From individuals to populations: changing scale in process algebra models of biological systems

Chris McCaig

October 2007

Abstract

The problem of changing scale in models of a system is relevant in many different fields. In this thesis we investigate the problem in models of biological systems, particularly infectious disease spread and population dynamics. We investigate this problem using the process algebra *Weighted Synchronous Calculus of Communicating Systems* (WSCCS). In WSCCS we can describe the different types of individual in a population and study the population by placing many of these individuals in parallel. We present an algorithm that allows us to rigorously derive mean field equations (MFE) describing the average change in the population. The algorithm takes into account the Markov chain semantics of WSCCS such that as the system being considered becomes larger, the approximation offered by the MFE tends towards the mean of the Markov chain.

The traditional approach to developing population level equations of a system involves making assumptions about the behaviour of the entire population. Our approach means that the population level dynamics explained by the MFE are a direct consequence of the behaviour of individuals, which is more readily observed and measured than the behaviour of the population. In this way we develop MFE models of several different systems and compare the equations obtained to the traditional mathematical models of the system.

Acknowledgements

I would like to thank a number of people who have helped me produce this work.

A special mention goes to my supervisors, Carron Shankland and Rachel Norman, for their wealth of knowledge, help, time and guidance in keeping me on the right track.

I am very grateful to all in the Department of Computing Science and Mathematics, in particular to the support network of my fellow PhD students.

My parents deserve thanks for their support and encouragement throughout the years.

Last, but certainly not least, thanks to Julie for her endless, unwavering belief and encouragement in everything I do.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The question of how best to model a system is relevant to many different fields. For example in chemistry a reaction may be described in terms of the overall concentrations of different chemicals in a well mixed solution or in terms of the individual interactions between molecules, which contribute to the change in concentrations. Similarly biological systems can be considered in terms of changes in the makeup of a population or in terms of the individual behaviours and interactions between individuals, which are fundamental to the way a population changes over time.

"Individual" and "population" have different meanings depending on the specific system being modelled. We may have individual whole organisms (insects, fish, mammals etc) in an ecological population with, for instance, disease spread modelled in terms of the number of infected individuals in the population or in terms of the interactions between individuals, which contribute to the spread of the disease. Similarly we can have individual cells in a tumour population that can be modelled in terms of the size of the tumour as a whole or in terms of the interactions between individual cells, which are fundamentally important to the growth of the tumour. Equations

that describe a system at the level of the population are amenable to a wide range of well established, algebraically tractable methods of analysis: however it is at the level of individual behaviour that systems can most easily be observed. Being able rigorously to relate population and individual level behaviour would allow us to take advantage of the benefits of both approaches.

In this thesis we address this problem of scaling from individuals to a population using the example of models of disease spread. We do this by investigating specific theoretical questions of how best to capture particular features of disease transmission and population dynamics in individual-based models. These individual-based models are then transformed into population level models for further investigation.

1.1 Mathematical modelling of ecological systems: population vs individual

1.1.1 Overview

Mathematical models of biological systems have been used for centuries to offer insight into the factors that govern the behaviour of a system. As far back as 1760 Bernoulli [13, 14] presented a model to study the effect of inoculation on the spread of smallpox. The most common approach to modelling disease spread uses ordinary differential equations (ODEs) (either singularly or systems of coupled ODEs), which model a system in continuous time [6, 52, 61, 91, 92] or difference equations (again either individually or systems of coupled equations), which model a system in discrete time [4, 75]. Specific models of disease are discussed in more detail below, in Section 1.1.2.

Models such as these can be analysed using a wide range of algebraic

techniques to study the system without the need to define values for the parameters in the model and also by simulating the model and producing the time series for a given initial population and biologically relevant parameter values. Although these models make some intrinsic assumptions about individual behaviour they do not explicitly model the behaviour of individuals. Specifically the interactions between individuals, which are fundamentally important in biological systems, are not modelled explicitly. Instead ODE models assume that individual behaviour translates predictably to population behaviour, however a number of papers have shown that this is not the case [19, 90]. ODE models describe changes in the number of individuals in the population, or in different subpopulations, but do not include any spatial information about systems. Since no spatial information is included these ODE models are based on the assumption that the population is subject to random mixing.

Another method of modelling biological systems is by the use of probabilistic cellular automata [2, 76, 81, 90]. Cellular automata capture behaviour at the individual level (individual-based models) by describing rules of behaviour for individuals and creating a population from many individuals. These models are generally studied by performing simulations of the system and only limited algebraic analysis is available through methods such as pair approximation [50]. Cellular automata are inherently spatial with the population defined on a grid that represents the spatial environment in which the population exists.

1.1.2 SIR disease models

While Bernoulli's smallpox model was the earliest disease model [13, 14], the classical ODE model was developed by Kermack and McKendrick [52,

53, 54]. This model of three coupled ODEs has two key features that are still used in many models today: the term chosen to describe transmission (discussed overleaf) and the subdivision of the population.

The population was divided into three distinct groups:

- Susceptibles - have never had the disease and may contract it after exposure.
- Infecteds - have the disease and can pass it on to susceptibles.
- Recovereds - have previously had the disease and are assumed to be immune to future infection.

Models that divide the population in this way are often referred to as SIR models. The SIR classification is appropriate for many diseases, and is widely used, but variations that have been used to model other diseases include SIS - recovery does not confer immunity and individuals become susceptible once more - and SIRS - conferred immunity lasts for a limited period and recovered individuals can once more become susceptible [27, 33, 43, 56, 66, 95].

In many disease systems there is a time delay between susceptibles coming into contact with an infected individual and becoming infected. This has been modelled by adding an exposed group (E) to the models, and these are known as SEIR models. As for SIR there are variations in which recovery does not confer immunity - SEIS - and immunity lasts for a limited time - SEIRS [3, 37, 72].

Transmission terms

The second feature introduced by Kermack and McKendrick [52] that is still widely used is the term used to model transmission [6]. The term chosen to

capture transmission was

$$\beta SI\ ,$$

with S and I being the numbers (or densities) of susceptible and infected individuals respectively. This term comes from the *Law of Mass Action* [39] from chemistry. In chemistry the underlying assumption is that the rate of a reaction increases with the concentration of either reactant, and similarly here the assumption is that the rate of contacts made by individuals increases as the population size increases. For this reason transmission of this form is often referred to as mass action transmission or more commonly *density dependent transmission*. This transmission term is commonly used for many wildlife and animal models and some human diseases.

Although density dependent transmission is still commonly used several other transmission terms have been suggested. Most notably frequency dependent transmission,

$$\frac{\beta SI}{N} \tag{1.1}$$

where β is different from the β in the density dependent term and N is the total number (or density) of individuals in the population, has been used to model certain types of disease. Frequency dependent transmission assumes that an individual makes a fixed number of contacts regardless of the population density and is used most commonly to model human diseases and vector borne diseases where contact saturation is assumed to have occurred [30, 85]. Begon et al. [10] described biological derivations for both density dependent and frequency dependent transmission, in particular suggesting that numbers of individuals, rather than densities, should always be used and therefore that density dependent transmission should more accurately

be written as

$$\frac{\beta SI}{A},$$

where A is the area occupied by the population.

Several studies have suggested that βSI may not be the term that most accurately describes transmission of disease systems where density dependent transmission is expected. Turner et al. [90] developed individual-based cellular automata models that displayed density dependent (contact with all nearby individuals) and frequency dependent (contact with a fixed number of individuals) transmission at the level of the individual. Fitting terms to the numerical results from their models they found that, irrespective of the individual level behaviour, frequency dependent transmission most accurately described the population level behaviour. This result was counterintuitive since it was assumed that behaviour at the individual level would translate to the same behaviour at the population level.

Other alternative transmission terms have also been suggested. Hochberg [48] proposed the term

$$\beta(S^p I^q)SI,$$

where p and q are parameters that can be chosen to give a variety of non-linear responses. Estimates of p and q for insect borne pathogens were calculated by Fenton et al. [31]. Briggs and Godfray [20] also proposed a transmission term,

$$\left[k\ln\left(1+\frac{\beta I}{k}\right)\right]S,$$

where k is a scaling constant. Both Hochberg and Briggs and Godfray

found that their transmission terms better fitted their experimental data than βSI . Knell et al. [55] fitted both terms to their experimental data and found they both provided a better fit than βSI: however it is difficult to measure transmission in any populations other than insect systems. The Hochberg and Briggs and Godfray models are more flexible than the more common density dependent term, which explains why they can be made to better fit experimental data.

Some studies have suggested transmission terms that seek to address the density/frequency dependent dichotomy by capturing both forms of transmission. Antonovics et al. [7] proposed the following term for contacts between susceptible and infected hosts:

$$N_e = \frac{aTS_tI_t}{1 + aT_hN_t} ,$$

where a is the area searched by infecteds in time T, S_t and I_t are the numbers of susceptibles and infecteds respectively, $N_t = S_t + I_t$, and T_h is the duration of each contact. Antonovics et al. demonstrated that for small N_t this term behaves like density dependent transmission and as N_t becomes large N_e asymptotically tends to a frequency dependent contact rate of T/T_h.

Ryder et al. [80] proposed the following transmission term for diseases that can be spread by both frequency dependent and density dependent contacts:

$$\frac{v(c + mN)SI}{N} .$$

For $c = 0$ this becomes the density dependent term $cmSI$ and for $m = 0$ we have the frequency dependent term $vcSI/N$. When $c > 0$ and $m > 0$ the term will captures aspects of both density dependent and frequency dependent transmission. This was proposed as being appropriate for diseases that

can be spread by both sexual (typically modelled and frequency dependent) and social (typically modelled as density dependent) contacts.

Which transmission term to use in different situations remains an open question. As such this is an area that lends itself to the rigorous investigation of the connection between behaviour at the individual and population levels.

Indirect Transmission

The models discussed so far all assume that the disease is transmitted by direct contact between infected and susceptible individuals. However, it is known that some infectious agents can persist in the environment for significant periods and therefore direct contact between individuals is not necessary. This persistence in the environment can be for differing periods depending on the disease, with a disease such as Feline Panleucopenia Virus surviving for up to a year outside of its host [15]. The common approach to modelling indirect disease transmission [6] is to introduce an additional equation to describe the quantity of infection present in the environment. The infectious agent is transmitted from infected individuals to the environment and susceptible individuals pick up the infection from the environment. This leads to a delay in the spread of the disease since there is a time cost associated with transmission of the disease from the infected individual to the environment, a time cost associated with the infectious agent persisting in the environment and a further cost associated with the disease being contracted by a susceptible individual. A challenge for our approach is to be able to model indirect transmission.

Superspreaders

A feature of disease transmission that is of growing interest is superspreaders [51]. This is the idea that the majority of new infections are caused by a small proportion of the infected individuals. There are two proposed mechanisms by which superspreaders can be responsible for a greater proportion of new infections:

- contact superspreader - superspreader individuals make more contacts than other infected individuals and therefore have more opportunity to pass on the disease
- infectiousness superspreader (supershedder) - superspreader individuals are more infectious than other infected individuals and susceptible individuals are more likely to become infected after contact with a superspreader

Lloyd-Smith et al. [60] studied data for epidemics for which contact tracing information was available and concluded that superspreaders are a common feature of disease transmission. They found that most individuals do not transmit the disease at all while a small proportion of infecteds pass the disease on to many susceptibles. Matthews and Woolhouse [64] proposed that the presence of superspreaders in a disease system could be expected to increase the variability of the system.

Models of superspreader systems feature separate groups of infecteds and superspreaders in the population [51]. Superspreaders either have an increased rate of contact or susceptibles have a greater rate of becoming infected if contact is with a superspreader [32, 94]. Since it is the behaviour of individuals that differs, this is an ideal proving ground for our modelling approach.

1.1.3 Disease free population dynamics

The way in which population size changes is important within disease models since it can affect the dynamics of the disease: however, the question of how best to describe population dynamics has long been of interest in its own right. The idea that populations cannot grow without bound has been of interest to modellers for centuries. Malthus, in 1798, [62] proposed a simple exponential growth model based on compound interest,

$$\frac{dN}{dt} = rN \ ,$$

but noted that this was unrealistic, since when a population becomes very large, access to resources will become restricted, restricting further growth in the population. In the Malthusian growth model r is the growth rate of the population: it is clear that for $r > 0$ the population will grow without bound; $r < 0$ will lead to extinction; and $r = 0$ will give a stable population size. This led Verhulst to propose the logistic growth model [91],

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) \ ,$$

where r is once again the growth rate and K is the carrying capacity of the environment in which the population exists i.e. the maximum population size (or density) that the environment can sustain. When N is small N/K will be close to 0 and the logistic equation displays exponential growth, similar to the Malthusian model. As N increases the quotient $N/K \rightarrow 1$ so that when $N = K$ $dN/dt = 0$ and N is stable.

Gompertz [34] also proposed a model that sought to address the short-

comings of the Malthusian model:

$$\frac{dN}{dt} = rN \ln\left(\frac{K}{N}\right) ,$$

with r and K once again the growth rate and carrying capacity respectively. This model displays similar behaviour to the logistic model, with the population growing rapidly when N is small and reaching a steady state when $N = K$. The Gompertz model has been shown to be particularly effective in modelling the growth of solid tumours [63].

The logistic model is the most commonly used model to describe population dynamics but several other models have been proposed [40, 78]. For example the Beverton-Holt model [16],

$$N_{t+1} = \frac{aN_t}{1 + bN_t} ,$$

is a discrete time model that was proposed to describe the dynamics of fish populations. The Beverton-Holt model has been widely used to study fish populations [71, 84] and was among the models for which Brännström and Sumpter [19] developed derivations from their discrete site-based framework.

Despite the prevalence of the logistic model in the literature it is still unclear which model is most appropriate to describe population dynamics in different situations and therefore this is another area that can benefit from a rigorous study of the connection between individual and population level behaviour.

Models of population dynamics are not merely interesting in isolation. For example in our field of interest, infectious disease spread, adding birth and death of individuals to a model of disease spread can alter the dynamics of the epidemic.

1.2 Process algebra

Baeten [8] defines process algebra as "the study of the behaviour of parallel or distributed systems by algebraic means". Many different process algebras, or process calculi, have been developed [44, 67, 69, 87] with different features, which lead to differing approaches in modelling a system. Process algebra traces its roots back to the development of Petri nets [74], which was developed to study distributed computer systems.

Milner developed the *Calculus of Communicating Systems* (CCS) [67] on which several other process algebras have been based, such as SCCS (Synchronous CCS) [68], WSCCS (Weighted SCCS) [87], the $\pi-$calculus [70] and PEPA (Performance Evaluation Process Algebra) [44]. Other non-CCS based process calculi include CSP (Communicating Sequential Processes) [47] and ACP (Algebra of Communicating Processes) [11].

Process algebras were originally developed to give formal semantics to parallel programming languages. In addition process algebras have been used to study a wide variety of systems; for instance hybrid systems [12], asynchronous systems [28], cryptographic protocols [1], stochastic musical systems [79], and biological systems [77]. There is much to be gained from using process algebra to study systems. Not only does the act of specification lead to deeper understanding of the system being described, through clarification of assumptions and explicit definition of the actions being performed and agent interaction, mathematical analysis can be carried out on the specification since it has a formal semantics. For WSCCS this means investigation of the underlying Markov chain, allowing the probabilities of states occurring to be calculated. As well as Markov chain analysis [46] such models can be studied by performing stochastic simulations [18].

