c[II]^m}{[I]}\right] \\
\frac{d}{dt}[S]^x &= -\rho f[S]^x + \sigma_m[SI]^m + \sigma_m[SS]^m + \nu[I]^x \\
&\quad -\lambda_c\left([SI]^c - \frac{[SI]^c[SS]^m}{[S]} - \frac{[SI]^c[SI]^m}{[S]}\right) \\
&\quad -\Delta\nu\left[\frac{2T^{cm}[I][S]}{N^2} - \frac{[SI]^c[SS]^m}{[S]}\right] + \Delta\nu\left[\frac{[II]^c[SI]^m}{[I]}\right] \\
&\quad -\Delta\nu\left[\frac{2T^{cm}[I]^2}{N^2} - \frac{[SI]^c[SI]^m}{[S]}\right] + \Delta\nu\left[\frac{[II]^c[II]^m}{[I]}\right]
\end{aligned}
\tag{3.36}
$$

where T^{cm} and T^{cc} are given by:

$$
\begin{aligned}
T^{cm} &= \frac{\rho^2 f(1-f)}{(\rho f + \sigma_m)\sigma_c}N \\
T^{cc} &= N\sum_{k=1}^{\infty}\left(\pi_k \times \sum_{n=1}^{k}(k-n)\right) \\
&= N\sum_{k=2}^{\infty}\left(\pi_k \times \frac{k(k-1)}{2}\right) \\
&= N\sum_{k=2}^{\infty}\left(\frac{e^{-k}(2C/N)^k}{k!} \times \frac{k(k-1)}{2}\right) \\
&= \frac{N}{2}\left(\frac{2C}{N}\right)^2 = \frac{N}{2}\left(\frac{\rho(1-f)}{\sigma_c}\right)^2
\end{aligned}
\tag{3.37}
$$

The expression for T^{cm} is the same as for the OTA. The first line in the derivation for T^{cc} comes from counting the contribution to T^{cc} from each node-type of degree k, according to its frequency π_k. T^{cm} and T^{cc} constrain the densities of casual-steady

and casual-casual triples:

$$\begin{aligned} T^{cm} &= [III]^{cm} + [IIS]^{cm} + [ISI]^{cm} + [ISS]^{cm} \\ &\quad + [SII]^{cm} + [SIS]^{cm} + [SSI]^{cm} + [SSS]^{cm} \\ 2T^{cc} &= [III]^{cc} + 2[IIS]^{cc} + [ISI]^{cc} + 2[ISS]^{cc} + [SIS]^{cc} + [SSS]^{cc} \end{aligned}$$

3.5.5 Comparing the Four Alternative Approximations

Figure 3.5 compares the alternative moment closure approximations to the simulation data and the OPA from figure 3.2. There are improvements for most of the new approximations, although some of them do not capture the qualitative aspects of the stochastic plot. The IVPA offers improvement for low recovery rates but lacks accuracy nearer to the critical point. However the pseudo-equilibrium approximation and particularly the OTA approximation offer significant improvements over the other approximations, although at the cost of increased complexity. The HPA deserves some special comment on account of the uncertainty in choosing Δ. Here we have rather arbitrarily set $\Delta = 0.6$. Other choices of Δ for the same parameters produce curves shown in figure 3.6; none of them really fits the stochastic data very well and so the conventional HPA as we have implemented does not work for the steady/casual model as well as it worked for the contact process. This certainly reflects some of the limitations of the HPA as we have implemented it, in particular our choice of only one mixing parameter δ instead of separate mixing parameters for each possible nearest-neighbour recovery event ($[III]^{cm} \rightarrow [SSI]^{cm}$, etc.). The difficulty in choosing a value for δ basically comes from the complexity of the steady/casual model and the inability to find simple solutions to the equations of motion. Perhaps for other parameter choices in the steady/casual model, HPA will produce agreement as good the other alternative approximations.

The overall conclusion from these comparisons is that taking triples into account is necessary for the steady/casual model and probably for any model where Q is low. This also reinforces our intuition about the relative importance of higher order correlations in such networks as in comparison with regular lattices with higher values of Q. The OTA is the most accurate approximation as long as f is not too small, and so we use it for the following analyses.

3.6 Analysis of the OTA of the Steady/Casual Model

This section consists of bifurcation diagrams and several time series of the equations of motion under the OTA. We vary the parameter f and study the properties of

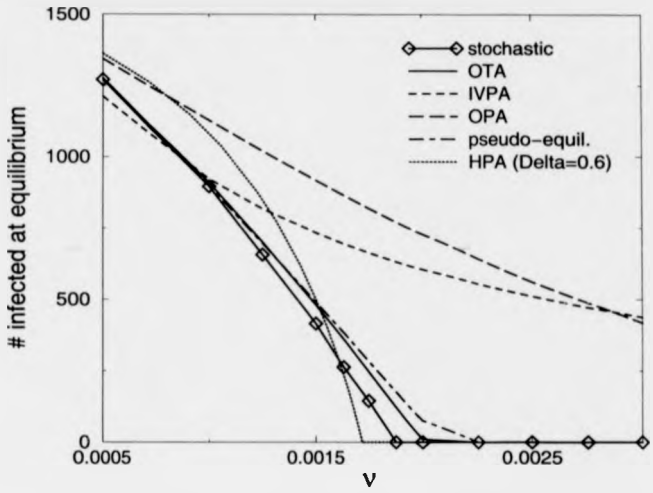


Figure 3.5: Comparison of results from stochastic model and five different moment closure approximations. The number of infecteds at equilibrium is plotted again the recovery rate ν. Other parameters are $\rho = 0.001$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $f = 0.95$, $\lambda_c = 0.14$, $\lambda_m = 0.14$, $N = 1500$.

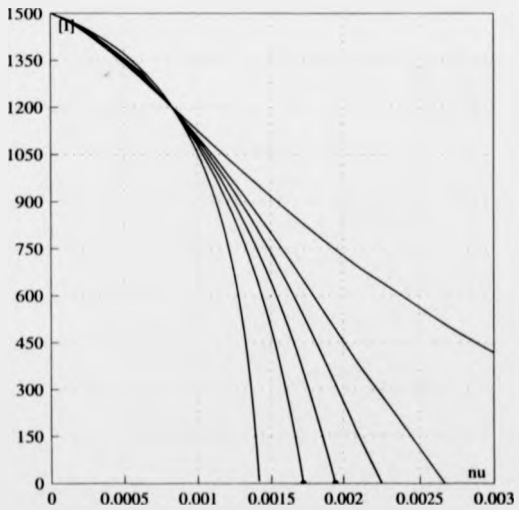


Figure 3.6: HPA for various values of Δ. Parameters: $\rho = 0.001$, $\lambda_c = \lambda_m = 0.14$, $\sigma_c = 0.005$, $\sigma_m = 0.00012$, $f = 0.95$. Values of Δ are, from top curve to bottom, $\Delta = 1.0, 0.8, 0.6, 0.5, 0.4, 0.2$.

nontrivial stable solution branches. As f is varied, the relative numbers of casual and steady partners are also varied. Changes in the nontrivial solution branches may reflect changes in the relative composition of steady and casual partners, or it may reflect changes in the absolute number of partnerships $P = C + M$. Ideally one should hold P constant while varying f to differentiate the effect of changing the distribution and number of casual partnerships from the change in the total number of partnerships. However for the parameters we choose, changing f has only a small impact on P so we do not hold P constant as f varies.

3.6.1 Early Time Evolution of Local Network Structure

Figure 3.7 shows the early time evolution of two measures of the local structure, $Q^c(I \mid S^mI)$ and $Q^c(S \mid S^mI)$, for the OTA. Also shown is the time series of $[I]$ for the same period of time. The population size is N so the illustrated time span covers only the earliest phase of the epidemic. The time span covers about one year and the establishment of pseudo-equilibrium requires approximately 50 days. The initial conditions are one infected person with a casual and a steady partner. The time scale over which these two individuals are infected is $o(1/\lambda_s) = o(1/\lambda_c) = 7$ days, explaining the speed of the initial jump in $[I]$. However the maxima for the local structure measures $Q^c(I \mid S^mI)$ and $Q^c(S \mid S^mI)$ occur somewhat later. The time scale for casual partnership dissolution is $o(1/\sigma_c) = 200$ days, and the location of the local maxima of $Q^c(I \mid S^mI)$ and $Q^c(S \mid S^mI)$ are influenced by the rate of turnover of casual partnerships. According to figure 2.9 it takes on the order of 100 days for the pseudo-equilibrium to be established, which is consistent with the fact that $1/\sigma_c = 200$ days. After the first approximately 100 days, the transients (whose characteristics depend on the initial conditions) have died down and the local network structure reaches a pseudo-equilibrium which changes only slowly over the course of the epidemic.

These data show that the description of early local structure evolution in regular static lattices also applies to dynamic, random lattices, and that this effect can be captured by an appropriate moment closure approximation. Also, although one must be careful in drawing conclusions about real-world epidemiology with these simplified models, we can say that the patterns of susceptible and infecteds in local areas of a sexual partnership network will remain relatively unchanged as the epidemic spreads through the population at large, unless there are disease-induced modifications in sexual behaviour. Certainly this conclusion would be altered by considering the high degree of clustering and the existence of core groups in real sexual networks.

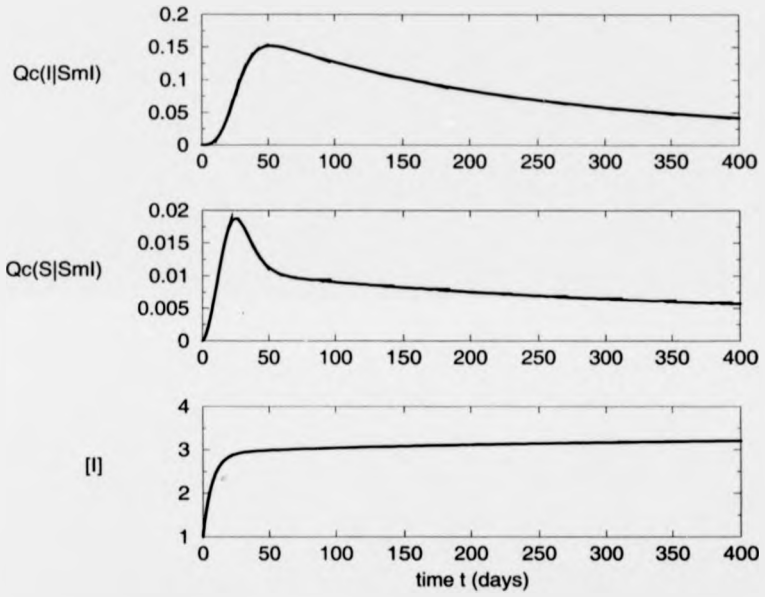


Figure 3.7: Evolution of local structure of OTA. Parameters are: $\rho = 0.001$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $f = 0.95$, $\lambda_m = 0.14$, $\lambda_c = 0.14$, $\nu = 0.002$, $N = 1500$. Initial conditions are one infected person with one steady and one casual partner: $[SI]^m = [SIS]^{cm} = [SI]^c = 1$.

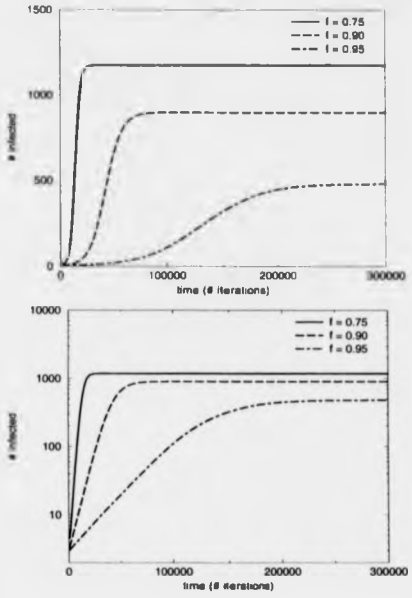


Figure 3.8: Time series of $[I]$ and log $[I]$ as a function of f. Parameters are: $\rho = 0.001$, $\sigma_s = 0.00012$, $\sigma_c = 0.005$, $\lambda_m = \lambda_c = 0.14$, $\nu = 0.0015$, $N = 1500$. Values of P are: $C = 7.8$, $M = 665.9$ for $f = 0.95$, $C = 15.0$, $M = 661.8$ for $f = 0.90$ and $C = 37.5$, $M = 646.6$ for $f = 0.75$. Initial conditions are one infected person with one casual and one steady partner: $[SI]^m = [SIS]^{cm} = [SI]^c = 1$.

3.6.2 Time Series of $[I]$ as a Function of f

Figure 3.8 shows several time series of $[I]_\infty$ for various values of f, starting from a single infected individual who is inoculated with the disease. As f decreases the growth rate and final size of the epidemic increase. Moreover, this large difference is caused by very small changes in the number C of casual partnerships.

3.6.3 Bifurcation Diagrams for $[I]$

Figure 3.9 shows data from a bifurcation diagram of $[I]$ versus f. The shape of the curve and the existence of the threshold are typical of epidemic systems. There is a stability-exchanging bifurcation where the nontrivial solution branch intersects the trivial solution branch. What is significant is that the bifurcation diagram shows how a small number of casual partnerships can have a very significant impact on $[I]_\infty$. Also shown in this figure is the total number of steady and casual partnerships

$P = C + M$ as a function of f. The total number of partnerships is roughly constant for most values of f but decreases significantly for small f. Thus we conclude it is the relative numbers of casual and steady partnerships which have this huge effect.

In figure 3.10 we consider a bifurcation diagram of $[I]$ versus λ_c. The motivation for this is the observation that individuals are more likely to practice safe sex with their casual partners than with their steady partners [66], and so studying the dependence of $[I]_\infty$ on λ_c can help us design control strategies for STDs. The two curves show two different recovery rates. In both cases there are regimes where reducing λ_c has little effect on $[I]_\infty$ and regimes where reducing λ_c can eradicate the disease. Thus safer sex practices will not necessarily reduce the incidence of STDs unless the endemicity is already close to the threshold. In practice one would have to be able to approximate the threshold value in a real population to assess the impact of more widespread condom use in casual partnerships. We also observe that as $\lambda_c \to \infty$, $[I]_\infty$ appears to asymptote, which reflects the restriction of the active disease spread to the small number of casual partnerships and those few steady partnerships which are concurrent to the infected casual partnerships.

In figure 3.11 we compare nontrivial solution branches for several values of λ_c. The impact of changing λ_c on the location of the critical point is not significant in comparison with considering the variation in $[I]_\infty$ in a vertical cross section for a fixed value of f.

There are two factors which could be allowing casual partnerships to impact the dynamics so much: their short lifespan and their concurrency. It is useful to distinguish the relative contribution of these factors. Chapter 4 deals in greater depth with the issue of concurrency, but here we present some analysis as well by considering specified regions of the parameter space. Such analysis also demonstrates the flexibility of this model. Figure 3.12 shows that for the case of no casual partnerships, $f = 1$, faster partnership dynamics still increase $[I]_\infty$. So the rapid turnover of casual partnership contributes to a larger final size and more rapid spread. However the concurrency of casual partnerships also has something to do with the increase in the final epidemic size, as shown in figure 3.13 where we set $\sigma_c = \sigma_m$, ensuring that casual partnerships turn over at the same rate as steady partnerships. Thus the only change in the network as f is decreased is the addition of more concurrency.

3.7 Computation of R_0

The calculation of R_0 for compartmental mean-field STD models is straightforward [26] but when we have the process of partnership formation and dissolution, or where

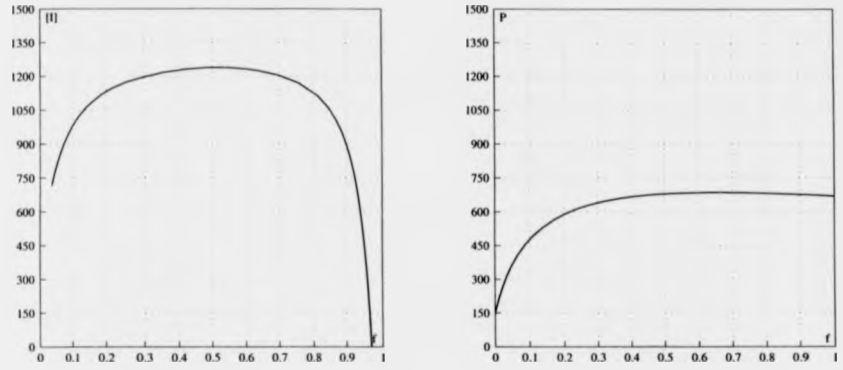


Figure 3.9: Bifurcation diagrams of $[I]$ versus f (left hand side) and $P = C + M$ versus f (right hand side), nontrivial stable branch only. Parameters are: $\rho = 0.001$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $\lambda_m = 0.14$, $\lambda_c = 0.14$, $\nu = 0.0015$, $N = 1500$.

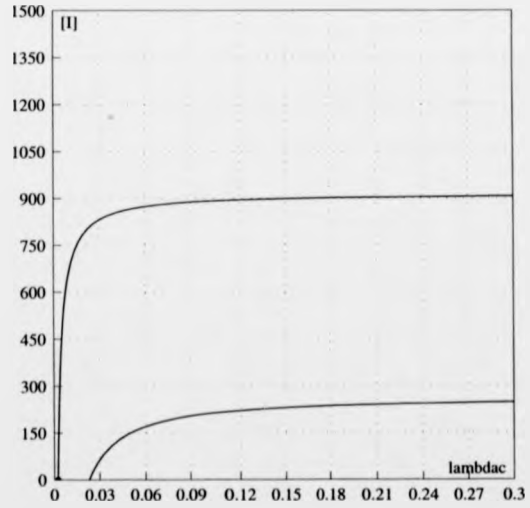


Figure 3.10: Bifurcation diagram of $[I]$ versus λ_c. Parameters are: $\rho = 0.001$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $\lambda_m = 0.14$, $f = 0.90$, $N = 1500$. For the upper branch, $\nu = 0.0015$ and for the lower branch, $\nu = 0.0028$.

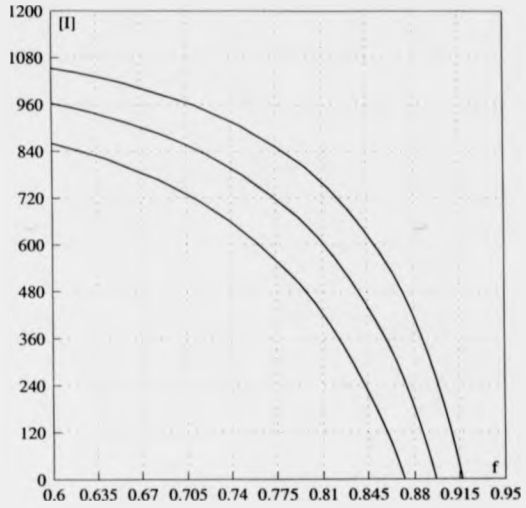


Figure 3.11: Bifurcation diagram of $[I]$ versus f for several values of λ_c. Parameters are: $\rho = 0.001$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $\lambda_m = 0.14$, $f = 0.90$, $\nu = 0.0028$, $N = 1500$, $\lambda_c = 0.01, 0.02, 0.14$ from bottom branch to top branch.

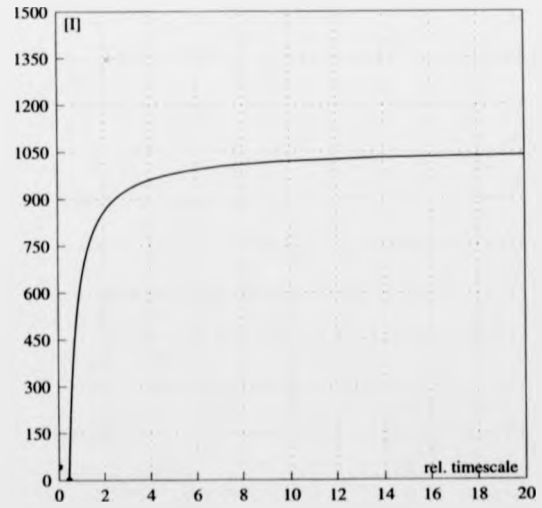


Figure 3.12: Bifurcation diagram of $[I]$ versus ρ for constant ratio $\sigma_s/\rho = 0.1$ and no casual partnerships. Parameters are: $\lambda_s = 0.15$, $\nu = 0.04$, $N = 1500$, $f = 1$.

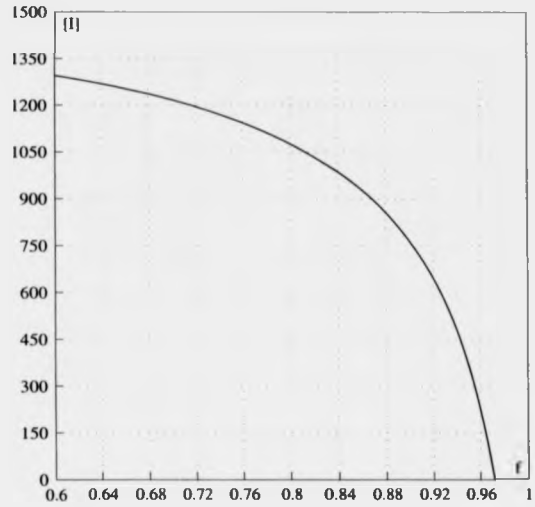


Figure 3.13: Bifurcation diagram of $[I]$ versus f for case where $\sigma_m = \sigma_c$. Parameters are: $\rho = 0.01$, $\sigma_m = \sigma_c = 0.005$, $\nu = 0.01$, $\lambda_m = \lambda_c = 0.14$, $N = 1500$.

there is concurrency, the analysis is more formidable.

