THE UNIVERSITY OF
WARWICK

AUTHOR: **Samuel Paul Clark Brand** DEGREE: **Ph.D.**

TITLE: **Spatial and Stochastic Epidemics: Theory, Simulation and Control**

DATE OF DEPOSIT: .....

I agree that this thesis shall be available in accordance with the regulations governing the University of Warwick theses.

I agree that the summary of this thesis may be submitted for publication.

I **agree** that the thesis may be photocopied (single copies for study purposes only).

Theses with no restriction on photocopying will also be made available to the British Library for microfilming. The British Library may supply copies to individuals or libraries. subject to a statement from them that the copy is supplied for non-publishing purposes. All copies supplied by the British Library will carry the following statement:

> "Attention is drawn to the fact that the copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's written consent."

AUTHOR'S SIGNATURE: .....

USER'S DECLARATION

1. I undertake not to quote or make use of any information from this thesis without making acknowledgement to the author.
2. I further undertake to allow no-one else to use this thesis while it is in my care.

DATE SIGNATURE ADDRESS

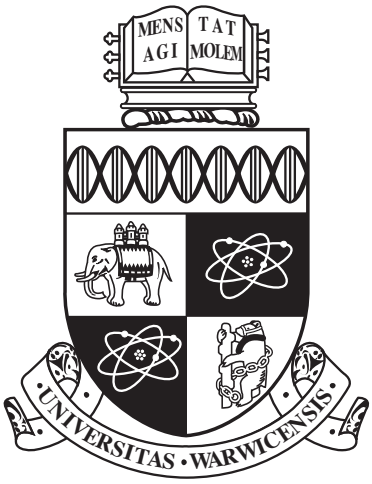
.....

.....

.....

.....

.....



Spatial and Stochastic Epidemics: Theory, Simulation and Control

by

Samuel Paul Clark Brand

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Centre for Complexity Science

September 2012

THE UNIVERSITY OF
WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

I would predominantly like to thank my supervisor Matt Keeling for all his guidance over the last few years. I wouldn't be here without him! Especial thanks go to everyone in the Complexity Science Centre, particularly Jamie, Quentin and Anthony for putting me up at their homes on numerous occasions.

And thanks to everyone in the WIDER research group for continual support and always providing a forum for me to discuss ideas.

Finally, thanks to my wonderful wife who has put up with a lot these past few years, and especially these last few months.

This work was supported by the Engineering and Physical Sciences Research Council as part of the University of Warwick Complexity Science Doctoral Training Centre.

Declarations

This work has been composed by myself and has not been submitted for any other degree or professional qualication.

The content for chapter 4 is in preparation for submission under the title, "A Fast Spectral Method for Spatial and Stochastic Epidemic Simulation".

Abstract

It is now widely acknowledged that spatial structure and hence the spatial position of host populations plays a vital role in the spread of infection. In this work I investigate an ensemble of techniques for understanding the stochastic dynamics of spatial and discrete epidemic processes, with especial consideration given to SIR disease dynamics for the Levins-type metapopulation.

I present a toolbox of techniques for the modeller of spatial epidemics. The highlight results are a novel form of moment closure derived directly from a stochastic differential representation of the epidemic, a stochastic simulation algorithm that asymptotically in system size greatly out-performs existing simulation methods for the spatial epidemic and finally a method for tackling optimal vaccination scheduling problems for controlling the spread of an invasive pathogen.

Chapter 1

Introduction to Modelling Infectious Diseases

"The Epidemiologist's unit is not a single human being, but an aggregate of human beings"

- Major Greenwood (Greenwood 1935)

1.1 Motivation and Aims

Disease has plagued mankind throughout recorded history, from biblically recorded plagues through to devastating epidemics such as Black Death in Europe (1347-50 and subsequent recurrences), the long persistence of smallpox and cholera and periodic intense influenza pandemics (the 1918-1920 being a particularly infamous example). Nor is the impact of disease on mankind contained to the direct epidemiology of humans. Diseases amongst commercial livestock, with the UK Foot-and-mouth disease (FMD) outbreak in 2001 being a notorious first-world example, can have a significant economic impact. In addition to FMD, bovine tuberculosis (bTB) remains extant in the UK with significant on-going control costs and blue tongue virus (BTV) has made a number of incursions into European farms since 2006. Both viral and fungal disease has had a considerable impact on commercially and environmental important plants. Important examples being Citrus Tristeza virus (CTV) spreading amongst citrus trees, sudden oak death which can effect a wide range of trees and stem rust which is a significant disease of cereal crops.

A common feature of the epidemiology of the disease examples mention above is that they are *infectious*. Whilst exogenous and environmental factors remain important fundamentally the disease-causing pathogen spreads from host to host each time recruiting a new individual (Human, animal or plant) to the epidemic. It has been recognised at least since the work of Ross, Kermack and McKendrick (Ross 1916; Kermack and McKendrick 1927) in the early twentieth century that the spread of infectious pathogens lent themselves to a mechanistic or model-based interpretation of their progression through time.

The relationship between models written in the language of mathematics and epidemiology has only strengthened since this early pioneering work. When the individuals carrying the infectious pathogen are approximately interchangeable and mix homogeneously amongst themselves then both deterministic and stochastic models for the dynamics of the epidemic have become classical (Bartlett 1956; Bailey 1975; Anderson and May 1992; Diekmann and Heesterbeek 2000; Keeling and Rohani 2008). In this thesis I will be concerned with varying away from this solid theoretical foundation by considering the disease-bearing population as being separated into discrete sub-populations each with a spatial habitat. I will assume more frequent mixing between proximate populations decaying to zero mixing between very distant populations. The assumption of spatially embedded dynamics leads to natural association with concepts of theoretical ecology, in particular the metapopulation concept of local populations in a fragmented landscape (Hanski and Gaggiotti 2004).

In addition to focusing on spatially structured epidemics I also focus on the SIR paradigm of disease (Kermack and McKendrick 1927; Anderson and May 1992). Each individual is assigned a discrete disease state or 'compartment' (Keeling and Rohani 2008), whether susceptible to the disease (S), currently infectious (I) or removed (R) from the epidemic in both the sense of personal risk of infection and contribution to recruiting new infected individuals to the epidemic via infectious transmission. The progression for the individual, or population if the Levins assumption is made, is unidirectional in time $S \rightarrow I \rightarrow R$. The SIR framework offers the ideal framework for investigating epidemic models where disease is non-repeating either through long lasting immunity for individuals or the effect of control for populations.

1.2 Outline of this Thesis

"Time is what stops everything happening at once. Space is what stops everything happening to me!"

- John Wheeler

I identify four key properties for a mathematical model of a real-world phenomenon that are important to the theoretical investigator,

- *The model should be at least partially theoretically understandable.* A model which represents a 'black box' does not promote understanding of the phenomena of interest. The investigator should have some theoretical expectation of the generic features of interest, and an idea of what the model is trying to predict. For models of epidemics the expectation of a threshold phenomenon and the concept of the reproductive ratio provide a unifying framework for otherwise rather disparate models of disease.

- *The model should be amenable to simulation/numerical solution.* For the modern investigator *in silico* methods have become invaluable for investigating the consequences of model selection beyond the limits of theoretical considerations. For epidemic models even the earliest deterministic models were in a form that required numerical solution. Modern theoretical epidemiology often trades in models involving millions of individuals and which include essential randomness. Numerical investigation requires large scale Monte Carlo simulation that push the limit of even modern computing resources.

- *The model should have understandable feedback properties.* Often the goal of the modeller is not just accurate prediction, but also to use her prediction to guide intervention. For epidemiologists there has been a long history of model-based disease management, with guiding vaccination being a classic example.

- *The model should be parametrisable by the available data.* Although the modeller might desire to include an every increasing set of features into her model, care should be taken as to whether the available data set can support a more complex model. For epidemic modelling this has been a consistently thorny issue; in particular for the emergence of a novel pathogen it might be impossible to produce a good model until substantially into the bulk of the outbreak.

The essential goal of this thesis is to present a tool-box of novel methods for a class of SIR epidemic model that focuses on the spatial properties of disease transmission. For this model I will be tacking the first three model properties listed above. I have left the hugely significant field of appropriate epidemic model parametrisation for future work. In each chapter my arguments and justifications have been given in mathematical form. However, in a number of cases this has only lead to convincing arguments rather than rigorous arguments. My general methodology throughout has been to use the definite underlying mathematical properties of the spatial epidemic in order to suggest novel applications. In some cases these applications are supported by simulation based evidence rather than proper mathematical rigour, which remains future work. I feel it is important to make this clear to the reader.

In **chapter 2** I will review some theoretical results from the calculus of pure jump processes, in particular the form of Itô lemma (a chain rule for stochastic processes) for pure jump processes. I will also introduce the basic spatial and stochastic model that I will then use throughout the rest of this work. The basic spatial model will be shown to admit representation as a jump process with an explicit stochastic differential equation (SDE) formulation. This formulation will be exploited in order to give a set of ODEs approximating the expected dynamics of the spatial epidemic.

In **chapter 3** the stochastic differential model for the spatial epidemic will be extended to include the case where the spatial locations of the habitats are themselves random variables. I use the classic approach of spatial moment equations to investigate the interplay between the stochastic dynamics of the infectious disease and the random, possibly spatially clustered, environment of habitat locations. The novelty of my approach lies in building the moment equations directly from the stochastic microscopic dynamics, which leads to a slightly corrected version of the moment equations compared to some extant in the literature. I will also give a theoretical and numerical investigation into the performance of different types of 'closing' approximations for the moment equation I derive.

Chapter 4 will be largely concerned with the Monte Carlo simulation of the spatial epidemic model. I will give a review of the stochastic simulation algorithms that exist in the literature and which are used throughout this thesis. For the spatial epidemic model used the simulation of even a single epidemic can become computationally burdensome using existing simulation algorithms when the size of the metapopulation becomes significant. This will be motivation for development

of a novel simulation algorithm tailored to the spatial epidemic which has excellent numerical efficiency compared to traditional stochastic simulations techniques as the population size grows large.

Finally, in **chapter 5** I will review the classic dynamic programming approach to solving continuous-time or multi-stage decision problems. I will introduce an optimal vaccination scheduling problem and demonstrate that for certain scenarios the space of optimal vaccination actions at any time is dramatically reduced to a small set. I will solve the optimal control problem to numerical exactness for toy problem, with a focus on the effect on optimality of the vaccinating agency lacking information about the epidemic. I also introduce a novel method for approximately solving optimal vaccination problems for the spatial epidemic, inspired by the field of approximate dynamic programming. I will demonstrate this method on an epidemic of sufficient size that classical dynamic programming would be numerically infeasible.

1.3 Historical context and the SIR model of Disease

"... the principal problems of epidemiology on which preventive measures largely depend, such as the rate of infection, the frequency of outbreaks, and the loss of immunity, can scarely ever be resolved by any other methods than those of analysis."

- Sir Ronald Ross 1916 (Ross 1916).

Ross[1] in his ground breaking work on the population level dynamics of infectious diseases, or pathometry in his terminology, expressed his surprise at the dearth of mechanistic *a priori* methodology for the study of infectious diseases. His solution to this lack was to develop what he referred to as the 'theory of happenings'; that the incremental change in the population levels of the 'affected part' and the 'unaffected part' over a short time interval are proportional to a rate and the size of the time interval. The rates he conceived were nativity, mortality, immigration and emigration for the overall population and for the disease dynamics a 'happening element' and a 'reversion element', which would now be referred to as respectively as transmission rate and recovery rate. Ross also embarked on a discussion of the type of recursion; whether the reverting individuals had gained long-term immunity

[1]Second noble prize winner in Physiology and Medicine (1902) for the identification of the mosquito as the vector for malaria.

to subsequent infection or if immunity was waning and whether they were capable of spreading further infection after reversion.

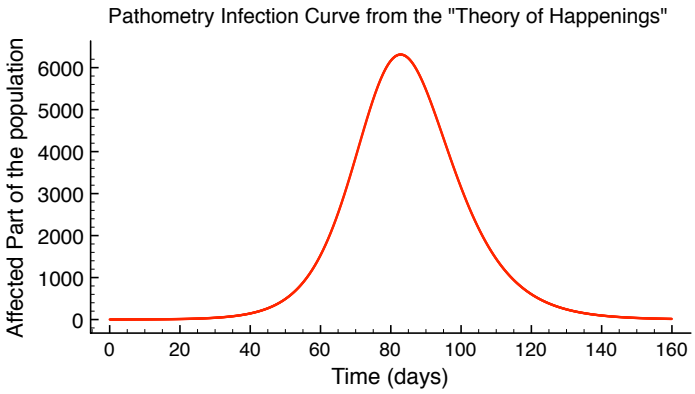


Figure 1.1: A graph in honour of Sir Ronald Ross. The numerical solution to the epidemic curve (infecteds number) according to his theory of happenings for when reversion (recovery) gives permanent immunity (SIR model) and there are no demographic effects. Population size is 10^5, the rate of happenings (transmission rate) is $0.375 \times 10^{-5} \times$ number of unaffected (susceptibles) $\times$ number of affected (infecteds) per day. The reversion rate (recovery) is $0.25 \times$ number of affected per day.

In this seminal work Ross identified the cessation of disease outbreak with the depletion of susceptible individuals available for sustaining the epidemic, a controversial concept at that time. He also found that the growth of numbers of infectious individuals would be exponential before this depletion effect took hold leading to an 'epidemic curve' that is almost symmetrical but somewhat heavier tailed at the end of the epidemic. These observations, which have now become standard, were revolutionary in the early 20th century since they demonstrated that a plausible model of human interaction and population dynamics could, by themselves, explain the intermittency of the epidemics that plague mankind. This has become a standard concept for the theoretical community of epidemiologists but was completely rejected at the time in certain sections (Brownlee 1909), where time-varying biological factors were favoured such as the pathogen losing infectiousness over time. Ross' differential model (Ross 1916) was ,

$$dP = (n - m + i - e)Adt + (N - M + I - E)Zdt, \tag{1.1}$$

$$dA = (n - m + i - e - h)Adt + (N + r)Zdt, \tag{1.2}$$

$$dZ = hAdt - (M - I + E + r)Zdt. \tag{1.3}$$

Where P, A and Z are respectively the population size, number of not affected (susceptible individual) and the number of affected (infectious individuals). n, m, i, e parameterise the nativity, mortality, immigration and emigration demographic rates for the susceptibles and their capital versions for the infecteds. h is the rate of happenings (recruitment rate of new infecteds) and r is the reversion rate (recovery rate) where the reversion here is back into the susceptible class. In honour of Ross' fundamental contribution I give a numerical solution to his theory of happenings for epidemics assuming permanent immunity reversion. The solution would have been a considerable numerical challenge in 1916, but is trivial with modern computing resource (figure 1.1).

1.3.1 Transmission and the Deterministic SIR Model

The next major contribution was from Kermack and McKendrick (Kermack and McKendrick 1927) who explored the differential model of epidemics from the perspective of *transmissibility*; that the recruitment of new infecteds was due to the infectiousness of the current infecteds, which mix homogeneously amongst the total population, a mass action assumption (Isham 2005). They left their model rather general, allowing the recruitment rate of new infecteds at each time to depend on the prior history of susceptible numbers leading back to the origin of the epidemic. Individuals recovering from the disease were assumed to have long lasting immunity and to play no further role (figure 1.2). There has been some interest in the general model (Fraser 2007) but their classic contribution was the special case now known as the *susceptible-infected-removed (SIR) model of disease* (Kermack and McKendrick 1927; Bartlett 1956; Bailey 1975; Anderson and May 1992; Diekmann and Heesterbeek 2000; Keeling and Rohani 2008), given by the following simple non-linear ordinary differential equation,

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

$$\frac{d}{dt}I(t) = \frac{\beta}{N}S(t)I(t) - \gamma I(t), \tag{1.5}$$

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

Where S, I, R are respectively the numbers of susceptibles, infecteds and removed individuals in the epidemic. β, γ are respectively the transmission rate and the recovery rate. The scaling $1/N$ in the transmission term describes *frequency depen-*

dent[2] transmission (Grassly and Fraser 2008). The assumption behind frequency dependent transmission is that each individual has an expected number of contacts per fixed time interval that does not depend on the total population size N, see (Diekmann and Heesterbeek 2000) chapter 10 for a more general discussion on infectious contacts. The solution given in figure 1.1 is for the SIR model choice as a specialisation of Ross' equations.

There are some immediate consequences of the basic SIR model, firstly that the total number of individuals $N = S + I + R$ is conserved, since there are no demographic effects considered in the model therefore $R(t) = N - S(t) - I(t)$ can be dropped. For the epidemic state $(S(t), I(t)) = (S_0, I_0)$ the number of infectious individuals is growing if, and only if,

$$\frac{S_0}{N} > \frac{\gamma}{\beta} = \frac{1}{R_0}. \tag{1.7}$$

This gives a condition for the number of infecteds to be growing in terms of a new quantity, $R_0 = \beta/\gamma$.

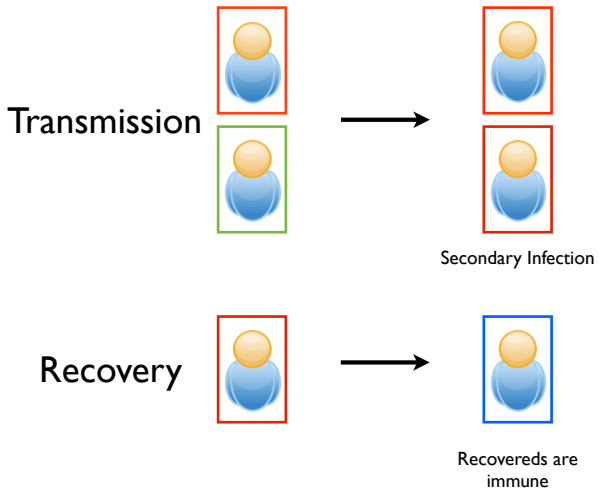


Figure 1.2: The basic SIR epidemic model. The two dynamical transitions are for currently infectious individuals (red box outline) to recruit susceptible individuals (green box outline) as their secondary cases and for infectious individuals to recover and become removed from the epidemic (blue box outline).

[2]Rather confusingly the alternative without the $1/N$ scaling is often referred to in epidemiological literature as *density dependent* transmission.

The *reproductive ratio* R_0[3] is one of the most important and robust quantities in epidemiology; it is written in terms of the transmission and recovery rates and hence is a measure for both the biological infectiousness and persistence of the infectious disease and the underlying mixing between infectious and susceptible individuals. R_0 governs the important threshold phenomenon known as *herd immunity* (Anderson and May 1985a; Fine 1993). Herd immunity is a direct consequence of (1.7) that an infectious disease can only recruit growing numbers of individuals if its reproductive ratio is greater than the inverse population density of susceptibles. This simple observation is crucial since it gives the threshold proportion, denoted V^*, of the population that must be made removed by some method, in most cases by a vaccination campaign, in order to incur herd immunity *for the entire population* as,

$$V^* = \max\left\{0, 1 - \frac{1}{R_0}\right\}. \tag{1.8}$$

Equation (1.8) is fundamental, the entire population can be effectively protected from an infectious disease by directly protecting just a fraction of the population since whatever (typically small) numbers of initial infecteds seed the epidemic will inevitably decline leading to a small outbreak. If $R_0 < 1$ then the infectious disease cannot invade even a naive[4] population. By contrast without vaccination the number of infecteds increases until the susceptibles are sufficiently depleted that (1.7) no longer holds which leads to a large epidemic.

1.3.2 Discrete and Stochastic Epidemic Models

So far in this introduction I have consistently referred to individuals but only discussed classical models that are deterministic and written in terms of ordinary differential equations (ODE). A model that predicts a non-integer population number of susceptibles and infecteds requires some further interpretation. Enforcing the discreteness of the SIR model has historically been done in conjunction with introducing randomness or stochasticity into the transmission and recovery of individuals for a continuous time model (Bailey 1950; Bartlett 1956; Bartlett 1960a; Bailey 1975), a discrete time model such as the Reed-Frost chain binomial model (Becker 1989) or

[3]Not to be confused notationally with R(0), the initial number of removed individuals.

[4]An (immunologically *naive* population is one in which every individual, bar the initial set of infecteds, is initially susceptible to the infectious disease under consideration.

a *generational time*[5] model such as a branching process (Griffiths 1973).

The 'work horse' of random models for the SIR process is the stochastic construction where each individual upon being recruited to the infected class remains infectious for the random length of time T, called the infectious duration (Conlan et al. 2010). During that time the infectious individual makes random infectious contacts, the standard model being contact at the points of a Poisson process (Ball and Neal 2002), with the rest of the population. When a susceptible is contacted that individual is recruited to the infectious class. The most popular choice is to have the infectious durations exponentially distributed since the stochastic epidemic construction with Poisson process contacts and exponential durations can be described as a Markov process (Ethier and Kurtz 1986) with transition probabilities given by the intensity of the Poisson contact β/N and the rate parameter of the exponential duration time γ,

$$\begin{aligned} \mathbb{P}[(S,I)(t+h)-(S,I)(t)=(-1,1)|(S,I)(t)] &= S(t)\frac{\beta}{N}I(t)h+o(h), \quad (1.9)\\ \mathbb{P}[(S,I)(t+h)-(S,I)(t)=(0,-1)|(S,I)(t)] &= \gamma I(t)h+o(h). \quad (1.10)\end{aligned}$$

Where $(S,I)(t)$ describes the random numbers of respectively susceptibles and infecteds at time t.

A major departure for the theory of the stochastic SIR epidemics compared to its deterministic counterpart is that due to randomness the threshold phenomenon is given a probabilistic interpretation in terms of infinite sized populations. If $Z_N(\infty)$ denotes the final number of cases[6] for a population of size N then the probability of a 'true epidemic' (Ball 1983) is said to be,

$$\mathbb{P}((\lim_{N\to\infty} Z_N(\infty))=\infty). \quad (1.11)$$

This is best interpreted as a probability of a non-zero fraction of a large population becoming infected (i.e. a significant epidemic occurring) with the other option being that an effectively zero fraction of a large population becomes infected (i.e. a small epidemic occurring). The probabilistic version of the threshold effect for the stochastic epidemic with dynamics given by is that for an initial population density of susceptibles $s_0=S(0)/N$ and an initial number of infecteds I_0 then the

[5]The generation of an infected individual is the number of infection events she is removed from the initial set of infecteds (the 0th generation).

[6]Cases are individuals who were infected at any point during the epidemic.

probability of a significant epidemic is (Ball 1983),

$$\mathbb{P}(\text{Significant sized epidemic}) = 1 - \left(\min\left\{1, \frac{1}{s_0 R_0}\right\}\right)^{I_0}. \tag{1.12}$$

The threshold vaccination coverage given by (1.8) can be reinterpreted for the stochastic SIR model as the minimum vaccination coverage that that makes the probability of a significant or 'true' epidemic zero. On other side of (1.12) is that even if the critical vaccination coverage (1.8) is not reached, or if vaccine has not been attempted, there is a probability that a small number initial infecteds will fail to cause a significant sized epidemic. This is not a feature of the deterministic SIR model, where the numbers of infected individuals always increases when $S(t)/R_0 < 1$. This bimodal character, i.e. large or small outbreaks, or *U-shape* (Isham 2005), of the distribution of epidemic final sizes is a classic feature of the stochastic SIR model. Even for diseases with SIR-type epidemiology for the individual but that persist for long periods of times due to recruitment of newly arriving susceptible individuals[7] (Anderson and May 1992), such as measles (Grenfell, Bjørnstad, and Kappey 2001; Jansen 2003), stochastic fluctuations can bring about population level 'fade-out' (Bartlett 1960b; Nasell 1999).

Given the above considerations, as a first approximation, the modeller would expect the average disease progression for a stochastic SIR model to be decreased from the deterministic model. More formally, the expected rate of recruitment of new infectious individuals is given as,

$$\frac{d}{dt}\mathbb{E}[I] = \frac{\beta}{N}\mathbb{E}[S]\mathbb{E}[I] - \gamma\mathbb{E}[I] + \frac{\beta}{N}cov[S, I]. \tag{1.13}$$

Where $cov[S, I]$ is the time varying covariance between the random numbers of susceptibles and individuals (Keeling 2000; Herbert and Isham 2000; Isham 2005). The basic SIR model this covariation will be negatively signed leading to, on average, retarded growth of infected numbers compared to the deterministic model. However this is not a universal effect, the interplay between stochasticity and an oscillatory approach to equilibrium for the SIR model with demography can lead stochastic resonance and wild oscillations in disease burden, for a more detailed discussion on stochastic amplification for disease models with demography see (Alonso, McKane, and Pascual 2007).

[7]The SIR model with demography, sometimes called the 'open' SIR model.

Since both the deterministic and stochastic SIR models have been extensively used for both the theoretical and practical understanding of the risk and dynamics of infectious disease it has been of interest to rigorously establish the link between the models. Kurtz established conditions for the convergence of Markovian dynamics with a state space of natural numbers with some maximum size N, say $X(t)$, to deterministic processes in *density* $x(t) = X(t)/N$ as the maximum size diverges $N \to \infty$; a variety of central limit result for Markov processes (Kurtz 1970; Kurtz 1971), which has now been further generalised, c.f. (Kallenberg 2002) theorem 17.15. Specialising Kurtz's results to the stochastic SIR epidemic gives convergence conditions,

$$\begin{aligned} \beta(N)S(t)I(t) \sim \mathcal{O}(N),\ t \geq 0 \qquad & \text{“Density dependence”}, \\ \frac{S(0)}{N} \to s_0,\ \frac{I(0)}{N} \to i_0, \quad & \text{Convergence of initial condition.} \end{aligned}$$

Where I have denoted the intensity of the individual to individual (Poisson) contact process $\beta(N)$ in order to emphasise that, in principle, the contact intensity's dependence on population size is a model choice. The standard choice $\beta(N) = \beta/N$ where β is a constant satisfies the density dependence criterion.

If the convergence conditions are satisfied then as the total population size diverges $N \to \infty$, the density SIR process $s(t) = S(t)/N$, $i(t) = I(t)/N$ converges[8] onto the deterministic density SIR process,

$$\frac{d}{dt}s(t) = -\beta si, \tag{1.14}$$

$$\frac{d}{dt}i(t) = \beta si - \gamma i, \tag{1.15}$$

$$s(0) = s_0,\ i(0) = i_0. \tag{1.16}$$

The standard SIR models given above were formulated in terms of identical and homogeneously mixing host individuals under-going epidemic invasion. This gives a solid theoretical foundation, however in light of real-life epidemic examples often the epidemic model must be modified in order to explain and predict observed phenomena. Examples of heterogeneity that have been incorporated into epidemic models include multi-pathogen models such as for malaria where it has been shown that treating different malaria strains as interchangeable can be mis-leading (Gupta,

[8]The type of convergence is uniform on compacts in probability (u.c.p); the probability that the greatest magnitude deviation between the stochastic density process and the deterministic model on any compact set in time is bigger than the arbitrary error $\epsilon > 0$ goes to zero as $N \to \infty$.

Swinton, and Anderson 1994). Age structured models have been shown to be particularly important in understanding childhood diseases (Anderson and May 1983); is due to both the epidemic sustaining itself by recruiting the individuals born in to the susceptible compartment and the heterogenous mixing between age-groups for human populations, which has been empirically verified (Mossong et al. 2008). Keeling and Rohani (Keeling and Rohani 2008) give large numbers of examples of further host and pathogen heterogeneities and appropriate model extensions.

The heterogeneity I will focus on in this thesis will be spatial heterogeneity. The transmission of a infectious disease between individuals often occurs at close proximity, with long range spreading in a population due to individual movement. However, realistically representing individual movement in a model can be difficult, not to mention leading to analytic intractability and problematic parametrisation. Nonetheless, the habitat location of individuals, e.g. households or farms, are often known and in conjunction with assumptions and arguments on the mixing pattern of the population one can hope to capture disease dynamics as the spatial spread on these sessile locations. Moreover, in the case of plant diseases the individual is stationary and disease transmission is inherently long-range, e.g. via wind-borne dispersal of fungal spores. In this case the individual and the 'habitat' can be directly associated with one another.

1.3.3 Spatial Epidemic Transmission and the Metapopulation Model

Understanding the dynamics of infection invading a spatially structured population is of both theoretical and practical importance. The theoretical importance is due to the relaxation of the homogenous mixing assumption common to classical models of epidemic dynamics discussed in the introductory chapter. Of primary practical concern is the effect that this introduction of heterogeneity in population mixing has on key model predictions, such as the rate of recruitment of new infecteds (Anderson and May 1992), the spatial variation of infection risk (Smith et al. 2002) and in particular the consequences of control measures (Keeling et al. 2003; Tildesley et al. 2006). In contrast to other forms of introducing population heterogeneity, spatial structuring focuses attention on the location of the initial infectious seeds and the complex dynamics of invasion, often characterised by a travelling leading 'edge' akin to population growth into a suitable environment (Ali and Grosskinsky 2010). The emergent spatiotemporal patterns of infection are dominated by this invasion mechanism, with epidemic risk to the underlying population being due to factors such as habitat location of individuals.

There are several approaches to modelling the dynamics of the spread of spatial infection, including PDE models (Murray, Stanley, and Brown 1986) and agent-based models (Barrett et al. 2008), but the *metapopulation model* is often the preferred formulation due to its relative simplicity and flexibility. The metapopulation model consists of discrete point habitats in continuous space that can interact with each other. Its origins date back over 40 years (Levins 1969), but it is experiencing increasing use in a wide number of problems in population dynamics (Hanski and Gilpin 1997; Hanski and Gaggiotti 2004). While the original metapopulation model assumed an equal rate of interaction between all habitats, the inclusion of more realistic rates that decay with spatial separation has lead to a greater understanding of the roles of spatial structure, with applications in as diverse settings as population biology (Ovaskainen and Cornell 2006a; Cornell and Ovaskainen 2008), conservation biology (Hanski), evolutionary biology (Heino and Hanski 2001) and epidemiology (Gibson 1997; Lloyd and Jansen 2004; Grenfell, Bolker, and Kleczkowski 1995).

This natural link between spatial ecology and spatial epidemiology has emphasised recently (Keeling, Bjørnstad, and Grenfell; Lloyd and Jansen 2004). In this thesis I will often make the classic Levins-type assumption (Levins 1969) that the dynamics within the habitat are sufficiently fast that the whole abiding population can be described as one individual, although I will also consider examples where there is explicit population dynamics within each point habitat. This Levins assumption has been used successfully for epidemic models with significant real-world impact such as the Keeling model for foot-and-mouth disease spread between farms (Keeling et al. 2001), which has undergone extensive model evaluation (Tildesley et al. 2008).

Chapter 2

A Spatial and Stochastic Epidemic Metapopulation Model

2.1 Introduction

The aim of this chapter is to incorporate stochastic transmission and spatial fragmentation into the general framework of the SIR compartmental model with a view to the analysis and understanding of their interplaying effect on disease invasion and the final burden of the epidemic. I will treat the time-varying risk of disease introduction to each naive population as Markovian; that is that the probabilistic rate of recruitment to the infectious class will depend only the current spatial distribution of infectious populations. I will exploit the assumed Markovian nature of the model in order to represent the epidemic dynamics as stochastic differential equations (SDE) driven by a set of Poisson processes.

SDEs driven by Brownian motion or diffusions (Karatzas and Shreve 1991) have become classical tool for modelling biological systems with a random component. Applications have included Feller's branching diffusion model of population growth (Feller 1951), as well as applications in population genetics (Fearnhead 2006) and epidemiology (Dargatz, Georgescu, and Held 2006). The fundamental convenience of diffusion models is that on one hand a plausible form of dynamic Gaussian noise can be incorporated into the system being modelled whilst on the other hand a chain rule for the stochastic calculus of the model still applies, namely the famous

Itô lemma[1] (Ito 1951; Oksendal 2004) result which allows functions of diffusion processes to be represented as diffusions themselves.

In this chapter the stochastic dynamics of the spatial epidemic will be analysed using the tools of stochastic calculus for processes which evolve in time only in jumps (Bichteler 2002; Klebaner 2005). Compared to diffusion models the sample paths of a pure jump process are constant except at the arrival of discrete points e.g. infectious transmission or recovery events; whereas nearly all diffusion sample pathes are sufficiently rough as to be nowhere differentiable by contrast the sample paths of pure jump process are nearly everywhere differentiable. This profoundly simplifies the integration theory for pure jump processes. Moreover, an Itô lemma remains accessible in a particularly simple form.

The goal will be to construct systems of ODEs where each deterministically evolving variable approximates a statistical quantity of interest, for example the expected burden of infectious individuals in a population at a certain time. In principle such a construction allows the investigator to analyse the expected behaviour of an epidemic according to a model that incorporates randomness and the fundamental discreteness of individuals within a population using the tools familiar from the theory of ODEs. The system of ODEs which approximates the dynamic moments and evolving covariances of the random epidemic is constructed by sequentially representing the stochastic dynamics of increasing product chains of disease state indicator functions, which I dub the *dynamical hierarchy* of the epidemic process.

I will first review some results for pure jump processes that will be useful in the sequel. I will then introduce the fundamental model of epidemic spread amongst a spatial metapopulation that I use throughout this work. The fundamental model will be shown to be equivalent to a pure jump process represented in a Stochastic differential equation (SDE) form which is amenable to analysis; the covariances between stochastic disease states at different populations is found to obey a hierarchy of ordinary differential equations which can be derived from the SDE formulation.

I am able to bound the size of the population disease state covariances driving the deviation between the expected dynamics of the stochastic model and related deterministic models in terms of an inverse length scale of disease transmission

[1]Also known as Itô's formula.

in a manner closely reminiscent of results found in theoretical population biology (Ovaskainen and Cornell 2006b). This allows the theory to be consider asymptotically exact, in the sense that it has the same limit property as the basic frequency dependent SIR model, albeit here both a large population limit and a diverging length scale of transmission are required.

2.2 Stochastic Calculus with Jumps

In this section I give a brief introduction to stochastic calculus for processes that change only in discrete jumps. The rather beautiful theory of diffusion processes is not relevant for the understanding of disease models where events occur discretely - I will only be concerned with models where an individual or population are considered to become infectious at a certain time in one 'jump'.

If I was considering a deterministic model of disease the natural first step in analysis would be to write the dynamics of the disease process $X = (X(t), t \geq 0)$ in a differential form,

$$dX(t) = F(X(t))dt. \tag{2.1}$$

The taxonomy of this equation is that infinitesimal change in time dt (the integrator) drives the infinitesimal evolution of the temporal disease state dX through the vector field at that time $F(X(t))$ (the integrand). The aim of this section is to establish a meaningful way of writing the dynamics of the stochastic jump model of disease transmission in an analogous differential form, and then harvest the useful theory that exists around such processes. The difficulty compared to classical calculus is that the driver of the differential equation won't be time, leading back to well understood classical calculus, but rather a set of discrete processes. The differential model will be in the form,

$$dX(t) = (F(X(t)), dN(t)). \tag{2.2}$$

Where $dN(t)$ is the infinitesimal change in a collection of Poisson processes, and $(\cdot, \cdot)$ indicates a sum over the Poisson drivers weighted by the set of vector fields, i.e. in the form of an inner product of vectors. This indicates that changes in state occur in jumps ΔX occurring only at the jumps of the driving processes N. Such a process is called a *pure jump process*. A pure jump process $N = (N(t), t \geq 0)$ is characterised by an almost surely increasing sequence of hitting times and their

associated jump sizes $\{(\tau^{(n)}, Z^{(n)})\}_{n\in\mathbb{N}}$, giving the sample path,

$$N(t) = N(0) + \sum_{n\in\mathbb{N}} \mathbf{1}(\tau^{(n)} \leq t) Z^{(n)}. \tag{2.3}$$

Where $\mathbf{1}(\cdot)$ is the standard boolean indicator function. Each sample path of a pure jump process is càdlàg[2]. The restriction on pure jump processes to all jump sizes $\{Z^{(n)} = 1\}_{n\in\mathbb{N}}$ with $N(0)$ is called a *counting process* that is that,

$$N(0) = 0,\ N(t) = \sum_{n\in\mathbb{N}} \mathbf{1}(\tau^{(n)} \leq t). \tag{2.4}$$

Defines the sample path of a counting process for some set of hitting times $\{\tau^{(n)}\}_{n\in\mathbb{N}}$.

In the following I cover a set of results that will be useful for analysing the behaviour of differential models driven by counting processes; the exposition is based upon (Applebaum 2004; Klebaner 2005; Di Nunno, Øksendal, and Proske 2009).

A stochastic processes $X = (X(t), t \geq 0)$ is a collection of random variables, $X(t)$ indexed by the time process t. Each random variable is a measurable map from a probability space to a state space which assigns a value to each sample element of the probability space, I will restrict to real multi-dimensional process so that the state space is $\mathbb{R}^n$ endowed with the standard Borel σ-algebra. The probability space is given as the triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the set of all possible samples for the stochastic model. $\mathcal{F}$ is a collection of sub-sets of Ω which define all events that can be assigned a probability within the model; by necessity $\mathcal{F}$ is endowed with the algebraic structure of a σ-algebra. It is worth mentioning that $\mathcal{F}$ is as much a model choice as the other elements, restricting $\mathcal{F}$ is equivalent to restricting the information accessible to an observer. $\mathbb{P}$ is a probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$. In addition, I define a filtration[3] $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F})$. A stochastic process X which is $\mathcal{F}_t$-measurable for $t \geq 0$ is called *adapted*. Whereas $\mathcal{F}$ represents all observable events for X, the filtration represents all information up to time t. Unless stated otherwise I use a σ-algebra which contains the complete information about the driving counting processes, i.e. all hitting times for the chosen sample element $\omega \in \Omega$, and a filtration that contains this information up until each time t. This is called the natural filtration of the counting processes. Notationally, I follow the majority of

[2] A sample path $N(t)$ is said to be càdlàg if $\lim_{u\downarrow t}$N(u) = N(t) and $\lim_{u\uparrow t} N(u)$ exists for $t \geq 0$.

[3] A filtration is an increasing family of σ-algebras which models the expanding information available to the observer as the process progresses in time.

authors on stochastic processes and write $X(t,\omega) = X(t),\ t \geq 0, \omega \in \Omega$, suppressing explicit dependence of the process on the sample.

A condition for defining 'sensible' (i.e. non-anticipatory) processes is if the integrand formally introduced in (2.2) can be completely determined by knowing only the events leading up to each time t. In other words if the instantaneously current or future random behaviour of the integrator $N(u)$ at times $u \geq t$ does not influence the integrand $F(X(t))$ at time t. This causal property is guaranteed by restricting $F(X(t))$ to be a *predictable* process. For my purpose it is sufficient to define a sub-class of predictable processes,

Definition 2.2.1. *The process X is a* predictable *process if X is left-continuous and adapted.*

Of crucial importance will be the ability to split processes into a *compensator* process, which governs the drift of the process, and a local[4] martingale process, which can be thought of as the fluctuations around the drift. This is referred to in the literature as the *semi-martingale* representation, but I find drift and fluctuation more explanatory and use that terminology. The martingale property for the 'fluctuation' part of the stochastic process will be crucial.

Definition 2.2.2. *An adapted stochastic process $M(t)$ is a martingale if*

- $\mathbb{E}[|M(t)|] < \infty$ *for* $t \geq 0$.
- $\mathbb{E}[M(t)|\mathcal{F}_s] = M(s)$ *almost sure for* $0 \leq s < t < \infty$.

A martingale is said to be square integrable *if also* $\sup_{t\geq 0} \mathbb{E}[M^2(t)] < \infty$.

For the 'drift' part of the stochastic process, the compensator, I restrict my definition to include only the examples useful for this work,

Definition 2.2.3. *The compensator $A(t)$ for a counting process $N(t)$ is the unique predictable process such that $\tilde{N}(t) = N(t) - A(t)$ is a local martingale.*

For the drift and fluctuation decomposition used in this work I will use,

$$X(t) = X(0) + A(t) + \tilde{X}(t), \qquad (2.5)$$
$$A(0) = \tilde{X}(0) = 0.$$

[4] A property for a stochastic process X is said to be local or to hold locally if there exists a sequence of stopping times (τ_n) leading to infinity where for each stopping time the property holds for $X(t \wedge \tau_n)$. In the general theory many basic processes such as Brownian motion exhibit the martingale property only locally. It will not be required for this work.

Hence, an expectation taken on the fluctuation martingale process $\tilde{X}(t)$ returns zero. This also explains the notation $A(t)$ for the compensator; it refers to the addition to the initial state $X(0)$.

I will be concerned with counting processes $N(t)$ where the events arrive at some deterministic average rate[5] $\nu(t)$ defined by the compensator,

$$A(t) = \int_0^t \nu(s)ds, \qquad t \geq 0.$$

Counting processes with deterministic and continuous compensators can be completely characterised as Poisson processes ((Klebaner 2005) Theorem 9.9),

Theorem 2.2.1. *Let $N(t)$ be a counting process with continuous deterministic compensator $A(t)$. Then it has independent Poisson distributed increments such that,*

$$(N(t) - N(s)) \sim Poisson(A(t) - A(s)), \qquad 0 \leq s < t. \tag{2.6}$$

Theorem 2.2.1 will be of great use in the sequel. I will henceforth refer to the driving counting processes as the driving Poisson processes. The local martingale $\tilde{N}(t) = N(t) - A(t)$ is called the *compensated Poisson process.*

Having established the Poisson character of the discrete processes which will play the role of the driver or integrator, I turn to defining the meaning of the differential model (2.2). I am restricting to one dimension and one Poisson driver for simplicity, multidimensional processes with many independent Poisson drivers extend naturally in 'inner product' form. For some vector field $F : \mathbb{R}^n \to \mathbb{R}^n$ and the one-dimensional Poisson process $N(t)$ we define the process X such that,

$$X(T) = X(0) + \int_0^T F(X(t))dN(t). \tag{2.7}$$

The integral on the right hand side of (2.7) is a *stochastic integral*, which is a random variable defined as a limit in probability[6],

$$\int_0^T F(X(t))dN(t) = \lim_{n\to\infty} \sum_{t_i \in \pi_n} F(X(t_i))(N(t_{i+1}) - N(t_i)). \tag{2.8}$$

[5]Often referred to as the *stochastic intensity* of the counting process.

[6]A sequence of random variables $\{X_n\}_{n\geq 0}$ is said to converge in probability to a limit (in probability) X if X is a random variable such that for every $\epsilon > 0$, $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$.

Where $(\pi_n)_{n\geq 0}$ is a sequence of ordered partitions $\pi_n = (t_0 = 0, t_1, \ldots, t_{n-1}, t_n = T)$ of the interval $[0, T]$ with vanishing mesh separation. If this limit exists then,

$$dX(t) = F(X(t))dN(t), \tag{2.9}$$

is the differential form of (2.7) with unique (in probability) solution $X(t)$, $t \geq 0$. Stochastic integration where the integrator is a Poisson process is considerably simpler than for the case where the integrator is a Brownian motion. For a Poisson integrator selecting $\omega \in \Omega$ specifies the set of hitting times and (2.8) gives,

$$\int_0^T F(X(t,\omega))dN(t,\omega) = \sum_{n\in\mathbb{N}} \mathbf{1}(\tau^{(n)} < T)F(X(\tau^{(n)},\omega)). \tag{2.10}$$

Where the integral has been interpreted as a Stieltjes integral (see (Klebaner 2005) chapter 1) completely determined by the choice of sample element $\omega \in \Omega$. This is analogous to the graphical construction of the contact process (c.f. (Liggett 1985)) and the Poisson hitting construction for the stochastic SIR epidemic (see (Ball and Neal 2002) and later in this chapter). Equation (2.10) also demonstrates directly from the definition (2.8) that stochastic processes driven by counting processes are pure jump processes with hitting times given by the integrator process and jump sizes given by the integrand process. For stochastic processes driven by Brownian motions this direct definition from the sample element turns out to be impossible, c.f. (Applebaum 2004) chapter 4. I state a classical existence theorem (c.f. (Di Nunno, Øksendal, and Proske 2009) chapter 9) for stochastic integrals, albeit restricted to Poisson process integrators and one dimension.

Theorem 2.2.2. *Let $N = (N(t), t \geq 0)$ be a Poisson process of intensity $\nu(t)$, let $F(t)$ be a predictable process. If*

$$\mathbb{E}\Big[\int_0^T (|F(t)| + F^2(t))\nu(t)dt\Big] < \infty, \qquad T > 0,$$

then the stochastic integral,

$$X(0) + \int_0^T F(t)dN(t),$$

exists and is called the solution to the stochastic differential equation (SDE) (2.9). Moreover, the process

$$\tilde{X}(T) = \int_0^T F(t)d\tilde{N}(t),$$

is a square integrable martingale.

The usefulness of theorem 2.2.2 is not just that it guarantees that a process $X(t)$ driven by Poisson processes satisfying the conditions above will not 'disappear' to infinity but also that the stochastic integral with respect to the compensated Poisson process is itself a martingale. Therefore, given a stochastic differential model one is able not only to split the integrator into drift and fluctuation form but also the whole process.

Finally, there is an Itô formula for pure jump processes (c.f. (Klebaner 2005) chapter 9) that will be very useful in the sequel for calculating the stochastic rate at which epidemic states jointly vary in time. Given any pure jump process X the Itô formula for pure jump processes allows one to define a new pure jump process for each locally bounded function f.

Theorem 2.2.3. The pure jump Itô formula. *Let $X = (X(t), t \geq 0)$, be an N-dimensional stochastic process driven by a set of counting processes $N(t) = (N_1(t), \ldots, N_n(t))^T$,*

$$X(t) = X(0) + \int_0^t F(X(s))dN(s).$$

For some set of predictable processes depending on X in matrix form $F(X(t)) = (F_{ij}(X(t)))_{1\leq i\leq N, 1\leq j\leq n}$. Let $f : \mathbb{R}^N \to \mathbb{R}$ be locally bounded and define

$$Y(t) := f(X(t)), \qquad t \geq 0.$$

Then the process $Y = (Y(t), t \geq 0)$ is a one-dimensional pure jump stochatic process and its differential form is given by

$$dY(t) = \sum_{j=1}^{n} [f(X(t) + F_j(X(t))) - f(X(t))]dN_j(t)$$

where $F_j(X(t))$ is the column number j of the $N \times n$ matrix $F(X(t))$.

2.3 The Spatial and Stochastic Epidemic Model

In this section I introduce the basic model of spatial disease transmission between discrete populations, which abide as a meta-population of sessile point habitats fragmented in $\mathbb{R}^d$. I define the stochastic rate at which an invasive pathogen spreads

from population to population as only depending on spatial distance between the habitats of the populations modulated by a transmission kernel. The basic model form is equivalent to a stochastic process driven by Poisson processes.

2.3.1 Epidemic Dynamics

The population undergoing epidemic invasion is considered to be distributed between a discrete set of N habitats labelled $i = 1, \ldots, N$. Each habitat is considered to be interchangeable and to have effective area zero compared to the background space; ergo the only information about the habitats required for this epidemic model is their spatial co-ordinates, $\{x_i\}_{i=1}^{N}$. When considering a finite number of habitats the spatial co-ordinates are constrained to be within a d-dimensional box of side length l, $\mathcal{A} = [-l/2, l/2]^d$. The density of habitats in the space is given by Nl^{-d}; in this work I choose units such that the spatial density of habitats is 1 i.e.

$$Nl^{-d} = 1. \tag{2.11}$$

Whenever a limit $N \to \infty$ is considered, implicit is also a limit $l \to \infty$ taken in such a manner as to maintain the unity of the spatial density. This choice of units has the benefit of removing habitat spatial density as an unnecessary parameter from the model. Moreover, it allows the definition of a natural length scale, i.e. the units in which (2.11) holds, useful for the comparison of theoretical predictions to field data.

In order to model the spatiotemporal development of an invasive pathogen spreading from population to population I assign a disease state to each population. Each population i, in addition to a spatial location x_i for its habitat, is assigned a state $\sigma_i(t) \in \{S, I, R\}$ according to the classic SIR compartmental model of disease (Kermack and McKendrick 1927; Bartlett 1956). The state is considered to represent the entire population abiding at habitat i; this is a modelling choice that is appropriate when intra-population transmission occurs at a much faster time scale than inter-population transmission, in particular subsequent transmission is considered to have negligible effect on the population subsequent to initial introduction of the pathogen. Treating the host population as a single host individual conforms to the standard Levins-type metapopulation (Levins 1969). The initial disease state $\sigma(0) = (\sigma_1(0), \ldots, \sigma_N(0))$ is distributed according to some initial law, $\mathbb{P}_0$.

A Spatial Transmission Model

I now give the details of the stochastic transmission model between populations abiding in the spatially fragmented meta-population. Central to this work is that each infectious (I) population introduces a risk per unit time of causing transmission to each susceptible (S) population. Each source of epidemic risk represents a coarse graining over all possible realistic routes of infection. Examples of transmission pathways might be the roaming of Badger individuals between cattle farm as a potential transmission pathway for bTB(Donnelly et al. 2005; Vial, Johnston, and Donnelly 2011) or the movement of potentially infectious livestock to a farm with a naive population (Vernon and Keeling 2009). Each transmission risk is assumed to operate independently of other sources with the consequence that the probabilistic rate of disease establishing itself in a susceptible population is the summed risk of transmission per unit time over the set of infectious populations.

The rate at which an infectious population introduces the pathogen into a susceptible population is modulated by their spatial separation. The modulation is governed by the *transmission kernel* K. Investigating the functional dependence of the epidemic dynamics on the choice of transmission kernel is one of the main goals of this work. I impose conditions on the selection of the spatial kernel:

- K is a smooth function on $\mathbb{R}^d$.
- The transmission is spatially isotropic and radially symmetric i.e. $K(x) \equiv K(|x|)\ \forall x \in \mathbb{R}^d$.
- K is a decreasing function of $|x|$ with a maximum at $K(0)$ (if it exists).
- There exists a characteristic length scale L to transmission, which is independent of the shape of the transmission kernel. To be exact I require that K is in the form $K(x) = \frac{1}{L^d}k(x/L)$ for some shape function k independent of L.

The final condition gives that the total 'effort' of the infection over space, defined as,

$$K_0 = \int_{\mathcal{A}} K(x)dx, \tag{2.12}$$

is invariant under variation in L, so long as $L \ll l$. This will be important since the invariance of the spatial integral plays an analogous role in the following as the frequency dependent transmission assumption does in deriving limiting large popu-

lation dynamics (Kurtz 1970; Kurtz 1971).

Almost surely each infectious population remains infectious, and thereby contributing to the ongoing epidemic, for a finite period of time. In this model once a population has ceased being infectious it is removed (R) from the epidemic dynamics, no subsequent re-infection can occur. The recovery mechanism potentially models either the population reaching a critical herd immunity (Anderson and May 1985a) after which infection declines rapidly, detection and population removal as is commonly the case with diseases of commercial livestock (Ferguson, Donnelly, and Anderson 2001b) or even the complete elimination of host by the disease as is the case with certain diseases of plants (Rizzo and Garbelotto 2003). Recovery occurs at a constant probabilistic rate γ during the infectious period of the population.

The model considerations above point towards the definition of Markovian dynamics for the symbolic process $\sigma = (\sigma(t) = (\sigma_1(t), \ldots, \sigma_N(t)), t \geq 0)$ on $\{S, I, R\}^N$ with initial law, $\mathbb{P}_0$ and probabilistic dynamics specified by,

$$\mathbb{P}(\sigma_i(t+h) = I | \sigma(t), \sigma_i(t) = S) = \Big[\sum_{j:\sigma_j(t)=I} K(x_i - x_j) \Big] h + o(h), \tag{2.13}$$

$$\mathbb{P}(\sigma_i(t+h) = R | \sigma_i(t) = I) = \gamma h + o(h). \tag{2.14}$$

$$i = 1, \ldots, N, \ h > 0.$$

This version of defining Markovian dynamics is possibly the most natural since one can consider the stochastic rates of new events arriving as setting the probability of a discrete disease event, either an infection or recovery event, arriving in the infinitesimal future interval $[t, t + dt]$. The probability of more than one event occurring simultaneously is zero; an event either occurs in the infinitesimal future or it doesn't.

2.3.2 An Equivalent Stochastic Model

In this section I present an alternate model of disease spread, based on selecting the transmission and recovery times from the hitting times of a set of Poisson process. This will be shown to be probabilistically equivalent to the coin tossing version of the Markovian dynamics presented above and allows the spatial epidemic to be represented as a stochastic process driven by Poisson processes and thereby give access to the theory presented in section 2.2.

Firstly, I define N^2 independent Poisson processes, using the notation $X \sim PP(\nu)$ to indicate that X is a Poisson process of intensity ν.

$$N_{ij}^I \sim PP(K(x_i - x_j)), \quad i, j = 1, \ldots, N, \ i \neq j, \tag{2.15}$$

$$N_i^R \sim PP(\gamma), \qquad i = 1, \ldots, N. \tag{2.16}$$

The spatial epidemic dynamics can now be subordinated to these Poisson processes. This is a technique that is common amongst the community of applied probabilists working in the field of mathematical epidemiology (Ball and Neal 2002) in order to better analyse thresholds for epidemic eradication (Ball, Britton, and Lyne 2004) or the probability of early extinction and distribution of final epidemic sizes (Ball, Sirl, and Trapman 2010). The essential idea is that the number of jumps of a homogeneous Poisson process $PP(\nu)$ on the time interval $[t, t+h)$ is always distributed as Poisson(νh) due to stationary independent increments. Hence, the probability of the number of jumps being ≥ 1 is

$$1 - \mathbb{P}(\#jumps \in [t, t+h) = 0) = 1 - e^{-\nu h} = \nu h + o(h).$$

Similarly for M independent homogenous Poisson processes each with some intensity $\{\nu_i\}_{i=1}^M$ the probability of there being no jumps for any of the M processes is

$$1 - \mathbb{P}(\#jumps \in [t, t+h) = 0) = 1 - \prod_{i=1}^{M} e^{-\nu_i h} = \sum_{i=1}^{M} \nu_i h + o(h).$$

These basic considerations demonstrate that probabilistic dynamics of the form (2.13) and (2.14) can be constructed using the hitting times for jumps of judiciously chosen Poisson processes.

I now give a mechanistic description of constructing a sample epidemic from Poisson processes. An initial state $\sigma(0)$ is drawn from $\mathbb{P}_0$. The hitting times of the Poisson processes form the complete set of times at which events can occur, the set of Poisson processes $\{N_{ij}^I\}$ defining the possible times of population i becoming infected by population j and $\{N_i^R\}$ defining the possible times of recovery for population i. The dynamics evolve thus, with a pictorial desciption being given in figure 2.1:

1. Generate a sufficiently large number of hitting times for each Poisson process, $\{(\tau_{ij}^I)^{(n)}\}_{n\geq 0}$ and $\{(\tau_i^R)^{(n)}\}_{n\geq 0}$ and set time $t = 0$.
2. Form an ordered set of the hitting times for all Poisson processes retaining their originating Poisson Process, $\tau^{(1)} < \tau^{(2)} < \tau^{(3)} < \ldots$

3. Sequentially, from the earliest hitting time onwards, check which Poisson process originated the hitting time. If the originator of the current hitting time $\tau^{(n)}$ was N_{ij}^I and an infectious transmission is feasible, that is that population i is susceptible and population j is infectious, then update the time to $\tau^{(n)}$ and the local state at population i to $\sigma_i(\tau^{(n)}) = I$. Similarly, for hitting times originating from N_i^R where the feasibility is given by population i being in the infectious state and update is given by $\sigma(\tau^{(n)}) = R$. At all other times the epidemic state is constant.

4. If $\sigma_i \neq I$ for all $i = 1, \dots N$. Then terminate the dynamics.

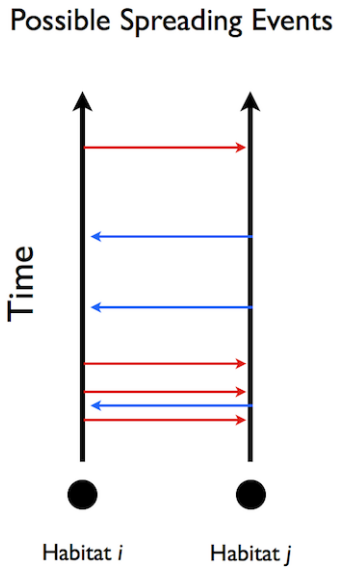


Figure 2.1: A schematic diagram of the *possible* random times of spread between two habitats labelled i and j. The red arrows indicate the hitting times of the Poisson process $N_{ij}^I(t) \sim PP(K(x_i - x_j))$. Transmission occurs at the hitting time if, and only if, habitat i is in an infectious state and habitat j is in a susceptible state. The blue arrows give the symmetric case of possible transmission times from habitat j to i.

In this stochastic construction an event arriving for the the Poisson process N_{ij}^I is interpreted as a contact time between the populations i and j. If the population j is infectious then this is an infectious contact and if in addition population i is susceptible this is a successful infectious contact that results in transmission.

2.3.3 From the Poisson construction to the Stochastic Integral

The purpose of this section is to demonstrate that the Poisson construction above, which is probabilistically equivalent to the defining Markovian dynamics of the spatial epidemic model, is also equivalent to a Stochastic process with the Poisson processes integrators, $\{N_{ij}^I\}$ and $\{N_i^R\}$.

Rather than work directly with the local symbolic state space, $\{S, I, R\}$, I define the indicator functions,

$$S_i(t) = \mathbf{1}(\sigma_i(t), S), \tag{2.17}$$

$$I_i(t) = \mathbf{1}(\sigma_i(t), I), \tag{2.18}$$

$$i = 1, \ldots, N,\ t \geq 0.$$

The indicator function $R_i(t)$ is redundant since $R_i(t) = 1 - S_i(t) - I_i(t)$. For any initial distribution of indicators, $\{S_i(0), I_i(0)\}_{i=1}^N$, the Poisson construction from the previous section can be compactly rewritten in terms of the hitting times of the Poisson drivers,

$$S_i(t) = S_i(0) - \sum_{j(\neq i)} \sum_{n \geq 0} \mathbf{1}((\tau_{ij}^I)^{(n)} < t) S_i((\tau_{ij}^I)^{(n)-}) I_j((\tau_{ij}^I)^{(n)-}), \tag{2.19}$$

$$\begin{aligned} I_i(t) = I_i(0) + \sum_{j(\neq i)} \sum_{n \geq 0} \mathbf{1}((\tau_{ij}^I)^{(n)} < t) S_i((\tau_{ij}^I)^{(n)-}) I_j((\tau_{ij}^I)^{(n)-}) \\ - \sum_{n \geq 0} \mathbf{1}((\tau_i^R)^{(n)} < t) I_i((\tau_i^R)^{(n)-}). \end{aligned} \tag{2.20}$$

$$i = 1, \ldots, N,\ t \geq 0.$$

Here, $f(t^-) = \lim_{s \uparrow t} f(s)$ is the left limit of the function f, and recalling that in this notation $\{(\tau_{ij}^I)^{(n)}\}_{n \geq 0}$ denotes the times at which habitat j makes a potentially infectious contact with habitat i. By comparison to (2.10) it is clear that equations (2.19) and (2.20) are solutions to a stochastic integral,

$$S_i(t) = S_i(0) - \int_0^t S_i(t^-) \sum_{j(\neq i)} I_j(t^-) dN_{ij}^I(t), \tag{2.21}$$

$$I_i(t) = I_i(0) + \int_0^t S_i(t^-) \sum_{j(\neq i)} I_j(t^-) dN_{ij}^I(t) - \int_0^t I_i(t^-) dN_i^R(t), \tag{2.22}$$

$$i = 1, \ldots, N.$$

From here on I refer to the stochastic process,

$$X = X(t) = (S_1(t), I_1(t), \ldots, S_N(t), I_N(t)), \tag{2.23}$$

as the spatial epidemic process, since it incorporates all the relevant information about the random state of epidemic as it progresses through time.

I fix the underlying probability space as in section 2.2 and use the natural filtration for the underlying Poisson drivers; that is the filtration that contains all information about their hitting times. The samples $\omega \in \Omega$ can be thought of as the combination of an initial epidemic state and all the hitting times of all the Poisson driving processes, this information becomes available to the observer as time t progresses. The construction above shows directly that $X(t)$ is adapted, moreover for each i, j $S_i(t^-)$ and $I_i(t^-)$ are left continuous by the definition of left limits. Therefore the integrand processes for (2.21) and (2.22) are predictable by definition 2.2.1. This establishes the existence of the spatial epidemic as a pure jump process.

Theorem 2.3.1. *Suppose that for the spatial meta-population* $\sum_{j(\neq)i} K(x_i - x_j) < \infty$ *for all* $i = 1, \ldots, N$. *Then the spatial epidemic process* X *defined by equation (2.23) is a pure jump process with vector valued jumps. The stochastic dynamics of local states given in differential form as the set of stochastic differential equations (SDEs),*

$$dS_i(t) = -\sum_{j(\neq i)} S_i(t^-) I_j(t^-) dN_{ij}^I(t), \tag{2.24}$$

$$dI_i(t) = \sum_{j(\neq i)} S_i(t^-) I_j(t^-) dN_{ij}^I(t) - I_i(t^-) dN_i^R(t), \tag{2.25}$$

$$i = 1, \ldots, N$$

with solutions defined in terms of the well-defined stochastic integrals (2.21) and (2.22).

Proof. Each indicator function $\{S_i(t), I_i(t)\}_{i=1}^N \in \{0, 1\}$ therefore at each hitting time τ for the driving Poisson processes the change in local state is one of $\{-1, 0, 1\}$ depending on whether an infection or recovery event occurred for a given habitat. Therefore a sample path for X can be defined from an initial configuration of local disease states, the complete set of all hitting times of the driving Poisson processes and the set of $2N$-length vectors corresponding to the local changes in disease state

due to an infection or recovery event at each habitat. Therefore X is a jump process with vector valued jumps. The left-continuous processes $S_i(t^-)$ and $I_i(t^-)$ are predictable, as established above. I focus on one term in the sum of the integrand process of (2.24) and consider the existence criteria for theorem 2.2.2,

$$\begin{aligned}\mathbb{E}\Big[\int_0^t K(x_i - x_j)(|S_i(s^-)I_j(s^-)| + (S_i^2(s^-)I_j^2(s^-))ds\Big] \\ = 2\mathbb{E}\Big[\int_0^t K(x_i - x_j)S_i(s^-)I_j(s^-)ds\Big] \\ \leq 2K(x_i - x_j)\mathbb{E}\Big[\int_0^t I_j(s^-)ds\Big] \\ \leq \frac{2K(x_i - x_j)}{\gamma} < \infty.\end{aligned}$$

Where I have used that the indicators are binary taking values 0 or 1 and used that $\int_0^\infty I_j(s^-)ds$ is either zero (if the jth population is never infected) or equal to a sample of the duration of infectiousness, which is of mean length $1/\gamma$. If also $\sum_{j(\neq)i} K(x_i - x_j) < \infty$ then the stochastic integral (2.21), which defines the stochastic dynamics (2.24), exists by theorem 2.2.2. A similar argument can be used for the existence of (2.22). □

It is common to present the SDEs (2.24) and (2.25) as decomposed into a drift term driven by time and a fluctuation term driven by a martingale process. First, we note that the compensator of the Poisson process coincides with its expectation at each time, $A(t) = \mathbb{E}[N_{ij}^I(t)] = K(x_i - x_j)t$ for all $t \geq 0$. Hence in differential form,

$$dN_{ij}^I(t) = dA(t) + d\tilde{N}(t) = K(x_i - x_j)dt + d\tilde{N}(t). \tag{2.26}$$

The differential form for the other Poisson integrators is similar. Introducing this

into equations (2.24) and (2.25) gives,

$$\begin{aligned} dS_i(t) &= -S_i(t^-)\sum_{j(\neq i)} K(x_i - x_j)I_j(t^-)dt \\ &\quad -S_i(t^-)\sum_{j(\neq i)} I_j(t^-)d\tilde{N}^I_{ij}(t), \qquad (2.27) \\ dI_i(t) &= \Big[S_i(t^-)\sum_{j(\neq i)} K(x_i - x_j)I_j(t^-) - \gamma I_i(t^-)\Big]dt \\ &\quad +S_i(t^-)\sum_{j(\neq i)} I_j(t^-)d\tilde{N}^I_{ij}(t) - I_i(t^-)d\tilde{N}^R_i(t), \qquad (2.28) \\ &\quad i = 1, \ldots, N. \end{aligned}$$

The terms proportional to the differential forms of the compensated Poisson processes are differential forms for a martingale process by theorem 2.2.2. Their sum defines the fluctuation part of $dS_i(t)$ and $dI_i(t)$. Hence, the terms proportional to dt are the differential forms for the compensators of $dS_i(t)$ and $dI_i(t)$. In drift and fluctuation form equations (2.27) and (2.28) are similar to those derived by Ovaskainen and Cornell (Ovaskainen and Cornell 2006b; Cornell and Ovaskainen 2008), albeit for the context SIR type epidemic dynamics rather than the equilibrium colonisation of unoccupied patches by an invasive species. My main departure is that I have given an explicit representation of the martingale part of the stochastic dynamics in terms of well defined integrals with respect to compensated Poisson processes.