1.2.1 Process algebra & biology

Over the past 10 years process algebra has increasingly been used to model a wide range of biological systems [24, 41, 73, 77, 82, 86]. The advantages of process algebras are that they are fully formal (with mathematical semantics), making them amenable to rigorous analysis, and the features they have for describing systems, particularly for creating larger systems from smaller identical components, are turning out to be useful in biological applications.

Regev et al. [77] developed the BioAmbients calculus, based on the π-calculus, for modelling biomolecular systems. In this work the BioAmbients calculus was used to model the hypothalamic system, which regulates body weight, at the level of individual neurons.

PEPA [44] has been applied to biochemical systems. Calder et al. [23, 24] used PEPA to model the effects of Raf Kinase Inhibitor Protein (RKIP) on the Extracellular signal Regulated Kinase (ERK) pathway, which has been shown to play an important role in tumour development.

WSCCS (Weighted Synchronous Calculus of Communicating Systems) [87] has been used in particularly diverse biological applications, ranging from insect behaviour [83, 86] through epidemiology [73] to genetics [41]. For example Hatcher and Tofts [41] presented WSCCS models of sex selection that incorporated genetic and environmental selection factors. Tofts [86] developed models of social insect colonies.

Further work on social insect colonies was presented by Sumpter [83]. These models addressed different aspects of social insect behaviour from activity synchronisation to a site based model of population dynamics in which individuals competed for breeding sites. In addition Sumpter introduced a simple model of population dynamics with food as a resource. Individuals die probabilistically and compete for food, using prioritised communication,

giving birth if they manage to eat. The competition for food makes birth density dependent. This means that the mean population tends to a steady state that is dependent on the quantity of food available and the probability of death. Sumpter also presented a simple SIR model of disease spread. For these models Sumpter used intuitive reasoning to derive mean field equations (MFE), which describe the mean behaviour of the system at the population level, and produced graphs to demonstrate that these closely approximated the mean of many simulations of the model. The models involve multiple distinct stages and one timestep in these MFE describe the mean behaviour of the model over a complete iteration. For example in the population dynamics model the system contains some number of $A1$ agents that become $A2$ before once again becoming $A1$ after two stages. During these two stages agents have been removed (death) and added (birth) so that the number of $A1$ agents is changed. The MFE describe the mean of this change so that one timestep in the MFE represent two stages in the WSCCS model.

Norman and Shankland [73] developed SIR models in WSCCS for which they derived MFE using the same intuitive reasoning method as Sumpter. These models sought to improve the biological realism of Sumpter's model [83] and the most realistic models lead to MFE featuring the density dependent transmission term

$$\frac{\beta S_t I_t}{N_t} .$$

In these models Norman and Shankland found that the decision to use prioritised or non-prioritised communication did not affect the resulting MFE. It is unclear whether this will always be true, or is a result of the specific models considered, and as such this is an interesting point for further investigation.

1.2.2 Deterministic equations from process algebra

In addition to the previous studies which derive MFE for WSCCS [73, 83] equations have been derived which approximate the mean behaviour of systems described in other process algebras. Cardelli produces a continuous time semantics in terms of ODEs from a subset of the π-calculus enriched with transition rates (the Chemical Ground Form) [26]. The process algebra is broadly similar to WSCCS, but has continuous rates instead of probabilistic choice. The translation to ODEs is given directly, but the proof is via translation to Chemical Reactions. A key observation is the translation between stochastic rates in the (discrete state) process algebra to kinetic rates in the (continuous state) ODEs multiplying by a factor related to the number of molecules in the solution. This allows use of the law of mass action, which applies only in the continuous setting. The work is motivated by chemical reactions, but can be applied in other settings. In particular, a related paper [25] contains the Kermack McKendrick SIR example [52], and the well known Lotka-Volterra predator-prey example [61, 92].

There is also some similarity with the work of Brodo et al. [21], who derive numeric rate information for π-calculus models. Their work is concerned with performance analysis, and relies on information about network topology, throughput, latency, protocol complexity; however, their systems appear not to be composed of many copies of the same agent (and therefore the rates do not take this into account).

Hillston [45], and Calder et al. [23] presented methods of deriving ODEs from a subset of the continuous time process algebra PEPA which makes use of a *numerical vector form* representation of the system. The methods are broadly the same but use different terms in the resulting ODEs to capture communication. The methods were extended by Bradley et al. [17] to

cover a greater range of potential PEPA models. The approach is applied to internet worm attacks [17], which have previously been modelled using SIR models similar to those used to describe infectious disease spread [59]. The form of communication used in these models does not allow communication directly between infected and susceptible computers, communication is instead modelled over a network. The infection passes from infected PCs to network channels and then from the network channels to susceptible PCs which is realistic for a worm attack. These models therefore capture behaviour which is analogous to indirect disease spread [6]: however they would not be suitable for modelling direct transmission as captured in existing WSCCS [73, 83] and π-calculus [25] models.

1.2.3 WSCCS

Syntax of WSCCS

In WSCCS the basic components are *actions* and the *agents* that carry out those actions. The actions are chosen by the modeller to represent activities in the system. For example, *infect*, *send*, *receive*, *throw dice*, and so on. Actions occur instantaneously and have no duration. Agents represent the different components of a system which can perform the actions e.g. *Infected, Susceptible, Die, Gambler, Router*. There is no measure of time in WSCCS but there is temporal ordering and synchronisation of events. If we think of the ticking of a universal clock, on each clock tick all agents must perform an action, though they need not change state as a result. Clock ticks are not necessarily evenly spaced in time and the wait between consecutive ticks can vary from being instantaneous to happening over longer periods of time.

WSCCS is a probabilistic process algebra, meaning that the decision to move from one state to another can be a probabilistic one. In Appendix A

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} pc.\surd : P2 + (1-pc).\surd : S2 \\
I1 &\stackrel{\text{def}}{=} pr.\surd : R2 + pa.\surd : T2 + (1-pr-pa).\surd : I2 \\
R1 &\stackrel{\text{def}}{=} 1.\surd : R2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
P2 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I1 + 1.\surd : S1 \\
I2 &\stackrel{\text{def}}{=} 1.\surd : I1 \\
T2 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : I1 + 1.\surd : I1 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} S1 \times S1 \times S1 \times S1 \times S1 \times I1 \lceil \{\surd\}
\end{aligned}
$$

Figure 1.1: Simple epidemic model of Sumpter [82]

we present the formal semantics of WSCCS (as defined in [87]) but here we present an informal overview.

To illustrate the different definitions we consider a basic model of disease spread by Sumpter [82], shown in Fig. 1.1. The behaviour captured in this model is as follows: susceptible $S1$ agents either make themselves available to be infected (by becoming $P2$ with probability p_c) or not (becoming $S2$ with probability $1-pc$); infected $I1$ agents make themselves available to pass on the disease (becoming $T2$ with probability p_a), recover from infection (becoming $R2$ with probability p_r) or do neither (becoming $I2$ with probability $1-p_r-p_a$); recovered $R1$ agents all remain recovered (becoming $R2$). In the second stage of the model the $S2, I2$ and $R2$ agents all have no choice to make (becoming $S1, I1$ and $R1$ respectively) while the $P2$ and $T2$ agents can be involved in transmission of the disease.$P2$ agents can either become

infected (becoming $I1$ agents) or not (becoming susceptible $S1$ agents) while the $T2$ agents all remain infected (becoming $I1$) irrespective of whether they pass on the disease or not.

Model structure The models presented in this thesis follow the same general structure used by Sumpter [82] and Norman and Shankland [73]. Activity is separated into different stages (*ticks*), which happen consecutively, and communication and probabilistic choice happen in different stages. For instance Sumpter's simple disease model [82] (Fig. 1.1) is a two stage model. In the first stage the system consists of $S1, I1$ and $R1$ agents that make probabilistic choices. In the second stage the system consists of $S2, P2, I2, T2$ and $R2$ agents with the numbers of each type of agent depending on the probabilistic choices made in the first stage. The $P2$ and $T2$ agents communicate to model transmission of the disease and the $S2, I2$ and $R2$ agents deterministically become $S1, I1$ and $R1$ respectively. After the two stages the population once again consists of $S1, I1$ and $R1$ agents although the numbers of each type of agent are different than before. The MFE describe the mean change in the numbers of $S1, I1$ and $R1$ agents over the two stages, so the *timestep* of the MFE covers two ticks in the WSCCS model. We refer to the two ticks of the model as an *iteration* of the model. In general for an $n-$stage model the MFE describe the average change in the numbers of each type of agent present at the first stage over an iteration, which consists of n stages.

A summary of the syntax is presented in Table 1.1. The operations of WSCCS are:

Prefix	$a : P$	Simple agent: does action a and becomes agent P at the next stage.
Weighted choice	$w_1.P1 + w_2.P2$	Behaves as $P1$ with probability $w_1/(w_1 + w_2)$, or behaves as $P2$ with probability $w_2/(w_1 + w_2)$.
Parallel coordination	$P1 \times P2$	Agents in parallel execute actions together at each stage.
Identity action	$\surd$	By default $\surd$ occurs without communication.
Restriction	$System \lceil A$	Only actions in the set A are allowed without communication.
Communication output action input action	 $\overline{act}$ act	In general actions can only happen by communication i.e. if one agent does act another can do $\overline{act}$. Neither action can happen independently.
Parallel actions same action different actions	 act^n $act1 \# act2$	Must communicate either with several agents, or with agent performing parallel actions.
Priority	$n\omega^k$	ω: infinite weight. Different levels of priority are allowed with ω^{k+1} chosen in preference to ω^k.
Null agent	0	An agent which becomes 0 is removed from the system.

Table 1.1: Summary of WSCCS syntax

prefix This is the simplest form of agent: $a : P$ where a is an action, and P is an agent. This agent can carry out the action a and then behave like agent P. Actions are as described above.

weighted choice The agent $w1.P1 + w2.P2$ offers a choice between the agents $P1$ and $P2$. Assuming both agents are able to progress, the branch chosen depends on the weights. Over a number of trials we observe $P1$ being chosen with a probability $w1/(w1+w2)$ and `P2` being chosen with a probability $w2/(w1 + w2)$. For example the agents $S1$ and $I1$ in Fig. 1.1 make choices based solely on the associated weights. For convenience these have been written as probabilities and we follow this example for all agents in our models which are governed only by weights (we refer to such agents as *probabilistic agents*). Weights are generally positive natural numbers, but may also incorporate the special weight ω which is greater than all natural numbers. This is used in *priority* and we can have different levels of priority. When different levels of priority are used the weight is written $m\omega^n$, where n is the priority and m is a weight used in determining between choices of equal priority. Options with priority $n + 1$ will always be chosen in preference to options with priority n

synchronous parallel coordination Obtaining more complex behaviour requires the use of coordination. Simple agents using the operators above may be combined with each other in parallel, e.g. $P1 \times P2$. Parallel agents operate in lock step; that is, if we imagine the ticking of a universal clock controlling the occurrence of actions, then all agents must execute some action together on the clock tick - but not necessarily the same action. For example in Fig. 1.1 the agent

Population is such a parallel agent which here represents the initial state of the population (five $S1$ and one $I1$). Here we consider only a small population (Sumpter [82], and Norman and Shankland [73] who also considered this model, did not specify the initial state of the population) since in the standard notation we must explicitly define each component of such parallel agents. In Chapter 2 we define notation that allows us easily to define very large populations, or to assign symbolic labels to the numbers of individuals.

Parallel agents can also be used either to describe individuals which have two different types of behaviour or to model changing population size i.e. birth.

communication Two agents in parallel may communicate when one carries out an output action and the other carries out the matching input action, e.g. $\overline{infect}$ and *infect*. Communication can be used to model passing of information from one agent to another, or to coordinate activity. Such communication is strictly two-way; that is, only two agents may interact on this action. For example in Fig. 1.1 for some of the agents $P2$ to be able to perform the input action *infect* an equal number of $T2$ must perform the output action $\overline{infect}$. Communication with several agents simultaneously is achieved by multiple actions. For example, $infect^3$ is shorthand for three *infect* actions in parallel (alternatively written as *infect*#*infect*#*infect*) and hence the possibility to synchronise with three other agents. The distinguished action $\surd$ can never communicate. Communication is enforced when the action is hidden from the environment using *restriction*.

restriction Without restriction, all agents may communicate with the environment as well as with each other. With restriction, we can force two (or more) agents to communicate with each other on chosen actions. For example, given the agent $(P1|P2)\lceil a$ where $P1$ and $P2$ can carry out actions a, b, then $P1$ and $P2$ must cooperate on b actions, but a actions are visible in the environment, and available to synchronise with other agents. Actions are hidden by default and only the actions which are explicitly allowed can happen independently. In Fig. 1.1 only the action $\surd$ is visible. All other actions are hidden (for this model the *infect* input and output actions) and can only happen by communication.

priority In a choice, the agent with infinite weight $n\omega^k$ will always be taken in preference to the one with a natural number weight. This can be used to force particular actions to occur (usually communications) if possible, allowing the alternative choice only if there is no other agent with which to communicate. There is a hierarchy of weights, with $\omega^{k+1} > \omega^k$. In Fig. 1.1 the agents $P2$ and $T2$ are both prioritised to communicate on the *infect* action ($P2$ the input action and $T2$ the output action) so that when there are sufficient $T2$ agents all of the $P2$ must communicate and vice versa. This means that the number of $P2$ agents which become infected $I1$ agents is either equal to the number of $T2$ agents, if there are more $P2$ than $T2$, or all of the $P2$ become $I1$, if there are as many (or more) $T2$ as $P2$.

In all models presented in this thesis the system as a whole is described by the system equation *Population* (or *Popn*), comprising multiple copies of each kind of agent in parallel.

Classification of Markov chains

In this section we comment on some of the types of Markov chains [38] that can arise from the WSCCS models presented in this thesis. We do not explicitly consider the Markov chains of our models although the algorithm for deriving MFE, which is introduced in Chapter 3, takes account of the Markov chain to calculate the mean behaviour of a model.

Models that do not feature birth and death, for example Sumpter's basic disease model [82] (Fig. 1.1) and the simple models presented in Chapter 3, have finite Markov chains. In Chapter 4 we consider models that feature birth and death of individuals and in subsequent chapters include this in our disease models. These models have infinite Markov chains. Whether the underlying Markov chain is finite or infinite does not affect our ability to derive MFE to describe the mean behaviour of the model.

In general for the $n-$stage models ($n > 1$) in the following chapters all states are periodic with period n, i.e. if the system is in state X it can only return to X after multiples of n ticks. Since all of the states that agents can take at each stage are different to the states they can be in at the previous and subsequent stages these models are never aperiodic. One-stage models are aperiodic since they can remain in the same state from one stage to the next.

The only model in this thesis that leads to a Markov chain with an absorbing state is a simple model in Chapter 2, included to demonstrate the use of functional probabilities (Fig. 2.1). In general $n-$stage models never feature absorbing states as all agents change state at each stage of the model.

Our $n-$stage models models without birth and death can lead to Markov chains that feature cycles. Consider, for example, Sumpter's basic model

(Fig. 1.1), if the disease dies out (i.e. all infecteds recover) the system will consist of some numbers of $S1$ and $R1$ agents. Every two ticks the system will consist of the same numbers of $S1$ and $R1$, and at the intermediate stage the system will consist of some numbers of $S2, P2$ and $R2$ agents. Models which feature birth and death will never feature cycles since it is always possible for the total number of agents in the model to change over $n-$stages.

The MFE which we derive for our system consider the mean behaviour of a model over $n - stages$. This means it is possible that the MFE will have a stable steady state, even though the Markov chain does not have an absorbing state. This would indicate that the expected state of the system tends to some state, X, even though the Markov chain will always be able to evolve to other states.