However there are some methods for calculating R_0 for other types of models which might be applicable to the calculation of R_0 for the steady/casual model. One way is to consider the invasion process in spatial systems. It is the structure of the localized invading patches which determines whether or not the infection can actually invade, a fact not accounted for in mean-field derivations. Taking this spatial phenomenon into account is a more realistic way of determining R_0 and modelling invasion, as has been discussed in connection with application of IPA to the contact process (see chapter 2). In the spatial approach, we work from the definition of the basic reproductive ratio given in chapter 1 and take R_0 to be:

$$R_0 = \frac{\lambda_c}{\nu} Q^c(S \mid I)_{inv} + \frac{\lambda_m}{\nu} Q^m(S \mid I)_{inv} \tag{3.38}$$

where $Q^c(S \mid I)_{inv}$ and $Q^m(S \mid I)_{inv}$ denote the values of $Q^c(S \mid I)$ and $Q^m(S \mid I)$ at the establishment of local pseudo-equilibrium (i.e. $t \approx 50$ days in figure 3.8).

We would like to find simple expressions for these two quantities. One way is to write down the equations of motion for $Q^c(S \mid I)$ and $Q^m(S \mid I)$ and exploit the fact that $\frac{d}{dt} Q^c(S \mid I) \approx 0$ and $\frac{d}{dt} Q^m(S \mid I) \approx 0$ after establishment. Then we can calculate the equilibrium of those equations of motion under the simplifying assumption $[I] \ll N$ which applies during invasion [9][60][84]. This procedure was

implemented for the steady/casual model under OPA but even after simplification the expression for R_0 is too unwieldy. This is because of the complexity of the steady/casual model. However the approach is successfully demonstrated in chapter 4 for a different STD model.

Another method for calculating R_0 for pair formation models has been derived by Diekmann et al. [26][23][50]. Their method involves linear stability analysis of the trivial stable solution. This method is applied by Kretzschmar et al. to a steady/casual model without concurrency, where an individual can have either a steady or a casual partnership, but not both at the same time [66][67]. In their derivation there is no recovery but there are the demographic processes of recruitment (the addition of new susceptibles to the population) and removal. They derive the following expression for R_0:

$$R_0 = \frac{\rho X'}{\kappa}\left(\frac{f(\mu+\sigma_m)\lambda_m}{(2\mu+\sigma_m)(2\mu+\sigma_m+\lambda_m)} + \frac{(1-f)(\mu+\sigma_c)\lambda_c}{(2\mu+\sigma_c)(2\mu+\sigma_c+\lambda_c)}\right) \quad (3.39)$$

Here μ is the removal rate of individuals from the sexually active population, κ is the constant recruitment rate, X' is the number of individuals in neither a steady nor a casual partnership, and the other parameters are defined as in our steady/casual model. In the case where demographic processes are slow compared to infection processes, $o(\mu) \ll o(\sigma_m)$ and $o(\mu) \ll o(\sigma_m)$, equation (3.39) simplifies to:

$$R_0 = \frac{\rho\lambda_c}{\mu(\sigma_c+\lambda_c)} + \frac{1}{\mu}\rho\left(\frac{\lambda_m}{\sigma_m+\lambda_m} - \frac{\lambda_c}{\sigma_c+\lambda_c}\right) f \quad (3.40)$$

Since $\sigma_c > \sigma_m$, this tells us that R_0 decreases linearly in f for slow demographic changes without recovery.

Of course, the limitation of the linear stability approach is that it cannot be applied to network models with concurrent edges. Also, the complexity of the resulting expressions for R_0 can be considerable. For the steady/casual model without concurrency, incorporating the recovery process makes the expression too large to analyze easily.

Our approach is to consider the spatialized invasion process and use a very ad-hoc approach to calculate the pseudo-equilibrium quantities $Q^c(S \mid I)_{inv}$ and $Q^m(S \mid I)_{inv}$. We draw detailed balance diagrams depicting the possible states and the transition rates between them, and we calculate the steady-state to determine $Q^c(S \mid I)_{inv}$ and $Q^m(S \mid I)_{inv}$. This derivation is somewhat related to the pseudo-equilibrium pair approximation of section 3.5.

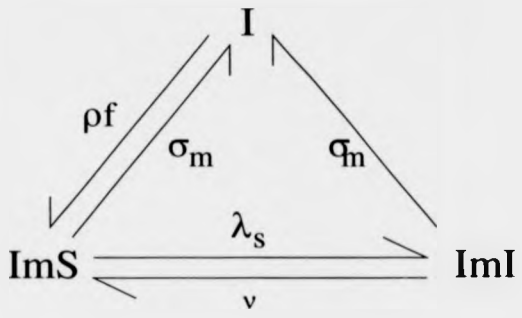


Figure 3.14: Detailed balance diagram for calculation of $Q^m(S \mid I)_{inv}$.

Calculation of $Q^m(S \mid I)_{inv}$

We want to know how many new cases an individual produces while infective. To estimate $Q^m(S \mid I)_{inv}$ we consider a *fixed individual* who remains infected for at least time $1/\sigma_c$. This individual may have a steady partnership with a *variable individual* who can be infected only by the fixed individual. We assume three possible states: let g_I be the expected proportion of time the fixed individual is single, and let g_{I_mI} (resp. g_{I_mS}) be the expected proportion of time the fixed individual is in steady partnership with an infected (resp. susceptible) individual. The transition rates between these states are determined by the model parameters. The detailed balance diagram for this is shown in figure 3.14. From the diagram we can calculate the steady-state and hence the quantities g_I, g_{I_mI} and g_{I_mS}. Because of the assumption of a fixed individual, we need the recovery rate ν to be sufficiently low to allow the configuration of figure 3.14 to come to equilibrium before time $1/\nu$ has passed. At equilibrium the the variables g_I, g_{I_mI} and g_{I_mS} satisfy the equations:

$$\begin{aligned} g_{I_mS}(\sigma_m + \lambda_m) &= g_{I_mI}(\nu) + g_I(\rho f) \\ g_{I_mI}(\sigma_m + \nu) &= g_{S_mI}\lambda_m \\ g_I(\rho f) &= g_{I_mS}(\sigma_m) + g_{I_mI}(\sigma_m) \\ 1 &= g_{I_mS} + g_{I_mI} + g_I \end{aligned}$$

Solving these for g_I, g_{I_mI} and g_{I_mS} and combining them gives us an expression for $Q^c(S \mid I)_{inv}$:

$$\begin{aligned} Q^m(S \mid I)_{inv} &= \frac{g_{S_mI}}{g_{S_mI} + g_{I_mI} + g_I} \\ &= \frac{1}{\frac{1}{1} + \frac{\lambda_m}{\sigma_m+\nu} + \frac{\sigma_m(\sigma_m+\nu+\lambda_m)}{\rho f(\sigma_m+\nu)}} \end{aligned}$$

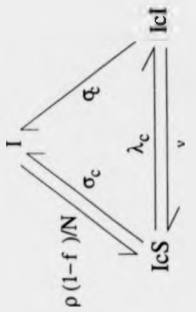


Figure 3.15: Detailed balance diagram for calculation of $Q^c(S \mid I)_{inv}$.

Calculation of $Q^c(S \mid I)_{inv}$

The derivation for $Q^c(S \mid I)_{inv}$ is similar to that of $Q^m(S \mid I)_{inv}$ except that we make a small adjustment to allow for concurrent casual partnerships. Consider the fixed individual and some other individual in the population. Let g_I be the expected proportion of time the fixed individual is not in partnership with the other individual. g_{I_cS} (resp. g_{I_cI}) be the expected proportion of time the fixed individual is in partnership with the other individual who is susceptible (resp. infected). Figure 3.15 shows the detailed balance diagram which describes this situation. The following equations can be derived from the diagram for the equilibrium:

$$\begin{aligned} g_{I_cS}(\sigma_c + \lambda_c) &= g_{I_cI}(\nu) + g_I\left(\frac{\rho(1-f)}{N}\right) \\ g_{I_cI}(\sigma_c + \nu) &= g_{I_cS}\lambda_c \\ g_I\left(\frac{\rho(1-f)}{N}\right) &= g_{I_cS}(\sigma_c) + g_{I_cI}(\sigma_c) \\ 1 &= g_{I_cS} + g_{I_cI} + g_I \end{aligned}$$

Leading to:

$$\begin{aligned} Q^c(S \mid I) &= N\frac{g_{I_cS}}{g_{I_cS} + g_{I_cI} + g_I} \\ &= \frac{N}{1 + \frac{\lambda_c}{\sigma_c+\nu} + \frac{\sigma_c(\sigma_c+\nu+\lambda_c)}{\frac{\rho(1-f)}{N}(\sigma_c+\nu)}} \end{aligned}$$

where the multiplicative factor of N comes from the fact that the fixed individual can form a casual partnership with any member of the population.

Combining these two derivations 3.41 and 3.41 for $Q^c(S \mid I)_{inv}$ and $Q^m(S \mid I)$,

we obtain the expression for R_0 based on the spatialized invasion process:

$$R_0 = \frac{1}{\nu}\left(\frac{\lambda_m}{1+\frac{\lambda_m}{\sigma_m+\nu}+\frac{\sigma_m(\sigma_m+\nu+\lambda_m)}{\rho f(\sigma_m+\nu)}}+\frac{N\lambda_c}{1+\frac{\lambda_c}{\sigma_c+\nu}+\frac{\sigma_c(\sigma_c+\nu+\lambda_c)}{\frac{\rho(1-f)}{N}(\sigma_c+\nu)}}\right) \tag{3.41}$$

The strengths of this method for deriving R_0 are that we can account for concurrent casual partnerships and that it is tractable for nonzero recovery rate $\nu > 0$, unlike the linear stability approach.

3.7.1 Analysis of R_0

We analyze equation (3.41) to see what the effects of casual partnerships are and whether or not our derivation gives sensible results.

Figure 3.16 shows plots of equation (3.41) against various model parameters. We have chosen a relatively slow recovery rate, $\nu = 0.001$, so that equation (3.41) should be valid and the assumption of a fixed individual not too unrealistic.

Equation 3.41 indicates that R_0 is linearly decreasing in f. Equation 3.40 for Kretzschmar's steady/casual model without concurrency also predicts that R_0 decreases linearly in f. We contrast this linear dependence with the nonlinear dependence of $[I]_\infty$ on f in figure 3.9. The linearity probably has to do with the linearity of the infection process during invasion. When $[I] \ll N$, the chance that two infecteds form a partnership is negligible. This has to do with the lack of clustering in the steady/casual model. Unlike square lattices, there is no clustering to interfere with the invasion process (notwithstanding the significant presence of higher-order correlations which still must be accounted for). This is because any individual can form a partnership with any other individual in the population. On the other hand, concurrency is still possible in the steady/casual network and this allows the disease to invade more easily, so R_0 increases as f decreases.

The other plots deserve some mention as well. The plot of R_0 versus σ_m exhibits a maximum. This is presumably caused by a balance of two competing effects. On the one hand we have a decreasing number of partnerships as we increase σ_m, and on the other hand we have an increased partner turnover (thus increasing the opportunities for infection spread) as σ_m increases. This maximum is not present in the plot of R_0 against σ_c but probably would be for a different choice of parameters. At any rate we cannot draw any epidemiological conclusions from the existence of this maximum since changing σ_m also changes the total number of partnerships. There is no sociological interpretation for this aspect of the model, i.e. there is no understanding of the relationship between changing partnership duration and

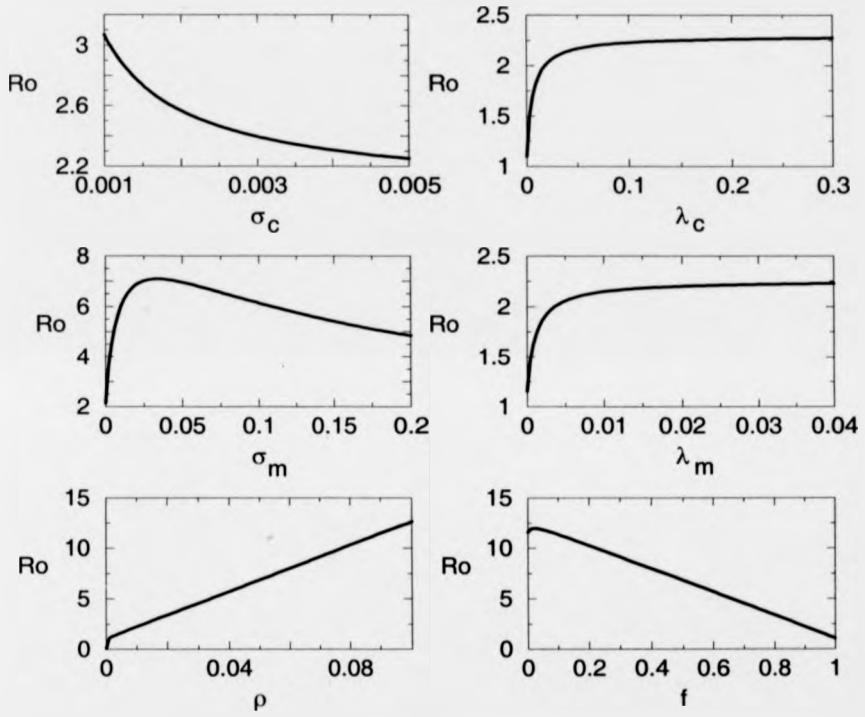


Figure 3.16: R_0 versus model parameters. Parameters are: $\rho = 0.01$, $\sigma_m = 0.00012$, $\sigma_c = 0.005$, $f = 0.90$, $\lambda_m = 0.15$, $\lambda_c = 0.15$, $\nu = 0.001$, $N = 1500$, except where they are the independent variables.

changing the total number of partnerships, or whether altering σ_m would reflect this accurately. Finally, the plots of R_0 versus λ_m and λ_c exhibit saturation as $\lambda_c \to \infty$ and $\lambda_m \to \infty$, with the plot for λ_m saturating more quickly for these parameter choices.

Finally, for the parameters of the simulation data of figure 3.2, equation 3.41 predicts $\nu_c = 0.00137$. This is obtained by setting $R_0 = 1$ and solving for $\nu = \nu_c$. This value is significantly lower than the (approximate) threshold value of $\nu_c = 0.0018$ of the simulation data.

In conclusion, although quantitatively the expression 3.41 is not very accurate, it at least produces sensible behaviour, and it tells us what happens to R_0 as f varies. So this derivation can supply many of the same answers of more laborious derivations of R_0. However it is very heuristic.

3.8 Discussion

In this chapter we defined a stochastic model for disease transmission through a network of steady and casual partnerships. We derived several moment closure approximations and after a discussion of their relative merits, carried out an analysis of the model using an ordinary triple approximation. We also derived R_0 by exploiting the characteristics of invasion in spatial models discussed in chapter 2. The derivation for R_0 is ad-hoc but much simpler than previous derivations for network STD models and pair formation STD models.

We saw that pair approximations are not sufficient for the steady/casual model, and we hypothesized that this is due to the nature of the network structure in the steady/casual model (particularly the low node degree Q). The data available suggest that for the steady/casual model, only triple approximations can capture the strong higher-order correlations present at epidemiologically realistic parameter values. However triple approximations are complicated and there is still a need for a pair approximation with a mechanistic foundation based on insights into dynamics of higher-order correlations on networks. HPA is one such example and works well for square lattices, but not for the steady/casual model.

We also explored the nature of invasions in dynamic networks and found some interesting differences between dynamic network models and static lattice models. In the steady/casual network, one still sees the establishment of a pseudo-equilibrium in the local structure of invading patches, and these patches slowly spread through the rest of the network. However in the case of the steady/casual network, clustering does not interfere in epidemic spread as it does in square lattices. This is because

of the lack of triangles in the steady/casual network (triangles, as discussed before, are network structures where three nodes are mutually connected by three edges). However we note that in a real sexual network one would find such structures, for instance in core groups.

This work shows that the impact of a sexually transmitted disease can depend strongly on a small number of concurrent casual partnerships in a predominantly steady population with a slow turnover of steady partnerships. The impact is both in terms of the rapidity of disease spread and the final size of the epidemic. As the number of casual partnerships increases (i.e. as the parameter f decreases) a threshold is passed beyond which $[I]_\infty$ increases quickly. Above the threshold (i.e. for fewer casual partnership) we have $R_0 < 1$ and $[I]_\infty = 0$. This dependence on the parameter f is caused both by the more rapid turnover of casual partnerships and by the fact that they are concurrent to steady partnerships and to one another. Another interesting result is that decreasing the transmissibility of the disease in casual partnerships, for instance through increased condom usage, is much more effective near the critical point. The OTA also seems to reproduce the early time evolution of the local network structure, something not possible with a mean-field model. Finally, this chapter further demonstrates the versatility of moment closure models. We have obtained a deterministic approximation which, in limiting cases, produces several deterministic models for STDs similar to those previously studied [2][67][92]. Although we have only been able to study a small part of the parameter space in figure 3.2, we have restircted attention to parameters which are biologically reasonable. Unfortunately, as noted beforehand, with moment closure approximations the only way to ascertain their accuracy is by comparison with results from simulations of the stochastic model. However the evidence presented in this chapter suggests that the OTA provides a more accurate approximation to the stochastic steady/casual model than the OPA can.

In the next chapter we will look at another STD pair approximation, so that we may consider in greater depth the issue of concurrency and also to contrast the MCA technique for a different network model to the one we have studied in this chapter.

Chapter 4

STD Dynamics in a Concurrent Partnership Network

The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" but "That's funny"...

–Isaac Asimov

4.1 Introduction

In the previous chapter we developed a moment closure approximation for STDs spreading through a steady/casual sexual partnership network. One of the interesting aspects of the model was the concurrency of the casual partnerships to one another and to the steady partnerships. We noted that this impacts the STD dynamics. This chapter develops further an existing model for concurrency in sexual partnership networks by deriving a pair approximation and determining R_0 using the same method for spatialized invasions that was used in the steady/casual model. Although there are no major mathematical differences between deriving MCA for the steady/casual model and the concurrency model, in this chapter we consider in greater depth the important issue of concurrency in sexual networks. Also we find that which moment closure approximation is best for the concurrency model depends on parameter choices. As with the steady/casual model, the epidemiology is of the SIS type, and we do not account for demographic processes, disease-induced mortality or any disease-induced modification of the underlying partnership network structure.

Intuitively, one might expect concurrency to increase the severity of a disease because, under serial monogamy, an infectious agent must wait for the current partnership to break before spreading the infection to a new partner [45] (see figure 4.1). This is not surprising if higher concurrency is associated with a general increase in the average number Q of partners per person. It is less clear how concurrency changes the invasion process if it is increased without increasing Q. In other words, we ask how the distribution of some constant number P of partnerships across the population affects the disease spread. When the issue is understood in this way it is less clear *a priori* whether concurrency will have a large effect, and what kind of effect that will be [68].

Some models for concurrency have already been developed and studied. Dietz et al. [28] derive a deterministic concurrency model by extending conventional pair models to allow for the case where someone who is paired gains a further single partner, thus creating triples. They conclude that the introduction of concurrency in this limited way does not have a large effect on the epidemic for the parameters they chose. Watts and May [92] construct a deterministic model with random mixing to understand the effect of concurrency on HIV transmission. In random mixing the number of partners an individual currently has does not influence the probability of gaining further partners. They find that concurrency can bring about complex dynamical patterns for HIV, including the existence of two time scales for the spread

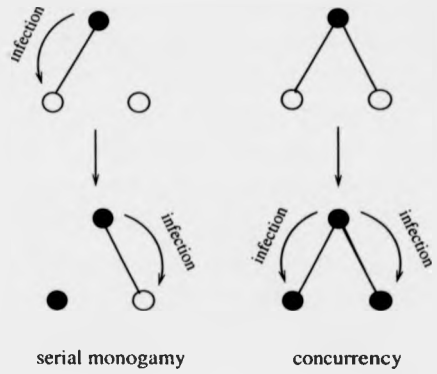


Figure 4.1: Diagram describing the effect of concurrency on STD epidemiology.

of the epidemic through a susceptible population: the fast time scale corresponds to the spread of infection through the existing set of partnerships and the slower time scale corresponds to spread due to the formation of new partnerships with susceptibles. Altmann has also considered an STD model with independent concurrent partnerships, similar to the model of Watts and May [2]. In his model infection can be transmitted during a partnership and also when the partnership is initiated. Altmann later formulates a model which, like our concurrency and steady/casual partnership models, can account for a range of possible sexual behaviours [3]. His simple formulation amounts to an ordinary pair approximation for STDs (he implicitly assumes conditional independence in partner disease states). An advantage of this model is that it is very general in its formulation. However there is no discussion of the accuracy of this model and the role of higher-order correlations, and we know from the previous chapter that the assumption of conditional independence, particularly when Q is low, can render the conventional OPA highly inaccurate. Therefore for deterministic approximations to STD network models it is important to apply the body of results which have been put forward since Katori [56], and to consider carefully which moment closure is best, if our approximations are to be useful.

Kretzschmar, Morris, Chick and others have analyzed stochastic models for concurrency [16][78]. Kretzschmar et al. have also suggested a concurrency measure κ_3 for networks [68]. They find concurrency increases the growth rate of the epidemic, and conclude that this is due to the growth in the size of the largest connected components as the concurrency is increased. One strength of the model of Kretzschmar and Morris is that it allows for a range of concurrent behaviour between the extremes of serial monogamy and random mixing. We consider a model like that of

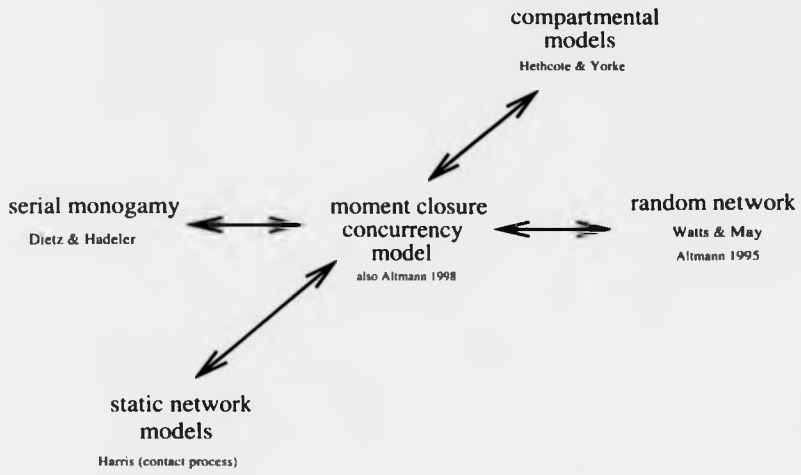


Figure 4.2: Placement of the pair approximation to the concurrency model with respect to the current menagerie of epidemic and STD models.