2.3.4 Covariance Closure and Covariance between Disease States

An important feature of stochastic models is the often observed fact that the predicted expected dynamics, $\mathbb{E}[X] = (\mathbb{E}[X](\mathrm{t}),\, t \geq 0)$, is not equivalent to solving the drift part of the decomposition as a related ordinary differential (ODE) problem. That is that for general stochastic differential models,

$$dX(t) = F(X(t))dt + \sigma(X(t))dM(t) \implies \frac{d}{dt}\mathbb{E}[X(t)] = F(\mathbb{E}[X](t)).$$

In particular, the above is true whenever the vector field $F(\cdot)$ depends non-linearly on the stochastic state $X(t)$. This has been emphasised in theoretical ecology and epidemiology since at least the nineties, c.f. (Isham 1995; Bolker and Pacala 1997; Isham 2005). However, it is worth investigating why the ODE approach to calculating the expected dynamics fails and the role of covariance between disease states. In this section I give the contribution of covariance between disease states to the expected dynamics of the spatial epidemic, give the most basic ODE approximation

to the expected dynamics and an upper bound for the supremum error of this basic approximation.

By taking expectations on equations (2.27) and (2.28) we see that for every disease state indicator process we can make the drift and fluctuation decomposition,

$$S_i(t) = \mathbb{E}[S_i(t)] + \tilde{S}_i(t) \tag{2.29}$$

$$I_i(t) = \mathbb{E}[I_i(t)] + \tilde{I}_i(t) \tag{2.30}$$

$$i = 1, \ldots, N.$$

The time-varying covariance between the random disease indicators, $S_i(t), I_j(t)$ is given by the joint expectation of their fluctuations,

$$\mathrm{cov}[S_i, I_j](t) = \mathbb{E}[\tilde{S}_i(t)\tilde{I}_j(t)] = \mathbb{E}[S_i(t)I_j(t)] - \mathbb{E}[S_i(t)]\mathbb{E}[I_j(t)].$$

Whenever $\mathrm{cov}[S_i(t), I_j(t)] \neq 0$ I say that populations i and j form an *SI correlated pair*. I also make the short-hand notation,

$$c_{XY}(i, j; t) = c_{YX}(j, i; t) = \mathrm{cov}[X_i(t), Y_j(t)] \tag{2.31}$$

$$i, j = 1, \ldots, N,\ X, Y \in \{S, I, R\},\ t \geq 0.$$

For the SDE representation of the spatial epidemic (2.24), (2.25), the expected dynamics are corrected by the covariance between the susceptibility of each population i and the infectiousness of each population j. This can be seen by introducing the drift and fluctuation decomposition to each disease indicator in equations (2.27) and (2.28), taking expectations and applying the martingale property,

$$\frac{d}{dt}\mathbb{E}[S_i](t) = -\mathbb{E}[S_i] \sum_{j(\neq i)} K(x_i - x_j)\mathbb{E}[I_j] - \sum_{j(\neq i)} K(x_i - x_j)c_{SI}(i, j; t) \tag{2.32}$$

$$\frac{d}{dt}\mathbb{E}[I_i](t) = \mathbb{E}[S_i] \sum_{j(\neq i)} K(x_i - x_j)\mathbb{E}[I_j] + \sum_{j(\neq i)} K(x_i - x_j)c_{SI}(i, j; t) - \gamma\mathbb{E}[I_i], \tag{2.33}$$

$$i = 1, \ldots, N.$$

We shall see that the formation of SI correlated pairs is driven by various triple configurations, and can indeed extrapolate that triple configurations are formed by quadruple configurations and so on. This observation plays a key role in so called *closure schemes* (Bolker and Pacala 1997) for approximating the statistical observables of random processes. Specialising to the spatial epidemic process equation we

will find that in order to calculate the contribution of SI correlated pairs at any time to the expected dynamics one must first calculate the expected contribution of SII and SSI correlated triple configurations. These considerations lead to a hierarchy of correlation dynamics that must be solved. A closure scheme is an approximation of higher order correlations in terms of lower order correlations and the expected dynamics.

The simplest correlation closure is to treat random variables as uncorrelated, i.e. that their covariance is zero. For the spatial epidemic process this is equivalent to the approximation,

$$\mathbb{E}[S_i(t)I_j(t)] \approx \mathbb{E}[S_i](t)\mathbb{E}[I_j](t), \qquad t \geq 0. \tag{2.34}$$

Applying closure (2.34) to equations (2.32) and (2.33) gives,

$$\frac{d}{dt}\mathbb{E}[S_i](t) \approx -\mathbb{E}[S_i](t) \sum_{j(\neq i)} K(x_i - x_j)\mathbb{E}[I_j](t), \tag{2.35}$$

$$\frac{d}{dt}\mathbb{E}[I_i](t) \approx \mathbb{E}[S_i](t) \sum_{j(\neq i)} K(x_i - x_j)\mathbb{E}[I_j](t) - \gamma\mathbb{E}[I_j](t), \tag{2.36}$$
$$i = 1, \ldots, N.$$

These equations are in ODE form for a closed set of dynamic variables, the expectation processes $\{\mathbb{E}[S_i], \mathbb{E}[I_i]\}_{i=1}^N$, and can therefore be solved efficiently using standard ODE integration techniques such as Runge-Kutta to give an approximation to the expected dynamics that does not require extensive Monte Carlo estimation to solve. However, it is not immediately clear what the size of the error due to using this approximation is, I turn to this question in the next section.

2.3.5 Truncation Error for Pair Covariances

The unfortunate aspect of correlation closure schemes is that the approximation error is generally uncontrolled (see (Murrell, Dieckmann, and Law 2004) for a general discussion on the reliability of closure schemes) and their performance can only be assessed through comparison to Monte Carlo simulation. However, it is clear that the magnitude of the error between the true expected dynamics and the approximations (2.35) and (2.36) is determined by the magnitude of the contribution due to SI correlated pairs. This can be made more explicit with a lemma, given without proof (see (Kurtz 1971) for a further generalised version of the lemma),

Lemma 2.3.2. *Suppose there exists a vector field $F(\cdot) : \mathbb{R}^m \to \mathbb{R}^n$ for some $m, n \in \mathbb{N}$, where F obeys Lipschitz continuity with Lipschitz constant C: $|F(x) - F(y)| \leq C|x - y|,\ \forall x, y \in \mathbb{R}^m$. Let $X(t)$ be the solution to*

$$\frac{d}{dt}X(t) = F(X(t)), \qquad X(0) = x_0.$$

Let $G : \mathbb{R}^m \to \mathbb{R}^n$ also be a vector field bounded on $\mathbb{R}^m$ and suppose $X'(t)$ is the solution to

$$\frac{d}{dt}X'(t) = G(X'(t)), \qquad X'(0) = x_0'.$$

then

$$\sup_{s \leq t} |X(s) - X'(s)| \leq \Big(|x_0 - x_0'| + t \sup_{x \in \mathbb{R}^m} |F(x) - G(x)|\Big)e^{Ct}. \tag{2.37}$$

The dependence on correlated SI pairs for the true expected dynamics (2.32) and (2.33) donates a time varying contribution to the vector field. I make the standard trick of converting into an autonomous ODE by incorporating the time process t into the state space. I denote the expected dynamics approximated by neglecting all the covariance processes $\{c_{SI}(i, j; \cdot)\}_{i,j=1}^N$, that is the approximate dynamics (2.35), (2.36) as $\mathbb{E}[X]'(t)$ and only consider the case when the initial distribution of local disease state is known, thereby allowing initial conditions to be matched. The difference between the vector fields for $\mathbb{E}[X]$ and $\mathbb{E}[X]'$ is the vector of length $2N$ giving the covariance contributions to each disease state indicator. Applying lemma 2.3.2 for the Euclidean difference between two $2N$ length vectors then gives that,

$$\sup_{s \leq t} |\mathbb{E}[X](t) - \mathbb{E}[X]'(t)| \leq \Big(\sup_{s \leq t} \sqrt{2 \sum_i \Big(\sum_{j(\neq i)} K(x_i - x_j) c_{SI}(i, j; s)\Big)^2}\Big) t e^{Ct}. \tag{2.38}$$

The uniform bound (2.38) depends directly on the magnitude of contribution of correlated pairs, if these decay to zero than we will obtain uniform convergence for the trajectory of the approximate dynamics (2.35), (2.36), onto the true expected dynamics.

2.3.6 The Dynamical Hierarchy of Local Disease States

In the previous section we saw that the error between the true expected dynamics and a basic ODE approximation was governed by the covariance between disease states; that is the joint expected dynamics of their fluctuations. The joint dynamics of products of disease indicators, and their fluctuations, are also stochastic processes. In fact they can be characterised as counting processes driven by the same set of N^2 Poisson process integrators via the pure jump Itô formula 2.2.3. I will call the joint dynamics of pair products the *pair dynamics*, triple products *triple dynamics* and so on. The pair and triple dynamics will be given explicitly. The most important observation will be that the product dynamics do not obey the chain rule familiar from classical calculus.

The deviation between the classical chain rule and the dynamics given by the Itô formula 2.2.3 is itself a stochastic process called the *quadratic covariation process*. In this section I will demonstrate that the quadratic covariation between disease states drives the formation of correlations between the disease states of spatially separate populations and enforces the discreteness of the dynamics. I will also explicitly calculate the joint stochastic dynamics for pairs and triples. In combination with the disease state dynamics I call this the *dynamical hierarchy* for the spatial epidemic. The basic idea is that taking expectations and applying a closure scheme to ever increasing sizes of the dynamical hierarchy leads to diminishing error between the true expectation dynamics and the 'closed' approximate expected dynamics. In this sense the approximate expected dynamics (2.35), (2.36), were a closure on the singlet hierarchy.

I construct the dynamics of pair diseases states as a real-valued process using products of the indicator functions (2.17) and (2.18). For example, the SI pair of populations i and j is value 1 at time t iff population i has the local disease state S and population j has local disease state I at time t, which is equivalent to $S_i(t)I_j(t)$. A direct application of the Itô formula 2.2.3 to the set of product functions

$$f(X(t)) = X_i(t)Y_j(t), \qquad X, Y \in \{S, I\},\ i, j, = 1, \ldots, N,\ t \geq 0,$$

gives the following stochastic pair dynamics. As a short-hand I leave the dependence of the integrand on the left-continuous version of the triples implicit and drop the explicit time dependence.

Pair Dynamics:

$$d(S_iS_j) = -S_iS_j\big(\sum_{k(\neq i)} I_k dN^I_{ik} + \sum_{k(\neq j)} I_k dN^I_{jk}\big) + \delta_{ij}S_i \sum_{k(\neq i)} I_k dN^I_{ik}, \tag{2.39}$$

$$d(S_iI_j) = S_iS_j \sum_{k(\neq j)} I_k dN^I_{jk} - S_iI_j dN^R_j - S_iI_j \sum_{k(\neq i)} I_k dN^I_{ik} - \delta_{ij}S_i \sum_{k(\neq i)} I_k dN^I_{ik}, \tag{2.40}$$

$$\begin{aligned} d(I_iI_j) &= S_iI_j \sum_{k(\neq i)} I_k dN^I_{ik} - I_iI_j dN^R_i + I_iS_j \sum_{k(\neq j)} I_k dN^I_{jk} \\ &\quad -I_iI_j dN^R_j + \delta_{ij}\big(S_i \sum_{k(\neq i)} I_k dN^I_{ik} + I_i dN^R_i\big), \end{aligned} \tag{2.41}$$

$i, j = 1, \ldots N.$

The origin of the Kronecker delta functions in the integrand processes above is from noting that the same arriving event for one Poisson process integrators drives the jump for both S_i and I_i; that is that a successful infectious contact event varies two of the system variables simultaneously. I make two additional observations on the pair dynamics. Firstly, the pair dynamics are now explicitly written as depending on triples and are driven by the N^2 set of Poisson processes. Secondly, note that the pair dynamics are in the form,

$$d(XY) = XdY + YdX + d[X, Y]. \tag{2.42}$$

Equation (2.42) presents the pair dynamics as the dynamics expected from the classic product rule plus a correcting process $[X, Y] = ([X, Y](t), t \geq 0)$. The correcting process is the *quadratic covariation process* between the processes X and Y. For example, the quadratic covariation process for the pair dynamics for $S_i(t)I_j(t)$ is,

$$d[S_i, I_j](t) = -\delta_{ij}S_i \sum_{k(\neq i)} I_k dN^I_{ik}.$$

Deterministic processes with differentiable trajectories obey the classic product rule of calculus, and hence for such deterministic processes X_{det} and Y_{det}: $[X_{det}, Y_{det}](t) = 0,\ t \geq 0$. For any stochastic process, X, let $\tilde{X}(t) = X(t) - A(t)$ be the martingale part of X, where $A(t)$ is the compensator of X. Since the expected

dynamics are deterministic, the covariance between two random processes' states at time t, $\text{cov}[X(t), Y(t)] = c_{XY}(t)$, is given by,

$$\begin{aligned} dc_{XY}(t) &= d\mathbb{E}[\tilde{X}\tilde{Y}] \\ &= \mathbb{E}[X dY] + \mathbb{E}[Y dX] + d\langle X, Y\rangle - \mathbb{E}[X]d\mathbb{E}[Y] - \mathbb{E}[Y]d\mathbb{E}[X] \\ &= \mathbb{E}[\tilde{X} dY] + \mathbb{E}[\tilde{Y} dX] + d\langle X, Y\rangle. \end{aligned} \tag{2.43}$$

Where $\langle X, Y\rangle = \mathbb{E}[[X, Y]]$ is the expected quadratic covariation process. At this juncture it is useful to introduce some further notation. Equations (2.27) and (2.28) gave the drift and fluctuation form for the disease state indicators at each population, I noted that the differential form of the compensator for each dS_i and dI_i was given in the form,

$$dX = dA(t) + d\tilde{X} = r_X dt + d\tilde{X}. \tag{2.44}$$

Since $A(t) = \int_0^t r_X(s)ds$ I refer to r_X as the intensity of process X in analogy to the terminology of Poisson processes. The intensity r_X is itself a stochastic process though ergo disjoint increments for $X(t)$ are expected to be neither independent nor Poisson distributed. Using the intensity of the process simplifies equation (2.43) to,

$$\frac{d}{dt}c_{XY}(t) = \mathbb{E}[\tilde{X} r_Y] + \mathbb{E}[\tilde{Y} r_X] + \mathbb{E}[r_{X,Y}]. \tag{2.45}$$

Where $r_{X,Y}$ is the intensity of the quadratic covariation process $[X, Y]$. Note that for X, Y representing disease states for the spatial epidemic $\mathbb{E}[\tilde{X} r_Y]$ depends on the disease state expectations, pair covariances and triple covariances; for example for

$X = S_i$ and $Y = I_j$,

$$\begin{aligned}
\mathbb{E}[\tilde{S}_i r_{I_j}] &= \mathbb{E}[\tilde{S}_i(S_j \sum_{k(\neq j)} K(x_j - x_k)I_k - \gamma I_j)] \\
&= \mathbb{E}[\tilde{S}_i(\mathbb{E}[S_j] \sum_{k(\neq j)} K(x_j - x_k)\mathbb{E}[I_k] - \gamma\mathbb{E}[I_j])] \\
&\quad +\mathbb{E}[\tilde{S}_i(\tilde{S}_j \sum_{k(\neq j)} K(x_j - x_k)\mathbb{E}[I_k])] \\
&\quad +\mathbb{E}[\tilde{S}_i(\mathbb{E}[S_j] \sum_{k(\neq j)} K(x_j - x_k)\tilde{I}_k)] \\
&\quad +\mathbb{E}[\tilde{S}_i(\tilde{S}_j \sum_{k(\neq j)} K(x_j - x_k)\tilde{I}_k)] - \mathbb{E}[\gamma\tilde{S}_i\tilde{I}_j] \\
&= c_{SS}(i,j;t) \sum_{k(\neq j)} K(x_j - x_k)\mathbb{E}[I_k] + \mathbb{E}[S_j] \sum_{k(\neq j)} K(x_j - x_k)c_{SI}(i,k;t) \\
&\quad + \sum_{k(\neq j)} K(x_j - x_k)c_{SSI}(i,j,k;t) - \gamma c_{SI}(i,j;t). \qquad (2.46)
\end{aligned}$$

Where I have used the drift and fluctuation representation for each disease indicator process, martingale property of the fluctuation terms and the definition of covariance. The triple covariance is given as $c_{XYZ} = \mathbb{E}[\tilde{X}\tilde{Y}\tilde{Z}]$. Equation (2.46) gives the rich correlation structure for the stochastic disease states of each population.

If the initial disease states of the spatial epidemic are chosen deterministically, or independently at random, then

$$\begin{aligned}
c_{XY}(i,j;0) &= 0, \\
c_{XYZ}(i,j,k;0) &= 0, \\
X,Y,Z \in \{S,I\},\ i,j,k &= 1,\ldots,N.
\end{aligned}$$

In the absence of quadratic covariation, i.e. if the processes are deterministic with differentiable trajectories, equation (2.45) gives that the covariance functions being zero is a fixed point for the covariance dynamics and therefore the covariance remains zero for all $t \geq 0$. For the stochastic spatial epidemic $[X,Y] \neq 0$ and covariance effects enter the dynamics irrespective of the choice of initial condition. We can now identify quadratic covariation as the fundamental source of the deviation between the true expected dynamics stochastic processes and approximating deterministic processes such as the correlation closed dynamics (2.35) and (2.36).

The quadratic covariation process is also the dynamical mechanism via which the fundamental discreteness of the population dynamics is enforced. From the definitions (2.17) and (2.18) we can see that $S_i(0)I_i(0) = 0$ for any physically admissible initial distribution; populations are either in disease state S or they are not. It is also obvious that this extends into the future, since the dynamics preclude population i being in both state S and I simultaneously. However, if the chain rule of classical calculus is naively applied to the pair dynamics we arrive at the false statement,

$$d(S_iI_i) = (S_i^2 - S_iI_i)\sum_{k(\neq i)} I_k dN_{ik}^I - S_iI_i dN_j^R,$$

which would predict that $S_i(t)I_i(t)$ is *not necessarily zero* for all time. The stochastic dynamics given by equation (2.40), with the correcting quadratic covariation process gives the correct result,

$$d(S_iI_i) = -S_iI_i dN_i^R - S_iI_i \sum_{k(\neq i)} I_k dN_{ik}^I.$$

So that for any proper initial conditions that respect that populations cannot be simultaneously infected and susceptible, $d(S_iI_i) = 0 \implies S_i(t)I_i(t) = 0,\ t \geq 0$. In the above I used that $S_i^2(t) = S_i(t)$, this is guaranteed for proper initial conditions by equation (2.39) which gives that $dS_i^2 = dS_i$ and hence $S_i^2(t) = S_i(t)\ t \geq 0$ since they are driven by the same underlying Poisson process hitting times. Finally, $I_i^2(t) = I_i(t),\ t \geq 0$. I call these (elementary) results the *discreteness conditions*. On one hand, the discreteness conditions are simply due to $S_i(t), I_i(t) \in \{0, 1\}$ for each i and $t \geq 0$. The point is that by using the correct chain rule in the construction of the dynamical hierarchy these conditions will be automatically satisfied without awkward post-calculation adjustments, see for example (Bolker 1999) for a treatment where the 'standard' chain rule is used and then corrected using an argument based on the fundamental discreteness of the spatial epidemic process.

I finish this section by giving the explicit triple dynamics and mentioning the *cubic covariation process*. The quadratic covariation defines the difference between the expected trajectory of the stochastic spatial epidemic jump process and an approximating deterministic processes; the cubic covariation defines the difference between the true expected trajectory and the expected trajectory of an approximating diffusion process. Applying the Itô formula for pure jump processes 2.2.3 to the

triple product functions,

$$f(X(t)) = X_i(t)Y_j(t)Z_k(t), \qquad X,Y,Z \in \{S,I\},\ i,j,k = 1,\ldots,N,\ t \geq 0,$$

gives the following triples dynamics represented as SDEs.

Triples Dynamics:

$$\begin{aligned} d(S_iS_jS_k) &= -S_iS_jS_k\big(\sum_{l(\neq i)} I_l dN_{il}^I + \sum_{l(\neq j)} I_l dN_{jl}^I + \sum_{l(\neq k)} I_l dN_{kl}^I\big) \\ &\quad +\delta_{ij}S_iS_k \sum_{l(\neq i)} I_l dN_{il}^I + \delta_{ik}S_iS_j \sum_{l(\neq i)} I_l dN_{il}^I \\ &\quad +\delta_{jk}S_iS_j \sum_{l(\neq j)} I_l dN_{jl}^I - \delta_{ijk}S_i \sum_{l(\neq i)} I_l dN_{il}^I. \end{aligned} \tag{2.47}$$

$$\begin{aligned} d(S_iS_jI_k) &= S_iS_jS_k \sum_{l(\neq k)} I_l dN_{kl}^I - S_iS_jI_k dN_k^R - S_iS_jI_k\big(\sum_{l(\neq i)} I_l dN_{il}^I + \sum_{l(\neq j)} I_l dN_{jl}^I\big) \\ &\quad +\delta_{ij}S_iI_k \sum_{l(\neq i)} I_l dN_{il}^I - \delta_{ik}S_iS_j \sum_{l(\neq i)} I_l dN_{il}^I - \delta_{jk}S_iS_j \sum_{l(\neq j)} I_l dN_{jl}^I \\ &\quad +\delta_{ijk}S_i \sum_{l(\neq i)} I_l dN_{il}^I. \end{aligned} \tag{2.48}$$

$$\begin{aligned} d(S_iI_jI_k) &= S_iS_jI_k \sum_{l(\neq j)} I_l dN_{jl}^I - S_iI_jI_k dN_j^R + S_iI_jS_k \sum_{l(\neq k)} I_l dN_{kl}^I \\ &\quad -S_iI_jI_k dN_k^R - S_iI_jI_k \sum_{l(\neq i)} I_l dN_{il}^I - \delta_{ij}S_iI_k \sum_{l(\neq i)} I_l dN_{il}^I \\ &\quad -\delta_{ik}S_iI_j \sum_{l(\neq i)} I_l dN_{il}^I + \delta_{jk}(S_iS_j \sum_{l(\neq j)} I_l dN_{jl}^I + S_iI_j dN_j^R) \\ &\quad -\delta_{ijk}S_i \sum_{l(\neq i)} I_l dN_{il}^I. \end{aligned} \tag{2.49}$$

$$
\begin{aligned}
d(I_iI_jI_k) &= S_iI_jI_k\sum_{l(\neq i)} I_l dN^I_{il} + I_iS_jI_k\sum_{l(\neq j)} I_l dN^I_{jl} + I_iI_jS_k\sum_{l(\neq k)} I_l dN^I_{kl} \\
&\quad -I_iI_jI_k dN^R_i - I_iI_jI_k dN^R_j - I_iI_jI_k dN^R_k \\
&\quad +\delta_{ij}(S_iI_k\sum_{l(\neq i)} I_l dN^I_{il} + I_iI_k dN^R_i) + \delta_{ik}(S_iI_j\sum_{l(\neq i)} I_l dN^I_{il} + I_iI_j dN^R_i) \\
&\quad +\delta_{jk}(S_jI_i\sum_{l(\neq j)} I_l dN^I_{jl} + I_iI_j dN^R_j) + \delta_{ijk}(S_i\sum_{l(\neq i)} I_l dN^I_{il} - I_i dN^R_i).
\end{aligned}
\tag{2.50}
$$

The triple dynamics depend on quadruple dynamics, as expected. The triple dynamics are in the form,

$$
d(XYZ) = XYdZ + XZdY + YZdX + Xd[Y,Z] + Yd[X,Z] + Zd[X,Y] + d[X,Y,Z].
\tag{2.51}
$$

Where $[X,Y,Z]$ is the *cubic covariation process*. The cubic covariation represents a major departure in the structure of the SDEs describing pure jump processes compared to the more familiar diffusion SDEs driven by Brownian motion. Processes driven by Brownian motions have zero cubic covariation, see for example (Oksendal 2004).

The reader should be convinced at this point that dynamics at the level of a size n configuration of disease states will generally depend on size $(n+1)$ configurations as has been found in a number of different contexts, again see (Murrell, Dieckmann, and Law 2004) for a longer discussion. For finite numbers of populations this is not in fact an infinite hierarchy of SDEs since a size N configuration cannot depend on a size $(N+1)$ configuration since there is no way to write a $(N+1)$-fold product of disease state indicators without the discreteness conditions implying that the $(N+1)$-fold product being equivalent to a configuration of at most N size.

2.3.7 The Approximating ODEs for the Spatial Epidemic

In this section I expand the covariance closure from section 2.3.4 by one order and include explicitly the dynamics using the expression for the dynamics of the pair covariances (2.45). The dynamical equation for each pair covariance will be rewritten, using the same method of expanding the dynamical hierarchy equations into drift and fluctuation form and taking expectations as in (2.46).

For N populations located at $\{x_i \in \mathcal{A}\}_{i=1}^N$, this gives $2N$ ODEs for the expected dynamics $\{\mathbb{E}[S]_i(t), \mathbb{E}[I]_i(t)\}_{i=1}^N$ and $3N^2$ ODEs for the pair covariance functions, $\{c_{SS}(i,j;t), c_{SI}(i,j;t), c_{II}(i,j;t)\}_{i,j=1}^N$. For better compactness of representation I define the *force of infection* at population i, $\lambda_i(t)$,

$$\lambda_i(t) = \sum_{j(\neq i)} K(x_i - x_j) I_j(t). \tag{2.52}$$

The approximation step is to truncate at the order of triples, that is to treat all triple covariances at zero.

Approximating ODEs for Spatial Epidemic

$$\frac{d}{dt}\mathbb{E}[S_i](t) = -\mathbb{E}[S_i](t)\mathbb{E}[\lambda_i](t) - \sum_{j(\neq i)} K(x_i - x_j) c_{SI}(i,j;t). \tag{2.53}$$

$$\frac{d}{dt}\mathbb{E}[I_i](t) = \mathbb{E}[S_i](t)\mathbb{E}[\lambda_i](t) + \sum_{j(\neq i)} K(x_i - x_j) c_{SI}(i,j;t) - \gamma\mathbb{E}[I_i](t). \tag{2.54}$$

$$\begin{aligned}\frac{d}{dt}c_{SS}(i,j;t) &= -c_{SS}(i,j;t)(\mathbb{E}[\lambda_j](t) + \mathbb{E}[\lambda_i](t) - \mathbb{E}[S_j](t)\sum_{k(\neq j)} K(x_j - x_k)c_{SI}(i,k;t) \\ &\quad -\mathbb{E}[S_i](t)\sum_{k(\neq i)} K(x_i - x_k)c_{SI}(j,k;t) + \delta_{ij}\mathbb{E}[S_i(t)\lambda_i(t)]. \end{aligned}\tag{2.55}$$

$$\begin{aligned}\frac{d}{dt}c_{SI}(i,j;t) &= c_{SS}(i,j;t)\mathbb{E}[\lambda_j](t) + \mathbb{E}[S_j](t)\sum_{k(\neq j)} K(x_j - x_k)c_{SI}(i,k;t) \\ &\quad -c_{SI}(i,j;t)\mathbb{E}[\lambda_i](t) - \gamma c_{SI}(i,j;t) - \mathbb{E}[S_i](t)\sum_{k(\neq i)} K(x_i - x_k)c_{II}(j,k;t) \\ &\quad -\delta_{ij}\mathbb{E}[S_i(t)\lambda_i(t)]. \end{aligned}\tag{2.56}$$

$$\begin{aligned}\frac{d}{dt}c_{II}(i,j;t) &= c_{SI}(i,j;t)\mathbb{E}[\lambda_i](t) + c_{SI}(j,i;t)\mathbb{E}[\lambda_j](t) - 2\gamma c_{II}(i,j;t) \\ &\quad +\mathbb{E}[S_i](t)\sum_{k(\neq i)} K(x_i - x_k)c_{II}(j,k;t) + \mathbb{E}[S_j]\sum_{k(\neq j)} K(x_j - x_k)c_{II}(i,k;t) \\ &\quad +\delta_{ij}\mathbb{E}[S_i(t)\lambda_i(t) + \gamma I_i(t)]. \\ &\quad i,j = 1,\ldots,N,\ i \neq j,\ t \geq 0. \end{aligned}\tag{2.57}$$

I solve the covariance equations for $i \neq j$ since the self covariance equations can be written in terms of the expected dynamics by using the discreteness conditions,

$$\begin{aligned} c_{SS}(i,i;t) &= \mathbb{E}[S_i^2](t) - \mathbb{E}[S_i]^2(t) = \mathbb{E}[S_i](t)(1 - \mathbb{E}[S_i](t)), \\ c_{SI}(i,i;t) &= \mathbb{E}[S_i I_i](t) - \mathbb{E}[S_i](t)\mathbb{E}[I_i](t) = -\mathbb{E}[S_i](t)\mathbb{E}[I_i](t), \\ c_{II}(i,i;t) &= \mathbb{E}[I_i^2](t) - \mathbb{E}[I_i]^2(t) = \mathbb{E}[I_i](t)(1 - \mathbb{E}[I_i](t)). \end{aligned}$$

Setting the 3-variate covariances to zero is equivalent to the approximation of third moments,

$$\mathbb{E}[XYZ] \approx \mathbb{E}[X]\mathbb{E}[YZ] + \mathbb{E}[Y]\mathbb{E}[XZ] + \mathbb{E}[Z]\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Z].$$

In the literature of dynamic approximations of statistical quantities using only moments rather than covariances, this approximation is referred to as the *power-1 symmetric moment closure* (Murrell, Dieckmann, and Law 2004). Therefore, I also refer to my covariance closure as a power-1 closure scheme. Having closed the correlation hierarchy using a power-1 closure the closed set of $2N + 3N^2$ approximating ODE equations can be solved using standard Runge-Kutta numerical solution techniques.

The pair covariances with $i \neq j$ can now be shown to be order $(1/L^d)$ magnitude quantities when the initial distribution is either determined or uniformly random. Let $c_{XY}(i,j;\cdot)$ be one of the $3N^2$ pair covariances given above, by inspection it has exactly two terms in its vector field that depend of a pair covariance of the form $c_{X'Y'}(k,k;\cdot)$ for some $X', Y' \in \{S, I\}$. Again by inspection the two terms themselves are in the form,

$$\begin{aligned} c_{XY}(i,j,\cdot) &= \pm\mathbb{E}[S_j]K(x_j - x_i)c_{X'Y'}(i,i;\cdot) \pm \mathbb{E}[S_i]K(x_i - x_j)c_{X''Y''}(j,j;\cdot) \\ &= \pm\mathbb{E}[S_j]\frac{k((x_j - x_i)/L)}{L^d}c_{X'Y'}(i,i;\cdot) \pm \mathbb{E}[S_i]\frac{k((x_i - x_j)/L)}{L^d}c_{X''Y''}(j,j;\cdot) + \dots \end{aligned} \tag{2.58}$$

Using the assumption for the transmission kernel that the shape function $k(\cdot)$ is $\mathcal{O}(1)$ we see that the two terms (2.58) are magnitude $\mathcal{O}(L^{-d})$. Therefore by lemma 2.3.2 the difference between each pair covariance $c_{XY}(i,j;\cdot)$ and $c'_{XY}(i,j;\cdot)$ which has the same ODE vector field with the terms (2.58) removed is $\mathcal{O}(L^{-d})$. On the other hand, $c'_{XY}(i,j;\cdot) = 0$ since $c'_{XY}(i,j;0) = 0$ for all $i \neq j$ (i.e. for determined or uniformly random initial conditions) is a fixed point for the covariance dynamics

less (2.58). Ergo, $c_{XY}(i,j;t) \sim \mathcal{O}(L^{-d})$.

By bounding the size of the covariance and considering equations (2.53) and (2.54) I give that leading correction to the expected dynamics due to stochasticity, which I denote χ for each disease indicator to be

$$\chi \sim \mathcal{O}(L^{-d} \sum_{j(\neq i)} K(x_i - x_j)). \tag{2.59}$$

Equation (2.59) reflects that the stochasticity alters the expected dynamics of $S_i(t), I_i(t)$ for each i, compared to related deterministic models, through the covariance between the susceptibility of the ith habitat $S_i(t)$ and the stochastic force of infection $\lambda_i(t)$ felt at that habitat.

Equation (2.59) also suggests that, in the 'mean-field' limit $L \to \infty$ the deterministic approximating equations are exact, so long as $\sum_{j(\neq i)} K(x_i - x_j)$ remains finite, e.g. the habitats are not all located at the same spatial location with $N \to \infty$. However, I have also fixed the effort of infection for the transmission kernel on $\mathcal{A}$, K_0, where

$$K_0 = \int_{\mathcal{A}} K(x)dx. \tag{2.60}$$

Therefore, for finite domains $\mathcal{A}$ the limit $L \to \infty$ for admissible transmission kernels gives

$$\lim_{L\to\infty} K(x) = \frac{K_0}{l^d}. \tag{2.61}$$

By comparison to the classical SIR model given in the introduction we see that the restriction to transmission kernels with L-invariant K_0 is analogous to frequency dependent transmission.

2.4 Numerical Examples

In this section I give two examples of using the dynamical hierarchy presented above to construct an approximation to the expected dynamics of a spatial epidemic process as a closed set of ODEs depending on the expected dynamics themselves and the pair covariances. This is essentially the same as the approach demonstrated in section 2.3.4 but rather than using the closure scheme of treating all pair covari-

ances as zero I instead close the dynamical hierarchy by assuming that all triple covariances are zero. This explicitly includes the contribution from quadratic covariation into the expected dynamics and as we have seen is a fundamentally better approximation to the expected dynamics.

As throughout this chapter I treat the habitat spatial locations $\{x_i \in \mathcal{A}\}_{i=1}^N$ as known data. For many applications this is a plausible assumption, a classic example being the spread of disease between livestock in UK farms where detailed spatial data is available. Another possibility is that the spatial interactions are considered to be homogeneous, in which case detailed knowledge of spatial locations is not important for making theoretical predictions. I will consider both this 'mean-field' scenario and a more interesting example modelling the invasion of a pathogen from an initial infection event into a spatial meta-population. In both cases I will demonstrate that the approximating ODEs can essentially recover the expected spatial dynamics and covariances.

I now apply the ODE approximating equations to the two models: the mean field dynamics and a spatially explicit epidemic.

2.4.1 Mean field Dynamics

In this section I consider a simple test case for the theory, the classic stochastic SIR process on N homogeneously mixing individuals (Bailey 1950). As we have seen removing space via taking the $L \to \infty$ limit fixed the transmission kernel as $K(x) = K_0/l^d$. If I also fix the density of the habitats in $\mathcal{A}$, $N/L^d = 1$ then the classic SIR model is recovered,

$$K(x) = K_0/l^d = K_0/N. \tag{2.62}$$

In this scenario the correlation hierarchy above describes *frequency dependent* transmission with a the reproductive ratio $R_0 = \frac{K_0}{\gamma}$. In the context of this model we can associate the space-less dynamics of the stochastic SIR model with the $L \to \infty$ limit when the 'local' interactions for each habitat become equivalent to the global 'mean field' interaction. For this reason I dub the stochastic SIR process: the *mean-field epidemic* process.

If a further limit $N \to \infty$ is taken then the dynamics of the disease state densities $S(t) = \frac{1}{N}\sum_i S_i(t)$ and $I(t) = \frac{1}{N}\sum_i I_i(t)$ convergence uniformly on com-

pacts in probability (u.c.p) to the classical deterministic density SIR dynamics, as was mentioned in the introduction,

$$\frac{d}{dt}S = -K_0 SI, \tag{2.63}$$

$$\frac{d}{dt}I = K_0 SI - \gamma I. \tag{2.64}$$

The power-1 closure in this context is equivalent to restoring an $\mathcal{O}(1/N)$ correction due to the stochasticity of the finite size SIR model. This can be seen from (2.59) and replacing the L^{-d} scaling with N^{-1},

$$\chi \sim \mathcal{O}(N^{-1} \sum_{j(\neq i)} K_0 N^{-1}) \implies \chi \sim \mathcal{O}(N^{-1}). \tag{2.65}$$

I demonstrate the correction due to using the approximating ODEs by comparing their numerical solution to the expected infection density dynamics, $\mathbb{E}[I] = (\mathbb{E}[I](t), t \geq 0)$, estimated from 10^6 Monte Carlo replicate simulations. Here and in the next section I used the tau-leap method (see (Gillespie 2001) and chapter 4) with a leap size of $\Delta t = 0.01$ to generate the Monte Carlo expected dynamics due to the convenience of taking empirical expectations over the MC replicates at each of the discrete time steps.

Very good agreement was found with the initial density fixed as $I(0) = 0.05$ and for epidemiological parameters, $K_0 = 1.5$, $\gamma = 1$, for both $N = 100$ and $N = 200$, as shown in figure 2.2. The agreement between the approximating ODEs and the observed stochastic behaviour was closer as N was larger; and both are in closer agreement with the prediction of the deterministic SIR equations. The units in time are in average infectious duration, $1/\gamma$.

2.4.2 Stochastic Spread from an Initial Source

In this section I generate $N = 400$ habitat locations on $\mathcal{A} = [-10, 10]^2$ such that $x_1 = 0$ and the other locations x_i for $i > 1$ are dispersed uniformly randomly (figure 2.3). Having generated this distribution of habitats locations they are considered fixed data and all expectations are with respect to the distribution of stochastic spatial epidemic with that habitat distribution. Population 1 is initially infected, and additionally all other habitats i satisfying $|x_i| < 19/\pi$ were initially infected (figure 2.3). This represents a super-spreading event as the initial source of the spatial epidemic, localised around the origin. For the spatial distribution considered

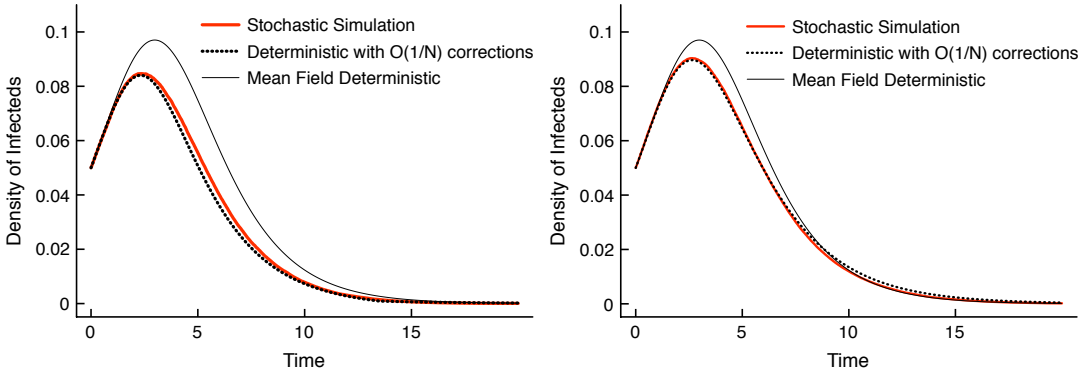


Figure 2.2: The mean-field epidemic with $K_0 = 1.5$, and $\gamma = 1$. A comparison is made between the infection density dynamics of the power-1 closure of the correlation hierarchy (black dotted), the deterministic mean field model (black lines) and the expected density dynamics estimated from 10^6 Monte Carlo replicates. *Left:* For $N = 100$. *Right:* For $N = 200$. Agreement between both the approximating ODEs and the deterministic mean field model with the stochastic simulation is better for larger N as expected from the theoretical considerations.

there were 25 populations initially infected by the super spreading event. Subsequent epidemic dynamics followed the dynamics (2.24) and (2.25). I chose the transmission kernel to be Gaussian shaped with a characteristic length scale L and total effort of infection K_0,

$$K(x) = \frac{K_0}{2\pi L^2} e^{-|x^2|/2L^2}. \tag{2.66}$$

The epidemiological parameters were set to $K_0 = 2$, $L = 3$ and $\gamma = 1$. From (2.59) error due to using the approximating ODEs is expected to be smaller than $\mathcal{O}(L^{-d})$ since for randomly dispersed habitats $\sum_{j(\neq i)} K(x_i - x_j) \sim K_0$.

I investigated the stochastic evolution of the spatial epidemic using both Monte Carlo (MC) simulation and the approximating ODEs. I refer to a case population as any population that became infected over the course of the epidemic. The probability for each population of becoming a case was estimated empirically from 1000 MC replicates of the spatial epidemic with the same fixed spatial distribution, epidemiological parameters and initial super spreading event. The expected dynamics of the susceptible and infected population densities, $\mathbb{E}[S](t)$ and $\mathbb{E}[I](t)$, and the population density variances $\text{var}(S)(t) = \mathbb{E}[S^2(t)] - \mathbb{E}[S]^2(t)$, $\text{var}(I)(t) = \mathbb{E}[I^2(t)] - \mathbb{E}[I]^2(t)$ were also estimated over 1000 MC replicates. The approximating ODE hierarchy was solved using a standard Runge-Kutta ODE numerical solving technique for the $2N$ approximate expected dynamics of the popula-

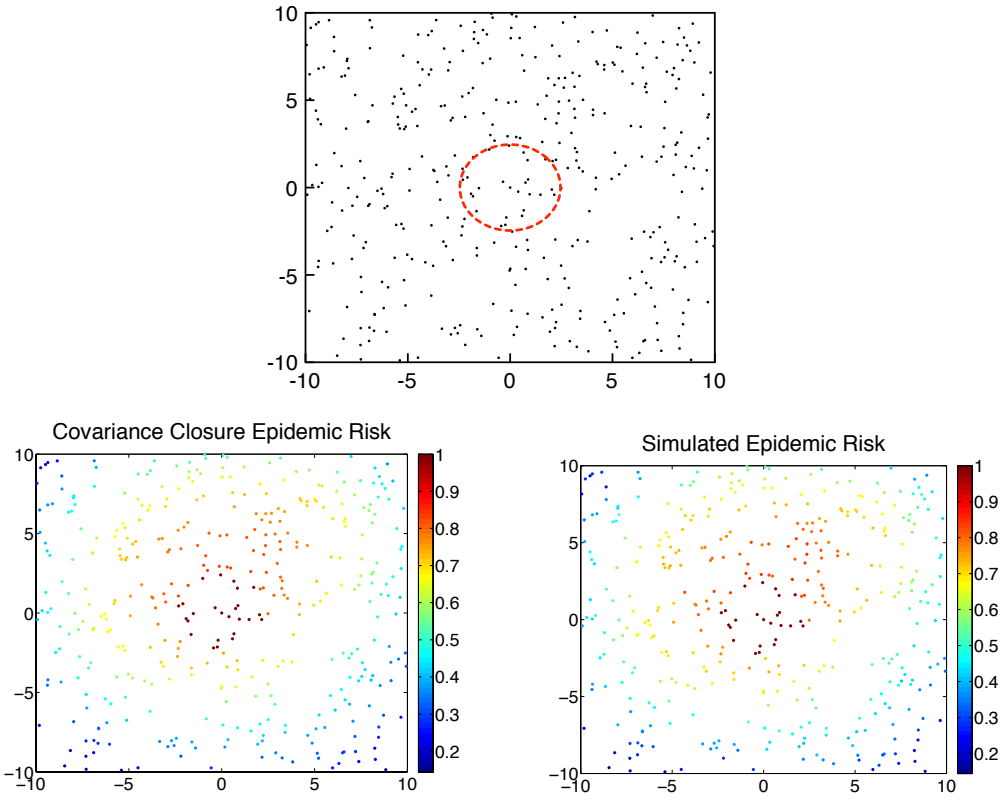


Figure 2.3: A spatial epidemic on $[-10, 10]^2$ with $N = 400$. Transmission rate was given by a gaussian shaped kernel, with effort of transmission, $K_0 = 2$, and characteristic length scale of transmission $L = 3$. The recovery rate was $\gamma = 1$. *Top:* The initially infected populations are those within the area delimited by the red dashed circle. This represents a super spreading event localised around the origin as the initial seed of the spatial epidemic. *Bottom:* The habitats are colour coded by their probability of being a case by the end of the epidemic. Left: The prediction from covariance closure. Right: Probability estimated from 1000 Monte Carlo replicates. Epidemic risk is predominantly governed by locality to the super spreading event with a secondary effect due to local clustering. The mean deviation for the spatial prediction using covariance closure ODEs from the Monte Carlo estimate was 0.0175.

tion disease states and the $3N^2$ covariances between population disease states. The corresponding variance estimates for the the power-1 closure were given by using the identities,

$$\text{var}(S(t)) = \frac{1}{N^2}\sum_{i,j} c_{SS}(i,j;t), \tag{2.67}$$

$$\text{var}(I(t)) = \frac{1}{N^2}\sum_{i,j} c_{II}(i,j;t). \tag{2.68}$$

For the chosen epidemiological parameters I found very good agreement between the approximating ODE hierarchy with power-1 type closure and the results of the MC estimates (Figure 2.4). The approximating ODEs predicted the expected population density dynamics for susceptibles and infecteds very accurately; including key random quantities such as the expected density of cases,

$$Z_\infty = 1 - \mathbb{E}[S](\infty), \tag{2.69}$$

the expected timing of the peak epidemic (2.1 $\times$ expected infection duration), and the expected peak severity ($I_{peak} = 0.105$). The variance of infecteds density was also well recovered, with peak variance closely following peak severity. The variance of the population density of susceptibles was slightly over estimated by the closure approximation; that is although the closure scheme ODEs approximated the expected density of cases well, it predicted that the stochastic spatial epidemic realisations of case burden would be more dispersed around the mean than was in fact found. The approximating ODEs were also able to recreate the detailed probabilities of each population becoming a case (figure 2.4). I define the approximate probability of the ith population becoming a case for the spatial epidemic according to the covariance closure approximating ODES as $p'_{case}(i)$. The mean deviation between the approximate and Monte Carlo estimated case probabilities was small $(1/N)\sum_{i=1}^{N} |p_{case}(i) - p'_{case}(i)| = 0.0175$.

Compared to the mean field prediction derived from the deterministic SIR equations, the spatial epidemic peaked marginally more rapidly and faded more slowly but the epidemic peak was significantly less severe than the mean field prediction. Consequently, the density of cases predicted by the mean field model was $Z_\infty^{MF} = 0.818$ whereas for the spatial epidemic the expected density of cases was found to be $Z_\infty = 0.591$ (Figure 2.4). Whereas a space-less mean field model predicts uniform dispersal of cases (apart from the initially infected populations), the

spatial epidemic predicts that the probability of becoming a case over the course of the epidemic is spatially distributed with the risk being high for populations located nearer the index infecteds and further amplified by local clustering (figure 2.3). The last statement was observed from the stochastic simulations and will be further investigated in the following sections.

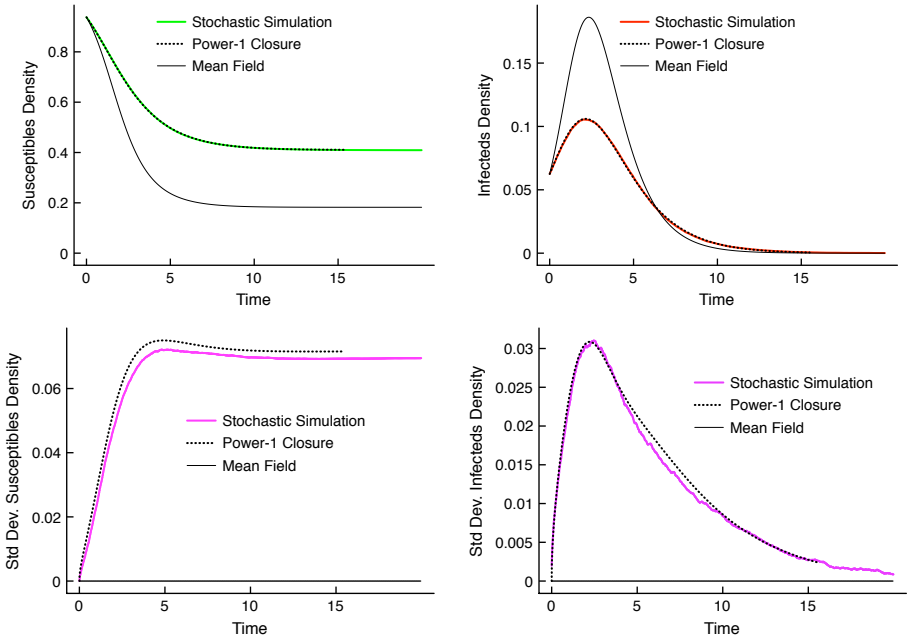


Figure 2.4: The expected dynamics for the spatial epidemic for a given habitat distribution with a super spreading initial event at the centre of the space. Spatial transmission was governed by a Gaussian shaped transmission kernel; epidemiological parameters were $K_0 = 2, \gamma = 1, L = 3$. Shown are $\mathbb{E}[S](t)$ (*top left*), $\mathbb{E}[I](t)$ (*top right*), the standard deviation of the susceptible density, $\sqrt{\text{var}(S)(t)}$ (*bottom left*), and the standard deviation of the infecteds density, $\sqrt{\text{var}(I)(t)}$ (*bottom right*). In all cases the expected dynamics deviated strongly from the space-less mean field prediction and agreed well with the prediction given by the approximating ODE hierarchy described as a power-1 closure, except for an over estimation of the variance of the susceptibles density.

The above gives an interesting example of predicting the expected dynamics of a stochastic spatial epidemic in the absence of common simplifying assumptions such as uniformly dispersed initial infecteds (Cornell and Ovaskainen 2008), periodic boundary conditions or taking a further average over a distribution of habitats. The good agreement between simulation studies and the power-1 type closure scheme of approximating ODEs I have developed are evidence of the solidity of the theoretical underpinnings of my approach. However, there are major practical drawbacks of using the this method for known habitat locations.

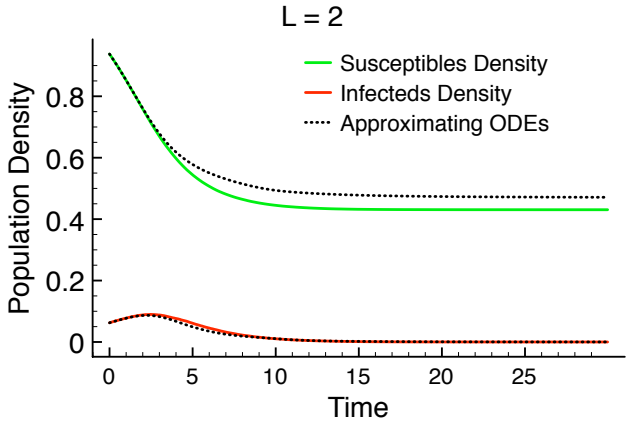


Figure 2.5: The expected dynamics for a spatial epidemic with fixed habitat distribution; characteristic length scale of transmission is $L = 2$, other epidemiological parameters are $K_0 = 2, \gamma = 1$. The ODE approximation to the expected dynamics is less accurate than for the $L = 3$ example.

Firstly, as transmission becomes more localised in space (L decreases) the quality of the approximation using the ODE hierarchy with power-1 type closure decreases (figure 2.5). This could be improved by including the triple dynamics into the truncated dynamical hierarchy and constructing a larger set of approximating ODEs but this run further foul of the second practical problem with this approach, that it is numerically inefficient. The dimensionality of the approximating ODE scheme grows as N^2 and in practice is not significantly more computationally efficient, or even less efficient, than running a large number of MC replicates for the system of interest. Including the triple dynamics would improve the approximation error, but lead to an ODE scheme growing as $\mathcal{O}(N^3)$ worsening the computational inefficiency.

2.5 Discussion

In this chapter I have introduced a stochastic differential equation model for a class of spatial epidemic models where transmission is regulated through a spatially decaying transmission kernel. I will use the model further in later chapters in the context of accelerated simulation and optimal control. Due to explicitly representation as a jump process the joint dynamics of products of disease state indicators could easily be solved. This gives a fresh insight into the so called 'moment hierarchy' that has been well described in the field of theoretical ecology and epidemiology

(Keeling 2000).

The main advantages of this approach are that the quadratic variation of the stochastic epidemic process is given a starring role linking endogenous stochasticity and the fundamental discreteness of the process. By observing that with determined or uniformly random initial conditions the absence of quadratic variation is equivalent to the absence randomness I could bound the magnitude of the covariation between susceptible and infectious populations in terms of the inverse length scale of transmission and thereby bound the effect of the stochasticity on the expected dynamics. The stochastic differential representation also allow me to write down the exact dynamics of the covariance functions; by truncating the hierarchy of their contribution I was able to construct a set of approximating ODEs for the expected dynamics, which in certain circumstances were highly successful at capturing both the metapopulation expected dynamics and the detailed spatial dynamics for individual populations.

The unfortunate aspect of the ODE approximations was that they formed a high dimensional system ($2N+3N^2$ equations), therefore solving the approximating ODEs was often less computationally efficient as a method for approximating the expected dynamics of the epidemic than recursive Monte Carlo simulation. In the next chapter I will use the general construction from this chapter but apply it to a situation where both the epidemic dynamics and the habitat locations are considered to be random. *Prima facie* this would appear to further confound the attempt to appoximate expected dynamics with ODEs, but in fact it remarkably simplifies matters. The cost is that detailed spatial information is lost from the moment dynamics. Detailed spatial modelling will be returned to in chapter 4, but the approach will be to accelerate the Monte Carlo simulation of the spatial epidemic rather than approximate it stochastic dynamics with ODEs.

Chapter 3

Moment Closure and Random Habitat Locations

3.1 Introduction

In chapter 2 using the dynamical hierarchy to approximate the expected dynamics for each habitat was only a partially successful application of the theory developed. For some epidemiological parameter regimes the hierarchy of ODEs for the expectation and pair covariance approximations gave an accurate prediction for the spatial epidemic; however as transmission was modelled as more local the approximation error became more severe. Moreover, the computational burden of solving the system of approximating ODEs meant that the approximate method for investigating the expected dynamics was not greatly more efficient than using standard Monte Carlo methods.

A different approach is to view the habitat locations as random spatial points with known spatial statistics (Diggle 1983; Diggle and Ribeiro 2003). In the following work I will attempt to interweave the effects of stochastic transmission from population to population with random variation in the underlying spatial landscape. The random time-varying quantities of greatest interest will be the epidemic spatial moment dynamics (Bolker and Pacala 1997; Bolker 1999; Herbert and Isham 2000; Keeling 2000). The spatial moments represent an average over the stochastic progression of the epidemic and the random distribution of habitats which characterise the ensemble behaviour of the spatial epidemic (Parham and Ferguson 2005). The moment approach largely avoids the problem of the hierarchy of approximating ODE becoming computationally unwieldily by summarising the set of expected

dynamics into a single spatial density. The price paid for this gain in efficiency is the loss of detailed spatial prediction.

Although detailed dependence on spatial structure will be 'averaged' out in the theory presented in this chapter, nonetheless both the spatial statistics of the assumed underlying habitat location distribution and the spatial characteristics of the transmission kernel will play a crucial, and intertwined, role (Keeling 1999). In particular the effect of spatial clustering on the dynamics of an invasive pathogen have been emphasised (Brown 2004; Tildesley, House, and Bruhn 2010). Moment dynamics, without the restriction to a single habitat location distribution, are the ideal framework within which to explore more general issues. In this chapter I will investigate the following questions:

1. How do the epidemic moment dynamics differ for spatially clustered location distributions compared to the uncorrelated case?

2. For a fixed characteristic length scale of transmission L how do the epidemic moment dynamics differ for a 'heavy-tailed' transmission kernel (e.g. power-law-shaped) compared to a spatially localised transmission kernel (e.g. Gaussian-shaped)?

3. What effect do combinations of the above factors have on the epidemic moment dynamics?

As in chapter 2 the moment dynamics will be approximated by a set of ODEs derived from the iterated expansion of the underlying stochastic model. For such a theory to be functional the dynamical hierarchy of underlying stochastic differential equations had to be truncated such that taking expectations led to a closed set of equations. It is important to be able to assess under which circumstances, e.g. choice of parameter regime, model selection etc., the ODE-based approximates have small error compared to the true statistical observables.

The first requirement is covered by choosing some *moment closure scheme* (Bolker and Pacala 1997). A moment closure scheme is a choice of truncation protocol for ridding the moment hierarchy of high order moments. Standard approaches are either treating a higher order moment as zero, for example the third central moment (Bolker 1999), or as a mixture of lower order moment, in analogy to pairwise closure schemes for network models (Rand 1999; Ellner 2001). The accuracy requirement is fulfilled when a moment hierarchy is chosen such that it is consistent

with the microscopic random epidemiological dynamics of the individuals of the population under consideration. A consistent moment hierarchy in this context can be loosely defined as one for which inclusion of increasingly many orders of moment leads to eventual convergence onto the true values of the moments in the hierarchy.

In this chapter I define the random spatial habitat location distribution as being drawn from a particularly simple form of spatial cluster process; a special case of the more general bivariate spatial point processes analysed by Diggle and Milne (Diggle and Milne 1983). I then turn to applying the theory developed in chapter 2. The application area will be to a spatial epidemic with cases that are initially dispersed uniformly over a random landscape of habitat locations. An approximating deterministic evolution equation for the spatial statistics of epidemic with bounded error is derived using a moment closure scheme built up directly from the microscopic stochastic dynamics. In the numerical investigation that follows three different choices of moment closure scheme are compared to simulation. It is found that a combination of a power-law shaped transmission kernel with a spatially clustered underlying habitat distribution was a severe amplifier for the epidemic, compared to both mean-field and more local disease dispersion, at virtually all length scales of transmission.

3.2 Derivation of the Moment Dynamics

In this section, I extend the correlation hierarchy methodology of chapter 2 in order to approximate the moment dynamics of the spatial epidemic. I use the term moment to mean a combination of taking an expectation over stochastic realisations of an epidemic and simultaneously over some random distribution of habitat locations, which may be a clustered distribution. Analogous to the population dynamics model of Cornell and Ovaskainen (Cornell and Ovaskainen 2008) the magnitude of the spatial pair correlations will be $\mathcal{O}(L^{-d})$ where L is the characteristic length scale of transmission. I will also present two possible improvements to the power-1 type correlation closure used in chapter 2: a power-2-type closure based on approximating triple covariance functions using a quadratic form of pair covariances and and an extension to include the dynamics of triple covariances but treating quadruple covariances as zero.

The precise order of the derivation will be,

1. Introduce the distribution form for the habitat locations, define a model for

random spatial coordinates and derive its spatial covariance function (section 3.2.1)

2. Introduce the distribution form for the epidemic as a marked version of the habitat distribution and define the spatial moments for the epidemic process (section 3.2.2).

3. Use the dynamical hierarchy derived in section 2.3.6 to derive analogous dynamics for the epidemic moments (section 3.2.4).

4. Transform the moment dynamics from the real space domain to the frequency domain in order to exploit improved performance (section 3.2.6).

5. Choose a closing approximation to truncate the correlation hierarchy; this involves deriving the dynamics of higher order covariances (section 3.3 and appendix B).

6. Solve approximating closure approximation on a collocation grid in the frequency domain using standard ODE techniques (section 3.4).

3.2.1 The Spatial Distribution of Habitats

In this section I introduce the random spatial model that I use to define the locations of the habitats for the spatial epidemic metapopulation model. By choosing a particularly simple model I can incorporate spatial clustering whilst retaining analytical traction and the ability to numerically generate random samples with considerable ease.

I encode the spatial location of the habitats into the *habitat distribution*, H, defined by,

$$H(x) = \sum_{i=1}^{N} \delta(x - x_i). \tag{3.1}$$

Here, $\delta(\cdot)$ is the *delta distribution* defined as a linear map from the space of smooth functions on the reals, $\mathcal{C}^\infty(\mathbb{R})$, to $\mathbb{R}$. The map is defined by the property that,

$$\int \delta(x-a)\phi(x)dx = \phi(a), \qquad \phi \in \mathcal{C}^\infty(\mathbb{R}),\ a \in \mathbb{R}.$$

When considering multidimensional spatial co-ordinates ($d > 1$) the delta distribution is to be understood as a product distribution over the number of dimensions. For

example with $d = 2$ then $x = (x^{(1)}, x^{(2)})$ in vector notation and $\delta(x) = \delta(x^{(1)})\delta(x^{(2)})$.

Representing the spatial locations of habitats as a sum of delta distribution is a popular modelling choice (Bolker 1999; Dieckmann and Law 2000; Ovaskainen and Cornell 2006b), its convenience lies in reducing spatial observables to integrals over spatial variables which are comparatively easy to manipulate. In particular, if the spatial locations $\{x_i\}_{i=1}^N$ are modelled as random variables then analysis of statistical quantities can be reduced to considering the (now) random measure $H(x)dx$.

Random Habitat Locations

For many real world applications the habitat spatial locations of interest are accessible data. An example might be the forecasting of epidemic dynamics of an invasive pathogen into the UK farming environment where farm locations are available to the theoretician (Ferguson, Donnelly, and Anderson 2001a). On the other hand there are many modelling applications where spatial information about the habitats is limited to statistical estimates such as density of habitats in observed areas, or estimates pertaining to spatial covariances. From a forecaster's point of view this might be due to incomplete field data. From a theoretician's perspective one might be interested in fundamental questions such as 'what role does spatial clustering play in amplifying epidemic size'(Tildesley, House, and Bruhn 2010). In either scenario forecasting risk for individual habitats separate from the population bulk becomes impossible; instead one concentrates on forecasting the dynamics of averaged quantities, such as the expected burden of the epidemic on the population averaged over the statistical distribution of habitat locations.

I denote the joint probability density of the random habitats, $p_N(x_1, \cdots, x_N)$, and make the following restrictions on the joint density,

- The joint density is smooth, $p_N(x_1, \cdots, x_N) \in \mathcal{C}^\infty((\mathbb{R}^d)^N)$.
- Each univariate marginal density is uniform on $\mathcal{A}$,

$$\begin{aligned} p_1(x_i) &= \int_{\mathcal{A}} p_N(x_1, \cdots, x_N) \prod_{j(\neq i)} [dx_j] \\ &= 1/l^d, \qquad i = 1, \ldots, N. \end{aligned} \tag{3.2}$$

- The habitats are interchangeable in the sense that $p_N(x_1, \ldots, x_N) = p_N(x_{\pi(1)}, \ldots, x_{\pi(N)})$ for any permutation $\pi \in S^N$.

I now define spatial statistical quantities of interest as averages of the habitat distribution with respect to the joint density of habitat locations. To avoid confusion with expectations over the random dynamics of epidemic progression, I denote averages over habitat locations by $\mathbb{E}_{\{x_i\}}[f] = \overline{f}$. Firstly, the expected number of habitats in a given spatial observation area, $\mathcal{A}_1 \subset \mathcal{A}$ of volume $|\mathcal{A}_1|$,

$$\begin{aligned}\overline{\int_{\mathcal{A}_1} H(x)dx} &= \sum_i \int_{\mathcal{A}_1} \int \delta(x - x_i) p_N(x_1, \cdots, x_N) \prod_i [dx_i]dx \\ &= \sum_i \int_{\mathcal{A}_1} p_1(x)dx \\ &= |\mathcal{A}_1|N/l^d = |\mathcal{A}_1|. \end{aligned} \tag{3.3}$$

This confirms that the expected density of habitats is isotropic on $\mathcal{A}$; the expected density of habitats in any observation area is equal to the global density which has been fixed to be 1. Also, one might consider the joint observation, (#Habitats in $\mathcal{A}_1$, #Habitats in $\mathcal{A}_2$).

$$\begin{aligned}\overline{\int_{\mathcal{A}_1}\int_{\mathcal{A}_2} H(x)H(y)dydx} &= \sum_{ij} \int_{\mathcal{A}_1}\int_{\mathcal{A}_2} \int \delta(x - x_i)\delta(y - x_j) p_N(x_1, \cdots, x_N) \prod_i [dx_i]dydx \\ &= \sum_i \int_{\mathcal{A}_1}\int_{\mathcal{A}_2} \delta(x - y)p_1(y)dydx + \sum_{i\neq j} \int_{\mathcal{A}_1}\int_{\mathcal{A}_2} p_2(x, y)dydx \\ &= |\mathcal{A}_1 \cap \mathcal{A}_2| + N(N-1)\int_{\mathcal{A}_1}\int_{\mathcal{A}_2} p_2(x, y)dydx. \end{aligned} \tag{3.4}$$

I define the two-point spatial covariance, $c_{HH}(x, y)$, between the density of habitats at two locations x, y, as the average joint density of the centred spatial distribution $\tilde{H}(x) = H(x) - 1$,

$$c_{HH}(x, y) = \overline{\tilde{H}(x)\tilde{H}(y)}. \tag{3.5}$$

In the case where each habitat location is independent (and N is large) then we can compute the spatial covariance between (#Habitats in $\mathcal{A}_1$, #Habitats in $\mathcal{A}_2$).

$$\overline{\int_{\mathcal{A}_1}\int_{\mathcal{A}_2} \tilde{H}(x)\tilde{H}(y)dydx} = \overline{\int_{\mathcal{A}_1}\int_{\mathcal{A}_2} H(x)H(y)dydx} - |\mathcal{A}_1||\mathcal{A}_2| = |\mathcal{A}_1 \cap \mathcal{A}_2|. \tag{3.6}$$

Unsurprisingly, the covariance between disjoint observation areas is zero for independent habitat locations. More importantly the result above implies that for

independently distributed habitat locations.

$$c_{HH}(x,y) = c_{HH}(x-y) = \delta(x-y). \tag{3.7}$$

The interpretation of (3.7) is that the observation of a habitat at some $x \in \mathcal{A}$ guarantees that one will find a habitat at $y = x$ (the observed habitat!) but does not effect the probability density of finding a habitat at any other spatial location. The delta distribution in the covariance function is ofter removed or corrected (Dieckmann and Law 2000; Murrell, Dieckmann, and Law 2004) as inconvenient. In this work I avoid corrected covariance equations; the delta distributions will be the spatial quadratic covariation for the epidemic spatial moment theory.