1.3 Thesis outline

In this thesis we address the problem of changing scale in terms of models of disease spread. In particular we address the following questions:

- How can we rigorously move from individual level to population level descriptions of a system?
- What individual level behaviours lead to different equations for population growth?
- Can individual level behaviour be defined which leads to the traditional density dependent transmission term?
- How can we capture indirect transmission in individual level models?
- What effect do superspreaders have on the variability of a system?

In Chapter 2 we introduce additional WSCCS notation which allows us simply to capture the kinds of density dependent behaviours which are common in biological systems. This additional notation does not extend the expressive capability of WSCCS but simplifies how we write such models. We demonstrate how models which make use of density dependent behaviour can be written using only the standard notation.

Chapter 3 addresses the question of changing scale from an individual-based WSCCS model of a system to population level equations. An algorithm is presented which formalises this process and its correctness is proved. The use of the algorithm is demonstrated for a basic WSCCS model of disease spread.

In Chapter 4 we consider the question of how to capture realistic growth in a population. This is crucial to be able to develop realistic disease models since fluctuating population dynamics can have a bearing on the dynamics of disease spread.

Chapter 5 considers different transmission terms which can arise from WSCCS disease models. The density dependent term, βSI , is most commonly used in ODE models of disease spread but the existing WSCCS models have naturally led to the frequency dependent term $\beta SI/N$. We investigate whether it is possible to define individual level behaviour which leads to equations featuring density dependent transmission.

Chapter 6 examines another aspect of disease transmission by considering models featuring indirect transmission. Many infectious micro-organisms can survive independently in the environment and this can have an effect on the spread of the disease.

In Chapter 7 we introduce superspreaders to our models. Studies have suggested that superspreaders play a vital role in the spread of some diseases.

We further study these models by performing simulations and comparing the variability with superspreaders to that of standard models in which all infected individuals are equally infectious.

In Chapter 8 we discuss the most important results of this thesis and propose some areas for further work.

Chapter 2

Non-canonical Notation

Some new WSCCS notation is introduced in this chapter to simplify the process of writing complex models. Everything we introduce is defined in terms of existing WSCCS constructs and macros can be defined that rewrite models using only the standard notation. Firstly notation is presented that aggregates many agents of the same type in parallel. This makes it simple to write models that consist of a large number of agents of only a few types and also models where the numbers of agents of each type is represented symbolically. Secondly notation is introduced that allows parameters in a model to be functionally dependent on the number of agents of a given type present at a given time. Both of these new forms of notation are additions to the rules for processes in the definition of the syntax in Section 1.2.3.

2.1 Aggregation Notation

The usual notation for writing down a number of WSCCS agents in parallel involves explicitly writing each of the agents involved, even if there are

many agents of the same type. For example 20 agents of type A in parallel is written as

$$A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \times A \ .$$

We wish to consider systems consisting of much larger numbers of individuals. By deriving MFE we can produce the time series of the mean of a system consisting of hundreds or thousands of individuals and also perform mathematical analysis on models where the numbers of individuals are not defined, with symbolic labels used in place of numbers of individuals. To aid us in writing large and symbolic models we define aggregated notation for a parallel agent consisting of multiple copies of a single agent type.

Definition 1 *The parallel agent $A\{a\}$ consists of a instances of agent A.*

By extension the agent $A\{a\} \times B\{b\} \times ... \times N\{n\}$ consists of a instances of the agent A, b instances of the agent B ... and n instances of the agent N, for example

$$X\{2\} \times Y\{3\} \times Z\{2\} \stackrel{\text{def}}{=} X \times X \times Y \times Y \times Y \times Z \times Z \ .$$

This aggregation fits within the standard WSCCS notation because the parallel operator ($\times$) is associative and commutative [87].

Traditional methods of analysis for process algebra models, such as exploring the Markov chain or performing simulation, require the initial numbers of each type of agent to be known - and to be small for Markov chain analysis - but the derivation of mean field equations for our models, described in Chapter 3, is done independently of the precise numerical values for the numbers of agents. For this reason all of the models presented in

the thesis will have the initial state of the system given in symbolic terms although numerical values may be assigned to the symbolic labels when analysis other than deriving the MFE is performed.

2.2 Functional Parameters

Many biological systems display density dependent behaviours, for example the infected individuals in disease systems may make more contacts as the population density increases [10]. To be able to model such systems efficiently we wish to be able to include functional parameters, which depend on the numbers of agents of a certain type present in the system at a given time. To facilitate this we introduce here a notation with which we can define functional parameters and also demonstrate how we can expand these to give a model using only the standard WSCCS notation [87]. Functional parameters do not extend the expressive capabilities of WSCCS since all models that make use of functional parameters could be written without them. However, what functional parameters do offer is a more concise and intuitive way to describe complex behaviours. We define two distinct types of functional parameter: functional probabilities for agents that evolve without interaction; and functional parallel agents in which the integer n, used in some parallel agent $X\{n\}$, depends on the numbers of some types of agents in the system.

To utilise functional parameters we define a notation that represents the number of agents of a given type currently present and we represent this using $\lfloor\ \rfloor$.

Definition 2 *$\lfloor S \rfloor$ is "the number of S agents that are present at the current stage".*

The parameters that are functional and the functions representing them are defined at the start of the model. The values of any constants required by the function can also be defined using standard mathematical notation e.g. $p = 0.5$.

In the expanded form a single stage in the model is replaced with three distinct stages. For example, if the agent Xi (which is present at the ith stage in the model) is to be functionally dependent on the numbers of some agents, it will be replaced by the agents Xia, Xib and several agents Xic_n, where n is the number of the agents counted, and each of which has the functional parameter set accordingly. At the first of these stages a *count* agent, $Cntia$, is used that interacts with all the agents we are interested in counting and becomes an agent where the state name encodes the number counted ($Cntib\_0, Cntib\_1$ etc). At the second stage the resulting agent once again interacts with all of the agents that were counted, performing the output action $\overline{countn}^n$. The agents that make use of the functional parameter then evolve to a the relevant state, Xic_n. This is similar to value-passing CCS as defined by Milner [68] with the count agent passing a parameter to other agents encoded in the name of the action performed.

In all other stages of the model the Cnt should be some agent that progresses to the next stage without any interaction or choice. For example

$$Cnt1 \stackrel{\text{def}}{=} 1.\surd : Cnt2$$

is the count agent at the first stage of a model where there is no functional parameter in the first or second stage (if there was a functional parameter at the second stage of the model $Cnt1$ would evolve to $Cnt2a$, which would perform the *count* action). Any agents in the model that are not counted

for calculation of the functional parameter behave in a similar manner at the additional stages when the functional parameter is calculated i.e. they do nothing while the count occurs.

Adding two stages to the model does not affect the time represented by an iteration of the model. In the biological system the density dependent behaviour arises instantaneously and since the stages in the model need not all be of the same duration the additional stages (for counting and value passing) can be thought of as instantaneous. In more complex models it may be the case that there are different functional parameters at different stages of the model (this is true of some of the models presented in Chapter 5). To implement the expanded form of such a model all of the stages featuring functional parameters would be expanded to three stages with counting happening each time. It may be true that the number of agents being counted will be the same each time but to make it possible to have a rigorous general method for expanding the functional parameters counting should be implemented each time.

Many of the models described in this thesis are for systems displaying density dependent behaviour. By utilising the notation presented here we can include density dependent properties simply, meaning that it is easy to study the effect of the density dependence without having to consider its implementation. In subsequent chapters models will be presented with functional parameters without any consideration of the expanded form, which would use only standard WSCCS notation.

2.2.1 Functional Probabilities

Functional probabilities are defined at the start of the model thus

$$label \stackrel{\text{prob}}{=} function \ .$$

Since the probability will always fall between 0 and some upper limit $pL \leq 1$ the function should be written in the form $p \stackrel{\text{prob}}{=} min(max(0, \mathrm{F}), p_L)$ where F is some function of the numbers of agents in the population. The probability p_L is the upper limit that the given probability can take. Probabilistic agents are generally written in the form

$$I1 \stackrel{\text{def}}{=} p_r.\surd : R2 + p_a.\surd : T2 + (1 - p_a - p_r).\surd : I2 \ ,$$

which is a typical infected agent from the disease models of Norman and Shankland [73]. In this example p_r is the probability with which the individual recovers, p_a is the probability that it is able to make an infectious contact and $1 - p_r - p_a$ is the probability it neither recovers nor is able to make an infectious contact. If the probability p_a is to be functional we require $pL = 1 - p\_r$ to ensure that $1 - p\_a - p\_r \geq 0$. Changing the probability p_a depending on the population size does not impact on the probability p_r, although $(1 - p_a - p_r)$ will change.

In WSCCS, weights can take any form and need not always be written as probabilities. For agents that behave purely probabilistically it is more convenient to choose the weights as probabilities and when using functional weights we have another reason to do so. This $I1$ agent could instead be written as

$$I1 \stackrel{\text{def}}{=} w1.\surd : R2 + w2.\surd : T2 + w3.\surd : I2 \ ,$$

where the weights $w1, w2, w3$ are not probabilities. In this case if we were to make the weight $w2$ functional in terms of $\lfloor I1 \rfloor$, in order that the likelihood that an individual is able to make an infectious contact depends on the number of infected individuals, we would also be altering the likelihood than an individual will recover. This is because the probability of recovery here is

$$p_1 = \frac{w1}{w1 + w2 + w3}$$

so that as $w2$ increases not only is an individual more likely to be able to make an infectious contact but it is also less likely to recover. In our models it would not be biologically realistic for one density dependent behaviour to have an inverse effect in all other probabilistic choices. It would be possible to write $I1$ in the form

$$I1 \stackrel{\text{def}}{=} w1.\surd : R2 + w2.\surd : T2 + (ws - w1 - w2).\surd : I2 \ ,$$

where ws is a fixed value for the sum of the three weights, meaning that the probability of becoming $T2$ would not be affected by changing $w1$. However, this is not an obvious way to write an agent, whereas if we are considering probabilities it is necessary that the probabilities sum to 1. For these reasons we always write agents that do not take part in communication using probabilities. One of the probabilities takes the form

$$1 - \sum_{i=1}^{n-1} p_i$$

where n is the number of choices. The choice that arises from the probability of this form should be the one that can be thought of as "do nothing": for example in the $I1$ agent considered above the choices can be thought of as

$$
\begin{array}{rcl}
p & \stackrel{\text{prob}}{=} & min(max(0, p0 + k * [N]), 1) \\
 & & \\
X & \stackrel{\text{def}}{=} & p.\surd : Y + (1-p).\surd : X \\
Y & \stackrel{\text{def}}{=} & 1.\surd : Y \\
 & & \\
Population & \stackrel{\text{def}}{=} & X\{5\}\lceil\{\surd\}
\end{array}
$$

Figure 2.1: Very simple model: X becomes Y with functional probability p

"recover", "attempt to pass on infection" and "do nothing". In the mean field case, for realistic parameter values, it should generally be true that $0 \leq F \leq p_L$ so that the functional parameters can be included in the MFE by substituting for $p = F$.

Example

Fig. 2.1 features a very simple model using functional probabilities with agents of type X becoming Y with probability p, which is a function of the number of X agents. A probability can be a function of the sum of the numbers of more than one type of agent if the sum changes. In this model there are only two types of agents, with $(\lfloor X \rfloor + \lfloor Y \rfloor)$ constant, therefore making p a function of $(\lfloor X \rfloor + \lfloor Y \rfloor)$ would make p constant. The agents X have only two options - become Y or remain X - so $(1-p)$ is a valid probability for $0 \leq p \leq 1$ and we set $p_L = 1$. Here a very small system is defined, consisting of only five agents. However we will demonstrate the changes necessary for larger systems and explain how the expanded description can be written for any finite sized system.

Fig. 2.2 is an expansion of the functional rates macro used in Fig. 2.1. On the first tick in the model priority is used to force the count agent, $Cnta$, to

$$
\begin{aligned}
Cnta &\stackrel{\text{def}}{=} \omega^5.\overline{count}^5 : Cntb5 + \omega^4.\overline{count}^4 : Cntb4 \\
&\quad +\omega^3.\overline{count}^3 : Cntb3 + \omega^2.\overline{count}^2 : Cntb2 \\
&\quad + \omega.\overline{count}^1 : Cntb1 + 1.\surd : Cntb0 \\
Xa &\stackrel{\text{def}}{=} \omega.count : Xb + 1.\surd : Xb \\
Ya &\stackrel{\text{def}}{=} 1.\surd : Yb \\
\\
Cntb5 &\stackrel{\text{def}}{=} \omega.\overline{count5}^5 : Cntc \\
Cntb4 &\stackrel{\text{def}}{=} \omega.\overline{count4}^4 : Cntc \\
Cntb3 &\stackrel{\text{def}}{=} \omega.\overline{count3}^3 : Cntc \\
Cntb2 &\stackrel{\text{def}}{=} \omega.\overline{count2}^2 : Cntc \\
Cntb1 &\stackrel{\text{def}}{=} \omega.\overline{count1}^1 : Cntc \\
Cntb0 &\stackrel{\text{def}}{=} 1.\surd : Cntc \\
Xb &\stackrel{\text{def}}{=} \omega^5.count5 : Xc5 + \omega^4.count4 : Xc4 \\
&\quad +\omega^3.count3 : Xc3 + \omega^2.count2 : Xc2 \\
&\quad + \omega.count1 : Xc1 + 1.\surd : Xc5 \\
Yb &\stackrel{\text{def}}{=} 1.\surd : Yc \\
\\
Cntc &\stackrel{\text{def}}{=} 1.\surd : Cnta \\
Xc5 &\stackrel{\text{def}}{=} p_5.\surd : Ya + (1-p_5).\surd : Xa \\
Xc4 &\stackrel{\text{def}}{=} p_4.\surd : Ya + (1-p_4).\surd : Xa \\
Xc3 &\stackrel{\text{def}}{=} p_3.\surd : Ya + (1-p_3).\surd : Xa \\
Xc2 &\stackrel{\text{def}}{=} p_2.\surd : Ya + (1-p_2).\surd : Xa \\
Xc1 &\stackrel{\text{def}}{=} p_1.\surd : Ya + (1-p_1).\surd : Xa \\
Yc &\stackrel{\text{def}}{=} 1.\surd : Ya \\
\\
Population &\stackrel{\text{def}}{=} Xa\{5\}\lceil\{\surd\}
\end{aligned}
$$

Figure 2.2: System described by Fig. 2.1 without using functional probability

interact with as many Xa agents as possible, up to a maximum of five. $Cnta$ then becomes one of the $Cntb_J$ agents $Cntb_0, Cntb_1, ...Cntb_5$ dependent on the number of contacts made. At the second stage of the model the count agent communicates with all of the Xb agents performing an action that depends on the number of Xa agents communicated with at the previous stage. Depending on the action performed the Xb agents become one of the agents $Xc_1, Xc_2, ...Xc_5$, which have the probability of becoming Ya, p, set to one of the values $p_1, p_2, ...p_5$. In this way p is calculated on each cycle of the model depending on the number of Xa agents present on the first tick of the cycle.