Kretzschmar and Morris but develop a pair approximation for it.

Few other deterministic STD models incorporate the wide range of network structures which our moment closure approximations for the concurrency and steady/casual models can, and this makes them useful tools for studying STDs on networks. Figure 4.2 summarizes the relationship of the concurrency pair approximation to other deterministic models developed thus far for STDs and more generally epidemics on networks. The moment closure model is capable of realizing, in four different limits, a mean-field compartmental homogeneous mixing model, a static network model, a pair formation model with serial monogamy and a model with complete partnership independence (the random network). Static network models such as the contact process have not to our knowledge been used specifically for STD modelling.

4.2 Description of Concurrency Model

We begin with a population of N individuals who form and break-off partnerships through which an STD is transmitted. The partnership dynamics are determined by the rate ρ/N at which any two singles form a partnership, the rate $\rho\theta/N$ at which any two individuals, at least one of whom is in a partnership, enter into a further partnership ($0 \leq \theta \leq 1$, θ controls concurrency) and the rate σ of partnership breakup. An infection is transmitted between an infected and a susceptible at rate λ, and infecteds recover at rate ν. Because we are interested in a partnership network

$[X_S]$	number of single susceptible individuals
$[X_I]$	number of single infected individuals
$[SI]$	number of infected-susceptible partnerships
$[II]$	twice the number of infected-infected partnerships
$[SS]$	twice the number of susceptible-susceptible partnerships
$[I]$	number of infected individuals
$[S]$	number of susceptible individuals

Table 4.1: The dynamical variables of the concurrency model.

ρ/N	rate at which any two singles form a partnership
$\rho\theta/N$	rate at which any two singles, at least one of whom is in a partnership, enter into a partnership ($0 \leq \theta \leq 1$). θ controls concurrency.
σ	partnership separation rate
λ	disease transmission rate in an SI partnership
ν	recovery rate
N	population size

Table 4.2: The parameters of the model. $\theta = 0$ corresponds to monogamy while $\theta = 1$ corresponds to independence of partnerships from one another.

and not some explicit spatial network, the relevant aspect of the network structure is which nodes are linked to which other nodes; the length of the links and their relative position are not considered. As with the steady/casual model, we allow the partnership dynamics (dictated by ρ, θ and σ) to equilibrate before the infection is inoculated into the population. Then we study the ensuing epidemic.

Tables 4.1 and 4.2 summarize the dynamical variables and static parameters for the model we have described. Note that the variables $[II]$ (resp. $[SS]$) are determined by counting each II (resp. SS) partnership twice, in keeping with previous conventions. Additionally, $[IS] = [SI]$.

Because we consider the partnership dynamics at equilibrium, the following constraints apply to the variables:

$$\begin{aligned} 2P &= 2[SI] + [II] + [SS] \\ X &= [X_S] + [X_I] \end{aligned} \tag{4.1}$$

where X is the equilibrium number of singles and P is the equilibrium number of partnerships. These are calculated in the next subsection.

4.2.1 Partnership Network Structure and the Index of Concurrency κ_3

We derive here the equilibrium structure of the partnership network and obtain exact expressions for P and X. We use the *detailed balance* method to calculate the proportion π_k of the population with k partners, and hence determine P and X.

Let $F_{m,n}$ be the rate at which individuals move from a state of having m partners to a state of having n partners. In equilibrium, $\pi_k F_{k,k+1} = \pi_{k+1} F_{k+1,k}$. Because of the appearance of singularities at $\theta = 0$ and $\theta = 1$ in the solution of π_k for $0 < \theta < 1$, it is convenient to calculate the equilibrium separately for the three cases $\theta = 0$ (serial monogamy), $\theta = 1$ (randomly distributed network), and $0 < \theta < 1$ (intermediate case). We can write down detailed balance diagrams for these three cases and derive the π_k distributions, applying the condition that $\sum_{k\geq 0} \pi_k = 1$ to get π_0:

$$\pi_0(\theta) = \begin{cases} (-1+\sqrt{1+4\eta})/2\eta & \theta = 0 \\ \frac{-\beta+\sqrt{\beta^2+4(\beta-1)(\theta^{-1}-1)}}{2(\beta-1)(\theta^{-1}-1)} & 0 < \theta < 1 \\ e^{-\eta} & \theta = 1 \end{cases} \tag{4.2}$$

where $\eta = \rho/\sigma$ and $\beta = \exp(\eta\theta)$. And for π_k, $k \geq 1$ we get:

$$\pi_k(\theta) = \begin{cases} 0 & \theta = 0, k \neq 1 \\ \frac{\pi_0(\theta)}{k!}(\eta\theta)^k \frac{(\pi_0(\theta)+\theta(1-\pi_0(\theta)))}{\theta} & 0 < \theta < 1 \\ \eta^k e^{-\eta}/k! & \theta = 1 \end{cases} \tag{4.3}$$

and finally for $\theta = 0$, $k = 1$ we have $\pi_1(0) = 1 - \pi_0(0)$. We can use the π_k distribution to obtain these expressions for X and P, valid for all θ:

$$\begin{aligned} X &= N\pi_0(\theta) \\ P &= \frac{N}{2}\sum_{k\geq 1} k\pi_k(\theta) = \frac{\rho(\theta N^2 + (1-\theta)X^2)}{2\sigma N} \end{aligned} \tag{4.4}$$

Kretzschmar and Morris' index of concurrency κ_3 [68] is the number of concurrent partnerships divided by the total number of partnerships, i.e. the number of triples divided by the number of pairs. It can be shown from the π_k distribution of equation (4.3) that

$$\kappa_3 = \frac{\rho\theta}{\sigma} = \eta\theta \tag{4.5}$$

for $\theta \geq 0$.

4.3 Derivation of Master Equations

We employ the usual approach for deriving the master equations, following the example of chapters 2 and 3 to derive the equation of motion for $[SI]$. The network events which affect $[SI]$ are:

$SS \rightarrow$	SS	infection
$SI \rightarrow$	II	infection
$SI \rightarrow$	SS	recovery
$II \rightarrow$	SI	recovery
$S + I \rightarrow$	SI	partnership formation
$SI \rightarrow$	$S + I$	partnership dissolution

We now derive the master equation for $[SI]$, using equation (1.12) and taking into account those six events which can possibly affect $[SI]$. The notation is similar to that of the previous chapter:

$$\begin{aligned}
[\dot{SI}] \;=\; & \sum_{\varsigma_{xy}=(S,X;I,X)} \frac{\rho}{N} + \sum_{\substack{\varsigma_{xy}=(S,P;I,P)\cup \\ (S,P;I,X)\cup(S,X;I,P)}} \frac{\rho\theta}{N} \\
& - \sum_{\varsigma_{xy}=(SI)} \sigma - \sum_{\varsigma_{xy}=(SI)} \nu + \sum_{\varsigma_{xy}=(II)} \nu \\
& + \sum_{\varsigma_{xy}=(SS)} \lambda Q_x(I) - \sum_{\varsigma_{xy}=(SI)} \lambda Q_x(I)
\end{aligned}$$

Next we apply the substitutions $Q_x(I) = Q(I \mid SI) + \eta_x(I \mid SI)$ for pairings (x, y) with $\varsigma_{xy} = (SI)$ and $Q_x(I) = Q(I \mid SS) + \eta_x(I \mid SS)$ for pairings (x, y) with $\varsigma_{xy} = (SS)$. We also rearrange the second sum for concurrent partnership formation to get:

$$\begin{aligned}
[\dot{SI}] \;=\; & \sum_{\varsigma_{xy}=(S,X;I,X)} \frac{\rho}{N} + \left\{ \sum_{\varsigma_{xy}=(S;I)} - \sum_{\varsigma_{xy}=(S,X;I,X)} \right\} \frac{\rho\theta}{N} \\
& - \sum_{\varsigma_{xy}=(SI)} \sigma - \sum_{\varsigma_{xy}=(SI)} \nu + \sum_{\varsigma_{xy}=(II)} \nu \\
& + \sum_{\varsigma_{xy}=(SS)} \lambda\{Q(I \mid SS) + \eta_x(I \mid SS)\} \\
& - \sum_{\varsigma_{xy}=(SI)} \lambda\{Q(I \mid SI) + \eta_x(I \mid SI)\}
\end{aligned}$$

Taking constants out of the sums and evaluating the sums gives us:

$$\begin{aligned}
[\dot{SI}] \;=\; & \frac{\rho}{N}[X_S][X_I] + \frac{\rho\theta}{N}\{(N - [I])[I] - [X_S][X_I]\} - \sigma[SI] - \nu[SI] \\
& + \nu[II] + \lambda[SS]Q(I \mid SS) - \lambda[SI]Q(I \mid SI) \qquad (4.6)
\end{aligned}$$

Thus we have derived the master equations and are left with the task of approximating the third-order correlations $Q(I \mid SI)$ and $Q(I \mid SS)$, The master equations for $[II]$ and $[SS]$ are derived similarly:

$$\begin{aligned}
\dot{[II]} &= \frac{\rho}{N}[X_I]^2 + \frac{\rho\theta}{N}([I]^2 - [X_I]^2) - \sigma[II] \\
&\quad -2\nu[II] + 2\lambda[SI]Q(I \mid SI) \qquad (4.7) \\
\dot{[SS]} &= \frac{\rho}{N}[X_S]^2 + \frac{\rho\theta}{N}\{(N - [I])^2 - [X_S]^2\} - \sigma[SS] \\
&\quad +2\nu[SI] - 2\lambda[SS]Q(I \mid SS) \qquad (4.8)
\end{aligned}$$

For $[X_I]$ the derivation is very similar, except for one term which requires us to estimate the number of monogamous individuals. In contrast to the steady/casual model it is not possible to get an identity for this:

$$\begin{aligned}
\frac{d}{dt}[X_I] &= -\sum_{\varsigma_{xy}=(I,X;X)} \frac{\rho}{N} - \sum_{\varsigma_{xy}=(I,X;P)} \frac{\rho\theta}{N} - \sum_{\varsigma_x=X_I} \nu + \\
&\quad \sum_{\varsigma_{xy}=(SI)} \sigma \times \left\{ \begin{array}{ll} 0 & \text{if y is not monogamous} \\ 1 & \text{if y is monogamous} \end{array} \right\} + \\
&\quad \sum_{\varsigma_{xy}=(II)} \sigma \times \left\{ \begin{array}{ll} 0 & \text{if y is not monogamous} \\ 1 & \text{if y is monogamous} \end{array} \right\} \\
&= -\frac{\rho}{N}[X_I]([X_I] + [X_S]) - \frac{\rho\theta}{N}[X_I](N - X) - r[X_I] + \\
&\quad \sigma \times (\text{number of monogamous infected individuals})
\end{aligned}$$

We approximate the number of monogamous infected individuals as:

$$([II] + [SI]) \times \left(\begin{array}{c} \text{probability I is monogamous} \\ \text{given I is in a relationship} \end{array} \right) \approx ([II] + [SI])\frac{\pi_1}{1 - \pi_0} \qquad (4.9)$$

The further approximation which appears in equation (4.9) is there because the variance in node degree in the network guarantees that knowledge of an individual's disease status and the disease status of his partners influences his expected node degree. For example, infected individuals are more likely to have a higher node degree.

So using approximation (4.9) produces the equation of motion:

$$\begin{aligned}
\frac{d}{dt}[X_I] &= -\frac{\rho}{N}[X_I](X) - \frac{\rho\theta}{N}[X_I](N - X) - \nu[X_I] \\
&\quad +\sigma([SI] + [II])\frac{\pi_1}{1 - \pi_0} \qquad (4.10)
\end{aligned}$$

Similarly for $[X_S]$:

$$\frac{d}{dt}[X_S] = -\frac{\rho}{N}[X_S](X) - \frac{\rho\theta}{N}[X_S](N - X) + \nu[X_I] + \sigma[SI]\frac{\pi_1}{1 - \pi_0} \qquad (4.11)$$

Finally for $[I]$:

$$\begin{aligned}[\dot{I}] &= \sum_{\varsigma_x=(I)} \nu + \sum_{\varsigma_{xy}=(SI)} \lambda \\ &= -\nu[I] + \lambda[SI] \end{aligned} \tag{4.12}$$

4.4 Moment Closures

We now turn to the task of closing the master equations (4.6) - (4.12). The only two higher order quantities we need to approximate are $Q(I \mid SI)$ and $Q(I \mid SS)$. The fact that one has only to approximate these two is general for epidemiological systems and means that the analytical treatment of such systems is not as difficult as is usually believed. We will present three moment closures: an OPA, an IVPA and a new pair approximation called simply *heuristic pair approximation* (HEPA).

Ordinary Pair Approximation

Working from the assumption of conditional independence we formulate the OPA for this system:

$$\begin{aligned} Q(I \mid SI) &= 1 + \frac{Q_{\geq 1} - 1}{Q} Q(I \mid S) \\ Q(I \mid SS) &= \frac{Q_{\geq 1}}{Q} Q(I \mid S) \end{aligned} \tag{4.13}$$

where Q is the average number of partners per person, a function of ρ, θ and σ, and where $Q_{\geq 1}$ is the expected number of partners of someone who has at least one partner. These expressions are derived from the π_k distribution, equation (4.3):

$$\begin{aligned} Q &= \frac{2P}{N} \\ &= \frac{1}{\sigma}\rho(\theta + (1-\theta)(X/N)^2) \\ Q_{\geq 1} &= \frac{2P}{N(1-\pi_0)} \\ &= \frac{\rho(\theta + (1-\theta)(X/N)^2)}{\sigma(1-\pi_0)} \end{aligned} \tag{4.14}$$

Invasory Pair Approximation

We also develop and test an IVPA for the concurrency model:

$$\begin{aligned} Q(I \mid SI) &= 1 + \frac{Q_{\geq 1} - 1}{Q} \frac{[SI]}{[S]_{inv}} \\ Q(I \mid SS) &= \frac{Q_{\geq 1} - 1}{Q} \frac{[SI]}{[S]} \end{aligned} \tag{4.15}$$

For $[S]_{inv}$ we take:

$$[S]_{inv} = [SI] - [SI](Q(I \mid SI) - 1) \tag{4.16}$$

where the negative $[SI](Q(I \mid SI) - 1)$ term helps compensate for multiple-counting of susceptibles who have more than one infected neighbour. Solving equations (4.15) and (4.16) for $[S]_{inv}$ produces:

$$[S]_{inv} = \frac{[SI]}{2}\left[1 + \left(1 - \frac{4Q_{\geq 1}}{Q} + \frac{4}{Q}\right)^{1/2}\right] \tag{4.17}$$

Heuristic Pair Approximation

Finally we develop what we call the heuristic pair approximation. To derive this we start from what we already know about the third-order correlations. Firstly we note that $Q(I \mid SI) > 1$ because the susceptible has at least one infected partner. We must also determine how many other infected partners the susceptible individual has on average. The number of extra (infected or uninfected) partners given that the individual has at least one partner is $Q_{\geq 1} - 1$. Also, the fraction of extra partners of the susceptible in an SI pair who are infected is $(Q(I \mid SI) - 1)/(Q(I \mid SI) - 1 + Q(S \mid SI))$. Combining this with the expression for the total number of extra partnerships $Q_{\geq 1} - 1$ we get:

$$Q(I \mid SI) \approx 1 + (Q_{\geq 1} - 1) \times \left\{\frac{Q(I \mid SI) - 1}{Q(I \mid SI) - 1 + Q(S \mid SI)}\right\} \tag{4.18}$$

This assumes no correlation between the total number of partners of an individual and the distribution of their disease status. We note that this is a slightly weaker assumption than the assumption of conditional independence used for the OPA. We rearrange $Q(S \mid SI)$ to get $Q(S \mid SI) = [SSI]/[SI] = Q(I \mid SS)[SS]/[SI]$. And for $Q(I \mid SS)$ we use the OPA calculation:

$$Q(I \mid SS) \approx \frac{Q_{\geq 1} - 1}{Q} Q(I \mid S) \tag{4.19}$$

Solving equation 4.18 for $Q(I \mid SI)$ and using equation 4.19 gives us the moment closure for the HEPA:

$$\begin{aligned} Q(I \mid SI) &= \frac{[SS] + (Q[S] - [SS])Q_{\geq 1}}{Q[S]} \\ Q(I \mid SS) &= \frac{Q_{\geq 1} - 1}{Q}\frac{[SI]}{[S]} \end{aligned} \tag{4.20}$$

Although equation 4.20 looks a bit mysterious, we can check partially if it has the behaviour it should. Firstly define Q_S (resp. Q_I) as the average number of partners

per susceptible (resp. infected). We expect that $Q_S < Q$ and $Q_I > Q$ since having more partners means one is more likely to be infected. It easily follows from this and the equality $[SI] + [SS] = Q_S[S]$ that $[SI] + [SS] < Q[S]$. Applying this inequality to equation (4.20) and doing some manipulation confirms that $Q(I \mid SI) > 1$. Also, it can be checked that in the case of serial monogamy, $Q(I \mid SI) = 1$.

Experience from the steady/casual model of chapter 3 tells us that a comparison with results from the stochastic model is in order, since we know that pair approximations tend to fail at low Q and since the characteristics of the dynamic and irregular network of the concurrency model are similar to those of the steady/casual model. More generally, comparison allows us to refine the deterministic model, understand the stochastic model better, and determine the error introduced by the moment closures. Therefore we compare the stochastic model and the three pair approximations in the next section.

4.5 Comparison of Stochastic Model and Pair Approximations

A Monte-Carlo simulation for the underlying stochastic model was run in order to compare results with the deterministic model. There are twelve cases, the parameters of which are shown in table 4.3 (in units of day^{-1}). For each case, $\rho = 0.01$, $\sigma = 0.005$ and $N = 1500$. Table 4.4 compares $[I]_\infty$ and $[SI]_\infty$ of the pair approximations and the long-time average of the stochastic model (the averaging starts after partnership and infection dynamics have equilibrated). There is only one stochastic run for each set of parameters, but the inoculation $[I]_0$ is large. The cases were chosen in order to compare the agreement as a function of concurrency, of *relative time scale* ν/σ of infection dynamics to partnership dynamics, and of λ/ν, ratio of the rate of transmission to the rate of recovery (giving us an idea of how likely one partner is to be infected given that the other partner is infected).

Out of the three pair approximations, the IVPA provides the best agreement with the stochastic model in cases 1 - 6 (with a slight tendency to overpredict) and fair agreement in cases 7 - 12 (tending to underpredict). From looking at the values for relative time scale ν/σ we can conclude that the IVPA performs best when infection dynamics are fast relative to partnership dynamics. This is usually the case in real sexual networks unless one considers a high density of short-lived casual partnerships. The OPA and HEPA do not differ significantly from one another but the HEPA performs slightly better, as expected on grounds of the weaker assumptions made

case	θ	λ	ν	κ_3	ν/σ	case	θ	λ	ν	κ_3	ν/σ
1	0.08	0.10	0.006	0.16	1.2	7	0.08	0.05	0.003	0.16	0.6
2	0.1	0.10	0.006	0.2	1.2	8	0.1	0.05	0.003	0.2	0.6
3	0.15	0.10	0.006	0.3	1.2	9	0.15	0.05	0.003	0.3	0.6
4	0.1	0.20	0.012	0.2	2.4	10	0.1	0.08	0.003	0.2	0.6
5	0.1	0.08	0.006	0.2	1.2	11	0.1	0.04	0.003	0.2	0.6
6	0.1	0.15	0.006	0.2	1.2	12	0.1	0.025	0.0015	0.2	0.3

Table 4.3: Parameter values for comparison of stochastic model and moment closure approximations.

measure	case	stoch	HEPA	OPA	IVPA	case	stoch	HEPA	OPA	IVPA
$[I]$	1	0	284	291	158	7	674	750	768	648
$[SI]$		0	17.1	17.5	9.5		40.4	45.0	46.1	38.8
$[I]$	2	275	469	487	296	8	748	811	840	680
$[SI]$		16.4	28.1	29.2	17.7		44.7	58.7	50.4	40.8
$[I]$	3	542	718	773	422	9	880	911	965	678
$[SI]$		32.3	43.1	46.4	25.3		52.8	54.6	57.9	40.7
$[I]$	4	0	111	116	0	10	865	920	963	783
$[SI]$		0	6.64	6.9	0		32.3	34.5	36.1	29.4
$[I]$	5	172	342	353	183	11	663	736	757	612
$[SI]$		12.7	25.6	26.4	13.7		49.5	55.2	56.8	45.9
$[I]$	6	378	627	660	442	12	1036	1002	1032	906
$[SI]$		14.9	25.1	26.4	17.7		61.8	60.1	61.9	54.3

Table 4.4: Comparison of stochastic and moment closure approximations at equilibrium. A zero entry in the columns for the simulation results means that the epidemic has died out.

in its derivation. However both are not very good in comparison with the IVPA in general, although their agreement is sufficiently good in cases 7 - 12 for faster partnership dynamics, and in some cases HEPA and OPA do better than IVPA.

In cases 1,2,3 and cases 7,8,9, the agreement of deterministic and stochastic models as a function of κ_3 is considered. Inspection of these cases shows that the agreement between stochastic models and pair approximations increases somewhat with increasing concurrency for HEPA and OPA, but not for IVPA.