Generating a sample of N habitat locations independently is a trivial application of the N-fold drawing from the uniform marginal densities $p_1(x) = 1/l^d$. However, I am also interested in spatial distributions with with more general covariance functions,

$$c_{HH}(x,y) = c_{HH}(x-y) = \delta(x-y) + C(x-y). \tag{3.8}$$

Where $C(\cdot)$ is a *clustering kernel* that, given the observation of a habitat at x shifts the probability density of finding further habitats at separations $x - y$.

Poisson Cluster Process

I choose a simple method for generating a collection of N habitat locations with a given two point spatial covariance function; a *simplified Poisson cluster process* (Diggle 1983). The method is based on a bivariate spatial process, one generating *parent* points and a second generating *child* points. The children points form the sample of N habitat locations, whereas the parent points are discarded. The parent points are generated uniformly in $\mathcal{A}$. Each parent point $\{p_i \in \mathcal{A}\}$ generates a deterministic number of children points (N_c), the spatial location of the jth child of the ith parent, c_{ij}, being dispersed around its parent according to a multivariate Gaussian of variance $\psi^2 \boldsymbol{I}_d$: $c_{ij} \sim \mathcal{N}(p_i, \psi^2 \boldsymbol{I}_d)$. The separation between two children points sharing the same parent is therefore also distributed as a Gaussian,

$$c_{ik_1} - c_{ik_2} = (c_{ik_1} - p_i) + (p_i - c_{ik_2}) \sim \mathcal{N}(0, 2\psi^2 \boldsymbol{I}_d). \tag{3.9}$$

Expression (3.9) gives that the conditional density of a child point location c_{ik_2} given the location of a child point dispersed from the same parent c_{ik_1} is,

$$p_{1|1}(c_{ik_2}|c_{ik_1}) = \frac{1}{(4\pi\psi^2)^{d/2}} e^{-(c_{ik_1}-c_{ik_2})^2/4\psi^2}.$$

Bayes' theorem and the independence of children dispersed from different parent points now give the bivariate spatial densities,

$$p_2(c_{ik_1}, c_{jk_2}) = \delta_{ij} p_{1|1}(c_{ik_2}|c_{ik_1}) l^{-d} + (1-\delta_{ij}) l^{-2d}. \tag{3.10}$$

In order to generate N randomly located habitat positions the number of parent points, N_p is always chosen such that $N_p N_c = N$. The union of all children points forms the set of habitat locations, the parent points are discarded. The indices of each child point are randomised to return a random vector of habitat locations $(x_1, \ldots, x_N)$ with the interchangeability property.

The important parameters for the simplified Poisson cluster process presented here in terms of determining the cluster kernel are the number of children per parent point, N_c, and the parent to child dispersal scale, ψ. Returning to the spatial covariance calculation we find that for a habitat locations drawn from the simplified Poisson cluster process the cluster kernel C can be derived analytically,

$$\begin{aligned} c_{HH}(x,y) &= \overline{\tilde{H}(x)\tilde{H}(y)} \\ &= \delta(x-y) + N(N_c-1)p_{1|1}(x|y)l^{-d} + N(N-N_c-2)l^{-2d} - 1 \\ &\underset{N\to\infty}{\to} \delta(x-y) + (N_c-1)p_{1|1}(x|y). \end{aligned} \tag{3.11}$$

Where I have used that for each of the N habitats only $N_c - 1$ children habitats of the same parent point are spatially correlated to the habitat. The final limit is a large N limit with N_c fixed as finite. From the derivation above it is now clear that generating a random set of habitat locations from the simplified Poisson cluster process is probabilistically equivalent to choosing the cluster kernel,

$$C(x-y) = \frac{(N_c-1)}{(4\pi\psi^2)^{d/2}} e^{-(x-y)^2/4\psi^2}. \tag{3.12}$$

Note that the choice $N_c = 1$, in addition to imposing wrap round boundary conditions in dispersal, is equivalent to the uncorrelated case (3.7). Two examples of the Poisson cluster process are given in figure 3.1.

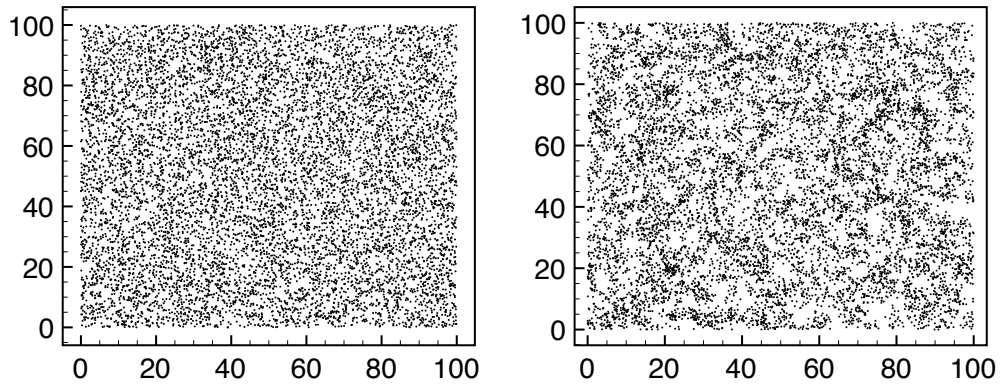


Figure 3.1: Two realisations of the Poisson cluster process with a deterministic number of children points (N_c) for each parent point. Dispersal of the children points from their parent point was Gaussian distributed with standard deviation ψ in each spatial direction. The total number of children points is $N = 10,000$ with units chosen such that $l = 100$. *Left:* $N_c = 1$. This is equivalent to the standard independent Poisson spatial process. *Right:* $N_c = 5$, $\psi = 1$. The children points are tightly correlated to their parent position, inducing strong clustering in the spatial distribution of habitats.

3.2.2 The Distribution Representation of the Epidemic

Having fixed the random spatial model for the fixed locations of habitats I turn to representing the spatial epidemic dynamics upon the random spatial model. The simplest method for this representation is to assign each habitat a disease state indicator function as a 'mark' (Stoyan and Penttinen 2000). The sum of delta distributions corresponding to the locations of the habitats for respectively susceptible and infectious populations will be the underlying stochastic dynamic variables from which I derive the spatial moment equations. The convenience of this representation is that the dynamics for the spatial epidemic distributions can be written in closed form with only two equations. Moreover, the spatial transmission rate for any location $x \in \mathcal{A}$ is compactly represented as a convolution integral.

I follow Bolker (Bolker 1999) and Ovaskainen and Cornell (Ovaskainen and Cornell 2006b; Ovaskainen and Cornell 2006a) in strongly reducing the dimensionality of the process, albeit at the cost of distribution representation, by describing the full solution to the $2N$ SDEs (2.24) and (2.25) as two coupled stochastic spatially

extended distributions,

$$S(x,t) = \sum_{i=1}^{N} \delta(x - x_i) S_i(t), \tag{3.13}$$

$$I(x,t) = \sum_{i=1}^{N} \delta(x - x_i) I_i(t), \qquad x \in \mathbb{R}^d, t \geq 0. \tag{3.14}$$

I call equations (3.13) and (3.14) the *distribution representation* of the epidemic; they are a marked version of the spatial distribution of habitat locations discussed in section 3.2.1 encoding respectively the spatial locations of habitats with a susceptible disease state and those with an infectious disease state. A spatial distribution for recovereds is unnecessary since,

$$R = H - S - I. \tag{3.15}$$

Note that despite the reduced representation no information has been discarded in this representation, for example the infectiousness of habitat i at time t can be, in principle, calculated via the relation $I_i(t) = \int_{\mathcal{B}_\delta(x_i)} I(x,t)dx$ where $\mathcal{B}_\delta(x_i) \subset \mathcal{A}$ is a ball around x_i of sufficiently small radius δ that only habitat i is contained within.

The distribution representation included the $2N$ disease states of the spatial epidemic, which I am treating respectively as solutions to the stochastic integrals (2.21) and (2.22). I use theorem 2.3.1 to define a differential form for the distribution representation,

$$dS(x,t) = \sum_{i=1}^{N} \delta(x - x_i) dS_i(t), \tag{3.16}$$

$$dI(x,t) = \sum_{i=1}^{N} \delta(x - x_i) dI_i(t). \tag{3.17}$$

In the model used in this work the habitat locations are fixed irrespective of the disease dynamics and I interchange expectation, conditioned on the habitat distribution H, and the delta distributions e.g.

$$\mathbb{E}[dS(x,t)|H] = \sum_{i=1}^{N} \delta(x - x_i) \mathbb{E}[dS_i(t)|H] = \sum_{i=1}^{N} \delta(x - x_i) \mathbb{E}[r_{S_i}(t)|H]dt. \tag{3.18}$$

I return to the stochastic dynamics written in drift and fluctuation form (2.27)

and (2.28); the expected stochastic dynamics for the infectiousness of population i condition on some habitat distribution H is,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[I_i|H](t) &= \mathbb{E}[S_i \sum_{j(\neq i)} K(x_i - x_j)I_j|H](t) - \gamma\mathbb{E}[I_i|H](t) \\ &= \mathbb{E}[S_i \sum_{j} K(x_i - x_j)I_j|H](t) - \gamma\mathbb{E}[I_i|H](t) \\ &= \mathbb{E}[S_i(t)K * I(x_i, t)|H] - \gamma\mathbb{E}[I_i|H](t). \qquad (3.19)\end{aligned}$$

In the above I have used the discreteness condition that $S_i(t)I_i(t) = 0$ to complete sum in the first line and that $*$ denotes a convolution over the spatial variables of a function and distribution,

$$K * I(x,t) = \int_{\mathcal{A}} K(x-y)\sum_j \delta(x - x_i)I_i(t)dx = \sum_j K(x - x_i)I_i(t).$$

The expected dynamics for susceptible disease states can be similarly rewritten in terms of a convolution over the spatial variables. Equation (3.19) allows the expected dynamics of the distribution representation to be written in closed form,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[S|H](x,t) &= -\mathbb{E}[SK * I|H](x,t), \qquad (3.20)\\ \frac{d}{dt}\mathbb{E}[I|H](x,t) &= \mathbb{E}[SK * I|H](x,t) - \gamma\mathbb{E}[I|H](x,t). \qquad (3.21)\end{aligned}$$

Note that S is a distribution but $(K * I)$ is a spatial (random) field with the same smoothness properties as K taking values on $\mathbb{R}^d$, therefore their product is well-defined (Friedlander 1998). The distribution dynamics (3.20) and (3.21) encode the full expected dynamics of the $2N$ indicator functions $\{S_i(t), I_i(t)\}_{i=1}^N$ in just 2 equations. However, the 'price' for the reduction in the number of dynamical equations is explicit dependence on the spatial location $x \in \mathcal{A}$.

Equations (3.20) and (3.20) describe spatial distributions that depend on $H = \{x_i \in \mathcal{A}\}_{i=1}^N$. Meaning is assigned to their time derivative by their action on any test function ϕ as described in section 3.2.1; that is for each $\phi \in \mathcal{C}^\infty$, and each

specific choice of landscape $H = h$,

$$\frac{d}{dt}\int_{\mathcal{A}} \mathbb{E}[S|h](x,t)\phi(x)dx = -\int_{\mathcal{A}} \mathbb{E}[SK * I|h](x,t)\phi(x)dx, \tag{3.22}$$

$$\frac{d}{dt}\int_{\mathcal{A}} \mathbb{E}[I|h](x,t)\phi(x)dx = \int_{\mathcal{A}} \Big(\mathbb{E}[SK * I|h](x,t) - \gamma\mathbb{E}[I|h](x,t)\Big)\phi(x)dx. \tag{3.23}$$

I can also define the *spatial moment* for the delta distributions with dynamic disease state 'marking' as the spatial average, as defined in section 3.2.1, of the conditional expectation form of the spatial epidemic distribution,

$$\langle S\rangle = \overline{\mathbb{E}[S|H]}, \tag{3.24}$$

$$\langle I\rangle = \overline{\mathbb{E}[I|H]}, \tag{3.25}$$

The disadvantage of this construction is that the order of expectations cannot be swapped since $\mathbb{E}[\overline{S}|H]$ makes no sense, but this will turn out not to be important for my purposes. I have not defined a probability space to incorporate both types of randomness but rather nested the dynamical stochasticity within the spatial randomness.

3.2.3 Translationally Invariant Initial Conditions and the Spatial Covariances

In section 3.2.1 I assumed that no sub-volume of $\mathcal{A}$ was more or less likely to contain a habitat; consequently the marginal 1 point spatial location density is flat on $\mathcal{A}$, $p_1(x) = l^{-d}$. For the model of clustering used in this work the 2 point spatial location density only depends on the spatial separation $p_2(x,y) = p_2(|x-y|)$. In fact the spatial densities inherit *translational and rotational invariance* from the dispersal Gaussian that defines the law of their location relative to the isotropic spatial parent process.

In this section I extend this to include isotropic initial disease state conditions; that is that no habitat in any sub-volume of $\mathcal{A}$ is more or less likely to be initially infectious. I define the initial probability of being infected as $\mathbb{P}(I_i(0) = 1) = I_0$ independently for $i = 1, \ldots, N$ and similarly for initial susceptibility, $S_0 = 1 - I_0$. This predicts that the initial moment densities of susceptible and infectious popu-

lations are uniform over the space,

$$\langle S(x,0)\rangle = S_0, \tag{3.26}$$
$$\langle I(x,0)\rangle = I_0. \tag{3.27}$$

I define the two-point *spatial covariance function* between the spatial distributions as in (3.5), except that I take the central moments, e.g. defining $\tilde{S}(x,t) = S(x,t) - \langle S\rangle(t)$, rather than the central spatial averages. This gives

$$c_{SS}(x,y;t) = \langle \tilde{S}(x,t)\tilde{S}(y,t)\rangle \tag{3.28}$$
$$c_{SI}(x,y;t) = c_{IS}(y,x;t) = \langle \tilde{S}(x,t)\tilde{I}(y,t)\rangle \tag{3.29}$$
$$c_{II}(x,y;t) = \langle \tilde{I}(x,t)\tilde{I}(y,t)\rangle. \tag{3.30}$$

as the two-point spatial covariance function for locations of habitats with susceptible populations. The spatial covariance functions unify the randomness of the underlying spatial distribution with the stochasticity of the transmission process on that spatial distribution. In the completely isotropic case considered above the initial covariance functions can be found using the same methods as section 3.2.1 whilst respecting the discreteness conditions, which affect the diagonal terms in the sum,

$$c_{SS}(x,y;0) = S_0\delta(|x-y|) + S_0^2 C(|x-y|), \tag{3.31}$$
$$c_{SI}(x,y;0) = S_0 I_0 C(|x-y|), \tag{3.32}$$
$$c_{II}(x,y;0) = I_0\delta(|x-y|) + I_0^2 C(|x-y|). \tag{3.33}$$

Where $C(\cdot)$ is the clustering function (3.12) derived in section 3.2.1. We see that the two-point covariance functions are translationally and rotationally invariant as expected.

3.2.4 The Spatial Moment Dynamics

We have seen that when restricted to isotropic initial conditions the initial spatial disease state covariance functions inherit the covariance structure of the underlying distribution of habitat locations. In this section I will derive the dynamics of the spatial moments, which inherit their dynamic structure directly from the dynamical hierarchy given in section 2.3.6. Deviation from mean field type dynamics will depend crucially on a spatial covariance quantity between susceptible and infectious populations, which gives a measure for the underlying average spatial aggregation of susceptible and infected populations due to both the clustering of the underlying

spatial distribution and the spatial progression of the epidemic process. I will also find that in the setting of spatial moment equations bounding the size of the deviation due to space and stochasticity is particularly simple.

Using the drift and fluctuation expansion such as $S(x,t) = \langle S\rangle(t) + \tilde{S}(x,t)$ and the translation invariance property of the spatial covariances we note that moment of the stochastic recruitment rate of new infecteds,

$$\begin{aligned}
\langle S(x,t)K * I(x,t)\rangle &= \left\langle S(x,t)\int K(x-y)I(y,t)dy\right\rangle \\
&= \langle S\rangle\langle I\rangle(t)\int K(x-y)dy + \int K(x-y)\langle\tilde{S}(x,t)\tilde{I}(y,t)\rangle dy \\
&= K_0\langle S\rangle\langle I\rangle(t) + \int K(x-y)c_{SI}(x-y,t)dy \\
&= K_0\langle S\rangle\langle I\rangle(t) + K * c_{SI}(0,t). \qquad (3.34)
\end{aligned}$$

Where I have used that for the translationally invariant spatial epidemics considered in this work $c_{SI}(x,y,t) = c_{SI}(x-y,t)$. The expected recruitment rate (3.34) allows me to write down the fundamental moment dynamics by taking the spatial averages for equations (3.20) and (3.21), which gives,

$$\begin{aligned}
\frac{d}{dt}\langle S\rangle &= -K_0\langle S\rangle\langle I\rangle - K * c_{SI}(0,t), \qquad (3.35)\\
\frac{d}{dt}\langle I\rangle &= K_0\langle S\rangle\langle I\rangle + K * c_{SI}(0,t) - \gamma\langle I\rangle. \qquad (3.36)
\end{aligned}$$

Compared to the mean field dynamics (2.63) and (2.64), the moment dynamics have a recruitment rate corrected by the term, $K * c_{SI}(0,t)$. The correction term is in the form of an weighted average over spatial covariances between susceptible and infected populations at distances $|x-y|$, for example in two spatial dimensions ($d=2$) with $r = |x-y|$ a transformation into polar co-ordinates gives,

$$K * c_{SI}(0,t) = 2\pi\int_0^\infty rK(r)c_{SI}(r,t)dr. \qquad (3.37)$$

Equation (3.37) is a measure of the spatial aggregation of susceptible and infected populations, as pointed out by Bolker (Bolker 1999). The null case is an uncorrelated distribution of habitat locations with no spatial pattern of disease states; this scenario is consistent with mean field dynamics and indeed predicts $K * c_{SI}(0,t) = 0$ and the moment dynamics (3.35) and (3.36) become indistinguishable from mean field dynamics. $K * c_{SI}(0,t) > 0$ describes spatial clustering between susceptible and

infected populations which accelerates the recruitment of new infected populations compared to mean-field dynamics since they cohabit at closer ranges than predicted by the null case. Conversely, $K * c_{SI}(0,t) < 0$ describes disaggregation between transmissible SI pairs which reduces the recruitment rate compared to the null case. Both clustering and disaggregation will be encountered in this work, broadly speaking clustering is due to the underlying clustering of the habitat locations whilst disaggregation is due to the dynamical effect of infectious populations preferentially depleting the density of susceptible host populations local to themselves compared to those further away.

I now give a derivation of the rate of change of the moment of the joint spatial distribution of susceptible and infected populations; the methodology carries over to all other joint moments used in this work.

$$\begin{aligned}
dS(x,t)I(y,t) &= d\Big(\sum_{i,j}\delta(x-x_i)\delta(y-x_j)S_i(t)I_j(t)\Big)\\
&= \sum_{i,j}\delta(x-x_i)\delta(y-x_j)d(S_i(t)I_j(t))\\
&= \sum_{i,j}\delta(x-x_i)\delta(y-x_j)\Big[S_iS_j\sum_{k(\neq j)}I_k dN^I_{jk}\\
&\quad -S_iI_j dN^R_j - S_iI_j\sum_{k(\neq i)}I_k dN^I_{ik} - \delta_{ij}S_i\sum_{k(\neq i)}I_k dN^I_{ik}\Big]\\
\implies \langle dS(x,t)I(y,t)\rangle &= \Big\langle\sum_{i,j}\delta(x-x_i)\delta(y-x_j)\Big[S_iS_j\sum_{k(\neq j)}K(x_j-x_k)I_k\\
&\quad -\gamma S_iI_j - S_iI_j\sum_{k(\neq i)}K(x_i-x_k)I_k - \delta_{ij}S_i\sum_{k(\neq i)}K(x_i-x_k)I_k\Big]\Big\rangle dt\\
&= \Big\langle\int\sum_{i,j,k}\delta(x-x_i)\delta(y-x_j)\delta(z-x_k)\Big[K(x_j-z)S_iS_jI_k\\
&\quad -K(x_i-z)S_iI_jI_k - \delta_{ij}K(x_i-z)S_iI_k\Big]dz\\
&\quad -\gamma\sum_{i,j}\delta(x-x_i)\delta(y-x_j)S_iI_j\Big\rangle dt\\
&= \Big[\int\Big(K(y-z)\langle S(x)S(y)I(z)\rangle - K(x-z)\langle S(x)I(y)I(z)\rangle\Big)dz\\
&\quad -\delta(x-y)\int K(x-z)\langle S(x)I(z)\rangle dz - \gamma\langle S(x)I(y)\rangle\Big]dt \qquad (3.38)
\end{aligned}$$

Here I have used the discreteness condition $S_i(t)K(0)I_i(t) = 0$ *a.s.* to complete the sums over habitats, the pair dynamics from section 2.3.6 and the drift and

fluctuation decomposition for Poisson integrators,

$$\begin{aligned} dN_{ij}^I(t) &= K(x_i - x_j)dt + \tilde{N}_{ik}^I(dt), \\ dN_i^R(t) &= \gamma dt + \tilde{N}_i^R(dt), \\ i, j &= 1, \ldots, N. \end{aligned}$$

We see that the equation (3.38) is analogous to the result found for $d\mathbb{E}[S_i I_j]$ in section 2.3.6, with sums over possible sources of infection being replaced with integrals over possible locations of infectious populations. The quadratic variation term,

$$[S, I](|x - y|, t) = -\delta(|x - y|) \int K(x - z)\langle S(x, t)I(z, t)\rangle dz,$$

no longer has a Kronecker delta pre-factor, which enforced the discreteness of the spatial epidemic dynamics; this term has been replaced with a delta distribution, which enforces discreteness for the spatial moments.

We see that the joint moment dynamics are analogous to the determined location case; that is the joint dynamics are of the form,

$$d(XY) = YdX + XdY + d[X, Y]. \tag{3.39}$$

Therefore the derivation for the dynamics of the two-variate covariances given by (2.43) can be carried analogously to the study of the two-point spatial moments,

$$\begin{aligned} dc_{XY}(|x - y|; t) &= \langle \tilde{X}(x, t)dY(y, t)\rangle + \langle \tilde{Y}(y, t)dX(x, t)\rangle + \langle d[X, Y](|x - y|, t)\rangle \\ & \qquad X, Y = \{S, I, R\},\ t \geq 0. \end{aligned} \tag{3.40}$$

Equation (3.40) allows one to calculate the dynamics for the two-point spatial covariances as a function of time and the distance $|x - y|$, that is $c_{XY}(x, y, t) = c_{XY}(|x - y|, t)$,

$$\begin{aligned} \frac{d}{dt}c_{SS} &= -2\Big[K_0\langle I\rangle c_{SS} + \langle S\rangle K * c_{SI} + \int K(x - z)c_{SSI}(x, y, z)dz\Big] \\ &\quad +\delta(x - y)[K_0\langle S\rangle\langle I\rangle + K * c_{SI}(0, t)]. \end{aligned} \tag{3.41}$$

$$\begin{aligned}\frac{d}{dt}c_{SI} &= -K_0\langle I\rangle c_{SI} - \langle S\rangle K * c_{II} - \int K(x-z)c_{SII}(x,y,z)dz \\ &\quad +K_0\langle I\rangle c_{SS} + \langle S\rangle K * c_{SI} + \int K(y-z)c_{SSI}(x,y,z)dz \\ &\quad -\gamma c_{SI} - \delta(x-y)[K_0\langle S\rangle\langle I\rangle + K * c_{SI}(0,t)]. \end{aligned} \tag{3.42}$$

$$\begin{aligned}\frac{d}{dt}c_{II} &= 2\Big[K_0\langle I\rangle c_{SI} + \langle S\rangle K * c_{II} + \int K(x-z)c_{SII}(x,y,z)dz \\ &\quad -\gamma c_{II}\Big] + \delta(x-y)[K_0\langle S\rangle\langle I\rangle + K * c_{SI}(0,t) + \gamma\langle I\rangle]. \end{aligned} \tag{3.43}$$

It is worth noting that the covariance dynamics written above are similar, but not identical to, those written down by other authors. This is due to an under-appreciation of the role of quadratic covariation, which as we have seen enforces the fundamental discreteness of the spatial epidemic. This leads authors to treat joint density dynamics of the form $d\langle XY\rangle$ using the familiar rule of calculus $d\langle XY\rangle = \langle XdY\rangle + \langle YdX\rangle$ see for example, (Bolker 1999) equation (16) or (Bolker and Pacala 1999) equation (A6).

3.2.5 Error Analysis

The error analysis of the spatial moment equation can proceed by bounding the size of the contribution due to space and stochasticity. Firstly, we note that for each range $|x-y|$ and each time a sum of the spatial covariances is conserved as the spatial covariance at that range for the underlying habitat distribution,

$$\begin{aligned}c_{HH}(|x-y|) &= \langle\tilde{H}(x)\tilde{H}(y)\rangle = \langle(\tilde{S}+\tilde{I}+\tilde{R})(x,t)(\tilde{S}+\tilde{I}+\tilde{R})(y,t)\rangle \\ &= c_{SS}(|x-y|,t) + c_{II}(|x-y|,t) + c_{RR}(|x-y|,t) + 2c_{SI}(|x-y|,t) \\ &\quad +2c_{IR}(|x-y|,t) + 2c_{SR}(|x-y|,t). \end{aligned} \tag{3.44}$$

Since the spatial covariance of the habitats is invariant in time the above describes a conservation of spatial covariance. Moreover, the initial spatial covariances have been solved in section 3.2.3. Note that,

$$\begin{aligned}\int K(x-y)c_{HH}(x-y)dy &= \frac{1}{L^d}\int k((x-y)/L)\Big(\delta(x-y) + \frac{N_c-1}{(4\pi\psi^2)^{d/2}}e^{-(x-y)^2/4\psi^2}\Big)dy \\ &\leq \frac{N_c k(0)}{L^d}. \end{aligned} \tag{3.45}$$

Where I have used the scaling of the transmission kernel and that its maximum is at zero. I avoid a mored detailed analysis by assuming that the covariance dynamics are

not degenerate in the sense that the covariance functions diverge symmetrically to large positive and negative values. In effect I am assuming that bounding the total contribution of covariance to the integral (3.45) implies a bound on the aggregation measure $K * c_{SI}(0,t)$ of the same order. Therefore neglecting the pair contributions to the moment equations lead to an uniform error of magnitude $\mathcal{O}(1/L^d)$ between the approximate spatial moment dynamics and the true moment dynamics by lemma 2.3.2. The time dependent pre-factor to this order bounding is unfortunately obscure.

This heuristic analysis can be repeated for the contribution of the triple moment dynamics to the pair moment dynamics, e.g.

$$\int K(y-z)c_{SSI}(x,y,z)dz,$$

the magnitude of this integral is also $\mathcal{O}(1/L^d)$ in a non-degenerate situation. Therefore discarding its contribution leads to an $\mathcal{O}(1/L^{2d})$ error for the moment equations.

This non-rigourous argument is essentially the same as that used by Cornell and Ovaskainen to bound the effect of correlations on the equilibrium patch occupancy density a colonising and dying species. Repeating the argument gives that truncating the spatial covariances for disease state configurations of order $n+1$ leads to an $\mathcal{O}(L^{-nd})$ sized uniform error.

3.2.6 The Frequency Domain Dynamics for Spatial Covariances

Physically the spatial covariance equations above describe the dynamic evolution of covariance between the existence of habitats of a given disease state at spatial coordinates $x, y \in \mathcal{A}$. I now take the limit $l \to \infty$, recalling that I also take the $N \to \infty$ limit in such a manner that $Nl^{-d} = 1$. In this section, I present the Fourier transformed dynamics of the spatial covariance functions,

$$\boldsymbol{F}[c_{XY}] = \hat{c}_{XY}(\omega_x, \omega_y, t) = \int c_{XY}(x,y,t)e^{-2\pi i[x\cdot\omega_x + y\cdot\omega_y]}dxdy. \tag{3.46}$$

Where ω_x, ω_y are spatial frequencies in the continuous spectrum, $\mathbb{R}^d$. It is notationally more common to denote the spatial frequency of a wave by k and the temporal frequency by ω, however since the letter k is overburdened in this work I swap notations since confusion with the notation for sample element seems less likely.

There are definite advantages to solving the two-point spatial covariance equations above in terms of their Fourier transforms in the frequency domain. The first advantage is that convolutions in real space are products in the frequency domain, which means that calculations in the Fourier domain are considerably simplified compared to numerically resolving the integral form a convolution over the spatial variables. The second advantage is that any sensible correlation closure I consider is expected to be more accurate, as $L \to \infty$. L sets the length scale of transmission in $\mathbb{R}^d$ since I only consider transmission kernels in families of functions of the form,

$$K(x) = \frac{1}{L^d} k(x/L), \tag{3.47}$$

where $k(\cdot)$ is the shape function of transmission, independent of L. For functions of this type we can identify their Fourier transform in terms of the shape function and L,

$$\begin{aligned} \boldsymbol{F}[K](\omega, t) &= \frac{1}{L^d} \int k(x/L) e^{-2\pi i [x \cdot \omega]} dx \\ &= \int k(x') e^{-2\pi i L [x' \cdot \omega]} dx' \\ &= \hat{k}(L\omega). \end{aligned} \tag{3.48}$$

Equation (3.48) gives that the characteristic scale of the transmission kernel in the frequency domain is $1/L$. This has the benefit that the large L regime calculations in the frequency domain can be done more efficiently since (effectively) a smaller region of spatial frequencies contribute.

Finally, the useful symmetries I derived from translationally invariant initial conditions are preserved in the frequency domain representation. For translation invariance we can directly define a single spatial frequency ω_{x-y} for the separation vector $x - y$, reducing from considering two spatial frequencies ω_x, ω_y,

$$\hat{c}_{XY}(\omega_{x-y}, t) = \int c_{XY}(x - y, t) e^{-2\pi i (x-y) \cdot \omega_{x-y}} d(x - y).$$

I write $\omega = \omega_{x-y}$ when there is no possibility of confusion. This reduces the dimensionality of the Fourier space needed for calculating the two-point spatial covariance dynamics from $2d$ to d. Rotation invariance further reduces dimensionality by one. I demonstrate this for two spatial dimensions by rewriting the spatial frequency $\omega = (|\omega| \cos\psi, |\omega| \sin\psi)$ and the spatial co-ordinate $x - y = (r\cos\theta, r\sin\theta)$ in stan-

dard polar form. If the spatial covariances are rotationally invariant in real-space, their Fourier transforms are in frequency space,

$$\begin{aligned}
\hat{c}_{XY}(\omega) &= \hat{c}_{XY}((|\omega|\cos\psi, |\omega|\sin\psi)) \\
&= \int c_{XY}(|x-y|)e^{-2\pi i|\omega|(\cos\psi(x_1-y_1)+\sin\psi(x_2-y_2))}d(x-y) \\
&= \int_0^\infty rc_{XY}(r)\int_0^{2\pi} e^{-2\pi i|\omega|r(\cos\psi\cos\theta+\sin\psi\sin\theta)}d\theta dr \\
&= \int_0^\infty rc_{XY}(r)\int_0^{2\pi} e^{-2\pi i|\omega|r\sin(\pi/2+\psi-\theta)}d\theta dr \\
&= 2\pi\int_0^\infty rc_{XY}(r)J_0(2\pi|\omega|r)dr \\
&= \hat{c}_{XY}(|\omega|).
\end{aligned}$$

Where I have used standard trigonometric identities. $J_0(\cdot)$ is the 0th Bessel function of the first kind (Abramowitz and Stegun 2006), which in this case only depends on the length of the separation $|x-y| = r$ and the magnitude of the spatial frequency $|\omega|$.

The preservation of symmetry into the Fourier domain allows the power-1 closure, that is neglecting all three-point spatial covariances, of the spatial covariance hierarchy given above to be solved in the frequency domain with great numeric efficiency. The integral form of the correction to the recruitment rate carries over directly,

Power-1 Closure Dynamics in Two Dimensional Frequency Domain:

$$\frac{d}{dt}\langle S\rangle = -K_0\langle S\rangle\langle I\rangle - 2\pi\int_0^\infty |\omega|\hat{k}(L|\omega|)\hat{c}_{SI}(|\omega|,t)d|\omega|, \tag{3.49}$$

$$\frac{d}{dt}\langle I\rangle = K_0\langle S\rangle\langle I\rangle + 2\pi\int_0^\infty |\omega|\hat{k}(L|\omega|)\hat{c}_{SI}(|\omega|,t)d|\omega| - \gamma\langle I\rangle. \tag{3.50}$$

The power-1 closure dynamics for the Fourier transformed covariance functions is

now,

$$\begin{aligned}
\frac{d}{dt}\hat{c}_{SS}(|\omega|,t) &= -2\Big[K_0\langle I\rangle\hat{c}_{SS} + \langle S\rangle\hat{K}\hat{c}_{SI}\Big] \\
&\quad + K_0\langle S\rangle\langle I\rangle + 2\pi\int_0^\infty |\omega|\hat{k}(L|\omega|)\hat{c}_{SI}(|\omega|,t)d|\omega|. \qquad (3.51)\\
\frac{d}{dt}\hat{c}_{SI}(|\omega|,t) &= -K_0\langle I\rangle\hat{c}_{SI} - \langle S\rangle\hat{K}\hat{c}_{II} + K_0\langle I\rangle\hat{c}_{SS} + \langle S\rangle\hat{K}\hat{c}_{SI} \\
&\quad - \gamma\hat{c}_{SI} - K_0\langle S\rangle\langle I\rangle - 2\pi\int_0^\infty |\omega|\hat{k}(L|\omega|)\hat{c}_{SI}(|\omega|,t)d|\omega|. \qquad (3.52)\\
\frac{d}{dt}\hat{c}_{II}(|\omega|,t) &= 2\Big[K_0\langle I\rangle\hat{c}_{SI} + \langle S\rangle\hat{K}\hat{c}_{II} - \gamma\hat{c}_{II}\Big] + K_0\langle S\rangle\langle I\rangle \\
&\quad + 2\pi\int_0^\infty |\omega|\hat{k}(L|\omega|)\hat{c}_{SI}(|\omega|,t)d|\omega| + \gamma\langle I\rangle. \qquad (3.53)
\end{aligned}$$

Where I have used that, $\int \delta(x-y)e^{-2\pi i(x-y)\cdot\omega}d(x-y) = 1$, in order to resolve the delta distributions in the frequency domain and not written the explicit dependence on $|\omega|$ and t for the covariance functions on the right-hand side for brevity of notation. In fact, the delta distributions which enter my theory through the consideration of the quadratic covariation of the underlying stochastic epidemic can now be seen to play a key role in the dynamics.

The closed covariance hierarchy given above has been written in terms of just the frequency magnitudes, $|\omega| \in [0,\infty)$. This gives a tremendous numerical benefit to their solution; the covariances for different frequency magnitudes only interact through the integral term $2\pi\int_0^\infty \hat{k}(L|\omega|)\hat{c}_{SI}(|\omega|,t)d|\omega|$. The spectrum of frequencies I solve for is defined by the accurate numerical integration of this integral term, and we see that for L large this spectrum can be left comparatively small.

3.3 Higher Order Closure Schemes

The error analysis above suggests that including higher order covariance functions into the solution hierarchy should lead to decreased error between the moment closure approximation and the true moment dynamics. Moreover, as we will see in the numerical investigation the power-1 closure scheme for the moment dynamics given above is not immune from problems of stability. In this section I introduce two methods for approximating the contribution of the covariances $\{c_{XYZ}\}$ to the covariance hierarchy in an attempt to construct a more accurate and stable theoretical approach to approximating the moment dynamics using ODEs. In this section I give two higher order closure schemes. The first is a closure scheme based on

attempting to make a good approximation of the dynamics of triple covariances using an approximation based on quadratic combinations of pair covariances; this will turn out to be equivalent to the symmetric power-2 closure already extant in the literature (Murrell, Dieckmann, and Law 2004). The second higher order closure is conceptually simpler. The dynamics of the triple covariances are included directly and the quadruple dynamics are neglected.

Dynamics of the Triple Covariances

Before introducing the higher order approximations I will need to be able to write down the dynamics of the three-point spatial covariance functions. By definition the dynamics are given by,

$$\begin{aligned} dc_{XYZ}(\cdot,t) &= \Big\langle d(\tilde{X}\tilde{Y}\tilde{Z})\Big\rangle(\cdot,t) \\ &= \Big\langle d[(X-\langle X\rangle)(Y-\langle Y\rangle)(Z-\langle Z\rangle)]\Big\rangle(\cdot,t) \\ &= \Big\langle d[XYZ-\langle X\rangle YZ-\langle Y\rangle XZ-\langle Z\rangle XY+\langle X\rangle\langle Y\rangle Z \\ &\quad +\langle X\rangle\langle Z\rangle Y+\langle Y\rangle\langle Z\rangle X-\langle X\rangle\langle Y\rangle\langle Z\rangle]\Big\rangle(\cdot,t). \end{aligned} \tag{3.54}$$

Equation (3.54) can be written in terms of spatial covariance using the stochastic dynamics for pairs and triples derived from dynamical hierarchy of the SIR type epidemic, these are respectively equations (2.42) and (2.51). It is convenient to use the notation from chapter 2 that the integrand of the drift part of a stochastic dynamical system $X = (X(t), t \geq 0)$ be written as r_X, the intensity of X, i.e.

$$\langle dX\rangle = \langle r_X\rangle dt. \tag{3.55}$$

A second notation I use is that the intensity of the quadratic covariation process $[X,Y] = ([X,Y](t), t \geq 0)$ should be written as V_{XY} and similarly for the cubic covariation $[X,Y,Z]$, e.g.

$$\langle d[X,Y]\rangle = \langle V_{XY}\rangle dt, \tag{3.56}$$

$$\langle d[X,Y,Z]\rangle = \langle V_{XYZ}\rangle dt. \tag{3.57}$$

Introducing the above into equation (3.54) and after some algebra gives,

$$\begin{aligned}\frac{d}{dt}c_{XYZ} &= \langle\tilde{X}\tilde{Y}r_Z\rangle - c_{XY}\langle r_Z\rangle + \langle\tilde{X}\tilde{Z}r_Y\rangle - c_{XZ}\langle r_Y\rangle \\ &\quad +\langle\tilde{Y}\tilde{Z}r_X\rangle - c_{YZ}\langle r_X\rangle + \langle\tilde{X}V_{YZ}\rangle + \langle\tilde{Y}V_{XZ}\rangle \\ &\quad +\langle\tilde{Z}V_{XY}\rangle + \langle V_{XYZ}\rangle, \\ &\quad X,Y,Z \in \{S,I,R\},\ t \geq 0.\end{aligned} \tag{3.58}$$

In principle the spatial covariances above are functions of time and also three spatial coordinates x, y, z which refer to the locations of three habitats; for the covariance functions I always order my spatial coordinates in the argument of the function to match the order of the disease state in the subscript before any dimension reduction due to symmetry is imposed. For example, $c_{SII}(x, y, z, t)$ refers to the central moment of finding a habitat at $x \in \mathcal{A}$ with local disease state for its population S and two populations with local disease state I at habitats with coordinates $y, z \in \mathcal{A}$, see figure 3.2.

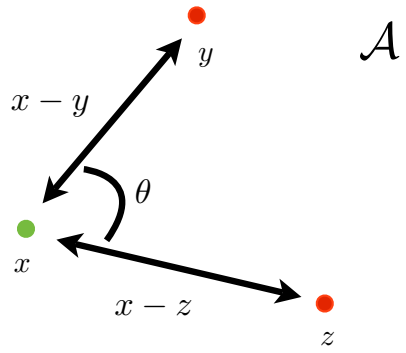


Figure 3.2: Three habitats in $\mathcal{A} = \mathbb{R}^2$. The local disease state of their abiding populations are either S (green) or I (red). Spatial covariance functions referring to both disease state and spatial location are always ordered so that disease state order matches the location order, e.g. the third central moment of finding the state depicted at time t is $c_{SII}(x, y, z, t) = c_{SII}(x, z, y, t) = c_{ISI}(y, x, z, t)$ etc. Translational invariance reduces the dimensionality required to $c_{SII}(x - y, x - z, t)$, if rotational invariance is also imposed we can reduce to $c_{SII}(|x-y|, |x-z|, \theta, t)$. Here θ is the unoriented angle in $[0, \pi]$ between vectors $x - y$ and $x - z$. Note that the triple $(|x - y|, |x - z|, \theta)$ specifies $|y - z|$ by standard trigonometry since θ is an interior angle of triangle with vertices at x, y, z.

3.3.1 The Power-2 Covariance Closure

The error analysis above suggested heuristically that the magnitudes of the two-point covariances in the length scale of transmission are $\mathcal{O}(L^{-d})$, whereas the magnitude

of the three-point covariances are $\mathcal{O}(L^{-2d})$. It is therefore natural to approximate the contribution of the three-point covariances in terms of quadratic combinations of two-point covariances since such an approximation would at least get the order of magnitude (in the inverse scale $1/L$) correct.

Since the space of possible quadratic approximations includes functions of various arguments it makes sense to restrict the set of possible approximations. The restriction I choose is that the three-point spatial covariance function of finding populations in disease states X, Y, Z abiding in habitats at the respective spatial locations x, y, z, will only be approximated by cross-products of the set of two-point spatial covariances for these disease state-location pairs: $\{c_{XY}, c_{XZ}, c_{YZ}\}$; the heuristic idea is that quadratic terms such as $c^2_{XY}(x, y, t)$ do not contain explicit information about the a habitat located at z in disease state Z. As we have seen the dynamics of the three-point covariances involves the formation of quartic-wise interactions between the disease states at four different populations, to allow sufficient approximation flexibility I therefore include time-varying pre-factors $\alpha(t), \beta(t), \gamma(t)$ into my approximation. Putting these considerations together gives a general approximation form:

$$\begin{aligned} c_{XYZ}(x, y, z, t) \approx \quad & \alpha(t)c_{XY}(x, y, t)c_{YZ}(y, z, t) + \beta(t)c_{XY}(x, y, t)c_{XZ}(x, z, t) \\ & +\gamma(t)c_{XZ}(x, z, t)c_{YZ}(y, z, t). \end{aligned} \tag{3.59}$$

The dynamics of the three-point spatial covariances are deterministic and differentiable in time, ergo I require that the time varying pre-factors share these regularity properties. I also require that the pre-factors are $\mathcal{O}(1)$ in terms of $1/L$ so that both the left and right hand sides of the approximation (3.59) match orders in $1/L$. Using equation (3.40) I can write down the time derivative of each term of the approximation, for example,

$$\begin{aligned} \frac{d}{dt}\Big(\alpha(t)c_{XY}c_{YZ}\Big) \quad = \quad & \frac{d\alpha}{dt}c_{XY}c_{YZ} + \alpha c_{YZ}(\langle \tilde{X} r_Y \rangle + \langle \tilde{Y} r_X \rangle + \langle V_{XY} \rangle) \\ & +\alpha c_{XY}(\langle \tilde{Y} r_Z \rangle + \langle \tilde{Z} r_Y \rangle + \langle V_{YZ} \rangle). \end{aligned} \tag{3.60}$$

I denote c'_{XYZ} the approximate three-point spatial covariance function described by equation (3.59) and consider the rate at which the true covariance function and its power-2 approximation vary in time,

$$\frac{d}{dt}\Big(c_{XYZ} - c'_{XYZ}\Big). \tag{3.61}$$

I assume for now that the initial power-2 approximation is more accurate than neglecting the three-point spatial covariance, that is,

$$|c_{XYZ}(x,y,z,0) - c'_{XYZ}(x,y,z,0)| \leq |c_{XYZ}(x,y,z,0)| \qquad x,y,z \in \mathcal{A}. \tag{3.62}$$

The aim is to expand (3.61) and match terms where a judicious choice of $\alpha(t), \beta(t), \gamma(t)$ will lead to partial cancellation. Although, the error due to approximating the three-point spatial covariances will still be $\mathcal{O}(L^{-2d})$ a comparison to just neglecting their contribution (the power-1 approximation) will show that the error must be decreased for the power-2 approximation. This will be an approximation for $c_{XYZ}(x,y,z,t)$ away from line $x = y = z$ and the planes such as $x = y$, so the terms due to quadratic covariation will not be considered.

There are a significant number of terms in the expansion of (3.61), however they collect symmetrically. I start by collecting all terms which dependent on the stochastic rate of change for random spatial distribution of disease states X, r_X, along with the term in the approximation $\frac{d\beta}{dt}c_{XY}c_{XZ}$,

$$\begin{aligned}\frac{d}{dt}(c_{XYZ} - c'_{XYZ}) = \\ \Big[\langle \tilde{Y}\tilde{Z}r_X\rangle - c_{YZ}\langle r_X\rangle\Big] - \Big[\alpha c_{YZ}\langle \tilde{Y}r_X\rangle + \beta c_{XY}\langle \tilde{Z}r_X\rangle + \beta c_{XZ}\langle \tilde{Y}r_X\rangle \\ +\gamma c_{YZ}\langle \tilde{Z}r_X\rangle + \frac{d\beta}{dt}c_{XY}c_{XZ}\Big] + \ldots\end{aligned}$$

Here I rewrite using $\tilde{X} = X - \langle X\rangle$ and similarly for the other random distributions, for example,

$$\langle \tilde{Y}\tilde{Z}r_X\rangle = \langle YZr_X\rangle + \langle Y\rangle\langle Z\rangle\langle r_X\rangle - \langle Y\rangle\langle Zr_X\rangle - \langle Z\rangle\langle Yr_X\rangle. \tag{3.63}$$

Again I match terms with the aim of choosing a set of pre-factors that will lead to

cancellations,

$$\begin{aligned}
\frac{d}{dt}(c_{XYZ} - c'_{XYZ}) = \\
\Big[\langle YZr_X\rangle - \Big(\alpha\langle YZ\rangle\langle Yr_X\rangle + \beta\langle XY\rangle\langle Zr_X\rangle + \beta\langle XZ\rangle\langle Yr_X\rangle + \gamma\langle YZ\rangle\langle Zr_X\rangle\Big)\Big] \\
+\Big[\langle Y\rangle\langle Z\rangle\langle r_X\rangle - \Big(\alpha\langle Y\rangle^2\langle Z\rangle\langle r_X\rangle + \beta\langle X\rangle\langle Y\rangle\langle Z\rangle\langle r_X\rangle + \beta\langle X\rangle\langle Y\rangle\langle Z\rangle\langle r_X\rangle + \gamma\langle Y\rangle\langle Z\rangle^2\langle r_X\rangle\Big)\Big] \\
-\Big[\langle Z\rangle\langle Yr_X\rangle - \Big(\alpha\langle Y\rangle\langle Z\rangle\langle Yr_X\rangle + \beta\langle X\rangle\langle Z\rangle\langle Yr_X\rangle\Big)\Big] \\
+\Big[\alpha\langle YZ\rangle\langle Y\rangle\langle r_X\rangle + \beta\langle XY\rangle\langle Z\rangle\langle r_X\rangle + \beta\langle XZ\rangle\langle Y\rangle\langle r_X\rangle + \gamma\langle YZ\rangle\langle Z\rangle\langle r_X\rangle\Big] \\
-\Big[\langle Y\rangle\langle Zr_X\rangle - \Big(\beta\langle X\rangle\langle Y\rangle\langle Zr_X\rangle + \gamma\langle Y\rangle\langle Z\rangle\langle Zr_X\rangle\Big)\Big] - \Big[c_{YZ}\langle r_X\rangle + \frac{d\beta}{dt}c_{XY}c_{XZ}\Big] \\
+\ldots
\end{aligned}$$

Two terms contained in the square brackets above are eliminated by choosing the time varying pre-factors,

$$\alpha(t) = \frac{1}{2\langle Y\rangle(t)},\ \beta(t) = \frac{1}{2\langle X\rangle(t)},\ \gamma(t) = \frac{1}{2\langle Z\rangle(t)},\ t \geq 0. \tag{3.64}$$

The time derivatives of these pre-factors are,

$$\frac{d\alpha(t)}{dt} = -\frac{\langle r_Y\rangle}{2\langle Y\rangle^2(t)},\ \frac{d\beta(t)}{dt} = -\frac{\langle r_X\rangle}{2\langle X\rangle^2(t)},\ \frac{d\gamma(t)}{dt} = -\frac{\langle r_Z\rangle}{2\langle Z\rangle^2(t)},\ t \geq 0. \tag{3.65}$$

Putting this approximation choice into the above gives,

$$\begin{aligned}
\frac{d}{dt}(c_{XYZ} - c'_{XYZ}) &= \Big[\langle YZr_X\rangle - \Big(\frac{\langle YZ\rangle}{2\langle Y\rangle}\langle Yr_X\rangle + \frac{\langle XY\rangle}{2\langle X\rangle}\langle Zr_X\rangle + \frac{\langle XZ\rangle}{2\langle X\rangle}\langle Yr_X\rangle \\
&\quad + \frac{\langle YZ\rangle}{2\langle Z\rangle}\langle Zr_X\rangle\Big)\Big] + \Big[\Big(\langle YZ\rangle + \frac{\langle XY\rangle}{2\langle X\rangle}\langle Z\rangle \\
&\quad + \frac{\langle XZ\rangle}{2\langle X\rangle}\langle Y\rangle\Big) - \langle Y\rangle\langle Z\rangle\Big]\langle r_X\rangle - \Big(c_{YZ} - \frac{c_{XY}c_{XZ}}{2\langle X\rangle^2}\Big)\langle r_X\rangle + \ldots.
\end{aligned} \tag{3.66}$$

The extra terms not written above can be derived from considering all terms involving r_Y and r_Z respectively along with the two terms involving $\frac{d}{dt}\alpha$ and $\frac{d}{dt}\gamma$, they have the same structure up to permutation of the disease states. As all covariance functions are at least $\mathcal{O}(L^{-d})$ the random distributions de-correlate in the limit $L \to \infty$ and the right hand side of equation 3.66 becomes zero as expected.

Moreover this allows manipulations such as,

$$\langle YZr_X\rangle = \frac{\langle YZ\rangle}{\langle Y\rangle}\langle Yr_X\rangle + \mathcal{O}(L^{-d}).$$

Putting the above into equation (3.66) gives,

$$\begin{aligned}\frac{d}{dt}(c_{XYZ} - c'_{XYZ}) &= \frac{1}{2}\langle YZr_X\rangle - \frac{1}{2}\langle Y\rangle\langle Zr_X\rangle - \frac{1}{2}\langle Z\rangle\langle Yr_X\rangle \\ &\quad + \frac{1}{2}\langle Y\rangle\langle Z\rangle\langle r_X\rangle - \frac{1}{2}c_{YZ}\langle r_X\rangle + \xi + \ldots. \end{aligned} \tag{3.67}$$

Where ξ is an $\mathcal{O}(L^{-d})$ quantity. Equation (3.67) together with equation (3.63) gives that for sufficiently small inverse transmission scale $1/L$ we have, including the permuted terms,

$$\begin{aligned}\Big|\frac{d}{dt}\Big(c_{XYZ} - c'_{XYZ}\Big)\Big| &= \frac{1}{2}\Big|\langle\tilde{Y}\tilde{Z}r_X\rangle - c_{YZ}\langle r_X\rangle + \langle\tilde{X}\tilde{Z}r_Y\rangle - c_{XZ}\langle r_Y\rangle \\ &\quad + \langle\tilde{X}\tilde{Y}r_Z\rangle - c_{XY}\langle r_Z\rangle\Big| + \xi' \\ &\leq \Big|\frac{d}{dt}c_{XYZ}\Big|. \end{aligned} \tag{3.68}$$

For some $\xi' > 0$.

Although hardly the last word in rigour the arguments above suggest that the power-2 approximation (3.59) with the pre-factors given by (3.64) has an error growing less quickly than the power-1 approximation. I have motivated my choice from the order of the approximation in the inverse scale $1/L$ and from the intrinsic dynamics of the third central moment, which can be calculated using the theory I have developed so far. Interestingly, the covariance closure I use,

$$c_{XYZ} = \frac{1}{2}\Big[\frac{c_{XZ}c_{YZ}}{\mathbb{E}[Z]} + \frac{c_{XY}c_{XZ}}{\mathbb{E}[X]} + \frac{c_{YZ}c_{XY}}{\mathbb{E}[Y]}\Big], \tag{3.69}$$

can be seen to be equivalent, after some algebra to the moment closure,

$$\langle XYZ\rangle \approx \frac{1}{2}\Big[\frac{\langle XZ\rangle\langle YZ\rangle}{\langle Z\rangle} + \frac{\langle XY\rangle\langle XZ\rangle}{\langle X\rangle} + \frac{\langle YZ\rangle\langle XY\rangle}{\langle Y\rangle} - \langle X\rangle\langle Y\rangle\langle Z\rangle\Big]. \tag{3.70}$$

This form of moment closure was first suggested in the context of spatial population dynamics by Dieckmann and Law (Dieckmann and Law 2000); the exact form that I use is equivalent to the symmetric case where each disease state is treated

equivalently with respect to the approximation. Moment closure approximations are notoriously difficult to analyse, although some effort has been made in the population dynamics literature to give criteria based on the properties of third moments and symmetry-type conditions on the relabelling of individuals, cf. (Murrell, Dieckmann, and Law 2004). I contend that analysing the dynamics of the third central moments (i.e. the covariance functions) using equation (3.58) is a more direct approach to understanding their role in a closure theory. But there is much more work to be done in this direction.

3.3.2 The Third Order Approximation

Equation (3.58) gives the dynamics of the triple covariance functions, c_{XYZ}. As has been mentioned these will generically depend on the four point spatial covariances, c_{XYZW}. The simplest covariance closure from the point of view of an error analysis is to just neglect these fourth order terms, which give a contribution of $\mathcal{O}(L^{-3d})$ to the spatial moment dynamics according to the error analysis above.

Translation invariance means that I have freedom to write the spatial covariances as functions of two separation vectors rather than three co-ordinates, see figure 3.2. I choose the representing separation vectors to be $(x-y)$ and $(x-z)$ for the standard ordering $c_{XYZ}(x,y,z,t)$ referring to the three-point spatial covariance between habitats at locations x,y,z each with an abiding population in the respective disease states X,Y,Z. Rotational invariance allows a further dimensionality reduction such that for $d=2$

$$c_{XYZ}(x,y,z,t) = c_{XYZ}(|x-y|,|x-z|,\theta,t). \tag{3.71}$$

Where θ is the interior angle between vectors $x-y$ and $x-z$ in the triangle formed by x,y,z as vertices (figure 3.2). The moment hierarchy can be constructed as analogous to the triple dynamics in 2.3.6 using essentially the same techniques as above, the spatial triple dynamics are given in appendix A.

3.4 Numerical Solutions to the Covariance Hierarchy

I finally return to the questions posed in the introduction to this chapter armed with three approximation schemes for the moment dynamics; the power-1 closure, the power-2 closure and a third order closure based on closing at the order of quadruples. In this section I numerically solve the approximating equations and compare

the different approximation schemes' ability to reproduce the spatial moment dynamics observed from Monte Carlo simulation. Having established the reliability of the closure schemes for different length scales of transmission I then turn to assessing the prediction the moment closure schemes give for the expected severity of the spatial epidemic. The base line severity compared to will be the mean field predictions (2.63) and (2.64) and the factors considered will be varying length scale of infection L, varying transmission kernel shape and the underlying clustering of the spatial metapopulation.

3.4.1 Resolving integrals in the Frequency Domain

The approximating deterministic moment and covariance equations were represented in the continuous frequency domain $\mathbb{R}^d$ (the frequency domain form for the third order closure are given in appendix B). For this section I restrict to infectious spread on the plane (d=2) and solve the moment equations using a discrete grid of frequencies, $\boldsymbol{G}$. For the power-1 closure, using translation and rotation invariance to reduce dimensionality as discussed in previous sections, each pair covariance functions for each disease state pair $(X, Y) \in \{S, I\}^2$ obeyed,

$$\hat{c}_{XY}(\omega_x, \omega_y; t) = \hat{c}_{XY}(|\omega|; t) \tag{3.72}$$

Where ω_x and ω_y are the spatial frequencies for the x and y directions, and ω is the spatial frequency for the $(x - y)$ separation vector. This meant that the solution grid for the dynamics of the pair covariances could be restricted to,

$$\boldsymbol{G}_{pairs} = \{n\Delta\omega \mid n = 0, \ldots, n_{max},\ \Delta\omega > 0\}. \tag{3.73}$$

The moment dynamics at each time were only effected by the pair covariance functions through the SI pair spatial covariance measure, given in $d = 2$ as,

$$2\pi \int_0^\infty |\omega| \hat{K}(|\omega|) \hat{c}_{SI}(|\omega|; t) d|\omega|. \tag{3.74}$$

For numerical solution I used the simple quadrature rule at each time step,

$$2\pi \int_0^\infty |\omega| \hat{K}(|\omega|) \hat{c}_{SI}(|\omega|; t) d|\omega| \approx 2\pi \sum_{i \in \boldsymbol{G}} \hat{K}(i) \hat{c}_{SI}(i, t)(i\Delta\omega)\Delta\omega. \tag{3.75}$$

Integration forwards in time was done using a Runge-Kutta scheme.

The important factors for the grid choice are to pick $\Delta\omega$ sufficiently small such that (3.75) is a good approximation to the integral up to $|\omega|_{max} = n_{max}\Delta\omega$, and to choose n_{max} sufficiently large that integration up to $|\omega|_{max}$ effectively covers the support of $\hat{K}$. The power-1 closure was especially numerically efficient so I was able to use a fine grid width $\Delta\omega = 10^{-3}$ with a large size $n_{max} = 10^4$. The power-2 closure was also solved using the grid $\boldsymbol{G}_{pair}$ although more integrals require numerical solution (see appendix B); each can be solved using the quadrature rule above. This meant that the power-2 closure was less numerically rapid than the power-1 closure.

For the third order closure scheme the triple covariance dynamics have a more complex Fourier representation (see appendix B). For d=2 dimensionality reduction due to translation and rotation invariance is that for each disease state triple,

$$\hat{c}_{XYZ}(\omega_x, \omega_y, \omega_z; t) = \hat{c}_{XYZ}(|\omega|, |\omega'|, \theta; t), \tag{3.76}$$

where θ is an angle and consequently the solution grid was defined as,

$$\boldsymbol{G}_{triples}((n\Delta\omega, m\Delta\omega, p\Delta\theta) \mid n, m = 0, \ldots, n_{max}, p = 0, \ldots, p_{max} - 1; \Delta p = (2\pi/p_{max})). \tag{3.77}$$

The grid for the triple dynamics has to be chosen so that $|\omega|_{max}$ covers the support of $\hat{K}$ and also the the additional frequency magnitude $|\omega+\omega'| = \omega^2+\omega'^2+2|\omega||\omega'|\cos\theta$ covers the support of $\hat{K}$. This is due to the geometry of the interactions between the spatial frequency components for the triples covariances (see appendix B). The third order closure was significantly slower to solve numerically than the other two closures.

3.4.2 Results for Gaussian Shaped Transmission Kernel

In this section I choose the transmission kernel to be Gaussian shaped,

$$K(x) = \frac{K_0}{2\pi L^2} e^{-x^2/2L^2}. \tag{3.78}$$

Which has the Fourier transform,

$$\hat{K}(|\omega|) = K_0 e^{-2\pi^2 L^2 |\omega|^2}. \tag{3.79}$$

I fix the effort of infection $K_0 = 1.5$, the recovery rate to be $\gamma = 1$ and the initial infecteds density $\langle I \rangle(0) = 0.05$ and investigate the performance of the three moment closure schemes versus spatial moments drawn from simulation averages as the length scale of transmission is varied.

Simulations were for a metapopulation of size $N = 10^4$ on a background space of $\mathcal{A} = [-50, 50]^2$. For each simulation iterate a new habitat distribution was drawn from the Poisson Cluster process defined in section 3.2.1 and initial infecteds chosen uniformly at random according to the chosen initial infecteds population density. All other populations were initialised as susceptible. The stochastic simulation method was the tau-leap method (see (D. Gillespie 2001) and next chapter) with time step $\Delta t = 0.01$. Simulation spatial moments were constructed by taking the empirical average of the population density of susceptibles and infecteds at each time. That is that the moments were an average over the habitat distribution and the translationally invariant initial conditions as required. Periodic boundary conditions were given. The choice of large population size and periodic boundary conditions was motivated so as to minimise finite size effects; recalling that the Fourier representation method for the pair and triples covariances is a $N \to \infty$ theory. Unless stated otherwise simulation spatial moments were taken over 500 iterations.

I defined my error measure as the maximum of the two supremum errors along time for respectively $\langle S \rangle$ and $\langle I \rangle$. Denoting $\langle X \rangle$ the simulation-based moment prediction and $\langle X \rangle'$ the covariance closure moment prediction this gives,

$$\text{error} = \max \Big\{ \sup_{t \leq T} |\langle S(t) \rangle - \langle S(t) \rangle'|, \sup_{t \leq T} |\langle I(t) \rangle - \langle I(t) \rangle'| \Big\}. \tag{3.80}$$

For L large the general results were as expected; all closures performed well, in congruence with the error analysis in section 3.2.5. However, as $L \to 2$ the power-1 closure became unstable, in the sense that that for shorter length scales of transmission it was possible for the power-1 closure to predict,

$$\frac{d}{dt} \langle S \rangle'_{pow-1} > 0, \tag{3.81}$$

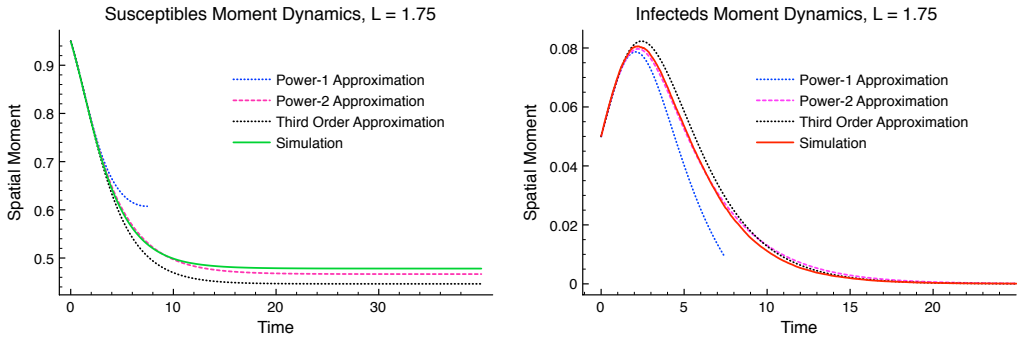


Figure 3.3: Moment dynamics for the spatial epidemic with comparison to covariance closure results. $K_0 = 1.5$, $\gamma = 1$, $L = 1.75$ and the habitat distribution was chosen as uncorrelated. At this length scale of transmission the Power-1 approximation (blue dots) was unstable and terminated at the first time it predicted unphysical dynamics. Both the power-2 closure and the third order approximation capture the essential dynamics, however the power-2 closure is better approximation to the true dynamics.

which is physically unrealistic and lead to a numerical explosion in the solving scheme. This instability could occur even when the power-2 closure and the third order closure predicted the moment dynamics well, for example when $L = 1.75$ (figure 3.3). My interpretation of this high error instability in light of lemma 2.3.2 is that the error analysis has overly focused on bounding in terms of powers of the inverse transmission scale $(1/L)$. Lemma 2.3.2 gives only a very loose bound in time; it is noticeable that the power-2 closure, which was motivated by a desire to match the dynamical behaviour of the triple covariances outperforms the third order closure, which has better static error properties in terms of $(1/L)$ despite that approximation explicitly including the triples covariance dynamics but treating the quadruples covariance as zero.