If we consider the function defined for p in Fig. 2.1

$$p \stackrel{\text{prob}}{=} \min(max(p_0 + k\lfloor X \rfloor, 0), 1) \ ,$$

and we wish to have, for instance, $p_0 = 0.1$ and $k = 0.01$ we would achieve this in the expanded model by adding the following definitions to the start of the model:

$$\begin{aligned} p_1 &= 0.11 \\ p_2 &= 0.12 \\ p_3 &= 0.13 \\ p_4 &= 0.14 \\ p_5 &= 0.15 \ , \end{aligned}$$

with the probabilities derived from the equation above. Consider now if the system were expanded to six agents. Note that if an Xb agent is unable to perform any of the *count* actions it becomes $Xc5$. This means that if there

are six Xb agents five of them will perform $count5$ and become $Xc5$ while the other will also become $Xc5$, performing $\surd$. However, if we have six X agents we want them to have $p = 0.16$ by the function for p in Fig. 2.1. To allow this we need to change the expanded model. Firstly we need to extend the $Cnta$ and Xb agents so that they are able to handle systems of six X agents:

$$\begin{aligned}
Cnta &\stackrel{\text{def}}{=} \omega^6.\overline{count}^6 : Cntb6 + \omega^5.\overline{count}^5 : Cntb5 + \\
&\qquad \omega^4.\overline{count}^4 : Cntb4 + \omega^3.\overline{count}^3 : Cntb3 + \\
&\qquad \omega^2.\overline{count}^2 : Cntb2 + \omega.\overline{count} : Cntb1 + 1.\surd : Cntb0 \,, \\
Xb &\stackrel{\text{def}}{=} \omega^6.count6 : Xc6 + \omega^5.count5 : Xc5 + \omega^4.count4 : Xc4 + \\
&\qquad \omega^3.count3 : Xc3 + \omega^2.count2 : Xc2 + \omega^1.count1 : Xc1 + \\
&\qquad 1.\surd : Xc5 \,.
\end{aligned}$$

In addition we require agents $Cntb6$ and $Xc6$ that will be used when there are six X agents, as well as defining the probability p_6:

$$\begin{aligned}
p_6 &= 0.16 \,, \\
Cntb6 &\stackrel{\text{def}}{=} \omega.\overline{count6}^6 : Cntc \,, \\
Xc6 &\stackrel{\text{def}}{=} p_6.\surd : Ya + (1 - p_6).\surd : Xa \,.
\end{aligned}$$

By extension of this approach we can extend the size of the system described by the model so that we can consider systems of any fixed finite number of agents. Considering a situation where the system can become very large highlights the advantages of writing the model in terms of functional parameters. For instance it is not unusual for biological systems to consist of thousands of individuals and if we were to extend the model in Fig. 2.2 to

feature even 100 agents the $Cnta$ and Xb agent definitions would each have 101 terms and there would have to be 100 different Xc agents and 101 different $Cntb$ agents defined. In contrast using the functional parameters and aggregation notation we can describe such a system with the model in Fig. 2.1, which features only two small agent definitions and the definition for the functional probability p. In Chapter 3 we present a rigorous method for deriving MFE and when doing this we use the compact form of the model.

2.2.2 Functional Parallel Agents

In a similar way to the functional selection of probabilities described in the previous section, the number of agents of a given type in a parallel agent can be made a function of the numbers of agents in the population. The function is defined at the start of the file thus

$$label \stackrel{\text{int}}{=} function\ .$$

Since the number of components in a parallel agent must always be an integer, I, we make use of the function $Round(x)$, which chooses the integer closest to x. In addition I must always be positive so if the function, F, is inversely proportional to the number of the agents being counted we need $I \stackrel{\text{int}}{=} max(Round(F), 0)$. Further, to have a practical implementation of the model it may be necessary to have an upper limit, i.e. $I \stackrel{\text{int}}{=} min(max(Round(F), Im), 0)$. However, in theory we could continue adding to the model at each iteration, while always having a finite system size.

$$
\begin{aligned}
I &\stackrel{\text{int}}{=} \text{Round}(I0 + K * \lfloor X1 \rfloor) \\
\\
X1 &\stackrel{\text{def}}{=} p.\surd : Y2 + (1-p).\surd : X2 \\
Y1 &\stackrel{\text{def}}{=} 1.\surd : Y2 \\
\\
X2 &\stackrel{\text{def}}{=} Xh2 \times Xc2\{I\} \\
Xh2 &\stackrel{\text{def}}{=} \omega.act : X1 + 1.\surd : X1 \\
Xc2 &\stackrel{\text{def}}{=} \omega.act : 0 + 1.\surd : 0 \\
Y2 &\stackrel{\text{def}}{=} \omega.act : X1 + 1.\surd : Y1 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} X1\{5\}\lceil\{\surd\}
\end{aligned}
$$

Figure 2.3: Two stage simple model: Y goes back to X after communication with X

Example

The model in Fig. 2.3 is a two stage extension of the model from Fig. 2.1, which makes use of parallel agents. In models of biological systems, such parallel agents can be used to model a situation where a single individual can interact with several other agents. For example, in a disease system an infected individual may interact with, and potentially pass the disease to, several susceptible individuals. If the overall population rises individuals may be expected to make more contacts in a fixed time and this would be captured in the model by making the parallel agent functional. The $X1$ agents here become $Y2$ with fixed probability p. In the second stage of the model the $X2$ agent is a parallel agent consisting of one $Xh2$ agent, which will become $X1$ at the next tick, and I $Xc2$ agents, which can perform the output action $\overline{act}$. The number of $Xc2$ agents in each $X2$, I, is a function of the number of $X2$ agents currently present. Again the function could depend on the numbers of multiple types of agents but in this example we

$$
\begin{aligned}
Cnt1 &\stackrel{\text{def}}{=} 1.\surd : Cnt2a \\
X1 &\stackrel{\text{def}}{=} p.\surd : Y2a + (1-p).\surd : X2a \\
Y1 &\stackrel{\text{def}}{=} 1.\surd : Y2a \\
\\
Cnt2a &\stackrel{\text{def}}{=} \omega^5.\overline{count}^5 : Cnt2b5 + \omega^4.\overline{count}^4 : Cnt2b4 \\
&\quad \omega^3.\overline{count}^3 : Cnt2b3 + \omega^2.\overline{count}^2 : Cnt2b2 \\
&\quad + \omega.\overline{count} : Cnt2b1 + 1.\surd : Cnt2b0 \\
X2a &\stackrel{\text{def}}{=} \omega.count : X2b + 1.\surd : X2b \\
Y2a &\stackrel{\text{def}}{=} 1.\surd : Y2b \\
\\
Cnt2b5 &\stackrel{\text{def}}{=} \omega.\overline{count5}^5 : Cnt2c \\
Cnt2b4 &\stackrel{\text{def}}{=} \omega.\overline{count4}^4 : Cnt2c \\
Cnt2b3 &\stackrel{\text{def}}{=} \omega.\overline{count3}^3 : Cnt2c \\
Cnt2b2 &\stackrel{\text{def}}{=} \omega.\overline{count2}^2 : Cnt2c \\
Cnt2b1 &\stackrel{\text{def}}{=} \omega.\overline{count1} : Cnt2c \\
Cnt2b0 &\stackrel{\text{def}}{=} 1.\surd : Cnt2c \\
X2b &\stackrel{\text{def}}{=} \omega^5.count5 : X2c5 + \omega^4.count4 : X2c4 \\
&\quad +\omega^3.count3 : X2c3 + \omega^2.count2 : X2c2 \\
&\quad + \omega.count1 : X2c1 + 1.\surd : X2c5 \\
Y2b &\stackrel{\text{def}}{=} 1.\surd : Y2c \\
\\
Cnt2c &\stackrel{\text{def}}{=} 1.\surd : Cnt1 \\
Y2c &\stackrel{\text{def}}{=} \omega.act : X1 + 1.\surd : Y1 \\
X2c5 &\stackrel{\text{def}}{=} Xh2c \times Xc2c\{C5\} \\
X2c4 &\stackrel{\text{def}}{=} Xh2c \times Xc2c\{C4\} \\
X2c3 &\stackrel{\text{def}}{=} Xh2c \times Xc2c\{C3\} \\
X2c2 &\stackrel{\text{def}}{=} Xh2c \times Xc2c\{C2\} \\
X2c1 &\stackrel{\text{def}}{=} Xh2c \times Xc2c\{C1\} \\
Xh2c &\stackrel{\text{def}}{=} \omega.act : X1 + 1.\surd : X1 \\
Xc2c &\stackrel{\text{def}}{=} \omega.\overline{act} : T + 1.\surd : T \\
\\
Population &\stackrel{\text{def}}{=} X1\{5\}\lceil\{\surd\}
\end{aligned}
$$

Figure 2.4: System described by Fig. 2.3 without using functional parallel agent

have a fixed system size in terms of $(\lfloor X2 \rfloor + \lfloor Y2 \rfloor)$, since no individuals are added or removed from the system. The $Xc2$ agents are prioritised to perform the *act* output action and if they communicate with a $Y2$ agent it will become $X1$ while if they communicate with an $Xh2$ it is unchanged by communication.

Full implementation of this model without the functional parallel agent, which can be seen in Fig. 2.4, once again makes use of a Cnt agent to establish the size of the population and choose from several $X2c$ agents. In this case the counting phase comes after the initial probabilistic stages. Depending on the values of the parameters $I0$ and K it is possible that multiple $X2c$ agents $(X2c\_0, X2c\_1, ..., X2c\_5)$ will feature the same number of $Xc2c$ agents. For example, $I0 = 0$ and $K = 0.5$, would mean that $C1 = C2$ and $C3 = C4$.

2.2.3 General Implementation

The expanded form of a model that uses functional parameters follows the same steps whether it is for a functional probability or functional parallel agents. The difference between the two types of functional parameter comes only in the restrictions that apply to the form of the function, and the form that the agents featuring the functional parameter must take. The general implementation of the expanded form follows the steps below. This general implementation assumes that the agent names include an index (1,2...n) to indicate at which stage of the model the agent is defined. If the ith stage features a functional parameter it will be replaced by three stages, ia, ib, ic.

- An agent $Cnti$ is added at each stage i. At stages that do not feature functional parameters this agent deterministically moves on to the next stage.

- The stages that do feature functional parameters are replaced by three separate stages.

    1. – The count agent, $Cntia$ is prioritised to perform as many instances as possible of the *count* output action.
       – The agents whose number is required by the functional parameter all perform the input action *count*.
       – All other agents at the first additional stage deterministically progress to the second additional stage.
    2. – The $Cntib$ agent performs n instances of the *countn* output action, where n is the number of agents interacted with at the previous stage.
       – Each of the agents that were counted at the previous stage can perform a number of input actions *countj* where the upper limit of j is $\geq n$.
       – The agents that make use of the functional parameter evolve to a state - e.g. Xic_n - where the parameter is set to satisfy the functional parameter for the value of n.
       – The agents that were counted but do not feature the functional parameter evolve to the same state irrespective of which *countj* action is performed.
    3. – The $Cntic$ agent deterministically progresses to the next stage.
       – There are a number of agents, Xic_n, that represent the agents with functional parameters.
       – Agents of these types are all in the states where the parameter satisfies the function.

  - All other agents are of the same form as in the functional parameters form of the model.

- During each iteration of the model the following steps may be necessary:
  - Rewrite the agent definition of $Cntia$ so that it can communicate with all of the agents to be counted.
  - Add a $Cntib_n$ agent where n is the number of agents contacted.
  - Add an Xic_n agent.

2.3 Summary

In this chapter some additional WSCCS notation has been defined that allows us to simplify the process of writing more complex models. This notation does not add to the expressive capabilities of WSCCS since all models that make use of the new notation could be written using the standard notation. However using the functional parameters notation density dependent behaviour can be captured in simple agent definitions, making it straightforward to define MFE for such models; and using the aggregation notation it is realistic to define models of systems consisting of hundreds or thousands of individuals. Throughout the remainder of the thesis the new notation will be used without comment on how the models would be implemented using only the standard notation.

Chapter 3

Deriving Mean Field Equations

Traditionally, process algebra models are studied either by exploring the Markov chain or by performing Monte Carlo simulations of the system. Exploring the Markov chain involves calculating the entire state space, which is computationally expensive, and for large systems is not possible. Probabilistic workbench [88], the tool for WSCCS, can handle systems up to 500 components [89] but imposes restrictions on how these models can be interpreted. More generally we can consider only small systems of up to 20 components. Individual simulations give only a single route through the state space so that to calculate the average behaviour of the system it is necessary to perform many simulations, which again is computationally expensive.

Another method of studying process algebra models is to develop models that describe the system at two levels of abstraction, and to show that the two descriptions are equivalent (for example using bisimulation). Although we do not explicitly develop WSCCS models for the population

level behaviour we take a similar approach by deriving mean field equations (MFE) that describe the average behaviour of the system at the population level. In Section 3.4.2 we make use of a limit theorem [58] to demonstrate that MFE are equivalent to the mean of the Markov chain. MFE offer a deterministic approximation to the mean behaviour of the model and are amenable to a wide range of algebraic and numerical analyses. The MFE are useful when we are interested in the average state of the system at each step of time. If instead we are interested, for instance, on the average time until completion of some system, or the likelihood of reaching a deadlock state, the other analysis techniques are more useful.

The previous studies that have derived MFE from WSCCS models [73, 82] have made use of intuitive reasoning to establish equations to describe the mean behaviour of the model. This was possible because those models consisted of only a small number of types of agents, with the average behaviour being relatively easy to identify; however it is preferable to have formal rules for deriving the MFE and also for larger, more complex models it is not so easy to intuitively determine the mean behaviour of the system.

This chapter introduces an algorithm for deriving MFE for a WSCCS description of a system, which will be used in subsequent chapters to obtain MFE for the models presented. The algorithm can be thought of as generating an alternative semantics for WSCCS allowing us to easily obtain the mean behaviour of a model in a way that is consistent with the standard Markov chain semantics. Our work differs from that of Cardelli [26], Brodo et al. [21] and Hillston [45] because of the nature of the process algebra used (discrete vs continuous, probabilistic choice vs stochastic rates) and the particular application area (epidemiology). The work was carried out independently. Essentially, the problem tackled here is the difficulty of

deriving transition rates from a calculus where choices are probabilistic (but there is no rate information on transitions).

The question addressed with all of the models presented in this thesis is "on average how many individuals of each type are present after each iteration of the model?" However, WSCCS models can be written to answer questions such as "what is the average time until completion?" or "what is the probability of reaching a deadlock state?" and deriving MFE for such models will not offer any insight into these questions.

The MFE that we obtain offer a way to circumvent the well known state space explosion problem. A model with a small definition and a moderately large number of agents (> 20 for WSCCS) leads to a Markov chain that is so large analysis becomes impractical. This is a problem that affects not only models of biological systems but any system that involves many instances of a small number of types of agents. Our algorithm is therefore useful for many computing science application areas where state space explosion is a problem.

3.1 Model building

The model featured in Fig. 3.1 will be used to illustrate the use of the algorithm. This model is based on the prioritised communication model of Norman and Shankland [73], with the main difference being that the order in which probabilistic choice and communication happen is reversed. To aid understanding of the differences between these two models, which we discuss in Section 3.3.5, we present flow diagrams which indicate the transitions that each type of agent can make. The flow diagram for Fig. 3.1 can be found in Fig. 3.2.