In cases 2,4,8 and cases 5,11,12 the agreement as a function of the relative time scale ν/σ is considered. One can see that the agreement is better for lower relative time scales ν/σ for HEPA and OPA, but that the reverse is true for the IVPA. In HEPA and OPA we assumed conditional independence, in some form, of the disease status of an individual's several partners. This obviously gives a better approximation when the partnership network does not change too slowly, and so there is a greater agreement for lower relative time scales is expected for HEPA and OPA. The derivation of IVPA, on the other hand, depends on the existence of active, correlated patches of infecteds in the partnership network; this spatial structuring is lessened when the partnership network changes too quickly and so the agreement for the IVPA worsens as the partnership network changes more quickly.

In cases 2,5,6 and cases 8,10,11 we look at the agreement as a function of λ/ν. For all pair approximations, the agreement is better for higher infectivity, as is usually the case for pair approximations further away from the critical point.

We can see that the IVPA is best when the partnership dynamics do not change too quickly. We usually consider this case when we analyze the model to answer questions about STD epidemiology, and so we will use mostly the IVPA in our analysis. When partnership dynamics are too fast the network becomes unimportant, and mean-field compartmental models can be used instead. Even for cases 7-12 when the partnership dynamics are faster, the IVPA produces agreement which is as good as HEPA and OPA, although it tends to underpredict instead of overpredict.

From these observations and the results from the steady/casual model it is clear that one must choose carefully which pair approximation to use when dynamic networks are considered, in fact more so than with static regular lattices. The data for these parameter values suggests that we do not have to go to third-order correlations to get an accurate moment closure model, unlike the case of the steady/casual model where we had to go to a triple approximation to get good accuracy.

4.6 Analysis of Equations of Motion under IVPA

Bifurcation diagrams of $[I]$ as a function of parameters θ and index of concurrency κ_3 were obtained for the IVPA. In considering the variation in endemicity with the parameter θ, we need to separate the effects of increasing concurrency *per se* from an increase accompanied by a larger equilibrium number of partnerships P, as discussed for the steady/casual model in chapter 3. There we did not hold P constant, but for this model, P is held at some constant value P_0 as concurrency is increased; for given values of N and P there is a unique value of ρ/σ such that $P(\rho, \sigma, \theta, N) = P_0$ for all θ on the interval $[0, 1]$. However this value can only be determined numerically. Also, it is clear that if P is held at the constant value P_0 as θ is increased, then π_0 must be monotone increasing in θ, since the proportion of individuals with multiple partners goes up as θ increases. Because of the monotonicity of π_0 in θ, $\pi_0(1)$ and $\pi_0(0)$ are upper and lower bounds respectively on $\pi_0(\theta)$.

Figure 4.3, a bifurcation diagram, shows the most typical behaviour (i.e. for parameters within a biologically realistic domain) of the $[I]$ versus κ_3. As concurrency is increased, a transcritical bifurcation occurs. Beyond this threshold the endemicity increases rapidly with increasing concurrency. Figure 4.4 shows that as ν is decreased, the critical point moves closer to $\kappa_3 = 0$. The shape of the nontrivial solution branch is mostly unchanged as ν varies, except that the slope at the bifurcation increases, and there appears to be a decrease in $[I]_\infty$ with increasing κ_3 for large κ_3 and small ν. This could be due to inaccuracy in the IVPA since we are dealing with low relative time scale ν/σ in this case. However the same kind of decrease occurs for cases where $\nu/\sigma \approx 1.5$ (figure 4.5) which should be within the regime of accuracy of IVPA. Also we note that HEPA also exhibits a decrease in $[I]_\infty$ in for some cases of low ν (figure 4.6). Therefore this decrease could be real.

We suggest this decrease is caused by clustering. As concurrency increases, some parts of the network have higher average node degree and some have lower average node degree; the infection is restricted to areas of the network with higher average node degree where $R_0 > 1$ locally. This effect, in the cases of high endemicity, is strong enough to cause a decrease in endemicity overall. This phenomenon is basically what happens in core groups. These data illustrate nicely this phenomenon and reinforce the theory behind core groups, but unfortunately do not tell us anything new about control strategies for STDs.

Figure 4.7 shows three examples for three different values of ν where the partnership dynamics are very fast ($\rho/\nu \approx 2000$). We used the HEPA in these examples since ν/σ is small. In all three examples, $[I]_\infty$ is relatively constant as concur-

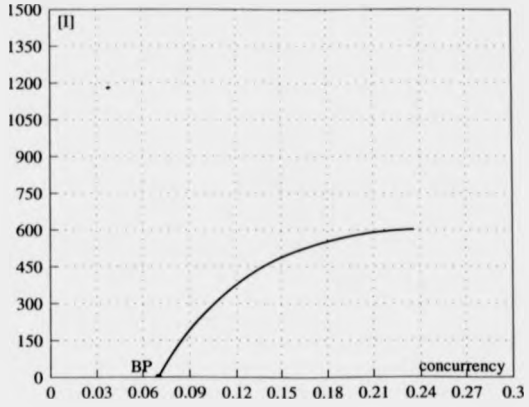


Figure 4.3: Bifurcation diagram for $[I]$ versus κ_3. The solid line indicates the stable nontrivial branch. 'BP' indicates the branching point where the trivial branch exchanges stability with the nontrivial branch. The trivial branch to the left of the branching point is stable, and to the right is unstable. $\rho = 0.01$, $P_0 = 440$, $\lambda = 0.15$, $\nu = 0.005$, $N = 1500$.

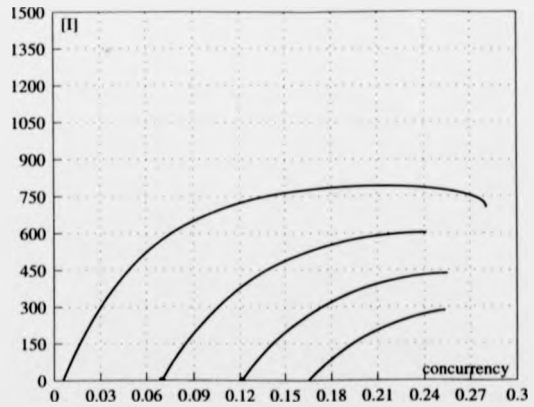


Figure 4.4: Bifurcation diagram for $[I]$ versus κ_3 showing the evolution of the nontrivial branch as ν is varied. From top branch to bottom branch the recovery rate is $\nu = 0.0035, 0.005, 0.0065, 0.008$. Other parameters are $\rho = 0.01$, $P_0 = 440$, $\lambda = 0.15$, $N = 1500$. The continuation ends where the step size became too small.

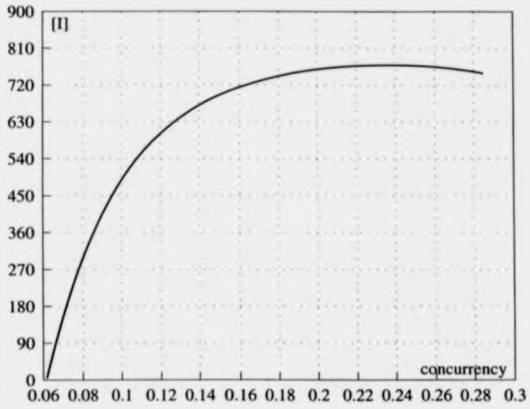


Figure 4.5: Bifurcation diagram for $[I]$ showing a decrease in endemicity with increasing concurrency κ_3. Parameters are: $\rho = 0.01$, $\lambda = 1$, $\nu = 0.006$, $P_0 = 500$, $N = 1500$. The continuation ends where the step size became too small.

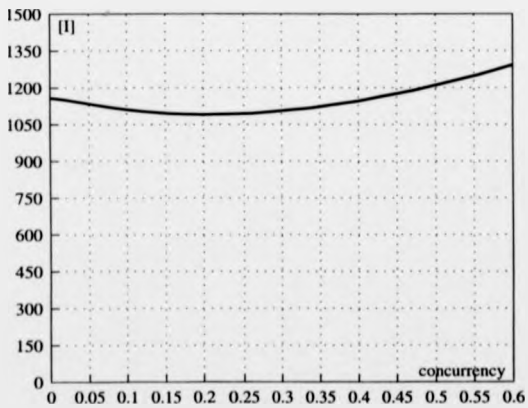


Figure 4.6: Bifurcation diagram showing the case of decreasing level of endemicity under HEPA $\rho = 0.01$, $\lambda = 0.01$, $\nu = 0.0006$, $P_0 = 442.4$, $N = 1500$.

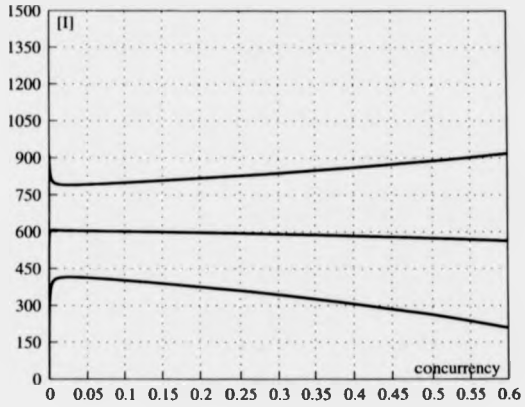


Figure 4.7: Bifurcation diagram under HEPA for $[I]$ for the case where partnership dynamics are fast relative to infection dynamics. Parameters are: $\rho = 0.2$, $P_0 = 365.7$, $\lambda = 0.0005$, $\nu = 0.0001, 0.00015, 0.0002$ from topmost branch to bottommost branch. $N = 1500$.

rency is increased, although it increases or decreases somewhat according to the recovery rate. One would not expect network structure to be important to the final epidemic size since in the limit of fast partnership dynamics we recover a compartmental model. Thus even when the partnership dynamics are occurring for a rapidly changing sexual partnership network, network structure has some impact. It is also interesting that the endemicity can either decrease or increase with increasing concurrency.

4.7 Calculation and Analysis of R_0

We use the method described in chapter 3 for the steady/casual model, accounting for the local structure evolution of invading patches. Here we are fortunate in that we can write down the equations of motion for the local structure variables $Q(S \mid I)$ and $Q(I \mid I)$ and solve explicitly for the steady-state solutions. This is easily done for square lattices [60][84] but is more difficult for network models like ours. By scaling time and the singleton and pair numbers, one can obtain a set of equations that depend only on parameters $\eta \equiv \rho/\sigma$, $\omega \equiv \lambda/\sigma$, $\delta \equiv \nu/\sigma$, and θ. It follows that R_0, which is coordinate-independent and independent of the unit of time, should only depend upon such dimensionless parameter combinations. After establishment of the localized infection patches at a pseudo-equilibrium we expect that quantities of the form $Q(i|I)$ and $Q(i|jI)$ will have reached pseudo-equilibrium values.

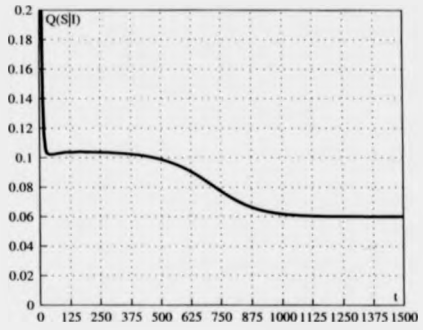


Figure 4.8: Time series of $Q(S \mid I)$ for $\rho = 0.01$, $\sigma = 0.005$, $\theta = 0.3$, $\lambda = 0.1$, $\nu = 0.02$.

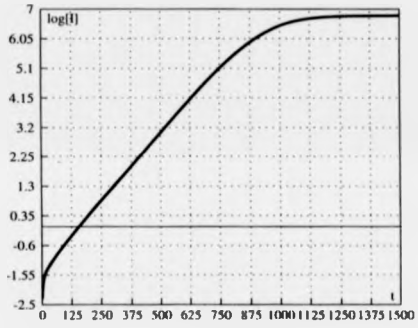


Figure 4.9: Time series of $\log[I]$ for $\rho = 0.01$, $\sigma = 0.005$, $\theta = 0.3$, $\lambda = 0.1$, $\nu = 0.02$.

Figure 4.8 shows a time series for $Q(S \mid I)$, illustrating the local structure evolution of the HEPA approximation to the concurrency model. The pseudo-equilibrium is established after the rapid decline of $Q(S \mid I)$ from its initial value of $Q(S \mid I) = Q$. Also, we see the gradual change in $Q(S \mid I)$ as the number of infected individuals increases. We compare this with figure 4.9 which shows the time series for $\log[I]$ for the same parameters. Note that immediately after $Q(S \mid I)$ has reached the quasi-equilibrium the growth of $[I]$ is exponential.

Using the equation $d[I]/dt = (-\nu + \lambda Q(S \mid I))[I]$ we see that after establishment, an infected individual transmits the infection to others at a rate $\lambda Q(S \mid I)_{inv}$. The expected length of time an individual remains infective is $1/\nu$. Thus we deduce that

$$R_0 = \frac{\lambda Q(S \mid I)_{inv}}{\nu}$$

We use the master equations (4.6) - (4.12) to derive equations for $dQ(S \mid I)/dt$,

$dQ(I \mid I)/dt$ and $d\Omega/dt$ where $\Omega = [X_I]/[I]$. These three quantities are the state variables which describe the time evolution of the local network structure. These equations still include the terms $Q(I \mid SI)$ and $Q(S \mid SI)$, but we use the HEPA to approximate them. We can then employ the quasi-equilibrium approximation $dQ(S \mid I)/dt = dQ(I \mid I)/dt = d\Omega/dt = 0$ and solve for $Q(S \mid I)_{inv}$, $Q(I \mid I)_{inv}$, and Ω_{inv}.

By taking the derivative of $Q(S \mid I) = [SI]/[I]$ we get $dQ(S \mid I)/dt = ([SI]^{-1}d[SI]/dt - [I]^{-1}d[I]/dt)Q(S \mid I)$. Inserting the expressions for $d[I]/dt$ and $d[SI]/dt$ from equations (4.12) and (4.6) produces:

$$\begin{aligned}\frac{1}{[I]}\frac{d}{dt}Q(S \mid I) &= \frac{\rho}{N}[X_S][X_I] + \frac{\rho\theta}{N}((N-[I])[I] - [X_S][X_I]) \\ &= -\sigma[SI] - \nu[SI] + \nu[II]\lambda[SS]Q(I \mid SS) \\ &= -\lambda[SI]Q(I \mid SI) - (\lambda Q(S|I) - \nu)[SI].\end{aligned}$$

Next we insert the expressions for $Q(I \mid SI)$ and $Q(I \mid SS)$ under HEPA, and also make the substitutions $[X_S] = X - [X_I]$, $[X_I] = \Omega[I]$, $[II] = Q(I \mid I)[I]$ and $[SI] = Q(S \mid I)[I]$ to produce equations in terms of $Q(S \mid I)$, $Q(I \mid I)$, Ω and $[I]$. Finally, we make the approximation that $[I]/N \approx 0$ (since the number of infecteds is initially small) to produce:

$$\begin{aligned}\frac{d}{dt}Q(S \mid I) &= \big[\rho\Omega QXN + \rho\theta QN^2 - \rho\theta Q\Omega XN - \sigma Q(S \mid I)N^2Q \\ &\quad + \nu Q(I \mid I)N^2Q + 4\lambda Q(S \mid I)NQ_{\geq 1}P - 4\lambda Q(S \mid I)NP \\ &\quad -\lambda Q(S \mid I)N^2QQ_{\geq 1} - \lambda Q(S \mid I)^2N^2Q)\big]/(N^2QQ(S \mid I))\end{aligned}$$

We can derive similar equations for $dQ(I \mid I)/dt$ and $d\Omega/dt$. Thus we have three differential equations in three unknowns. At establishment we can set $dQ(S \mid I)/dt = dQ(I \mid I)/dt = d\Omega/dt = 0$. Then we solve for the unknowns giving us expressions for $Q(S \mid I)_{inv}$, $Q(I \mid I)_{inv}$ and Ω_{inv} and hence R_0. The solution for $Q(S \mid I)_{inv}$ is a large third-order polynomial which is difficult to simplify except in special cases. Because of its length we do not write the full solution here. The complication of the solution reflects the complexity of the model rather than any unnecessary complication introduced by our method. However we can study it numerically.

Additionally, we are able to obtain a simple analytical expression when partnership dynamics are slow (when the $o(\rho)$ and $o(\sigma) \ll o(\lambda)$ and $o(\nu)$). In this case, R_0 reduces to:

$$R_0 = -1 - \chi + \xi + \sqrt{(\xi - \chi)^2 + 2\chi} \tag{4.21}$$

where $\chi = \lambda/\nu$, $\xi = 1/2 + \lambda P/(N\nu(1-\pi_0))$, and π_0 is as in equation (4.2).

Taking the derivative of R_0 with respect to π_0 shows that R_0 is monotone increasing in π_0. Also, previously it was shown that π_0 is itself an increasing function of θ on the interval $0 < \theta < 1$ (if P is held constant by adjusting σ). Thus, R_0 is also an increasing function of θ. As the proportion of individuals in the population with multiple partners increases, R_0 also grows and the disease is able to spread more quickly through the population. We conclude that concurrency always increases R_0 for slow partnership dynamics.

Because R_0 is monotone increasing in θ, we know that R_0 is bounded above and below on the interval $0 < \theta < 1$ according to the bounds for π_0. It is difficult to infer the shape of the R_0 curve on the $0 < \theta < 1$ interval because of the complicated dependence of π_0 on θ. However numerical results suggest that the dependence of R_0 on the index of concurrency κ_3 is roughly exponential.

We also carried out some numerical analysis of the full solution for R_0. Firstly, there is the question of how fast R_0 grows with increasing concurrency κ_3. For the parameters of figure 4.10 we see a strong increase in R_0 with concurrency in all cases, although the shape of the curve depends on the density of partnerships in the population. Figure 4.11, for the case of slow partnership dynamics, on the other hand suggests exponential growth of R_0 with concurrency (at least up to $\kappa_3 = 0.8$) for a range of values of P. However because we have not compared the expression for R_0 to results from the stochastic model we cannot be sure how accurate our expression is. In future, work should be carried out to compare the expression for R_0 to data from simulations.

4.8 Discussion

In this chapter we defined a stochastic model for STD spread through a sexual partnership network with concurrency and we obtained three moment closure approximations for it. We compared the approximations with stochastic simulation data, analyzed the effect of concurrency using the IVPA and HEPA approximations, and used a recently-developed method for deriving R_0 in spatial systems to get an explicit expression for R_0 for the concurrency model.

As with the steady/casual model, we see that choosing an appropriate moment closure approximation for dynamic network models is very important, especially in comparison with static lattice models. This is because dynamic networks have added complexity on account of the formation and dissolution of edges. We saw that several types of pair approximations have different accuracies in different parameter regimes. Based on our experiences with the steady/casual and concurrency models,

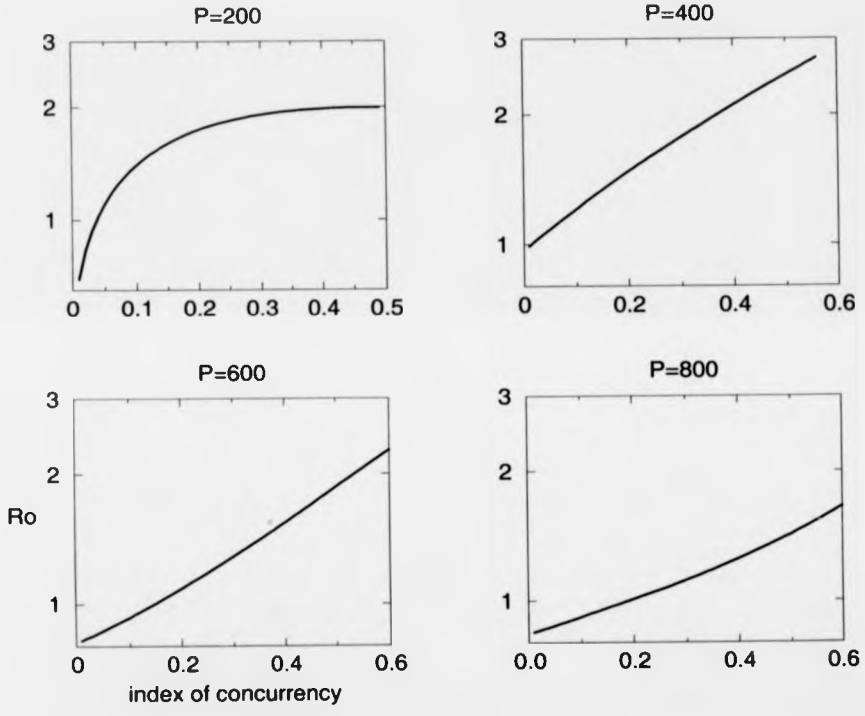


Figure 4.10: Linear-log plot of R_0 against κ_3. $\rho = 0.02$, $\lambda = 0.05$ and $r = 0.005$. Values for P are shown above diagrams.

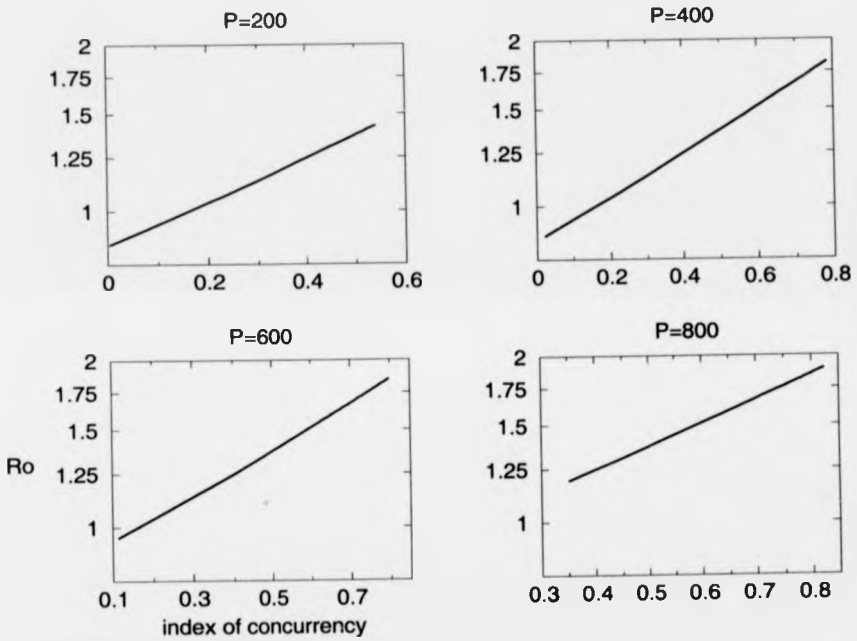


Figure 4.11: Linear-log plot of R_0 against κ_3 for the special case of slow partnership dynamics (see equation (4.21)). Parameters are: $\rho = 0.001$, $\lambda = 0.12$ and $\nu = 0.01$. Values for P are shown above diagrams.

it seems each problem should be considered individually in order to find the optimal moment closure scheme for that problem. Nonetheless there may exist some general moment closure scheme for network models which would circumvent this complication, although it may be difficult to find.