I extended the numerical error analysis to both uncorrelated and correlated habitat spatial distributions (correlated spatial distributions were drawn from the Poisson Cluster process defined in section 3.2.1 with children number $N_c = 5$ and clustering scale $\psi = 1$) varying the transmission length scale such that $(1/L) \in [0, 1]$. I found that for $L > 2$ all the closures performed well but error diverged for $L < 2$; the power-2 closure outperformed the other closure schemes at each length scale for uncorrelated habitats (figure 3.4). For the correlated spatial distribution case the power-1 closure was accurate and stable over a longer interval of transmission lengths, until $L \approx 1.75$. The performance of the third order approximation was a much better match to the power-2 closure for the correlated habitat distribution

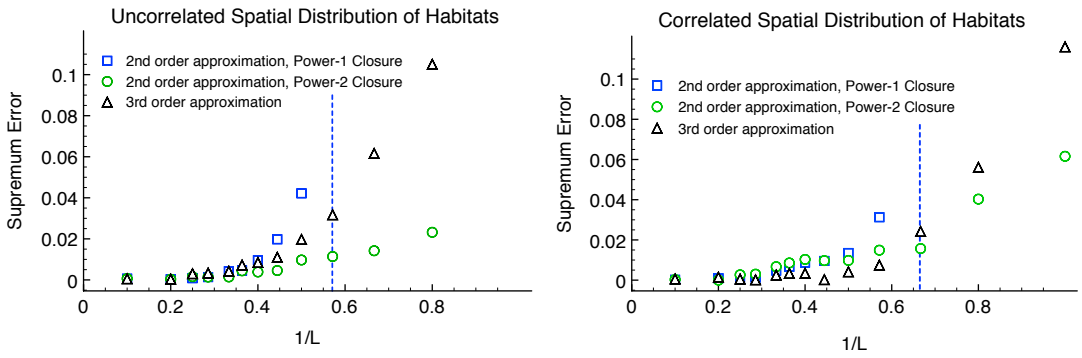


Figure 3.4: A comparison of the discrepancy between the power-1 closure (blue squares), power-2 closure (green circles) and the third order closure (black triangles) and the epidemic spatial moment dynamics were estimated from 500 Monte Carlo iterates. The epidemiological parameters were $K_0 = 1.5$, $\langle I \rangle(0) = 0.05$ and $\gamma = 1$. The transmission kernel was Gaussian shaped. The error measure used was larger supremum error over time until effective depletion of $\langle S \rangle$ and $\langle I \rangle$. *Left:* Error for uncorrelated habitat locations against inverse transmission length scale $(1/L)$. For $L > 2$ all closures perform well; for $L \approx 2$ and smaller the closures based on neglecting higher order terms (power-1 and 3rd order) have rapidly increasing error. The blue dashed lines mark the instability line for the power-1 closure. The power-2 closure out-performs the other closures at each L. *Right:* Error for correlated habitat locations against inverse transmission length scale $(1/L)$. The clustering parameters were $N_c = 5$ and clustering scale $\psi = 1$. All closures perform well for $L > 1.75$ with rapid divergence is seen for smaller length scales. The third order approximation out-performs the power-2 closure until $L \approx 1.75$.

than for the uncorrelated case. The third order closure outperformed the power-2 closure until a cross over point at $L = 1.5$ (figure 3.4).

In order to better understand the dynamical mechanisms underlying the stochastic spread amongst the habitats, I also investigated the time-varying SI pair covariance measure (3.74) which directly influences the moment dynamics. Recalling the discussion in section 3.2.4 the S-I spatial covariance can be interpreted as a measure for spatial aggregation at the length scales where transmission occurs.

I found that for uncorrelated habitat distributions the SI pair covariance was strictly negative for all L, in each case decreasing to negative peak and then tailing to zero as the infecteds are removed from the epidemic (figure 3.5). This reveals the dynamic mechanism of the initially uniformly dispersed epidemic; the infecteds proceed to recruit locally to themselves, which acts as disaggregation mechanism making susceptible and infected populations negatively spatially correlated. A consequence of this effect is that the stochastic and spatial epidemic consistently predicts a lower expected recruitment rate than a mean field model with matched

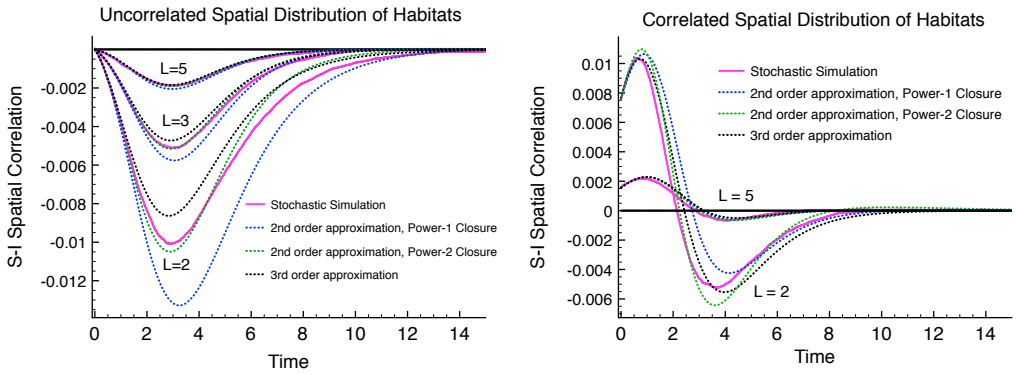


Figure 3.5: The time-varying spatial correlation measure between susceptible and infected populations, $K * c_{SI}(0,t)$. This measures the effect of space and stochasticity on the recruitment rate of new infected at each time compared to the mean field rate ($K_0\langle S\rangle\langle I\rangle(\mathrm{t})$). The epidemiological parameters were $K_0 = 1.5$, $\langle I\rangle(0) = 0.05$ and $\gamma = 1$. The transmission kernel was Gaussian shaped. *Left:* The underlying spatial distribution is uncorrelated; spatial disaggregation at the length scales of the transmission kernel is detected by the negative sign of $K*c_{SI}(0,t)$. This is induced by infected populations depleting their local environment of susceptibles via secondary infection. In order of decreasing magnitude the solid lines give simulation results for $L = 2, 3, 5$. Dashed lines indicate the approximation predictions. For $L = 2$ the power-1 closure overestimates the disaggregation effect whilst the third order approximation underestimates the disaggregation. *Right:* $K * c_{SI}(0,t)$ for correlated habitat locations. The clustering parameters were $N_c = 5$ and clustering scale $\psi = 1$. In order of decreasing magnitude the black curves give the simulation result for $L = 2, 5$. The effect of positive spatial covariance increases as the expected number of infecteds in each local cluster increases before the disaggregation mechanism dominates later in the 'average' epidemic.

K_0 and γ, with this effect more pronounced for smaller L. Comparison to simulation also revealed that for short length scales the power-1 closure significantly overestimated the effect of aggregation which is the source of the instability (3.81) (figure 3.5).

For spatially correlated habitats the dynamics of the SI pair covariance was found to be substantially different to the uncorrelated case with the susceptible and infected populations starting the epidemic in an aggregated state due to the clustering in the habitats. For approximately the first generation time the susceptible and infected populations further aggregate as the infecteds recruit preferentially within their local cluster. On a longer time scale the disaggregation mechanism rapidly dominates as the infecteds exhaust their local environment.

3.4.3 Results for Power-Law Shaped Transmission Kernel

The Gaussian was used as an archetype of transmission where the characteristic length scale strongly determines the locality of spread since very little 'mass' of the transmission kernel will be beyond $4L$ from the origin. An alternative is to consider transmission with matched length scale of transmission but a much more disperse shape function. The archetype choice for this is a power-law shaped kernel. I choose,

$$K(x) = \frac{K_0}{L^2(2\pi(1+(x/L)^2)^{3/2}}. \tag{3.82}$$

This was a convenient choice because its Fourier representation for $d = 2$ can be given analytically,

$$\hat{K}(|\omega|) = K_0 e^{-2\pi L|\omega|}. \tag{3.83}$$

The performance of the moment closure methods for the power-law transmission kernel was substantially improved for each L. In fact even for very local transmission ($L = 1$) the power-2 closure had an error of 0.0152 for uncorrelated habitats and 0.014. In light of the success of the power-2 closure even into the regime of small L I proceeded using this closure as my investigation tool (figure 3.6).

The disaggregation dynamics for the spatial epidemic driven by a power-law shaped transmission kernel were similar to the Gaussian kernel case when the underlying habitat distribution remained uncorrelated. Albeit with substantially reduced magnitude; the SI covariance dynamics for the Gaussian kernel case with

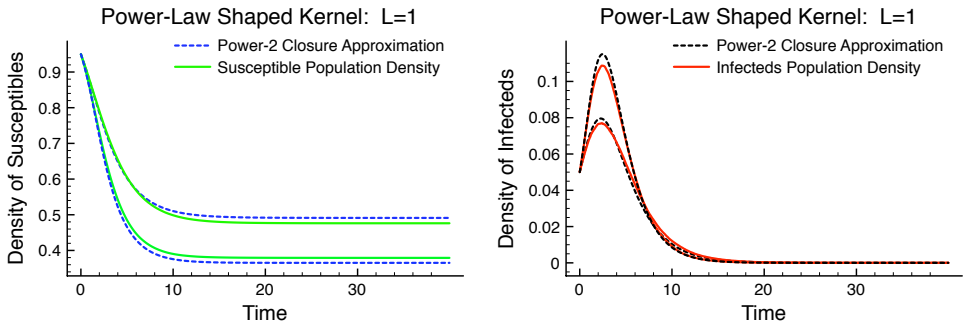


Figure 3.6: The moment dynamics for the spatial epidemic with power-law shaped transmission kernel and $L = 1$. The shape function is $k(x) = K_0/(2\pi(1+x^2)^{3/2})$. $\langle S \rangle$ is given as a green line for uncorrelated habitat locations and correlated habitat locations with the 3σ error bars for the end of the epidemic. The more severe epidemic is the one for correlated habitat locations. $\langle I \rangle$ is given as a red line, the error bars were approximately the width of the line and are not shown. Dashed lines are the power-2 closure approximation. The power-2 closure is still very accurate at this short length scale for the power-law shaped kernel.

$L = 2$ being very similar to the those of the power-law shaped kernel case with half that transmission length scale. This accounts for the improved performance of the closure scheme, since the error in the method is fundamentally connected to the magnitude of the covariance contribution to the expected dynamics (figure 3.7).

However, for the situation where the habitat spatial distribution is correlated is substantially different. For power-law shaped transmission with $L \geq 2$ in combination with the cluster habitat distribution used in this section the susceptible and infected populations remain spatially aggregated, i.e. amplifying the epidemic compared to the mean field case, throughout the epidemic until infecteds are removed. This is rather different to the Gaussian case, the power-law dispersed pathogen is amplified by local clustering of the habitats, and is sufficiently heavy tailed in its dispersal that local depletion is not a significant factor. Only at local transmission scales (L=1) does one see the recovery of the disaggregation mechanism.

The failure of the disaggregation mechanism for epidemics with power-law shaped transmission kernels had a significant impact on the expected severity of the epidemic, that is the percentage of the metapopulation predicted to become infected at some point during the epidemic. Using the mean-field prediction as a baseline, epidemics with Gaussian shaped dispersal can essentially predict at worst the severity of a mean-field epidemic for the clustering model used in this work. On the other hand, power-law shaped transmission could increase the expected

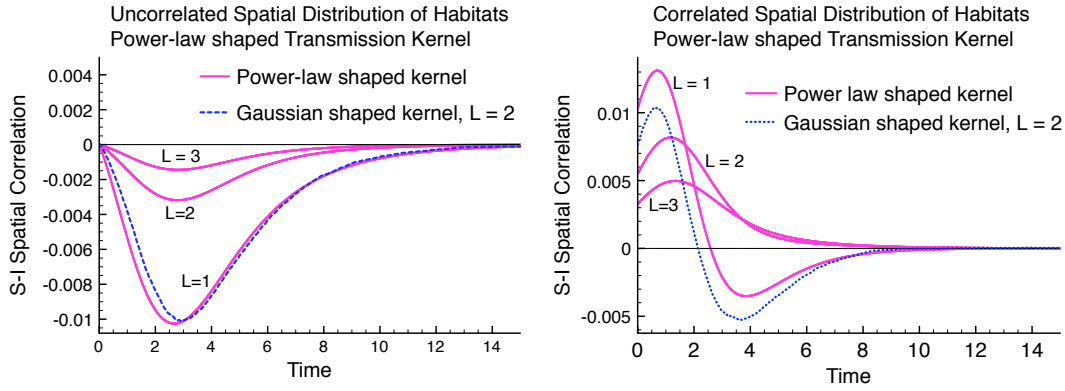


Figure 3.7: The time-varying spatial correlation measure between susceptible and infected populations, $K * c_{SI}(0,t)$, for the power-law shaped transmission kernel. For both uncorrelated and correlated habitat locations the power-2 approximation is shown with purple lines. In decreasing correlation magnitude the black curves correspond to the transmission length scales $L = 1, 2, 3$. *Left:* For uncorrelated habitat locations the correlation dynamics are similar to the scenario with Gaussian shaped transmission kernel, albeit with much decreased magnitude of $K * c_{SI}(0,t)$. The spatial correlation measure for a Gaussian shaped kernel with $L = 2$ is given as a comparison (purple dashed line) and is similar in magnitude to the spatial correlation measure for the power-law shaped kernel with $L = 1$. *Right:* For correlated habitat locations, the power-law shaped kernel does not cause disaggregation until transmission is localised to length scales $L < 2$. This is a significantly different phenomenon to that found for Gaussian shaped kernels; local depletion of susceptibles does not play a role in decreasing the rate of recruitment compared to mean-field model unless transmission is strongly localised. The Gaussian comparison with $L = 2$ is shown, and is again similar in magnitude to the spatial correlation measure for the power-law kernel with $L = 1$.

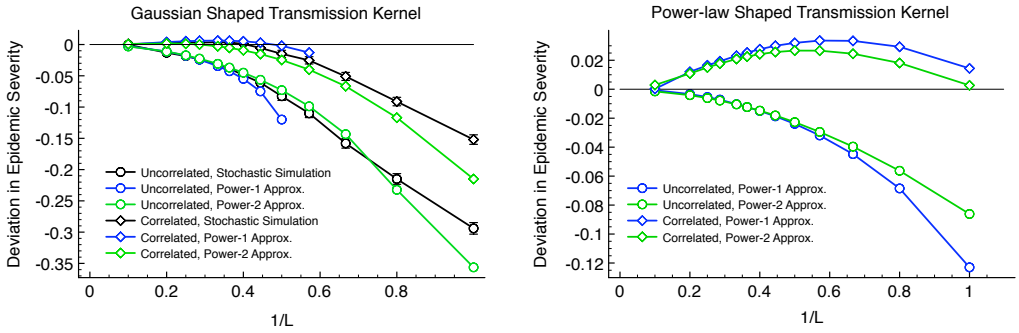


Figure 3.8: Deviation in epidemic severity measured versus the mean field prediction i.e. positive deviation predicts spatial epidemics worse than the mean field prediction and vice versa. 3σ error bars are given when larger than the data point. *Left:* Gaussian shaped kernel. For uncorrelated habitat locations (circles) both the simulation results (black) and the power-2 closure approximation (green) predict that epidemic severity is strongly decreasing with more localised transmission ($1/L$ greater). The power-1 closure agrees until it becomes unstable at $L \approx 2$. For correlated habitat locations (diamonds) there is a very mild increase in severity predicted by simulation, power-1 closure and power-2 closure until $L \approx 2$. At more local scales local depletion effects cause the epidemic to be less severe than the baseline mean field prediction. *Right:* Power-law shaped kernel. Both the power-1 and power-2 closures are stable over all the inverse length scales considered $1/L \in [0, 1]$. The power-2 closure is accurate over this scale. For uncorrelated habitat locations epidemic severity is decreasing with inverse length scale $1/L$, albeit significantly less rapidly than for the Gaussian shaped transmission kernel. For correlated habitat locations the epidemic severity is noticeably *increasing* with $1/L$ until $1/L \approx 0.6$. The combination of spatial and stochastic effects is predicted to be a net amplifier (compared to space-less models) of epidemic with strongly clustered habitats and power-law shaped transmission over a wide range of length scales.

burden of case as much as 3% of the entire population compared to the mean-field prediction, this is agreed with by the power-1 closure predictions. For both shapes of transmission kernel stochastic and spatial transmission into an uncorrelated habitat distribution is expected to lead to less severe epidemics than the mean-field baseline (figure 3.8).

3.5 Discussion

I have developed the spatial and stochastic model of epidemic spread within a Levins-type metapopulation used in this thesis to include random habitat locations. Following a rich literature on spatial moment equations from ecology and epidemiology (Bolker and Pacala 1997; Herbert and Isham 2000; Keeling 2000) I built a set of approximating ODEs for the spatial moments, meaning a joint average over the stochastic dynamics of the epidemic and the randomness of the underlying habitat's spatial distribution. As in chapter 2 the ODE approximations were motivated by being able to write down the exact spatial covariance dynamics using the tools of stochastic calculus for jump processes. In this sense the model used shares a number of features with the invasive population model of Cornell and Ovaskainen (Cornell and Ovaskainen 2008), including ability to bound the contribution of spatial covariance on the overall moment dynamics.

The analogous closure scheme to the one used in chapter 2, the power-1 closure, where the spatial covariance of triples was neglected was found to be particularly well suited to representation in the Fourier domain and was also very efficient for numeric computation. However, the power-1 closure was also found to be unstable for situations where the typical range of transmission was small. A power-1 closure for one order higher, that is neglecting quadruple covariances, was therefore developed. Although, novel in the field[1] the third order closure was disappointing in its performance, as well as being considerably more computationally burdensome than the power-1 closure it was generally less accurate than a closure based upon the quadratic approximation of the spatial triple covariances by spatial pair covariances, the power-2 closure.

The exact form of the power-2 closure was motivated by attempting to best fit the dynamical evolution of the spatial triple covariance derived from the stochastic differential representation of the epidemic, which interestingly arrived at a closure

[1] I think.

scheme already extant in the literature and confirmed to have good analytic properties (Murrell, Dieckmann, and Law 2004). Although the analysis given in this chapter is at best a first approximation, I believe this to be a fruitful direction for further study. This is especially the case since the power-2 closure outperforms the third order approximation despite having notationally worse error in terms (1/L). This suggests that the error bounding approach of Cornell and Ovaskainen (Cornell and Ovaskainen 2008) is better suited to the analysis of static equilibrium properties than dynamic quantities.

I consider my investigation into the spatial clustering and disaggregation of susceptible and infected populations as following on from Bolker (Bolker 1999). My finding that spatial disaggregation of susceptibles and infecteds retards epidemic spread for uncorrelated habitat distributions was known to him, as was the amplification effect of spatial clustering. Where I feel that I have contributed further is by appreciating the role of quadratic covariation in the underlying stochastic discrete epidemic process, which leads to corrected power-1 closed moment equations. I have also identified a further amplification effect from using a heavy-tailed transmission kernel in the presence of clustered habitats; the heavy-tailed nature of transmission means that the epidemic avoids the retarding effect of local depletion completely unless transmission is strongly localised. Since a number of real-world infectious pathogens appear, or have appeared, to exhibit power-law shaped transmission (Gibson 1997; Keeling et al. 2001), this effect is worthy of further investigation.

Finally, a direction that I would like to pursue in the future is to investigate the possibilities of using moment equations to guide control strategies. Tildesley has extensively investigated the optimal ring size vaccination around farms with confirmed cases of FMD (Tildesley et al. 2006), using a model that fits into the broad category considered in this work. However, the order to vaccinate local farms to the focal infected was unclear. A possibility could be to introduce a vaccination 'kernel' around discovered infected populations which 'transmits' vaccination, akin to the contact tracing model of Eames and Keeling (Eames and Keeling 2003). Would considering the optimal vaccination kernel, with constraints, be an optimisation problem well posed in the language of moment equations?

Chapter 4

Accelerated Simulation for Spatial Epidemics

4.1 Introduction

The results of chapters 2 and 3 gave general insights into the complex interplay between the role of stochastic transmission and the spatial fragmentation of the susceptible meta-population. However the greatest insights were towards averaged dynamics of pathogenic invasion of the meta-population, where the averaging was over both stochastic realisations of the epidemic process and the spatial statistics of the habitat landscape. Comparatively little progress was made towards covariance closure methods in the scenario where the habitat locations are given data due to the burdensome number of equations required and the possibility of numerical instability in the solution method. In light of this I return to the stochastic differential equation model of chapter 2 and attempt to integrate them directly. Stochastic integration methods we consider (Milstein 1994; Bichteler 2002) are better known as *Monte Carlo* (MC) simulation methods.

Simulation is a universally applicable approach to investigating these models, allowing flexibility in adding further realism to the general model framework. If scientific or statistical confidence in the relevant parameters can be achieved, then Monte Carlo (MC) simulation methods can have 'real-world' predictive power (Tildesley et al. 2008). However, for stochastic simulation the MC convergence of the observables of interest can require a very large number of independent replicates of simulated epidemic realisations. Moreover, this problem is compounded whenever insight requires the sweeping of large sections of plausible parameter space, or when

the results of simulation are required for parameter inference, such as the novel use of the particle filter by Iondies et al (Ionides, Bret, and King 2006). The requirement to perform very many independent realisations puts great emphasis on the development of novel and efficient algorithms for stochastic simulation, even at the cost of mild inexactness (an example of such an approach is (D. Gillespie 2001)). Even a small saving in computation time per replicate can make large scale computation significantly quicker.

The essential non-linearity in epidemic models is due to the interaction between susceptible and infectious individuals. In this context each susceptible habitat becomes infected due to the summed risk of transmission from all infectious habitats. Therefore, naively a simulation must calculate (and sum) the transmission strength between all susceptible - infectious pairs in the population; and this value must be recalculated after any event that changes the infection status of the population. It is this calculation that is the necessary but time-consuming ingredient for event based stochastic spatial simulation, and it is improving the efficiency of this calculation that is the subject of this work.

Here I present a method for simulating spatial epidemics via the rapid estimation of the force of infection rather than resolving the sum over infected habitats directly. Inspired by Galerkin methods common to the numerical investigation of PDEs (Hesthaven, Gottlieb, and Gottlieb 2007), this approach is motivated by representing the force of infection as a smooth field over space. After approximating the habitat locations as unit mass Gaussian distributions in space, the spectral projection of this field onto a uniform grid of size M can be calculated efficiently using FFT (Fast Fourier Transform) at a computational cost of $\mathcal{O}(M \ln M)$. The infection recruitment rate for each susceptible habitat can then be determined by interpolation on nearest grid points. This compares favourably with the potentially $\mathcal{O}(N^2)$ cost of calculating the infection rate for all susceptibles from the direct definition. Although this methodology is an approximate scheme, the error due to the spectral projection can be made negligible by decreasing grid separations. In the case of large population size where many infectious habitats contribute to the risk of infection for each susceptible habitat (that is the transmission kernel is relatively wide) the spectral method can be implemented with low error whilst still retaining significant speed advantage over simulations using direct summation to calculate force of infection.

The basic methodology is introduced in stages; firstly the basic spatial model framework and state space of the model are briefly recalled. I then describe existing methods for simulating the dynamics of infection, in terms of solving the stochastic differential equation (SDE) representation of the system and demonstrating how the force of infection on a habitat can be represented as a convolution on the spatial locations. This leads to a novel algorithm (termed the Fast Spectral Rate Recalculation or FSR) to simulate spatial epidemics. Later in the results section I focus on the accuracy and computational efficiency of this new algorithm.

4.2 Stochastic Simulation Techniques

In this section I briefly recall the spatial epidemic model used through out this work; the dynamics of which can be represented as a stochastic differential model driven by independent Poisson processes, see chapter 2. In the previous chapters I used this representation to motivate the construction of ODE systems in order to approximate the expected dynamics of the spatial epidemic. In this chapter I concentrate on integrating the underlying stochastic differential equations representing the stochastic progression of the epidemic directly.

4.2.1 A Spatial Metapopulation Epidemic Model

I consider N habitats, labelled $i = 1, \ldots, N$ with associated vector locations in the d-dimensional box ($\{x_i\}_{i=1}^N \in [-l/2, l/2]^d = \mathcal{A} \subset \mathbb{R}^d$). Each habitat is considered to be inhabited by a single host individual with the state alphabet of 'Susceptible' (S), 'Infected and Infectious' (I) and 'Removed' (R) with respect to a infectious disease following the classic SIR paradigm. (In principle the methodology outlined here readily extends to stochastic populations at each habitat, but the single host assumption makes the formulation more transparent.) Each habitat is treated as sufficiently small compared to the background space that it can be modelled as a point location. The underlying epidemic is treated as a Markovian stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined in the previous chapter, with the local disease state at habitat i given by the boolean-valued vector $X_i(t) = (S_i(t), I_i(t), R_i(t))$. The functions $\{S_i(t), I_i(t), R_i(t) \in \{0, 1\}\}_{i=1}^N$ are indicators for, respectively, the disease states S, I, R at the time t for the population abiding at habitat i. The rate at which an infectious (I) habitat j transmits to a susceptible (S) habitat i is governed by a spatial transmission kernel K, which I assume depends only on the (Euclidean) distance between the two habitats e.g. $K(x_i, x_j) \equiv K(|x_i - x_j|)$. (Note that this assumption restricts us to considering translational and rotational invariant kernels,

such that the precise locations of the habitats are irrelevant and transmission only depends on their relative positions). Hence, the *force of infection* on habitat i for a given epidemic state $X(t)$ is

$$\lambda_i(t) = \sum_j K(x_i - x_j) I_j(t). \tag{4.1}$$

and this quantity determines the rate at which susceptible individuals become infected. All infectious habitats recover and become removed (R) from the epidemic process at rate γ. In the continuous time case, these two processes lead to the familiar probabilistic dynamics for each habitat i,

$$\begin{aligned} \mathbb{P}(\text{Inf. event at hab. } i \in [t, t+h] | X(t)) &= S_i(t)\lambda_i(t)h + o(h) \\ \mathbb{P}(\text{Rec. event at hab. } i \in [t, t+h] | X(t)) &= \gamma I_i(t)h + o(h). \end{aligned} \tag{4.2}$$

The above holding for arbitrary $h > 0$. The initial disease states are drawn from some probability measure $\mathbb{P}_0$.

In chapter 2 I proved that given an initial disease state

$$X(0) = (S_1(0), I_1(0), \ldots, S_N(0), I_N(0)) \tag{4.3}$$

then for all $t \geq 0$ the disease state indicators that describe the complete state of the epidemic are stochastic processes with the stochastic integral solutions,

$$S_i(t) = S_i(0) - \sum_{j(\neq i)} \int_0^t S_i(s^-) I_j(s^-) dN_{ij}^I(s), \tag{4.4}$$

$$I_i(t) = I_i(0) + \sum_{j(\neq i)} \int_0^t S_i(s^-) I_j(s^-) dN_{ij}^I(s) - \int_0^t I_i(s^-) dN_i^R(s), \tag{4.5}$$

$$i = 1, \ldots, N, \ t \geq 0.$$

Where the $N(\cdot)$ are N^2 independent Poisson processes, $\{N_{ij}^I \sim PP(K(x_i - x_j))\}_{i \neq j}$ and $\{N_i^R \sim PP(\gamma)\}_{i=1}^N$.

The stochastic integral representation suggests a number of methods for generating a sample epidemic process $X = (X(t), t \geq 0)$. Since the generation method for a sample epidemic will in each case be best described as recursively applying a rule I call each generation method a *stochastic simulation algorithm* (SSA). The most natural SSA is to subordinate the epidemic process to the driving Poisson

processes, which was used to originally derive the stochastic integral representation (4.4) and (4.5). This method is described in chapter 2 section 2.3.2. I do not use Poisson process subordination in this chapter since it is numerically inefficient; many more random variables (the complete set hitting times of the N^2 Poisson processes up until some large cut-off time) are generated than for the other SSAs to be considered. Instead, I present two exact SSAs based on the probabilistic qualities of the hitting times of the driving Poisson processes and an approximate SSA based on an Euler-type numerical approximation of the integrals (4.4) and (4.5).

Equations (4.4) and (4.5) describe a piece-wise constant process In the continuous time. SSAs generate the next epidemic event (either a successful infectious recruitment or a recovery) and then step time forward by the random time to next event. A popular example of such an SSA is the Gillespie algorithm (Gillespie 1977), which exploits that for a given complete epidemic state $X(s)$ at some time s the random time until the next event ΔT is exponentially distributed. To see this consider the set $\{\tau_{ij}^I, \tau_i^R\}$ denoting the next contact times for the driving Poisson processes of the epidemic; the waiting time between each contact being exponentially distributed. Using the memoryless property of the exponential distribution and that the minimum of a set of exponential random variables is also exponential with rate parameter given by the sum of over the set of rate parameters gives,

$$\mathbb{P}(\Delta T < \delta t) = 1 - e^{-(\sum_i S_i(s)\lambda_i(s) + \sum_i \gamma I_i(s))\delta t}.$$

The total rate parameter $\sum_i S_i(s)\lambda_i(s) + \sum_i \gamma I_i(s) = \xi(s)$ is called the *exit rate* of state $X(s)$. The rate at which any given event, either an infectious recruitment or a recovery event, occurs is likewise given by the rate parameter of the exponential distribution of the minimum of the set of all next hitting times that given epidemic state $X(s)$ cause that event. For any event occurring at rate $r(s)$ at some time s the probability that it will be the next event to occur is $r(s)/\xi(s)$. These probabilistic considerations suggest the following SSA suggested by Gillespie (Gillespie 1977),

Gillespie Direct Algorithm:

1. Generate the initial epidemic state $X(0)$ from $\mathbb{P}_0$ and set time $t = 0$.
2. For each possible event $m = 1, \ldots, M$ of the current state $X(t)$ calculate the rate $r_m(t)$ and exit rate $\xi(t) = \sum_m r_m(t)$.
3. Randomly select the next event from the set of possible events. The mth event being chosen with probability $r_m(t)/\xi(t)$.

4. Update the current state by implementing the selected event.

5. Draw time increment $\Delta T \sim \exp(\xi(t))$ and set the new time $t \to t + \Delta T$.

6. If infected populations remain return to 2, else stop.

I have taken the view point that the sample elements $\omega \in \Omega$ are the sample paths of the N^2 Poisson drivers of the epidemic. The Gillespie algorithm is probabilistically exact; that is for each sample element $\omega \in \Omega$ a sample stochastic epidemic $X(\omega)$ is calculated exactly *and* the sample element ω is drawn from Ω in a manner consistent with the measure $\mathbb{P}$.

The Gillespie algorithm is the 'work-horse' for stochastic simulations, however for situations where habitat disease state is non-repeating (eg SIR-type dynamics rather than SIS) the *Sellke construction* is an attractive, probabilistically equivalent, choice. The Sellke construction was first introduced as an analytic tool to investigate the distribution of final sizes of a stochastic epidemic (Sellke 1983); its use as a simulation technique has been under-recognised although there has been some recent work using very similar constructions (Dhersin and Decreusefond 2011). The idea behind the Sellke construction is that no disease state $X(t)$ ever repeats since the disease progression is unidirectional $S \to I \to R$. Therefore for each initially susceptible population i a local infection event occurs at most once and using the independence of the Poisson drivers one can write down the conditional distribution function for the time of infection T_i^I, where $T_i^I = \infty$ is a valid outcome,

$$\mathbb{P}(T_i^I < t) = F(t) = 1 - \prod_{j(\neq i)} e^{-K(x_i - x_j)\int_0^t I_j(s)ds} = 1 - e^{-\int_0^t \lambda_i(s)ds}. \tag{4.6}$$

Here the distribution function is conditioned on the infection dynamics of the other $N-1$ populations. It is well known that any random variable transformed by its distribution function is a uniform random variable on $[0,1]$; that is $F(T_i^I) \sim \mathcal{U}[0,1]$. Therefore, for each initially susceptible habitat T_i^I is equal in distribution to the random variable $(T_i^I)'$ generated from the uniform random variable $U \sim \mathcal{U}[0,1]$,

$$\begin{aligned} U = 1 - e^{-\int_0^{(T_i^I)'} \lambda_i(s)ds} \implies \int_0^{(T_i^I)'} \lambda_i(s)ds = -\ln(1-U) \\ \implies \int_0^{(T_i^I)'} \lambda_i(s)ds = Z_i. \end{aligned} \tag{4.7}$$

Where, $Z_i \sim \exp(1)$; that $-\ln(1-U) \sim \exp(1)$ is also well known. The usefulness of equation (4.7) is that the random variables $\{Z_i\}_{i:S_i(0)=1}$ do not depend on the

epidemic dynamics and therefore only require generation once at the beginning of the stochastic simulation. The dynamic dependence of the left hand side of equation (4.7) can be solved sequentially in time by determining the order in which populations become infected. Sellke interpreted the random variable Z_i physically as the epidemic resistance for the ith individual. Whilst habitat i is susceptible this resistance is eroded at the time varying rate $\lambda_i(t)$, such that the time of infection, T_i^I, is the solution to

$$Z_i - \int_0^{T_i^I} \lambda_i(s)ds = 0. \tag{4.8}$$

It is clear that the duration of infectiousness for the ith population D_i is exponential distributed $D_i \sim \exp(\gamma)$ independently to each other population and can also be pre-generated. In fact the duration of infectiousness can be pre-generated from any required distribution; the exponential distribution corresponds to the Poisson recovery drivers. The above justifies the following SSA,

Sellke Construction Algorithm:

1. Generate the initial epidemic state $X(0)$ from $\mathbb{P}_0$ and set time $t = 0$.
2. For each initially susceptible population generate the epidemic resistance $Z_i \sim \exp(1)$, for each population i generate the infectiousness duration $D_i \sim \exp(\gamma)$. For each initially infectious population i store the recovery time $T_i^R = D_i$.
3. For each susceptible population i calculate the force of infection $\lambda_i(t)$.
4. For each susceptible population i use that the force of infection $\lambda_i(t)$ is constant until the next event to solve for the *possible* time until infection δt_i^I from
$$\delta t_i^I = \frac{Z_i}{\lambda_i(t)}.$$
5. Construct an ensemble of times until next event, $\{\{\delta t_i^I\}_{i:S_i(t)=1}, \{T_i^R - t\}_{i:I_i(t)=1}\}$.
6. Let $\delta t^* = \min\{\{\delta t_i^I\}_{i:S_i(t)=1}, \{T_i^R - t\}_{i:I_i(t)=1}\}$ and let E be the event corresponding to δt^*.
7. Deplete epidemic resistances, $Z_i \to Z_i - \lambda_i(t)\delta t^*$ for each remaining susceptible population. If E is an infection event for the ith population then generate the recovery time $T_i^R = t + \delta t^* + D_i$.
8. Update $t \to t + \delta t^*$ and state $X(t) \to X'(t)$ by implementing E.

9. If infected populations remain return to 3, else stop.

We see that once all random numbers are chosen (I follow Cook et al (Cook et al. 2008) in referring to these as the *latent variables* of the process) the resultant epidemic, $X = (X(t), t \geq 0)$, can be constructed from the deterministic solution of the infection process (4.8) and the recovery times. The solution to (4.8) can be calculated exactly since $\lambda_i(t)$ is piecewise constant in time, jumping only at infection and recovery times. For both the Gillespie and Sellke SSAs the size of time-step, in this case the random time to next event, of the algorithm will be $\mathcal{O}(1/N)$ or smaller. I note that at each of these time-steps it is generally necessary to recalculate the force of infection (λ_i) for each susceptible in the population (or at the very least consider how each force of infection has been modified by the event that has occurred).

Both of the probabilistically exact SSAs described can become computationally burdensome as the number of populations is allowed to grow; the expected duration of the epidemic is expected to increase if the number of initial infecteds does not grow proportionally with total population numbers and the number of rates to calculate also increases. This motivates using approximate SSAs to simulate epidemic outbreak. The *Tau-leap* method (Gillespie 2001; Gillespie and Petzold 2003) is an SSA where the time-step is fixed to be $\tau = \mathcal{O}(1)$. The underlying motivation is the Euler-type approximation of the stochastic integrals (4.4) and (4.5),

$$S_i(t+\tau) = S_i(t) - S_i(t) \sum_{j(\neq i)} I_j(t)(N_{ij}^I(t+\tau) - N_{ij}^I(t)), \tag{4.9}$$

$$\begin{aligned} I_i(t+\tau) &= I_i(t) + S_i(t) \sum_{j(\neq i)} I_j(t)(N_{ij}^I(t+\tau) - N_{ij}^I(t)) \\ &\quad - I_i(t)(N_i^R(t+\tau) - N_i^R(t)). \end{aligned} \tag{4.10}$$

Obviously care must be taken not to violate fundamental constraints of the process, that only one indicator function can be active per population and that they return boolean values. Also, if one is not interested in the exact source of each infectious transmission $j \to i$ but rather just the infection event for population i then it is useful to note that,

$$\sum_{j(\neq i)} I_j(t)(N_{ij}^I(t+\tau) - N_{ij}^I(t)) \sim \text{Poisson}(\lambda_i(t)\tau). \tag{4.11}$$

This leads to the SSA,

Constrained Tau-leap Algorithm:

1. Generate the initial epidemic state $X(0)$ from $\mathbb{P}_0$ and set time $t = 0$ and choose a time-step parameter τ.

2. Calculate the force of infection $\lambda_i(t)$ for each susceptible population i.

3. Draw the stochastic changes, $\Delta I_i \sim \text{Poisson}(S_i(t)\lambda_i(t)\tau)$ and $\Delta R_i \sim \text{Poisson}(\gamma I_i(t))$.

4. Set $S_i(t) \to S_i(t) - \Delta I_i$, $I_i(t) \to I_i(t) + \Delta I_i - \Delta R_i$.

5. Enforce population constraints, i.e. if $S_i(t) < 0$, then set $S_i(t) = 0$.

6. Increase time $t \to t + \tau$.

7. If infected populations remain return to 2, else stop.

The Tau-leap method is inexact in the sense that although the random sample $\omega \in \Omega$ is drawn consistently with the measure $\mathbb{P}$, the Tau-leap SSA predicts an epidemic process $X'(\omega)$ which differs from the true process $X(\omega)$. However, in a general analysis of the path-wise computation of stochastic integrals Bichteler (see (Bichteler 2002) Chapter 5 section 4) demonstrated the uniform convergence of the Euler approximate $X'(\omega) \to X(\omega)$ as $\tau \to 0$. The convergence conditions are exactly that of the existence of the stochastic integrals (4.4) and (4.5) which I guaranteed in chaper 2.

In this chapter I will also be concerned with related discrete time Markov processes. This can either be viewed as an approximation to the continue time process or an independent model in its own right where the processes occur on a natural cycle (eg daily). For many natural systems there is a clear daily cycle, and transmission rates may vary substantially between day and night time, in such cases a discrete time model may be viewed as more realistic than a continuous (time homogeneous) model. In this discrete time model infection and recovery events do not have instantaneous effect on the stochastic rates of all other events but impact on the subsequent time step. Having chosen the cycle period $\delta t^* > 0$ as a model choice, I use the Keeling-type (Keeling et al. 2001) model, which has been successfully deployed for the simulation of epidemics between farms,

$$\mathbb{P}(\text{Inf. at hab. } i \in [t_n, t_{n+1}] | X(t_n), S_i(t_n) = 1) = 1 - \exp(-\lambda_i(t_n)\,\delta t^*) \quad (4.12)$$

$$\mathbb{P}(\text{Rec. at hab. } i \in [t_n, t_{n+1}] | X(t_n), I_i(t_n) = 1) = 1 - \exp(-\gamma\,\delta t^*). \quad (4.13)$$

Where $t_n = n(\delta t^*), n \in \mathbb{N}$. Discrete time epidemics can be simulated directly from (4.12) and (4.13), or via a modified Sellke construction where all infection times

that satisfy condition (4.8) during a temporal cycle lead to an infection event at the end of that cycle, similarly for recovery events. In contrast to continuous time SSAs this allows more than one event per time-step and hence may reduce the number of times the force of infection needs to be calculated.

The Keeling-type model (4.12) and (4.13) can be seen to be equivalent to the constrained Tau-leap algorithm where the time-step τ is equal to the natural cycle δt^* since for identical states the two models predict infection and recovery events with identical probability at each time step. Ergo, the constrained Tau leap SSA can also be used for the simulation of the discrete time Keeling model. I include the constrained Tau leap algorithm because it will be useful for simulating related models to my basic metapopulation models where I incorporate habitat population dynamics, that is the local disease state for a habitat will refer to numbers of infectious and susceptible hosts abiding at the habitat.

4.2.2 The Force of Infection as a Spatial Convolution

When the number of habitats under consideration is very large, the necessary recalculation of the force of infection, $\lambda_i(t)$, for each susceptible habitat after each time-step becomes computationally burdensome. A brute force approach to this task requires a sum over all infected habitats for each susceptible habitat at a computational cost of $\mathcal{O}(N^2)$ per time-step. An immediate saving is made by including $\{\lambda_i(t)\}_{i\in S}$ as part of the epidemic state, updating for each susceptible habitat upon each event. Hence, for E events occurring in a time-step, the force of infection recalculation is of cost $\mathcal{O}(EN)$. I call this approach *SSA with rate updating*. For probabilistically exact continuous time SSAs rate updating grants an order of magnitude saving in computational cost compared to the brute force approach, however for the Tau-leap method and the discrete time models the recalculation still scales as $\mathcal{O}(N^2)$, due to the fact that $E \sim \mathcal{O}(N)$, although there is a significant saving compared to full recalculation.

I propose a method of accelerating stochastic simulation by reducing the computational burden involved in recalculation of the force of infection for each susceptible habitat after E events have occurred in a time-step of our simulation method. The basic idea is to treat $\lambda_i(t)$ as a single point of a field, which is in the form of a convolution between the transmission kernel and a spatial distribution of infected habitats. I have already used the convolution representation of the force of infection in chapter 3; it was analytically convenient for Fourier transformation

into the continuous frequency domain. In this chapter I am concerned with efficient simulation of stochastic models of spatial epidemic transmission and a notional limit $l \to \infty$ is not taken. For this chapter the analytical advantage of considering the force of infection to be a convolution is to access the numerical efficiency of convolution solving using the discrete transform known as the *fast Fourier transform* (FFT). The computational effort of convolution solving using FFT on M collocation points grows $\mathcal{O}(M \log M)$, which is potentially very efficient compared to brute force calculation of the force of infection. I recall the formalism of Ovaskainen and Cornell (Ovaskainen and Cornell 2006a; Cornell and Ovaskainen 2008) used in chapter 3 and represent the spatial spread of the infection across the habitats in distribution form. I call this the spatially extended *image* of the infectious sources,

$$I(x,t) = \sum_{i=1}^{N} \delta(x - x_i) I_i(t), \qquad x \in \mathcal{A}. \tag{4.14}$$

In a similar manner, the force of infection is extended to a field *for every point in* $\mathbb{R}^d$. This field has the compact expression

$$\begin{aligned} \lambda(x,t) &= \sum_{j=1}^{N} K(x - x_j) I_j(t) \\ &= \sum_{j=1}^{N} \int K(x-y)\delta(y - x_j) I_j(t) dy \\ &= (K * I)(x,t) \end{aligned} \tag{4.15}$$

Here, $*$ represents convolution over the spatial variables, $(f * g)(x,t) = \int f(x - y, t) g(y,t) dy$. For a given epidemic state, $X(t)$, the force of infection at each habitat i is recovered through the identity $\lambda_i(t) = \lambda(x_i, t)$.

The conceptual power behind our approach is to use this convolution representation to rapidly calculate the force of infection field λ on M collocation points and use local interpolation to estimate λ_i (for all susceptibles) at each time-step of the stochastic integration scheme.

4.2.3 Spectral versus Pseudo-spectral projection

In this subsection I review elements of the standard theory of spectral approximations to periodic functions. I make explicit the difference between 'true' spectral methods based on functional representation in terms of a weighted sum over a finite

Parameter/Function	Definition
$\{x_i\}_{i=1}^N$	Habitat locations
K	Spatial Transmission kernel
λ_i	Force of infection, Habitat i
γ	Recovery rate
I	Image of infection
ϵ	Width of Gaussian Approximation
Δx	Separations in collocation grid
I^ϵ	Blurry Image of infection
λ	Force of infection field
$\Phi_{\Delta x}^{l,d}$	Approximation grid
$\lambda_{PS}^{M,\epsilon}$	FFT estimated Force of infection field

Table 4.1: Model Parameters and Functions

set of basis wave functions and the pseudo-spectral method that makes use of efficient summation over a defined collocation grid.

I consider all functions $f \in \mathcal{C}^1$ on $\mathcal{A} = [-l/2, l/2]^d$ to be l-periodically extended to $\mathbb{R}^d$, i.e. $f(x+l\boldsymbol{n}) = f(x)$, $\forall x \in \mathcal{A}$, $\forall \boldsymbol{n} \in \mathbb{Z}^d$. In the context of the epidemic this is equivalent to imposing periodic boundary conditions on the dynamics. The requirement for periodic boundary conditions can be effectively relaxed using zero padding at the boundaries hence I do not consider this a significant restriction.

It is well known that l-periodically extended functions $f \in \mathcal{C}^1$ have a Fourier series representation, in particular

$$\lambda(x,t) \sim \frac{1}{l^d} \sum_{\omega \in \mathbb{Z}^d} \hat{\lambda}(\omega,t) e^{2\pi i \omega \cdot x/l}. \tag{4.16}$$

Where for each $\omega \in \mathbb{Z}^d$, the *Fourier coefficients*: $\hat{\lambda}(\omega,t)$[1] are defined as,

$$\hat{\lambda}(\omega,t) = \int_{\mathcal{A}} \lambda(x,t) e^{-2\pi i x \cdot \omega/l} dx. \tag{4.17}$$

The relationship $\sim$ indicates uniform convergence on $\mathbb{R}^d$ as the partial sum of the right hand side of (4.16) includes vectors $\omega \in \mathbb{Z}^d$ of increasing magnitude, subsequently I shall use $=$ without confusion since I only consider force of infection fields with continuous first derivatives. Equation (4.16) indicates that the set of plane

[1] I use ω to denote spatial frequency rather than the more common k to avoid confusion with the transmission kernel in our notation. I rarely use the ω to mean a sample element in this chapter.

waves $\{e^{2\pi\omega\cdot x/l}\}_{k\in\mathbb{Z}^d}$ forms a basis for the vector space of l- periodic $\mathcal{C}^1$ functions. Note that the definition above differs from the one used in the main work. Definition (4.17) is the true Fourier coefficient used in constructing spectral approximations to periodic functions, the Fourier coefficient calculated via FFT is a quadrature approximation to the integral in (4.17).

The formal connection back to the continuous Fourier transform used in chapter 3 can be justified by treating the Fourier coefficients as functions from the set $\omega' \in \{\omega/l \in \mathbb{R}|\omega \in \mathbb{Z}^d\}$. From this view point equation (4.17) coincides with the continuous Fourier transform when $l \to \infty$. On the other hand, equation (4.16) defines a quadrature approximation of the integral form of the continuous inverse Fourier transform,

$$\lambda(x,t) = \frac{1}{l^d}\sum_{\omega'} \hat{\lambda}(\omega',t)e^{2\pi i\omega'\cdot x} \approx \int \hat{\lambda}(\omega',t)e^{2\pi i\omega'\cdot x}d\omega'. \tag{4.18}$$

This approximation becomes correct as $l \to \infty$.

For each plane wave, ω is the d-dimensional vector of spatial frequency, or *wave-vector*, for the wave in each standard Euclidean dimension. Since this basis set is infinite in size I truncate and only use $M < \infty$ plane waves, with associated Fourier coefficients, to estimate the force of infection field, λ,

$$\lambda(x,t) \approx \lambda_S^M(x,t) = \frac{1}{l^d}\sum_{\omega\in\Omega_M^d} \hat{\lambda}(\omega)e^{2\pi i\omega\cdot x/l}. \tag{4.19}$$

The sum is over the set

$$\Omega_M^d = \{\omega = (\omega_1,\ldots,\omega_d) \in \mathbb{Z}^d \big| -m/2 \le \omega_i \le m/2-1 \ i = 1,\ldots,d\}, \tag{4.20}$$

with $m^d = M$. It is convenient to restrict to cases where m is even. I call this estimate, λ_S^M, the *spectral projection* of the force of infection field. The projection is onto the subspace of the square integrable functions on $\mathcal{A}$, $L^2(\mathcal{A})$, spanned by $\{e^{2\pi i\omega\cdot x/l}\}_{\omega\in\Omega_M^d}$.

As we have seen the force of infection field can be represented as a convolution between the transmission kernel (K) and the image of infecteds (I),

$$\lambda(x,t) = K * I(x,t) = \int_{\mathcal{A}} K(x-y)I(y,t)dy. \tag{4.21}$$

The representation as a convolution is very convenient for our analysis, since the Fourier coefficients $\hat{\lambda}$ are simply the the product of the Fourier coefficients of the transmission kernel $\hat{K}$ and the image of infecteds $\hat{f}_I$. This can be seen by applying definition (4.17) and the Fourier decomposition of K and I,

$$\begin{aligned}
\hat{\lambda}(\omega,t) &= \int_{\mathcal{A}} K * I(x,t) e^{-2\pi i x\cdot\omega/l} dx \\
&= \int_{\mathcal{A}}\int_{\mathcal{A}} K(x-y) I(y,t) e^{-2\pi i x\cdot\omega/l}\, dydx \\
&= \frac{1}{l^{2d}} \int_{\mathcal{A}}\int_{\mathcal{A}} \sum_{\omega'\in\mathbb{Z}^d} \sum_{\omega''\in\mathbb{Z}^d} \hat{K}(\omega')\hat{f}_I(\omega'',t) e^{2\pi i y\cdot(\omega''-\omega')/l} e^{2\pi i x\cdot(\omega'-\omega)/l}\, dydx \\
&= \hat{K}(\omega)\hat{f}_I(\omega,t). \qquad (4.22)
\end{aligned}$$

The final line is due to using the orthogonality of the plane waves,

$$\int_{\mathcal{A}} e^{2\pi i(\omega-\omega')\cdot q/l} dx = l^d \delta_{\omega,\omega'}. \qquad (4.23)$$

Equation (4.22) will be the basis of the fast spectral rate recalculation method; it splits the time invariant and time varying parts of the force of infection field in the spatial frequency domain.

The fastest method for calculating the Fourier coefficients required for the projection approximation is by using the fast (discrete) Fourier transform (FFT) algorithm, which scales in computational complexity as $\mathcal{O}(M \log M)$ for M Fourier coefficients. The discrete transform approximates the integral (4.17) using sum of samples of λ at regular intervals in $\mathcal{A}$. To this end I introduce a regular grid on $\mathcal{A}$ of separation Δx, $\Phi^{l,d}_{\Delta x}$. By varying Δx I can restrict the grid to size M, but I always choose Δx such that

$$l = m\Delta x.$$

Due to this I fix $M = m^d$ and hence,

$$\Phi^{l,d}_{\Delta x} = \{\boldsymbol{i} = (i_1,\ldots,i_d)\Delta x \mid (i_1,\ldots,i_d) \in \mathbb{Z}^d,\ 0 \le i_j < m\ j = 1,\ldots,d\}. \qquad (4.24)$$

The grid-points are the *collocation points* for the discrete transformation. I recall that the quadrature rules for FFT and inverse fast Fourier transform (IFFT) are

respectively,

$$F\big[\lambda\big](\omega,t) = \tilde{\lambda}(\omega,t) = (\Delta x)^d \sum_{\boldsymbol{i}\in\Phi^{l,d}_{\Delta x}} \lambda(\boldsymbol{i},t) e^{-2\pi i(\boldsymbol{i}/\Delta x)\cdot\omega/m}, \qquad \omega \in \Omega^d_M \tag{4.25}$$

$$F^{-1}\big[\tilde{\lambda}\big](\boldsymbol{i},t) = \frac{1}{(m\Delta x)^d} \sum_{\omega\in\Omega^d_M} \tilde{\lambda}(k,t) e^{2\pi i(\boldsymbol{i}/\Delta x)\cdot\omega/m}, \qquad \boldsymbol{i} \in \Phi^{l,d}_{\Delta x}. \tag{4.26}$$

The definition above is not completely standard, it is more common in the literature not to include the factor $(\Delta x)^d$. However, when subsequently I consider products of FFT coefficients the factor $(\Delta x)^d$ becomes important.

Using Fourier coefficients estimated from a discrete quadrature rule I define the *pseudo-spectral* (discrete) projection of the force of infection field,

$$\begin{aligned} \lambda^M_{PS}(\boldsymbol{i},t) &= \frac{1}{l^d} \sum_{\omega\in\Omega^d_M} \tilde{\lambda}(\omega,t) e^{2\pi i\omega\cdot\boldsymbol{i}/l} \\ &= \frac{1}{l^d} \sum_{\omega\in\Omega^d_M} \tilde{K}(\omega)\tilde{I}(\omega,t) e^{2\pi i\omega\cdot\boldsymbol{i}/l} \end{aligned} \tag{4.27}$$

This projection onto the grid $\Phi^{l,d}_{\Delta x}$ is the approximation to the force of infection field (λ) used in this work.

4.3 Fast Spectral Rate Recalculation

My goal is to reduce the problem of simulating stochastic epidemic dynamics in continuous space to a closely approximating problem on a discrete grid. The grid-based dynamics can be simulated much more efficiently than the full system dynamics using FFT techniques, since the force of infection on each habitat can be rapidly recalculated after infection or recovery events. The use of FFT requires periodic boundary conditions for $\mathcal{A}$, for this section I use this assumption. In the following section I mention how this assumption can be relaxed.

Two ideas underpin accelerating the calculation of the force of infection (Figure 1). Firstly, the image of the infectious sources is 'blurred' by approximating each delta distribution as a Gaussian shaped function of width $\epsilon > 0$; this is so that the infection potential can be approximated as originating from grid points. Secondly, I solve an analogous equation to (4.15) defined on a discrete collocation grid using efficient FFT convolution solving and read off $\{\lambda_i(t)\}_{i=1}^N$ directly from a local inter-

polation scheme.

Since our goal is to utilise the efficiency of FFT algorithms I consider the transmission kernel and image of infection supported on the collection grid (4.24), noting that the image of infection evolves in time as infection and recovery events occur,

$$K(\boldsymbol{i}),\ I(\boldsymbol{i},t), \qquad \boldsymbol{i} \in \Phi_{\Delta x}^{l,d},\ t \in \mathbb{R}_+.$$

Having constructed the grid support for these functions their Fourier coefficients can be estimated using FFT. An FFT algorithm calculates all the required Fourier coefficients with $\mathcal{O}(M \ln M)$, via the efficient decomposition of the sum (4.25),

$$\tilde{K}(\omega),\ \tilde{I}(\omega,t), \qquad \omega \in \Omega_M^d,\ t \geq 0.$$

The Fourier coefficients for the transmission kernel do not vary in time (although this is a possible model extension) and hence require calculation solely once. The image of infection evolves as the simulated epidemic progresses.

A natural approximation to the integral form of the force of infection field (4.15) on the periodic domain $\mathcal{A}$ is to use a sum over grid points, whilst respecting the periodic boundary conditions. My method of fast calculation of the force of infection exploits that this natural approximation is equivalent to λ_{PS}^M given by (4.27),

$$\begin{aligned} \lambda_{PS}^M(\boldsymbol{i},t) &= (\Delta x)^d \sum_{\boldsymbol{j} \in \Phi_{\Delta x}^{l,d}} K(\{\boldsymbol{i}-\boldsymbol{j}\}_l) I(\boldsymbol{j},t) \\ &= F^{-1}[\tilde{K}\tilde{I}](\boldsymbol{i},t), \qquad \boldsymbol{i} \in \Phi_{\Delta x}^{l,d},\ t \geq 0 \end{aligned} \tag{4.28}$$

Where $F^{-1}[\cdot]$ is the inverse fast Fourier transform (IFFT). $\{\boldsymbol{i}-\boldsymbol{j}\}_l$ represents that on the periodic domain $\mathcal{A}$, in principle infection can be transmitted at a number of ranges - from the shortest route between the infectious habitat and the susceptible to any addition of 'wrapping round' routes. In practise I am concerned with transmission that is short range compared to the size of the background space, in which case $\{\boldsymbol{i}-\boldsymbol{j}\}_l$ is the shortest vector between points $\boldsymbol{i}$ and $\boldsymbol{j}$ on the periodic space $\mathcal{A}$.

One potential difficulty is that although in theory a well-defined Fourier series exists for I, the numerical performance of the discrete transform FFT is very poor

in this situation. This is due to the delta distributions in the definition of I not being functions in the traditional sense, but rather distributions in a more general sense (Friedlander 1998). I resolve with this problem by introducing *blurred images*, where the delta distributions in the image definition (4.14) are replaced by tight Gaussians of width ϵ (> 0),

$$\delta_\epsilon(x) = \frac{1}{(2\pi\epsilon^2)^{d/2}} e^{-\frac{-|x|^2}{2\epsilon^2}}. \tag{4.29}$$

To provide even greater computational efficiency, I treat these Gaussians as having finite range, cutting off at a distance of 5ϵ. The number of collocation points local to habitat i (that is are involved with the blurring of the image of i) is denoted $n_i = |\{\boldsymbol{i} \in \Phi^{l,d}_{\Delta x} \; s.t. \; |\boldsymbol{i} - x_i| < 5\epsilon\}|$. Note that $n_{max} = \max_{i \in N}\{n_i\} \ll M$ if $\epsilon = \mathcal{O}(1)$.

The Gaussian approximated version of the epidemic image is denoted, I^ϵ. Putting this together I have the following approximation for the local force of infection,

$$\lambda(\boldsymbol{i}, t) \approx \lambda^{M,\epsilon}_{PS}(\boldsymbol{i}, t) = F^{-1}[\tilde{K}\tilde{I}^\epsilon](\boldsymbol{i}, t), \qquad \boldsymbol{i} \in \Phi^{l,d}_{\Delta x}, t \geq 0. \tag{4.30}$$

Where $\lambda^{M,\epsilon}_{PS}$ is the pseudo-spectral approximation to the force of infection field with Gaussian blurring; essentially combining (4.28) and (4.29). Once, I have estimated the force of infection on the grid of points, $\Phi^{l,d}_{\Delta x}$, I can read off an approximation for $\lambda_i(t)$ from the local grid points using an interpolation rule. My examples will be in $d = 2$ and I shall use bilinear interpolation from the 4 grid points nearest x_i which is an $\mathcal{O}(\Delta x^2)$ interpolation scheme.

I call simulation using recalculation of $\lambda^{M,\epsilon}_{PS}$ and local interpolation the *fast spectral rate recalculation* (FSR) method, which can be represented by the following algorithm.

Fast Spectral Rate Recalculation (FSR) algorithm:

1. Choose a suitable SSA for epidemic simulation, a FFT algorithm and a local interpolation rule.

2. Before simulation calculate and store local points for each habitat i, $\{\boldsymbol{i} \in \Phi^{l,d}_{\Delta x} \; s.t. \; |\boldsymbol{i} - x_i| < 5\epsilon\}$.

3. Draw initial epidemic state from $\mathbb{P}_0$ and construct the initial projection of

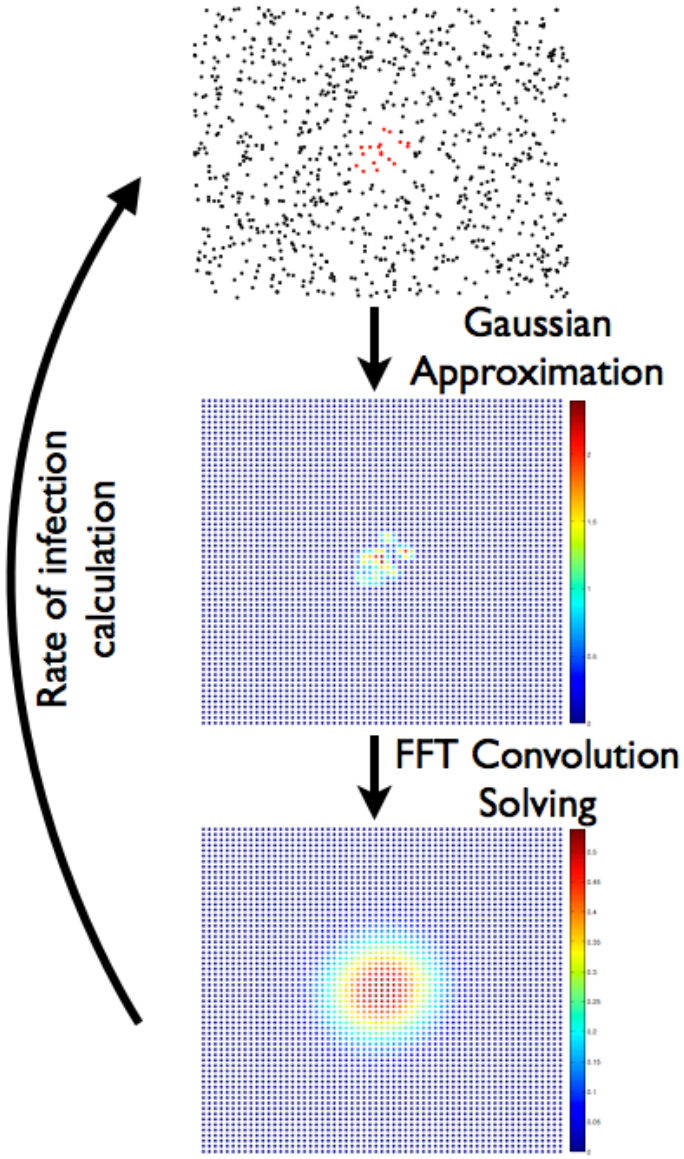


Figure 4.1: A schematic diagram of the FSR method. *Top:* The image of the infection. Infected habitat locations in $\mathcal{A}$ are given as red dots in Euclidean space, other disease state habitats are given in black. *Middle:* The 'blurry' image of the infection. A regular grid of separation Δx is laid over $\mathcal{A}$ with an additional zero-padding layer to impose non-periodic boundary conditions. Infected habitats are given support on the grid according to an ϵ width spatial Gaussian approximation. Support is truncated at 5ϵ from the Euclidean location of the habitat, hence infection or recovery events cause update only at the local points to the event on the grid. *Bottom:* The pseudo-spectral projection of the force of infection. Using the FFT convolution solution at computational cost $\mathcal{O}(M \ln M)$, the force of infection is projected onto the grid. Stochastic rates of infection are then read off from bilinear interpolation on the 4 nearest grid points to the susceptible habitat.

force of infection field, $\lambda_{PS}^{M,\epsilon}(\boldsymbol{i},0)\ \forall \boldsymbol{i} \in \Phi_{\Delta x}^{l,d}$.

4. Solve the sum $\lambda_{PS}^{M,\epsilon}(\boldsymbol{i},t)$, $\forall \boldsymbol{i} \in \Phi_{\Delta x}^{l,d}$, by using an FFT algorithm to construct the product function $\tilde{K}\tilde{I}$ and taking IFFT, $F^{-1}[\tilde{K}\tilde{I}]$.

5. Re-scale $\lambda_{PS}^{M,\epsilon}$ by the factor $(\Delta x)^2$, if this is not performed by the FFT algorithm.

6. Perform SSA time-step update using rates of infection $\lambda_i(t)$ read off from local interpolation of $\lambda_{PS}^{M,\epsilon}(\boldsymbol{i},t)$.

7. For each of E events that occurred in step 6., update I^ϵ on $\Phi_{\Delta x}^{l,d}$ at the n_i local points to the habitat where the event has occurred.

8. If there are remaining infected habitats return to 4. If epidemic has finished, stop simulation.

9. If further simulations are required return to 3. The information calculated in 2. can be reused.

Figure 4.1 gives a schematic diagram of the FSR algorithm. Conveniently, there are a multiplicity of high quality FFT algorithms publicly accessible, for example the FFTw library for C programming. MATLAB also uses a high quality FFT routine. Hence, there is no necessity for the reader to construct their own transformation algorithm as required by the FSR method for rate recalculation.

Although I have chosen to illustrate the use of the FSR method with the motivating example of the spatial SIR-type epidemic, it should be noted that this methodology holds for any type of stochastic spatial process where the stochastic rate for events can be written in the form (4.15). Examples therefore include a range of spatial epidemiological models (eg SEIR-type and SIS-type dynamics) and spatial ecological models where competition or colonisation is governed by a local interaction process. In particular, the necessary ingredients for using FSR are,

- (Euclidean) Spatially distributed individuals or densities,
- Smooth and translation invariant spatial interactions.

Choosing the Spectral Parameters

In this section I briefly discuss the interplay between choosing grid separation Δx and Gaussian width, ϵ. For smooth transmission kernels, K, I have for any state

$X(t)$ a rigorous upper bound on the uniform error (the worst error anywhere in $\mathcal{A}$) between the true (λ) and estimated ($\lambda_{PS}^{M,\epsilon}$) force of infection field,

$$\|\lambda_{PS}^{M,\epsilon} - \lambda\|_\infty \leq \Psi\|\lambda\|_\infty + \frac{\epsilon^2}{2}\|\Delta\lambda\|_\infty + Ce^{-\alpha M} + \mathcal{O}(\epsilon^4). \tag{4.31}$$

Where $C > 0$ is a constant, $\alpha > 0$ is dependent on the smoothness of the transmission kernel K, Δ is the Laplacian operator, and Ψ is an error factor due to using the discrete grid $\Phi_{\Delta x}^{l,d}$ to approximate the Fourier coefficients. $\Psi(r)$ depends on the ratio of Gaussian width to grid width, $r = \Delta x/\epsilon$, with

$$\Psi(r) \to 0, \qquad \text{as } r \to 0. \tag{4.32}$$

The proof for the error analysis equation (4.31) involves an investigation into the joint error due to using a finite number of waves (spectral projection error), estimating the Fourier coefficients using a discrete transform (aliasing error) and the error due to using the Gaussian approximation for the delta distributions. The main feature of interest in the error analysis will be that for fixed grid separations Δx the aliasing error and the Gaussian approximation error 'trade off' against each other. The proof can be found in full in appendix C.