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} \omega.infect : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} T1 \times Trans \\
T1 &\stackrel{\text{def}}{=} \omega.infect : I2 + 1.\surd : I2 \\
Trans &\stackrel{\text{def}}{=} \omega.\overline{infect} : 0 + 1.\surd : 0 \\
R1 &\stackrel{\text{def}}{=} \omega.infect : R2 + 1.\surd : R2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
SI2 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
I2 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1 - p_r).\surd : I1 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
\\
Population &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 3.1: Disease model with contact followed by probabilistic choice

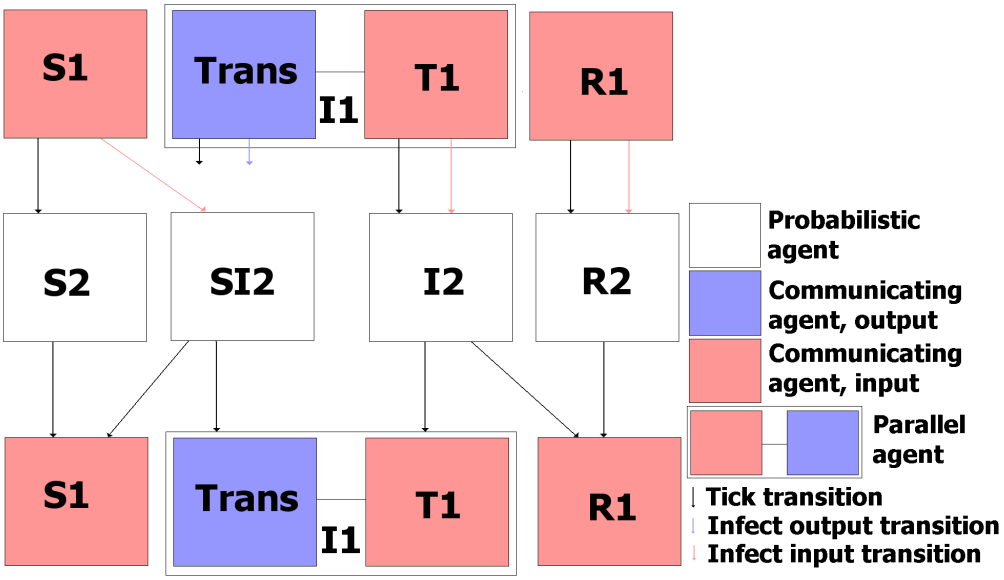

Figure 3.2: Flow diagram for Fig. 3.1

The first stage in Fig. 3.1 is the contact stage where the infected individuals are represented by a parallel agent consisting of one $T1$ and one $Trans$ agent. The $Trans$ agents are prioritised to perform the *infect* output action and can do so by communicating with an $S1, T1$ or $R1$ agent. If an $S1$ agent interacts with a $Trans$ it will become the agent $SI2$, which subsequently make a probabilistic choice to become infected or not. The $T1$ and $R1$ agents can also interact with the $Trans$ agents but their future behaviour is unaffected by doing so, since they always evolve to $I2$ and $R2$ respectively. This is included to capture the situations where infecteds make contact with other infecteds, which cannot become more infected, or recovereds, which are immune to future infection. (In the flow diagram the transitions from the $Trans$ agent do not lead to any other agent. This represents transition to the null agent 0.)

The second stage in this model features probabilistic choice. In addition to the $SI2$ agents making the choice to become infected or not the $I2$ agents also recover probabilistically. Defining the model with communication first, followed by probabilistic choice, has the advantage that making an infectious contact will not automatically lead to infection. In many disease systems contact does not guarantee infection and this is captured in the model by the probability p_i: diseases where contact always leads to infection can be modelled by setting $p_i = 1$. For this reason all of the subsequent models in this thesis will have the contact stage followed by a probabilistic stage, where the susceptible individuals that have made contact with an infected individual make a probabilistic choice to become infected.

In addition this model has all infected individuals able to pass on the disease in each iteration of the model, and, since communication is prioritised, if they can make an infectious contact they will. In most disease systems

infection leads an individual to make potentially infectious contacts: for example a fox with rabies tends to roam outside of its normal territory and become more aggressive, leading it to fight with other foxes and potentially pass on the disease. It may still be desirable to have the mean number of contacts that individuals make < 1. This could be achieved by having a probabilistic stage before the contact stage, as well as the probabilistic stage after contact, with infected individuals making a choice to make contact.

3.2 Restrictions

The algorithm presented here cannot be used to obtain mean field equations for every WSCCS model that could possibly be written. Firstly the models should be designed to investigate the numbers of individuals of each type that are present after each iteration of the model.

Secondly the system being considered must be sufficiently large in terms of the numbers of agents. It is a well known result that deterministic models do not accurately capture the behaviour of small systems where stochastic effects can have a great influence. Most importantly in disease systems it is known that the initial number of infected individuals greatly affects the convergence of deterministic equations to a discrete stochastic system [93].

In addition to these more obvious restrictions we also place some restrictions on the way that models must be written to be amenable to the algorithm. These restrictions make it more straightforward to write a model that describes a system and in turn reason about the mean behaviour of the system. In Chapter 7 we consider models of superspreader systems that do not fall within this framework. For these specific cases we are able to derive equations by carefully considering the one step behaviour over the communication stage, without extending the algorithm to include models

featuring communication of forms not currently covered. Other models for superspreaders are considered and we see that these lead to the same MFE, which can be derived directly using our algorithm. Therefore at present we can say that the restrictions have not proved limiting in terms of the systems that we have been able to model.

1. The algorithm is constructed under the assumption that the model takes the form $P\{p\} \times Q\{q\} \times ... \times Z\{z\}\lceil\{\surd\}$ where the components can be sequential or parallel processes, and may include priority. This is a big restriction but it is sensible in terms of the kind of questions that are addressed.

2. All weights associated with communication must be 1, and for single actions, there should be only one alternative action to the communication action. A consequence of this is that probabilistic choice steps must be separate from communication steps. Generally systems can be reformulated to fit this restriction, therefore, it can be thought of as a renormalisation step rather than a real restriction.

3. There should be at most one communicating action in each agent in any stage. This does not hamper expressivity, since it is possible to put two different communicating actions on different stages. An example of the behaviour we do not allow is found in the agent

$$X1 \stackrel{\text{def}}{=} 1.actionA\#actionB : Y2 + 1.\surd : X2 \ ,$$

which must perform the actions $actionA$ and $actionB$, which may require communication with two separate agents, to become the agent

$Y2$. This behaviour can instead be captured by agents on different stages, for instance

$$\begin{aligned} X1 &\stackrel{\text{def}}{=} 1.actionA : Y1b + 1.\surd : X1b \\ X1b &\stackrel{\text{def}}{=} 1.\surd : X2 \\ Y1b &\stackrel{\text{def}}{=} 1.actionB : Y2 + 1.\surd : X2 \,, \end{aligned}$$

which gives the same requirement for $X1$ to perform both $actionA$ and $actionB$ to become $Y2$, and become $X2$ otherwise. This alternative formulation is amenable to the algorithm described here and MFE can be derived for models including communication of this form. It is, however, unclear if this requirement for individual agents to perform two communicating actions to make one transition will be necessary in describing any biological system.

4. Agents performing the input action perform only a single instance, and may evolve to different states depending on whether it performs the input action or the free action. This is a special case of restriction 3, where $actionB = actionA$, and can be handled in the same way.

5. Agents performing a single instance of the output action may evolve to different states, depending on whether they communicate or not; however, agents that perform multiple instances of the action must evolve to the same state, regardless of whether they communicate or not (and irrespective of how many instances of the action they perform). Biologically there seems to be little need to allow evolution to different states depending on the number of instances of an action performed.

6. Processes should not include nested permission sets, i.e. all communication takes place between all processes (potentially), and not between subgroups defined by restriction. The reason for this is that the restriction operator cannot be distributed over parallelism. From a modelling perspective, this appears to be a reasonable restriction, being equivalent to assuming random mixing since all agents can (potentially) communicate with all others. It would, however, be possible to develop models that circumvent this restriction by renaming the action being performed. For instance, agents X_A and Y_A could communicate on $actionA$ while agents X_B and Y_B communicate on $actionB$.

Restrictions 2, 4 and 5 make the definition of the general terms for changing agents defined in Section 3.3 simpler; however, it should be possible to remove these restrictions in future work.

3.3 Mean Field Equations

In this section we look at how models evolve and present the algorithm with which we derive MFE. Using the algorithm we derive equations that describe the numbers of each type of agent in terms of the agents in the population at the previous stage (one stage equations). These equations can be algebraically manipulated to give the final MFE describing the behaviour of the model over the several stages that make up an iteration of the model.

The MFE that are derived from our models will always be first order difference equations, i.e. of the general form

$$\mathbf{X}_{t+1} = f(\mathbf{X}_t) \,, \tag{3.1}$$

where $\mathbf{X}$ is the vector of the different agent types for which we are deriving equations,

$$\mathbf{X} = \begin{pmatrix} S \\ I \\ R \end{pmatrix}$$

for simple epidemic models. The fact that the MFE are first order is a consequence of the Markovian nature of WSCCS, which means that the future state of the system depends only on the current state of the system and not on the previous states.

It is always possible to obtain higher order equations by substitution. In the general case a second order difference equation can be constructed by noting that (3.1) implies that

$$\mathbf{X}_{t+2} = f(\mathbf{X}_{t+1}) \ ,$$

and substituting for $\mathbf{X}_{t+1}$ to find

$$\begin{aligned} \mathbf{X}_{t+2} &= f(\mathbf{X}_{t+1}) \\ &= f(f(\mathbf{X}_t)) \ . \end{aligned}$$

Such second order equations, however, ignore important information about $\mathbf{X}_{t+1}$, the state of the system after the intermediate timestep, and are likely to be algebraically more complicated. In Chapter 6 we derive MFE for which we can eliminate equations that describe the quantity of infection in the environment. This leads to a system of second order difference equations,

$$\mathbf{X}_{t+1} = f(\mathbf{X}_t, \mathbf{X}_{t-1}) \ ,$$

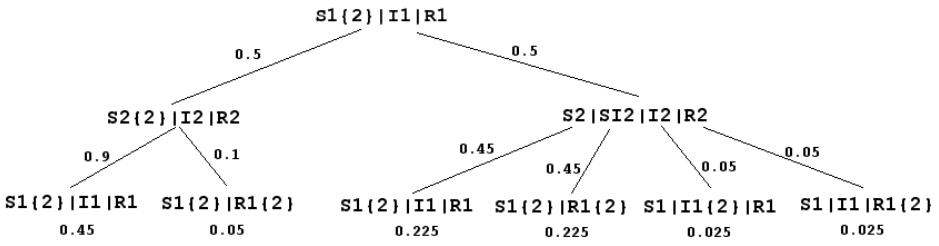

Figure 3.3: Two stage Markov chain for Fig. 3.1 with initial population of $S1\{2\} \times I1 \times R1$, $p_i = 0.5$, $p_r = 0.1$

which describe the population. In that case the individual second order equations are not significantly more complicated and the system of equations is less complicated since fewer equations are required to describe the population.

For models featuring only probabilistic choice the derived equations will be first order linear difference equations (i.e. each term in f is linear in one of the components of $\mathbf{X}$) and for models that feature communication they will be non-linear.

The diagram in Fig. 3.3 illustrates, by a numeric example for a small initial population of $S1\{2\} \times I1 \times R1 = S1\{2\} \times T1 \times Trans \times R1$, the progress of the model in Fig. 3.1 over two stages. After the first stage the system can either be in the state $S2\{2\} \times I2 \times R2$, with probability 0.5 , or $S2 \times SI2 \times I2 \times R2$, with probability 0.5 . From $S2\{2\} \times I2 \times R2$ the system can progress to $S1\{2\} \times I1 \times R1$, with probability 0.9 , or $S1\{2\} \times R1\{2\}$, with probability 0.1 . Alternatively, from $S2 \times SI2 \times I2 \times R2$ the system can progress to $S1\{2\} \times I1 \times R1$, with probability 0.45 , $S1\{2\} \times R1\{2\}$, with probability 0.45 , $S1 \times I1\{2\} \times R1$, with probability 0.05 , or $S1 \times I1 \times R1\{2\}$, with probability 0.05 . There are therefore 4 distinct states that the system can enter after two stages - $S1\{2\} \times I1 \times R1$, $S1\{2\} \times R1\{2\}$, $S1 \times I1\{2\} \times R1$

and $S1 \times I1 \times R1\{2\}$, with respective probabilities of reaching these states of 0.675 , 0.275 , 0.025 and 0.025 - with multiple routes by which the system can reach some of these states. Taking into account the probabilities of entering the different states after two stages we can calculate the mean population, which consists of $S1\{1.95\} \times I1\{0.75\} \times R1\{1.3\}$.

The algorithm allows us to deduce this last fact based only on the syntactical description of the model. No state space calculation is required. For the example considered in Fig. 3.3, calculation of the state space is relatively simple so that the advantage offered by the MFE is minimal: however we are interested in studying much bigger systems, consisting of potentially thousands of agents. In such cases calculating the mean behaviour from the MFE does not involve any more work than for a small system. In contrast computing the mean of the Markov chain requires calculation of the entire state space, which is not possible for such large systems using the tool for WSCCS [88]. The alternative method of calculating the mean is to perform many simulations of the system and calculate the mean of those. Compared to calculating the mean from the MFE performing sufficiently many simulations to accurately find the mean (at least several hundred) is computationally expensive; therefore, MFE can be said to offer a straightforward method of describing the mean of a system.

3.3.1 State transition table

The algorithm makes use of a state transition table that represents the mean evolution of the system, and from which we can construct a one stage MFE for each type of agent in the model. These one stage equations can be manipulated by standard algebraic techniques to produce a single system of MFE to describe the mean behaviour of the system over an iteration of the

model.

The rows of the state transition table are the agent types in the model and the actions that they can perform. For instance the agent

$$S1 \stackrel{\text{def}}{=} 1.\mathit{infect} : SI2 + 1.\surd : S2 \ ,$$

from Fig. 3.1, gives two rows labelled $S1 \ \ \mathit{infect}$ and $S1 \ \ \surd$. For agents that always evolve to the same state and for parallel agents, whose contribution to the model does not depend on performing actions, $*$ is used in place of action names. This means that the agent

$$T1 \stackrel{\text{def}}{=} 1.\mathit{infect} : I2 + 1.\surd : I2$$

gives a row labelled $T1 \ \ *$, since the $T1$ agents all evolve to $I2$.

The columns of the table are labelled with the names of the agent types in the model and represent the state that the system evolves to, at time$= t$. The content of the cell $(Ax \ \ aj, Ai)$ is the expression $AxajAi_{new}$, representing the number of agents of type Ax that perform the action aj and become Ai. By summing the column for the agent Ai we obtain an equation for Ai_t, the number of Ai agents at time t, in terms of the numbers of the agents present in the population at time $t-1$, and for some types of agent the numbers present at time t. These one stage equations are then combined and manipulated using standard algebraic techniques to give a system of MFE for the behaviour of the agents of interest in the model.

		X	Y
X	$\surd$	XtX_{new}	XtY_{new}
Y	$\surd$		YtY_{new}

Table 3.1: State transition table for Fig. 2.1

For example if we consider the simple model from Fig. 2.1, which consists of the agents

$$\begin{aligned} X &\stackrel{\text{def}}{=} p.\surd : Y + (1-p).\surd : X \\ Y &\stackrel{\text{def}}{=} 1.\surd : Y \ , \end{aligned}$$

the state transitions are given by Table 3.1. We then sum columns to give the following generalised system of MFE for the model:

$$\begin{aligned} X_t &= XtX_{new} \ , \\ Y_t &= XtY_{new} + YtY_{new} \ . \end{aligned} \tag{3.2}$$

No substitution is necessary for this model since it is a one stage model, i.e. we are interested in the numbers of X and Y agents at any given time.