Pursuing an accurate moment closure approximation is important, and sometimes OPA simply will not do. We do not have a guarantee that results are robust for different moment closures because they make different assumptions about the underlying dynamics. If some important aspect of the model dynamics is excluded, such as conditional dependence or clustering, it is not clear whether or not the approximation has anything to do with the model it is intended to reproduce.

For the steady/casual model at least, it appears that HEPA and OPA work best for fast partnership dynamics, whereas the IVPA works best for slow partnership dynamics. Indeed, in network models the IVPA marks a significant improvement over the OPA, and it is simple and intuitively appealing. We are able to model accurately systems with low average node degree without using an OTA as in the steady/casual model. This is in contrast to the lesser improvement IVPA provides in the case of the contact process (chapter 2). The demonstration of the new method for deriving R_0 (where we exploit knowledge about the early time evolution of the network structure) is also interesting. The concept is a very natural one for such systems. We believe that this approach will be of considerable use in a wide range of invasion problems in spatial and network models.

Finally, the results show that concurrency in sexual partnership networks greatly increases the impact of an STD, both in terms of endemicity and the basic reproduction ratio R_0. This increase occurs even when the total number of partnerships P is held constant as κ_3 is varied.

One possible elaboration of this model would be to tailor it more closely to particular STDs. This could be done just by looking at particular areas of the parameter space or by introducing elements such as disease-induced mortality, age-structure, core groups, etc. Although the sexual network in this model is not sociologically realistic, we believe that the results are robust for the other types of concurrency found in real sexual networks. One way to check this would be to introduce other types of concurrency models and compare the effect they have on the disease epidemiology.

In the next chapter we move to a completely different application of moment closure approximations, namely the search for deterministic approximations of epidemics in populations living in continuous space. Much like the differences in the dynamics of network models and the contact process, we will see there are some interesting differences between the continuous space model and network models with

respect to clustering.

Chapter 5

Spatial Randomness and Epidemic Spread

I haven't failed; I've found 10,000 ways that don't work.

–Albert Einstein

5.1 Introduction

In the previous chapters we considered infectious diseases on dynamic networks. These models are spatial in the sense that different individuals 'see' different environments. We now turn to moment closure approximations for models which are spatial in the more usual sense of the word; i.e. the population is embedded in a two-dimensional continuous space where the relative location of individuals influences the disease spread.

Our main goals in this preliminary work are a) to devise the simplest moment closure approximation which can capture certain aspects of a continuous space model and b) to explore the feasibility of applying moment closure approximations to epidemics occurring in populations embedded in continuous space, instead of the traditional lattice-based IPS models. We also want to understand through these models how epidemics in randomly-distributed populations are different from epidemics on a regular lattice.

5.2 Previous Work on Continuous Space Models

In the real world, individuals do not live on square grids. The spatial distribution of individuals can affect epidemic spread and therefore there is a need for epidemic models which incorporate continuous space. Most moment closure approximations are intended to model regular lattices, or perhaps indirectly some generic spatial distribution in space. These approximations are simple to derive, but in real biological systems there may be some dependence of disease dynamics on the distance between individuals and this is not accounted for explicitly in square lattice models. For instance the infection process for a plant disease may be described by a kernel function of the distance from the infected individual; the disease vector might be a wandering insect. Secondly, there is no possibility for clustering of individual's locations on a square lattice. Although some moment closure approximations for network models can get around this to some degree, they still require a constant number of neighbours Q per person.

For instance, Morris et al. [75][84][90] study clustering within the framework of pair approximations by taking triangles into account in their moment closure. The motivation is that in real networks, two neighbours of a given individual may also be neighbours, thus forming highly-correlated triangles. This can significantly influence dynamics and makes conditional independence a less viable assumption. Normally

the number of triples in the network is approximated as

$$[ijk] = \frac{m-1}{m}\frac{[ij][jk]}{[j]}$$

however triangles are taken into account by using the closure:

$$[ijk] = \frac{m-1}{m}\frac{[ij][jk]}{[j]}\left(1-\phi+\frac{\phi N}{Q}\frac{[ik]}{[i][k]}\right)$$

where ϕ is the proportion of triples which are triangles. The density of triangles is taken to indicate the amount of clustering in the population. Of course there can be clustering in a population in more than one sense. This closure can model one aspect of spatial clustering, but clustering can be expressed more explicitly in terms of the number of neighbours of an individual and their distance to that individual. To model this kind of clustering we should develop moment closures for networks with varying Q or for continuous spatial models.

Some of the most interesting work in the area of spatial modelling in recent years is by Bolker and Pacala [12] [13]. They derive a pair approximation in one dimension describing the auto- and cross-correlation of species densities at points in continuous space a distance x apart. This tells us the effect a certain density of species A at point z has on the density of species A, B, etc. at point w, where $\|z-w\| = x$. The equations are derived by considering the dynamics in discrete space made of units of length h, and then taking the limit $h \to 0$ to obtain a model for continuous space. Their application is ecological competition models, but a model describing disease spread could also be written down using this framework. In the same vein, Gandhi et al. [42] derive moment closure approximations to RDE models, using a Gaussian approximation similar to that originally used in turbulence [82] which is well-suited to populations in a continuous spatial environment. They use the moment closure approximation mostly as a supplement to the full RDE model.

Our MCA is similar to that of Bolker and Pacala in that we want to find a moment closure description of continuous space, except that they consider correlations of densities of species at difference points in space after taking the limit of a discrete case, whereas our model will retain individual-based measures. Our model can be thought of as a modified network model where the relative location of individuals in two-dimensional Euclidean space determines the strength of interaction.

5.3 Definition of the Stochastic Model

We randomly distribute N individuals over a two-dimensional Euclidean space of size $H \times H$ with periodic mapping at the edges. Once placed, individuals cannot move.

Infection dynamics follow an SIS scheme (however for plant diseases for instance one should include disease-induced mortality). Infected individuals recover at a rate ν. We define a distance-dependent infection kernel which determines the rate at which an infected individual transmits the disease to a susceptible. In this chapter we take an exponentially decaying kernel:

$$\text{rate of infection} = \begin{cases} e^{-k\|x-y\|} & \|x-y\| \leq R \\ 0 & \|x-y\| > R \end{cases} \tag{5.1}$$

where R is a parameter intended to restrict the distance over which an infected can transmit an infection, and where $\|x-y\|$ is the distance between the infected individual x and susceptible individual y. The kernel is truncated for simplicity although a discontinuous kernel is probably not natural. However individuals very far from the infected individual do not face significant infection risk from him since de^{-kd} is small for large distances d, and we can often choose our truncation R accordingly. This simplification will make it easier to run simulations and write down a deterministic approximation of this stochastic model.

Individuals within a radius R of an individual are called *neighbours* of that individual. We can speak of there being an edge in the network sense between an individual and his neighbours, but of course this is not an edge in the sense of network models because the distance between individuals matters. The expectation value of the number of individuals falling within a circle of radius R is denoted Q. Because this is a Poisson process, Q is equal to the expected number of neighbours per individual. The overall population density N/H^2 is denoted by $\varrho = Q/\pi R^2$.

5.4 Some Preliminary Analysis of the Stochastic Model

Before considering how to capture elements of continuous space in a deterministic approximation, we analyze the data from the stochastic model in order to better understand the process. With this information we can formulate a better deterministic approximation.

Three probability functions (PFs) are particularly useful in understanding the stochastic model: $\Omega_{SI}(m)$ is the probability function of the number m of infected neighbours per susceptible, and $\Omega_{II}(m)$ and $\Omega_{SS}(m)$ are similarly defined. We also need three probability density functions (PDFs): $\psi_{SI}(r)$ is the PDF of the number of infected individuals at a distance r from a susceptible individual, and $\psi_{II}(r)$ and

$\psi_{SS}(r)$ are similarly defined. The PFs and PDFs follow the usual normalization rules:

$$\begin{aligned} \int_0^R \psi_{SI}(r) r &= 1 \\ \sum_{m=0}^{\infty} \Omega_{SI}(m) &= 1 \end{aligned} \tag{5.2}$$

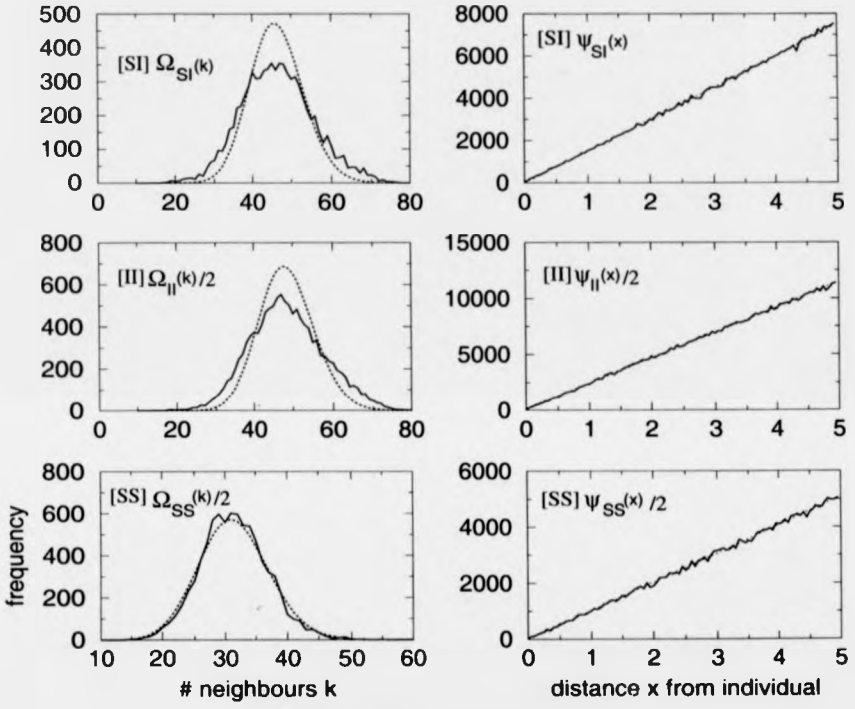
Finally, we need $\psi(r)$, the PDF of the number of individuals at a distance r from an individual, and $\Omega(m)$, the PF of the number of neighbours of an individual. $\psi(r)$ and $\Omega(m)$ are determined by the initial spatial distribution of the population.

Figures 5.1 through 5.4 show examples of these six functions taken from the stochastic model data, in addition to information about $[I]$ at equilibrium and the mean μ_{SI} (resp. μ_{II}, μ_{SS}, μ) of the $\psi_{SI}(r)$ (resp. $\psi_{II}(r)$, $\psi_{SS}(r)$, $\psi(r)$) PDF. All data represent a long-time average taken after the infection dynamics have equilibrated.

Roughly speaking, $\psi_{SI}(r)$, $\psi_{II}(r)$ and $\psi_{SS}(r)$ are linear in r, although there are some interesting deviations. Figures 5.3 and 5.4 in particular show significant deviations of $\psi_{SI}(r)$ and $\psi_{II}(r)$ from the linear regression. Nonlinearity in these functions is also just apparent in figure 5.2. Where deviation in $\psi_{SI}(r)$ and $\psi_{II}(r)$ exists it indicates clustering of infecteds. Since the force of infection declines with distance, we expect to find that in areas where the local population density is high, there are more infected individuals. Hence ψ_{II} should be concave. For a similar reason we expect ψ_{SI} to be convex. The reason these effects are more apparent in figures 5.2, 5.3 and 5.4 than in figure 5.1 is that the number of neighbours per person Q in the latter case is very high which makes the relative locations of susceptibles and infecteds less important. Thus, clustering is less. We can use the nonlinearity of the ψ_{SI} and ψ_{II} probability density functions as measures of spatial clustering.

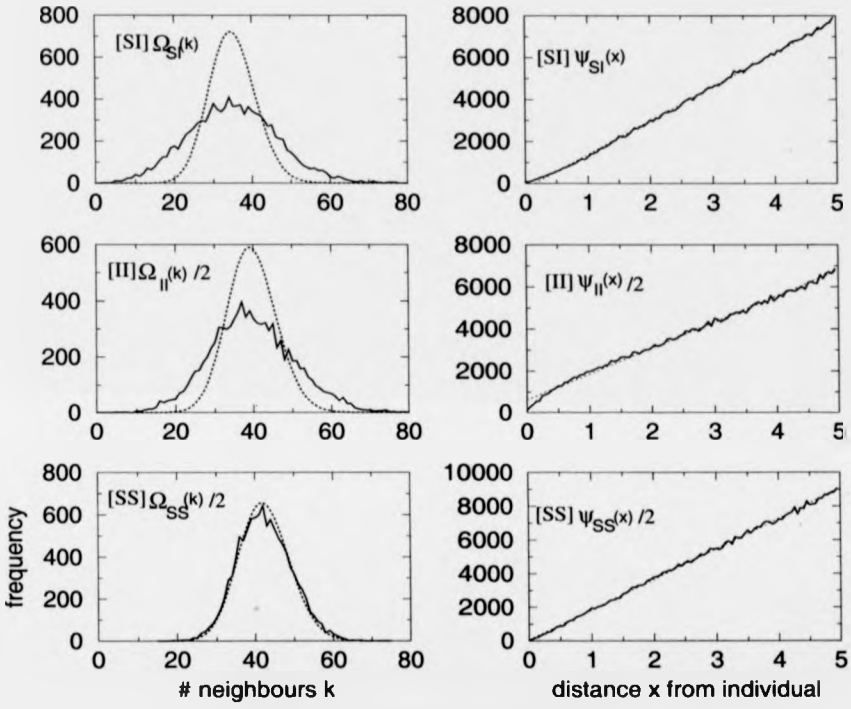
Interestingly the nonlinearity is not observable in any of the $\psi_{SS}(r)$ data even when it is obvious in the $\psi_{II}(r)$ and $\psi_{SI}(r)$ data. This reflects the fact that SS pairs are not created from any neighbour-dependent process but rather from the recovery of infected individuals in SI pairs. Since recovery is a random process with respect to population structure, we expect linearity in the $\psi_{SS}(r)$ PDF. We can use this observation in the derivation of a deterministic approximation to reduce the dimensionality of the equations of motion.

Clustering is also reflected in the means μ_{SI} and μ_{II}. The concavity of $\psi_{SI}(r)$ guarantees that $\mu_{SI} > \mu$ always and similarly we know that $\mu_{II} < \mu$. For a linear function hr in r between 0 and R and slope h, the mean μ is $\mu = 2R/3$. Thus $\mu_{SS} = \mu = 2R/3$ in all cases. μ_{SI}/μ or μ_{II}/μ can serve as a measure of clustering of infecteds.



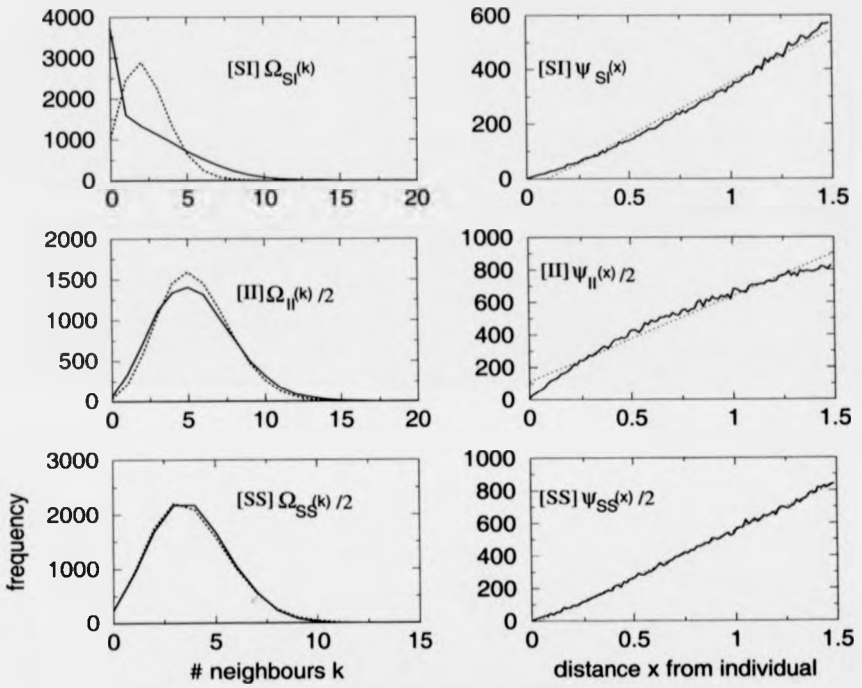
$\mu_{SI} = 3.294$	$\mu_{SS} = 3.292$
$\mu_{II} = 3.268$	$\mu = 3.284$
$[I] = 11966$	

Figure 5.1: Six distributions from the simulation data; case of high Q and gradual kernel decay. Each data point is the average of eight runs. Parameters are $k = 0.5$, $R = 5.0$, $\nu = 7.0$, $\varrho = 1.0$, $N = 20000$. The dashed line in graphs on the left is a Poisson distribution with mean equal to that of the stochastic data. The dashed line in the graphs on the right is a linear regression fitted to simulation data. Bin size for RHS $= R/100$.



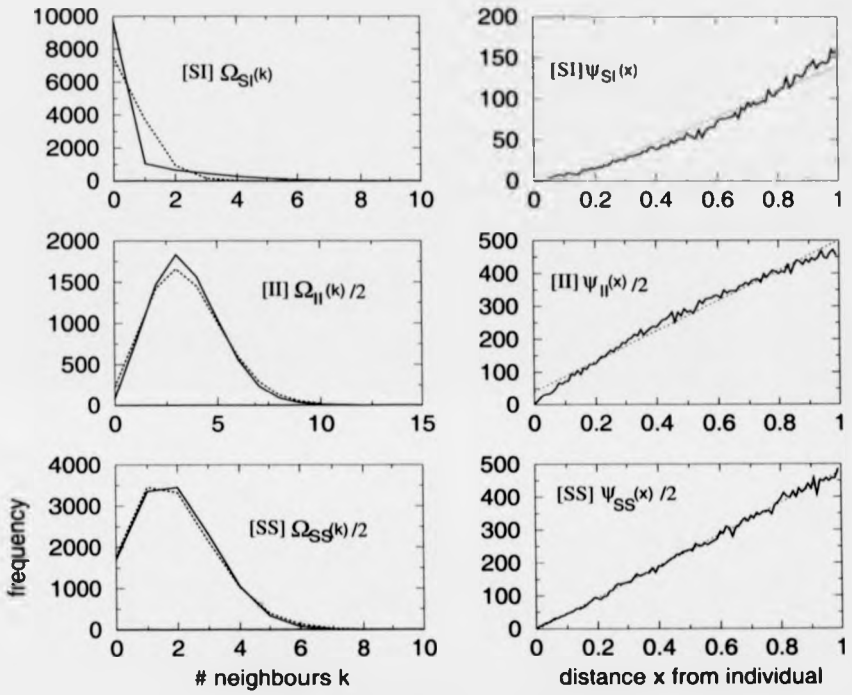
$\mu_{SI} = 3.348$	$\mu_{SS} = 3.284$
$\mu_{II} = 3.150$	$\mu = 3.283$
$[I] = 9297$	

Figure 5.2: Six distributions from the simulation data; case of high Q and rapid kernel decay. Each data point is the average of eight runs. Parameters are $k = 2.5$, $R = 5.0$, $\nu = 0.4$, $\varrho = 1.0$, $N = 20000$. The dashed line in graphs on the left is a Poisson distribution with mean equal to that of the stochastic data. The dashed line in the graphs on the right is a linear regression fitted to simulation data. Bin size for RHS $= R/100$.



$\mu_{SI} = 1.025$	$\mu_{SS} = 0.998$
$\mu_{II} = 0.930$	$\mu = 0.985$
$[I] = 9162$	

Figure 5.3: Six distributions from stochastic model data; small Q and rapid kernel decay. Parameters are $k = 2.5$, $R = 1.5$, $\nu = 0.3$, $\varrho = 1.0$, $N = 20000$. Each data point is the average of eight runs. The dashed line in graphs on the left is a Poisson distribution with mean equal to that of the stochastic data. The dashed line in the graphs on the right is a linear regression fitted to simulation data. Bin size for RHS = $R/100$.



$\mu_{SI} = 0.657$	$\mu_{SS} = 0.657$
$\mu_{II} = 0.631$	$\mu = 0.655$
$[I] = 7489$	

Figure 5.4: Six distributions from stochastic model data; small Q and rapid kernel decay. Parameters are $k = 3.0$, $R = 1.0$, $\nu = 0.15$, $\varrho = 1.0$, $N = 20000$. Each data point is the average of eight runs. The dashed line in graphs on the left is a Poisson distribution with mean equal to that of the stochastic data. The dashed line in the graphs on the right is a linear regression fitted to simulation data. Bin size for RHS $= R/100$.