Equations (4.31) and (4.32) allow key insights into the FSR method:

- Smooth transmission kernels, K, lead to the small uniform norms and large values of α in (4.31), and hence greater expected accuracy for the FSR method.
- For any $\epsilon > 0$, $\lim_{\Delta x \to 0} \|\lambda_{PS}^{M,\epsilon} - \lambda\|_\infty = \frac{\epsilon^2}{2}\|\Delta\lambda\|_\infty$; therefore, the uniform error (4.31) can be set arbitrarily small by reducing both Δx and ϵ.
- As I increase the width of the space I consider (l) the uniform norms in (4.31) depend on how the habitats are distributed since they represent the maximum rate of infection anywhere in $\mathcal{A}$ for the disease state $X(t)$. For randomly distributed habitats this maximum rate will grow very slowly with l.
- For a given collocation grid with $\Delta x > 0$, I can use that $\|\lambda\|_\infty \leq I(t)\|K\|_\infty$ where $I(t) = \sum_j I_j(t)$ to write a condition for the error (4.31) that can be calculated *before* any simulations are run,

$$\|\lambda_{PS}^{M,\epsilon} - \lambda\|_\infty \leq I(t)\big[\Psi\|K\|_\infty + \frac{\epsilon^2}{2}\|\Delta K\|_\infty\big] + Ce^{-\alpha M} + \mathcal{O}(\epsilon^4). \tag{4.33}$$

  Typically, for our applications $\|K\|_\infty = K(0)$ and $\|\Delta K\|_\infty = |\Delta K(0)|$. Hence,

a good choice for ϵ will be one which makes the following quantity small

$$\Psi K(0) + \frac{\epsilon^2}{2}|\Delta K(0)|. \tag{4.34}$$

From numerical investigation choosing $\epsilon = 0.8\Delta x$ has been a broadly successful choice, although I prefer in the numerical experimentation section to pick ϵ as the minimiser of (4.33). The minimiser can be found by a closer investigation of the error factor Ψ in the following sections.

A final consideration is choice of M, which in turn determines Δx. FFT algorithms operate via the factorisation of M, with greater operational speed if M has a factorisation into small primes. Hence, having tuned parameters for acceptable error it is better to pick M such that it has a factorisation into a product of powers of $2, 3, 5$. This typically constitutes a small modification of Δx.

Reducing Error via Kernel Correction

As mentioned in the previous section the error caused by using the FSR simulation method compared to a direct simulation method is jointly due to spectral projection error, aliasing error and the approximation of the spatial delta distributions by Gaussians. In fact, the supremum error (over the spatial variables) between the true force infection and its FSR approximation can be bounded above by a sum of contributions from each of these three sources of error (cf. appendix C). In this section I introduce a method of correcting the transmission kernel so as to reduce the error due to Gaussian approximation; this method will be particularly numerically effective for the special case where the transmission kernel is itself also Gaussian shaped.

The fundamental idea is based upon minimising the absolute difference between the force of infection field projected using M waves and it's Gaussian approximation; this difference being a contribution to the error upper bound. The absolute difference, at an arbitrary spatial location $x \in \mathcal{A}$ is,

$$\left|\lambda_S^{M,\epsilon}(x,t) - \lambda_S^M(x,t)\right| \;=\; \left|\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N} I_j(t)[e^{-2\pi^2\omega^2\epsilon^2/l} - 1]e^{2\pi i\omega\cdot(x-x_j)/l}\right|. \tag{4.35}$$

Equation (4.35) is suggestive; the absolute difference for all $x \in \mathcal{A}$ would be zero if the spectral approximation with Gaussian blurring $\lambda_S^{M,\epsilon}$ was calculated using a

corrected kernel K_c defined by its Fourier coefficients,

$$\hat{K}_c(\omega) = \hat{K}(\omega)e^{2\pi^2\omega^2\epsilon^2/l}, \qquad \omega \in \mathbb{Z}^d. \tag{4.36}$$

Equation (4.36) leads to a well defined transmission kernel whenever the Fourier representation (4.16) converges. For Gaussian shaped transmission kernels this will be true so long as $L > \epsilon$, and the correct kernel is simply the true transmission kernel under the transformation $L \to L - \epsilon$.

Whilst the above considerations guides ones intuition, in fact greater numerical accuracy is found by correcting the transmission using the discrete Fourier coefficients of the truncated Gaussians used in the FSR method rather than the true coefficients as in (4.36), leading to the alternative definition of the corrected kernel as,

$$\hat{K}_c(\omega) = \hat{K}(\omega)/\tilde{\delta}_\epsilon(\omega), \qquad \omega \in \Omega_M^d. \tag{4.37}$$

Using the corrected kernel defined by (4.37) will allow greater numerical accuracy in using FSR for the Gaussian shaped transmission kernel used in section 4.4, however it is not a panacea since for many kernels of possible interest (e.g. the power-law kernel considered in chapter 3) the definition (4.37) does not give a convergent corrected transmission kernel. It is possible that replacing the Gaussian approximation with a smaller scale copy of the transmission kernel of interest is a generalisation of this approach, but for heavy-tailed kernels this will come at the cost of numerical performance, and is not explored in this work. Therefore, I will present results using both a corrected and uncorrected transmission kernel.

Non-periodic Boundary Conditions

In many real-world applications physical boundaries to transmission play an important role in containing epidemic spread, or the habitats exist in a finite domain isolated from external infection. However, FSR implicitly uses periodic boundary conditions in solving the convolution sum (4.30); essentially modelling the dynamics on a torus rather than a finite plane. To circumvent this problem and yet retain the efficiency of using FFT I embed our original space, $\mathcal{A}$, within an extended space, $\mathcal{A}'$, setting the new length $l' > l$ sufficiently large such that $l'-$periodic extensions of λ do not interact. This is done by placing the habitats in the smaller space $\mathcal{A}$, and then subsequently creating a zero padded space for I^ϵ on $I_{l'}$. On this space an infection event has a zero probability of 'wrapping' round the space $I_{l'}$. I achieve

this either by restricting to transmission kernels of compact support, or by imposing compact support by finite range truncation. I therefore set the epidemic on $\mathcal{A}'$ and use the FSR method for fast calculation of $\lambda(\boldsymbol{i}, t)$ on $\boldsymbol{i} \in \Phi^{l'}_{\Delta x}$ as above, see Figure 4.2 for a schematic.

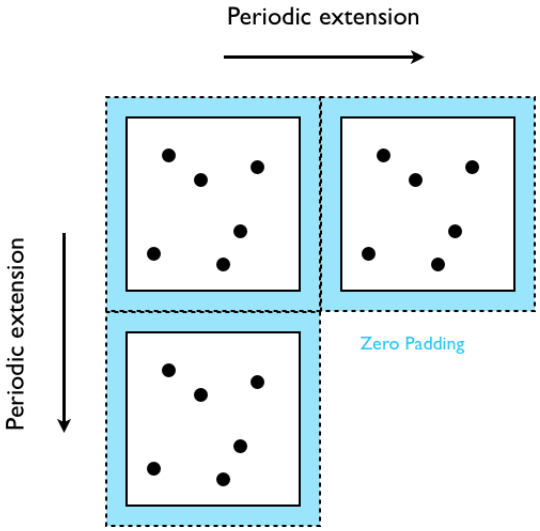


Figure 4.2: A schematic diagram for effectively imposing non-periodic boundary conditions. FFT treats both the real and frequency domain as periodic, to eliminate the possibility of unphysical 'wrap around' infection events the space $\mathcal{A}$ is augmented with a sufficiently large zero-padded area. The larger space $\mathcal{A}'$ is treated as the periodic real domain for the FFT transforms.

4.4 Numerical Experimentation using FSR method

I present results and analyse the accelerated performance due to using the FSR method in two different toy models: a simplified version of the Keeling-type model used for Foot-and-Mouth disease transmission between farms (Keeling et al. 2001; Keeling et al. 2003) with time discretised onto a daily cycle and a generalisation to the Gibson model for the spread of Citrus Tristeza virus in continuous time (Gibson 1997). I find that in all investigated parameter regions the error in epidemic simulation due to using spectral methods for recalculating the force of infection followed the supremum error (4.31) in the sense that when $\|\lambda^{M,\epsilon}_{PS} - \lambda\|_\infty$ was small so was the error in epidemic simulation. I also found that for stochastic simulation algorithms that use discretise simulation time-step then, asymptotically as the number of habitats becomes large, FSR delivers significant time saving even whilst main-

taining high accuracy. For the Keeling-type model with a Gaussian shaped kernel both results for uncorrected and corrected kernels will be demonstrated.

4.4.1 Comparing Simulation Methods by Matching Latent Variables

The baseline methodology for stochastic simulation of Markov processes is to use rates calculated directly from the current state. However, even for a fixed initial distribution of infected and susceptible habitats and fixed parameters for the epidemic dynamics I still expect potentially significant variation between multiple simulated epidemics due to stochasticity. This complicates defining an appropriate performance measure for the accuracy of the FSR method. For example a crude comparison of the epidemic curves averaged over many samples is inadequate since this measure might miss significant spatial variation between methods that produce very similar population dynamics; while any method that compares a finite number of simulations must account for the variability in epidemic outcomes.

Although in principle this comparison problem can be approached from the point of view of classical multi-variate statistics, I use an alternative which exploits the special nature of the Sellke construction. The Sellke construction, as discussed earlier, uses the fact that the inherent randomness of a epidemic realisation can be encoded at the beginning of the simulation in a random vector Z (termed the latent variables of the epidemic process) of exp(1) distributed resistances and infection durations. Essentially, once the random vector Z is calculated, the epidemic process is then deterministic. Therefore, knowledge of these random latent variables allows direct comparison between simulation output using direct rate recalculation (DR) and using FSR, by matching values of Z at the start of the epidemic. Such matching of latent processes/variables has been used in the context of evaluating retrospective outcomes of intervention (Cook et al. 2008). Here I use latent variable Z matching to compare simulation techniques. Figure 4.3 gives a schematic description of latent variable matching.

Fixing the latent variables and initial condition is similar in spirit to fixing the sample element $\omega \in \Omega$. Recall that if a stochastic process exists (existence for the spatial epidemic was established in chapter 2) then each sample element ω specifies an epidemic process $X(\omega) = (X(t;\omega), t \geq 0)$. In principle each ω used in this work encodes an initial disease state and a realisation on $[0, \infty)$ of the underlying Poisson drivers of the epidemic process. This constructs the path $X(\omega)$ via solving

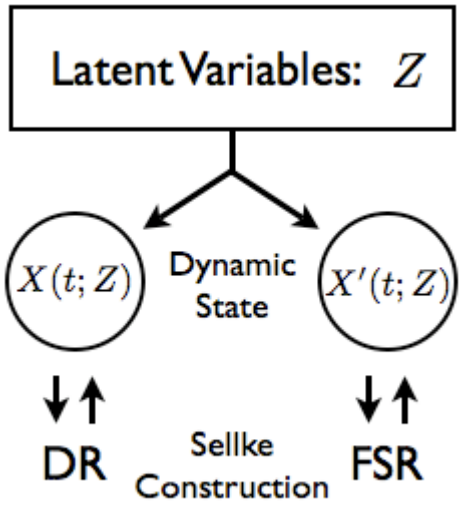


Figure 4.3: A schematic diagram for Sellke Construction simulation using direct rate recalculation and fast spectral recalculation with matched latent variables. The latent variables Z are pre-generated and stored before simulation. For each choice of rate recalculation technique, FSR or DR, Z completely specifies the path of the epidemic.

(4.4) and (4.5). The Sellke construction allows me to reduce from considering ω to considering the random vector Z: an initial disease state and just $2N$ latent 'epidemic resistances'. Each possible Z also specifies an epidemic process through the Sellke construction.

The major difference in fixing the latent variable Z compared to fixing the sample element ω only becomes apparent when using an approximation scheme for the stochastic rates. The FSR method approximates the infectious pressure at each habitat which has no meaning in terms of the possible times of infection/recovery fixed by choosing ω. This is because drawing ω depends directly on the epidemic parameters that are being approximated whereas Z can be drawn independently for any SIR type epidemic and it is only the rate at which the epidemic resistances are being depleted that is being approximated.

I use as our measure of accuracy the expectation of the following supremum norm on the difference between the epidemic simulated using direct rate (DR) calculation, $X(\cdot;Z)$, and that simulated using FSR, $X'(\cdot;Z)$, using identical latent

variables Z in both methods.

$$\text{Error} = \mathbb{E}_Z[\text{Error}(Z)] = \mathbb{E}_Z\Big[\sup_{0\leq s\leq T}\Big\{\frac{1}{N}\sum_{i=1}^{N} d(X_i(s;Z), X_i'(s;Z))\Big\}\Big]. \tag{4.38}$$

Where X_i gives the local disease state for population i and $d(x, y)$ is the discrete metric returning 0 iff $x = y$ and 1 otherwise. This error therefore measures the maximal difference between the spatial epidemic pattern predicted by the two methods over the entire epidemic process. This error measure has the useful property that if FSR returns identical estimates for $\{\lambda_i(t)\}_{i:S_i(t)=1}$ to DR then $\text{Error}(Z) = 0,\ \forall Z$. Such a relationship would not hold without matching latent variables Z, due to inherent variability between epidemic samples; instead, the error would approach zero only in a limiting sense with some appropriate statistical test being required to interpret small error values. The error measure (4.38) is reminiscent of almost sure convergence, that is,

$$\mathbb{P}(\omega :\ X'(\omega) = X(\omega)) = 1, \tag{4.39}$$

but with the latent variables Z replacing the role of the sample elements $\omega \in \Omega$.

4.4.2 Epidemic Spread amongst Farms

In this model the N habitats represent a population of farms, that are experiencing an epidemic invasion of foot-and-mouth disease that rapidly infects their livestock. The transmission kernel represents the hazard per unit time of the introduction of infection from a farm with some infected animals into a farm with naive livestock, through any of multiple possible transmission route, e.g. animal transport between farms, wind dispersal etc. The pathogen is considered to be sufficiently infectious that once introduced into a naive farm it spreads rapidly amongst the livestock there. This motivates the decision to represent the epidemic state of the farm by a single meta-population state, such that the entire farm is considered to by either susceptible, infectious or removed. An infection event in this context is the spread of the disease from one farm to another. A recovery event is the removal of farm livestock after detection as an infected premise, although it could refer to the affects of quarantine and vaccination if such control measures were used. In order to focus on the spatial characteristic of the spread no attempt is made to model a farm latency period, or introduce heterogeneity between farms, although the addition of these realistic heterogeneities is feasible within the FSR framework.

I simulate the epidemic dynamics as occurring on a daily cycle, using the discrete time Sellke construction (DTSC) described in section 4.2. Non-periodic boundary conditions as described above were imposed to replicate the fixed spatial scale of the at-risk farm population. I assume standard SIR-type dynamics with recovery rate γ, as such the model time scale is related to the day time scale through the standard result,

$$\mathbb{E}[\text{Days spent infectious}] = \frac{1}{\gamma}. \tag{4.40}$$

I rescale time in my simulations to fix $\gamma = 1$, and assume that the average time for detection and complete response to disease is 10 days, which is loosely comparable to the detection and removal period observed during the 2001 Foot-and-mouth epidemic (Ster and Ferguson 2007). Therefore, choosing daily updating dynamics corresponds to fixing our updating time step $\delta t^* = 0.1$ in time units defined by the mean infectious period. The farm location coordinates where chosen as spatial Poisson points on $\mathcal{A}$, which is embedded in two spatial dimensions ($d = 2$). The length scales are fixed by choosing units such that the farm density $N/l^2 = 1$. The length scale of transmission, L, can then considered to be in those units. For each N considered the random distribution of habitat locations was realised once, with further simulations performed on identical landscapes.

The transmission kernel is chosen from the family of Gaussian shaped functions with width L, $\{K(L;\cdot)\}_{L>0}$ such that

$$K(L;|x|) = \frac{K_0}{2\pi L^2} e^{-|x|^2/2L^2}. \tag{4.41}$$

For $l \gg L$, K_0 sets the integral of the transmission kernel over $\mathcal{A}$. As a reference, in the $L \to \infty$ limit spatial position becomes irrelevant and the epidemics collapses into the well understood non-spatial or mean-field closed N populated stochastic SIR model with $R_0 = K_0(N/l^2)\gamma = K_0/\gamma$, since I choose the spatial units to be such that $N/l^d = 1$. In light of this I choose $K_0 = 2$, which leads to a supercritical epidemic (ie one where I expect invasion to be successful) except at small values of L.

4.4.3 Performance of FSR method

As an initial trial I investigated the ability of the FSR method to recreate the expected dynamics and spatial distributions of a spatial epidemic amongst farms

seeded by a very localised infection event and perpetuated by fairly localised transmission. To this end I constructed an ensemble of $10,000$ farm locations from the hitting points of a spatial Poisson process on the box $\mathcal{A} = 100 \times 100$. Spatial units having been fixed such that the density $N/l^2 = 1$. The initial index case was drawn at random, with the nearest other 9 farms also being infected at time 0. The transmission length scale was chosen to be $L = 2$.

Spatial distributions and expected total population dynamics were calculated from 1000 simulated epidemic path iterates, each using the same epidemiological parameters, farm locations and initial cases. The resultant expected epidemic dynamics were characterised by a very slow recruitment of farm populations reaching an expected epidemic peak of 200 farms at 280 days after the initial disease introduction (i.e. an average recruitment rate of less than one farm a day until peak). The expected final burden of the epidemic was 6688 removed farms over the course of the epidemic (Figure 4.4).

Since the total population expectations neglect detailed spatial information I also considered the spatial distribution of risk at fixed time snapshots through the empirical probabilities, $p(\cdot)$ of finding each farm in a given disease state. For example,

$$p(I_i(t)) = \frac{1}{R}\sum_{r=1}^{R} I_i^{(r)}(t), \tag{4.42}$$

gives the empirical probability of the ith farm being infected at time t calculated over R simulation replicates and superscript (r) denoting the result of the rth replicate. The spatio-temporal risk of farm infection was characterised by a roughly 'wave-like' spread of risk from the initial cases, although boundary effects impede transmission as well as local variation in farm spatial density either amplifying or retarding the wave-like expected dynamics (Figure 4.4).

The fast spectral rate recalculation (FSR) method was implemented for this spatial epidemic using a grid separation width of $\Delta x = 0.5$ and a Gaussian blurring width of $\epsilon = 0.4$. Hard boundary conditions were recovered by further embedding $\mathcal{A}$ in $\mathcal{A}' = 112 \times 112$ box giving a grid, $\Phi_{0.5}^{100,2}$, of 224×224 uniformly spaced collocation points for FFT transformation. I found that the FSR methodology could recover the expected total population dynamics, with maximum difference in the predicted number infecteds being 3 farms at epidemic peak (Figure 4.4 top). The

FSR prediction could also recover detailed spatial information accurately. Choosing a time snapshot at $T = 200$ days I found that the total population mean absolute deviation $\langle|\Delta p(I(T = 200))|\rangle = 4.55 \times 10^{-3}$. Compared to the expected density of infected farms (from the direct method) this is still a small quantity: $\langle|\Delta p(I_{T=200})|\rangle / \langle I_{direct}(T = 200)\rangle = 0.2796$. The maximum deviation found was $|\Delta p(I_{T=200})|_{max} = 0.062$. For $N = 10,000$ the FSR method is already delivering an order of magnitude time saving per simulation. The computation time per simulation (averaged over serial computation) for the direct simulation method was 30.194 secs on a 3.2 GHz machine; for the FSR method the computation time per simulation was 2.668 secs on the same machine.

The investigation above is evidence that the distribution of spatial epidemics given by the FSR method can resemble the distribution given by using a non-approximate SSA; this is approximation in a distribution (or weak) sense. However, the interpretation of 'closeness' between distributions is not entirely straight forward, with measures including Kullbeck-Leibler divergence (Coolen, Kuehn, and Sollich 2005), the total variation metric (Capinski and Kopp 2007) and the 'Dobrushin-MacKay' metric (MacKay 2008). For each each of these choices the meaning of being a small but non-zero 'distance'[2] in distribution sense is problematic. Moreover, the final two mentioned above are given in variational form and require, respectively, the determination of a maximising set or function; even their calculation is problematic! By contrast the error measure (4.38) given in section 4.4.1 is a measure for the strong convergence of epidemic paths using the latent variables Z in place of the sample elements $\omega \in \Omega$. The maximum number of errors due to using FSR at any point in an epidemic path is a natural number depending on Z. The error measure (4.38) used in this work is simply the sum of these possible maximum errors, re-scaled as a density, weighted by the measure of the set of Zs that give that number of maximum errors. For the rest of this work I use this error measure.

For a systematic investigation of the error due to using FSR, measured using (4.38) I took as default parameters a population of $N = 10000$ farms in a 100×100 box, with $K_0 = 2$, $L = 3$ defining the transmission kernel, the FSR method used a grid separations at $\Delta x = 0.5$ and Gaussian width $\epsilon = 0.4$; averages were taken over 100 realisations of the latent variables Z. To initiate the epidemic, a single source farm was chosen, moved to the centre of the space and then it and the surrounding

[2]It should be noted that the Kullbeck-Leibler divergence is not a metric on distributions, although convergence of divergence to zero between two distributions does imply their mutual identity.

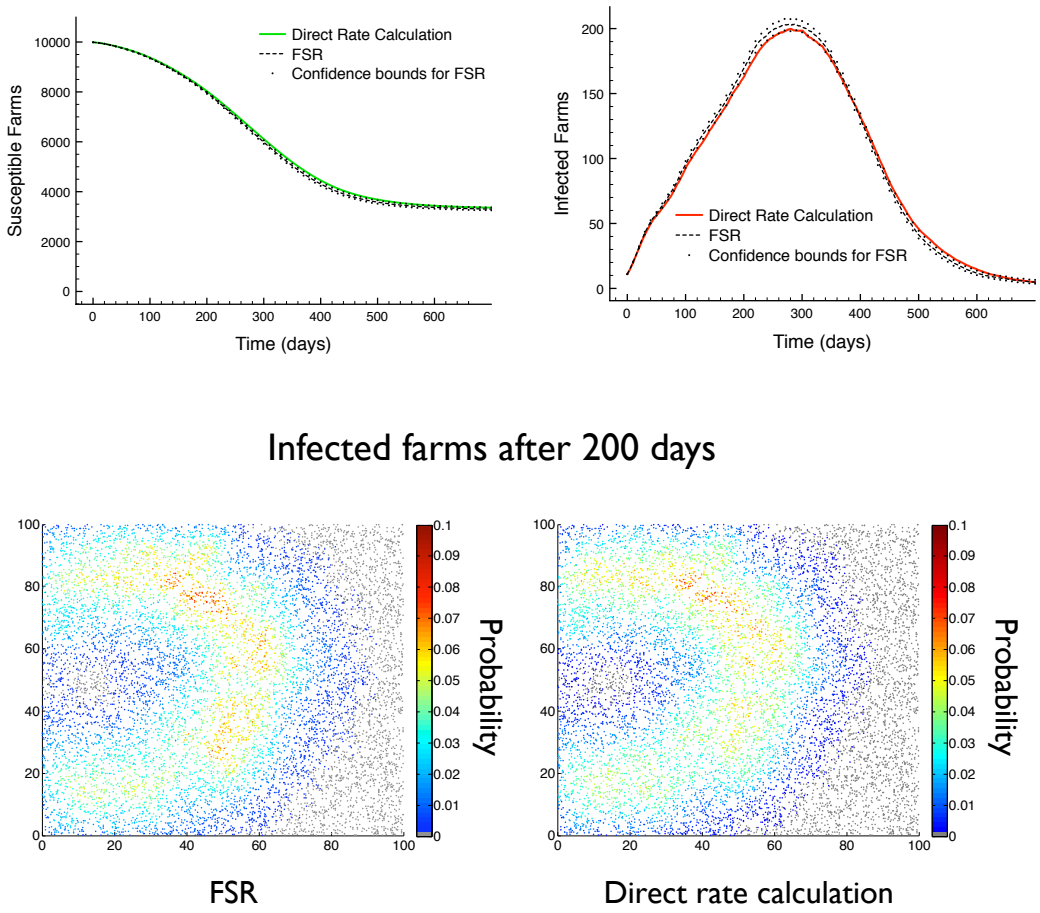


Figure 4.4: A comparison of spatial distributions of a slow travelling farm epidemic between direct simulation and fast spectral simulation. $N = 10,000$, distributed uniformly independent over $\mathcal{A} = 100 \times 100$ box. Initial seed was chosen at random, with 9 closest farms also infected at $T = 0$. Spatial transmission parameters were $L = 2$, $K_0 = 2$. The FSR parameters were $\Delta x = 0.5$ and $\epsilon = 0.4$. *Top Row:* Expected dynamics of number of susceptible (left) and infected (right) farms with the time scale given in days. Coloured lines give the dynamics calculated using direct summation of rates. Black dashed lines give the FSR results. Dots indicate 95% Bias corrected bootstrapped confidence intervals for the FSR result. In both cases good agreement is found between the FSR prediction and the prediction from direct rate calculation. The maximum error between the expected number of infected farms was 3 farms. *Bottom Row:* A snapshot at $T = 200$ days, of the spatial probability distribution of infected farms. I found the mean deviation in empirical probability of being infected to be $\langle |\Delta p(I(T = 200))| \rangle = 4.55 \times 10^{-3}$.

farms where infected such that the initial density of of infecteds $I/N = 0.02$. This density of infection means that stochastic extinction of infection in the early stages is unlikely. For all simulations using FSR non-periodic boundary conditions were imposed by the addition of $6L$ width zero-padded boundary as described above. For the default parameters, the average supremum error using standard FSR method was 2.19×10^{-2} (with confidence intervals constructed using 95% bias corrected bootstrapping $[2.12 \times 10^{-2}, 2.26 \times 10^{-2}]$). Using the corrected kernel demonstrated the increased possible accuracy when using a combination of Gaussian shaped kernels and Gaussian approximation for the delta distributions; the average supremum error for the default parameters using the corrected kernel (4.37) dropped to 6.7×10^{-3} ($[6.2 \times 10^{-3}, 7.3 \times 10^{-3}]$).

As expected from equation (4.32) the error decayed further at finer grid separation, Δx, with ϵ chosen as to minimise (4.34) (Figure 4.5A). Similarly, average error sharply decreases with increasing length scale of infection, L (Figure 4.5B). This is related to the uniform norms in equation (4.31) which asymptote to 0 as $L \to \infty$. Fixing both transmission length scale (L) and grid width (Δx), and taking a sequence of epidemic models in increasing N (while keeping a constant density of farms) I found that average error increased only weakly with N over two orders of magnitude and then appears to saturate. The same trend of error under parameter variation was found using the corrected kernel; the largest simulation I considered being size $N = 105625$ on a box of width $l = 325$ for which the error using the uncorrected kernel was 2.23×10^{-2} ($[2.18 \times 10^{-2}, 2.28 \times 10^{-2}]$), using the corrected kernel gave an average supremum error of 5.1×10^{-3} ($[4.9 \times 10^{-3}, 5.4 \times 10^{-3}]$) .

The error observed between epidemic realisations using DR and FSR (Figure 4.5) follows the same structure as the error in the pseudo-spectral projection of the force of infection field, equations (4.31) and (4.32); that is that error decayed as L increases, and/or as Δx decreased. In addition, there was weak dependence on N for the error between epidemics, which is suggested by the weak dependence on l in (4.31). It should be emphasised that the exact relationship between the uniform convergence in force of infection (4.31) and the rate of convergence of epidemic paths in our error measure (4.38) is unclear. I can demonstrate that the probability density of epidemics using the grid based approximation converges onto the true probability density of epidemics; ergo I postulate that when uniform error in the pseudo-spectral projection of the force of infection is small then so is the error in paths. This is certainly supported by the numerical error analysis.

In addition to supremum error over the epidemic path, also of interest is the forecasting error of number and spatial distribution of affected farms at the final (absorbing) state of the epidemic. Again matching Z, but only considering error at the final state of the epidemic I found the final state error again matched the supremum error in trend, but to be significantly lower than the supremum error over the epidemic path in every part of parameter space (Figure 4.5). For our default scenario, the error drops to 5.4×10^{-3} ($[4.8 \times 10^{-3}, 6.6 \times 10^{-3}]$); using a corrected kernel reduces the average final error for the default scenario to 3.8×10^{-4} ($[3.0 \times 10^{-4}, 5.8 \times 10^{-4}]$). Since (4.38) is an expected uniform norm on *time* it returns the average greatest error over the entire time of the epidemic. This suggests that (i) small errors in force of infection calculation due to FSR don't have a significant compound effect on the dynamic evolution of the epidemic and (ii) that the large majority of errors are *timing* errors; that is errors in *when* an event occurs rather than *whether* an event occurs. Broadly, as larger epidemic simulations are considered the largest error weakly increases to saturation whereas the final state error decreases - the density of occurrence of events becomes more certain.

The algorithmic efficiency of the FSR was measured by the time taken to simulate (serially) a number of replicate complete epidemic realisations; results are presented as an average simulation time per replicate epidemic (Figure 3D). This was done separately from simulating for accuracy, which was slower due to using multiple simulation techniques simultaneously. For most data points 1000 MC replicates were performed for speed estimates, however, for some of the large simulations this number was reduced, but never decreased below 20. For large N there was not great variation in simulation time between individual epidemics.

For the discrete-time Sellke construction (DTSC) simulation method with FSR, fixing $\Delta x = \mathcal{O}(1)$ and $N/l^2 = 1$ imposed the scaling $M = \mathcal{O}(N)$. Asymptotically, using $\Delta x = 0.5$ and $L = 3$ in addition to the other default parameters, simulation time, capped at 1000 time steps for fair comparison between differently sized epidemics, scaled as $\mathcal{O}(N^2)$ for DTSC with DR, as evidenced by a log-log slope fit of 1.997, and the predicted sub-$\mathcal{O}(N^2)$ scaling for DTSC with FSR was supported by a best log-log slope fit of 1.503 (Figure 4.5D). All simulations used for speed comparison were performed on the same 3.2 GHz desk-top computer.

The maximum population size for which I considered both the DR and FSR

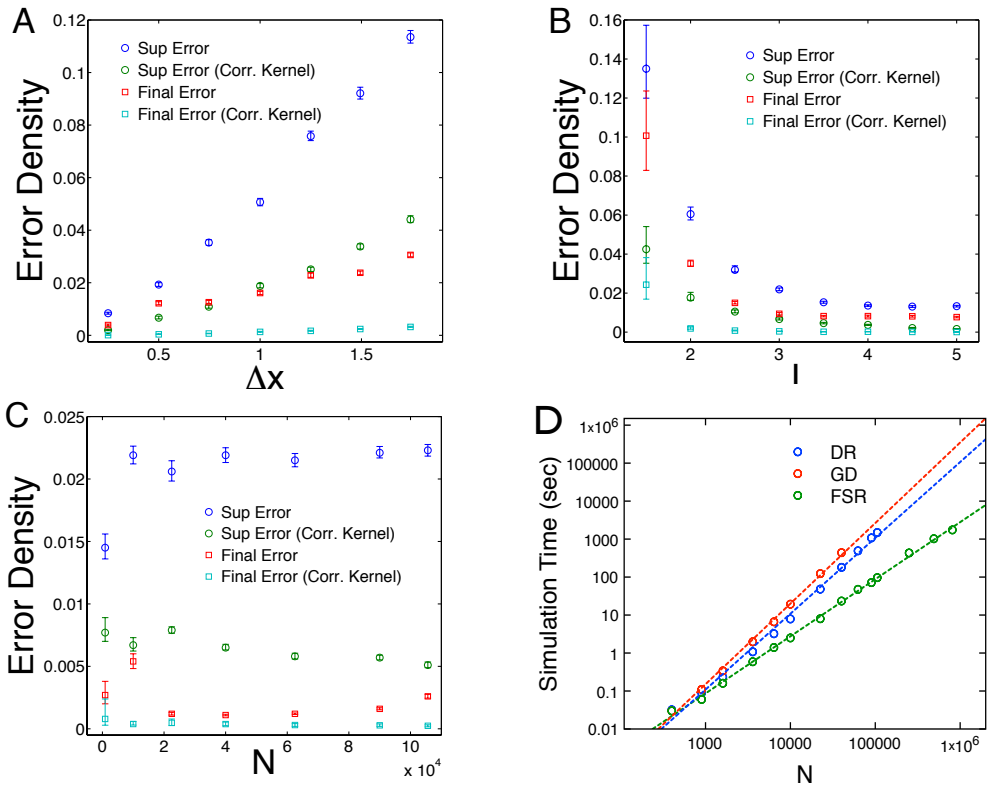


Figure 4.5: Results for model of epidemic spread amongst livestock on farms. For the discrete time Sellke construction comparison is made between simulations using direct rate recalculation and using fast spectral rate recalculation. Latent variables are matched in order to directly compare simulation methods, and comparisons taken over 100 matched MC replicates. All confidence intervals are due to bias corrected 95% confidence boot-strapping. The base set of parameter is $N = 1000$, $\Delta x = 0.5$, $L = 3$, $K_0 = 2$, $\gamma = 1$. **A-C**: Varying the base parameter set in, respectively, Δx, L, and N. As expected from theoretical considerations, the error in the force of infection field error is increasing with Δx, decreasing with L and only weakly dependent on N. Also shown are the results using a corrected transmission kernel, which follow the same trend but with lower error in each case. **D**: Time taken per MC replicate for the base set of parameters. For small value of N there is very little gain in using FSR, but as N increases the time saving becomes increasingly favourable. The computational cost for the FSR method (green circles) was well fitted by a log-log slope of 1.503 (green dashed line) which indicates sub-$\mathcal{O}(N^2)$ scaling, in contrast direct rate recalculation (blue circles) appeared to scale as $\mathcal{O}(N^2)$ (blue dashed log-log slope of 1.997). As a comparison the Gillespie Direct algorithm was also considered (red circles) and found to scale as $\mathcal{O}(N^2)$ (red dashed log-log slope of 2.121).

methods was 105625 farms, which is in the same order of magnitude as the number of farms in Great Britain [29]. At this large population size the FSR method considerably outperformed using direct rate recalculation in terms of computational efficiency, running simulations in approximately 6.6% of the time. As a comparison I also simulated sequences of epidemics using the popular Gillespie Direct algorithm (GD) with rate-updating up to a maximum size of $40,000$ farms, and extrapolated to larger values of N using a log-log slope fit of 2.121 (Figure 4.5D). Unsurprisingly, the continuous time simulation algorithm was considerably slower than either of the other two methods; using GD to, say, perform 1000 MC replicates for $\sim 10^5$ farms would require well over a month of dedicated computer time on the machine used for this work, *per parameter choice*. It is therefore clear that for large scale simulation of spatial epidemic outbreak together with wide sweeps of parameters space and large numbers of replicates necessary, the standard GD based simulation is only feasible when dedicated high performance computing resources are available. However, the use of discrete-time Selke construction methods with FSR brings this type of calculation within the realm of powerful desktop computers. In order to further explore the simulation time scaling I also ran FSR method simulations up to a maximum size of 8.1×10^5 farms, which is the same order of magnitude as the number of US cattle farms (M. J. Keeling, private correspondence), and found the scaling a good fit over the range.

I found that for more disperse transmission kernels the collocation grid can be coarsened (Δx can be increased) without substantially impairing accuracy and time saving is then even more dramatic. On the other hand, for very local transmission, which demands very fine grid mesh for good accuracy, time saving due to using FSR will only become apparent when simulating a very large populations. Although I emphasise that due to the $\mathcal{O}(N^2)$ versus $\mathcal{O}(M \ln M)$ scaling in time the FSR method will be faster for a sufficiently large simulation size. In the localized transmission situation, when effectively disease can spread to only a small number of other Farms, other simulation methods might be more appropriate, such as the subdivision of $\mathcal{A}$ into cells and exploiting the property that transmission to distant cells is highly unlikely, an example of such a method is given by Keeling and Rohani (Keeling and Rohani 2008).

4.4.4 Geographic Spread of a Plant Disease

I now focus on the situation where the N habitats are locations of subpopulations for some plant species. Each habitat is assumed to contains a random number of

individual plants. Again I assume SIR-type infection dynamics for each host plant. Removal is due to either human intervention to control infection or due to death of the plant, and occurs at rate $\gamma = 1$, although this is no longer directly related to a day-by-day cycle and occurs in continuous time. As in the previous farm-based model I consider Gaussian shaped transmission kernels, but with $K_0 = 0.1$. (I note that this low value of K_0 does not imply that R_0 is less than one, due to the presence of multiple plants at each location.) I define the transmission rate within a habitat as the value of the transmission kernel at a distance of zero. This form of model has been used in investigating the spatio-temporal spread of Citrus Canker amongst citrus trees (Gibson 1997).

The interchangeability amongst plants of identical disease state within a habitat allows us to extend our basic spatial epidemic model to one with population dynamics by reinterpreting the state at each habitat i, $X_i(t) = (S_i(t), I_i(t), R_i(t))$ as numbers of individuals in each disease state. Each habitat has a local population size N_i and the conservation of individuals is expressed as,

$$S_i(t) + I_i(t) + R_i(t) = N_i, \qquad t \geq 0.$$

Otherwise, equations (4.1) and (4.2) are unaffected. The Sellke construction is unable to gain a speed benefit from this symmetric clumping of individuals abiding at the same habitat since each individual still requires the drawing of an epidemic resistance and infectious duration. By contrast the Tau-leap method can now be used with much relaxed population size constraints which allows efficiency gains compared to the Sellke construction.

I fixed $N = 900$ and drew the local population (N_i) for each habitat independently from a binomial distribution ($\mathcal{B}(40, 0.5)$); this was chosen to reflect demographic randomness around a mean local population size of 20 plants with a hard carrying capacity of 40 plants per habitat. The total randomly generated population consisted of $17,909$ plants. At this population size the continuous time Sellke construction (CTSC) is extremely computationally costly to implement (Figure 4.6B); I therefore chose to consider the discrete-time equivalent (DTSC) and the Tau-leap (TL) approximation with a time discretisation of $\delta t = 0.05$ (taking care to respect the population constraints of non-negativity and $S_i(t) + I_i(t) \leq N_i$ where N_i is the local population size at habitat i). The convergence of Markov processes simulated via the Tau-leap approximation onto those simulated via a probabilistically

exact algorithm has been discussed in the literature (D. Gillespie 2001; D. Gillespie and Petzold 2003), while for a more detailed discussion on using Euler-type time discretisation in the context of stochastic integration see (Bichteler 2002).

Simulations were begun by completely infecting all plants within a randomly chosen habitat and in the nearest two habitats to this seed location. For the realised distribution of habitats and plant numbers within each habitat, the total number of initial infectious plants was 63. For transmission length scale $L = 3$ matching latent variables Z between the DTSC algorithm with direct rate calculation and with FSR methodology gave a similar profile of errors with varying Δx as in the farm model, albeit with the magnitude of the error reduced Figure 4.6A]. Variation in L and N also had a similar pattens of error to that found in the farm model (not shown). The lower error in this plant-pathogen model compared to the farm-model can be related to equation (4.31) by considering the reduction the uniform norms due to reducing K_0.

As an aside, an area of concentrated plant population, such as a small forest of trees, might well *not* be well modelled as a point location. The common choice of habitat as a area-less point is convenient for the calculation of the force of infection as a weighted sum over infectious individuals, but using the FSR framework it is actually more natural to consider habitats as having a spatial extent. Here I have chosen Gaussian shaped habitats as an approximation to point locations, but other choices are possible.

I trialled the five stochastic simulation algorithms (SSAs) that were considered by computing the time taken to run 1000 MC iterates of the epidemic from identical initial seed. I found that the time required in computing an epidemic realisation was mildly improved by using the tau-leap (TL) method rather than the discrete-time Sellke construction(DTSC); this is because although TL requires drawing two random variables per time-step for each habitat, the DTSC requires drawing two random variables for each plant and in this example there are multiple plants per habitat. Therefore as the local population at each habitat gets larger the more efficiently the TL SSA performs compared to the DTSC SSA. However, the more dramatic time saving was due to using FSR compared to using DR (Figure 4.6B). In this particular scenario, using FSR reduced the simulation time to 18.3% of the time for DR. At this point it is interesting to compare to the performance of FSR relative to DR for the farm model considered above with $N = 900$; using the FSR

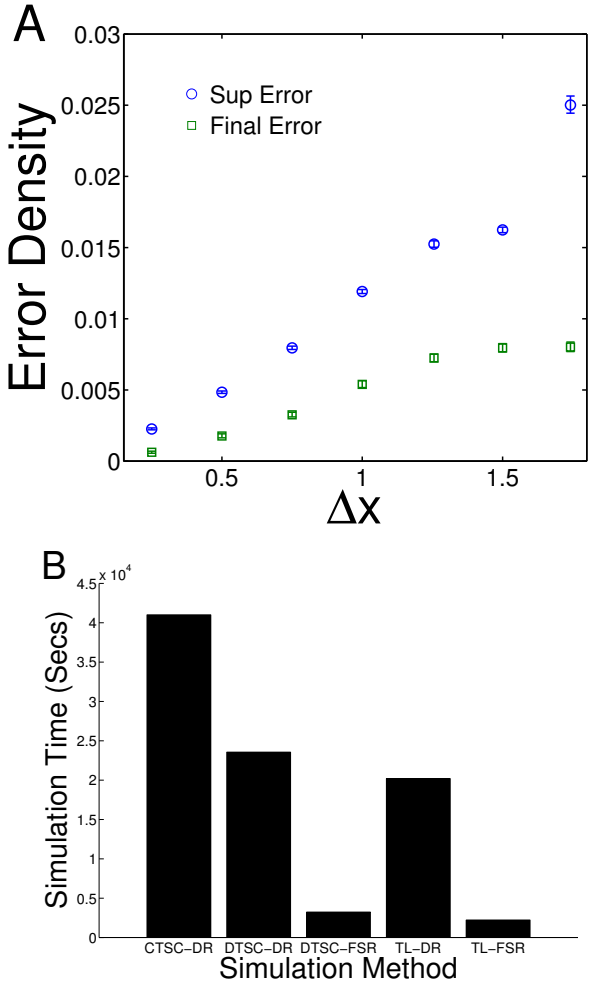


Figure 4.6: Results for model of disease spread amongst populated plant habitats. For the discrete time Sellke construction comparison is made between simulations using direct rate recalculation and using fast spectral rate recalculation. Latent variables are matched in order to directly compare simulation methods, and comparisons taken over 100 match MC replicates. All confidence intervals are due to bias corrected 95% confidence boot-strapping. The base set of parameter is $N = 900$, $\Delta x = 0.5$, $L = 3$. The number of plants at each habitat is distributed $\mathcal{B}(40, 0.5)$, in this example total population of plants was $17,909$. **A**: Varying Δx from the base value. The error profiles as in the farm model, but is reduced for each choice of parameter. Other parameter variation follows the profile of the farm model, but also with reduced error [not shown]. **B**: A comparison of time taken by five spearate methods in simulating 1000 MC replicates of the plant model, data for the FSR method includes set-up time. Continuous time Sellke construction with directly recalculated rates (CTSC-DR), Discrete time Sellke construction with directly recalculated rates (DTSC-DR), Tau-leap method with directly recalculated rates (TL-DR) and the last two also with fast spectral rate recalculation (FSR). I note that in the more populated plant model, the FSR method outperforms DR more dramatically than in the farm model.

method for the plant model with population dynamics gives higher accuracy relative to the DR method than in the farm model at a greater relative speed advantage.

This improvement can be understood in terms of large population limits, where the local population size at each habitat (N_i) becomes large. If I assume the that K_0 is suitably scaled (such that transmission is frequency dependent) then the dynamics of the densities $\{S_i(t)/N_i, I_i(t)/N_i\}_{i=1}^{N}$ converge onto a deterministic trajectory (Kurtz 1970; Kurtz 1971). This can be viewed as the limit of *infinite* events per time-step for any $\delta t > 0$. Integration of such deterministic dynamics, for example using Runge-Kutta, requires at every time-step the calculation of the force of infection at each habitat via a sum over all habitats at cost $\mathcal{O}(N^2)$. In contrast, the FSR method spatially extends the image of infected densities at habitats to a smooth field on $\mathcal{A}$. The time-dynamics are therefore given by an integro-differential equation; the FSR method then becomes equivalent to using a highly accurate Spectral-Galerkin method (Hesthaven, Gottlieb, and Gottlieb 2007), the cost per time-step being $\mathcal{O}(N \ln N)$ due to FFT.

4.5 Discussion

I have treated various stochastic simulation methods for a very common class of models for spatial dispersion of an invasive pathogen. In particular, I have demonstrated a novel method for recalculating stochastic transition rate upon arriving epidemic events using a convolution solution that, in conjunction with time discretising simulation algorithms, can deliver significant time saving to large scale Monte Carlo investigations. The convolution solution uses only 'out-of-the-box' software for implementation that is readily and freely accessible, moreover the analytic properties of the error in spectral convolution solving are well understood.

Since the FSR method is an addendum to commonly used stochastic simulation algorithms it is flexible and is not restricted to solely SIR type spatial epidemic modelling. Rather, the FSR method is a possible tool for accelerating simulation of models concerned with spatial dispersal where the dispersion kernel is translation invariant and smooth. Additionally, for the case where the transmission kernel is Gaussian shaped the FSR error can be reduced by making an appropriate correction to the kernel. It should be noted that the treatment of habitats as point locations is a modelling assumption in order to allow the force of infection to be written as a weighted sum over infected habitats. The FSR method allows the very natural

relaxation of this assumption to modelling scenarios where habitats have spatial extent.

FSR is particularly successful for large population simulations, whether that refers to the existence of many habitat locations or large local populations at each habitat, with long-range dispersion. When dispersion is highly localised potentially a different method for accelerating rate recalculation should be considered, such as space subdivision of $\mathcal{A}$. Heuristically, for very localised transmission, the force of infection for a susceptible habitat can be compactly represented as a truncated sum over only a comparatively small number of nearby infectious habitats but would require many discrete modes for description in the spatial frequency domain used for FSR convolution solving. The converse is that for disperse transmission a significant fraction of the total population of infectious habitats contribute to the infection hazard for each susceptible habitat, whereas in the frequency domain only a comparatively small number of modes are required for a good description of the force of infection field. In this situation FSR performs best compared to more direct methods.

Monte Carlo updating is not just used for investigating simulated outcomes of a stochastic model; a popular methodology for parameter estimation in epidemiology is the use of Markov chain Monte Carlo (MCMC) (O'Neill 2002). The likelihood calculations necessary in a number of variants of MCMC involve sequential calculations of transition rates between (proposed) events, which is where I have focused my investigation presented in this chapter. Another intriguing avenue to parameter estimation is through the numerical solution of filtering problems via Sequential Monte Carlo (aka. the particle filter) (Bain and Crisan 2009), which relies on simulating large numbers of interacting realisations of the process of interest. In addition to the intrinsic interest in filtering noisy data to estimate current state given confidence in model parameterisation, sequences of filtering operations have been used in parameter inference for dynamical systems (Ionides, Bret, and King 2006; A. A. King et al. 2008). FSR could play a role in increasing the numerical efficiency of these computationally intensive but potentially very useful techniques.

Chapter 5

Optimal Control of Epidemics using Dynamic Programming

5.1 Intervention and Epidemiology

In previous chapters I have investigated the dynamic properties of a stochastic model for the temporal and spatial evolution of an invasive pathogen amongst a population of potentially susceptible populations spatially stratified into patchy habitats. In this chapter I investigate intervention into the spatial epidemic in such a manner as to minimise a form of economic cost, which covers both the cost induced by the disease status of the spatially separated populations and the costs associated with deploying intervention. I will review the dynamic programming approach to solving multi-stage or dynamic optimisation problems (Bellman 1952; Bellman 1954) and apply this theory to the stochastic and spatial model of epidemic spread that I have used through out this work. The goal is to construct policies for deciding upon optimal interventions into the ongoing epidemic which are reactive to the arrival of new information about epidemic progression. Firstly I will contextualise this approach within the much wider literature of using epidemiological models to guide intervention strategies.

On one hand there is intrinsic scientific interest in investigating mathematical epidemiological models as a paradigm for clear thinking about epidemic risk (Roberts and Heesterbeek 1993; Brauer 2009). On the other hand from a disease management perspective any given epidemiological model must past two crucial *relevance tests*: does the model correspond to a real-world situation of interest, and if so can it be parameterised in such a manner as to have predictive power? The

guide here for the theoretician is to construct models with maximal parsimony; that is to use the simplest model that can explain available data (but no simpler!). This is a common principle across the entirety of applied science; it is discussed in Keeling and Rohani with direct reference to intervention measures for epidemic control (Keeling and Rohani 2008) chapter 8. A parsimonious model, but sufficiently complex to explain observation, can be viewed as more likely from a Bayesian point of view given *a priori* uncertainty (Coolen, Khn, and Sollich 2005) chapter 6.

Having passed the relevance tests a mathematical model is an invaluable tool for comparing the effectiveness of intervention strategies. This necessarily involves a certain degree of subjectivity on the part of the agency intervening against the epidemic since there are a multiplicity of plausible control objectives. Plausible objectives include reduction of transmissibility leading to epidemic eradication (Anderson and May 1985b), minimisation of infectious incidence (Galvani and Reluga 2007), reduction in peak epidemic severity (House et al. 2011) or the minimisation of economic measures of disease impact such as the quality adjusted life year (QALY) (Baguelin et al. 2010). Moreover, it is possible for these objectives to be contradictory to one another. A classic example of such a trade-off can be see for highly infectious diseases that are typically caught by the young after the protective effect of their maternally derived antibodies has dissipated such as measles, mumps and rubella (MMR). Models of these childhood diseases and their most efficient age group targeted vaccination policy have been extensively studied (Anderson and May 1983; Mclean and Anderson 1988); it is paradoxically possible for vaccination coverage to decrease total incidence per year whilst increasing incidence amongst pregnant women for whom rubella infection can lead to severe complications.

Probably the most important class of intervention for disease control is the use of vaccination as a prophylactic measure for 'removing' individuals, or even whole populations through the herd immunity effect (Kermack and McKendrick 1927; Anderson and May 1992), from epidemic risk. The vaccination of individuals does not just imply their protection, but also removes potential future secondary cases caused by infectious transmission from the individual. This observaiton gives rise to the classic threshold vaccination level for eradication in a well mixed population demonstrated in the introductory chapter. When the susceptible population is not well mixed the uniform coverage threshold is generically higher; however a targeted campaign of vaccination can lead to good results (Anderson and May 1992). For diseases of humans the understanding of age-stratified mixing has been greatly

advanced by the POLYMOD study (Mossong et al. 2008) into pan-european contact patterns. At an even more detailed level where individual to individual contact networks can be partially reconstructed there is the potential to target vaccination at individuals expected to contribute more strongly in the event of epidemic outbreak (Salath and Jones 2010). For a holistic treatment of vaccination policy during an influenza pandemic aimed at considering age, risk and spatially stratified mixing Keeling and White (Keeling and White 2011) give a good overview.

As was mentioned in chapter 4 the spatial meta-population model considered throughout this work is closely related to models used to represent the dynamics and effect of control measures for the 2001 Foot-and-mouth (FMD) outbreak in the UK (Keeling et al. 2001; Ferguson, Donnelly, and Anderson 2001b). Due to the obscurity of the detailed within farm contact structure of individual livestock and the rapid within farm spread observed after initial introduction a compartment model for the disease status of the entire population is quite appropriate. A metapopulation model with spatially decaying transmission rate was a successful model for farm-to-farm transmission after a ban on livestock movement restricts transmission to be a spatially local process; before such a ban a network transmission model based upon livestock movement is likely a better model (Kiss, Green, and Kao 2006). The primary control mechanism was a veterinary discovery and investigation process for assessing potential transmission risk and the location of infectious individuals amongst the livestock herds at each farm. The analogy to the spatial metapopulation model used in this work is clear; habitats are farms, populations are the aggregate of the farm livestock herds and removal is detection and culling intervention removing the possible of future transmission from the infectious premise. The ability of such a spatial metapopulation model, augmented with heterogenous population infectiousness and susceptibility depending on the composition of the livestock numbers and species, to recreate and predict the 2001 FMD epidemic has been well investigated (Ster and Ferguson 2007; Tildesley et al. 2008) and not only for FMD but also other diseases of commercial livestock such as bovine tuberculosis (bTB) (Gilbert et al. 2005; Donnelly et al. 2005).

Culling for livestock plays an intervening role in the spatial epidemic that is similar to a successful vaccination - removing the population and thereby eradicating future risk of their secondary infectious contacts. However, it is obviously an extreme and potentially costly intervention to be used judiciously, for a reterospective investigation of the efficacy of culling for the 2001 FMD outbreak c.f. (Tildesley

et al. 2009). In 2001 culling was also used as a secondary control mechanism aimed at removing livestock herds local to confirmed infectious farm premises and thereby preempting the spread of FMD (Keeling et al. 2001). Given the costly nature of culling as an epidemiological intervention there has been interest in efforts to use vaccination, assuming a viable vaccine is available, to retard the invasion of FMD amongst the farms of the UK (Keeling et al. 2003; Tildesley et al. 2006) and thereby minimise required culling. In these cases a spatial metapopulation model was useful for comparing and optimising intervention strategies to be used throughout the epidemic outbreak. This is of obvious practical and theoretical significance, but it does leave open the question of whether it might be possible to design intervention strategies with a great flexibility to react in real-time to incoming information about infectious spread using the techniques of optimal control theory.

Optimal control theory arose from a desire to choose an optimal control process in order to minimise a cost functional of the path of a given dynamical system. For deterministic dynamical systems Pontryagin and his collaborators in the 1950s developed the most commonly used solution method, Pontryagin's principle (Pontryagin et al. 1962), which gives necessary conditions for the optimal control process. Bellman was working at the same time for RAND corporation and developed dynamic programming as a solution method for multi-stage decision making with respect to stochastic dynamics (Bellman 1954). The essential ideas behind dynamic programming, applied to deterministic systems, transforms the optimal control problem into solving a non-linear PDE. This PDE was known, in the context of the calculus of variations, to Hamilton and Jacobi in the 19th century and now is commonly referred to as the Hamilton-Jacobi-Bellman (HJB) equation. Pontryagin's principle and dynamic programming have continued to dominate the field of optimal control theory since the 1950s. In a sense the twin methodologies have respected their origins; although it is possible to formulate a Pontryagin principle for a wide class of stochastic optimal control problems, c.f. (Øksendal and Sulem 2005) chapter 3, the implicit form of the necessary conditions is difficult to calculate. On the other hand, the elegance of Pontryagin's method for deterministic systems renders tackling the HJB equation partially obsolete.

There has been intermittent interest in the use of optimal control theory for vaccine scheduling decisions. For the homogeneously mixing deterministic SIR epidemic essentially the solution was found to be to deploy maximally at the beginning of the epidemic and decrease vaccination effort later as it became cost ineffective.

This has been found using a dynamic programming approach (Hethcote and Waltman 1973), and Pontryagin's principle (Morton and Wickwire 1974; Behncke 2000). More recently there has been some advance in constructing optimal controls for the case of two sub-populations; the optimal deployment of cures for an SIS type epidemic spreading between two spatially separated population groupings (Rowthorn and et al 2009) and tree culling for a model of sudden oak death (Rizzo and Garbelotto 2003) with two species (Mbah et al 2010). In both cases the disease progression was modelled as deterministic and the class of solutions, switching strategies in both examples, were derived using Pontryagin's principle. The optimal control of epidemic models with discrete population numbers and stochastic transmission has not received a great deal of attention, with such interest seemingly restricted to the homogeneously mixing case (Lefvre 1981; Clancy 1999) in continuous time or optimal vaccination decisions taken on a year-by-year basis (Viet et al. 2012). In each case a variant on dynamic programming was used in order to solve for the optimal control process. In this chapter I will extend the current literature by numerically solving a stochastic and discrete SIR epidemic process with populations separated into two spatially separated groupings and investigate the resultant optimal decision processes for both a scenario where vaccination is relatively cheap but cannot be deployed rapidly and where vaccination deployment is potentially fast but the vaccine is relatively expensive. I will also investigate the cost impact of lacking information such as the controlling agency being unable to assess the exact disease status of each population or the invasive pathogen being able to invade 'silently' until random detection.

The fundamental difficulty with using dynamic programming to solve stochastic optimal control problems for epidemics is that the HJB equation requires solution for every possible configuration of local disease states for the population. For any spatial epidemic model of serious practical importance the number of possible configurations or epidemic states is effectively innumerable, and therefore numerically exact solutions for the stochastic optimal control problem are restricted to 'toy' epidemic models. This observation possibly explains the dearth of work on optimal intervention for epidemics. However since at least the work of Pardoux and Peng (Pardoux and Peng 1992), which has been extended by Bouchard et al (Bouchard, Elie, and Touzi 2009), the solution to non-linear PDEs (a class the HJB equation falls into whenever the underlying process is deterministic or a diffusion) has been interpreted as the expectation of a corresponding stochastic process represented as a system of stochastic differential equations evolving both forwards and back-

wards in time (FBSDE). This is the non-linear analogous formula to the famous Feynman-Kac formula for linear PDEs, c.f. for example (Klebaner 2005) chapter 6. The benefit of the FBSDE representation is computational, a problem of unfeasible computation size can be reduced to the recursive simulation of sets of SDEs and a reliance on Monte Carlo convergence. This kind of Monte Carlo approach is closely related to approximate dynamic programming (ADP) (Powell 2007) which describes the broad field of Monte Carlo techniques used to estimate the solution to HJB equations that involve too many states as to be directly solvable. ADP methods are commonly used to solve real-world problems of resource management (e.g. (Zhang and Adelman 2009)) but their use for epidemic control is very sparse, the only example the author is aware of is the use of the Longstaff-Schwartz algorithm (Longstaff and Schwartz 2001), which is related to ADP methods, to solve an optimal stopping problem for a homogeneously mixing stochastic SIR epidemic (Ludkovski and Niemi 2010).

In this chapter I will introduce the field of optimal control theory and give a formal derivation of the HJB equation which is essential for the dynamic programming treatment. Having established the ideas behind dynamic programming I give rigorous conditions for its application to the solution of optimal control problems, these can be found in Fleming and Soner (Fleming and Soner 2006). Specialising to the spatial epidemic I give the HJB equation for the spatial epidemic and introduce the value iteration method of numerically exact solution as well as introducing a novel ADP algorithm for the estimation of the solution to the HJB equation and generating approximately optimal vaccination decisions. After investigating a toy model of the spatial epidemic where exact solution is accessible I will demonstrate that for the toy model the ADP method can closely recreate the cost minimisation effect of exactly solved optimal vaccination allocation. Finally, I will demonstrate that the ADP method can be extended to a spatial epidemic models much too large to be broached by exact numerical methods.

5.2 Introduction to Optimal Control Theory

A protagonist (whom we dub the *Controller*) exerts control on some process through the ability to influence it's dynamics, although this influence is usually considered to be constrained in certain ways. If the controller's goal is to minimise some associated cost functional on paths and the terminal state then the controller is attempting to solve an *optimal control* problem. In this section I will re-introduce Markov pro-

cesses as a solution to a martingale problem characterised by an initial distribution of states and a generator operator which encodes the probabilistic dynamics. This form for the Markov process sets the scene for the control theoretic approach of the next section. The particular form of cost functional, which is the target for minimisation will also be introduced.

We denote the process to be controlled $X = (X(t))_{t\geq 0}$ which takes values the polish space[1] Σ. I will only consider process dynamics that are Markovian, but I use a different construction than in previous chapters. The underlying probability space remains the collection $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, which refer to, respectively, the sample space, a valid σ-algebra of sets in Ω, a valid filtration on $\mathcal{F}$, taken to be the natural filtration for X, and a probability measure on $\mathcal{F}$. Previously, the state dynamics were constructed from sample elements $\omega \in \Omega$ using a stochastic integral representation. For investigating control theoretic concepts for Markovian processes it is simpler to specify a Markov process via the pair $\mathcal{G}$, an operator mapping bounded functions of Σ to bounded functions of Σ and a probability distribution $\mathbb{P}_0$ for the initial state $X(0)$. The pair $(\mathcal{G}, \mathbb{P}_0)$ specifies the process X through the condition that every bounded function on Σ, f, must satisfy,

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{G}f(X(s))ds = M(t), \qquad t \geq 0. \tag{5.1}$$

Where $M(t)$ is a $\mathcal{F}_t$-measurable martingale process, cf definition 2.2.2. This has been referred to as the *martingale problem* version of characterising a Markov process e.g. (Kurtz 2007) Chapter 1. The *control process* is denoted $u = \{u(t)\}_{t\geq 0}$, and takes values in some space $\mathcal{U}$ called the *admissible control space*. The control process is the time varying sequence of control implementations by the controller; potentially the controller decision to intervene with some control measure at some time t might be based on all the available information of the random process X up until the time time t, that is the space of $\mathcal{F}_t$-measurable random variables.

The decision function or *policy* that the controller uses to decide upon admissible interventions is denoted, π. By allowing the controller to be reactive to the stochastic evolution of X the policy couples the control process to the dynamics. On the other hand for each policy π the state of the system is influenced by the control process and hence we have a feedback loop between state influencing controller decision which in turn influence the state and so forth.

[1] A polish space is a complete, separable metric space, e.g. $\mathbb{R}^n$.

The martingale problem form (5.1) implies the following differential limit form for the generator, that for all bounded functions on Σ, f,

$$\mathcal{G}f(x) = \lim_{t \to 0_+} \frac{\mathbb{E}_{t,x}[f(X(t))] - f(x)}{t}. \tag{5.2}$$

Where the expectation is conditioned on $X(0) = x \in \Sigma$. For controlled processes the control decisions of the controller influences the expected dynamics. If the controller is following policy π I denoted the controlled process X^π, the corresponding control process u^π and expectations $\mathbb{E}^\pi$; however since it is the actual control measure implemented that effect the dynamics, I denote the controlled generator at time s, $\mathcal{G}^{u^\pi(s)}$. Note that the generator is no longer time homogeneous due to the potentially time varying action of the controller in influencing the dynamics through control decisions. I only consider control policies where $\mathcal{G}^{u^\pi(s)}$ exists in the sense of (5.2) for each time s.

The task is to control the process in such a way as to minimise (over admissible control policies) the accumulated cost along the path of the process up until the final time T plus an additional cost for the final state $\psi(\cdot) : \Sigma \to \mathbb{R}$. Given a particular sample path of states and controls (X^π, u^π) the associated cost is,

$$\int_0^T L(s, X^\pi(s), u^\pi(s))ds + \psi(X^\pi(T)). \tag{5.3}$$

The control process $u^\pi(\cdot)$ takes admissible values specified by the dynamics of the process and the policy π. There is no requirement that the final time T must be finite. L is the *Lagrangian* of the problem, the instantaneous cost associated with the triple $(t, X(t), u(t))$. Each admissible control policy gives an expected cost,

$$J(x;\pi) = \mathbb{E}_x^\pi\Big[\int_0^T L(s, X^\pi(s), u^\pi(s))ds + \psi(X^\pi(T))\Big] \tag{5.4}$$

where $\mathbb{E}_x^\pi$ is the expectation over the outcomes of the stochastic process X^π controlled by policy π and conditioned on $X(0) = x \in \Sigma$. The optimal control process associated with the optimal policy π^* is denoted $u^{\pi^*} = u^* = \{u^*(t)\}_{t \geq 0}$. Under optimal control the cost function is called the *value function* V of the control problem. For each $x \in \Sigma$ the value function is given by,

$$V(x) = J(x;\pi^*) = \inf_{\pi \in \Pi} J(x;\pi) \qquad \forall x \in \Sigma. \tag{5.5}$$

Where $V(x)$ is called the value of the state $x \in \Sigma$. The value function will play a crucial role in the sequel. Π is the space of admissible policies; if there is some constraint on the *instantaneous* control decision at time t, $u(t)$, then I call the set of possible instantaneous controls $\mathcal{U}$. I have implicitly treated the space of policies as containing only deterministic decision functions; that is knowledge of the process completely determines the controller's action. Compared to other applications of decision theory such as Game theory this seems a big restriction since it is well-known that the best strategies in many games involve a chance decision element (Binmore 2007). However, we will see that even in a more general setting the optimal policy will be determined by a surprisingly small amount of information about the controlled process.

5.3 Dynamic Programming

Dynamic programming is a method for solving for the value function of a control problem, moreover we will see that solving the value function will allow us to construct the optimal policy by considering actions that consistently 'push' $X^{\pi^*} = X^*$ towards lower value states. The basic idea is to breakdown the problem of solving (5.5) directly into a sequence of sub-problems; firstly solving for optimal intervention when the first intervention time t is very close to the final time T and then using this to construct solutions sequentially backwards in first intervention time towards the time 0.