In order to use MFE to describe the evolution of the mean of the system we need to express, for example, XtY_{new} and YtY_{new} in terms of X_{t-1} and Y_{t-1}, the numbers of X and Y agents at the previous timestep. In general the form of these terms depends on the type of agent involved. In Section 3.3.2 we examine the form of the $AxajAi_{new}$ terms for the different types of agents in our models.

3.3.2 How agents evolve

There are three basic types of agent utilised in our models: probabilistic agents, communicating agents and parallel agents. In this section we look at the ways in which each of these types of agent evolve and how they contribute to the MFE.

Probabilistic agents

Calculation of $Axaj Ai_{new}$ is straightforward for steps involving only probabilistic choice. Probabilistic agents take the form

$$A0 \stackrel{\text{def}}{=} w_1.a : A1 + w_2.a : A2 + ...w_m.a : Am \ ,$$

and proceed independently without communicating with any other agent. (This assumes that a is a free action. In all of the models presented in this thesis $\surd$ is the only free action but any action could be defined as free and the same principle applies.) We generally write the weights in such an agent as probabilities, i.e.

$$\forall i : \qquad 0 \leq w_i \leq 1 \qquad \text{and} \qquad \sum_{i=1}^{m} w_i = 1 \ ,$$

to simplify the process of writing models and deriving MFE, although this is not a restriction that is required to use the algorithm. When the weights are not written in this form, the probability that $A0$ will become one of its destination processes Ai is

$$p_i = \frac{w_i}{\sum_{j=1}^{m} w_j} \ .$$

The evolution of each agent of this type is independent of the evolution

of the others and the evolution of a large number of these is governed by the multinomial distribution. This is a generalisation of the binomial distribution, which is the probability distribution of the number of "successes" from n independent Bernoulli (yes/no) trials with the same probability of "success". From the theory of multinomial distributions we know that the mean number of $A0$ agents that become one of the agents Ai is

$$A0tAi_{new} = p_i A0_{t-1} \ ,$$

where p_i is the probability that $A0$ will become Ai. In terms of the algorithm, this means that for each Ai the term $p_i A0_{t-1}$ replaces $A0tAi_{new}$ in the cell $(A0 \ \ t, Ai)$ of the state transition table. If any of the probabilities is functional the functional form should be used in the table, with, for example, $\lfloor A0 \rfloor = A0_{t-1}$.

Example We can now return to Table 3.1 since the model in Fig. 2.1 features only probabilistic agents. The Y agents remain as Y so that

$$YtY_{new} = Y_{t-1} \ .$$

The X agents however make a probabilistic choice, either becoming Y with probability p or becoming Y with probability $(1-p)$, which gives us

$$\begin{aligned} XtY_{new} &= pX_{t-1} \ , \\ XtX_{new} &= (1-p)X_{t-1} \ . \end{aligned}$$

Substituting for YtY_{new}, XtX_{new} and XtY_{new} in (3.2) gives the following system of MFE for this model:

$$\begin{aligned} X_t &= (1-p)X_{t-1}\ , \\ Y_t &= pX_{t-1} + Y_{t-1}\ . \end{aligned}$$

Parallel agents

Parallel agents take the form

$$A0 \stackrel{\text{def}}{=} A1\{n1\} \times A2\{n2\} \times ...Am\{nm\}\ .$$

This means that the agent $A0$ consists of $n1$ agents of type $A1$, $n2$ agents of type $A2$ ... nm agents of type Am, which all behave independently of one another. With such an agent the term $n_i A0_t$ is added to the cell $(A0\ \ *, Ai)$, since the agent $A0$ is instantly replaced by each of the agents $A1, A2...An$. For example the term added to the cell $(A0\ \ *, A1)$ will be $n_1 A0_t$. For any of the numbers of agents $(n1, n2, ..., nm)$ that are functional the functional form should be used in the table, with, for example, $\lfloor A0 \rfloor = A0_t$.

Communicating agents

For communicating agents, the mean number of agents that successfully communicate, and evolve to a different state by communicating than if they had not communicated, depends on the mix of agents available to perform the input and output actions. We consider a general system with agents S, Ti and Wi. S is the agent for which the communicating proportion is calculated, i.e. the Ax in the table row, or the state we are moving from. Ti are the agents that interact with S, e.g. the infecteds, or the agents

who have the output action. Wi are the other agents that interact with Ti. These may be regarded as being in competition with the S since they may absorb instances of the action. For example, in the SIR system of Fig. 3.1, this is equivalent to communication between an infected and a recovered. An opportunity to infect a susceptible has been missed.

To calculate mean behaviour of the system we must consider all possible ways in which the system can evolve. The Operational Rules of WSCCS (Table A.1) include the rule

$$\frac{E \stackrel{w}{\longmapsto} E' \quad F \stackrel{v}{\longmapsto} F'}{E \times F \stackrel{wv}{\longmapsto} E' \times F'}, \tag{3.3}$$

which says that if agent E becomes E' with weight w and agent F becomes F' with weight v, then the parallel agent $E \times F$ becomes $E' \times F'$ with weight wv. By extension the weight with which a population of agents makes any transition is the product of the weights of the individual transitions that occur. This means that, since all weights in agents that can perform communication are 1, the weight with which any combination of actions can occur will be 1. This is unaffected by priority, which merely constrains the actions that are possible.

Different weights of population level changes come about because of the condition (described in Section A.2) that processes are multi-related by weight. This means that if a process (in this case the system process) can evolve to the same state in more than one way, the cumulative weight with which it makes the transition is the sum of all the weights with which the transition can occur. Since all of the weights in our communicating stages are 1 we must calculate the number of possible unordered choices

of which agents communicate to make a given transition. Formally, the binomial coefficient,

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} ,$$

represents the number of unordered ways to choose m objects from a group of n distinct objects. In cases where m lies outside of the range $n \geq m \geq 0$ the binomial coefficient is defined as

$$\binom{n}{m} = 0 .$$

The generalised form of the binomial coefficient is the multinomial coefficient

$$\binom{n}{k_1, k_2, ..., k_m} = \frac{n!}{k_1!k_2!...k_m!} ,$$

which is the number of unordered ways of dividing n distinct objects into m groups with k_i the number of objects in the ith group. If $k_i < 0$ for any of the k_i or if $\sum_i k_i > n$ the multinomial coefficient is defined as

$$\binom{n}{k_1, k_2, ..., k_m} = 0 ,$$

and the particular form of the multinomial coefficient depends on the type of communication being considered.

Given these definitions, there are four general cases covering all the types of model for which we can currently derive terms for the number that communicate, arising from: prioritised or non-prioritised communication, and single or multiple instances of the output action.

Prioritised, Single This is the form of communication used in Fig. 3.1. We will look in detail at the number of the $S1$ agents that communicate for

a small population size and then generalise this to find an expression for the number of $S1$ agents that communicate. This is then further generalised to give a general term for models featuring communication of this form.

The communicating agents in Fig. 3.1 are:

$$\begin{aligned} S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \,, \\ I1 &\stackrel{\text{def}}{=} T1 \times Trans \,, \\ T1 &\stackrel{\text{def}}{=} \omega.infect : I2 + 1.\surd : I2 \,, \\ Trans &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : 0 + 1.\surd : 0 \,, \\ R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \,. \end{aligned}$$

Relating this to our generalised system, $S1$ are the S agents for which we are interested in calculating the number that communicate; $Trans$ are the only Ti, the agents that can communicate with $S1$; and $T1$ and $R1$ are the Wi agents, which can also communicate with $Trans$.

Numerical example If we consider the specific situation where the system is made up of two $S1$ agents, two $I1$ agents (giving two $T1$ and two $Trans$ to take part in the communication) and one $R1$ agent we can have either none, one or two of the $S1$ agents communicating to become $SI2$. The weight with which population level transitions occur will come from the product of binomial coefficients

$$\binom{2}{s}\binom{2}{t}\binom{1}{r}\binom{2}{x} \,,$$

where s, t, r and x are respectively the numbers of $S1, T1, R1$ and $Trans$ agents that must communicate for the transition to occur. The case where none of the $S1$ communicate can come about in two different ways: both of

the $Trans$ communicate with a $T1$, meaning the $R1$ cannot communicate, which happens with weight

$$\binom{2}{0}\binom{2}{2}\binom{1}{0}\binom{2}{2} = 1\ ;$$

or one of the $Trans$ communicates with the $R1$ and the other $Trans$ communicates with one of the $T1$, which happens with weight

$$\binom{2}{0}\binom{2}{1}\binom{1}{1}\binom{2}{2} = 2\ .$$

Similarly one $S1$ can communicate in two ways: one of the $Trans$ agents communicates with one of the $S1$ agents and the other communicates with one of the $T1$ meaning that the $R1$ agent cannot communicate, which happens with weight

$$\binom{2}{1}\binom{2}{1}\binom{1}{0}\binom{2}{2} = 4\ ;$$

or the second $Trans$ agent communicates with the $R1$ with neither of the $T1$ communicating, which gives the weight

$$\binom{2}{1}\binom{2}{0}\binom{1}{1}\binom{2}{2} = 2\ .$$

Finally both of the $S1$ agents communicate, which means that none of the $T1$ or $R1$ agents can communicate and this happens with the weight

$$\binom{2}{2}\binom{2}{0}\binom{1}{0}\binom{2}{2} = 1\ .$$

The mean number of $S1$ agents that communicate is the weighted average of the numbers communicating in these different options,

$$\frac{0 \times (1+2) + 1 \times (4+2) + 2 \times 1}{(1+2) + (4+2) + 1} = \frac{4}{5} \, .$$

Example of symbolic size In general the average number of successful communications, e.g. new infections, will be

$$\frac{\sum_{i=0}^{n} i \times f_i}{\sum_{i=0}^{n} f_i}, \tag{3.4}$$

where n is the maximum possible number of new infections and f_i is the combined weight with which i new infections will occur. We now generalise to consider a population that consists of a $S1$ agents, b $I1$ agents (giving b $T1$ and b $Trans$) and $n - a - b$ $R1$ agents where n is the total number of individuals in the population. For this model there will always be sufficiently many agents for all of the $Trans$ agents to communicate with, because for each $Trans$ there is a $T1$ and therefore we can discount the influence of the number of $Trans$ that communicate, since it will always lead to a factor of

$$\binom{b}{b} = 1 \, .$$

This is the source of the term

$$\binom{2}{2}$$

when calculating the weight for each way in which the population can evolve for the numerical example above. The mean number of $S1$ agents that communicate is now given by

$$S1infectSI2_{new} = \frac{\sum_{r=0}^{b} (b-r) \binom{a}{b-r} \sum_{k=0}^{r} \binom{b}{r-k} \binom{n-a-b}{k}}{\sum_{r=0}^{b} \binom{a}{b-r} \sum_{k=0}^{r} \binom{b}{r-k} \binom{n-a-b}{k}} \, , \tag{3.5}$$

which is the generalised form of equation (3.4) for this model. Here r is the number of the $Trans$ agents that communicate with the Wi agents ($T1$ or $R1$) and k is the number of those that communicate with $R1$. We can see that $(b-r)$ is the number of $Trans$ that communicate with $S1$ and

$$\binom{a}{b-r}$$

is the number of ways of choosing which of the a $S1$ agents communicate for a particular population change to occur. Similarly

$$\binom{b}{r-k}$$

is the number of ways of choosing which of the $T1$ communicate and

$$\binom{n-a-b}{k}$$

is the number of ways of choosing which of the $R1$ communicate.

Equation (3.5) is algebraically intractable in this form but can be simplified by several applications of *Vandermonde's convolution* [35],

$$\sum_{k} \binom{j}{m+k}\binom{s}{i-k} = \binom{j+s}{m+i} .$$

The term

$$\sum_{k=0}^{r} \binom{b}{r-k}\binom{n-a-b}{k} ,$$

which appears in the numerator and denominator of $S1infectSI2_{new}$, can be rewritten as

$$\sum_{k} \binom{b}{r-k}\binom{n-a-b}{k} ,$$

since outside of the range $r \geq k \geq 0$ either of the binomial coefficients will be 0. By applying Vandermonde's convolution the term becomes

$$\binom{n-a}{r}.$$

This leads to the denominator becoming

$$\sum_{r=0}^{b}\binom{a}{b-r}\binom{n-a}{r} = \sum_{r}\binom{a}{b-r}\binom{n-a}{r} = \binom{n}{b}.$$

Similarly the numerator of (3.5) becomes

$$\begin{aligned}
\sum_{r=0}^{b}(b-r)\binom{a}{b-r}\binom{n-a}{r} &= \sum_{r}(b-r)\frac{a!}{(b-r)!(a-(b-r))!}\binom{n-a}{r} \\
&= a\sum_{r}\frac{(a-1)!}{(b-r-1)!(a-(b-r))!}\binom{n-a}{r} \\
&= a\sum_{r}\binom{a-1}{b-r-1}\binom{n-a}{r} \\
&= a\binom{n-1}{b-1}.
\end{aligned}$$

It can now be seen that the average number of $S1$ agents that make contact with a $Trans$ agent is

$$\begin{aligned}
S1infectSI2_{new} &= \frac{a\binom{n-1}{b-1}}{\binom{n}{b}} \\
&= \frac{a\frac{(n-1)!}{(b-1)!(n-b)!}}{\frac{n!}{b!(n-b)!}} \\
&= a\frac{(n-1)!}{(b-1)!}\times\frac{b!}{n!} \\
&= \frac{ab}{n} \\
&= \frac{S1_{t-1}I1_{t-1}}{S1_{t-1}+I1_{t-1}+R1_{t-1}}.
\end{aligned}$$

This term applies in the case where we have four types of agents able to communicate on the action: one whose future behaviour depends on whether it performs the input action or not ($S2$), one that can perform the output action ($Trans$), and two that can perform the input action thereby absorbing an instance of the output action ($I2$ and $R2$). We generate a term for each type of agent that has their future behaviour altered by communication: but what happens if we vary the number of absorbing agents or the number of agents performing the output action?