We can also learn something from the $\Omega_{SI}(m)$, $\Omega_{II}(m)$ and $\Omega_{SS}(m)$ probability functions. $\Omega_{SI}(m)$ and $\Omega_{II}(m)$ exhibit flattening relative to a Poisson distribution in figures 5.1 and 5.2 even though $\psi_{SI}(x)$ and $\psi_{II}(r)$ are more or less linear in these cases. In figures 5.3 and 5.4, $\Omega_{SI}(m)$ and $\Omega_{II}(m)$ are poorly described by a Poisson distribution although $\Omega_{II}(m)$ and $\Omega_{II}(m)$ fit a Poisson distribution quite well (probably because $Q(I \mid I) > Q(I \mid S)$). The flattening observed in the first two figures suggests the clustering of infecteds, although it is a clustering in terms of numbers of infected neighbours instead of their distribution in space. We can thus distinguish between *distance clustering* where infecteds are closer to one another, and *number clustering* where clustering is reflected in the relative numbers of SI, II and SS edges. Whether or not distance clustering is present depends on the relative value of k and R, which determines the amount of decline of the kernel function between $r = 0$ and the truncation $r = R$. The graphs for $\psi_{SI}(r)$ and $\psi_{II}(r)$ in figure 5.1 and 5.2 do not show distance-clustering while the graphs for $\Omega_{SI}(m)$ and $\Omega_{II}(m)$ do show number clustering, showing that it is possible to have one form of clustering without the other for the right choice of parameters. In figures 5.3 and 5.4 the large deviation from a Poisson distribution is because when Q is small, the infection states of an individual's neighbours are highly correlated and so the distribution is not well-described by assuming a Poisson distribution. $\Omega_{SS}(m)$ is modelled by a Poisson distribution very well in all four figures, and this is again because the recovery process which generates SS pairs is independent of network structure.

Accounting for the deviation in $\Omega_{SI}(m)$ from a Poisson distribution at low Q in a deterministic model is the same problem that plagues other moment closure models because it has to do with conditional dependence. The truncations used for low Q cases in the chapters on STDs would probably come in useful here (for $Q < 4$). However incorporating the flattening effect seen in cases of high Q should be easier. For large Q the distribution is approximately normal. Usually, in deriving pair approximations, the mean of the distribution is incorporated into equations of motion as a dynamical variable (through terms such as $Q(I \mid S) = [SI]/[S]$). In addition to the mean one could introduce the variance of the distribution as a dynamical variable. Since a normal distribution is completely characterized by the mean and variance, such a deterministic approximation should be pretty accurate while not being too complicated. To our knowledge, incorporating the variance of $\Omega_{SI}(m)$ as a state variable in deterministic approximations has not been done for ecological systems, although in [82] Orszag considers a Gaussian truncation of a hierarchy of moment equations derived from the Navier-Stokes equation). At any

rate, in this preliminary work we will only try to model the nonlinearity in $\psi_{SI}(r)$ which represents distance clustering.

5.4.1 Snapshots of the Simulation at Equilibrium

Figure 5.5 shows snapshots of the population after the epidemic has reached equilibrium. There are two different cases, one with large Q and one with small Q (compare to figures 5.1 and 5.3 respectively). It is clear that there is more distance clustering in the bottom picture (which has smaller Q) than the top picture, a fact we also gleaned from figures 5.1 through 5.4.

5.5 Derivation of the Pair Approximation

We are trying to find a deterministic approximation to a network model where the force of infection on a susceptible depends on the distance to neighbours as well as their number. We also want an approximation with the smallest number of assumptions which can still reproduce the interesting behaviour found in the simulation data, and perhaps point out behaviour we missed in the stochastic model. Previously in this thesis we looked at network models where the relative location of individuals in space does not matter. Here, since the lengths of edges comes into the picture, the obvious choice for our state variables are $\psi_{SI}(r)$, $\psi_{II}(r)$ and $\psi_{SS}(r)$ which are the probability density functions of the number of SI (resp. II and SS) edges of length $0 < r \leq R$.

Deriving the equations of motion for these state variables is a slight extension to our previous MCA derivations. We start with the variable $[SI](r, dr)$ which is the number of SI edges of length $l \in [r, r+dr]$. Similarly, $[II](r, dr)$ (resp. $[SS](r, dr)$) is twice the number of II (resp. SS) edges of length $l \in [r, r+dr]$ We will also need the usual state variables $[SI]$, $[II]$ and $[SS]$. We denote the force of infection on node x as Λ_x. Also, $Q_x^{r,dr}(I)$ (resp. $Q_x^{r,dr}(S)$) is the number of length $l \in [r, r+dr]$ edges of node x connecting x to infected (resp. susceptible) neighbours. Finally, $\mathcal{N}_x(I)$ denotes the set of infected neighbours of node x and ς_x denotes the state of node x. Notation such as $\varsigma_{xy} = SI_{r,dr}$ means that the pair of nodes x and y are connected by an edge of length $l \in [r, r+dr]$. The equation of motion for

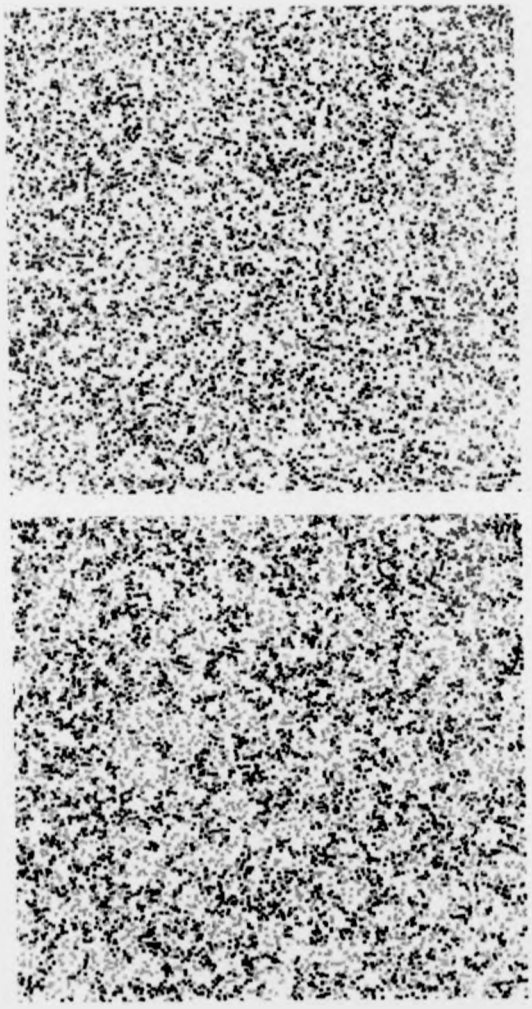


Figure 5.5: Snapshot of stochastic model at equilibrium. Dark gray individuals are infected and light gray are susceptible. Size of grid is 100×100. Parameters for top picture are $k = 0.5$, $R = 5.0$, $\nu = 7.0$, $\varrho = 1.0$, $N = 10000$; $[I] = 5923$. Parameters for bottom picture are $k = 2.5$, $R = 1.5$, $\nu = 0.25$, $\varrho = 1.0$, $N = 10000$; $[I] = 5448$.

$[SI](r, dr)$ is:

$$\frac{d}{dt}[SI](r,dr) = \sum_{\varsigma_{xy}=SI_{r,dr}} \nu(-1) + \sum_{\varsigma_{xy}=II_{r,dr}} \nu(+1) + \sum_{\varsigma_{xy}=SI_{r,dr}} \Lambda_x(-1) + \sum_{\varsigma_{xy}=SS_{r,dr}} \Lambda_x(+1)$$

For Λ_x we use:

$$\begin{aligned} \Lambda_x = \; & e^{-kr} + \textstyle\sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} e^{-kd(x,z)} && \text{for } \varsigma_{xy} = SI_{r,dr} \\ \Lambda_x = \; & \textstyle\sum_{\varsigma_z \in \mathcal{N}_x(I)} e^{-kd(x,z)} && \text{for } \varsigma_{xy} = SS_{r,dr} \end{aligned} \tag{5.3}$$

where $d(x, z)$ is the length of the edge connecting x and z, and we have written down the contribution of node y explicitly and subtracted it from $\mathcal{N}_x(I)$ Inserting this into equation (5.3) gives:

$$\begin{aligned} \frac{d}{dt}[SI](r,dr) = & \sum_{\varsigma_{xy}=SI_{r,dr}} \nu(-1) + \sum_{\varsigma_{xy}=II_{r,dr}} \nu(+1) \\ & + \sum_{\varsigma_{xy}=SI_{r,dr}} \left\{ e^{-kr} + \sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} e^{-kd(x,z)} \right\} (-1) \\ & + \sum_{\varsigma_{xy}=SS_{r,dr}} \left\{ \sum_{\varsigma_z \in \mathcal{N}_x(I)} e^{-kd(x,z)} \right\} (+1) \end{aligned} \tag{5.4}$$

At this point we have to invoke another substitution for the exponential terms of the neighbourhood of x in terms of means and fluctuations:

$$\begin{aligned} e^{-kd(x,z)} = \; & \lambda_{SI} + \eta^{\mathsf{exp}}_{xz}(SI) && \text{for } \varsigma_{xy} = SI_{r,dr} \\ e^{-kd(x,z)} = \; & \lambda_{SI} + \eta^{\mathsf{exp}}_{xz}(SI) && \text{for } \varsigma_{xy} = SS_{r,dr} \end{aligned} \tag{5.5}$$

where λ_{SI} is the force of infection averaged over all SI edges in the population, and $\eta^{\mathsf{exp}}_{xz}(SI)$ is again the deviation of $e^{-kd(x,z)}$ from the average value at node pair xz.

Now we simplify the sums in equation (5.4) which involve the neighbourhoods $\mathcal{N}_x(I)$, applying conditional independence at several points. The first double sum in

$[SI](r, dr)$ is:

$$\frac{d}{dt}[SI](r,dr) = \sum_{\varsigma_{xy}=SI_{r,dr}} \nu(-1) + \sum_{\varsigma_{xy}=II_{r,dr}} \nu(+1) + \sum_{\varsigma_{xy}=SI_{r,dr}} \Lambda_x(-1) + \sum_{\varsigma_{xy}=SS_{r,dr}} \Lambda_x(+1)$$

For Λ_x we use:

$$\begin{aligned} \Lambda_x &= e^{-kr} + \textstyle\sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} e^{-kd(x,z)} && \text{for } \varsigma_{xy} = SI_{r,dr} \\ \Lambda_x &= \textstyle\sum_{\varsigma_z \in \mathcal{N}_x(I)} e^{-kd(x,z)} && \text{for } \varsigma_{xy} = SS_{r,dr} \end{aligned} \tag{5.3}$$

where $d(x, z)$ is the length of the edge connecting x and z, and we have written down the contribution of node y explicitly and subtracted it from $\mathcal{N}_x(I)$ Inserting this into equation (5.3) gives:

$$\begin{aligned} \frac{d}{dt}[SI](r,dr) = & \sum_{\varsigma_{xy}=SI_{r,dr}} \nu(-1) + \sum_{\varsigma_{xy}=II_{r,dr}} \nu(+1) \\ & + \sum_{\varsigma_{xy}=SI_{r,dr}} \left\{ e^{-kr} + \sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} e^{-kd(x,z)} \right\} (-1) \\ & + \sum_{\varsigma_{xy}=SS_{r,dr}} \left\{ \sum_{\varsigma_z \in \mathcal{N}_x(I)} e^{-kd(x,z)} \right\} (+1) \end{aligned} \tag{5.4}$$

At this point we have to invoke another substitution for the exponential terms of the neighbourhood of x in terms of means and fluctuations:

$$\begin{aligned} e^{-kd(x,z)} &= \lambda_{SI} + \eta^{\mathsf{exp}}_{xz}(SI) && \text{for } \varsigma_{xy} = SI_{r,dr} \\ e^{-kd(x,z)} &= \lambda_{SI} + \eta^{\mathsf{exp}}_{xz}(SI) && \text{for } \varsigma_{xy} = SS_{r,dr} \end{aligned} \tag{5.5}$$

where λ_{SI} is the force of infection averaged over all SI edges in the population, and $\eta^{\mathsf{exp}}_{xz}(SI)$ is again the deviation of $e^{-kd(x,z)}$ from the average value at node pair xz.

Now we simplify the sums in equation (5.4) which involve the neighbourhoods $\mathcal{N}_x(I)$, applying conditional independence at several points. The first double sum in

equation (5.4) is simplified as follows:

$$\begin{aligned}
\sum_{\varsigma_{xy}=SI_{r,dr}} \sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} e^{-kd(x,z)} &= \sum_{\varsigma_{xy}=SI_{r,dr}} \sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} (\lambda_{SI} + \eta_{xz}^{\exp}(SI)) \\
&= \lambda_{SI} \sum_{\varsigma_{xy}=SI_{r,dr}} \sum_{\varsigma_z \in \mathcal{N}_x(I)\setminus y} \\
&= \lambda_{SI} \sum_{\varsigma_{xy}=SI_{r,dr}} (Q_{xy}(I) - 1) \\
&= \lambda_{SI} \sum_{\varsigma_{xy}=SI_{r,dr}} (Q(I \mid SI_{r,dr}) + \eta_{xy}(I \mid SI_{r,dr}) - 1) \\
&= \lambda_{SI} Q(I \mid S) \sum_{\varsigma_{xy}=SI_{r,dr}} \\
&= \lambda_{SI} \frac{[SI]}{[S]} [SI](r, dr) \qquad (5.6)
\end{aligned}$$

where $Q(I \mid SI_{r,dr})$ is the expectation value of the number of infected neighbours of a susceptible, subject to the condition that the susceptible has at least one neighbour a distance $l \in [r, r+dr]$ away. We have applied conditional independence to obtain line five. In line five we have used the approximation $Q(I \mid SI_{r,dr}) \approx 1 + Q(I \mid S)$ on the basis that the number of neighbours per individual is Poisson-distributed. We proceed in similar fashion for the other double sum over $SS_{r,dr}$ pairs to get:

$$\sum_{\varsigma_{xy}=SS_{r,dr}} \sum_{\varsigma_z \in \mathcal{N}_x(I)} e^{-kd(x,z)} = \lambda_{SI}[SS](r, dr) \frac{[SI]}{[S]} \qquad (5.7)$$

Therefore after evaluating the other sums equation (5.4) becomes:

$$\begin{aligned}
\frac{d}{dt}[SI](r, dr) &= -\nu[SI](r, dr) + \nu[II](r, dr) - e^{-kr}[SI](r, dr) \\
&\quad -\lambda_{SI}[SI](r, dr) \frac{[SI]}{[S]} + \lambda_{SI}[SS](r, dr) \frac{[SI]}{[S]} \qquad (5.8)
\end{aligned}$$

We want to consider the PDFs $\psi_{SI}(x)$, $\psi_{II}(x)$ and $\psi_{SS}(x)$ which were measured for the simulation data in the first part of this chapter. These can be obtained by taking the limits:

$$\begin{aligned}
\psi_{SI}(r) &\equiv \lim_{dr \to 0} \frac{[SI](r, dr)}{[SI]} \\
\psi_{II}(r) &\equiv \lim_{dr \to 0} \frac{[II](r, dr)}{[II]} \\
\psi_{SS}(r) &\equiv \lim_{dr \to 0} \frac{[SS](r, dr)}{[SS]} \qquad (5.9)
\end{aligned}$$

The state variables $\psi_{SI}(r)$, $\psi_{II}(r)$ and $\psi_{SS}(r)$ are referred to as *edge-length variables*, whereas $[SI]$, $[II]$ and $[SS]$ are referred to as *edge variables*. We divide

equation (5.8) by $[SI]$, rearrange the $[SS](r,dr)$ and $[II](r,dr)$ terms slightly and take the limit $dr \to 0$ to get:

$$\begin{aligned}\frac{d}{dt}\psi_{SI}(r) &= -\nu\psi_{SI}(r) + \nu\frac{[II]}{[SI]}\psi_{II}(r) - e^{-kr}\psi_{SI}(r) \\ &\quad -\lambda_{SI}\frac{[SI]}{[S]}\psi_{SI}(r) + \lambda_{SI}\frac{[SS]}{[S]}\psi_{SS}(r) \qquad (5.10)\end{aligned}$$

For $\psi_{II}(r)$ and $\psi_{SS}(r)$ the derivation is similar:

$$\frac{d}{dt}\psi_{II}(r) = -2\nu\psi_{II}(r) + 2e^{-kr}\frac{[SI]}{[II]}\psi_{SI}(r) + 2\lambda_{SI}\frac{[SI]}{[II]}\frac{[SI]}{[S]}\psi_{SI}(r) \quad (5.11)$$

$$\frac{d}{dt}\psi_{SS}(r) = 2\nu\frac{[SI]}{[SS]}\psi_{SI}(r) - 2\lambda_{SI}\frac{[SI]}{[S]}\psi_{SS}(r) \quad (5.12)$$

We now derive the constraint for the state variables. We know that the initial geometry constrains the total number of links, and more specifically, the frequency of links of a given length l. Let $L(r,dr)$ be the number of links of length $l \in [r, r+dr]$. Then:

$$2L(r,dr) = 2[SI](r,dr) + [II](r,dr) + [SS](r,dr)$$

Rearranging terms produces:

$$2L\frac{L(r,dr)}{L} = 2[SI]\frac{[SI](r,dr)}{[SI]} + [II]\frac{[II](r,dr)}{[II]} + [SS]\frac{[SS](r,dr)}{[SS]}$$

Taking the limit as $dr \to 0$ produces:

$$2L\psi(r) = 2[SI]\psi_{SI}(r) + [II]\psi_{II}(r) + [SS]\psi_{SS}(r) \quad (5.13)$$

where $\psi(r)$ is the PDF of the number of links of length r and L is the total number of links, i.e. $QN/2$. $\psi(r)$ is normalized to one:

$$\int_0^R dr\psi(r) = 1$$

The constraint (5.13) reduces the dimensionality by one. We could reduce the dimensionality further by using the empirical observation that $\psi_{SS}(r)$ seems to be linear always, but because this observation is not well-confirmed we do not attempt this in our preliminary study. However it might be possible to solve exactly the equilibrium solutions of $\psi_{SS}(r)$, $\psi_{SI}(r)$ and $\psi_{II}(r)$, and thus deduce the dependence of $\psi_{SS}(r)$ on r, at least at equilibrium.

If we integrate expression (5.13) we recover the constraint for the total number

of edges in a network (compare with equation (1.18)):

$$\begin{aligned} 2L\int_0^R dr\psi(r) &= 2[SI]\int_0^R dr\psi_{SI}(r) + [II]\int_0^R dr\psi_{II}(r) \\ &\quad +[SS]\int_0^R dr\psi_{SS}(r) \\ \Rightarrow 2L &= 2[SI] + [II] + [SS] \\ \Rightarrow QN &= 2[SI] + [II] + [SS] \end{aligned} \tag{5.14}$$

We can calculate $\psi(r)$ explicitly from the normalization requirement and the assumption of linearity of $\psi(r)$ (which follows from the geometry and the fact that the population distribution is uniform and random):

$$\psi(r) = \frac{2r}{R^2} \tag{5.15}$$

One can take the time derivative of equation (5.13) and use the equations of motion (5.10), (5.11) and (5.12), and the equations of motion (5.18) for the edge variables (derived in the following paragraphs) to confirm that $\frac{d}{dt}\psi(r) = 0$, as expected.

The most important term in these equations is the $e^{-kr}\psi_{SI}(r)$ term because it supplies the explicit dependence on the length of edges. The ratio $1/Q$ is also important because it tells us how much of an effect the $e^{-kr}\psi_{SI}(r)$ term has relative to the population-averaged contributions of λ_{SI}.

Further dependence on the network structure is introduced by λ_{SI}, a quantity we now derive. To calculate λ_{SI} we need the probability distribution function $\psi_{SI}(r)$:

$$\lambda_{SI} = \int_0^R e^{-kr}\psi_{SI}(r)dr \tag{5.16}$$

We know that $\psi_{SI}(r)$ is almost linear (with the nonlinearity due to clustering) so in the first instance we take $\psi_{SI}(r) = \psi(r) = 2r/R^2$:

$$\begin{aligned} \lambda_{SI} &= \int_0^R e^{-kr}\frac{2r}{R^2}dr \\ &= \frac{2}{k^2R^2} - \frac{2e^{-kR}}{kR}\left(1 + \frac{1}{kR}\right) \end{aligned} \tag{5.17}$$

In future we can increase the accuracy of the approximation by using a perturbation expansion of $\psi_{SI}(r)$ in r, but initially we just use equation (5.17).

Now we have a closed set of equations and the only further requirements are to know the edge variables $[SI]$, $[II]$ and $[SS]$. To derive these we can integrate equation (5.8) and the corresponding equations for $[II](r, dr)$ and $[SS](r, dr)$ from

$r = 0$ to $r = R$ to produce:

$$\begin{aligned}
\frac{d}{dt}[SI] &= \nu[II] - \nu[SI] + \lambda_{SI}[SS]\frac{[SI]}{[S]} - \lambda_{SI}[SI]\left(1 + \frac{[SI]}{[S]}\right) \\
\frac{d}{dt}[II] &= -\nu[II] + \lambda_{SI}[SI]\left(1 + \frac{[SI]}{[S]}\right) \\
\frac{d}{dt}[SS] &= +\nu[SI] - \lambda_{SI}[SS]\left(\frac{[SI]}{[S]}\right)
\end{aligned} \tag{5.18}$$

Also, the equation for the total number of infecteds $[I]$ is:

$$\frac{d}{dt}[I] = -\nu[I] + \lambda_{SI}[SI] \tag{5.19}$$

We note that the equations of motion for the edge variables have no dependence on the edge-length variables and so the network structure only comes in through λ_{SI} which in turn depends on R and k.

With our constraints (5.13) and (5.14) we can re-express $[SS]$ and $\psi_{SS}(r)$ in terms of other state variables, giving us four equations of motion in terms of both the edge variables $[SI]$ and $[II]$ and the edge-length variables $\psi_{SI}(r)$ and $\psi_{II}(r)$. The edge-length variables are most important because they contain information on the continuous spatial structure. In the next section we analyze this system of equations.