In his seminal work in the 1950s (Bellman 1953; Bellman 1954), Bellman introduced the *Bellman principle of optimality* which underlies the dynamic programming approach to solving the optimal control problems introduced above. The principle of optimality allows one to couple the solutions going backwards in first intervention time.

PRINCIPLE OF OPTIMALITY [Bellman]. *An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decisions.*

In order to solve for the optimal policy I extend our interest to a time-dependent family of control problems. Instead of considering the first intervention time to be 0 and controlling up until some later time T I allow the first intervention

time to vary as $t < T$. The goal is to be able to write down an evolution equation for the value function; where the time considered is in fact the first intervention time. The cost function for first intervention at time t following policy π and conditioned on the initial state $X(t) = x$ is given by,

$$J(t, x; u) = \mathbb{E}^{\pi}_{t,x}\Big[\int_t^T L(s, X^{\pi}(s), u^{\pi}(s))ds + \psi(X^{\pi}(T))\Big], \tag{5.6}$$

The value functions for the time-dependent family are denoted $V(\cdot, \cdot)$ with the value of a state and intervention time being $V(t, x)$. The solution to (5.5) coincides with $V(0, \cdot)$, which relates the time dependent problem with the original time independent problem of finding the infimum control for (5.4). It is important to note that the value of states for $t = T$ is known by construction,

$$V(T, x) = \psi(T, x), \qquad x \in \Sigma. \tag{5.7}$$

Bellman's principle can be thought of as a self-consistency relation for the first intervention time dependent value functions. If I divide the time interval from the first intervention time until the final time in two: $[t, T) = [t, r) \bigcup [r, T)$ where $t \leq r < T$ then the self-consistency relationship, often referred to in the literature as the *dynamic programming principle* (DPP) (Fleming and Soner 2006) Chapter 1, gives that for all intermediate times r,

$$V(t, x) = \inf_{\pi \in \Pi} \mathbb{E}^{\pi}_{t,x}\Big[\int_t^r L(s, X^{\pi}(s), u^{\pi}(s))ds + V(r, X^{\pi}(r))\Big], \ 0 \leq t \leq T, x \in \Sigma. \tag{5.8}$$

Equation (5.8) is a reformulation of Bellman's principle that states that the minimum expected cost for the initial data (t, x) is given by accumulating cost up until time r by following the optimal policy and then proceeding optimally with the new initial data $(r, X^*(r))$. Whilst this is intuitive; equation (5.8) describes balancing for all intermediate times r the myopic reward up until r, which would be optimised by using a greedy policy, with the forecasted reward after r of proceeding optimally. Since the value function is known for $t = T$ by (5.7) it is possible to exploit the DPP to solve V for all initial data.

5.3.1 Formal Derivation of the Hamilton-Jacobi-Bellman Equation for dynamic programming

In this section I give a formal derivation of a dynamical equation that the value function must satisfy. The role of the time variable being taken by the first in-

tervention time t. This dynamical equation is called the *Hamilton-Jacobi-Bellman* (HJB) equation. I focus on the interesting concepts behind deriving the HJB equation without complete rigour, as well as noting that the optimal decision at any time s will only depend on the process state $X(s)$. Consequently, the optimal policy will be in the form $\pi^* : [t, T) \times \Sigma \to \mathcal{U}$. In the next section I will introduce rigourous conditions for the application of the approach given here.

The DPP (5.8) implies that at each intermediate time r we can't pick an admissible policy that delivers a lower expected cost than the choice(s)[2],

$$\pi^* \in \arg\inf_{\pi\in\Pi} \Big\{\mathbb{E}^{\pi}_{t,x}\Big[\int_t^r L(s, X^{\pi}(s), u^{\pi}(s))ds + V(r, X^{\pi}(r))\Big\}. \tag{5.9}$$

Hence, for an arbitrary control policy $\pi \in \Pi$ we have for r,

$$\mathbb{E}^{\pi}_{t,x}[V(r, X^{\pi}(r))] - V(t, x) + \mathbb{E}^{\pi}_{t,x}\Big[\int_t^r L(s, X^{\pi}(s), u^{\pi}(s))ds\Big] \geq 0. \tag{5.10}$$

With equality iff $\pi = \pi^*$ satisfies (5.9). Since (5.10) holds for arbitrary r we have the freedom to choose $r = t + \delta t$ for some $\delta t > 0$ and rescale by $1/\delta t$ giving,

$$\frac{1}{\delta t}\Big[\mathbb{E}^{\pi}_{t,x}[V(t + \delta t, X^{\pi}(t + \delta t))] - V(t, x) + \mathbb{E}^{\pi}_{t,x}\Big[\int_t^{t+\delta t} L(s, X^{\pi}(s), u^{\pi}(s))ds\Big]\Big] \geq 0. \tag{5.11}$$

From equation (5.2) taking the limit $\delta t \to 0_+$ in (5.11) for the first two terms is equivalent to the action of the generator of a process, in this case the generator the time augmented process $Y(t) = (t, X(t))$, upon the value function. Since the time process is smooth and deterministic by applying partial derivatives we see that,

$$\lim_{\delta t\to 0_+} \frac{\mathbb{E}^{\pi}_{t,x}[f(Y(t + \delta t))] - f(t, x)}{\delta t} = \frac{\partial f}{\partial t}(t, x) + \mathcal{G}^{u^{\pi}(t)} f(t, x). \tag{5.12}$$

Introducing this into equation (5.11) gives a differential inequality that must be satisfied for all admissible policies,

$$\frac{\partial}{\partial t}V(x, t) + \mathcal{G}^{u^{\pi}(t)}V(t, x) + L(t, x, u^{\pi}(t)) \geq 0,\ u^{\pi}(t) \in \mathcal{U}, 0 \leq t < T, x \in \Sigma. \tag{5.13}$$

On one hand choosing $\pi = \pi^*$ recovers the equality in (5.11), but on the other hand establishing the equality in (5.11) is a minimisation over the possible instantaneous

[2]There is no necessity for the optimal control to be unique.

values of $u^\pi(t) \in \mathcal{U}$. Ergo, we have, formally, established the *Hamilton-Jacobi-Bellman* (HJB) equation that the value function for the abstract control problem (5.5) must satisfy

$$\frac{\partial V}{\partial t}(t,x) + \min_{u\in\mathcal{U}}[\mathcal{G}^u V(t,x) + L(t,x,u)] = 0, \qquad 0 \leq t < T, x \in \Sigma. \tag{5.14}$$

A particular solution to (5.14) is specified by using the known $V(T,\cdot)$ as a boundary condition,

$$V(x,T) = \psi(x), \qquad x \in \Sigma. \tag{5.15}$$

The HJB equation (5.14) defines a non-linear evolution equation for the value function V, evolving *backwards* in intervention time from the end-state condition (5.15).

As a consequence of the reasoning above the optimal policy π^* defines an optimal control process $u^* = (u^*(t), t \geq 0)$ that must satisfy

$$u^*(t) \in \arg\min_{u\in\mathcal{U}}[\mathcal{G}^u V(t, X^*(t)) + L(t, X^*(t), u)], \qquad 0 \leq t < T, x \in \Sigma. \tag{5.16}$$

Equation (5.16) is extremely valuable for both the theoretical treatment and numerical solution of optimal control problems, since it gives the optimal control process in *feedback* or *Markov* form. The minimisation in (5.16) is over a function only of state and intervention time. Hence for each intervention time and state pair $(t, X^*(t))$ arrived upon the optimal policy π^* is to choose the control decision $u^*(t)$ according to (5.16), which can be derived by solving the HJB equation (5.14). This is a rather remarkable result due to the essentially Markovian character of the dynamics and form of cost functional; despite the controller in principle having access to the complete information of the process up until some time t the optimal decision is always dependent only on the *current* time and state, not some more general $\mathcal{F}_t$-measurable random variable. Moreover, the chance element in optimal decision making is restricted to the stochastic dynamics; conditioned on the arrival of the pair $(t, X^*(t))$ the optimal decision is deterministic. In functional terms we can state that,

$$\pi^* : [0,T) \times \Sigma \to \mathcal{U}. \tag{5.17}$$

Control towards an Infinite Time Horizon

In some applications there is no natural time horizon over which control occurs. In this case it is common to consider classes of cost functional over processes that

extend over infinite time, e.g.

$$J(x;\pi) = \mathbb{E}_x\Big[\int_0^\infty e^{-\beta s} L(X^\pi(s), u\pi(s))ds\Big], \qquad \beta \geq 0. \tag{5.18}$$

The positive parameter β is used to ensure convergence of cost. Intuitively, it reflects the controller's preference for minimising cost over the nearer future and her discount of the distant future costs. Thus β is called the *discount factor*.

For the formal derivation of a dynamic programming equation for the infinite time horizon case we use the same 'trick' of considering a family of time-dependent controls up until a terminal time T, but then take T to infinity. In this part we assume that this does not cause J to explode, rigourous conditions will be given in the following section. Following the same reasoning as above gives,

$$J(t,x;\pi) \quad = \quad \mathbb{E}_{t,x}\Big[\int_t^T e^{-\beta s} L(X^\pi(s), u^\pi(s))ds\Big]$$

Shifting the time variable such that the first intervention time t is again the origin time gives,

$$J(t,x;\pi) \quad = \quad e^{-\beta t}\mathbb{E}_{t,x}\Big[\int_0^{T-t} e^{-\beta s} L(X^\pi(s+t), u^\pi(s+t))ds\Big].$$

Taking $T \to \infty$ and considering the time shifted control policies $\hat{\pi}(s,x) = \pi(s+t,x),\ \forall s \geq 0$ gives

$$J(t,x;\pi) = e^{-\beta t} J(0,x;\hat{\pi}). \tag{5.19}$$

The time-shifted policies also inhabit Π hence the infimum control over the expression (5.19) suggests that the time dependent value function takes the special form.

$$V(t,x) = e^{-\beta t} V(x). \tag{5.20}$$

Where $V(\cdot)$ is the value function for infinite time horizon control problem,

$$V(x) = \inf_{\pi\in\Pi} J(x;\pi),\ \forall x \in \Sigma. \tag{5.21}$$

Substitution into the HJB (5.14) gives,

$$\beta V(x) = \min_{u\in\mathcal{U}}[\mathcal{G}^\pi V(x) + L(x,u)]. \tag{5.22}$$

Equation (5.22) is the dynamic programming equation for the infinite time horizon control problem defined by finding the infimum control for (5.18). Since time dependence has vanished from equation (5.22) I refer to it as an 'elliptic'-type equation, in analogy to the defining equation for the stationary measure of a diffusion. The Markov feedback control process u^* is given by at each time solving

$$u^*(t) \in \arg\min_{u \in \mathcal{U}} [\mathcal{G}^\pi V(X^*(t)) + L(X^*(t), u)]. \tag{5.23}$$

We note for the elliptic case the control decision is governed by finding the minimiser for a function of state alone. Hence, only the state arrived at is required to determine the optimal policy,

$$\pi^* : \Sigma \to \mathcal{U}. \tag{5.24}$$

5.3.2 Classical Solutions to the HJB Equation and Verification Theorems

The arguments above give the value function as the solution to the non-linear evolution equation (5.14) in the finite time horizon case or the elliptic-type equation (5.22). Importantly for applications a feedback control policy can be inferred from the solution to the value function. In this section I give conditions for the arguments of the previous sections to be made rigourous and state *verification theorems* for the solution V to the HJB equation. The purpose of the verification theorems is to confirm that the value functions constructed as a solution to the HJB equation are in fact minimal for each pair (t, x) over the set of admissible controls. Additionally, the verification theorems confirm that the the feedback control of the process is given as the minimiser of a function that can be constructed from the solution of the value function. Proofs for the below can be found in Fleming and Soner chapter 3 sections 8 and 9 (Fleming and Soner 2006).

I denote the domain of the operator $(\frac{\partial}{\partial t} + \mathcal{G}^u)$ by $\mathcal{D}(\frac{\partial}{\partial t} + \mathcal{G}^u)$. In principle this domain depends on the control decision $u \in \mathcal{U}$, hence I assume that there exists a space $\mathcal{D} \subseteq \mathcal{D}(\frac{\partial}{\partial t} + \mathcal{G}^u)$ $\forall u \in \mathcal{U}$. We call V a *classical solution* of the Hamiltonian-Jacobi-Bellman equation if V satisfies (5.14) with terminal data (5.15) $\forall (t, x) \in [0, T] \times \Sigma$ and $V \in \mathcal{D}$.

I now extend the concept of an admissible policy $\pi \in \Pi$, which I have treated simply as some decision rule, and introduce the concept of the *admissible control*

system. In the control system interpretation each π is a probability space for state and decisions; which is a much more general space than treating each π as a fixed decision policy. To be precise, for each initial condition (t, x),

$$\pi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, X, u) \tag{5.25}$$

is an admissible control system if $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, \mathbb{P})$ is a filtered probability space and the following set of conditions on the state process and control process are satisfied:
Conditions for Admissible Control System (Finite Time Horizon) :

1. $X(s) \in \Sigma$, $t \leq s \leq T$, $X(t) = x$.
2. X is *càdlàg* and adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$.
3. $u(s) \in \mathcal{U}$, $t \leq s \leq T$.
4. u is adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$.
5. The feedback form of the control process $u(\cdot, \cdot) \to \mathcal{U}$ is a measurable function on $[t, T] \times \Sigma$.
6. For all $W \in \mathcal{D}$, where $\mathbb{E}_{t,x}[|W(T, x)|] < \infty$ and

$$\mathbb{E}_{t,x}\Big[\int_t^T \Big|\Big(\frac{\partial}{\partial t} + \mathcal{G}^{u(s)}\Big)W(s, x)\Big|ds\Big] < \infty$$

then the differential limit holds

$$\lim_{\delta t \to 0_+} \frac{\mathbb{E}_{s,x}[W(s + \delta t, X(s + \delta t)] - W(s, X(s))}{\delta t} = \Big(\frac{\partial}{\partial t} + \mathcal{G}^{u(s)}\Big)W(s, X(s)).$$

There are now sufficient ingredients to guarantee that a classical solution to the finite time horizon HJB equation (5.14) with end state condition (5.15) is, in fact, the value function of the control problem from which the optimal feedback control can be derived.

Theorem 5.3.1. (Verification Theorem: Finite Time Horizon)
Let $V \in \mathcal{D}$ be a classical solution to (5.14) *with end state condition* (5.15)*. Then $\forall (t, x) \in [0, T] \times \Sigma$:*
(a) $V(t, x) \leq J(t, x; \pi)$ for every admissible control system.
(b) If there exists an admissible system $\pi^* = (\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, \mathbb{P}, X^*, u^*)$ such that

$$u^*(s) \in \arg\min_{u\in U} \Big\{\Big(\frac{\partial}{\partial t} + \mathcal{G}^u\Big)V(s, X^*(s)) + L(s, X^*(s), u)\Big\}$$

for $\tilde{\mathbb{P}}$-almost all $(s,\omega) \in [t,T] \times \Omega$, where $\tilde{\mathbb{P}}$ is the product measure of Lebesgue measure and $\mathbb{P}$, then $V(t,x) = J(t,x;\pi^*)$.

I have generally avoided using the general form of a control system of the policy π preferring to interpret the policy as a decision function rather than the probability space required for the general treatment; Theorem 5.3.1 justifies this treatment since the space of deterministic decision functions will effectively contain the optimal control system as the probability space where the control process u^* obeys the optimal decision function type policy π^* with probability 1.

For the infinite time horizon case there is a concern that the associated expected cost for each admissible policy might blow-up as the control period expands to infinity despite a given discount factor β. Hence, a further set of conditions are introduced in order to exclude this possibility from consideration. These conditions for admissibility of π (in the control system interpretation) are in addition to those for the finite time horizon case where $T < \infty$.
Additional Conditions for Admissible Control System (Infinite Time Horizon) :

1. $\mathbb{E}_x\Big[\int_0^\infty e^{-\beta s}|L(X(s),u(s))|ds\Big] < \infty$.
2. Condition 6. for the finite time horizon case holds with

$$W(t,x) = e^{-\beta t}\phi(x)$$

for all $\phi \in \mathcal{D}$ and $t \geq 0$.

For the infinite time horizon case in principle it is necessary to consider classical solutions V to the elliptic HJB equation (5.22) where for some admissible control system and discount β, $\lim_{T\to\infty} e^{-\beta T}\mathbb{E}_x^\pi[V(X(T))] \neq 0$; that is the value of being in the current state grows exponentially. This leads to further complications that will not be required for this work. Therefore, I only consider control systems where

$$\lim_{T\to\infty} e^{-\beta T}\mathbb{E}_x^\pi[V(X(T))] = 0, \qquad x \in \Sigma, \beta \geq 0. \tag{5.26}$$

Where including the non-discounted case $\beta = 0$ is intentional. This considerably simplifies the verification theorem for the infinite time horizon case.

Theorem 5.3.2. (Verification Theorem: Infinite Time Horizon)
Let $V \in \mathcal{D}$ be a classical solution to (5.22) *which satisfies (5.26). Then for $(s,x) \in [0,\infty) \times \Sigma$:*

(a) $V(x) \leq J(x;\pi)$ *for every admissible control system.*
(b) If there exists an admissible system π^* *such that*

$$u^*(s) \in \arg\min_{u\in U}\Big\{\mathcal{G}^u V(X^*(s)) + L(X^*(s),u)\Big\}$$

for $\tilde{\mathbb{P}}$*-almost all* $(s,\omega) \in [t,T]\times\Omega$*, where* $\tilde{\mathbb{P}}$ *is the product measure of Lebesgue measure and* $\mathbb{P}$*, then* $V(x) = J(x;\pi^*)$*.*

As for theorem 5.3.1 the optimal control policy is to choose effectively deterministically an admissible action based on minimising a function of current state value and instantaneous cost.

5.4 Optimal Control for the Spatial Epidemic

In this section I apply the dynamic programming approach to optimal control of the spatial epidemic outbreak. Key to this is the construction of the HJB equation for the spatial epidemic. The exact form of controlled generator for the epidemic process depends on the type of control desired to be modelled and the model choice for the epidemic dynamics. For the discrete jump process I have considered throughout this thesis the generator is in the form of a difference operator that can be solved explicitly using the stochastic dynamics derived in chapter 2. The controller is treated as a central authority with the capability of deploying a vaccination campaign aimed at removing susceptible populations from epidemic risk in order to reduce economic cost. The number of doses/interventions is not limited but rather the maximum rate at which they can be deployed is subject to a hard limit, moreover faster deployment entails increasing costs.

The Lagrangian I use to model the economic impact of the epidemic is scaled in a currency such that the cost per unit time incurred by each infected population is one. In this currency scale the cost due to deploying intervention methods such as vaccination is given by a cost function $\kappa(\cdot) : \mathcal{U} \to \mathbb{R}$. Hence, by following some admissible policy π the expected cost of the epidemic up until time T is

$$J(t,x;\pi) = \mathbb{E}^{\pi}_{t,x}\Big[\int_t^T \sum_i I_i(X^{\pi}(s)) + \kappa(u^{\pi}(s))ds + \psi(X^{\pi}(T))\Big]. \tag{5.27}$$

Where $I_i(\cdot)$ indicates the local infectious state of the ith population. I will not write the explicit dependency on state of indicator functions used below wherever it will not cause confusion. The exit state cost function $\psi(\cdot)$ is chosen so that $\psi(x) = 0$

for all epidemic states $x \in \Sigma$ with no infectious populations and is increasing in increased numbers of infectious populations. The optimal control problem for the spatial epidemic is to find an admissible policy π^* such that,

$$V(t,x) = J(t,x;\pi^*) \leq J(t,x;\pi), \qquad x \in \Sigma,\ t < T,\ \pi \in \Pi. \tag{5.28}$$

5.4.1 Vaccination and the Controlled Generator

In the previous chapters the starting point for my investigation of the stochastic spatial epidemic was the stochastic integral form given by theorem 2.2.2. However, in this chapter it was more appropriate to specify the controlled process X through the martingale problem form (5.1) for the initial distribution of states $\mathbb{P}_0$ and the generator operator $\mathcal{G}$. Below I start from the stochastic integral representation and infer the generator for the martingale problem form. I will also introduce a stochastic model for the vaccination efforts of the controller and include their effect in the generator.

The Generator for the Spatial Epidemic

I briefly re-familiarise the reader with the basic stochastic spatial epidemic model used in this work. As before I consider a meta-population model of N populations each with a disease state summarised by SIR compartmental paradigm. Each population abides at a habitat; the habitat for the ith population being spatially located at $x_i \in \mathcal{A} \subset \mathbb{R}^d$. The stochastic rate of disease transmission from population to population is governed by their spatial separation through the the transmission kernel $K(\cdot)$. The rate of spread from an infectious population at x_i to a susceptible population at x_j being $K(x_i - x_j)$. If populations share a habitat location they are considered to cohabit and that habitat has internal disease dynamics defined by zero-range local transmission, $K(0)$. The stochastic rate of recovery is γ. The epidemic state at time t will be given in vector form via the familiar $2N$ indicator functions for the local disease status for each population,

$$X(t) = (S_1(t), I_1(t), \ldots, S_N(t), I_N(t)). \tag{5.29}$$

The stochastic evolution of the indicator functions is given in stochastic integral form as,

$$S_i(t) = S_i(0) - \sum_{j(\neq i)} \int_0^t S_i(s^-) I_j(s^-) dN_{ij}^I(s), \tag{5.30}$$

$$I_i(t) = I_i(0) + \sum_{j(\neq i)} \int_0^t S_i(s^-) I_j(s^-) dN_{ij}^I(s) - \int_0^t I_i(s^-) dN_i^R(s), \tag{5.31}$$

$$i = 1, \ldots, N, \ t \geq 0.$$

Where the integrators above are Poisson processes $\{N_{ij}^I \sim PP(K(x_i - x_j))\}_{i \neq j}$ and $\{N_i^R \sim PP(\gamma)\}_i$.

First I define the following elementary vectors for encoding the change in $X(t)$ due to the arrival of an infection or recovery event. The elementary vector for an infection event at the ith habitat is

$$e_i^I = (0, 0, \ldots, -1, 1, \ldots, 0, 0). \tag{5.32}$$

Where the non-zero components are in the $(2i-1)$th and $(2i)$th positions, reflecting that the local state for the ith population is altered by losing a susceptible and gaining an infected after an infection event. Similarly, the elementary vector for a recovery at i is

$$e_i^R = (0, 0, \ldots, 0, -1, \ldots, 0, 0). \tag{5.33}$$

Where the non-zero component is in the $(2i)$th position effects the local infectious status for the ith population.

Now I will derive the generator of the spatial epidemic process from the stochastic integral form. Let f be a bounded function of Σ and let the stochastic process $Y_f = (Y_f(t) = f(X(t)), t \geq 0)$. Since f is bounded then by theorem 2.2.3

Y_f is a pure jump process obeying,

$$\begin{aligned} Y_f(t) &= Y_f(0) + \int_0^t dY_f(s) \\ &= Y_f(0) + \int_0^t \sum_{i\neq j} S_i(s^-)I_j(s^-)(f(X(s^-)+e_i^I) - f(X(s^-)))dN_{ij}^I(s) \\ &\quad + \int_0^t \sum_i I_i(s^-)(f(X(s^-)+e_i^R) - f(X(s^-)))dN_i^R(s) \\ &= Y_f(0) + \int_0^t \sum_i S_i(s^-)\lambda_i(t)(f(X(s^-)+e_i^I) - f(X(s^-)))ds \\ &\quad + \int_0^t \sum_i \gamma I_i(s^-)(f(X(s^-)+e_i^R) - f(X(s^-)))ds + M_f(t). \end{aligned} \tag{5.34}$$

Where $\lambda_i(t) = \sum_{j(\neq i)} K(x_i - x_j)I_j(t)$ is the force of infection on the ith population due to the distribution of infectious populations at time t. I have separated the Poisson processes into their compensator and martingale decomposition, recall that for $N \sim PP(\nu)$ then $N(t) = \nu t + \tilde{N}(t)$ where (νt) is the compensator of the Poisson process and $\tilde{N}(t)$ is the compensated Poisson martingale process. This gives that $M_f(t)$ is a martingale with integral representation,

$$\begin{aligned} M_f(t) &= \int_0^t \sum_{i\neq j} S_i(s^-)I_j(s^-)(f(X(s^-)+e_i^I) - f(X(s^-)))d\tilde{N}_{ij}^I(s) \\ &\quad + \int_0^t \sum_i I_i(s^-)(f(X(s^-)+e_i^R) - f(X(s^-)))d\tilde{N}_i^R(s). \end{aligned} \tag{5.35}$$

M_f is $\mathcal{F}_t$-measurable by its integral construction, that it is a valid martingale is given by theorem 2.3.1. By comparison to the martingale problem characterisation of Markov processes we see that the spatial epidemic process is a Markov process with initial distribution $\mathbb{P}_0$ and the generator is a difference operator (the domain of which is bounded functions on Σ augmented with unphysical states that get accessed with rate 0) given by,

$$\mathcal{G}f(x) = \sum_i \Big[S_i\lambda_i(f(x+e_i^I) - f(x)) + \gamma I_i(f(x+e_i^R) - f(x)) \Big]. \tag{5.36}$$

Where $S_i = S_i(x)$ indicates the $(2i-1)$th position of the arbitrary state vector $x \in \Sigma$, similarly for λ_i and I_i. If no additional control measures are possible, apart from a rate γ of removing infectious populations, then the HJB equation (5.14) can be used to solve the expected cost of the epidemic over the time period $[0, T]$ starting in the

initial state $x \in \Sigma$: $J(0,x)$. This can be solved from the affine evolution equation backwards in intervention time,

$$\begin{aligned}\frac{\partial}{\partial t}J(t,x) \quad + \quad \sum_i \Big[& S_i\lambda_i(J(t,x+e_i^I) - J(t,x)) \\ & +\gamma I_i(J(t,x+e_i^R) - J(t,x))\Big] + \sum_i I_i = 0, \qquad (5.37)\end{aligned}$$

$x \in \Sigma, 0 \leq t < T$.

with terminal data $J(T,x) = 0$ for $x \in \Sigma$.

Vaccination as a Control Measure for Meta-populations

For the meta-population models considered in this thesis 'vaccination' takes on a rather general meaning. By vaccination control I mean any control measure designed to remove susceptible populations from the epidemic; that is to introduce a dynamical pathway for the local disease state of each population $S \to R$ thus avoiding the infectious state and incurred cost. This might be a model for deploying sufficient actual vaccine amongst the individuals of a population so that herd-immunity is achieved or any other method that achieves similar success. Crucially, the vaccination effort has no effect on infectious or removed populations although if erroneously deployed for such populations the effort still accrues some associated cost.

I will now define the vaccination model used in this work and give the controlled generator, which will be in the form (5.36) with additional terms corresponding to the possibility of vaccination events. The elementary vector for a vaccination event at the ith habitat is defined as,

$$e_i^V = (0, 0 \ldots, -1, 0, \ldots, 0, 0). \qquad (5.38)$$

Where the non-zero component is at the $(2i-1)$th position. Compared to infection events a vaccination event depletes the susceptibility of the ith population without causing infectiousness.

The controlling agency has freedom to determine both the target of vaccination effort and the rate of vaccination in order to minimise the economic impact of the epidemic (5.29), however the event times will be random. The stochastic model I use for the vaccination is closely related to the underlying stochastic model for the spatial epidemic. Let $\{N_i^V \sim PP(u_i(t))\}_{i=1}^N$ be a set of non-homogeneous Poisson

processes with time-varying intensities $\{u_i(t)\}_{i=1}^N$. The possible times of vaccination for the ith population is then given by the hitting times of the N_i^V. The controllable variables for the epidemic are the intensities of the vaccination effort for each population. For each policy π the control process takes values in the admissible control space,

$$u^\pi(t) \in \mathcal{U} = \Big\{ u = (u_1, \cdots, u_N) \in [0,\infty)^N \Big| \sum_{i=1}^N u_i \leq u_{max} \Big\}. \tag{5.39}$$

Where u_{max} is the maximum rate of vaccinating populations. This gives that the expected time taken to vaccinate all N populations is $T_{vac} = N/u_{max}$ for any policy that always deploys maximum resource without wastage of effort, such as erroneously targeting non-susceptible populations. This sets a useful time-scale for the duration of the epidemic. The cost per unit time at time t of the vaccination decision made according to policy π is $\kappa(u^\pi(t))$ where $\kappa(\cdot)$ is assumed to be an increasing function in each vaccination rate $u_i^\pi(t)$.

The compensator and martingale decompositions for the vaccination Poisson drivers if the the policy π is followed are,

$$N_i^V(t) = \int_0^t u_i^\pi(s)ds + \tilde{N}_i^V(t) \implies dN_i^V(t) = u_i^\pi(t)dt + d\tilde{N}_i^V(t)$$

Where $\tilde{N}_i^V$ has the martingale property. The differential stochastic dynamics for the controlled spatial epidemic with policy π is then given by,

$$\begin{aligned} dS_i(t) &= -\sum_{j(\neq i)} S_i(t^-)I_j(t^-)dN_{ij}^I(t) - S_i(t^-)dN_i^V(t), & (5.40)\\ dI_i(t) &= \sum_{j(\neq i)} S_i(t^-)I_j(t^-)dN_{ij}^I(t) - I_i(t^-)dN_i^R(t), & (5.41)\\ & \quad i = 1, \ldots, N. \end{aligned}$$

The collection of the indicators with dynamics given above

$$X^\pi = (X^\pi(t) = (\ldots, S_i(t), I_i(t), \ldots) \in \Sigma, t \geq 0$$

is the càdlàg and adapted controlled spatial epidemic process. By considering bounded functions f I can replicate the argument of section 5.4.1 which gives the

the controlled generator,

$$\begin{aligned}\mathcal{G}^{u^\pi(t)}f &= \sum_i \Big[S_i\lambda_i(f(x+e_i^I)-f(x)) + \gamma I_i(f(x+e_i^R)-f(x)) \\ &\quad + u_i^\pi(t)S_i(t)(f(x+e_i^V)-f(x))\Big]. \qquad (5.42)\end{aligned}$$

5.4.2 The HJB equation for the Spatial Epidemic with Vaccination

In this section I give the HJB equation for the stochastic spatial epidemic process with constrained and costly vaccination. The solution of the HJB equation gives the value for each epidemic state and allows the construction of the optimal vaccine allocation policy as a reaction to knowledge of the current disease state. In principle the solution differs if the controller can intervene up to only some finite time T compared to allowing an unlimited period of intervention, but when T is the cessation time of the epidemic outbreak then I will show that the finite-time horizon control problem is equivalent to the infinite-time horizon problem.

The stochastic construction of the controlled spatial epidemic process given above guarantees that choosing the vaccination distribution $u^\pi(t) \in \mathcal{U}$ for each $t \geq 0$ according to a policy π that can be determined by the current state and time allows policy π to be extended to an admissible control system in the sense of section 5.3.2. The HJB equation for the stochastic spatial epidemic with costs given by (5.27) and admissible vaccination controls (5.39) up until the finite-time horizon T is given by equation (5.14) with controlled generator given by (5.42),

$$\begin{aligned}\frac{\partial}{\partial t}V(t,x) + \sum_i \Big[S_i\lambda_i(V(t,x+e_i^I)-V(t,x)) + \gamma I_i(V(t,x+e_i^R)-V(t,x))\Big] \\ + \sum_i I_i + \min_{u\in\mathcal{U}}\Big\{\kappa(u) + \sum_i u_i S_i(V(t,x+e_i^V)-V(t,x))\Big\} = 0, \qquad (5.43)\end{aligned}$$

$$x \in \Sigma,\ 0 \leq t < T,$$

$$V(T,x) = \psi(x). \qquad (5.44)$$

By theorem 5.3.1 the solution to (5.43) with terminal data (5.44) is the value function for the spatial epidemic with vaccination and the optimal policy π^* is given by choosing a minimiser control decision u^* for the non-linear minimum part of equation (5.43) at each time-state pair $(t, X^*(t))$ arrived upon. Since I consider only finite N-sized meta-populations then directly from the definition (5.27) we see that the value function is bounded for all states.

Control up until Epidemic Cessation

The finite-time horizon I will be most interested in will be the stopping time[3] T^* which is the random first time at which the epidemic ceases with all local disease states being either susceptible or recovered. If the controller has adopted a policy π where the complete depletion of infectious populations implies the cessation of vaccination effort, which I dub a *sensible policy*, then each of these states are absorbing for the dynamics; that is the time spent in the state is infinite. Every non-absorbing disease state is transient since the local disease states are non-repeating for the SIR-type compartmental model, hence by standard finite state space Markov process theory the first arrival time for an absorbing disease-free state is almost surely finite; that is that $T^* < \infty$ almost surely.

Restricting to sensible policies we see that choosing the cessation time T^* as the finite time horizon has the nice property that the value function can be solved as an elliptic form problem (5.22) by allowing the intervention time t to be treated as time 0 without loss of generality. To see this we note that for all T and all sensible policies,

$$\begin{aligned} J(t,x;\pi) &= \mathbb{E}^{\pi}_{t,x}\Big[\int_t^T \sum_i I_i(s) + k(u^{\pi}(s))ds + \psi(X^{\pi}(T))\Big] \\ &= \mathbb{E}^{\pi}_{t,x}\Big[\int_t^{T\wedge T^*} \sum_i I_i(s) + k(u^{\pi}(s))ds + \mathbf{1}_{T<t^*}\psi(X^{\pi}(T))\Big] \end{aligned}$$

Where I have used that no costs are incurred after epidemic cessation for sensible policies and that the exit cost for ceased epidemics is zero. On the other hand we have seen above that taking the limit $T \to \infty$ gives the infinite time non-discounted expected cost of the epidemic given the controller follows policy π. Hence for all sensible policies and all intervention times t control up to cessation is given by,

$$J(x;\pi) = \mathbb{E}^{\pi}_{x}\Big[\int_t^{T^*} \sum_i I_i(s) + k(u^{\pi}(s))ds\Big] = \mathbb{E}^{\pi}_{x}\Big[\int_0^{\infty} \sum_i I_i(s) + k(u^{\pi}(s))ds\Big] \quad (5.45)$$

The infinite time horizon HJB equation with zero discount rate for the spatial epi-

[3] A *stopping time* is a real non-negative random variable T where the event $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.

demic with vaccination is,

$$\sum_i \Big[S_i\lambda_i(V(x+e_i^I)-V(x))+\gamma I_i(V(x+e_i^R)-V(x))\Big]+\sum_i I_i \\ +\min_{u\in\mathcal{U}}\Big\{\kappa(u)+\sum_i u_i S_i(V(x+e_i^V)-V(x))\Big\}=0, \qquad (5.46)$$
$$x\in\Sigma.$$

The particular solution to (5.46) is specified by the boundary conditions that $V(x) = 0$ for absorbing states x. The almost sure finiteness of T^* entails boundedness for expression (5.45) despite the zero discounting rate ($\beta = 0$). Theorem 5.3.2 gives that the solution to the elliptic-type HJB equation is the value function for spatial epidemic controlled up until epidemic cessation and the optimal policy is to choose a minimiser control decision u^* for the non-linear minimum part of equation (5.46) for each state $X^*(t)$ arrived upon.

Optimal Solutions for Control until Cessation

I will now briefly discuss the type of optimal solutions expected from considering the HJB equation (5.46). Without loss of generality I choose time units such that the rate of removal of infectious populations is unity ($\gamma = 1$), in these units each population that becomes infected will contribute an expected cost to the epidemic of unity. With this in mind, I decompose the epidemic costs (5.45) into a contribution for the initially infected populations, a contribution from initially susceptible populations at risk from the epidemic and a contribution from the vaccination campaign. With this decomposition the value function is,

$$V(x)=J(x;\pi^*)=I_0+\sum_{i:S_i(0)=1}\mathbb{P}^*(R_i(\infty)=1|X(0)=x)+\mathbb{E}_x^*\Big[\int_0^\infty k(u^*(s))ds\Big] \qquad (5.47)$$

Where I_0 is the total number of infected populations for state x, and $\mathbb{P}^*$ is the controlled epidemic measure for π^*, the optimal control system given by theorem 5.3.2. Therefore, for each state x where the ith population is susceptible,

$$V(x+e_i^V)-V(x)\leq 0, \qquad (5.48)$$

that is the optimal expected cost of an epidemic initially in the state after vaccination of the ith population ($x+e_i^V$) is less than for the state without the vaccination (x). This is due to the removal of risk of the ith population becoming a case, the reduction of risk of potential secondary cases due to the ith population and reduc-

tion of the expected time over which the vaccination campaign runs.

Theorem 5.3.2 requires that the controller fix his policy such that for each state x arrived upon the control decision is given by,

$$u^* \in \arg\min_{u \in \mathcal{U}} \Big\{ \kappa(u) + \sum_i u_i S_i (V(x + e_i^V) - V(x)) \Big\}. \tag{5.49}$$

If $\kappa(\cdot)$ is a continuous function then it obtains its minimum on the compact space $\mathcal{U}$ given by (5.39). For increasing vaccination rates, condition (5.49) describes a balance between increasing instantaneous cost of deployment, $\kappa(u)$, and the negative signed term, $\sum_i u_i S_i (V(x + e_i^V) - V(x))$, which gives the rate at which the expected future costs are decreased by deploying a removal rate for each population. The ability to correctly balance these rates of cost is the purpose of solving the HJB equation to obtain the value function V.

5.5 Solving the HJB Equation and Optimal Controls

The dynamic programming approach to solving the optimal control problem for spatial epidemics (5.28) gives the optimal control decision as the minimiser of a functional depending on the cost model of vaccine deployment and the value function. Instead of exploring the large space of possible vaccination control policies, one can focus on solving the Hamilton-Jacobi-Bellman equation for the value function whereupon the optimal policy is found in feedback form for any possible state the stochastic epidemic might enter. In this section I present two methods for solving the value function.

The first solution method is derived from considering an equivalent controlled discrete time Markov chain, the value function can be derived using *value iteration*; a classical approach from the theory of Markov decision processes. This method is numerically exact, however the solution requires the construction of a convergent sequence of value approximations for each state $x \in \Sigma$. For systems with a large number of states this can mean that the value iteration technique is effectively insolvable due to the very large dimensionality of Σ.

The second solution method is similar in spirit to value iteration but instead of implementing a sequence of value approximations for each state I restrict to updating a lower dimensional polynomial approximation of the value function. I

perform a number of forward sweep simulations of spatial epidemics using a policy based on optimising for the current belief about epidemic state value. The value function approximation is updated for each forward sweep according to a backwards sweep of state value estimates using a stochastic gradient method. The algorithm I have developed for the Monte Carlo approximation of the value function lies in the class of so-called *approximate dynamic programming* methods (Powell 2007).

As a consequence of the high dimensionality of the state space for the spatial epidemic I will forced to aggregate individual populations into area groupings. For aggregated populations within an area the rates of transmission between populations internal and external to the area are identical. This is equivalent to the population model for plant diseases introduced in chapter 4. In particular the intra-area transmission is treated as zero-range[4] and the inter-area transmission is treated as centroid to centroid range between the areas. The area aggregates define a mesoscopic scale between the individual populations abiding at habitats treated as point locations and the full meta-population. The number of areas is given as N_{area} and the ith area contains N_i populations; the total number of populations is conserved as,

$$\sum_{i=1}^{N_{area}} N_i = N. \tag{5.50}$$

The new, reduced, state space is defined as the possible configurations of numbers of susceptible and infectious populations at each area. The indicators functions $\{S_i(x), I_i(x)\}_{i=1}^{N_{area}}$ now return the epidemic numbers for each area.

5.5.1 Value Iteration for the Embedded Markov Chain

It was noted in chapter 4 when describing the Gillespie algorithm for simulating continuous time finite state Markov processes that for each state x the probability that the next state arrived at is y is given by $r(x, y))/\xi(x)$ where $r(x, y)$ is the probabilistic rate of the one step transition $x \to y$ and $\xi(x)$ is the exit rate of the state: $\xi(x) = \sum_{y(\neq x)} r(x, y)$. The expected time spent in state x conditioned on arrival at some time, also dubbed *sojourn* time of x, $\Delta t(x)$, is exponentially distributed with mean $\mathbb{E}[\Delta t(x)] = \xi^{-1}(x)$. From this point of view an epidemic

[4]This eliminates the possibility of using a Bessel function transmission kernels with a degeneracy at zero.

path can be reconstructed from the ordered sequence of state-sojurn times, $\hat{X}$:

$$\hat{X} = \Big((x_0, \Delta t(x_0)) \to (x_1, \Delta t(x_1)) \to \dots \Big| x_n \in \Sigma, \Delta t(x_n) \in [0, \infty)\Big).$$

via $X(0) = x_0$ and whenever $\sum_{k=0}^{n-1} \Delta t(x_k) \leq t < \sum_{k=0}^{n} \Delta t(x_k)$ then $X(t) = x_n$. The cost incurred by an epidemic realisation X can be calculated by solving the cost integral (5.3) directly,

$$\int_0^\infty L(X(s))ds = \sum_{k\geq 0} L(x_k)\Delta t(x_k). \tag{5.51}$$

Conditioned on the state the sojourn times are independent therefore the expected cost of the epidemic given a policy π is

$$\mathbb{E}_x^\pi\Big[\int_0^\infty L(X(s))ds\Big] = \hat{\mathbb{E}}_x^\pi\Big[\sum_{k\geq 0} L(x_k)\xi^{-1}(x_k)\Big]. \tag{5.52}$$

Where $\hat{\mathbb{E}}$ is expectation with respect to the probability distribution of the *embedded Markov Chain* of spatial epidemic; that is the Markov chain on Σ with initial distribution $\mathbb{P}_0$ and transition probabilities defined by one-step transitions of the continuous time spatial epidemic.

The discrete time control problem of choosing a policy π to minimise (5.52) for the embedded Markov chain of the spatial epidemic is an example of a Markov decision process (MDP)(Puterman 2005). The value function for this MDP problem is exactly that for the continuous time control problem and can be solved using the discrete time analogue of dynamic programming. The *Bellman*[5] equation for the value function is similar in form to the HJB equation given above and can be found from Fleming and Soner Chapter 9 (Fleming and Soner 2006) or Bertsekas Chapter 5 (Bertsekas 1987),

$$V(x) = \min_{u\in\mathcal{U}} \Big\{\hat{L}(x,u) + \sum_{y(\neq x)} p^u(x,y)V(y)\Big\} \qquad x \in \Sigma. \tag{5.53}$$

Where the state and control dependent costs are,

$$\hat{L}(x,u) = \begin{cases} L(x,u)/\xi(x,u) & \xi(x,u) > 0 \\ 0 & \xi(x,u) = 0 \end{cases} \tag{5.54}$$

[5]It is more common in the MDP literature to neglect the reference to Hamilton and Jacobi.

and $p^u(x, y)$ is the one step probabilities for the embedded Markov chain of the spatial epidemic process which depend on the control decisions. Note that equation (5.54) now depends on the control decision through the Lagrangian of the continuous time control problem and through varying the exit rate of the state x.

The solution to the Bellman equation (5.53) is identical state-by-state to the solution of the HJB equation (5.46). Therefore solving the discrete time problem will deliver the solution to the continuous time control problem. Fortunately, the Bellman equation (5.53) is solvable using an iterative method called *value iteration*. Writing the right hand side of (5.53) as an operator, $F(\cdot)$, gives the Bellman equation in compact form, $V(x) = F(V)(x)$ for $x \in \Sigma$.

I define an iterative sequence of value function approximations

$$W^{(n)} = F(W^{(n-1)}), \qquad n \geq 0. \tag{5.55}$$

Fleming and Soner (Fleming and Soner 2006) give that $V = F(V)$ is the unique attractive fixed point of this sequence and that the operator $F(\cdot)$ is monotone non-decreasing in the sup-norm. Choosing an initial guess $W^{(0)} = 0$ for the value function gives an increasing sequence of value function approximations converging to V. In algorithmic form the solution method is:

Value Iteration Algorithm:

1. For the continuous time control problem calculate the transition probabilities $p(\cdot, \cdot)$ and state-action costs $\hat{L}(\cdot, \cdot)$ for the equivalent MDP. Define a uniform error tolerance $\epsilon > 0$.
2. Set the initial value estimate $W^{(0)} = 0$.
3. Derive next estimate by solving for each $x \in \Sigma$,
$$W^{(n+1)}(x) = \min_{u \in \mathcal{U}} \Big\{ \hat{L}(x, u) + \sum_{y(\neq x)} p^u(x, y) W^{(n)}(y) \Big\}.$$
4. If the uniform error $\|W^{(n+1)} - W^{(n)}\|_\infty < \epsilon$ then return value function $V = W^{(n+1)}$. Else return to 3.

5.5.2 Approximate Dynamic Programming

Whilst conceptually simple the value iteration approach to solving the value function runs foul of the infamous 'curse of dimensionality'; a general term referring to

the exponential (or faster) explosion in the number of configurations that must be considered in order to solve some class of problems, parameterised by a size N, as the size $N \to \infty$. As a concrete example, the value iteration method utilised above relies directly on the ability to loop over the entire state space. If the state space is too large to enumerate then looping over states is also clearly unfeasible. For this case it is customary to turn to Monte Carlo methods which avoid solving the HJB equation for all states simultaneously in favour of sampling states according to some random scheme and updating a sequence of approximations.

In this section I will describe a novel algorithm for approximating the value of states for the spatial epidemic and thereby constructing an approximate optimal policy based on applying (5.16) to the approximate value function. The algorithm involves sampling states according to their visitation by stochastic simulations going forward in time and updating the value approximation using a cost process assessed retrospectively backwards in time along the path of the forward simulation. The final ingredient is updating the policy used for the forward simulations in concordance with each update of the value function. This will effectively be a stochastic form of value iteration, or an approximate dynamical programming (ADP) method and I call the algorithm the ADP method for solving the optimal vaccination problem for the spatial epidemic; unfortunately I have to date been unable to rigourously establish the convergence properties of the solution method. In the following I establish the plausibility of the ADP method and subsequently I will give numerical examples that demonstrate that it is as effective as the exact solution method at least for sufficiently 'small' epidemic models that either method can be applied.

Recursive Cost Estimation

I will first briefly cover an approach to estimating the cost of states for some fixed policy π using recursive Monte Carlo simulation. Although in a sense intuitive I think the exposition clarifies where the ADP method progresses from estimating fixed policy cost to approximating the optimal value of states. Without the requirement to optimise over the space of admissible policies the general expected cost function for following policy π up until epidemic cessation is given by,

$$\mathcal{G}^{u^\pi} J(x, u^\pi(x); \pi) + L(x, u^\pi(x)) = 0, \qquad x \in \Sigma. \tag{5.56}$$

If the state space for the process is sufficiently small and the generator is a difference operator then equation (5.56) is an affine matrix equation that can solved directly

using linear algebra. This relation between general cost and equation (5.56) has been noted in the literature of stochastic epidemics (J. V. Ross, House, and Keeling 2010) where it was interpreted as an efficient method of calculating epidemiological quantities of interest that can be represented as expectations of integrals along the path of the epidemic X.

However, if equation (5.56) is the object of interest and the state space is effectively innumerable then the chain of inference can be reversed and $J(x;\pi)$ estimated from stochastic simulations of X. One method for achieving this is to choose an initial distribution $\mathbb{P}_0$ and independently draw a sample sequence of M initial states $\{\hat{x}^{(m)} \in \Sigma\}_{m=1}^{M}$. For each of these sampled initial states a sample path for the process X can be drawn via stochastic simulation,

$$\left\{\hat{X}^{\pi,m} = (\hat{X}^{\pi,m}(t), 0 \leq t \leq T, \hat{X}^{\pi,m}(0) = \hat{x}^{(m)})\right\}_{m=1}^{M}. \tag{5.57}$$

The expected cost of originating in each intervention time-state pair visited by the sample path $\hat{X}^{\pi,m}$ is given by the conditional expectation,

$$J(\hat{X}^{\pi,m}(t);\pi) = \mathbb{E}\Big[\psi(X^{\pi}(T)) + \int_t^T L(X^{\pi}(s), u^{\pi}(s))ds\Big|\mathcal{F}_t\Big], \qquad 0 \leq t \leq T. \tag{5.58}$$

This expected cost can estimated using a cost processes evolving backwards in time from the final cost condition. For each of the M sampled paths of X there is an associated cost process $\hat{Y}^{\pi,m} = (\hat{Y}^{\pi,m}(t), 0 \leq t \leq T)$ given by,

$$\hat{Y}^{\pi,m}(t) = \psi(\hat{X}^{\pi,m}(T)) + \int_t^T L(\hat{X}^{\pi,m}(s), u^{\pi}(s))ds, \qquad 0 \leq t \leq T. \tag{5.59}$$

For control up until epidemic cessation ($T = T^*$) the mth sample cost process $\hat{Y}^{\pi,m}$ gives an unbiased sample estimate for the cost under policy π for each state visited by $\hat{X}^{\pi,m}$. If one is in particular interested in the cost associated with some state $x \in \Sigma$ one could choose the initial distribution to be a delta mass at x: $\mathbb{P}_0 = \delta_x$ and then each $\{\hat{Y}^{\pi,m}(0)\}_{m=1}^{M}$ is an unbiased sample estimate of the cost associated with the state x and the policy π. It is then possible to recursively improve one's estimate for $J(x;\pi)$ using a *stochastic approximation*[6] method (Robbins and Monro 1951; Kushner and Yin 2003). I denote the estimate of the cost of state x after m sample estimates by $\bar{J}^{(m)}$ and the estimate is recursively updated using the

[6]Often referred to as a stochastic gradient method e.g. (Powell 2007)

stochastic approximation rule,

$$\bar{J}^{(m)} = \bar{J}^{(m-1)} - \alpha_m(\bar{J}^{(m-1)} - \hat{Y}^{\pi,m}(0)). \tag{5.60}$$

Choosing the *step-size rule* $\{\alpha_m\}_{m\geq 1}$ selects different stochastic approximations. It is perfectly permissible to pick a random step-size rule that changes according the random estimates being recursively considered although general necessary conditions (c.f. (Kushner and Yin 2003) Chapter 5) for the convergence of the recursive approximation scheme (5.60) are:

- $\alpha_m \geq 0$ for $m = 1, 2, 3, \ldots$ almost surely.
- $\sum_{m\geq 0} \alpha_m = \infty$ almost surely.
- $\sum_{m\geq 0} \alpha_m^2 < \infty$ almost surely.

For independently and identically distributed unbiased sample estimates it is well known that the optimal step-size rule, in the sense of providing an unbiased estimate of the cost with minimal estimate variance, is $\alpha_m = 1/m$ for each m, c.f (Powell 2007) Chapter 6. This simple deterministic step-size rule is just a recursive method for calculating the very familiar empirical mean,

$$\bar{J}^{(M)} = \frac{1}{M}\sum_{m=1}^{M} \hat{Y}^{\pi,m}(0). \tag{5.61}$$

As $M \to \infty$ we have $\bar{J}^{(M)} \to J(x;\pi)$ in probability as required.

An Approximate Dynamic Programming Algorithm for the Spatial Epidemic

The simple methodology described above use above essentially describes the method I intend to use to estimate the value (not the cost!) of states. The modifications I use are:

- The value function V will be approximated by a polynomial function on the space of local epidemic numbers for each state $x \in \Sigma$: $\{S_i(x), I_i(x)\}_{i=1}^{M}$. The estimates for the coefficients of the polynomial approximation will be updated for each state visited by the forward simulation sample $\hat{X}$ using the sample cost process $\hat{Y}$. The polynomial value approximation after M recursive estimate updates will be denoted, $\bar{V}^{(M)}$.

- No particular policy π will be followed, rather for the mth recursion the approximate optimal policy will be $\bar{\pi}^{(m)}$ given by following the optimal strategy given by (5.49) using the current best approximation of the value function $\bar{V}^{(m-1)}$. That is for each state visited $\hat{X}^{\pi,m}(t)$ the decision will be given by,

$$\bar{u}^{(m-1)}(t) \in \arg\min_{u\in\mathcal{U}} \Big\{ \kappa(u) + \sum_i u_i \delta_{S_i(t),0} (\bar{V}^{(m-1)}(\hat{X}^{\pi,m}(t) + e_i^V) - \bar{V}^{(m-1)}(\hat{X}^{\pi,m}(t))) \Big\}. \quad (5.62)$$

  Where the Kronecker delta function enforces that vaccinating in an area with no susceptibles has no effect.

- I will use a stochastic approximation method to estimate V from the sample cost process $\hat{Y}^{\bar{\pi}^{(m-1)},m}$. However due to the policy potentially altering at each recursion the sample estimates will not be identically distributed and a different step-size rule from the standard $\alpha_m = 1/m$ rule should be used. Moreover, the early sample estimates $\hat{Y}^{\bar{\pi}^{(m-1)},m}$ will be poor sample estimates for the value function due to using a (probably) very sub-optimal approximations for the optimal decisions when simulating forwards. The stochastic approximation method I choose puts much less weight on the early sample estimates.

From a recursive estimation point of view the value iteration scheme for solving the optimal control problem introduced above was in the form,

$$\bar{V}^{(m+1)}(x) = \min_{u_1\in\mathcal{U}} \Big\{ L(x,u_1) + \mathbb{E}^{u_1,x}_{1step}[\bar{V}^{(m)}] \Big\}. \quad (5.63)$$

Where $\mathbb{E}^{u_1,x}_{1step}[f] = \sum_{y(\neq x)} p^{u_1}(x,y) f(y)$ is the next step from state x expectation for the embedded Markov chain given that the control action u_1 is performed. The next step expectation estimates the future costs for starting in state x using the current estimate of the value function $\bar{V}$. The stochastic updating rule is very simple, the last 'sample' provides the current estimate. Value iteration demonstrates that recursively optimising and updating the value estimate for all states converges onto the true value function.

The ADP algorithm I have developed for the spatial epidemic uses as its sample estimates of future cost the sample process $\hat{Y}$ where the forward sample simulations $\hat{X}$ have been optimised according the current value approximation. However, not all states will be visited for updating instead a global polynomial approximation

to the value function is updated. The idea is that in a limit with sufficiently disperse states updated using a very large number of recursions and for a sufficiently high dimensional polynomial approximation the value approximation will be very 'close', in some sense, to the true value function for each epidemic state. Unfortunately, I have not managed to rigorously establish this for the algorithm I have developed although for certain related algorithms such as the Longstaff-Schwartz method for pricing Bermudan options[7] in discrete time (Longstaff and Schwartz 2001; Egloff 2005) such a proof is available.

Polynomial approximation for the value function: With aggregated collections of populations in mind I choose the following polynomial approximation for the value function of the spatial epidemic,

$$V(x) \approx \sum_{i,j=1}^{N_{area}} \sum_{k=1}^{K} \sum_{l=1}^{L} C_{ijkl} \Big(\frac{S_i(x)}{N_i} \Big)^k R_{ij} I_j(x) \Big(\frac{I_j(x)}{N_j} \Big)^{l-1}. \tag{5.64}$$

Where R_{ij} is the expected number of infectious contacts made by a single population in area j to any population in area i[8]. If the respective centroids of areas i, j are $x_i, x_j \in \mathcal{A}$ then we can define,

$$R_{ij} = \frac{N_i K(x_i - x_j)}{\gamma}. \tag{5.65}$$

In the sequel I will typically choose $K(\cdot)$ so that $R_{ij} \sim \mathcal{O}(1)$.

The coefficients $\{C_{ijkl} \in \mathbb{R}\}$ are the targets for estimation. The motivation for the approximation choice (5.64) is therefore quite simple; it is an attempt to fit a multiplier to each currently observed infected population which takes into account their expected cost and the expected future infectious events due their infectious status using the transmission parameters and a hierarchy of non-linear dependence on the current population densities of susceptible and infecteds.

Stochastic Approximation Scheme: For each state visited each polynomial coefficient is updated according to a gradient descent stochastic approximation rule. Writing the estimate of the coefficient set $\{C_{ijkl}\}$ after m recursions as a vector

[7] Bermudan options are characterised by the owners right to excise at set of discrete time points. The optimal control problem is then of a stopping variety, picking the discrete excise time which maximises expected return.

[8] Specifically, the expected number of hitting times for the underlying Poisson drivers $N^I_{i'j'}$ for any population i' contained in area i during the infectious period of population j' in area j.

$\bar{\boldsymbol{C}}^{(m)}$ the estimate update for arriving at the forward sampled state at time t, $\hat{X}(t)$ is,

$$\bar{\boldsymbol{C}}^{(m)} = \bar{\boldsymbol{C}}^{(m-1)} - \alpha_m(\bar{V}^{(m-1)} - \hat{Y}^{(m)}(t))\nabla_{\boldsymbol{C}}\bar{V}^{(m-1)}. \tag{5.66}$$

Where the value estimate at each recursion is set by the parameters. Equation (5.66) can be viewed as an iterative scheme for finding the minimiser of the error measure $\frac{1}{2}(V - \bar{V}(\boldsymbol{C}))^2$, c.f. (Kushner and Yin 2003) Chapter 1.

For the early recursions of the ADP scheme the sample costs used to update the value estimates are likely to be much larger than the true value of the states visited by the forward sample simulations due to sub-optimal decisions being made. These early recursions are important for exploring the space of value function approximations, but should not be given great statistical weight for estimating the true value of the states visited. On the other hand later recursions should have approximately equal proportions of sample costs being higher or lower than the current value estimate, if the scheme is successful and the later value estimates are close to the true values of the states visited. For statistical convergence it is important to incorporate the later sample costs into the value estimate. Taking this into account I chose a stochastic step-size rule known as *Kesten's rule* (Kesten 1958). Kesten's idea was very simple, the step-size α_m should only decrease when the sign of the $(\bar{V}^{(m-1)} - \hat{Y}^{(m)})$ oscillates over the previous two recursions,

$$\alpha_m = \frac{a}{(a + \lambda^{(m)} - 1)N_{loc}}. \tag{5.67}$$

Where $a \geq 0$ is a tuning parameter for the step-size and $\lambda^{(m)}$ counts the number of sign oscillations up until the mth recursion,

$$\lambda^{(m)} = m, \qquad m = 1, 2, \tag{5.68}$$

$$\lambda^{(m)} = \lambda^{(m-1)} + \frac{1}{2}(1 - \mathrm{sgn}[\bar{V}^{(m-1)} - \hat{Y}^{(m)}]\mathrm{sgn}[\bar{V}^{(m-2)} - \hat{Y}^{(m-1)}]), \ m > 2. \tag{5.69}$$

I have introduced the scaling $1/N_{loc}$ to the step-size rule where $N_{loc} = \langle N_i \rangle$ is the mean number of populations per area. The reason for that scaling was to balance the gradient term $\nabla_{\boldsymbol{C}}\bar{V}^{(m-1)}$ which is $\mathcal{O}(N_{loc})$.

Having given a largely heuristic justification for the approximate dynamic programming approach I now make explicit the algorithm I have developed to approximately solve the optimal control problem for the spatial epidemic with vaccination up until epidemic cessation. It is much more numerically efficient to use a time-stepping approximate stochastic simulation procedure, the tau-leap method described in chapter 4, but it is not necessary.

ADP algorithm for the Stochastic Epidemic:

1. Decide upon underlying optimal control problem for spatial epidemic with epidemiological parameters and costs. If a time-stepping stochastic simulation method is used then fix time step δt.

2. Choose maximum power of $\{S_i\}$ (K), maximum power of $\{I_i\}$ (L) for polynomial fit to value function. Choose Kesten tuning parameter (a). Choose the number of recursive steps (M) for fitting the value function approximation (5.64). Choose initial distribution of states for fitting recursions $\mathbb{P}_0$. Initialise the fitting coefficients $\boldsymbol{C}^{(0)} = 0$, i.e. $\bar{V}^{(0)} = 0$.

3. For $m = 1, \ldots, M$,

   (a) Draw an initial state $\hat{x}_0^{(m)}$ from $\mathbb{P}_0$.

   (b) Use stochastic simulation method to construct a sample epidemic path in state, sojourn time and control decision form. At each time step solve (5.62) to decide upon optimal control given current value function estimate $\bar{V}^{(m-1)}$,

   $$\hat{X}^{(m)} = \Big((\hat{x}_0^{(m)}, \Delta\hat{t}_0^{(m)}, \hat{u}_0^{(m)}) \to (\hat{x}_1^{(m)}, \Delta\hat{t}_0^{(m)}, \hat{u}_1^{(m)}) \to \cdots \to (\hat{x}_F^{(m)}, \infty, 0) \Big).$$

   Where x_F is the first disease free state located.

   (c) Solve for cost process $\hat{Y}^{(m)}$ backwards along $\hat{X}^{(m)}$. Denoting $\hat{y}_k^{(m)} = \hat{Y}^{(m)}(t)$ for $t = \inf\{s | \hat{X}^{(m)}(s) = \hat{x}_k^{(m)}\}$. Solution is given recursively,

   $$\begin{aligned} \hat{y}_F^{(m)} &= 0, \\ \hat{y}_k^{(m)} &= L(\hat{x}_k^{(m)}, \hat{u}_k^{(m)}) \Delta\hat{t}_k^{(m)} + \hat{y}_{k+1}^{(m)}, \qquad 0 \le k \le F. \end{aligned}$$

   (d) Update coefficient vector $\boldsymbol{C}$ for each state visited using the stochastic gradient method with Kesten rule step-sizes,

   $$\bar{\boldsymbol{C}}^{(m-1)} \to \bar{\boldsymbol{C}}^{(m-1)} - \alpha_m (\bar{V}^{(m-1)}(\hat{x}_k^{(m)}) - \hat{y}_k^{(m)}) \nabla_{\boldsymbol{C}} \bar{V}^{(m-1)}(\hat{x}_k^{(m)}), 0 \le k \le F.$$

Note that the coefficients are updated $F+1$ times for each sample path $\hat{X}^{(m)}$.

(e) Update $\bar{\boldsymbol{C}}^{(m)} = \bar{\boldsymbol{C}}^{(m-1)}$.

4. Return final coefficient vector $\bar{\boldsymbol{C}}^{(M)}$ and hence estimated value function $\bar{V}^{(M)}$.

5.6 Numerical Examples

In this section I will give a numerical investigation into the optimal vaccination control of an stochastic epidemic invasion with SIR-type dynamics into a spatial metapopulation stratified into areas as described in section 5.5. Throughout I will present results for two vaccination scenarios; a scenario where vaccine deployment rate cost per unit time is cheap compared to the expected cost of each infected population (10%) but the maximum rate of deployment is low and the alternative where both the maximum constraint on deployment but also the cost per unit time of deployment are a multiple of five greater. I will demonstrate the theoretical maximum expected cost savings dues to using an optimally controlled vaccination policy and give the alteration to the expected dynamics due to optimal control. I will also investigate some cost issues associated with lacking full information about the current epidemic state; namely the cost associated with random late discovery of epidemic invasion and inefficient population targeting of vaccination resources. Since uniform deployment of vaccination effort is a natural default strategy I will compare this strategy's performance versus the optimal, and also how the extra saving due to optimal control varies as the symmetry of the transmission parameters is broken.

Having considered a model sufficiently small as to be solvable using direct value iteration I shall turn my attention to the ADP method for approximating the value function for the spatial epidemic process. I compare the time-varying expected control processes for both the exactly solved and approximately controlled stochastic epidemic and demonstrate that the ADP method is close to being a cost minimiser. I progress to a numerical study on the expected costs associated with different ADP algorithm parameterisations and finally give an example of using ADP to approximately optimally target a vaccination campaign against an aggressively invasive pathogen in the setting of a meta-population aggregated into multiple area with large area populations.

5.6.1 A Two Area Numerically Solvable Control Problem

The two area population aggregation model has received attention over the last decade due to being the simplest model within which it is possible to test hypothesises about the nature of spatial coupling between population groupings (Keeling and Rohani 2002). There has also been some recent interest in solving optimal control problems for epidemic spread between two coupled population aggregates (Rowthorn and et al 2009; Klepac, Laxminarayan, and Grenfell 2011). In both these examples the disease dynamics were assumed to be deterministic with optimal controls found by either an application of Pontryagin's principle or numerical investigation. Here I give the basic model a stochastic formulation and use value iteration to solve the optimal vaccination delivery problem.

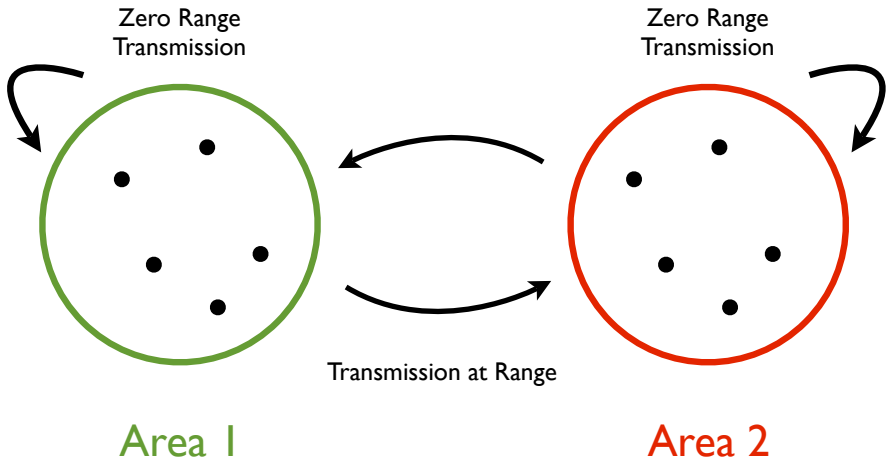


Figure 5.1: A schematic of the basic control problem. The habitats are dispersed into two geographically distinct areas with intra-area transmission rate treated as zero range and inter-area transmission at some lower rate. The optimal control problem is to set the vaccination rate for the abiding populations in such a manner as to minimised the total cost of the epidemic up until cessation.