General term We further generalise this process to consider any model that utilises prioritised communication with agents performing the output action able to perform only one. The term that arises is

$$Sa_jAi_{new} = \frac{\sum_r \left(\sum_i T_i - r\right) \binom{S}{\sum_i T_i - r} \sum_{k_1} \binom{W_1}{k_1} \sum_{k_2} \binom{W_2}{k_2} \cdots \sum_{k_{m-1}} \binom{W_{m-1}}{k_{m-1}} \binom{W_m}{r - \sum_{j=1}^{m-1} k_j}}{\sum_r \binom{S}{\sum_i T_i - r} \sum_{k_1} \binom{W_1}{k_1} \sum_{k_2} \binom{W_2}{k_2} \cdots \sum_{k_{m-1}} \binom{W_{m-1}}{k_{m-1}} \binom{W_m}{r - \sum_{j=1}^{m-1} k_j}},$$

where $(T_i - r)$ is the number of the T_i agents that interact with S agents, and hence the number of S that communicate, r is the number of T_i that communicate with the W_j agents and m is the number of types of W_j agents. This term can also be simplified using Vandermonde's convolution to give us

$$Sa_jAi_{new} = \frac{S \sum_i Ti}{S + \sum_j Wj} .$$

This term is valid when $S + \sum_j Wj \geq \sum_i Ti$. This is always true in Fig. 3.1 since the only Ti are the $Trans$ agents, which are matched in the parallel $I1$ agent by a $T1$ agent and $T1$ is one of the Wj in this model. In a model where this condition is not guaranteed the situation can arise where $\sum_i Ti > S + \sum_j Wj$, which implies that all of the S and Wj will be contacted

so the number of S agents that are contacted is S_{t-1}. This means that the fully general form of the term is

$$Sa_jAi_{new} = \min\left(S, \frac{S\sum_i Ti}{S+\sum_j Wj}\right) . \tag{3.6}$$

Non-prioritised, Single For the case where non-prioritised communication is employed and agents can perform only one instance of the action aj the general term arises in much the same way. The main difference is that the agents performing the output action can choose not to communicate even when there are sufficient agents available to perform the input action to allow them all to do so. This means that when we are considering binomial coefficients that contribute the weight of a particular population change we must consider a binomial coefficient for each of the T_i agents in the model. The general term that comes about is then

$$Sa_jAi_{new} = \frac{\sum_c \binom{\sum_i T_i}{c} \sum_r (c-r) \binom{S}{c-r} \sum_{k_1} \binom{W_1}{k_1} \sum_{k_2} \binom{W_2}{k_2} \cdots \sum_{k_{m-1}} \binom{W_{m-1}}{k_{m-1}} \binom{W_m}{r-\sum_{j=1}^{m-1} k_j}}{\sum_c \binom{\sum_i T_i}{c} \sum_r \binom{S}{c-r} \sum_{k_1} \binom{W_1}{k_1} \sum_{k_2} \binom{W_2}{k_2} \cdots \sum_{k_{m-1}} \binom{W_{m-1}}{k_{m-1}} \binom{W_m}{r-\sum_{j=1}^{m-1} k_j}} ,$$

where c is the number of the T_i agents that communicate for a particular transition to occur and all other terms have the same meaning as in the prioritised case. Once again this term can be simplified by Vandermonde's convolution to give the tractable form

$$Sa_jAi_{new} = \frac{S\sum_i Ti}{S+\sum_i Ti+\sum_j Wj} . \tag{3.7}$$

Similarities By considering (3.6) and (3.7) we can understand why Norman and Shankland [73] found the same MFE for models with prioritised and non-prioritised communication. The only difference between these

models, other than the choice to use priority or not, comes in the form of the agents that perform the output action. In their prioritised model the agents that perform the output action come from a parallel agent featuring an agent that can perform the input action. In the non-prioritised model, the agent that performs the output action cannot explicitly perform the input action. If in a general case we consider these agents separately to other W_j agents performing the input action (say as X_i) we can rewrite (3.6) as

$$Sa_j Ai_{new} = \min\left(S, \frac{S\sum_i Ti}{S + \sum_i X_i + \sum_j Wj}\right) . \tag{3.8}$$

Since the numbers of each X_i are the same as the numbers of the relevant T_i

$$\frac{S\sum_i Ti}{S + \sum_i X_i + \sum_j Wj} \leq S ,$$

and we can rewrite (3.8) as

$$Sa_j Ai_{new} = \frac{S\sum_i Ti}{S + \sum_i T_i + \sum_j Wj} ,$$

which is the same as 3.7, the general equation for non-prioritised communication. In subsequent chapters we will investigate whether these two approaches (prioritised communication with explicit input action versus non-prioritised communication) always lead to the same MFE.

Parallel actions

The existing WSCCS models for which MFE have been derived [73, 82] all feature agents that perform only single instances of the communicating actions. However, WSCCS does allow for agents to perform multiple instances of an action and we want to be able to derive equations for such models.

We place restrictions on how agents must be written for these forms of communication, so that we can reason about the model and derive MFE. As with communication where all agents can perform only a single instance of the action, we require that all actions happen with weight 1. In addition, we allow only the agents performing the output action to perform multiple instances. Agents should perform only a single instance of the input action. Also if an agent can perform n instances it should also be able to perform $n-1$, $n-2$, ... 1 instances and also perform a free action so that it can perform 0 instances. Finally agents performing multiple instances of an output action must evolve to the same state irrespective of how many instances of the action are performed.

Prioritised, Multiple If individuals can perform multiple instances of the action then the general terms become more complex. Agents that can perform multiple instances of an output action with prioritised communication should take the general form

$$A \stackrel{\text{def}}{=} \omega^n.\overline{action}^n : B + \omega^{n-1}.\overline{action}^{n-1} : B + ... + \omega.\overline{action} : B + 1.\surd : B \ ,$$

so that A must perform n instances of $\overline{action}$ where possible, only performing fewer instances where there are insufficient agents to perform the input action. Similarly to the single instance prioritised communication case, either all of the agents performing the output action perform the maximum number of contacts where there are sufficient numbers of agents that can perform the input action with which they can communicate. Otherwise, all of the agents that can perform the input action do so. This makes only a small change to the transmission term for that case, introducing a factor c_i

where c_i is the maximum number of instances of a_j that Ti can perform, leading to

$$Sa_jAi_{new} = \frac{\sum_r \left(\sum_i c_iT_i - r\right)\binom{S}{\sum_i c_iT_i - r}\sum_{k_1}\binom{W_1}{k_1}\sum_{k_2}\binom{W_2}{k_2}\cdots\sum_{k_{m-1}}\binom{W_{m-1}}{k_{m-1}}\binom{W_m}{r-\sum_{j=1}^{m-1}k_j}}{\sum_r \binom{S}{\sum_i c_iT_i - r}\sum_{k_1}\binom{W_1}{k_1}\sum_{k_2}\binom{W_2}{k_2}\cdots\sum_{k_{m-1}}\binom{W_{m-1}}{k_{m-1}}\binom{W_m}{r-\sum_{j=1}^{m-1}k_j}},$$

which can once again be simplified to give us

$$Sa_jAi_{new} = \frac{S\sum_i c_iTi}{S + \sum_j Wj}\ .$$

When there are fewer agents available to perform the input action than agents able to perform the output action, priority means that all of the agents that can perform the input action do so. This means that all of the S agents communicate and taking this into account the general term is

$$Sa_jAi_{new} = \min\left\{S, \frac{S\sum_i c_iTi}{S + \sum_j Wj}\right\}\ . \tag{3.9}$$

Non-prioritised, Multiple When non-prioritised communication is employed, the general term is even more complicated. We make use of the Multinomial coefficient

$$\frac{Ti!}{\prod_{v=1}^{c_i} n_{i,v}!(Ti - \sum_{k=1}^{c_i} n_{i,k})!}\ ,$$

for each of the T_i agents that perform the output action, where $n_{(i,k)}$ is the number of Ti agents performing k instances of a_j at a particular time. The binomial coefficients

$$\binom{S + (\sum_{l=1}^{w} Wl) - 1}{(\sum_{m=1}^{p}\sum_{q=1}^{t_m} q \times n_{m,q}) - 1} \quad \text{and} \quad \binom{S + (\sum_{l=1}^{w} Wl)}{\sum_{m=1}^{p}\sum_{q=1}^{t_m} q \times n_{m,q}}$$

come from the simplification of the product of the individual binomial coefficients for the numbers of S and Wj that communicate. The overall general term for this case is

$$Sa_jAi_{new} = S\frac{f\Big(\Big(\prod_{i=1}^{p}\frac{Ti!}{\prod_{v=1}^{c_i}n_{i,v}!(Ti-\sum_{k=1}^{c_i}n_{i,k})!}\Big)\binom{S+(\sum_{l=1}^{w}Wl)-1}{(\sum_{m=1}^{p}\sum_{q=1}^{t_m}q\times n_{m,q})-1}\Big)}{f\Big(\Big(\prod_{i=1}^{p}\frac{Ti!}{\prod_{v=1}^{c_i}n_{i,v}!(Ti-\sum_{k=1}^{c_i}n_{i,k})!}\Big)\binom{S+\sum_{l=1}^{w}Wl}{\sum_{m=1}^{p}\sum_{q=1}^{t_m}q\times n_{m,q}}\Big)}\ , \tag{3.10}$$

where

$$f(X) = \sum_{n_{p,c_p}=0}^{Tp}\sum_{n_{p,c_{p-1}}=0}^{Tp-n_{p,c_p}}\ldots\sum_{n_{p,1}=0}^{Tp-\sum_{i=1}^{cp}n_{p,i}}\sum_{n_{p-1,c_{p-1}=0}}^{T(p-1)}\ldots\sum_{n_{1,1}=0}^{T1-\sum_{j=1}^{c_1}n_{1,j}}X\ ,$$

p is the number of types of agent that can perform a_j and c_i is the maximum number of instances of a_j that Ti can perform. Due to the agents performing the input action being able to make more than two choices we are left with multinomial coefficients rather than only binomial coefficients. These cannot be simplified in the same way and we are left with (3.10) as the general term for this form of communication. If we do wish to use this form of communication, the specific term will be simpler since the number of types of agent is generally small.

The four cases given in (3.6), (3.7), (3.9) and (3.10) provide the general cases to describe what proportion of the agents that can communicate do so.

3.3.3 Algorithm

Preliminaries

Processes can be *serial* or *parallel*. Given a serial process

$$A \quad w_1.a_1 : A1 + w_2.a_2 : A2 + \ ... \ + w_n.a_n : A_n$$

we make the following definitions

$$\begin{aligned}
derivatives(A) &= \{w_1.a_1 : A_1, w_2.a_2 : A_2, ..., w_n.a_n : A_n\} \\
&\quad also\ denoted\ \{D1, D2, ..., Dn\} \\
sumw(0, n, A) &= w_1 + w_2 + ... + w_n \\
process(D) &= process(w.a : A) = A \\
process(D1, D2, ..., Dn) &= \{A1, A2, ..., An\} \\
action(D) &= action(w.a : A) = a \\
weight(D) &= weight(w.a : A) = w
\end{aligned}$$

Given a parallel processes

$$A \ \ A1\{n_1\} \times A2\{n_2\} \times \ ... \ \times Am\{n_m\}$$

we define

$$components(A) = \{A1.n_1, A2.n_2, ..., Am.n_m\}$$

Finally, an action a is a *communicating* action if there is a restriction set L and $a \notin L$. A process is a communicating agent if it is one that can perform a communicating action.

Pseudo code

The pseudo code of the algorithm is presented in Fig. 3.4. The input that must be given to the algorithm is: the agents of interest (those for which the final MFE must be derived); the number of ticks in the WSCCS model that represent a timestep in the MFE; and the WSCCS description of the model.

3.3.4 Example (Fig. 3.1)

Applying the algorithm to Fig. 3.1 we must note that the agents for which we wish to derive equations are $S1, I1$ and $R1$ and that one timestep in the equations should represent two ticks. This means that the transition table will lead to two sets of equations (one for each tick), which can be algebraically manipulated to obtain the two stage MFE for the system. As for most of our models the state transition table will be sparse since, for example, the agents $S1, I1, R1$ never evolve to $S1, I1, R1$. For this reason we present the non-empty sections of the full table in Tables 3.2 and 3.3. We first consider the evolution of the $S1, T1, Trans$ and $R1$ agents, which evolve to $S2, SI2, I2$ and $R2$, represented in Table 3.2. The $Trans\ *$ row is empty since the $Trans$ agents do not evolve to any of the other agents in the model, instead becoming the null agent 0. Table 3.3 represents the evolution of the agents $S2, SI2, I2, R2$ and $I1$, which evolve to $S1, I1, R1, T1$ and $Trans$.

Following the construction of the state transition table, the next stage in the algorithm is the construction of the $AiajAj_{new}$ terms for the communicating agents, which evolve differently depending on whether they perform the action or not. For this model we only require one such term, $S1infectSI2_{new}$, which represents the number of $S1$ agents that interact

Pseudo Code

1. */*Construct transition table*/*
   For each process Ai {
   if serial(Ai) then {
   if process(derivatives(A_i))={Aj} then
   */*single derivative */*
   add_entry((Ai,),Aj)=Ai_{t-1}*
   else
   */*more than one derivative*/*
   For each derivative $D = (w_j.a_j : Aj)$ {
   if Ai is communicating process then
   if action(D)∈ communicating then
   add_entry((Ai, a_j),Aj)= Aia_jAj_{new}
   else
   add_entry((Ai, a_j), Aj) = $Ai_{t-1} - Aia_jAj_{new}$
   *else{ /*simple probabilistic choice*/*
   $p_j = w_j/sumw(0, nAi)$
   *add_entry((Ai, a_j),Aj)=$p_j * Ai_{t-1}$ } } }
   *else /*process is parallel*/ {*
   For each component Aj{nj}
   add_entry((Ai,),Aj)= n_jAi_t } }

2. */*Construct the change from communication*/*
   For each communicating action aj {
   For each communicating agent{
   construct Aia_jAj_{new} } }

3. */*Construct equations*/*
   For each Ak
   For each action a_j
   For each Ai
   MFE_Ak := MFE_Ak + lookup((Ai,a_j),Ak)
   */*Simplify equations*/*
   For each AgentOfInterest Ai
   For each tick
   replace Aj in MFE_Ai by MFE_Aj

Figure 3.4: Algorithm

		$S2$	$SI2$	$I2$	$R2$
$S1$	$infect$		$S1infectSI2_{new}$		
$S1$	$\surd$	$S1_{t-1} - S1infectSI2_{new}$			
$T1$	$*$			$T1_{t-1}$	
$Trans$	$*$				
$R1$	$*$				$R1_{t-1}$

Table 3.2: State transition table for $S1, T1, Trans$ and $R1$ agents in Fig. 3.1

		$S1$	$I1$	$R1$	$T1$	$Trans$
$S2$	$\surd$	$S2_{t-1}$				
$SI2$	$\surd$	$(1-p_i)SI2_{t-1}$	p_iSI2_{t-1}			
$I2$	$\surd$		$(1-p_r)I2_{t-1}$	p_rI2_{t-1}		
$R2$	$\surd$			$R2_{t-1}$		
$I1$	$*$				$I1_t$	$I1_t$

Table 3.3: State transition table for $S2, SI2, I2, R2$ and $I1$ agents in Fig. 3.1

and become $SI2$. We saw in Section 3.3.2 that this term is

$$S1infectSI2_{new} = \frac{S1_{t-1}Trans_{t-1}}{S1_{t-1} + T1_{t-1} + R1_{t-1}} \ .$$

By summing the columns of Table 3.2 we find the following equations for the evolution of $S1, T1, Trans$ and $R1$:

$$\begin{aligned} S2_t &= S1_{t-1} - S1infectSI2_{new} \ , \\ SI2_t &= S1infectSI2_{new} \ , \\ I2_t &= T1_{t-1} \ , \\ R2_t &= R1_{t-1} \ . \end{aligned} \tag{3.11}$$

Similarly from 3.3 we find equations for the evolution of $S2, SI2, I2, R2$ and

$I1$:

$$\begin{aligned} S1_t &= S2_{t-1} + (1-p_i)SI2_{t-1} \ , \\ I1_t &= p_i SI2_{t-1} + (1-p_r)I2_{t-1} \ , \\ R1_t &= p_r I2_{t-1} + R2_{t-1} \ , \\ T1_t &= I1_t \ , \\ Trans_t &= I1_t \ . \end{aligned} \tag{3.12}$$

From these one stage equations we can substitute for $T1$ and $Trans$ to give

$$S1infectSI2_{new} = \frac{S1_{t-1}I1_{t-1}}{S1_{t-1} + I1_{t-1} + R1_{t-1}} \ ,$$

and

$$I2_t = I1_{t-1} \ .$$

The final step in the algorithm is to create a system of MFE that describes the agents of interest over a full iteration of the model. This is done by writing the equations for the agents of interest at time$= t+n$, where n is the number of stages in the model, and substituting in the right hand side of the equations with expressions in terms of the agents of interest. For our example this involves writing equations for $S1, I1$ and $R1$ at time$= t+2$:

$$\begin{aligned} S1_{t+2} &= S2_{t+1} + (1-p_i)SI2_{t+1} \ , \\ I1_{t+2} &= p_i SI2_{t+1} + (1-p_r)I2_{t+1} \ , \\ R1_{t+2} &= p_r I2_{t+1} + R2_{t+1} \ . \end{aligned}$$

By now substituting for $S2, SI2, I2$ and $R2$ we get

$$\begin{aligned} S1_{t+2} &= S1_t - p_i \frac{S1_t I1_t}{S1_t + I1_t + R1_t} \ , \\ I1_{t+2} &= p_i \frac{S1_t I1_t}{S1_t + I1_t + R1_t} + (1 - p_r) I1_t \ , \\ R1_{t+2} &= p_r I1_t + R1_t \ . \end{aligned}$$

To simplify these equations further we note that they represent an iteration of the model, which, in terms of the MFE, we can think of as a single timestep. Further we can drop the indices from the state names because we are no longer interested in the fact that the $S2, SI2, I2$ and $R2$ agents ever existed, giving us

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i I_t S_t}{N_t} \ , \\ I_{t+1} &= (1 - p_r) I_t + \frac{p_i I_t S_t}{N_t} \ , \\ R_{t+1} &= R_t + p_r I_t \ , \end{aligned} \tag{3.13}$$

where $N_t = S_t + I_t + R_t$ is the total population size. These are the same MFE derived by Norman and Shankland [73] for their models, with p_a (the probability that an infected individual is able to pass on the disease) replaced by p_i (the probability of becoming infected having made contact with an infected individual). Recall that in the model of Norman and Shankland, the choice to infect came first, then the contact stage (and infection was guaranteed following contact). This leads us to ask: what is the effect on the MFE of changing the order of stages in our models? We address this question in Section 3.3.5.