5.6 Comparison of Deterministic Results and Simulation Data

Figure 5.6 compares the steady-state predictions of the stochastic model and the deterministic approximation for various values of the recovery rate ν. Shown in table 5.1 is a comparison for other parameter choices. Surveying table 5.1, we see that the error is high in cases where edge length becomes important, i.e. when the value of k is high enough that the magnitude of the force of infection varies significantly between $r = 0$ and $r = R$. In general the agreement improves as R increases. This is because as Q increases, each node has more edges and so we move away from the regime where higher order correlations are significant and towards the regime where the mean field description applies. Interestingly, the agreement in case one is very good even though from figure 5.1 we know that the $\Omega_{SI}(m)$ distribution is poorly modelled by a Poisson distribution. The worst error occurs for case ten, which has both low Q and high k relative to R.

We conclude that in cases where k is large for a given R, and in cases where Q is low, the pair approximation can be inaccurate and caution must be used in these

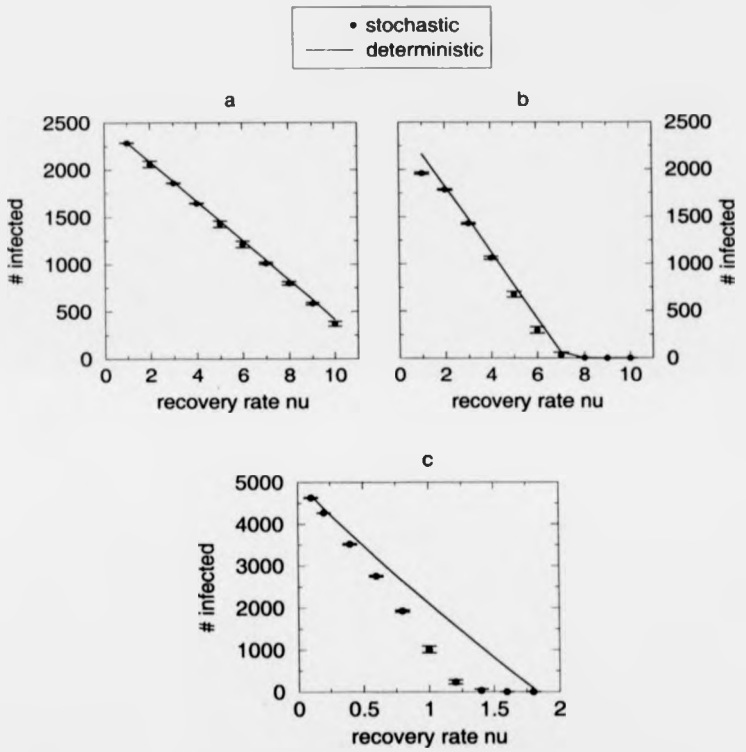


Figure 5.6: $[I]_\infty$ versus ν. Case a: $R = 5$, $\varrho = 2$, $N = 2500$, $k = 1$; Case b: $R = 2$, $\varrho = 2$, $N = 2500$, $k = 1$; Case c: $R = 1.5$, $\varrho = 1$, $N = 5000$, $k = 1.5$. The solid line denotes the deterministic prediction. Each circle is an average of eight stochastic runs, and each run is a running average after equilibrium has been reached. Initial conditions: 2000 infected. Total number of iterations = 500000. Error bars denote one standard deviation.

case	k	R	ν	ϱ	Q	$[I]_\infty$ det.	$\langle[I]\rangle$ stoch.	% err. [I]
1	0.5	5.0	7.0	1.0	78.5	3030.8	3013.0	0.6
2	2.5	5.0	0.4	1.0	78.5	2995.2	2356.7	27.1
3	0.5	5.0	21.0	3.0	235.6	3041.0	3020.9	0.7
4	2.5	5.0	1.2	3.0	235.6	3005.4	2822.9	6.5
5	0.5	1.5	1.6	1.0	7.1	2993.9	2682.4	11.6
6	2.5	1.5	0.3	1.0	7.1	3155.6	2324.5	35.8
7	0.5	1.5	4.5	3.0	21.3	3224.6	3181.4	1.4
8	2.5	1.5	1.0	3.0	21.3	3077.9	2875.6	7.0
9	1.5	1.5	0.5	1.0	7.1	3489.2	3145.5	10.9
10	3.0	1.0	0.12	1.0	3.1	3607.0	2008.1	79.6

Table 5.1: Comparison of $[I]$ at equilibrium for simulation data and pair approximation. N = 5000 in all cases, number of iterations = 500000, and there is only one stochastic run per case. The value of ν is varied between cases so that $[I]_\infty$ is approximately the same in each case; this allows for a better comparison.

cases if accuracy is an issue. However the pair approximation still serves fairly well for a wide range of parameters.

5.7 Analysis and Discussion of the Pair Approximation

As mentioned, the dynamics of the edge variables do not depend on the dynamics of the edge-length variables. This is because the distribution function $\psi_{SI}(r)$ which determines λ_{SI} is assumed to be approximately equal to $\psi(r)$ instead of using $\psi_{SI}(r)$. The dynamics of $[SI]$, $[II]$ and $[SS]$ are therefore not very interesting except insofar as λ_{SI} depends on k and R. The bifurcation diagrams for the edge variables do not tell us anything new (see figures 5.9 through 5.11).

On the other hand we find that edge-length variables can capture interesting behaviour found in the simulation data. Figures 5.7 and 5.8 show plots of $\psi_{SI}(r)$ against r which illustrate how the pair approximation correctly reproduces the non-linear dependence of $\psi_{SI}(r)$ on r. For each plot there is a linear regression to make visual estimation of the nonlinearity easier. This nonlinear dependence was interpreted as evidence of clustering in the simulation data of figures 5.1 through 5.4. Figures 5.7 and 5.8 also illustrate how clustering depends on parameters k and ν. To see this let r_d denote the value of r for which the deviation of $\psi_{SI}(r)$ from its linear regression is most negative. In figure 5.7 we see that higher k for a given value of R will shift r_d left and increase the overall deviation from a linear regression. This is

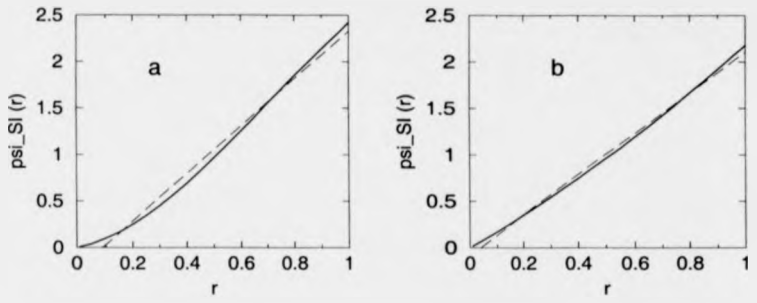


Figure 5.7: $\psi_{SI}(r)$ versus r showing effects of clustering for different values of k. Parameters for graph a are: $k = 4$, $\nu = 0.1$, $N = 5000$, $R = 1$, $\varrho = 1$. Parameters for graph b are: $k = 1$, $\nu = 0.2$, $N = 5000$, $R = 1$, $\varrho = 1$. The solid line is pair approximation data and dashed line is a linear regression.

because the decay of the force of infection kernel is stronger, therefore most of the deviation occurs where r is small. More interesting is the effect of varying ν on the shape of the plot of $\psi_{SI}(r)$ versus r. Figure 5.8 shows that when the final size of the epidemic is higher (i.e. for lower ν), the plot of $\psi_{SI}(r)$ versus r becomes *less* linear, indicating increased clustering. This seems at first to contradict the observations of the behaviour of spatial models studied thus far (in particular regular lattice models) which show that clustering increases for a lower final epidemic size on account of the development of long-range correlations near the critical point [89]. The unexpected behaviour might be due to some limitation of the pair approximation for the parameters we have chosen, particularly as we have chosen a low Q and a high k. However we feel this is not likely because in this case the PA deviation from linearity for high ν should be less, not more. A more likely explanation lies in the distinction between distance clustering exhibited here and the number clustering which is more pervasive near the critical point in network and spatial models. Distance clustering and number clustering can be quite independent processes, as figures 5.1 through 5.4 demonstrate. For future work simulations should be conducted to see if this property really exists and a more comprehensive analytical framework should be built to answer this and other questions. Also, we should investigate more thoroughly the relationship between the nonlinearity of $\psi_{SI}(r)$ and distance clustering.

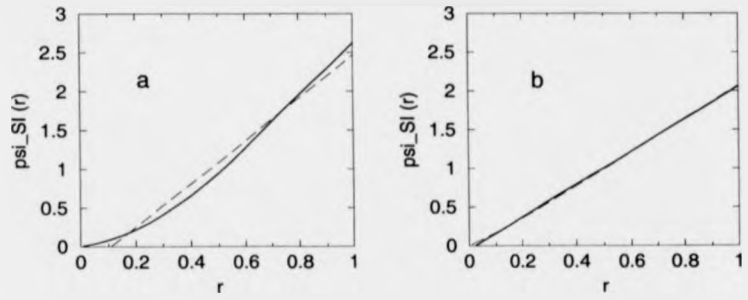


Figure 5.8: $\psi_{SI}(r)$ versus r showing effects of clustering for different values of ν. Parameters for graph a are: $k = 4$, $\nu = 0.03$, $N = 5000$, $R = 1$, $\varrho = 1$. $[I]_\infty = 4415$. Parameters for graph b are: $k = 4$, $\nu = 0.3$, $N = 5000$, $R = 1$, $\varrho = 1$. $[I]_\infty = 622$. The solid line is pair approximation data and dashed line is a linear regression.

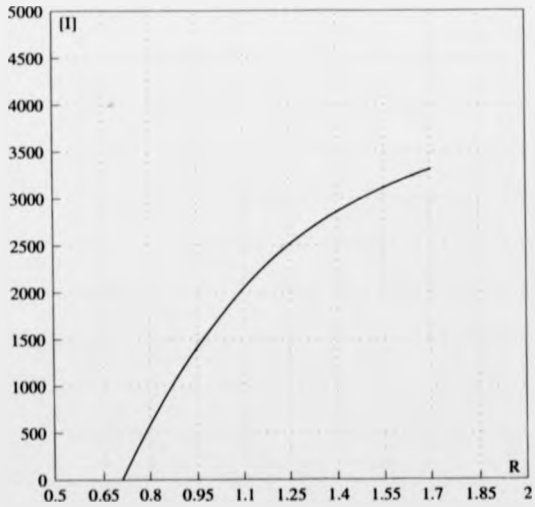


Figure 5.9: Bifurcation diagram of $[I]$ versus R. Other parameters are: $k = 1$, $N = 5000$, $\nu = 1$, $\varrho = 1$.

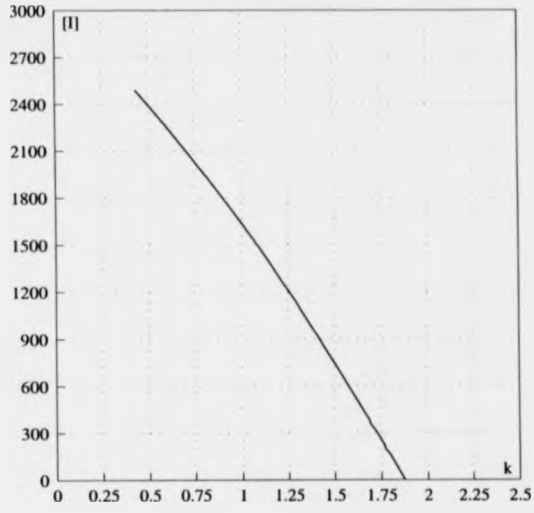


Figure 5.10: Bifurcation diagram of $[I]$ versus k. Other parameters are: $R = 1$, $N = 5000$, $\nu = 1$, $\varrho = 1$.

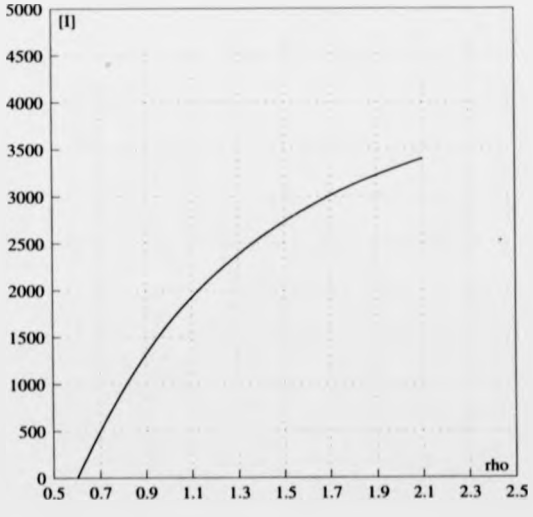


Figure 5.11: Bifurcation diagram of $[I]$ versus ϱ. Other parameters are: $R = 1$, $N = 5000$, $\nu = 1$, $k = 1$.

5.8 Modifications to the Pair Approximation

The error in the pair approximation comes from two sources: the assumption of conditional independence and the approximation of equation 5.17 imposing linearity of $\psi_{SI}(r)$ in r. Most of the error probably comes from the assumption of conditional independence, and could be partially removed by IPA, IVPA or HPA. The error which comes from assuming $\psi_{SI}(r)$ is linear could be removed by incorporating $\lambda_{SI}(r) = \int_0^R e^{-kr}\psi_{SI}(r)dr$ directly into the equations of motion, producing an integro-differential equation. A third approach could be to estimate $\psi_{SI}(r)$ more accurately than by assuming a linear distribution. The advantage of this is that we would have an ODE formulation instead of an integro-differential equation formulation.

The third option was implemented by trying to fit a curve to the simulation data. We found that data can be fitted quite well with a two-parameter function of the form $A(e^{Br} - 1)$. Figure 5.12 shows the data from the $\psi_{SI}(x)$ distributions in figures 5.1 through 5.4 fitted to such an equation. The curve-fitting formula seems to work well and is relatively simple, with only two parameters controlling the shape of the curve. Given the mean μ_{SI} of the $\psi_{SI}(r)$ distribution, it should be possible to calculate A and B by applying the constraints:

$$\begin{aligned} \int_0^R x\psi_{SI}(x)dx &= \mu_{SI} \\ \int_0^R \psi_{SI}(x)dx &= 1 \end{aligned}$$

However to do so we must solve transcendental equations of the form:

$$e^{BR}(BR - 1 - \mu_{SI}B) = \frac{1}{2}B^2R^2 - 1 - \mu_{SI}B - \mu_{SI}B^2R \qquad (5.20)$$

The solution can be arrived at through numerical recursion. and the results can be incorporated into the equations of motion, which are then analyzed numerically. The two sides of equation (5.20) are shown in figure 5.13.

We did not implement the approach of using a numerical solution of the transcendental equation for the pair approximation described in this chapter, but we did implement this approach in a deterministic approximation of the same continuous spatial model which tracks the time evolution of means μ_{SI}, μ_{II} and μ_{SS}. The convergence time was extremely large and thus we decided this scheme was too impractical to determine $\psi_{SI}(x)$. We also tried using a polynomial approximation $A(e^{Bx} - 1) \approx A(Bx + (Bx)^2/2 + (Bx)^3/6 + \cdots)$ to obtain explicit expressions for A and B when deviations are not large, however the improvements of this approximation were washed out by the errors caused by the assumption of conditional

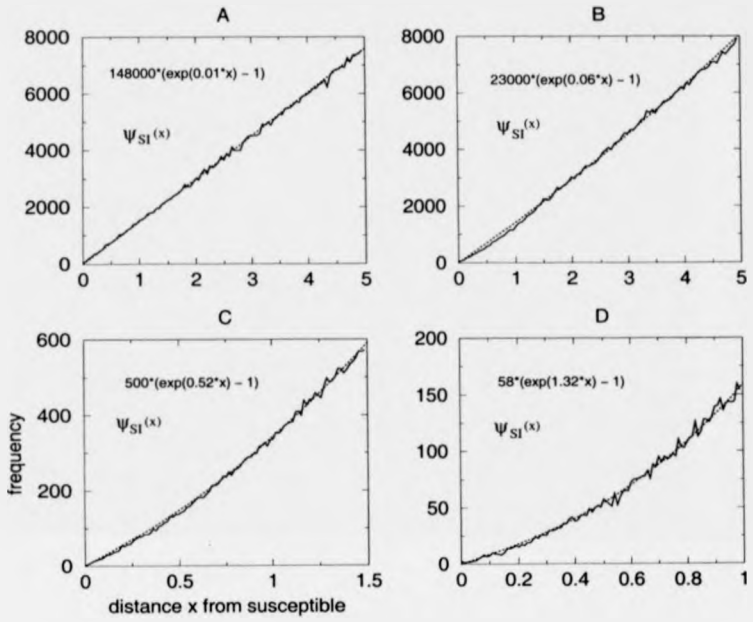


Figure 5.12: Curve-fitting for $\psi_{SI}(x)$. Cases A,B,C and D show data from figures 5.1, 5.2, 5.3 and 5.4 respectively. The dashed line is the curve-fit and the solid line is the simulation data.

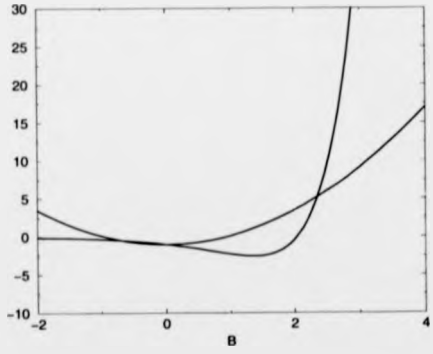


Figure 5.13: Solution branches (LHS and RHS) of equation 5.20. $\mu_{SI} = 1.01$, $R = 1.5$.

independence. Given the difficulties of the curve-fitting approach and the undesirability of introducing two extra parameters which themselves must be estimated, it is better in future simply to stick with the integro-differential approach combined with IPA, IVPA or HPA, if we want a more accurate model.

5.9 Discussion

In this chapter we derived a simple pair approximation for a randomly-distributed population on a plane with SIS infection dynamics. The infection dynamics are described by an exponentially decreasing force of infection kernel. The pair approximation tracked the time evolution edge and edge-length variables and was able to capture some interesting aspects of the stochastic model's epidemic dynamics, particularly the nonlinearity in $\psi_{SI}(r)$ which implies clustering of infecteds. Apparently, the stochastic model can exhibit two distinct kinds of clustering: distance clustering and number clustering. It is possible to have one without the other, depending on the decay of the kernel k, the radius of truncation R and the population density ϱ. Our pair approximation is designed to capture distance clustering only. We found that distance clustering actually increases as the density of infecteds increases, a surprising fact in light of the results about number clustering in other models which increases for a lower density of infecteds. Although we have suggested that the nonlinearity in $\psi_{SI}(r)$ versus r reflects distance clustering, more work must be done to establish this on a firm analytical foundation or reject it.

Although this pair approximation is a small step it shows that pair approximations might be applicable to modelling populations living in continuous space as well as populations living on networks and square lattices. Our pair approximation is simple and makes few assumptions, and yet can capture interesting behaviour. The advantages of developing pair approximations for continuous spatial models instead of using PDE models is that we retain an individual-based description and one can describe the neighbourhood of single individuals. In real ecologies of course the individual is the fundamental unit of measurement.

For future work it is important to understand in more depth the relationship between the moment closure approximations of Bolker, Gandhi et al. [12] [42] which take the limit from an individual-based model to a density-based model, and moment closures such as ours which retain an individual-based formulation. The two most important questions are: what is gained and/or lost in taking the limit, and what results from one type of model also apply to the other? Also in future we need to calculate R_0 for this continuous spatial model and compare it to the

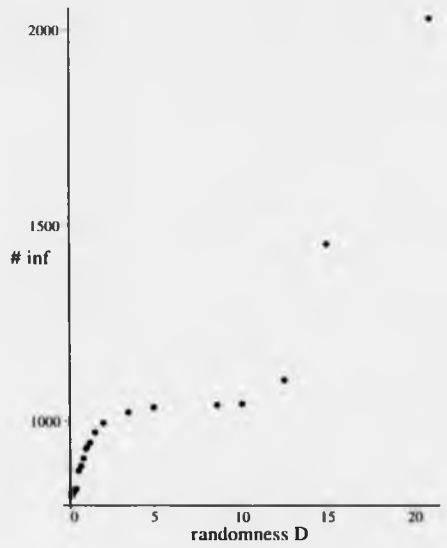


Figure 5.14: Plot of simulation data; $[I]$ at equilibrium versus spatial randomness D. $\nu = 0.7$, $k = 1.0$, $\rho = 0.2$, no. iterations $= 2000000$, $N = 2500$, d(grid size) $= 1$. Each point represents a long-time average from eight simulation runs.

expressions for R_0 in populations based on lattices. The stochastic data of figure 5.6 suggest that the intersection of the stable solution branches of $[I]_\infty$ versus ν should be smooth, unlike what is predicted for most moment closure approximations for network models which do not have a background force of infection.

Although we have not paid much attention to the distributions $\Omega_{SI}(k)$, $\Omega_{II}(k)$ and $\Omega_{SS}(k)$, an obvious extension would be to derive equations of motion with these quantities as state variables. This should be quite easy for cases of high Q since inspection of figures 5.1 through 5.4 suggests that the distributions are well-described by a normal distribution and so we can use the variance as a state variable. Also, as already discussed, it might be worthwhile to get more accurate estimates for $\psi_{SI}(r)$ than that of equation (5.17), and we should also cross-apply better pair approximations such as IVPA, IPA and HPA to this model.

Finally, models can be developed which vary the network structure in other ways. For instance, one could start with a square grid and place one individual at a randomly chosen location within a radius D of each node. Then we could introduce our infection kernels as before to produce an epidemic system. Thus we have a single parameter D which controls the amount of randomness in the spatial population.

Figure 5.14 shows an example of data we obtained from a simulation for such a model. For $D = 0$ we have a regular lattice and for large D we have a randomly-distributed population as in the model we have studied in this chapter. There is an increase in the final size of the epidemic as D is increased. Deterministic approximations to such models could be developed to study the dependence of epidemic characteristics on D, with applications to disease control in agriculture. Another approach to capturing continuous space using a modified cluster formulation is being developed by Kirkilionis and Bauch [8] [65].