Model Definition, HJB Equation and Possible Optimal Vaccination Rates

I now define a two area model which is sufficiently small that I can numerically solve the optimal control problem using value iteration. I give the stochastic dynamics, present the HJB equation and heavily restrict the space of possible optimal strategies which makes the numerical solution more efficient.

I consider $N = 100$ populations aggregated into two spatial areas with local populations $N_1 = N_2 = 50$. Transmission parameters are considered to be identical between each population sharing an area with intra-area transmission given by the zero-range rate $K(0)$ and inter-area transmission given by $K(r)$ where r is the Euclidean distance between the two centroids of the areas, figure 5.1 gives a schematic of the basic set-up. Since there are only two transmission rates due to symmetry I write $K(0) = \beta_{in}$ and $K(r) = \beta_{out}$. Throughout this work I have left the mechanistic origins of the transmission kernel deliberately vague since many different microscopic models of disease transmission pathways such as mixing due to commuting for Human diseases or air-bourne spore dispersal for certain plant diseases can lead to effectively the same spatial coupling models, c.f. (Keeling and Rohani 2002) for a more detailed discussion. In this case I do not want to consider an amplification effect simply due to area aggregation size and re-scale the transmission from population to population by $1/N_{loc}$ where $N_{loc} = \langle N_i \rangle$ is the mean size of the local aggregation sizes. The removal rate sets the time scale of the process; that is a single time unit is the expected generation time $1/\gamma$, and I set $\gamma = 1$. All times used in this section are in these units.

As with the models given in chapter 4, I treat the removal rate as due to the endogenous efforts of embedded local disease management resource at each population. The additional vaccination campaign I consider to be an exogenous mobile resource with flexible scheduling and the ability to react rapidly to real-time information about disease progression. It operates until all infectious populations are removed from the meta-population; that is I consider control until epidemic cessation throughout the following examples. The space of admissible controls is,

$$\mathcal{U} = \{(u_1, u_2) \in [0, \infty)^2 \mid u_1 + u_2 \leq u_{max} N_{loc}\}. \tag{5.70}$$

Where the maximum vaccination rate is scaled by the mean size of the population aggregates. With this scaling I can define u_{max} via the expected time required to completely vaccinate a mean sized area given maximum resource, $\langle T_{vac} \rangle$,

$$u_{max} = \frac{1}{\langle T_{vac} \rangle}. \tag{5.71}$$

I model the cost associated with the mobile vaccination resource as growing linearly with the rate of vaccination deployment. The cost of deploying at unit rate for a unit interval being a cost constant κ. This specifies the general cost function given

previously to,

$$\kappa(u) = \kappa \sum_i u_i, \qquad \kappa \in [0, \infty). \tag{5.72}$$

As with the plant model in chapter 4 I give the stochastic dynamics of the aggregate populations at each area with $\{S_i, I_i\}_{i=1}^{N_{area}}$ now indicating total numbers of susceptible and infectious populations in each area. The SDE formulation for the dynamics with some admissible sensible[9] control policy π is,

$$dS_i(t) = -\Big[\frac{S_i(t^-)}{N_{loc}}\Big(\beta_{in} I_i(t^-) + \beta_{out} I_j(t^-)\Big) + u_i^{\pi}(t^-)\Big]dt + dM_{S_i}(t), \tag{5.73}$$

$$dI_i(t) = \Big[\frac{S_i(t^-)}{N_{loc}}\Big(\beta_{in} I_i(t^-) + \beta_{out} I_j(t^-)\Big) - \gamma I_i(t^-)\Big]dt + dM_{I_i}(t). \tag{5.74}$$

$$i, j = 1, 2, i \neq j.$$

Where $\{M_{S_i}, M_{I_i}\}$ are martingale processes that can either be constructed from Poisson driver representation of the epidemic or defined via the martingale problem representation of the epidemic. It is worth noting the implicit assumption that because the infectious status of each population is known in addition to the aggregate population state that each vaccination campaign is targeted effectively. Therefore, the rate of removing susceptible populations via vaccination is independent of the current disease state of the area being targeted, so long as there are any susceptible populations left in the area. The target cost functional to be minimised is,

$$J(x; \pi) = \mathbb{E}_x\Big[\int_0^{T^*} \sum_{i=1}^{2} (I_i(t) + \kappa u_i^{\pi}(t))dt\Big]. \tag{5.75}$$

In time units where $\gamma = 1$ the expected cost per infected is 1 which defines the cost unit, similarly κ defines the expected cost of the effort required to deploy one vaccine.

I can redefine the elementary event vectors to be length N_{area} rather than N; the HJB equation for the value function V of this two area model is given by

[9] Recall the criteria for sensible policies included not vaccinating any non-susceptible population.

(5.46) which specialises to,

$$\sum_{i=1}^{2}\Big[\frac{S_i}{N_{loc}}(\beta_{in}I_i+\beta_{out}I_j)(V(x+e_i^I)-V(x))+\gamma I_i(V(x+e_i^R)-V(x))\Big]$$
$$+\sum_{i=1}^{2}I_i+\min_{u\in\mathcal{U}}\Big\{\sum_{i=1}^{2}u_i(\kappa+\delta_{S_i,0}(V(x+e_i^V)-V(x))\Big\}=0, \quad (5.76)$$
$$i\neq j, x\in\Sigma.$$

Where the Kronecker delta function enforces that areas with no susceptible populations receive no benefit from vaccination effort. Theorem 5.3.2 gives that the optimal policy π^* is for each state $X^*(t)$ visited by the stochastic epidemic is to choose the vaccination rate given by,

$$u^*(t)\in\arg\min_{u\in\mathcal{U}}\Big\{\sum_{i=1}^{2}u_i\Big(\kappa+\delta_{S_i,0}(V(X^*(t)+e_i^V)-V(X^*(t)))\Big)\Big\}. \quad (5.77)$$

Equation (5.77) immediately restricts the possible optimal actions for each state arrived at by X^* to three options. The optimal choice from these three options depending on the value of $X^*(t)$ and the cost per unit rate of vaccine deployment $\kappa>0$.

Possible Optimal Vaccination Rates: There are three scenarios:

1. If

$$\kappa>\delta_{S_1,0}(V(X^*(t))-V(X^*(t)+e_1^V))$$

AND

$$\kappa>\delta_{S_2,0}(V(X^*(t))-V(X^*(t)+e_2^V)).$$

The vaccination cost out-weighs the vaccination benefit for both areas for state $X^*(t)$. It is optimal not to vaccinate after the epidemic enters $X^*(t)$.

2. If 1. is not true AND

$$\delta_{S_1,0}(V(X^*(t))-V(X^*(t)+e_1^V))>\delta_{S_2,0}(V(X^*(t))-V(X^*(t)+e_2^V)).$$

It is cost effective to vaccinate in at least one area and the area delivering most benefit per vaccination is area 1. It is optimal to vaccinate area 1 with

maximum effort; after epidemic enters $X^*(t)$ the optimal vaccination rate is

$$u^*(t) = (u_{max}N_{loc}, 0).$$

3. If 1. is not true AND

$$\delta_{S_2,0}(V(X^*(t)) - V(X^*(t) + e_2^V)) > \delta_{S_1,0}(V(X^*(t)) - V(X^*(t) + e_1^V)).$$

It is cost effective to vaccinate in at least one area and the area delivering most benefit per vaccination is area 2. It is optimal to vaccinate area 2 with *maximum* effort; after epidemic enters $X^*(t)$ the optimal vaccination rate is

$$u^*(t) = (0, u_{max}N_{loc}).$$

Optimal Trajectories for the Two Area Model

The expected costs for the two area epidemic model without vaccination, $J(\cdot)$ can be found either by using a matrix solution or by using value iteration with no minimisation step. As one would expect $J(\cdot)$ is an increasing function in the number of infecteds and susceptibles. This allowed me to calculate the benefit of an optimally targeted vaccination campaign compared to no intervention for each state $x \in \Sigma$ via $J(x) - V(x)$. The value function $V(\cdot)$ was calculated using value iteration. This took a couple of minutes per solution due to the reduction to three of the number of possibly optimal actions to consider at each iteration, even with the error tolerance set to numerically zero.

The relative cost saving of optimal control does not vary monotonically in the local disease state indicators. As a concrete example, for transmission parameters $\beta_{in} = 1.4$ and $\beta_{out} = 0.1$ with a relatively expensive vaccine $\kappa = 0.5$ that can be deployed sufficiently quickly that a complete area can be vaccinated in 4 time units ($u_{max} = 0.25$), the peak benefit due to optimal vaccination control occurred when the initial disease state was 7 infectious populations isolated in one area. This maximum benefit was found to be 9.726 cost units (figure 5.2). Less infecteds meant that the chance of early stochastic fade-out, which is likely to be only weakly effected by the vaccination campaign, was higher and hence the benefit of vaccination less meaningful. On the other hand for greater numbers of infecteds the number of remaining susceptible populations is lower and the recruitment rate of new infecteds is higher ergo the vaccination effort is less effective (figure 5.2). This emphasises that

vaccination as a mechanism for controlling an invasive pathogen is best deployed as early as possible in an epidemic.

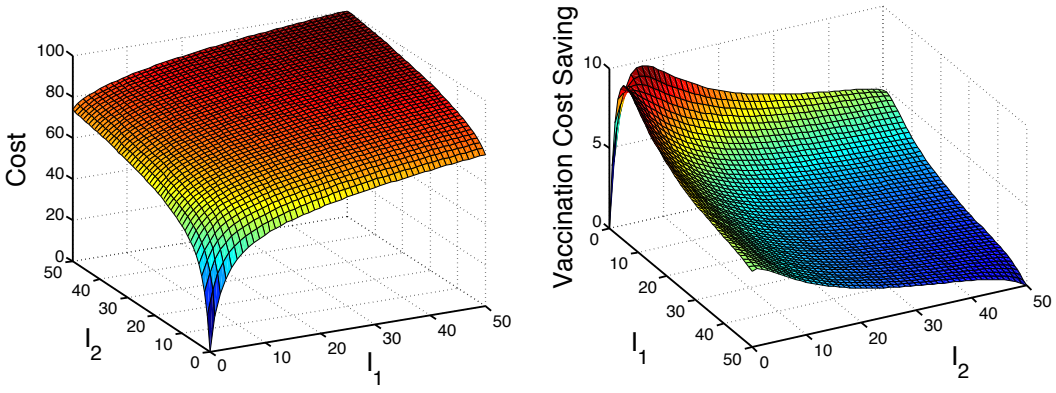


Figure 5.2: Expected costs for an uncontrolled epidemic and benefit analysis of a vaccination campaign. *Left:* Expected cost of epidemic for uncontrolled case for initial numbers of infected populations (I_1, I_2). There were no initially removed populations. Epidemiological parameters were $\beta_{in} = 1.4$ and $\beta_{out} = 0.1$, $\gamma = 1$. *Right:* Cost saved by a vaccination campaign with unit cost of vaccine deployment $\kappa = 0.5$ and maximum deployment rate $u_{max} = 0.25$. Maximum benefit was accrued when $I_1 = 7$, $I_2 = 0$ and vice versa. The maximal benefit 9.726 cost units (equivalent to avoiding infection for an average of 9.726 populations). The surface plot is rotated for easier viewing of the peaks.

I also considered a cheaper vaccine ($\kappa = 0.1$) but where the maximum deployment rate was correspondingly less ($u_{max} = 0.05$). The optimally controlled dynamics for both the slow and faster vaccination controls were characterised by spending some time focusing on the area containing the initially infected populations and thereby reducing the initial growth of the epidemic before then spending considerable time targeting the other area, which contained a great number of initially susceptible populations. Although due to the stochasticity of the process there were occasional switch backs in reaction to infection events (figure 5.3, top row). Intriguingly, the control pressure was exerted to increase the similarity between the infection curves rather than other plausible strategies such as increasing protection for areas with a lower infectious burden as a prophylactic measure. This can be seen in figure 5.3 for an epidemic scenario with greater mixing between the areas ($\beta_{in} = 1.5$, $\beta_{out} = 0.5$) seeded by 5 initial infecteds in area 1. For insight into the switching focus of the optimal vaccine control I calculated the time-varying empirical probabilities of each of the three possible optimal actions: no vaccination, vaccinate area 1 or vaccinate area 2. The idea is to investigate the trend of the

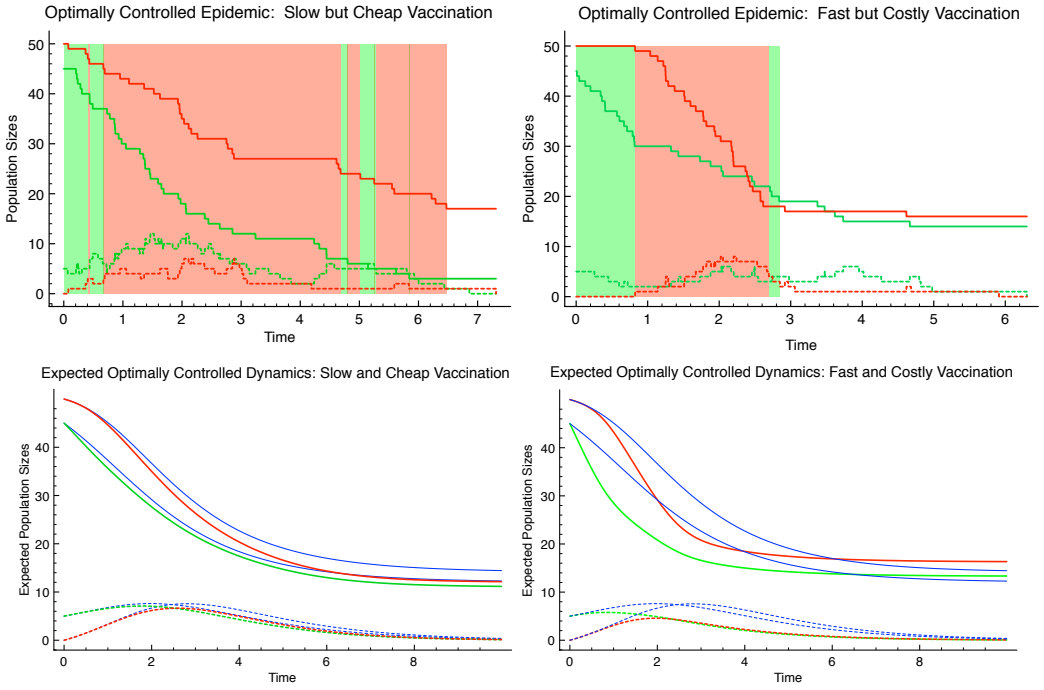


Figure 5.3: Two samples of an optimally controlled epidemic. Epidemiological parameters were $\beta_{in} = 1.5$, $\beta_{out} = 0.5$, $\gamma = 1$. The epidemic seed was chosen to be 5 initially infected in area 1. Colour coding is green for area 1, red for area 2. Solid lines give susceptible numbers, dashed line give infected numbers. Background shading indicates that the optimal vaccination policy was to deploy at maximum rate at the area of indicated colour. No background shading indicates that it was optimal not to deploy vaccine. *Top Left:* Cheap vaccine deployment ($\kappa = 0.1$) but where the maximum deployment rate is slow ($u_{max} = 0.05$). The sample cost was 57.2847 cost units compared to an expected cost of 62.5937. *Top Right:* Expensive vaccine deployment ($\kappa = 0.5$) but a significantly faster maximum deployment rate ($u_{max} = 0.25$). The sample cost was 52.4261 cost units, compared to an expected cost of 54.374. *Bottom Row:* The optimally controlled expected dynamics. Blue lines give the uncontrolled expected dynamics as comparison. For both the slow and fast vaccination scenarios optimal control leads to greater similarity between the expected numbers of infecteds at each time than for the controlled dynamics.

optimal vaccination effort. For both the slow and fast vaccination campaigns the optimal vaccination allocation is more likely to be targeting the area with the initial infectious populations (area 1) for approximately the first generation time (1 time unit) but be substantially more likely to be targeting the area with the greatest number of initial susceptibles by second generation time. For the faster but more costly vaccination campaign further intervention becomes likely to be cost ineffective shortly after this. The cheaper vaccination campaign remains cost effective for a longer period, and the likelihood of returning to re-target area 1 increases in order to traduce the epidemic in its original outbreak area. This was due to either the initial vaccination effort at area 1 being insufficient to preclude an ongoing epidemic in area 1 or rescue effects of reintroduction from area 2 (figure 5.7).

Comparison to Uniform Vaccination and Transmission Symmetry Breaking

Encoding the transmission rates into a transmission kernel that only depends on distance enforces symmetric transmission between populations abiding in a spatially fragmented meta-population. In this section I break this symmetry for the only time in this thesis and investigate the interplay between breaking transmission symmetry, the location of the initially infected population and the performance of an optimal vaccination policy versus a uniformly distributed vaccination effort.

The transmission rates between areas one and two are now allowed to be given by an arbitrary matrix of positive real values,

$$\beta = \frac{1}{N_{loc}} \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}. \tag{5.78}$$

Where β_{ij} sets the transmission rate from a population in area j to area i. For each of choice of transmission parameters the optimal vaccination strategy can be solved using dynamic programming. On the other hand, one could define a simple uniform vaccination policy π_u where the maximal vaccination effort is divided evenly between both areas until epidemic cessation. The optimal vaccination strategy has two benefits compared to the uniform strategy; that the vaccine is deployed to the area faster where it gives optimal cost benefit and that the optimal strategy can assess when it has become cost ineffective to vaccinate further.

For the symmetric case with $\beta_{in} = 1.4$ and $\beta_{out} = 0.1$ I compared the

value of the initial epidemic state x_0 of having one initial infected population to a Monte Carlo estimate of the expected cost of using the uniform vaccination strategy, $J(x_0; \pi_u)$. For slow, cheap vaccination this saving was a modest 0.922 cost units (± 0.0696 3σ confidence intervals) but for faster more expensive vaccination the benefits of optimality were much greater, the saving was 7.3727 cost units (± 0.0870 3σ confidence intervals). I broke from this symmetric transmission scenario by varying β_{22} and β_{12} in such a manner that $\beta_{22} + \beta_{12} = 1.5$ was fixed. The transmission parameters for area were also left fixed $\beta_{11} = 1.4$ and $\beta_{21} = 0.1$. In this manner the expected number of infectious contacts made by a population in area 2 was left constant while the distribution of where those contacts were made was varied.

Estimating cost saving compared to uniform vaccination I found that if the index population was in area 1 the relative saving due to optimality was increasing with more asymmetric transmission. For the index population being in area 2 the relative saving decreased as transmission $2 \to 1$ became more rapid, attained a minimum and was increasing in the parameter region where area 2 was unable to sustain an epidemic locally (figure 5.4). I interpret this as being a consequence of the populations in area 1 being fundamentally more at risk than those of area 2 due to the unfavourable asymmetry $\beta_{11} + \beta_{12} > \beta_{21} + \beta_{22}$. Therefore if the initial infected population is in area 1 there is a combination of the missed opportunity cost of not targeting the area with the initial infecteds maximally, which as we have seen is the common optimal strategy, and the unfavourable asymmetry. If the initial infected population is in area 2 then these factors balance to an extent. Whilst area 2 is able to sustain an epidemic strongly locally (the threshold being $\beta_{22} > 1$) there is missed opportunity cost for not maximally targeting area 2 early in the epidemic, this missed opportunity cost becomes less significant as infectious spread within area 2 lessens. If area 2 can't sustain an epidemic locally then the missed opportunity is to target area 1 maximally despite the initial infected population originating in area 1.

5.6.2 Pricing the Lack of Information for the Two Area Model

Throughout this chapter I have assumed that the controller has had complete information with which to enact optimal vaccination controls. I now relax this assumption with two small case studies. In both cases the goal is to calculate the difference between the value of states with and without complete information. I call this difference a 'price' because it defines the fair amount the optimally acting controller would pay (in the currency units defined by the expected cost of one

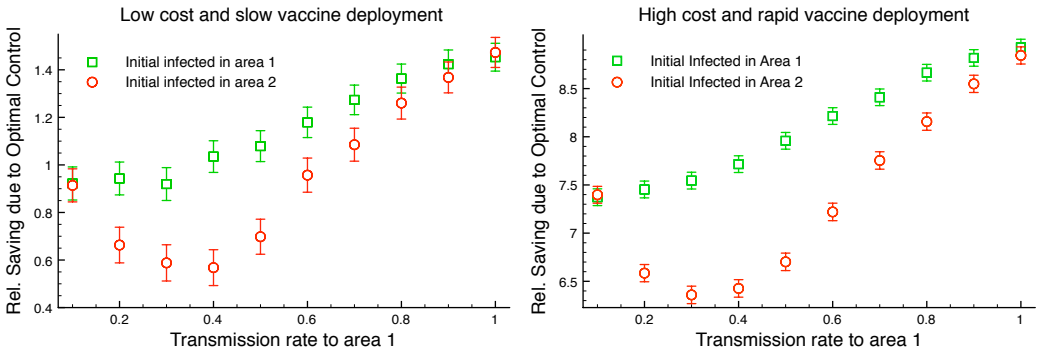


Figure 5.4: A cost comparison between optimally controlled epidemics and a uniform vaccination policy for symmetric and asymmetric transmission with one initially infected population. The spatially symmetric epidemiological parameters were $\beta_{in} = 1.4$, $\beta_{out} = 0.1$, $\gamma = 1$. Symmetry breaking was introduced by varying the area 2 to area 1 transmission parameter β_{12}. $\beta_{12}+\beta_{22}$ was kept fixed. The expected epidemic cost with uniform vaccination (estimated using 5000 MC replicates with 3σ error bars) was compared to the exact optimal cost. For the symmetric case optimal control: slow, cheap vaccination lead to a saving of 0.922 ± 0.0696; rapid but costly vaccination lead to a greater saving of 7.3727 ± 0.0870. For broken symmetry: relative saving depended significantly on location of the index population.

infectious population) for that extra information.

Erroneous Vaccination Effort due to Missing Population Information

This study considers when the vaccinating agency can successfully estimate the aggregate level of susceptible and infectious populations in each area but is unable to assess the exact disease status of each population. Therefore selecting a population in a target area can potentially lead to wasting vaccination resource on an infectious or removed population. I model the risk of mis-directing vaccination resource by the controller selecting target habitats within that the targeted area at random. Therefore the stochastic rate of vaccination at area i is transformed,

$$u_i(t) \to \frac{S_i}{N_i} u_i(t). \tag{5.79}$$

Since there is a $(1 - S_i/N_i)$ probability of the current deployment of vaccination effort delivering an effective zero rate of removals due to mis-targeting of vaccination resource. This model of ineffective vaccination is appropriate for migrating communities of wild animals where populations found at identified habitats can be targeted for vaccination coverage and there are estimates of disease burden over wider geographic areas, but it remains uncertain if individual population groups have been vaccinated due in ineffective tagging within the area. In this model the areas play

the dual role of population aggregations and as defining the resolution at which the controller has good statistics.

The optimal control problem with possibly mis-directed vaccination resource can be solved using dynamic programming by applying the transformation (5.79) to the HJB equation (5.76). The three possible optimal vaccination rates remain the same but within the inequalities governing the decisionw considered above we have,

$$\delta_{S_i,0} \to \frac{S_i}{N_i}, \qquad i = 1, 2. \tag{5.80}$$

The new optimal strategies therefore place more emphasis on targeting areas with greater numbers of susceptibles in order to reduce the chance of ineffective vaccination effort, whereas with complete information susceptible populations within an area with considerable infectious burden could be rescued by accurate and speedy vaccination.

This change in emphasis can be seen by considering the optimal policy π^* as a decision function mapping to the space of the three possible optimal vaccination rates. As an example I consider the subspace of the domain of π^* corresponding to all epidemic states where there are no removed populations. This can be represented as a two dimensional decision matrix; for vaccination without detailed population information both the region where it was optimal to not vaccinate and the region where it was optimal to target the area with greater numbers of susceptibles were greatly enlarged compared to the complete information decision matrix (figure 5.5). This was representative of the decision matrices for other two-dimensional subspaces (not shown). For both the slower, cheaper and the faster, more expensive vaccination campaigns the relative cost between the value of each state under complete information optimal control and optimal control with less information was calculated by comparison of the solutions to their respective HJB equations, I call this the 'price' of the missing information for each epidemic state.

In contrast to the price of controlling optimally compared to not controlling (figure 5.2) the peak price for lack of information occurred when the initial infecteds were equally distributed between the two areas; the maximal price for lack of information with a slow, cheap vaccination being 2.5458 cost units with the initial epidemic state being $I_1 = I_2 = 6$ and there being no removed populations. For a fast, expensive vaccination campaign the maximal price was 4.0867 at $I_1 = I_2 = 7$

(figure 5.5).

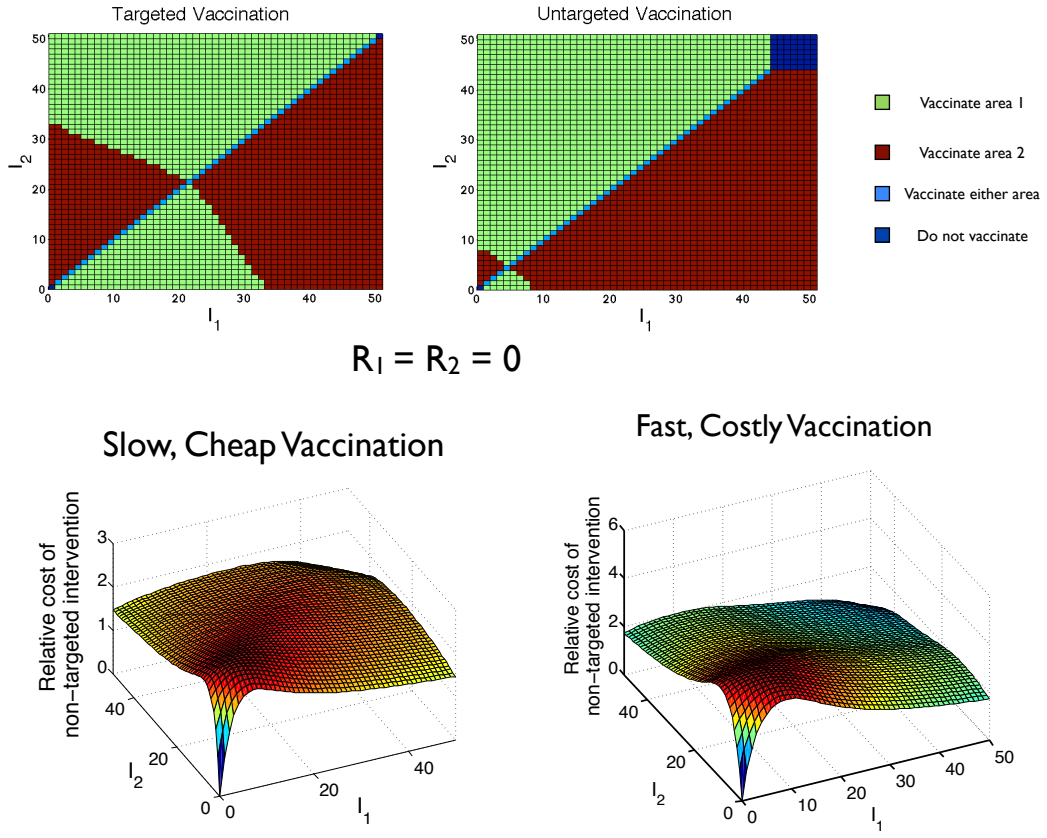


Figure 5.5: A comparison between targeted and untargeted vaccination, where there is the risk that the deployed effort is wasted on a population already removed from the epidemic. Epidemiological parameters were $\beta_{in} = 1.5$, $\beta_{out} = 0.5$, $\gamma = 1$. *Top:* The decision diagram for the initial ($R_1 = R_2 = 0$) persistent control decision with slow, cheap vaccination. When susceptible populations can be identified it is optimal to target the area with greatest infectious burden for a large part of the decision space. Controller inability to target susceptible populations enlarges the region where concentrating on the area with more susceptibles is optimal as well as the decision region where it is optimal to not vaccinate. *Bottom:* The extra cost incurred by inability to target populations. For slow, cheap vaccination the maximum extra cost is 2.5458 at $I_1 = I_2 = 6$. For fast, more expensive vaccination the maximum extra cost is 4.0867 at $I_1 = I_2 = 7$.

Delayed Starting of Vaccination Campaign

The second study into lack of information considers delayed initialisation of vaccination effort; the vaccination begins after the first infectious population is detected by the controlling agency which starts the epidemic is a 'passive' stance. After first detection the agency switches to an 'active' stance and initiates optimal vaccination. In 'active' stance the agency is presumed to have complete knowledge of the

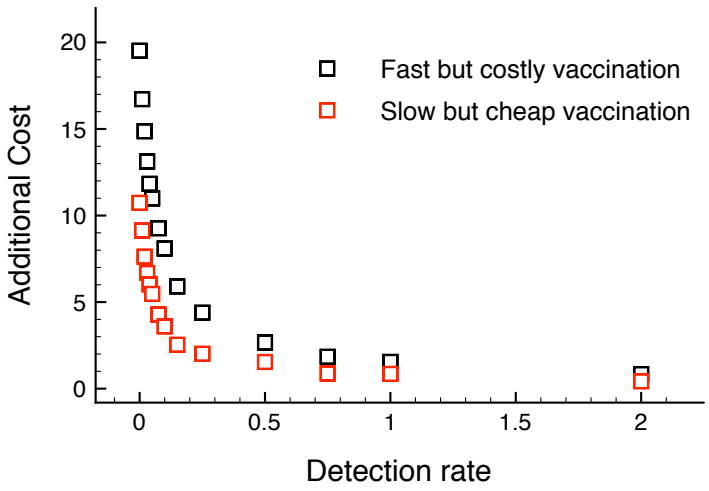


Figure 5.6: The epidemic is uncontrolled until the first infectious individual is detected. After first detection the epidemic is controlled using an optimal targeted vaccination policy. The additional cost of detection is the difference between the expected cost (estimated using 10^5 MC replicates - error bars were negligible) and the optimal expected cost without detection. Epidemiological parameters were $\beta_{in} = 1.5$, $\beta_{out} = 0.5$, $\gamma = 1$. There were 5 initially infectious populations. For both vaccination scenarios the additional cost decreases most rapidly over $D \in [0, 0.25]$. For fast but costly vaccination the additional cost is always greater.

epidemic state.

I model the change from 'passive' to 'active' for the controlling agency as a random boolean switch with the switching time being the first hitting time of a time, T_D, of an inhomogeneous Poisson process with intensity,

$$D\Big(\sum_i I_i(t)\Big). \tag{5.81}$$

Where D is the detection rate for each infectious population. Given that an initial epidemic state x_0 I define the 'price' for not having immediate information about epidemic invasion as,

$$\mathbb{E}^0_{x_0}\Big[\int_0^{T_D}\sum_i I_i(t)dt + V(X(T_D))\Big] - V(x_0). \tag{5.82}$$

Where $\mathbb{E}^0_{x_0}$ is the expectation for the epidemic conditioned on $X(0) = x_0$ and with no intervention. The price can be readily solved by solving the HJB equation and using Monte Carlo simulation to sample T_D and $\int_0^{T_D}\sum_i I_i(t)dt$. The price depends strongly on the detection rate D. For the case where the transmission parameters

were $\beta_{in} = 1.5$ and $\beta_{out} = 0.5$ with initially 5 infectious populations the price associated with each detection rate decayed on the interval $D \in [0, 0.25]$ and by $D = 2$ was small for both slow, cheap vaccination (0.43 cost units) and fast, expensive vaccination (0.8431 cost units). The price for not detecting immediately was always greater for the faster more expensive vaccination campaign compared to slower, cheaper vaccination (figure 5.6).

5.6.3 Approximately Controlled Epidemics

In this section I use the approximate dynamic programming method presented in section 5.5.2 to investigate approximate optimal control for epidemic systems with sufficiently many possible disease configurations as to preclude the use of the numerically exact value iteration method. Since the ADP method I use remains heuristic only in its justification I start by demonstrating that it can perform virtually as well as the exactly optimally controlled vaccination allocation at reducing epidemic cost. I will follow this by investigating the performance of the ADP algorithm for a more populous two area epidemic model under variation of parametrisation. Finally, I will give an example of the ADP method applied to a much more populous multi-area spatial epidemic as a motivating example for future work.

ADP Results for the Small Two Area Epidemic

I implemented the ADP algorithm for the small sized two area population model introduced above ($N_1 = N_2 = 50$) using symmetric transmission parameters $\beta_{in} = 1.5$, $\beta_{out} = 0.5$. For the polynomial approximation to the value function I set the maximum power of (S_i/N_i) to be considered as 10 and the maximum power of (I_i/N_i) to be 2. Since there are two areas being considered this meant that the number of parameters used to fit the value function $|\{C_{ijkl}\}| = 80$; this is a significant dimensionality reduction from $|\Sigma| = 1,758,276$ for the state space of the two area model with $N_1 = N_2 = 50$. The Kesten tuning parameter was set to be $a = 3$. 200 learning recursions ($M = 200$) were used for fitting the value approximation $\bar{V}$. The forward simulations for the learning recursions were initialised randomly with $I_i(0) = 0, 1, \ldots, 5$ chosen uniformly and $S_i(0) = 15, \ldots, 45$ chosen uniformly for $i = 1, 2$. The full learning sweep took approximately 5 minutes of computation time.

I investigated the performance of the ADP method via 10^4 Monte Carlo replicate simulations using (5.62) at each state visited to approximately optimise

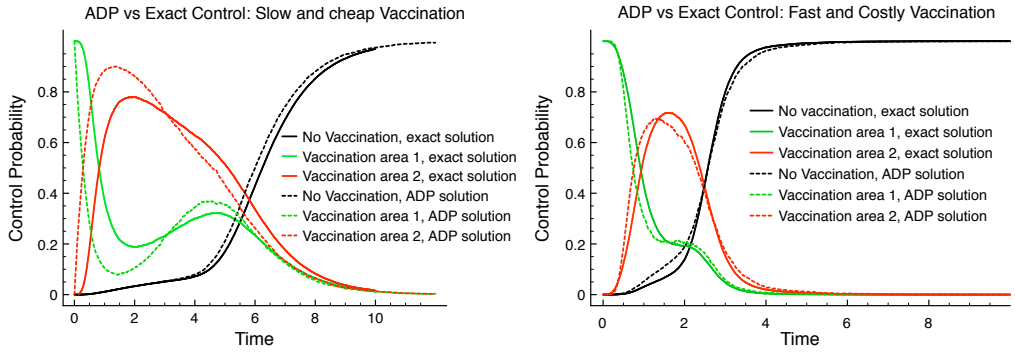


Figure 5.7: The time-varying probability of the control decision of the exact control dynamics for the two area epidemic compared to the time-varying probability of the control decision for the ADP approximated controls. The epidemiological parameters were $\beta_{in} = 1.5$, $\beta_{out} = 0.5$, $\gamma = 1$, $N_1 = N_2 = 50$. The epidemic seed was 5 initially infected populations in area 1. For faster, more expensive vaccination ($\kappa = 0.5$, $u_{max} = 0.25$) the ADP algorithm (dashed lines) recovers the expected control dynamics of the exact solution (solid lines). For slow, cheap vaccination ($\kappa = 0.1$, $u_{max} = 0.05$) the ADP expected control dynamics favoured switching from focus on area 1 to area 2 faster than the expected control dynamic of the exact solution but was still close to cost optimal.

the vaccination effort. The results were very good; for the slow, cheap vaccination ($\kappa = 0.1$, $u_{max} = 0.05$) scenario the expected cost of an epidemic initiated by 5 infected populations in area 1 was 62.64 (± 0.5499 3σ confidence interval) compared to the exact value 62.5937. For the faster deploying but more costly vaccination scenario ($\kappa = 0.5$, $u_{max} = 0.25$) the expected cost was 54.48 (± 0.627) compared to the exact value 54.374. In both cases the empirically calculated expected cost was within confidence interval of the exactly calculated value. I also calculated the time-varying probability of each control decision from the empirical probability of the controller performing one of the three possible optimal vaccination actions at each time, for simplicity this was done using a tau-leap stochastic simulation method with time step $\delta t = 0.01$. The idea was to get a better sense of the 'likely' progression of optimal control decisions.

For faster, more expensive vaccination the expected control dynamics derived using the approximate value function $\bar{V}$ very closely resembled its exactly controlled counterpart. For slower, cheaper vaccination the ADP expected control dynamics favoured switching from focus on area 1 to area 2 significantly faster than the expected exact control dynamics, but also favoured returning to area 1 in the latter stages of the epidemic more than its exact counterpart (figure 5.7). For expensive but potentially very effective vaccination it is likely that there are only a tight band of near optimal strategies that deliver significant cost savings compared to (say)

uniform vaccination; for slower cheaper vaccination there is probably a much wider space of strategies that deliver near optimal cost savings.

ADP for Larger Spatial Epidemics

I start by investigating the performance of the ADP algorithm under variation of the fitting parameters for the two area model but with much larger population numbers $N_1 = N_2 = 1000$. I chose the transmission parameters $\beta_{in} = 1.5$ and $\beta_{out} = 0.5$ and an intermediate vaccination campaign ($\kappa = 0.25$, $u_{max} = 0.15$). Since this is far too large a system to use value iteration upon I am unable to compare the resulting expected approximately controlled epidemic costs to the true values. Instead, I chose a baseline set of fitting parameters to be varied from: maximum power of (S_i/N_i) fitted $K = 10$, maximum power of (I_i/N_i) fitted $L = 2$, initial susceptible number for forward simulation $S_i(0) = s_{min}N_i, \ldots, 0.9N_i$ chosen uniformly random where the baseline density was $s_{min} = 0.3$, initial infecteds number for forward simulation $I_i(0) = 0, \ldots, 5$ chosen uniformly random, $a = 3$ and 500 forward simulation learning recursions $M = 500$. Different parameter choices were compared by using Monte Carlo estimates using the resultant value approximation $\bar{V}$ to construct approximately optimal vaccination controls. The expected cost for the base line fitting parameters was calculated from 1000 MC replicates for an epidemic initiated by 5 infectious populations at area 1 was 649.4 ± 17.5. By comparison the uncontrolled epidemic cost was 1544.0 ± 27.9 and the uniformly vaccinated epidemic cost was 776.9 ± 20.8. Both the learning simulations and the MC estimates were performed using the tau-leap stochastic simulation algorithm with time step $\delta t = 0.01$.

Increasing the number of learning recursions (M) improved the ADP performance with plateauing improvement after approximately $M \approx 100$. This was a surprisingly small numbers of recursions for plateaued performance, although one should recall that the global value polynomial approximation was updated for each state visited by the forward simulation learning sweeps. The other most sensitive fitting parameter was the maximum power of (S_i/N_i) for $i = 1, 2$ (K). Expected epidemic cost reduced strongly with increasing K up to $K = 10$ and then mildly increased. This is probably a sign that the number of parameters and the numbers of learning rules should not be varied independently since more parameters typically require larger data sets for good fitting, c.f. (Bender and Steiner 2012) for a discussion of this with respect to least-squares regression with MC simulation. There was also a noticeable effect from varying the initial distribution of states for the forward learning simulations. If $s_{min} > 0.6$ then the expect cost was significantly higher than

for the baseline fitting parameters. I interpret this as the ADP algorithm requiring a selection of states where the number of susceptibles is too low for the epidemic to take off in order to learn that having a low numbers of susceptibles is a strong indicator of a low value state. Variation of the maximum power of (I_i/N_i) for fitting (L) seemed to have no significant effect of the expected cost. See figure 5.8 for plots of the above. Variation of the Kesten tuning parameter a seemed to have little effect (not shown).

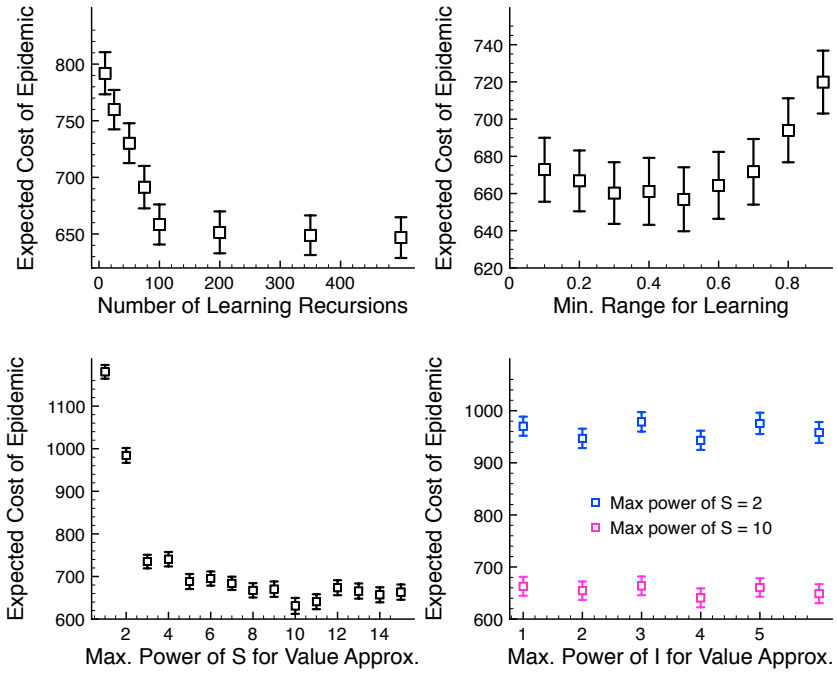


Figure 5.8: Variation of the ADP value function fitting parameters. For the two area epidemic model with $N_1 = N_2 = 1000$ with epidemiological parameters $\beta_{in} = 1.5$, $\beta_{out} = 0.5$, $\gamma = 1$. The vaccination campaign was an intermediate between the slow and fast vaccinations considered before ($\kappa = 0.25$, $u_{max} = 0.15$). The base-line fitting parameters varied from were maximum power of (S_i/N_i) fitted $K = 10$, maximum power of (I_i/N_i) fitted $L = 2$, initial susceptible number for forward simulation $S_i(0) = 0.3N_i, \ldots, 0.9N_i$ chosen uniformly random, initial infecteds number for forward simulation $I_i(0) = 0, \ldots, 5$ chosen uniformly random, $a = 3$ and 500 forward simulation learning recursions $M = 500$. Expected epidemic costs are for 5 initial infecteds at area 1 controlled using value approximation $\bar{V}$ estimated with 3σ confidence intervals from 1000 Monte Carlo replicates. *Top Left:* Varying M. The expected cost reduces rapidly to a plateau after 100 learning recursions. *Top Right:* Varying the minimum range in density of the randomly selected initial susceptible number for each area. A minimum density range of 0.3-0.5 seems to produce the best results. *Bottom Left:* Varying K. The expected cost reduces strongly as the maximum power of S_i considered for fitting increases to 10 and mildly increases. *Bottom Right:* Varying L for $K = 2$ and $K = 10$. The effect of varying the maximum power of I_i seems to be negligible.

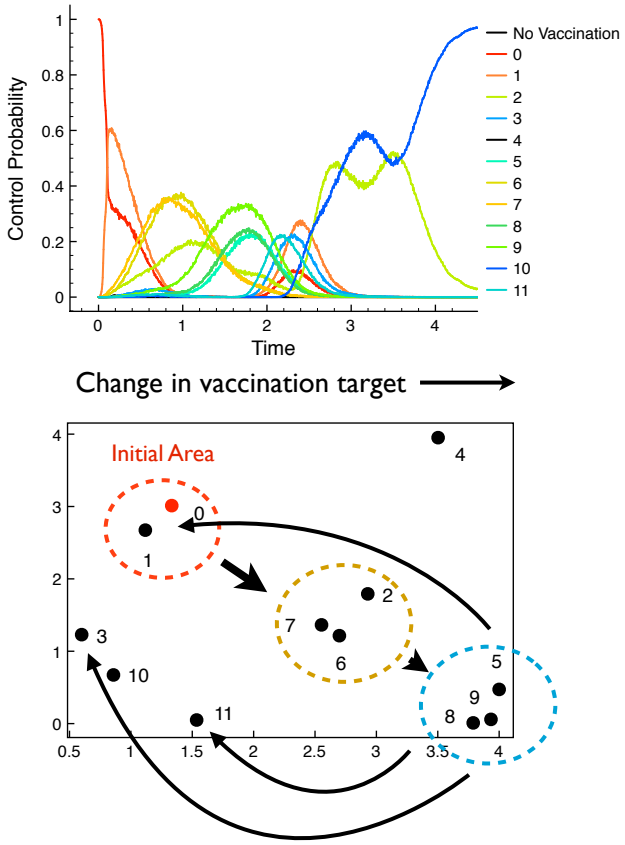


Figure 5.9: An example of ADP-based control for a large epidemic. The populations were distributed into 12 areas, each containing 10,000 populations. Transmission was governed by a Gaussian shaped kernel with $K_0 = 10$ and $L = 1$. The initial epidemic seed was 5 infected populations in area 0. *Top:* The expected control dynamics were given by 5000 MC replicates using the optimal vaccination strategy for the ADP approximate value function $\bar{V}$. The curves each represent the time-varying probability of each area being targeted for vaccination. They are colour coded red to blue according to the timing of the first strong likelihood of being targeted. Targeting arrived in five distinct groupings. *Bottom:* The control target groups represented in space. Black arrows indicate the sequence of likely targeting. First areas 0 and 1 were targeted (red dashed circle), then areas 7 and 6 with a smaller likelihood of targeting area 2 (yellow dashed circle), then area 9 with a smaller likelihood of targeting areas 5 and 8 (blue dashed circle). These represent the initially infected area, its near neighbour, and the two areas of spatial clustering. Subsequently, the ADP-based vaccination schedule targeted either areas 1,3, or 11 before finally switching to area 2 or 10. Area 4, which is spatially isolated, was never deemed cost effective to target.

The plateauing of the expected cost under variation of fitting parametrisation indicates that the ADP algorithm has reached its optimal value, which one hopes is close to the true optimal value given by the solution to the HJB equation. To finish the numerical investigations I consider optimally controlling a much large meta-population of $N_{area} = 12$ areas each with a local population $N_i = 10,000$ labelled $i = 0, \ldots, 11$. Each area's centroid was distributed randomly in $\mathcal{A} = [0, 4]^2$ (figure 5.9) and the meta-population is invaded by a strongly invasive pathogen. The transmission kernel is a scaled by the transmission length scale L and is Gaussian shaped, exactly as considered for the plant epidemic in chapter 4. The transmission kernel parameters were $K_0 = 10$ and $L = 1$. The vaccination parameters were $\kappa = 0.5$ and $u_{max} = 0.25$. For estimating the value function I used the baseline fitting parameters considered above except with $M = 5000$ learning sweeps to reflect the larger system size. The epidemic was seeded in area 0 with 5 initially infected populations.

The expected control dynamics estimated from 5000 MC replicates using the ADP approximate value function follow a distinct pattern. First the initial area and the nearest neighbouring area were targeted, then sequentially the groupings of areas further from the initial epidemic seed followed by a 'mopping up' operation of the more spatial isolated area. One area (area 4) is not targeted at any point (figure 5.9). For this highly invasive pathogen the expected uncontrolled cost was very near maximal $117,288 \pm 173.2$ and the cost with uniform vaccination up until cessation was even worse $122,521 \pm 222.3$, this counter-intuitive result was due to the large expense of ineffectively vaccinating the tail of the epidemic. The approximately optimally controlled epidemic cost was $115,220 \pm 179.4$, but by contrast a uniform vaccination policy that operated until the population reached the herd immunity level of $S_i \leq 0.1N_i$ and then ceased, thereby avoiding ineffective tail costs, delivered nearly the same expected cost, $115,765 \pm 209.2$. The ADP controlled dynamics presented here are probably not optimal, however I give them to show that approximate dynamic programming is on the cusp of informing optimal control problems for realistic sized systems.

5.7 Discussion

In this chapter I have presented a novel application for classical dynamical programming, the vaccination control problem for two spatially stratified population aggregates undergoing a Markovian stochastic epidemic. Vaccination control is highly non-standard since the controller can only influence the dynamics of the costly part

of the system (the infected populations) through the second order effect of depleting susceptibles. Since it is not common for the mobile resources assumed to exist in this work to be available in the real-world, nor complete information about the invading epidemic, my treatment has focused on quantifying the cost associated with not having complete information or using a non-optimal but simple strategy such as uniform vaccination. By understanding these missed opportunity costs it is possible to assess whether it is cost effective to expend capital on mobile resources and more aggressive information gathering.

In this chapter only vaccination has been considered as an intervention methodology, there are other interesting possibilities such as targeted treatment with contact tracing, c.f. (Eames and Keeling 2003), where the treated individuals 'transmit' their treatment to infected individuals thereby modelling the admission of their infectious contacts into the treatment program. Interesting optimal control problems here include solving for a cost optimal contact tracing rate and optimally deciding upon when to switch between a contact tracing based intervention against a novel pathogen to a mass vaccination effort. Another interesting direction would be to reformulate the timing of school closures during an emerging epidemic as an optimal stopping problem so as to optimally minimise excess demand on spatially local intensive care units, as has been considered by House et al, (House et al. 2011).

The applicability of numerically exact dynamic programming to such questions will probably remain confined to 'toy' sized models due to the prohibitively large state space of realistic epidemic models. Therefore I have also presented a novel algorithm for approximately solving dynamic programming problems based on fitting a polynomial form to the value function of the spatial epidemic using a stochastic gradient method. At least for the two area epidemic model this was successful and the extension to larger epidemic systems has been shown to be computationally feasible.

The major problem with the application of the ADP algorithm is its heuristic nature; there is no solid guarantee of optimality. Future analytic directions include making closer connection to the theory of Forward-Backwards SDEs which, from the ADP perspective, represent the coupled processes of a forward epidemic process and a backwards remaining cost process. At least in some cases there exist rigorous error criteria for such processes (Bender and Steiner 2010). Another possibility is to investigate more closely the stochastic Pontryagin principle given for a wide class of

stochastic control problems (including those driven by Poisson processes) given by Øksendal and Sulem (Øksendal and Sulem 2005) chapter 3. This gives a necessary condition for the SDE form of the change in value given an arriving event along an optimally controlled trajectory; an interesting next step would be to see if it was feasible to numerically check this rather abstract necessary condition for optimality.

Missing from my treatment of optimal vaccination problems has been the Levins-type spatial meta-population model, which has been the focus of the previous chapters. The value approximation form used for my ADP algorithm is not suitable for this scenario since it is written in terms of area disease state densities and therefore is inappropriate for single spatially isolated populations. One direction for tackling this problem is to approximate the control problem by aggregating local population into artificial area groupings. The value function would be solved for the aggregated model, but the controls would be applied to the true model. It is unclear how effective this would be. Another possibility would be to use the vaccination ring method of Tildesley et al (Tildesley et al. 2006) which was applied to finding the optimal size of ring around each confirmed FMD infectious farm within which to target vaccination. A one step extension could be to solve the optimal vaccination scheduling of the farms within each ring.

Certainly the possibility of tackling optimal decision problems much larger than previously considered in the literature of theoretical epidemiology is intriguing, albeit fraught with difficulty. I look forward to investigating these possibilities more throughly in the future.

Chapter 6

Conclusion and Outlook

In this thesis I have investigated a widely used class of spatial and stochastic epidemic model. My treatment has focused on theoretical aspects, increasing simulation efficiency and giving a control theoretic interpretation to vaccination scheduling amongst the disjoint populations of the model. I have used rigorous mathematical methods wherever I was capable of doing so; in some sections of the thesis the results I construct are based on heuristic arguments in mathematical form supported by simulation-based evidence.

My theoretical approach was based on appreciating the role that dynamical features well-known in the literature of stochastic calculus have on the expected stochastic dynamics of the spatial epidemic, constructing ordinary differential models to approximate these expected dynamics and bounding their contribution. In particular the role of quadratic covariation between disease state indicators was emphasised in **chapters 2** and **3**; where it was found to play the dual role of enforcing the discreteness of the epidemic dynamics and introducing covariation between the stochastic dynamics of disease states for separate populations. In fact, the stochastic differential approach taken in these chapters made it possible to write down exact dynamics for the spatial covariances between populations, which in turn drove the variation between expected dynamics of the stochastic epidemic model from related deterministic models. So long as the transmission kernel was chosen so that its total volume on the space was fixed this variation could be bounded in terms of powers of the inverse characteristic length of transmission.

Since the exact covariance dynamics for each order were found to depend on higher order covariance I was led to consider the classic moment and covariance hi-

erarchy approach to approximating the expected dynamics of the spatial epidemic. Here I gave a novel set of moment equations for the spatial SIR epidemic based on constructing the approximation ODES directly from the underlying stochastic differential processes of infection and recovery. I also introduced a closure based on truncating the spatial quadruple covariances. Ultimately, this was comparatively unsuccessful since this third order approximation was generally out-performed by a power-2-type closure already known to the literature, although I was able to give the power-2 closure a novel interpretation as a closure choice 'likely' to be a good approximation of the dynamics of spatial triple covariance. The successful power-2-type closure allowed me to follow Bolker in revealing the spatial clustering and disaggregation mechanisms underlying the spatial epidemic dynamics and their effect on the expected disease progression, as well as identifying an alarming amplification effect due to the combination of heavy-tailed transmission and spatially clustered habitats.

In **chapter 4** I was able to present a novel algorithm for accelerating the Monte Carlo simulation of the spatial epidemic model where the computational effort required scaled asymptotically as $N \ln N$ compared to N^2 for more traditional simulation methods. Since the method was based on the spectral decomposition of the spatially varying instantaneous risk of infection I called this the *Fast Spectral Rate Recalculation* (FSR) algorithm. Its novelty lay in exploiting the efficiency of the fast Fourier transform in order to efficiently incorporate the change in stochastic rates of the process upon the arrival of new random events. Its asymptotic speed savings and accuracy were demonstrated using extensive Monte Carlo simulation with matched latent variables (Cook et al. 2008), in support of a rigorous analysis of the sources of error for the approximation. Additionally, for Gaussian shaped kernels greater accuracy could be obtained by a kernel correction technique. The performance of the algorithm was demonstrated for both the Levins-type metapopulation model and a metapopulation model with explicit within habitat population dynamics.

In **chapter 5** I applied the framework of dynamic programming to solving an optimal vaccination scheduling amongst the sub-populations problem. For a toy sized model I was able to demonstrate numerically exact optimal vaccination deployment between two distinct geographic areas. The optimal policies were solved in feedback form, so that the vaccinating agency was able to react optimally to incoming information about the progression of the disease. I also investigated the effect that lacking information had on the optimality of the epidemic management.

For fast deploying, but expensive, vaccination optimality in targeting the available resource was found to be particularly important.

I also introduced a novel method for approximating the the optimal vaccination policy for the spatial epidemic. This method was based on ideas from the field of approximate dynamic programming, and is based on using Monte Carlo simulations with recursively improved vaccination decisions to estimate a polynomial approximation to the value function of the epidemic problem; that is the minimum possible expected cost for each possible disease configuration.

The future direction I could take my research feels broad, as a limited subset I would like to,

- Further clarify the dynamical reasons why the power-2 approximation performs so well in predicting the spatial moment dynamics of the spatial epidemic. I intend to continue using the covariance dynamics method, and investigate how quadratic approximations in pair covariances perform under different scenarios.
- Explore other approximations of farm point locations other than as spatial Gaussians; in particular with the goal of developing kernel corrections for non-Gaussian shaped transmission kernels.
- Investigate the possibility of using the FSR algorithm in conjunction with MCMC for accelerated performance, or the novel parameter estimation method of Ionides et al (Ionides, Breto, and King 2006).
- Push forward the ADP-type algorithm for optimal intervention. Can this approach be extended to the true Levins-type metapopulation, or indeed truely realistic disease management? If so, what are the essential modification to the value approximation and what computational resource will be required?

There seems a lot to do. I better crack on.

Thank you for reading.

Appendix A

Triple Covariance Dynamics

Applying equation (3.58) and expanding in terms of three and four point covariance functions by using the standard decomposition $X(x,t) = \langle X \rangle + \tilde{X}(x,t)$ for each random spatial distribution of disease states gives the triple covariance dynamics,

$$\begin{aligned}
\frac{d}{dt}c_{SSS}(x-y,x-z,t) &= -3K_0\langle I\rangle c_{SSS} \\
&-\langle S\rangle \int K(z-w)c_{SSI}(x-y,x-w,t)dw \\
&-\langle S\rangle \int K(y-w)c_{SSI}(x-z,x-w,t)dw \\
&-\langle S\rangle \int K(x-w)c_{SSI}(y-z,y-w,t)dw \\
&+\delta(x-y)\langle \tilde{S}(z,t)S(x,t)\lambda(x,t)\rangle + \delta(x-z)\langle \tilde{S}(y,t)S(x,t)\lambda(x,t)\rangle \\
&+\delta(y-z)\langle \tilde{S}(x,t)S(y,t)\lambda(y,t)\rangle - 3\int K(x-w)c_{SSSI}(x,y,z,w,t)dw \\
&-\delta(x-y)\delta(x-z)\langle S(x,t)\lambda(x,t)\rangle. \qquad \text{(A.1)}
\end{aligned}$$

$$\frac{d}{dt}c_{SSI}(x-y,x-z,t) = K_0\langle I\rangle c_{SSS} + \langle S\rangle \int K(z-w)c_{SSI}(x-y,x-w,t)dw$$
$$+\int K(z-w)c_{SSSI}(x,y,z,w,t)dw - \gamma c_{SSI}$$
$$+\delta(x-y)\langle \tilde{I}(z,t)S(x,t)\lambda(x,t)\rangle - 2K_0\langle I\rangle\Big[c_{SSI}(x-y,t,x-z,t)\Big]$$
$$-\langle S\rangle \int K(y-w)c_{SII}(x-z,x-w,t)dw$$
$$-\langle S\rangle \int K(x-w)c_{SII}(y-z,y-w,t)dw$$
$$-\delta(x-z)\langle \tilde{S}(y,t)S(x,t)\lambda(x,t)\rangle - \delta(y-z)\langle \tilde{S}(x,t)S(y,t)\lambda(y,t)\rangle$$
$$-2\int K(x-w)c_{SSII}(x,y,z,w)dw$$
$$+\delta(x-y)\delta(x-z)\langle S(x)\lambda(x)\rangle. \quad \text{(A.2)}$$

$$\frac{d}{dt}c_{SII}(x-y,x-z,t) = 2K_0\langle I\rangle c_{SSI} + \langle S\rangle \int K(z-w)c_{SII}(x-y,x-w,t)dw$$
$$+\langle S\rangle \int K(y-w)c_{SII}(x-z,x-w,t)dw$$
$$+2\int K(y-w)c_{SSII}(x,y,z,w)dw - 2\gamma c_{SII}$$
$$-\delta(x-y)\langle \tilde{I}(z,t)S(x,t)\lambda(x,t)\rangle - \delta(x-z)\langle \tilde{I}(y,t)S(x,t)\lambda(x,t)\rangle$$
$$-K_0\langle I\rangle c_{SII} - \langle S\rangle \int K(x-w)c_{III}(y-z,y-w,t)dw$$
$$-\int K(x-w)c_{SIII}(x,y,z,w,t)dw + \delta(y-z)\langle \tilde{S}(x,t)(S(y,t)\lambda(y,t)$$
$$+\gamma I(y,t))\rangle - \delta(x-y)\delta(x-z)\langle S(x,t)\lambda(x,t)\rangle.$$

$$\frac{d}{dt}c_{III}(x-y,x-z,t) = 3K_0\langle I\rangle c_{SII} + \langle S\rangle \int K(z-w)c_{III}(x-y,x-w,t)dw$$
$$+\langle S\rangle \int K(y-w)c_{III}(x-z,x-w,t)dw$$
$$+\langle S\rangle \int K(x-w)c_{III}(y-z,y-w,t)dw - 3\gamma c_{III}$$
$$+3\int K(x-w)c_{SIII}(x,y,z,w,t)dw$$
$$+\delta(x-y)\langle \tilde{I}(z,t)(S(x,t)\lambda(x,t)+\gamma I(x,t))\rangle$$
$$+\delta(x-z)\langle \tilde{I}(y,t)(S(x,t)\lambda(x,t)+\gamma I(x,t))\rangle$$
$$+\delta(y-z)\langle \tilde{I}(x,t)(S(y,t)\lambda(y,t)+\gamma I(y,t))\rangle$$
$$+\delta(x-y)\delta(x-z)\langle S(x,t)\lambda(x,t)-\gamma I(x,t)\rangle.$$

Where I have written the stochastic force of infection at $x \in \mathcal{A}$, $\lambda(x) = \int K(x-w)I(w,t)dw$. I have also not expanded the quadratic and cubic variation terms, these expand in terms of moments, two and three point covariance functions and therefore can be represented within the closed covariance hierarchy after I neglect the contribution of four-point spatial covariances.

It is important to note that the interaction between covariance functions at different separation distances is more complicated than for the power-1 closure. For example the dynamics of c_{SSS} involve the term,

$$\left\langle \langle S\rangle \tilde{S}(x,t)\int K(z-w)\tilde{S}(y,t)\tilde{I}(w,t)dw\right\rangle = \langle S\rangle \int K(z-w)c_{SSI}(x,y,w,t)dw$$
$$= \langle S\rangle \int K(z-w)c_{SSI}(x-y,x-w,t)dw,$$

as seen above.

Appendix B

Frequency Domain Representation for Higher Order Approximations

Solving the extended covariance hierarchy in terms of Fourier transformed covariance functions is still numerically attractive because the size of the relevant spectrum of spatial frequencies to solve over scales as $1/L$ rather than L for real space calculations. The Fourier representation also permits the reduction of convolution terms into products. In this section I write $\omega = \omega_{x-y}$ and $\omega' = \omega_{x-z}$ for the spatial frequency along respectively the $x - y$ separation vector and the $x - z$ separation vector.

B.1 The Power-2 Closure

The two integral terms dropped by the power-1 closure and approximated by the power-2 closure are,

$$\int K(x-z)c_{SII}(x,y,z)dz \tag{B.1}$$

$$\int K(y-z)c_{SSI}(x,y,z)dz \tag{B.2}$$

The power-2 closure used with translational invariance is,

$$\begin{aligned} c_{XYZ}(x,y,z;t) &\approx c_{XYZ}(x-y,x-z;t) \\ &= \frac{1}{2}\Big[\frac{c_{XZ}(x-z;t)c_{YZ}(y-z;t)}{\langle Z\rangle} + \frac{c_{XY}(x-y;t)c_{XZ}(x-z;t)}{\langle X\rangle} \\ &\quad + \frac{c_{YZ}(y-z;t)c_{XY}(x-y;t)}{\langle Y\rangle}\Big]. \end{aligned} \tag{B.3}$$

Introducing this estimate into (B.1) and (B.2) gives,

$$\int K(x-z)c_{SII}(x,y,z)dz = \frac{1}{2}\Big(\frac{c_{SI}K * c_{II}}{\langle I\rangle} + \frac{c_{SI}[K * c_{SI}(0,t)]}{\langle S\rangle} + \frac{(Kc_{SI}) * c_{SI}}{\langle I\rangle}\Big) \tag{B.4}$$

$$\int K(y-z)c_{SSI}(x,y,z)dz = \frac{1}{2}\Big(\frac{c_{SS}K * c_{SI}}{\langle S\rangle} + \frac{c_{SI}[K * c_{SI}(0,t)]}{\langle I\rangle} + \frac{(Kc_{SI}) * c_{SS}}{\langle S\rangle}\Big) \tag{B.5}$$

Where the dependencies not given are all on the distance $(x-y)$ and the time t. These expressions are in product-convolution form and therefore can be easily transformed into the Fourier domain,

$$\begin{aligned} \boldsymbol{F}\Big[\int K(x-z)c_{SII}(x,y,z)dz\Big] &= \frac{1}{2}\Big(\frac{\hat{c}_{SI} *_\omega (\hat{K}\hat{c}_{II})}{\langle I\rangle} + \frac{\hat{c}_{SI}[\int \hat{K}\hat{c}_{SI}d\omega']}{\langle S\rangle} \\ &\quad + \frac{(\hat{K} *_\omega \hat{c}_{SI})\hat{c}_{SI}}{\langle I\rangle}\Big), \end{aligned} \tag{B.6}$$

$$\begin{aligned} \boldsymbol{F}\Big[\int K(y-z)c_{SSI}(x,y,z)dz\Big] &= \frac{1}{2}\Big(\frac{\hat{c}_{SS} *_\omega (\hat{K}\hat{c}_{SI})}{\langle S\rangle} + \frac{\hat{c}_{SI}[\int \hat{K}\hat{c}_{SI}d\omega']}{\langle I\rangle} \\ &\quad + \frac{(\hat{K}\hat{c}_{SI}) *_\omega \hat{c}_{SS}}{\langle S\rangle}\Big). \end{aligned} \tag{B.7}$$

For numerical calculation the convolutions over frequencies were solved by a numerical quadrature on their integral definition.