In the final chapter we will tie in the results of this chapter with the results from the rest of the thesis.

Chapter 6

Conclusions

He said: "You're as clever as a monkey,
Juggling greedily with futile notions
And the luxury of doubt."
I became an idiot, impoverished, a loose end.

– Molana Jalal-e-Din Mohammed Molavi Rumi

6.1 Discussion

The main theme of this work is the extension of moment closure approximations to stochastic epidemic models on dynamic networks and in continuous space. This is in contrast to the usual application of MCA to regular static networks. We conclude that good moment closure approximations can be derived for these systems and can serve as valuable tools to understand the epidemiology of the systems under study. This is a fertile and interesting area of research and should be further pursued. Some others have also started to cross-apply MCA in new ways, for instance in continuous space epidemic models (e.g. Bolker and Pacala [13]). However this is the first work to apply MCA specifically to dynamic network models, which are especially useful for infectious diseases such as STDs. We have also designed a few new moment closure techniques in this thesis.

Because of the nature of our main theme of research, the body of results is multi-faceted. We have considered several different types of models which are united, for the most part, only by being different from conventional regular lattice models. However there are also several important sub-themes which run through all the chapters and serve to unite them. These include

1. the development of several new, more accurate and generally-applicable moment closure approximations;
2. the study of the accuracy of moment closure approximations for modelling invasion and their behaviour and accuracy near critical points;
3. the study of invasion in spatially structured populations;
4. the phenomenon of clustering and its effect on epidemics.

We have limited ourselves by considering only SIS epidemic models. However much of what we have learned in this thesis applies to ecological models generally, such as our discussion of spatialized invasion and the new moment closure approximations. Non SIS types of infection can, as discussed in chapter 1, produce very different dynamics from SIS infections such as oscillatory behaviour. Also many real diseases are not adequately described by an SIS process. However our major goal in this thesis was not to find novel behaviour or predict the course of real diseases, but rather to answer more theoretical questions about the usefulness of MCA. We also limited ourselves by preferring numerical treatment over pen-and-paper analysis. Although the possibilities for more rigorous analytical treatment were not explored in as much depth as they could have been, in most cases the complexity of the models would have made this approach less useful than numerical analysis.

We found that the new moment closure approximations, particularly IVPA, often have significantly improved accuracy. We also found that for some systems, such as sexual partnership networks where the number of neighbours Q per person is low, it is essential to go to a triple approximation to get good agreement, or to find much better pair approximations than the ones which are currently in use. Applying moment closure approximations to dynamic network models presents extra difficulty. Which MCA we choose depends on the interaction of the two dynamical processes of disease transmission and partnership network formation, a complicating factor not present in static regular lattice models. This complication was illustrated in the pair approximations for the concurrency model where different pair approximations applied in different parameter regimes. However the extra effort necessary for deriving MCA for dynamic networks is worthwhile since a mean-field or a static network description is inaccurate for many diseases. This is especially true when the transmission occurs at timescales similar to the partnership dynamics or when repeated contacts are an important features of the social network structure. Another apparent difference between the two types of models is the nature of invasion. In our network models clustering did not interfere with the spread of the disease as it would on a regular lattice because of the lack of spatially localized patches of infecteds in the network model (i.e. no triangles and squares). However both the concurrency model and the steady/casual model exhibit the initial establishment of a pseudo-equilibrium which determines whether or not the disease can invade, and which can be exploited to calculate R_0.

Our approximations for STD network models generally reinforced the hypothesis that casual partnerships and concurrency, even in small amounts, significantly impact STD growth rate and final size $[I]_\infty$, and thus require special attention in efforts for disease control and eradication. However we found that in some extreme cases the moment closure approximations predicted behaviours contrary to what was expected, such as a decrease in final size with an increase in concurrency. The curves we ended up with in the plots of $[I]_\infty$ versus κ_3 and $[I]_\infty$ versus f were standard epidemic curves, exhibiting a transcritical bifurcation and with a steep slope for the nontrivial stable solution. In fact one criticism of our results might be that we have not learned anything new since we have ended up with the same standard epidemic curve for $[I]_\infty$ versus transmissibility. However the significance of our results lies in the parameters which are varied in order to produce such curves. The parameters we vary have to do with network structure, and it is not clear *a priori* how $[I]_\infty$ depends on them. With this model we discover there is a threshold in STD models which depends not on population density in a mean-field treatment, as in Hamer's

model for malaria transmission, but on the density-independent network structure. This is a new insight.

Our preliminary research on continuous spatial models showed some interesting results, with a simple pair approximation capturing the effect of clustering found in the simulation data. Also, we distinguished between distance clustering and number clustering, and found that sometimes one can be present without the other. However the research is too undeveloped to draw any concrete conclusions about the usefulness of moment closure approximations in this area. Additionally there is already a well-established body of research on the same topic using PDE-based models [13][80], and the mathematical theory of PDEs is more well-developed than that for ODEs. And a PDE approximation to individual-based models can often be found. So the usefulness of MCA network models for capturing aspects of continuous spatial structure is less clear.

6.2 Further Research

A common criticism of pair approximations is that they are just too complicated for practical use and will not gain wide popularity among theoretical epidemiologists. There is some truth in this, particularly if one considers the mathematical bulkiness of OTA or the involved derivations of HPA and IVPA. Accurate, simple and robust closures are a 'holy grail', in the sense that it would make MCA more widely useful and accessible (and perhaps also in the sense that they might not exist). However in the long run, as mathematics becomes more and more a familiar tool in the biological sciences, arguments against pair approximations on grounds of their complexity will gradually lose ground. Even simple OPA models can provide significant improvements over mean-field approximations for many ecological systems. Also, much practical modelling does not require a heavy mathematical background; often user-friendly software is used where one simply has to specify the number and type of compartments and the transition rates between them. It would be easy to write similar software for pair approximations, since the meaning of the state variables and model parameters in pair approximations is easily understood and models can be readily defined. In many cases the gain in realism of the deterministic approximation would be considerable.

Notwithstanding their greater complexity, moment closure approximations continue to present opportunities for interesting research. MCA especially for dynamic networks should be further studied. The rise of network models and the recognition of their usefulness in epidemic modelling has stimulated a need for new analytical

frameworks to help us interpret what is happening in the simulations. The requirement for good analytical frameworks for network models is even more important than for regular lattices on account of their greater diversity and generality. Along this vein, more research could be done to apply existing results and terminology from graph and network theory to MCA for the dynamic networks we have studied. This will speed formulation of good moment closure approximations and will clarify its presentation.

Appendix A

Abbreviations

The following abbreviations are used throughout the thesis.

MCA	moment closure approximation
OPA	ordinary pair approximation
IPA	improved pair approximation
IVPA	invasory pair approximation
HPA	hybrid pair approximation
HEPA	heurstic pair approximation
OTA	ordinary triple approximation
PDE	partial differential equation
RDE	reaction-diffusion equation
ODE	ordinary differential equation
IPS	interacting particle system
CA	cellular automata
CP	contact process
SEIR	susceptible-exposed-infected-removed
SIR	susceptible-infected-removed
SIS	susceptible-infected
PDF	probability density function
PF	probability function

Appendix B

Software and Computation

The numerical analysis of deterministic approximations was done with CONTENT 1.5 and Maple V.

Simulations were written in C and were run on SGI O_2 workstations, a Sun Ultra 60 workstation and an Origin 2000 supercomputer. In the case of the contact process and the pair approximation of chapter 5, the space in which the population is embedded always maps toroidally at the edges.

Graphics were produced using XMGR, Maple V and CONTENT 1.5, and gui software for simulations was written by Keith Yates in C.

Bibliography

[1] Helen Abbey. An examination of the reed-frost theory of epidemics. *Hum. Biology*, 24:201 – 233, 1952.

[2] M. Altmann. Susceptible-infected-removed epidemic models with dynamic partnerships. *Journal of Mathematical Biology*, 33:661–675, 1995.

[3] M. Altmann. The deterministic limit of infectious disease models with dynamic partners. *Mathematical Biosciences*, 150:153–175, 1998.

[4] R.M. Anderson and R.M. May. Population biology of infectious diseases i. *Nature*, 280:361 – 367, 1979.

[5] R.M. Anderson and R.M. May. Population biology of infectious diseases ii. *Nature*, 280:455 – 461, 1979.

[6] R.M. Anderson and R.M. May. *Infectious Diseases of Humans*. Oxford Science Publications, OUP, 1991.

[7] Hakan Andersson. Epidemic models on graphs and lattices: A short survey. *Manuscript*, 1998.

[8] Chris Bauch and Markus Kirkilionius. Cluster approximations for infection processes on a perturbed lattice, 2000.

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

[10] Daniel Bernoulli. Essai d'une nouvelle analyse de la mortalite causee par la petite verole et des avantages de l'incubation pout la prevenir. *Mem. Math. Phys. Acad. R. Sci. Paris*, pages 1 – 45, 1760.

[11] H.A. Bethe. *Proceedings of the Royal Society of London*, A150:552 +, 1935.

[12] Benjamin Bolker and Stephen W. Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, 52:1 – 19, 1997.

[13] Benjamin M. Bolker and Stephen W. Pacala. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *The American Naturalist*, 153(6):575 – 602, June 1999.

[14] J. Brownlee. Statistical studies in immunity: The theory of an epidemic. *Proceedings of the Royal Society of Edinburgh*, 26:484 – 521, 1906.

[15] D.M. Burley. Closed form approximations for lattice systems. In C. Domb and M.S. Green, editors, *Phase Transitions and Critical Phenomena*, pages 329 – 374. Academic Press, London, 1972.

[16] Stephen E. Chick, Andrew L. Adams, and James S. Koopman. Analysis and simulation of a stochastic, discrete-individual model of std transmission with partnership concurrency. *Mathematical Biosciences*, 166:45–68, 2000.

[17] Andrew Cliff and Peter Haggett. *Atlas of Disease Distributions: Analytic Approaches to Epidemiological Data*. Oxford: Basil Blackwell, 1988.

[18] James J. Collins and Carson C. Chow. It's a small world. *Nature*, pages 409–410, June 1998.

[19] K.L. Cooke and J.A. Yorke. Some equations modelling growth processes and gonorrhea epidemics. *Mathematical Biosciences*, 16:75 – 101, 1973.

[20] R.F. Costantino, R.A. Desharnais, J.M. Cushing, and B. Dennis. Chaotic dynamics in an insect population. *Science*, 275(5298):389–391, 1997.

[21] J.T. Cox and Richard Durrett. Limit theorems for the spread of epidemics and forest fires. *Stochastic Processes and their Applications*, 30:171 – 191, 1988.

[22] D.J. Daley and J. Gani. *Epidemic Modelling:An Introduction*. Cambridge University Press, 1999.

[23] Odo Diekmann, Klaus Dietz, and J.A.P. Heesterbeek. The basic reproduction ratio for sexually transmitted diseases. theoretical considerations. *Mathematical Biosciences*, 107:325–339, 1991.

[24] Odo Diekmann, Hans Heesterbeek, and Hans Metz. The legacy of kermack and mckendrick. In Denis Mollison, editor, *Epidemic Models*, pages 95 – 115. CUP, 1995.

[25] Odo Diekmann and J.A.P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases*. Wiley Series in Mathematical and Computational Biology. John Wiley & Sons, 2000.

[26] Odo Diekmann, J.A.P. Heesterbeek, and J.A.J. Metz. On the definition and the calculation of the basic reproduction ratio R_0 in models for infectious disease in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382, 1990.

[27] Klaus Dietz and K.P. Hadeler. Epidemiological models for sexually transmitted diseases. *Journal of Mathematical Biology*, 26:1 – 25, 1988.

[28] Klaus Dietz and D. Tudor. Triangles in heterosexual HIV transmission:. *AIDS Epidemiology: Methodological Issues*, pages 143–155, 1992.

[29] Richard Durrett. Spatial epidemic models. In Denis Mollison, editor, *Epidemic Models*, pages 187 – 201. CUP, 1995.

[30] W.J. Edmunds, C.J. O'Callaghan, and D.J. Nokes. Who mixes with whom? a method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proceedings of the Royal Society B*, 264:949–957, 1997.

[31] S.P. Ellner, A. Sasaki, Y. Haraguchi, and H. Matsuda. Speed of invasion in lattice population models: Pair-edge approximation. *Journal of Mathematical Biology*, 36:469 – 484, 1998.

[32] P.D. En'ko. On the course of epidemics of some infectious diseases. *Vrach.*, X:1008 – 1010, 1039 – 1042, 1061 – 1063, 1889.

[33] Elizabeth Pisani et al. Report on the global HIV/AIDS epidemic. Technical report, UNAids, 2000.

[34] J. Cleland et al. Looking deeper into the HIV epidemic: A questionnaire for tracing sexual networks. Technical report, UNAids, 1998.

[35] Jeanette Rodrigues et al. Risk factors for HIV infection in people attending clinics for sexually transmitted diseases in india. *British Medical Journal*, 311:283 – 286, July 1995.

[36] Sheldon Cohen et al. Social ties and susceptibility to the common cold. *Journal of the American Medical Association*, 277(24):1940 – 1944, 1997.

[37] W. Farr. Progress of epidemics. In *Second Report of the Registrar General for England*, 1840.

[38] J.A.N. Filipe. Hybrid closure-approximation to epidemic models. *Physica A*, 266:238 – 241, 1999.

[39] J.A.N. Filipe. personal communication, 2000.

[40] J.A.N Filipe and G.J. Gibson. Comparing approximations to spatio-temporal models for epidemics with local spread. submitted.

[41] J.A.N. Filipe and G.J. Gibson. Studying and approximating spatio-temporal models for epidemic spread and control. *Phil. Trans. R. Soc. Lond. B*, 353:2153–2162, 1998.

[42] Amar Gandhi, Simon Levin, and Steven Orszag. Moment expansions in spatial ecology models and moment closure through gaussian approximation. *Bulletin of Mathematical Biology*, 62:595 – 632, 2000.

[43] Geoff Garnett. The natural history of syphilis: Implications for the transmission dynamics and control of infection. *Sexually Transmitted Diseases*, 24(4):185–200, April 1997.

[44] Azra Ghani. *Sexual Partner Networks and the Epidemiology of Gonorrhoea*. PhD thesis, University of Oxford, December 1996.

[45] Azra Ghani. The role of sexual partnership networks in the epidemiology of gonorrhoea. *Sexually Transmitted Diseases*, 24:227–238, 1997.

[46] Paul Glendinning. Island chain models and gradient systems. *Journal of Mathematical Biology*, 32:171 – 178, 1994.

[47] Z. Grossman. Oscillatory phenomena in a model of infectious diseases. *Theoretical Population Biology*, 18:204 – 243, 1980.

[48] W.H. Hamer. Epidemic disease in england – the evidence of variability and persistency of type. *Lancet*, ii:733 – 739, 1906.

[49] T.E. Harris. Contact interactions on a lattice. *Ann. Prob.*, 2:969, 1974.

[50] Hans Heesteerbeek. R_0. PhD thesis, Centrum voor Wiskunde en Informatica, Amsterdam, 1992.

[51] Herbert W. Hethcote and Pauline van den Driessche. An sis epidemic model with variable population size and a delay. *Journal of Mathematical Biology*, 34:177 – 194, 1996.

[52] Herbert W. Hethcote and Pauline van den Driessche. Two sis epidemiologic models with delays. *Journal of Mathematical Biology*, 40:3 – 26, 2000.

[53] F.C. Hoppensteadt. An age dependent epidemic model. *Journal of the Franklin Institute*, 297:325 – 333, 1974.

[54] M. Huston, D.L. DeAngelis, and W. Post. New computer models unify ecological theory. *Bioscience*, 38:682 – 691, 1988.

[55] Olivia P. Judson. The rise of the individual-based model in ecology. *Trends in Ecology and Evolution*, 9:9 – 14, 1994.

[56] M. Katori and N. Konno. Upper bounds for survival probability of the contact process. *Journal of Statistical Physics*, 63:115 – 130, 1991.

[57] Matthew Keeling and Bryan Grenfell. Disease extinction and community size: Modeling the persistence of measeles. *Science*, 275:65–67, 1997.

[58] Matthew Keeling, David Rand, and Andrew Morris. Correlation models for childhood epidemics. *Proceedings of the Royal Society B*, 264:1149 – 1156, 1997.

[59] Matthew J. Keeling. Correlation equations for endemic diseases. *Proceedings of the Royal Society of London B*, 266:953–960, 1999.

[60] Matthew J. Keeling. The effects of local spatial structure on epidemiology invasions. *Proceedings of the Royal Society B*, 266:859–867, 1999.

[61] M.J. Keeling and D.A. Rand. Two methods to model the persistence of measles in a population. manuscript, 1996.

[62] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics – i. *Proceedings of the Royal Society*, 115A:700 – 721, 1927.

[63] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics – i. *Bulletin of Mathematical Biology*, 53:33 – 55, 1991.

[64] Ryoichi Kikuchi. A theory of cooperative phenomena. *Physical Review*, 81(6):988 – 1003, 1950.

[65] Markus Kirkilionis, Chris Bauch, and David Rand. Cluster dynamics: Population models with stationary individuals. in preparation, 2000.

[66] M. Kretzschmar, J.C. Jaeger, D.P. Reinking, G. van Zessen, and H. Brouwers. The basic reproduction ratio R_0 for a sexually transmitted disease in a pair

formation model with two types of pairs. *Mathematical Biosciences*, 124:181–205, 1994.

[67] Mirjam Kretzschmar. Deterministic and stochastic pair formation models for the spread of sexually transmitted diseases. *Journal of Biological Systems*, 3:789–801, 1995.

[68] Mirjam Kretzschmar and Martina Morris. Measures of concurrency in networks and the spread of infectious disease. *Mathematical Biosciences*, 133:165–195, 1996.

[69] Mirjam Kretzschmar, Yvonne van Duynhoven, and Anton Severijnen. Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *American Journal of Epidemiology*, 144:306–317, 1996.

[70] S.A. Levin and R. Durrett. From individuals to epidemics. *Philosophical Transactions of the Royal Society of London B*, 351:1615 – 1621, 1996.

[71] Simon A. Levin. The problem of pattern and scale in ecology. *Ecology*, 73(6):1943–1967, 1992.

[72] A.L. Lloyd and R.M. May. Spatial heterogeneity in epidemic models. *Journal of Theoretical Biology*, 179:1 – 12, 1996.

[73] Hirotsugo Matsuda, Naofumi Ogita, Akira Sasaki, and Kazunori Sato. Statistical mechanics of population. *Progress in Theoretical Physics*, 88:1035–1049, 1992.

[74] Denis Mollison. Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society B*, 39:283 – 326, 1977.

[75] Andrew J. Morris. *Representing Spatial Interactions in Simple Ecological Models*. PhD thesis, University of Warwick, September 1997.

[76] M. Morris, D. Serwadda, M. Kretzschmar, N. Sewankambo, and M. Wawer. Concurrent partnerships and HIV transmission in uganda. In *XI International Conference on AIDS*, July 1996.

[77] Martina Morris. Sexual networks and HIV. *AIDS*, 11:S1 – S8, 1997.

[78] Martina Morris and Mirjam Kretzschmar. Concurrent partnerships and the spread of HIV. *AIDS*, 11(5):1–7, April 1997.

[79] J. Mueller, M. Kretzschmar, and K. Dietz. Contact tracing in stochastic and deterministic epidemic models. *Mathematical Biosciences*, 164:39–64, 2000.

[80] James D. Murray. *Mathematical Biology*. Springer-Verlag, 1989.

[81] R.S. Nord and J.W. Evans. Irreversible immobile random adsorption of dimers, trimers,...on 2d lattices. *Journal of Chemical Physics*, 82(6):2795 – 2810, 1985.

[82] Steven A. Orszag. Analytical theories of turbulence. *Journal of Fluid Mechanics*, 41:363 – 386, 1970.

[83] Mercedes Pascual and Simon A. Levin. From individuals to population densities: Searching for the intermediate scale of nontrivial determinism. *Ecology*, 80(7):2225–2236, 1999.

[84] David Rand. Correlation equations and pair approximations for spatial ecologies. In J. McGlade, editor, *Advanced Ecological Theory. Principles and Applications*, pages 100–142. Blackwell Science, 1999.

[85] David Rand, Matthew Keeling, and H. Wilson. Invasion, stability and evolution to criticality in spatially extended host-pathogen systems. *Proc. R. Soc. Lond. B*, 259:55–63, 1995.

[86] R. Ross and H.P. Hudson. An application of the theory of probabilities to the study of a priori pathometry, part iii. *Proceedings of the Royal Society of London A*, 436:225 – 240, 1917.

[87] Ronald Ross. *The Prevention of Malaria*. Murray, London, 1909.

[88] Kazunori Sato, Hirotsugu Matsuda, and Akira Sasaki. Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology*, 32:251–268, 1994.

[89] Dietrich Stauffer and Amnon Aharony. *Introduction to Percolation Theory*. London: Taylor & Francis, 1992.

[90] Minus van Baalen and David Rand. The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology*, 193:631–648, 1998.

[91] V. Volterra. Variazionie fluttuazioni del numero d'individui in specie animali conviventi. *Mem. Acad. Lincei.*, 2:31 – 113, 1926.

[92] Charlotte H. Watts and Robert M. May. The influence of concurrent partnerships on the dynamics of HIV/AIDS. *Mathematical Biosciences*, 108:89–104, 1992.

[93] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, pages 440–442, June 1998.

[94] E.B. Wilson and M.H. Burke. The epidemic curve. *Proc. Nat. Acad. Sci., Washington*, 28:361 – 367, 1942.

The gods did not reveal from the beginning
All things to us; but in the course of time
Through seeking, men found that which is better.
But as for certain truth, no man has known it,
Nor will he know it; neither of the gods,
Nor yet of all the things of which I speak.
And even if by chance he were to utter
The final truth, he would himself not know it;
For all is but a woven web of guesses.

– Xenophanes, c. 570 - c. 480 B.C.E.