B.2 The Third Order Approximation

There are three different convolution forms influencing the the covariance function dynamics for the third order approximation in the frequency domain,

$$\int K(z-w)c_{XYZ}(x-y,x-w,t)dw, \tag{B.8}$$

$$\int K(y-w)c_{XYZ}(x-z,x-w,t)dw, \tag{B.9}$$

$$\int K(x-w)c_{XYZ}(y-z,y-w,t)dw. \tag{B.10}$$

Each convolution (B.8), (B.9), (B.10), describes a different interaction between the Fourier dynamics at different frequencies. For the third order approximation these can be written down by using identity

$$f(x,y) = \int \hat{f}(\omega_1,\omega_2) e^{2\pi i[\omega_1\cdot x+\omega_2\cdot y]} d\omega_1 d\omega_2$$

and the identity

$$\delta(\omega_1-\omega_2) = \int e^{-2\pi i(\omega_1-\omega_2)\cdot x} dx$$

in order to solve the integrals directly. For example,

$$\begin{aligned}
&\boldsymbol{F}\Big[\int K(x-w)c_{XYZ}(y-z,y-w,t)dw\Big](\omega,\omega',t) \\
&= \int K(x-w)c_{XYZ}(y-z,y-w,t)e^{-2\pi i[\omega\cdot(x-y)+\omega'\cdot(x-z)]}dwd(x-y)d(x-z) \\
&= \int \hat{K}(\omega_3)\hat{c}_{XYZ}(\omega_1,\omega_2,t)e^{-2\pi i[-\omega_1\cdot(y-z)-\omega_2\cdot(y-w)-\omega_3\cdot(x-w)+\omega\cdot(x-y)+\omega'\cdot(x-z)]} \\
&\qquad \times dwd(x-y)d(x-z)d\omega_1 d\omega_2 d\omega_3 \\
&= \int \hat{K}(\omega_3)\hat{c}_{XYZ}(\omega_1,\omega_2,t)\delta(\omega_3+\omega_2)\delta(\omega_1+\omega_2+\omega)\delta(\omega'-\omega_1)d\omega_1 d\omega_2 d\omega_3 \\
&= \hat{K}(\omega+\omega')\hat{c}_{XYZ}(\omega',-(\omega+\omega'),t) \qquad \text{(B.11)}
\end{aligned}$$

The other convolution forms have the Fourier representation given by,

$$\boldsymbol{F}\Big[\int K(y-w)c_{XYZ}(x-z,x-w,t)dw\Big](\omega,\omega',t) = \hat{K}(\omega')\hat{c}_{XYZ}(\omega',\omega,t), \quad \text{(B.12)}$$

$$\boldsymbol{F}\Big[\int K(z-w)c_{XYZ}(x-y,x-w,t)dw\Big](\omega,\omega',t) = \hat{K}(\omega')\hat{c}_{XYZ}(\omega,\omega',t). \quad \text{(B.13)}$$

Any other Fourier calculation can be solved in similar manner prior to numerical integration.

I have reduced the dimensionality of the necessary numerical calculations by reducing to two frequencies by using translational invariance; however each spatial frequency is a d-dimensional vector. Therefore for calculations in $d = 2$ the three-point covariance functions must be solved as a four dimensional object. Fortunately, rotational invariance in real space allows the reduction of a further dimension; all the relevant spatial information can be encoded as a function of the triple $(|x-y|, |x-z|, \theta)$, where θ is the interior angle between the edge vectors $x-y$ and $x-z$ of the triangle in figure 3.2.

As we have seen rotational invariance carries over into the frequency domain; in order to numerically integrate the three-point covariance functions it is sufficient to solve them as functions of $(|\omega|, |\omega'|, \theta)$ where θ is the interior angle between the position vectors in the frequency domain ω and ω'. Armed with this information I can identify the contribution of the Fourier transformed transmission kernel in two spatial dimensions at the $\omega + \omega'$ spatial frequency using standard geometry,

$$\hat{K}(\omega + \omega') = \hat{K}(|\omega + \omega'|) = \hat{K}(\sqrt{\omega^2 + \omega'^2 + 2|\omega||\omega'|\cos\theta}). \tag{B.14}$$

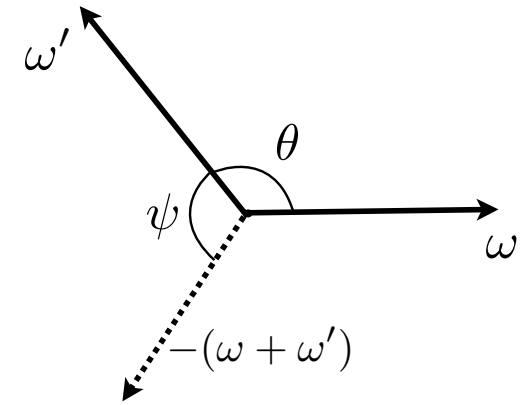
I can also identify the contribution of the Fourier transformed three-point spatial covariance function at the joint frequency $(\omega', -(\omega + \omega'))$,

$$\hat{c}_{XYZ}(\omega', -(\omega + \omega')) = \hat{c}_{XYZ}(|\omega'|, |\omega + \omega'|, \psi). \tag{B.15}$$

Where ψ is the interior angle between the two position vectors ω' and $-(\omega + \omega')$. This angle can also be found in terms of the triple $(|\omega|, |\omega'|, \theta)$,

$$\cos\psi = -\frac{\omega'^2 + |\omega||\omega'|\cos\theta}{|\omega'||\omega + \omega'|}. \tag{B.16}$$

See figure B.1.



$$|\omega+\omega'| = \sqrt{\omega^2+\omega'^2+2|\omega||\omega'|\cos\theta}$$

$$\cos\psi = -\frac{\omega'^2+|\omega||\omega'|\cos\theta}{|\omega'||\omega+\omega'|}$$

Figure B.1: The spatial frequency geometry of three interacting populations. The three relevant position vectors in the frequency domain and their interior angles. The interior angle between ω' and $-(\omega+\omega')$ can be found by noting that the third edge of the triangle they form is length $|\omega+2\omega'|$ and using the law of cosines.

Appendix C

Theoretical Basis for Fast Spectral Simulation

Fundamentally, I am concerned with comparing two Markovian epidemic processes; X, with some defined probabilistic dynamics, and X' defined by using the same model stochastic rates approximated via FSR. In general demonstrating convergence of the FSR estimated rates used to generate X' to the rates used to generate X is insufficient to show $X' \to X$ in any sense. However, our situation is sufficiently simple (both X and X' take values on the same finite state space $\{S, I, R\}^N$) that if I can show,

- The initial probability distribution $\mathbb{P}(X'(0)) \to \mathbb{P}(X(0))$ in probability,
- $\lambda_i'(X'(t)) \to \lambda_i(X(t))$ for all states $X(t) \in \{S, I, R\}^N$ and for all habitats $i = 1, \ldots, N$ (λ_i' is the force of infection estimated using FSR),

then I have $X' \to X$ in probability, or equivalently the distribution functions of X' and X are identical. For justification see Kallenberg Theorem 17.15 [Kallenberg]. The first condition is easily fixed as a model choice. The second condition requires a spectral analysis of the error in using the grid based approximation $\lambda_{PS}^{M,\epsilon}$ to estimate the force of infection on each susceptible habitat.

I demonstrate my convergence result for the pseudo-spectral approximation to the force of infection field. The methods used are standard in the theory of spectral projection. However I reveal an interesting interplay between error induced by approximating the image of infecteds using Gaussians, the 'blurry' image approximation, and the aliasing error[1] of the pseudo-spectral projection. That is that

[1] Aliasing error is the commonly used term for error due to using a sum approximation for the Fourier coefficients. The term originates from high frequency waves becoming indistinguishable

for a given set of approximation and epidemiological parameters there is an optimal choice of Gaussian approximation width ϵ^* which minimises the uniform error between the force of infection field and its pseudo-spectral projection.

C.1 Gaussian approximation and the Pseudo-Spectral Projection as a Convolution sum

One potential difficulty is that although in theory a well-defined Fourier series exists for I, the numerical performance of the discrete transform FFT is very poor in this situation. Essentially, the delta distributions in the infection image can be expressed as,

$$\delta(\boldsymbol{i} - q) = \frac{1}{l^d} \sum_{\omega \in \mathbb{Z}^d} e^{2\pi i \omega \cdot (\boldsymbol{i} - q)/l}. \tag{C.1}$$

However any finite truncation of this sum does not provide a good approximation to the delta distribution. I resolved with this problem by introducing *blurred images*, where the delta distributions in the image definition are replaced by tight Gaussians of width ϵ (> 0),

$$\delta_\epsilon(x) = \frac{1}{(2\pi\epsilon^2)^{d/2}} e^{-\frac{-|x|^2}{2\epsilon^2}}. \tag{C.2}$$

This approximation is justified by the convergence,

$$\epsilon \to 0_+ \implies \delta_\epsilon(x - x_i) \to \delta(x - x_i).$$

Where $\to$ is weak convergence in distribution, that is for any smooth test function ϕ,

$$\lim_{\epsilon \to 0_+} \int \phi(x) \delta_\epsilon(x - x_i) dx = \phi(x_i). \tag{C.3}$$

Assuming that $\epsilon \ll l$ the Fourier coefficients for the blurring Gaussians is given by factorising ,over the d spatial dimensions and completing the square in the exponent,

$$\hat{\delta}_\epsilon(\omega) = \frac{1}{(2\pi\epsilon^2)^{d/2}} \int_{\mathcal{A}} e^{-\frac{-|x|^2}{2\epsilon^2}} e^{-2\pi i \omega \cdot x} dx = e^{-2\pi^2 \omega^2 \epsilon^2 / l^2}. \tag{C.4}$$

from or an 'alias' of lower frequency waves on a discrete grid.

The image of infecteds with delta distributions approximated by Gaussians is denoted

$$\begin{aligned} I^{\epsilon}(x,t) &= \sum_{j=1}^{N} \delta_{\epsilon}(x-x_i) I_j(t) \\ &= \frac{1}{l^d} \sum_{\omega \in \mathbb{Z}^d} \sum_{j=1}^{N} I_j(t) e^{-2\pi^2\omega^2\epsilon^2/l^2} e^{2\pi i \omega \cdot (x-x_j)/l}. \end{aligned} \tag{C.5}$$

Note that the effect of the Gaussian approximation on the image of infecteds Fourier representation is to decrease each Fourier coefficient by a factor $e^{-2\pi^2\omega^2\epsilon^2/l^2}$. This causes the sum over wave vectors to converge, by filtering out the effect of the high frequency contributions from the delta distributions.

As mentioned in section 4.3 the natural sum approximation of the force of infection field (4.15) is, in fact, identical to the pseudo-spectral projection of the force of infection field using Gaussian approximations for the delta distributions of habitat locations. Applying definition (4.25) to the pseudo-spectral projection (4.27) with the Fourier coefficients of the image of infected replaced by their 'blurry' Gaussian approximation gives that,

$$\begin{aligned} \lambda(\boldsymbol{i},t) &\approx (\Delta x)^d \sum_{\boldsymbol{j} \in \Phi_{\Delta x}^{l,d}} K(\{\boldsymbol{i}-\boldsymbol{j}\}_l) I^{\epsilon}(\boldsymbol{j}) \\ &= \frac{(\Delta x)^d}{m^{2d}(\Delta x)^{2d}} \sum_{\boldsymbol{j} \in \Phi_{\Delta x}^{l,d}} \sum_{\omega \in \Omega_M^d} \sum_{\omega' \in \Omega_M^d} \tilde{K}(\omega) \tilde{I}^{\epsilon}(\omega',t) e^{2\pi i [\boldsymbol{i}\cdot\omega + \boldsymbol{j}\cdot(\omega'-\omega)]/(m\Delta x)} \\ &= \frac{1}{l^d} \sum_{\omega \in \Omega_M^d} \tilde{K}(\omega) \tilde{I}^{\epsilon}(\omega,t) e^{2\pi i \omega \cdot \boldsymbol{i}/l} \\ &= \lambda_{PS}^{M,\epsilon}(\boldsymbol{i},t), \qquad \boldsymbol{i} \in \Phi_{\Delta x}^{l,d},\ t \in \mathbb{R}_+. \end{aligned} \tag{C.6}$$

Where I have used that for $\boldsymbol{i} \in \Phi_{\Delta x}^{l,d}$, $\boldsymbol{i}/\Delta x \in \mathbb{Z}^d$, by construction. Now, by considering mth roots of unity I get the analogous discrete orthogonality relationship to (4.23),

$$\sum_{\boldsymbol{j} \in \Phi_{\Delta x}^{l,d}} e^{2\pi i (\boldsymbol{j}/\Delta x)\cdot(\omega'-\omega)/m} = m^d \delta_{\omega', \omega+mp}, \qquad p \in \mathbb{Z}^d. \tag{C.7}$$

Note that this orthogonality relation is periodic, it does not distinguish between the equivalence classes of the form $[\omega]_m = \{\omega + m\boldsymbol{n} | \boldsymbol{n} \in \mathbb{Z}^d\}$.

It is the discrete projection $\lambda_{PS}^{M,\epsilon}$ that I can solve for each point $\boldsymbol{i} \in \Phi_{\Delta x}^{l,d}$ with computation cost $\mathcal{O}(M \ln M)$ using the FSR algorithm described in section 4.3.

C.2 Convergence of Pseudo-Spectral Approximation

Having described the grid based approximation underlying the FSR method in terms of Fourier coefficients I turn to a detailed error analysis. There are three sources of error in the approximation scheme $\lambda \approx \lambda_{PS}^{M,\epsilon}$ on the grid $\Phi_{\Delta x}^{l,d}$.

- *Aliasing error* due to FFT being a discrete transformation with a quadrature on discrete spatial points replacing an integral on continuous space.
- *Projection error* due to projecting K and I^ϵ onto the subspace $L_M^2(\mathcal{A})$ constructed using $M < \infty$ plane waves.
- The approximation of delta distributions by Gaussians.

The smoothness of the force of infection field, λ, is guaranteed by the smoothness of the transmission kernel. This in turn guarantees the existence of the spectral projection of the force of infection field (λ_S^M) and its counterpart with delta distributions replaced by approximating Gaussians,

$$\lambda_S^{M,\epsilon}(x,t) = \frac{1}{l^d} \sum_{\omega \in \Omega_M^d} \hat{K}(\omega) \hat{I}^\epsilon(\omega,t) e^{2\pi i \omega \cdot q/l} \tag{C.8}$$

I can decompose the uniform error between the true force of infection field and its pseudo-spectral projection using Gaussian approximation into the useful upper bound,

$$\|\lambda_{PS}^{M,\epsilon} - \lambda\|_\infty \leq \|\lambda_{PS}^{M,\epsilon} - \lambda_S^{M,\epsilon}\|_\infty + \|\lambda_S^{M,\epsilon} - \lambda_S^M\|_\infty + \|\lambda_S^M - \lambda\|_\infty. \tag{C.9}$$

Where $\|f\|_\infty = \sup_{x \in \mathcal{A}}\{|f(x)|\}$. Each term on the right hand side of above is respectively, the aliasing error, the Gaussian approximation error and the spectral error. I consider each case separately. The projection error is given by a standard result, however the error analysis for the error due to aliasing and Gaussian approximation is more involved. I make the general point that the smaller ϵ the better for the Gaussian approximation, but the less support $\lambda_{PS}^{M,\epsilon}$ has on $\Phi_{\Delta x}$ and hence, for fixed Δx, the worse the aliasing error. This suggests that for a given resolution Δx care should be taken to choose ϵ such that (C.9) is small.

Projection Error

Due to the smoothness of the force of infection field, it is well known that I have convergence as $M \to \infty$ for the spectral projection λ_S^M to λ [cite]. Moreover, the rate of convergence is exponential,

$$\|\lambda_S^M(\cdot,t) - \lambda(\cdot,t)\|_\infty \sim Ce^{-\alpha M}, \tag{C.10}$$

for some $C > 0$ and $\alpha > 0$ dependent on the smoothness of λ. In practise this is the least significant part of (C.9).

Aliasing Error

Recall that in each dimension I implement FFT on $m = (l/\Delta x)$ discrete frequencies, i.e. $M = (l/\Delta x)^d = m^d$. A classic manipulation is the decomposition (Hesthaven, Gottlieb, and Gottlieb 2007; Xiu 2010) using (4.16) and (4.25),

$$\begin{aligned}
\tilde{\lambda}(\omega,t) &= (\Delta x)^d \sum_{\boldsymbol{j}\in\Phi_{\Delta x}^{l,d}} \lambda(\boldsymbol{j}) e^{-2\pi i\omega\cdot(\boldsymbol{j}/\Delta x)/m} \\
&= \frac{(\Delta x)^d}{l^d} \sum_{\boldsymbol{j}\in\Phi_{\Delta x}^{l,d}} \sum_{\omega'\in\mathbb{Z}^d} \hat{\lambda}(\omega',t) e^{2\pi i\omega'\cdot(\boldsymbol{j}/\Delta x)/m} e^{-2\pi i\omega\cdot(\boldsymbol{j}/\Delta x)/m} \\
&= m^d \frac{(\Delta x)^d}{l^d} \sum_{\omega'\in\mathbb{Z}^d} \hat{\lambda}(\omega',t)\delta_{\omega',\omega+mp}, \qquad p\in\mathbb{Z}^d, \\
&= \hat{\lambda}(\omega,t) + \sum_{p\neq 0} \hat{\lambda}(\omega+mp).
\end{aligned} \tag{C.11}$$

Where $\hat{f}(\omega)$ is the Fourier coefficient for f at discrete frequency $\omega \in \mathbb{Z}^d$. This relates the true Fourier coefficient to its FFT estimate at each frequency. The second term in the sum is the aliasing error induced by the indistinguishability of modes within the equivalence classes $[\omega]_m$.

I make the reasonable assumption that the transmission kernel has a known Fourier series $\hat{K}$, or has negligible aliasing error; recall that $\hat{K}$ only requires calculation once hence an arbitrarily accurate numerical scheme can be used without drawing significant computation effort. Using equation (C.11) the aliasing error, for

a given state $X(t)$ and $x \in \mathcal{A}$, can be written as,

$$\begin{aligned} |(\lambda_{PS}^{M,\epsilon} - \lambda_S^{M,\epsilon})(x,t)| &= \frac{1}{l^d}\Big| \sum_{\omega \in \Omega_M^d} \hat{K}(\omega)(\tilde{I}^\epsilon(\omega) - \hat{I}^\epsilon(\omega))e^{2\pi i\omega\cdot x/l}\Big| \\ &= \frac{1}{l^d}\Big| \sum_{\omega \in \Omega_M^d} \hat{K}(\omega) \sum_{j=1}^{N} I_j(t) e^{2\pi i\omega\cdot(x-x_j)/l} \\ &\quad \times \sum_{p\neq 0} e^{-2\pi i m p\cdot(\frac{x_j}{\Delta x})/m} e^{-2\pi^2(\omega+mp)^2/r^2m^2}\Big| \end{aligned} \tag{C.12}$$

Where

$$r = \Delta x/\epsilon$$

is the ratio between grid width and Gaussian width. I emphasis that the Gaussian approximation to the delta distribution was necessary for good numerical performance since the sum $\sum_{p\neq 0} e^{-2\pi i p\cdot(\frac{x_j}{\Delta x})}$ does not converge, causing significant aliasing error in the pseudo-spectral approach to force of infection calculation with $\epsilon = 0$, however big M is allowed to become. However, the sum $\sum_{p\neq 0} e^{-2\pi^2(\omega+mp)^2/r^2m^2}$ converges for all $r > 0$.

An obvious progression from (C.12) is to use the Cauchy-Schwarz inequality, however this leads to rather crude estimates for the aliasing error for a given choice of Δx, ϵ. Instead, I exploit the special structure of the force of infection field's spectral representation. Firstly, that $(\lambda_{PS}^{M,\epsilon} - \lambda_S^{M,\epsilon})$ is a real valued function, and secondly that for each wave-vector $\omega \in \mathbb{Z}^d$, the Fourier coefficient of the transmission kernel $\hat{K}(\omega)$ is also real-valued. To see this I use the even symmetry of the definition $K(x_i - x_j) = K(|x_i - x_j|)$,

$$\begin{aligned} \hat{K}(\omega) &= \int_{\mathcal{A}} K(x) e^{-2\pi i\omega\cdot q/l} dx, \qquad \omega \in \mathbb{Z}^d \\ &= \int_{\mathcal{A}} K(x)\cos(2\pi\omega\cdot q/l)dx - i\int_{\mathcal{A}} K(x)\sin(2\pi\omega\cdot q/l)dx \\ &= \int_{\mathcal{A}} K(x)\cos(2\pi\omega\cdot q/l)dx \implies \hat{K}(\omega) \in \mathbb{R}, \qquad \omega \in \mathbb{Z}^d. \end{aligned}$$

The last line exploits the odd symmetry of the sine function, and the even symmetry

of K, recalling that $\mathcal{A}$ is a zero-centred space. Using this result in (C.12) gives that,

$$\begin{aligned}(\lambda_{PS}^{M,\epsilon} - \lambda_{S}^{M,\epsilon})(x,t) &= \frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\sum_{p\neq 0}e^{-2\pi^2(\omega+mp)^2/r^2m^2} \\ &\quad\times\cos(2\pi[\omega\cdot x-(\omega+mp)\cdot x_j]/l) \\ &\quad+i\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\sum_{p\neq 0}e^{-2\pi^2(\omega+mp)^2/r^2m^2} \\ &\quad\times\sin(2\pi[\omega\cdot x-(\omega+mp)\cdot x_j]/l) \\ \implies \quad &\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\sum_{p\neq 0}e^{-2\pi^2(\omega+mp)^2/r^2m^2} \\ &\quad\times\sin(2\pi[\omega\cdot x-(\omega+mp)\cdot x_j]/l)=0. \qquad \text{(C.13)}\end{aligned}$$

Where I have separated real and imaginary parts. This allows a tighter assessment of the point-wise error between the spectral and pseudo-spectral projections of the force of infection field.

$$\begin{aligned}\left|(\lambda_{PS}^{M,\epsilon} - \lambda_{S}^{M,\epsilon})(x,t)\right| &= \Big|\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\sum_{p\neq 0}e^{-2\pi^2(\omega+mp)^2/r^2m^2} \\ &\quad\times\cos(2\pi\omega\cdot(x-x_j)/l-2\pi(p\cdot x_j)/\Delta x)\Big| \qquad \text{(C.14)}\end{aligned}$$

I note that on the periodic domain $\mathcal{A}$ the contributing transmission range is, $(x - x_j)^* = \{x - x_j\}_l \leq l/2$. Therefore, $\cos(2\pi\omega\cdot(x-x_j)/l) \geq 0$ for all x. Hence, expression (C.14) obtains its maximum under variation of habitat locations when each habitat is located on the grid $\Phi_{\Delta x}^{l,d}$ when

$$\cos(2\pi\omega\cdot(x-x_j)/l-2\pi(p\cdot x_j)/\Delta x)=\cos(2\pi\omega\cdot(x-x_j)/l), \qquad p\in\mathbb{Z}^d.$$

Other habitat arrangements will lead to negative signed terms in the sum over the

aliasing wave-vectors. This permits me to develop the following inequality,

$$\begin{aligned}
\Big|(\lambda_{PS}^{M,\epsilon} - \lambda_{S}^{M,\epsilon})(x,t)\Big| &\leq \Big|\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\sum_{p\neq 0}e^{-2\pi^2(\omega+mp)^2/r^2m^2} \\
&\quad \times\cos(2\pi\omega\cdot(x-x_j)/l)\Big| \\
&= \frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\sum_{p\neq 0}e^{-2\pi^2(\omega+mp)^2/r^2m^2} \\
&\quad \times\cos(2\pi\omega\cdot(x-x_j)/l) \\
&\leq \Big(\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\cos(2\pi\omega\cdot(x-x_j)/l)\Big) \\
&\quad \times\Big(\sum_{p\neq 0}e^{-2\pi^2(\omega^*+mp)^2/r^2m^2}\Big) \\
&\leq \|\lambda_S^M\|_\infty\Big(\sum_{p\neq 0}e^{-2\pi^2(\omega^*+mp)^2/r^2m^2}\Big) \\
&\leq \|\lambda\|_\infty\Big(\sum_{p\neq 0}e^{-2\pi^2(\omega^*+mp)^2/r^2m^2}\Big). \qquad \text{(C.15)}
\end{aligned}$$

For the second line I have used that the right hand side of the inequality is positive. For the third line ω^* is the wave-vector which maximises the positive bounded function $\alpha(\omega) = \sum_{p\neq 0} e^{-2\pi^2(\omega+mp)^2/r^2m^2}$ over Ω_M^d. For the last line I use the real-valued non-negativity of each Fourier coefficient $\hat{K}(\omega)$, $\omega\in\mathbb{Z}^d$,

$$\begin{aligned}
\|\lambda_S^M\|_\infty &= \frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\cos(2\pi\omega\cdot(x^*-x_j)/l), \qquad x^*\in\arg\max_{x\in\mathcal{A}}\{\lambda_S^M(x,t)\}, \\
&\leq \frac{1}{l^d}\sum_{\omega\in\mathbb{Z}^d}\hat{K}(\omega)\sum_{j=1}^{N}I_j(t)\cos(2\pi\omega\cdot(x^*-x_j)/l) \\
&\leq \|\lambda\|_\infty.
\end{aligned}$$

Result (C.15) allows us to use the Gaussian approximation to control the aliasing error by tuning the ϵ width parameter, since I have for $\omega^*\in\Omega_M^d$,

$$\begin{aligned}
\sum_{p\neq 0}e^{-2\pi^2(\omega^*+mp)^2/r^2m^2} &= \sum_{p\neq 0}e^{-2\pi^2(\frac{\omega^*}{m}+p)^2/r^2} \\
&= \sum_{p\in\mathbb{Z}^d}e^{-2\pi^2(\frac{\omega^*}{m}+p)^2/r^2} - e^{-2\pi^2(\frac{\omega^*}{m})^2/r^2}. \qquad \text{(C.16)}
\end{aligned}$$

I write the wave-number in vector form $\omega^* = (\omega_1^*, \ldots, \omega_d^*)$ and split the sum in (C.16) into a product of sums over the integers,

$$\sum_{p\in\mathbb{Z}^d} e^{-2\pi^2(\frac{\omega^*}{m}+p)^2/r^2} = \prod_{i=1}^{d}\Big(\sum_{p_i=-\infty}^{\infty} e^{-2\pi^2(\frac{\omega_i^*}{m}+p_i)^2/r^2}\Big). \tag{C.17}$$

I recall that each component of the wave-number has been computed for $-m/2 \leq \omega_i \leq m/2-1$, and hence $|\omega_i^*/m| \leq 1/2$ for each term in (C.12). Our aim is to provide an upper bound for (C.16). Without loss of generality I can in fact take $-1/2 \leq \omega_i^*/m \leq 0$, if $\omega^* > 0$ then reversing the order of summation in (C.17) has the same effect.

I recall the standard result of analysis that for non-negative monotonically decreasing functions f,

$$\sum_{n=N}^{M} f(n) \leq f(N) + \int_N^M f(x)dx.$$

I also note that for $-1/2 \leq \omega_i^*/m \leq 0$ then $f(x) = e^{-2\pi^2(x-\omega_i^*/m)^2/r^2}$ is monotonically decreasing on $[0,\infty)$ and $g(x) = e^{-2\pi^2(x+\omega_i^*/m)^2/r^2}$ is monotonically decreasing on $[1,\infty)$. For each sum in the product term in $(C.16)$ I have, via standard sum splitting and change of variable,

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} e^{-2\pi^2(n+\omega_i^*/m)^2/r^2} &= \sum_{j=0}^{\infty} e^{-2\pi^2(j-\omega_i^*/m)^2/r^2} + \sum_{l=1}^{\infty} e^{-2\pi^2(l+\omega_i^*/m)^2/r^2} \\
&\leq e^{-2\pi^2(\omega_i^*/m)^2/r^2} + e^{-2\pi^2(1+\omega_i^*/m)^2/r^2} \\
&\quad + \int_0^{\infty} e^{-2\pi^2(x-\omega_i^*/m)^2/r^2}dx \\
&\quad + \int_1^{\infty} e^{-2\pi^2(x+\omega_i^*/m)^2/r^2}dx.
\end{aligned} \tag{C.18}$$

The integrals in (C.18) have unnormalised Gaussian probability density functions

as their integrand. This permits the following inequality,

$$\begin{aligned}
&\int_0^\infty e^{-2\pi^2(x-\omega_i^*/m)^2/r^2}dx + \int_1^\infty e^{-2\pi^2(x+\omega_i^*/m)^2/r^2}dx \\
&\qquad = \frac{r}{\sqrt{2\pi}}\Big[1 - \frac{\sqrt{2\pi}}{r}\int_{-1}^0 e^{-2\pi^2(x-\omega_i^*/m)^2/r^2}dx\Big] \\
&\leq \frac{r}{\sqrt{2\pi}}\Big[1 - \frac{\sqrt{2\pi}}{r}\int_{-1}^0 e^{-2\pi^2x^2/r^2}dx\Big], \qquad -1/2 \leq \omega_i^* \leq 0, \\
&\qquad = \frac{1}{r\sqrt{2\pi}}\mathbb{P}(Y \geq 0 \ \vee \ Y \leq -1).
\end{aligned} \tag{C.19}$$

Where the random variable $Y \sim \mathcal{N}(0, r^2/4\pi^2)$. I can use Chebyshev's inequality and the symmetry of the Gaussian distribution to bound the probability above in terms of r.

$$\begin{aligned}
\mathbb{P}(Y \geq 0 \ \vee \ Y \leq -1) = 1/2 + \mathbb{P}(Y \leq -1) &= 1/2(1 + \mathbb{P}(|Y| \geq 1)) \\
&\leq 1/2\Big(1 + \frac{r^2}{4\pi^2}\Big).
\end{aligned} \tag{C.20}$$

Introducing (C.17), (C.18), (C.19) and (C.20) into (C.16) gives,

$$\begin{aligned}
\sum_{p\in\mathbb{Z}^d} e^{-2\pi^2(\frac{\omega^*}{m}+p)^2/r^2} - e^{-2\pi^2(\frac{\omega^*}{m})^2/r^2} &\leq \prod_{i=1}^d \Big(e^{-2\pi^2(\omega_i^*/m)^2/r^2} + e^{-2\pi^2(1+\omega_i^*/m)^2/r^2} \\
&\quad + \frac{r^3 + 4\pi^2 r}{2(2\pi)^{5/2}}\Big) - e^{-2\pi^2(\frac{\omega^*}{m})^2/r^2} \\
&\leq \prod_{i=1}^d \Big(e^{-2\pi^2(\omega_i^*/m)^2/r^2} + e^{-\pi^2/2r^2} \\
&\quad + \frac{r^3 + 4\pi^2 r}{2(2\pi)^{5/2}}\Big) - e^{-2\pi^2(\frac{\omega^*}{m})^2/r^2}, \\
&\leq \Big(1 + e^{-\pi^2/2r^2} + \frac{r^3 + 4\pi^2 r}{2(2\pi)^{5/2}}\Big)^d - 1.
\end{aligned} \tag{C.21}$$

Where the above holds for $-1/2 \leq \omega_i^* \leq 0$. For compactness of representation I now write,

$$\psi = 1 + e^{-\pi^2/2r^2} + \frac{r^3 + 4\pi^2 r}{2(2\pi)^{5/2}}.$$

Introducing this estimate into equation (C.12) gives an upper bound for the error due to aliasing,

$$\begin{aligned} |(\lambda_{PS}^{M,\epsilon} - \lambda_S^{M,\epsilon})(x,t)| &\leq (\psi^d - 1)\Big|\frac{1}{l^d}\sum_{\omega\in\boldsymbol{\omega}_M^d}\hat{K}(\omega)\sum_{j=1}^{N} I_j(t)e^{2\pi i\omega\cdot(q-x_j)/l}\Big| \\ \implies \|\lambda_{PS}^{M,\epsilon} - \lambda_S^{M,\epsilon}\|_\infty &\leq (\psi^d - 1)\|\lambda\|_\infty \end{aligned} \tag{C.22}$$

Gaussian Approximation Error

For the uniform Gaussian approximation error I can exploit the definition of the exponential function and use a standard result for spectral representation of derivatives of smooth functions, f,

$$\Delta^k f = \frac{1}{l^d}\sum_{\omega\in\mathbb{Z}^d}\frac{(-4\pi^2\omega^2)^k}{l^{2k}}\hat{f}(\omega)e^{2\pi i\omega\cdot q/l}.$$

Where Δ^k is the k-fold operator mapping on suitably regular functions f of the Laplacian ($\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_{(i)}^2}$ where $x = (x_{(1)},\ldots,x_{(d)})$),

$$\Delta^k f = \Delta(\Delta^{k-1}f) = \cdots = \Delta(\Delta(\ldots\Delta(\Delta f)\ldots).$$

Hence, by using the definition of the spectral projection with Gaussian approximation I can write, denoting $x^* \in \arg\max_{x\in\mathcal{A}}\{|(\lambda_S^{M,\epsilon} - \lambda_S^M)(x,t)|\}$,

$$\begin{aligned} \|\lambda_S^{M,\epsilon} - \lambda_S^M\|_\infty &= \Big|\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)[\hat{I}^\epsilon(\omega) - \hat{I}(\omega)]e^{2\pi i\omega\cdot x^*/l}\Big|, \\ &= \Big|\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N} I_j(t)[e^{-2\pi^2\omega^2\epsilon^2/l} - 1]e^{2\pi i\omega\cdot(x^*-x_j)/l}\Big| \\ &= \Big|\frac{1}{l^d}\sum_{\omega\in\Omega_M^d}\hat{K}(\omega)\sum_{j=1}^{N} I_j(t)\sum_{k\geq 1}\frac{(-4\pi^2\omega^2/l^2)^k\epsilon^{2k}}{2^k k!}e^{2\pi i\omega\cdot(x^*-x_j)/l}\Big| \\ &\leq \Big\|\sum_{k\geq 1}\frac{\epsilon^{2k}}{2^k k!}\Delta^k\lambda\Big\|_\infty \end{aligned} \tag{C.23}$$

Since I am concerned with the regime where the width parameter ϵ is small I will use the estimate for (C.23),

$$\|\lambda_S^{M,\epsilon} - \lambda_S^M\|_\infty \leq \frac{\epsilon^2}{2}\|\Delta\lambda\|_\infty + \mathcal{O}(\epsilon^4) \tag{C.24}$$

C.3 Combined Error and choice of ϵ:

Putting our three estimates into the decomposed uniform error (C.9) gives the result in the main work,

$$\|\lambda_{PS}^{M,\epsilon} - \lambda\|_\infty \leq \frac{\epsilon^2}{2}\|\Delta\lambda\|_\infty + (\psi^d - 1)\|\lambda\|_\infty + Ce^{-\alpha M} + \mathcal{O}(\epsilon^4). \tag{C.25}$$

The error factor (Ψ) referred to in equation 4.31 is,

$$\Psi = \psi^d - 1. \tag{C.26}$$

For a given state $X(t)$ the worst possible configuration, in terms of maximising the RHS of (C.25), is to have each habitat located on exactly the same coordinate, this gives a worst possible error of,

$$\|\lambda_{PS}^{M,\epsilon} - \lambda\|_\infty \leq I(t)\Big(\frac{\epsilon^2}{2}|\Delta K(0)| + (\psi^d - 1)K(0)\Big) + Ce^{-\alpha M} + \mathcal{O}(\epsilon^4). \tag{C.27}$$

Where $I(t) = \sum_{j=1}^{N} I_j(t)$. One approach to choosing ϵ, given some grid resolution Δx, is to minimise the maximum error in this worst possible scenario; that is to choose $\epsilon = \epsilon^*$ where

$$\epsilon^* \in \arg\min_{\epsilon>0} \Big\{\frac{\epsilon^2}{2}|\Delta K(0)| + (\psi^d - 1)K(0)\Big\}. \tag{C.28}$$

This was the approach that was used successfully in the numerical experimentation section, see Figure C.1 for a plot of $f(\epsilon) = \frac{\epsilon^2}{2}|\Delta K(0)| + (\psi^d - 1)K(0)$ as a function of ϵ for a few choices of Δx. For the baseline case considered in this work, that is Gaussian shaped transmission kernel of width $L = 3$, this consists of picking ϵ such that it was not too small. For more local transmission, e.g. $L = 1$, a distinct minimum appears in $f(\epsilon)$ which guides the choice of ϵ. In all cases smaller values of Δx allow smaller values of ϵ to be chosen.

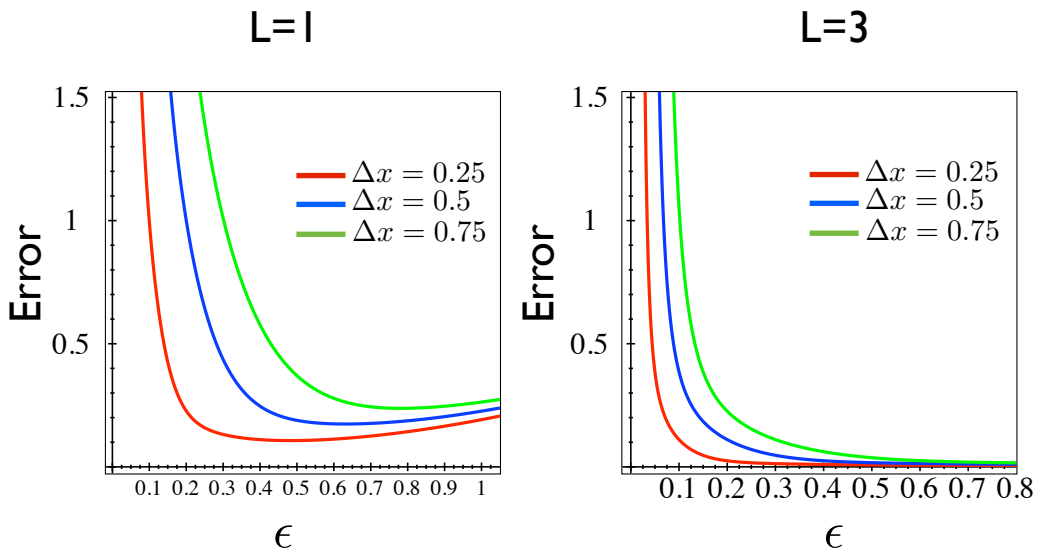


Figure C.1: Plots of the 'worst case' error $f(\epsilon) = \frac{\epsilon^2}{2}|\Delta K(0)| + (\psi^d - 1)K(0)$ against width of Gaussian approximation ϵ for $d = 2$. Curves from left to right are for increasing $\Delta x = 0.25, 0.5, 0.75$. The transmission kernel, K, is chosen to be Gaussian shaped with width L. *Left*: For transmission length scale $L = 1$. I note that for this very local transmission range a distinctive minimiser ϵ^* exists. *Right*: L = 3. Due to increased smoothness of the transmission kernel $f(\epsilon)$ is close to zero for a wide range of ϵ. Any choice of ϵ^* not too small is acceptable.

Bibliography

Abramowitz, M, and I Stegun. 1964. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York

Ali, A, and Grosskinsky, S. 2010. Pattern Formation Through Genetic Drift at Expanding Population Fronts. Advances in Complex Systems 13 (03): 349-366.

Alonso, D, McKane, AJ, and Pascual, M. 2007. Stochastic Amplification in Epidemics. Journal of the Royal Society Interface 4 (14): 575.

Anderson, R M, and May, RM. 1983. Vaccination Against Rubella and Measles: Quantitative Investigations of Different Policies. Journal of Hygiene 90 (02) (May 15): 259-325.

Anderson, R M, and May, RM. 1985a. Vaccination and Herd Immunity to Infectious Diseases. Nature, 318: 323 - 329 .

Anderson, R M, and May, RM. 1985b. Age-Related Changes in the Rate of Disease Transmission: Implications for the Design of Vaccination Programmes. Journal of Hygiene 94: 365-436.

Anderson, R M, May, R M. 1992. Infectious Diseases of Humans: Dynamics and Control. Oxford: Oxford University Press.

Applebaum, David. 2004. Lévy Processes and Stochastic Calculus. Cambridge University Press.

Baguelin, M, Van Hoek, AJ, Jit,M, Flasche, S, White, PJ, and Edmunds, WJ. 2010. Vaccination Against Pandemic Influenza a/H1N1v in England: a

Real-Time Economic Evaluation. Vaccine 28 (12) (March 11): 2370-2384.

Bailey, N T J. 1950. A Simple Stochastic Epidemic. Biometrika: 37 (3-4):193-202.

Bailey, N T J. 1975. The Mathematical Theory of Infectious Diseases and Its Applications (2nd Edition). Charles Griffin & Co Ltd.

Bain, A, and D Crisan. 2009. Fundamentals of Stochastic Filtering. UK: Springer.

Ball, F. 1983. The Threshold Behaviour of Epidemic Models. Journal of Applied Probability: 227-241.

Ball, F, and P Neal. 2002. A General Model for Stochastic SIR Epidemics with Two Levels of Mixing. Mathematical Biosciences 180 (1-2): 73-102.

Ball, F, D Sirl, and P Trapman. 2010. Analysis of a Stochastic SIR Epidemic on a Random Network Incorporating Household Structure. Mathematical Biosciences 224: 53-73.

Ball, F, T Britton, and O Lyne. 2004. Stochastic Multitype Epidemics in a Community of Households: Estimation and Form of Optimal Vaccination Schemes. Mathematical Biosciences, 191(1):19-40.

Barrett, C.L., Bisset K.R. , Eubank S.G. , Feng X., and Marathe M.V. 2008. Episimdemics: an Efficient Algorithm for Simulating the Spread of Infectious Disease Over Large Realistic Social Networks. Proceedings of the 2008 ACM/IEEE Conference on Supercomputing: 1-12.

Bartlett, M S. 1956. Deterministic and Stochastic Models for Recurrent Epidemics. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, University of California Press.

Bartlett, M S. 1960a. Stochastic Population Models in Ecology and Epidemiology. New York: Wiley.

Bartlett, M S. 1960b. The Critical Community Size for Measles in the

United States. Journal of the Royal Statistical Society. Series a (General), 123 (1) 37-44.

Becker, N G. 1989. Analysis of Infectious Disease Data. 1st ed. Chapman and Hall.

Behncke, H. 2000. Optimal Control of Deterministic Epidemics. Optimal Control Applications and Methods.

Bellman, R. 1954. The Theory of Dynamic Programming. Bull. Amer. Math. Soc 60 (6): 503-515.

Bellman, R. 1952. On the Theory of Dynamic Programming. Proceedings of the National Academy of Sciences of the United States of America 38 (8) (August 1): 716.

Bellman, Richard Ernest. 1953. An Introduction to the Theory of Dynamic Programming. RAND Corporation.

Bender, C, and J Steiner. 2010. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42.

Bender, C, and J Steiner. 2012. Least-Squares Monte Carlo for Backward SDEs. Numerical Methods in Finance: 257-289.

Bertsekas, Dimitri P. 1987. Dynamic Programming: Deterministic and Stochastic Models. Prentice Hall.

Bichteler, K. 2002. Stochastic Integration with Jumps - Klaus Bichteler - Google Books. USA: Cambridge University Press.

Binmore, Ken. 2007. Game Theory: a Very Short Introduction. Oxford University Press.

Bolker, B.M. 1999. Analytic Models for the Patchy Spread of Plant Disease. Bulletin of Mathematical Biology 61 (5): 845-874.

Bolker, B.M., and Pacala, S. 1997. Using Moment Equations to Understand Stochastically Driven Spatial Pattern Formation in Ecological Systems. Theoretical Population Biology 52: 179-197.

Bolker, B.M., and S W Pacala. 1999. Spatial Moment Equations for Plant Competition: Understanding Spatial Strategies and the Advantages of Short Dispersal. The American Naturalist 153 (6): 575-602.

Bouchard, B, R Elie, and N Touzi. 2009. Discrete-Time Approximation of BSDEs and Probabilistic Schemes for Fully Nonlinear PDEs. Advanced Financial Modelling 8: 91.

Brauer, F. 2009. Mathematical Epidemiology Is Not an Oxymoron. BMC Public Health.

Brown, D. 2004. The Effects of Disease Dispersal and Host Clustering on the Epidemic Threshold in Plants. Bulletin of Mathematical Biology 66 (2) (March): 341-371.

Brownlee, J. 1909. Certain Considerations on the Causation and Course of Epidemics. Proceedings of the Royal Society of Medicine.

Capinski, M, and Kopp, PE. 2007. Measure, Integral and Probability. Springer.

Clancy, D. 1999. Optimal Intervention for Epidemic Models with General Infection and Removal Rate Functions. Journal of Mathematical Biology 39 (4): 309-331.

Conlan, AJK, Rohani, P, Lloyd AL, Keeling, M, and Grenfell, BT. 2010. Resolving the Impact of Waiting Time Distributions on the Persistence of Measles. Journal of the Royal Society Interface 7 (45): 623.

Cook, AR, Gibson, GJ, Gottwald, TR, and Gilligan, CA. 2008. Constructing the Effect of Alternative Intervention Strategies on Historic Epidemics. Journal of the Royal Society Interface 5 (27): 1203-1213.

Coolen, A C C,Khn, R, and Sollich, P. 2005. Theory of Neural Informa-

tion Processing Systems. Oxford University Press, USA.

Cornell, S, and Ovaskainen, O. 2008. Exact Asymptotic Analysis for Metapopulation Dynamics on Correlated Dynamic Landscapes. Theoretical Population Biology 74: 209-225.

Dargatz, C, Georgescu V, and Held, L. 2006. Stochastic Modelling of the Spatial Spread of Influenza in Germany. Austrian Journal of Statistics 35: 5-20.

Dhersin, J, and L Decreusefond. 2011. Large Graph Limit for an SIR Process in Random Network with Heterogeneous Connectivity. Annals of Applied Probability, 22 (2): 541-575.

Di Nunno, G., B Øksendal, and F Proske. 2009. Malliavin Calculus for Lvy Processes with Applications to Finance. Springer.

Dieckmann, U, and Law, R. 2000. Relaxation Projections and the Method of Moments. The Geometry of Ecological Interactions: Symplifying Spatial Complexity (U Dieckmann, R. Law, JAJ Metz, Editors). Cambridge University Press, Cambridge: 412-455.

Diekmann, O, and J A P Heesterbeek. 2000. Mathematical Epidemiology of Infectious Diseases. Wiley.

Diggle, P, and P Ribeiro Jr. 2003. An Introduction to Model-Based Geostatistics. Spatial statistics and computational methods, Mller, J. (ed.), Springer.

Diggle, P.J., and R.K. Milne. 1983. Bivariate Cox Processes: Some Models for Bivariate Spatial Point Patterns. Journal of the Royal Statistical Society. Series B (Methodological): 11-21.

Diggle, Peter. 1983. Statistical Analysis of Spatial Point Patterns. A Hodder Arnold Publication.

Donnelly, Christl A, Rosie Woodroffe, D R COX, F John Bourne, C L Cheeseman, Richard S Clifton-Hadley, Gao Wei, et al. 2005. Positive and Negative Effects of Widespread Badger Culling on Tuberculosis in Cattle. Nature

439 (7078): 843-846.

Eames, K T D, and M J Keeling. 2003. Contact Tracing and Disease Control. Proceedings of the Royal Society B: Biological Sciences 270 (1533): 2565-2571.

Egloff, Daniel. 2005. Monte Carlo Algorithms for Optimal Stopping and Statistical Learning. The Annals of Applied Probability 15 (2) : 1396-1432.

Ellner, S P. 2001. Pair Approximation for Lattice Models with Multiple Interaction Scales. Journal of Theoretical Biology 210 (4) : 435-447.

Ethier, S, and T Kurtz. 1986. Markov Processes : Characterization and Convergence. Wiley.

Fearnhead, Paul. 2006. The Stationary Distribution of Allele Frequencies When Selection Acts at Unlinked Loci. Theoretical Population Biology 70 (3) (November): 376-386.

Feller, W. 1951. Feller: Diffusion Processes in Genetics. Proc Second Berkeley Symp Math Statist Prob.

Ferguson, N M, C A Donnelly, and R M Anderson. 2001a. Transmission Intensity and Impact of Control Policies on the Foot and Mouth Epidemic in Great Britain. Nature 413 (6855): 542-547.

Ferguson, N M, C A Donnelly, and R M Anderson. 2001b. The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions. Science 292 (5519): 1155-1160.

Fine, P E PE. 1993. Herd Immunity: History, Theory, Practice.. Epidemiology Review 15 (2) : 265-302.

Fleming, W, and H M Soner. 2006. Controlled Markov Processes and Viscosity Solutions. Springer.

Fraser, C. 2007. Estimating Individual and Household Reproduction Num-

bers in an Emerging Epidemic. PLoS One 2 (8): 758.

Friedlander, FG. 1998. Introduction to the Theory of Distributions. 2nd ed. Cambridge: Cambridge University Press.

Galvani, A P, and T C Reluga. 2007. Long-Standing Influenza Vaccination Policy Is in Accord with Individual Self-Interest but Not with the Utilitarian Optimum. Proceedings of the National Academy of Sciences, 104 (13): 5692-5697.

Gibson, G J. 1997. Investigating Mechanisms of Spatiotemporal Epidemic Spread Using Stochastic Models. Phytopathology 87 (2) : 139-146.

Gilbert, M, A Mitchell, D Bourn, J Mawdsley, R Clifton-Hadley, and W Wint. 2005. Cattle Movements and Bovine Tuberculosis in Great Britain. Nature 435 (7041): 491-496.

Gillespie, Daniel T. 1977. Exact Stochastic Simulation of Coupled Chemical Reactions. The Journal of Physical Chemistry 81 (25): 2340-2361.

Gillespie, DT. 2001. Approximate Accelerated Stochastic Simulation of Chemically Reacting Systems. The Journal of Chemical Physics 115: 1716-1733.

Gillespie, DT, and LR Petzold. 2003. Improved Leap-Size Selection for Accelerated Stochastic Simulation. The Journal of Chemical Physics 119: 8229-8234.

Grassly, Nicholas C, and Christophe Fraser. 2008. Mathematical Models of Infectious Disease Transmission. Nature Reviews Microbiology: 1-11.

Greenwood, Major. 1935. Epidemics and Crowd-Diseases : an Introduction to the Study of Epidemiology. Williams & Norgate.

Grenfell, B, O Bjørnstad, and J Kappey. 2001. Travelling Waves and Spatial Hierarchies in Measles Epidemics. Nature 414: 716-723.

Grenfell, BT, BM Bolker, and A. Kleczkowski. 1995. Seasonality and Extinction in Chaotic Metapopulations. Proceedings of the Royal Society of London. Series B. Biological Sciences 259: 97-103.

Griffiths, D A. 1973. Multivariate Birth-and-Death Processes as Approximations to Epidemic Processes. Journal of Applied Probability: 10: 15-26.

Gupta, S, J Swinton, and R M Anderson. 1994. Theoretical Studies of the Effects of Heterogeneity in the Parasite Population on the Transmission Dynamics of Malaria. Proceedings of the Royal Society B: Biological Sciences 256 (1347): 231-238.

Hanski, I. 199. a Practical Model of Metapopulation Dynamics. Journal of Animal Ecology 63 (1): 151-162.

Hanski, I, and OE Gaggiotti. 2004. Ecology, Evolution and Genetics of Metapopulations. New York: Academic Press.

Hanski, Ilkka A, and Michael E Gilpin, eds. 1997. Metapopulation Biology: Ecology, Genetics, and Evolution. 1st ed. New York: Oxford University Press.

Heino, Mikko, and Ilkka Hanski. 2001. Evolution of Migration Rate in a Spatially Realistic Metapopulation Model. The American Naturalist 157 (5): 495-511.

Herbert, Julian, and Valerie Isham. 2000. Stochastic Host-Parasite Interaction Models. Journal of Mathematical Biology 40 (4) : 343-371.

Hesthaven, JS, S Gottlieb, and David Gottlieb. 2007. Spectral Methods for Time-Dependent Problems. Cambridge: Cambridge University Press.

Hethcote, Herbert W, and Paul Waltman. 1973. Optimal Vaccination Schedules in a Deterministic Epidemic Model. Mathematical Biosciences 18 (3-4) : 365-381.

House, T, M Baguelin, A J Van Hoek, P J White, Z Sadique, K Eames, J M Read, et al. 2011. Modelling the Impact of Local Reactive School Closures on Critical Care Provision During an Influenza Pandemic. Proceedings of the Royal Society B: Biological Sciences 278 (1719): 2753-2760.

Ionides, EL, C Bret, and AA King. 2006. Inference for Nonlinear Dynamical Systems. Proceedings of the National Academy of Sciences 103 (49): 18438-18443.

Isham, V. 1995. Stochastic Models of Host-Macroparasite Interaction. The Annals of Applied Probability: 5 (3), 720-740.

Isham, V. 2005. Stochastic Models for Epidemics: Current Issues and Developments.Celebrating Statistics 33: 27-54.

Ito, Kiyosi. 1951. On Stochastic Differential Equations. Memoirs, American Mathematical Society.

Jansen, V A A. 2003. Measles Outbreaks in a Population with Declining Vaccine Uptake. Science 301 (5634): 804-804..

Kallenberg, O. 2002. Foundations of Modern Probability Berlin: Springer.

Karatzas, Ioannis, and Steven E Shreve. 1991. Brownian Motion and Stochastic Calculus. Springer Verlag.

Keeling, M J. 2000. Metapopulation Moments: Coupling, Stochasticity and Persistence. Journal of Animal Ecology 69 (5): 725-736.

Keeling, M J, O N Bjornstad, and B T Grenfell. Metapopulation Dynamics of Infectious Diseases. In Ecology, Genetics and Evolution of Metapopulations, ed. I Hanski and OE Gaggiotti. Elsevier.

Keeling, Matt J, and Pejman Rohani. 2002. Estimating Spatial Coupling in Epidemiological Systems: a Mechanistic Approach. Ecology Letters 5 (1) : 20-29.

Keeling, Matt J, and Pejman Rohani. 2008. Modelling Infectious Diseases in Humans and Animals. Princeton University Press.

Keeling, Matt J, and Peter J White. 2011. Targeting Vaccination Against Novel Infections: Risk, Age and Spatial Structure for Pandemic Influenza in Great Britain. Journal of the Royal Society Society interface 8 (58) 661-670.

Keeling, MJ. 1999. The Effects of Local Spatial Structure on Epidemiological Invasions. Proceedings: Royal Society of London B (Biological Sciences) 266 (1421): 859-867.

Keeling, MJ, MEJ Woolhouse, DJ Shaw, L Matthews, M Chase-Topping, DT Haydon, SJ Cornell, J Kappey, J Wilesmith, and BT Grenfell. 2001. Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. Science 294 (5543): 813-817.

Keeling, MJ, MEJ Woolhouse, RM May, G Davies, and BT Grenfell. 2003. Modelling Vaccination Strategies Against Foot-and-Mouth Disease. Nature 421 (6919): 136-142.

Kermack, W O, and A G McKendrick. 1927. A Contribution to the Mathematical Theory of Epidemics. Proc. Roy. Soc. Lond. A 115 (772): 700-721.

Kesten, H. 1958. Accelerated Stochastic Approximation. The Annals of Mathematical Statistics 29 (1): 41-59.

King, Aaron A, Edward L Ionides, Mercedes Pascual, and Menno J Bouma. 2008. Inapparent Infections and Cholera Dynamics. Nature 454 (7206) : 877-880.

Kiss, I Z, D M Green, and R R Kao. 2006. The Network of Sheep Movements Within Great Britain: Network Properties and Their Implications for Infectious Disease Spread. Journal of the Royal Society Interface 3 (10): 669-677.

Klebaner, Fima C. 2005. Introduction to Stochastic Calculus with Applications. 2nd ed. Imperial College Press.

Klepac, P, R Laxminarayan, and B T Grenfell. 2011. Synthesizing Epidemiological and Economic Optima for Control of Immunizing Infections. Proceedings of the National Academy of Sciences of the United States of America 108 (34): 14366-14370.

Kurtz, T. 1970. Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes. Journal of Applied Probability 7: 49-58.

Kurtz, T. 1971. Limit Theorems for Sequences of Jump Markov Processes Approximating Ordinary Differential Processes. Journal of Applied Probability 8: 344-356.

Kurtz, T. 2007. Large Deviations for Stochastic Processes (Mathematical Survey & Monographs, N 131). Lavoisier.France.

Kushner, Harold J, and George Yin. 2003. Stochastic Approximation and Recursive Algorithms and Applications. Springer.

Lefévre, C. 1981. Optimal Control of a Birth and Death Epidemic Process. Operations Research 29 (5): 971-982.

Levins, Richard. 1969. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. Bull. Entomol. Soc. Am. 15: 237-240.

Liggett, T M. 1985. Interacting Particle Systems. New York: Springer.

Lloyd, Alun L, and Vincent A A Jansen. 2004. Spatiotemporal Dynamics of Epidemics: Synchrony in Metapopulation Models. Mathematical Biosciences 188 (1-2) : 1-16.

Longstaff, FA, and ES Schwartz. 2001. Valuing American Options by Simulation: a Simple Least Squares Approach. Review of Financial Studies, 14: 113-148

Ludkovski, M., and J. Niemi. 2010. Optimal Dynamic Policies for Influenza Management. Statistical Communications in Infectious Diseases 2 (1): 5.

MacKay, R. Robustness of Markov Processes on Large Networks. Journal of Difference Equations and Applications, 17 (8): 1155-1167.

Mbah N , Martial L, and Christopher A Gilligan. 2010. Optimization of Control Strategies for Epidemics in Heterogeneous Populations with Symmetric and Asymmetric Transmission. Journal of Theoretical Biology 262 (4) (February

21): 757-763.

Mclean, A R, and R M Anderson. 1988. Measles in Developing Countries. Part II. the Predicted Impact of Mass Vaccination. Epidemiology and Infection 100 (03): 419-442.

Milstein, G N. 1994. Numerical Integration of Stochastic Differential Equations. Springer.

Morton, R, and K H Wickwire. 1974. On the Optimal Control of a Deterministic Epidemic. Advances in Applied Probability.

Mossong, Jol J, Niel N Hens, Mark M Jit, Philippe P Beutels, Kari K Auranen, Rafael R Mikolajczyk, Marco M Massari, et al. 2008. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases.. PLoS Medicine 5 (3) e74-e74.

Murray, JD, EA Stanley, and DL Brown. 1986. On the Spatial Spread of Rabies Among Foxes. Proceedings of the Royal Society of London. Series B. Biological Sciences 229 (1255): 111-150.

Murrell, David J, Ulf Dieckmann, and Richard Law. 2004. On Moment Closures for Population Dynamics in Continuous Space. Journal of Theoretical Biology 229 (3) (August): 421-432.

Nasell, I. 1999. On the Time to Extinction in Recurrent Epidemics. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61 (2) (April): 309-330.

O'Neill, Philip D PD. 2002. A Tutorial Introduction to Bayesian Inference for Stochastic Epidemic Models Using Markov Chain Monte Carlo Methods.. Mathematical Biosciences 180: 103-114.

Øksendal, Bernt. 2004. Stochastic Differential Equations: an Introduction with Applications. 6th ed. Springer.

Øksendal, B K, and A Sulem. 2005. , Applied Stochastic Control of Jump

Diffusions. Springer.

Ovaskainen, O, and SJ Cornell. 2006a. Space and Stochasticity in Population Dynamics. Proceedings of the National Academy of Sciences 103 (34): 12781-12786.

Ovaskainen, O, and SJ Cornell. 2006b. Asymptotically Exact Analysis of Stochastic Metapopulation Dynamics with Explicit Spatial Structure. Theoretical Population Biology 69 (1): 13-33.

Pardoux, E, and S Peng. 1992. Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations. Vol. 176. Berlin/Heidelberg: Lecture Notes in Control and Inform. Sci.

Parham, Paul E, and Neil M Ferguson. 2005. Space and Contact Networks: Capturing the Locality of Disease Transmission. Journal of the Royal Society Interface 3 (9) (December 19): 483-493.

Pontryagin, L S, R V Boltyanskii, R V Gamkriledze, and E F Mischenko. 1962. The Mathematical Theory of Optimal Processes. Halsted Press. Powell, Warren. 2007. Approximate Dynamic Programming: Solving the Curses of Dimensionality. John Wiley & Sons.

Puterman, Martin L. 2005. Markov Decision Processes: Discrete Stochastic Dynamic Programming. New York: John Wiley & Sons.

Rand, DA. 1999. 4: Correlation Equations and Pair Approximations for Spatial Ecologies. Advanced Ecological Theory: Principles and Applications: 100.

Rizzo, David M, and Matteo Garbelotto. 2003. Sudden Oak Death: Endangering California and Oregon Forest Ecosystems. Frontiers in Ecology and the Environment 1 (4) (May): 197-204.

Robbins, H, and S Monro. 1951. A Stochastic Approximation Method. The Annals of Mathematical Statistics.

Roberts, Mick, and Hans Heesterbeek. 1993. Bluff Your Way in Epidemic Models. Trends in Microbiology 1 (December 12): 1-6.

Ross, J V, T House, and M J Keeling. 2010. Calculation of Disease Dynamics in a Population of Households. PLoS One 5 (3): e9666.

Ross, R. 1916. An Application of the Theory of Probabilities to the Study of a Priori Pathometry. Part I. Proceedings of the Royal Society of London. Series A 92 (638): 204-230.

Rowthorn, RE, and et al. 2009. Optimal Control of Epidemics in Metapopulations.J. R. Soc. Interface 6 December 2009 6 (41): 1135-1144.

Salathè, Marcel, and James H Jones. 2010. Dynamics and Control of Diseases in Networks with Community Structure.. PLoS Computational Biology 6 (4) : e1000736-e1000736. .

Sellke, T. 1983. On the Asymptotic Distribution of the Size of a Stochastic Epidemic. Journal of Applied Probability 20: 390-394.

Smith, D.L., B. Lucey, L.A. Waller, J.E. Childs, and L.A. Real. 2002. Predicting the Spatial Dynamics of Rabies Epidemics on Heterogeneous Landscapes. Proceedings of the National Academy of Sciences of the United States of America 99 (6): 3668-3672.

Ster, IC, and NM Ferguson. 2007. Transmission Parameters of the 2001 Foot and Mouth Epidemic in Great Britain. PLoS One 2 (6): e502. .

Stoyan, D, and A Penttinen. 2000. Recent Applications of Point Process Methods in Forestry Statistics. Statistical Science 15 (1): 61-78.

Tildesley, M J, P R Bessell, M J Keeling, and M E J Woolhouse. 2009. The Role of Pre-Emptive Culling in the Control of Foot-and-Mouth Disease. Proceedings of the Royal Society of London. Series B. Biological Sciences 276 (1671): 3239-3248.

Tildesley, M, T House, and M Bruhn. 2010. Impact of Spatial Clustering

on Disease Transmission and Optimal Control. Proceedings of the National Academy of Sciences 107 (3): 1041-1046.

Tildesley, MJ, NJ Savill, DJ Shaw, R Deardon, SP Brooks, MEJ Woolhouse, BT Grenfell, and MJ Keeling. 2006. Optimal Reactive Vaccination Strategies for a Foot-and-Mouth Outbreak in the UK. Nature 440 (7080): 83-86.

Tildesley, MJ, R Deardon, NJ Savill, PR Bessell, SP Brooks, MEJ Woolhouse, BT Grenfell, and MJ Keeling. 2008. Accuracy of Models for the 2001 Foot-and-Mouth Epidemic. Proceedings of the Royal Society B: Biological Sciences 275 (1641): 1459-1468.

Vernon, MC, and MJ Keeling. 2009. Representing the UK's Cattle Herd as Static and Dynamic Networks. Proceedings of the Royal Society B: Biological Sciences 276 (1656): 469.

Vial, Flavie, W Thomas Johnston, and Christl A Donnelly. 2011. Local Cattle and Badger Populations Affect the Risk of Confirmed Tuberculosis in British Cattle Herds. PLoS One 6 (3) (March 28): e18058.

Viet, A F, L Jeanpierre, M Bouzid, and A I Mouaddib. 2012. Using Markov Decision Processes to Define an Adaptive Strategy to Control the Spread of an Animal Disease. Computers and Electronics in Agriculture 80: 71-79.

Xiu, Dongbin. 2010. Numerical Methods for Stochastic Computations. Princeton: Princeton University Press.

Zhang, D, and D Adelman. 2009. An Approximate Dynamic Programming Approach to Network Revenue Management with Customer Choice. Transportation Science 43 (3) (August 26): 381-394.