3.3.5 Order matters?

We may naïvely assume that merely changing the order of stages in a model will lead to the same overall mean behaviour, and therefore the same MFE, since all of the same actions happen in each iteration of the model. The model in Fig. 3.5 is the prioritised model of Norman and Shankland [73], with the corresponding flow diagram in Fig. 3.7. We consider the effect of changing order in Fig. 3.5 by deriving MFE for the two stage behaviour of $S2, T2, Trans$ and $R2$. The MFE that are then derived from this model,

$$\begin{aligned} S_{t+1} &= S_t - \frac{S_t Trans_t}{S_t + I_t + R_t} , \\ I_{t+1} &= (1-p_r) I_t + \frac{(1-p_r) S_t Trans_t}{S_t + I_t + R_t} , \\ Trans_{t+1} &= p_a I_t + \frac{p_a S_t Trans_t}{S_t + I_t + R_t} , \\ R_{t+1} &= R_t + p_r I_t + \frac{p_r S_t Trans_t}{S_t + I_t + R_t} , \end{aligned} \tag{3.14}$$

are quite different to the equations that Norman and Shankland found for their model, (3.13), most notably because we now require four equations to describe the system. This difference arises because by changing the order in which the steps occur, we have also changed the underlying biological assumptions of the model. The probabilistic choice for the infected individuals to be able to pass on the infection happens in the second stage of the model, which means the number of $Trans$ agents that can spread the disease in the current iteration of the model is set in the previous iteration. This leads to a separate equation being required to describe the number of $Trans$ agents in the population. In addition the newly infected individuals, represented by

$$\frac{S_t Trans_t}{S_t + I_t + R_t} ,$$

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} p_r.\surd : R2 + p_a.\surd : T2 + (1 - p_r - p_a).\surd : I2 \\
R1 &\stackrel{\text{def}}{=} 1.\surd : R2 \\
\\
S2 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I1 + 1.\surd : S1 \\
T2 &\stackrel{\text{def}}{=} I2 | Trans \\
I2 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I1 + 1.\surd : I1 \\
Trans &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : T + 1.\surd : T \\
R2 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R1 + 1.\surd : R1 \\
\\
Population &\stackrel{\text{def}}{=} S2\{s\} \times I2\{i\} \times R2\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 3.5: Model of infectious disease spread: Fig. 5 from Norman and Shankland [73]

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} T1 \times Trans \\
T1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
Trans &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : 0 + 1.\surd : 0 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
SI2 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
I2 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1 - p_r).\surd : I1 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
\\
Population &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 3.6: Repetition of disease model from Fig. 3.1, included here to allow comparison with Fig. 3.5

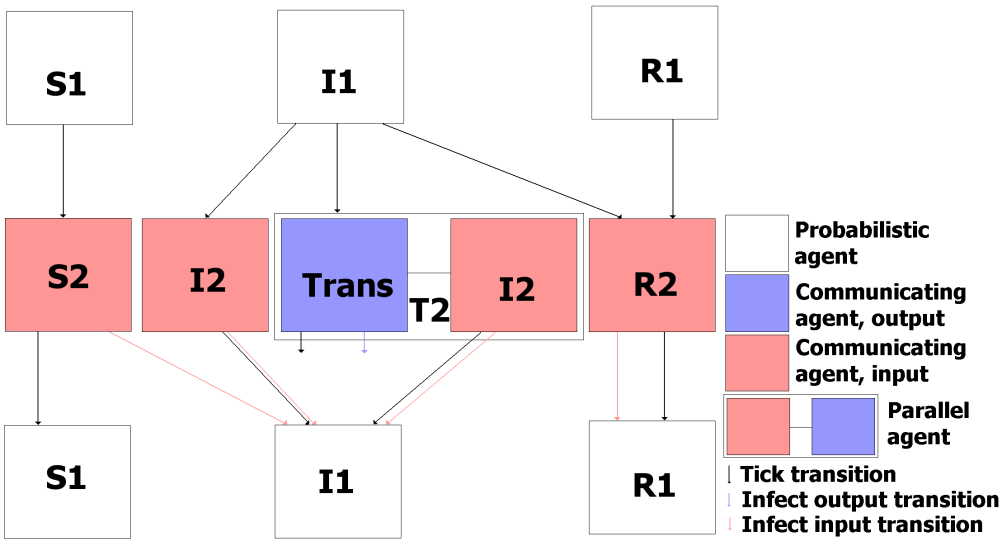

Figure 3.7: Flow diagram for Fig. 3.5

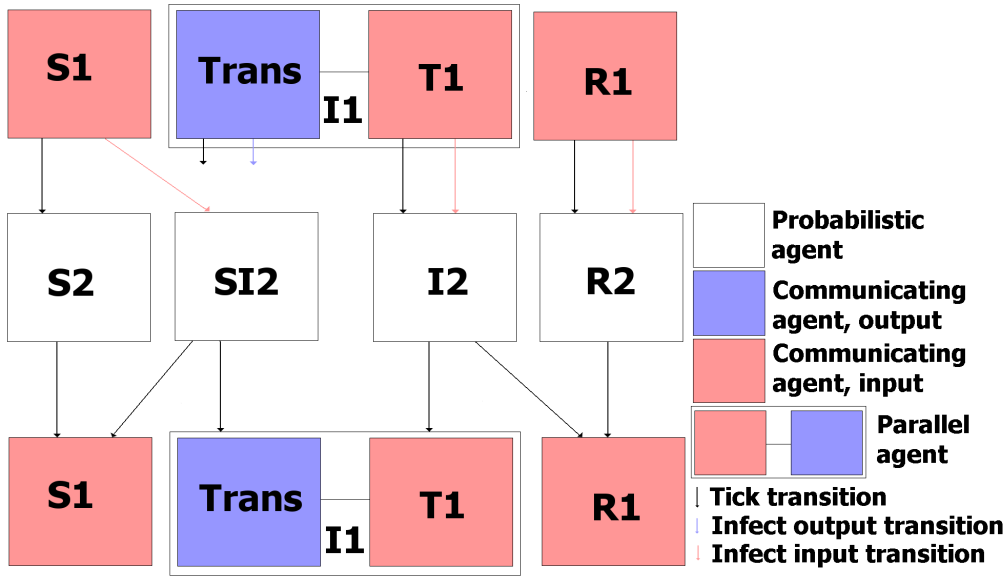

Figure 3.8: Flow diagram for Fig. 3.6

can recover immediately, leading to an extra factor of $(1 - p_r)$ in the transmission term as well as an extra term,

$$\frac{p_r S_t Trans_t}{S_t + I_t + R_t},$$

in the equation for R_{t+1}. The model in Fig. 3.1 (repeated here in Fig. 3.6, along with the flow diagram in Fig. 3.8, to allow comparison with Fig. 3.5) demonstrates that it is possible, at least in this case, to switch the order of communication and choice while maintaining the same overall mean behaviour. However the equations that arise from the model in Fig. 3.5 illustrate the importance of thinking carefully about the biological implications of any changes made to the model.

3.4 Correctness

In this section we consider the correctness of this approach, first by investigating how well the MFE fit to mean behaviour of the example model from Fig. 3.1 and then by relating our approach to the conditions of the limit theorems presented by Kurtz [58].

3.4.1 Accuracy of MFE

We have seen that the MFE are derived by considering the mean of all the possible ways in which the system can evolve. We now consider how well the MFE approximate the average behaviour of the system. This is done by comparing the time series of the MFE (choosing parameter values and an initial population) with the average of a large number of simulations. The simulations were performed using the computational software package *Mathematica* [49]. For each stage of the model the simulation iterates through

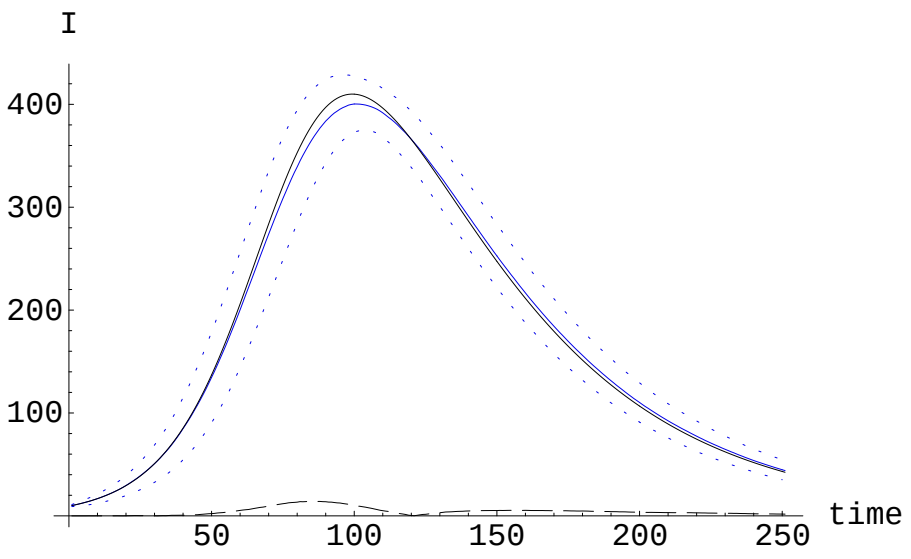

Figure 3.9: Infecteds (I) of Fig. 3.1 for $p_i = 0.08, p_r = 0.02$ and initial population $S1\{990\} \times I1\{10\} \times R1\{0\}$: — MFE, Simulations — mean, ... mean±SD, --- |mean−MFE|

each individual present and uses random numbers, along with the probabilities, to determine how each agent will evolve. For the communication stage we think of the agents performing the output action ($Trans$) as being 'active' and the agents that perform the input action ($S1, T1, R1$) as being 'passive'. This means that the numbers of $S1, T1$ and $R1$ that communicate is determined by the probabilistic choices of the $Trans$ agents. In Fig. 3.9 the infected MFE is plotted along with the mean of 1000 simulations and the mean ± one standard deviation. This graph was produced with $p_i = 0.08$, $p_r = 0.02$ and an initial population of $S1\{990\} \times I1\{10\} \times R1\{0\}$. We can see that the MFE is close to the mean of the simulations for the duration of the epidemic and lies within the standard deviation. Fig. 3.9 also features the absolute value of the difference between the MFE and the mean of the simulations. This shows that the difference is small relative to the mean, reaching a peak slightly before the peak of infection.

To investigate the effect of varying the initial numbers of infecteds we consider systems with the same total population size and parameter values but with different initial numbers of infected individuals. In Fig. 3.10 we consider an initial population featuring only one infected individual. In this case we can see that the average of the simulations fits less well to the MFE. This occurs because, with only one infected individual initially, the probability of the disease dying out before it becomes established is much greater than for the previous example. This means that many of the simulations will be disease free by the time of the peak and therefore the distribution of the number of infecteds in the individual simulations is skewed. For this reason we use the median and quartiles to denote the average and spread, rather that the mean and standard deviation. In this case the MFE fit the average of the simulations less well because of the significant proportion of simulations that are disease free.

In Fig. 3.11 we consider an initial population featuring 20 infected individuals. Here we can see that the MFE and the mean of the simulations are indistinguishable for the majority of the epidemic and the MFE offer an excellent approximation to the mean behaviour of the system. Although the graphs in Figs. 3.9, 3.10 and 3.11 are produced for a single set of parameter values, by investigating a wide range of parameters we find similar results, which show that only for very small initial numbers of infected individuals do the MFE not offer a good approximation to the mean behaviour of the system.

We have demonstrated the accuracy of the MFE by choosing parameter values and computing the time series of the MFE and simulations: however, one of the advantages of MFE is that we can perform some analysis without having to set values for the parameters. For example, we can calculate

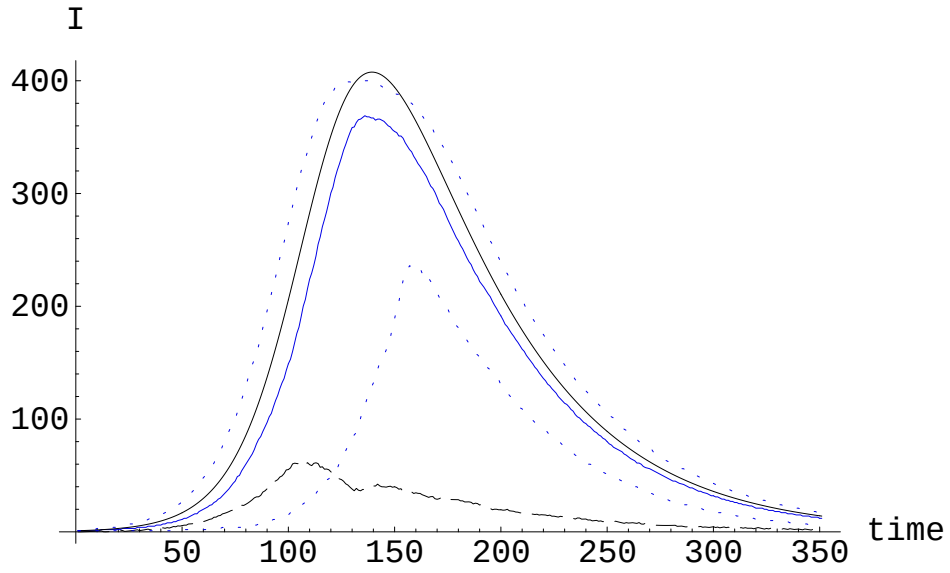

Figure 3.10: Infecteds (I) of Fig. 3.1 for $p_i = 0.08, p_r = 0.02$ and initial population $S1\{999\} \times I1\{1\} \times R1\{0\}$: — MFE, Simulations — median, ... upper and lower quartiles, --- |median−MFE|

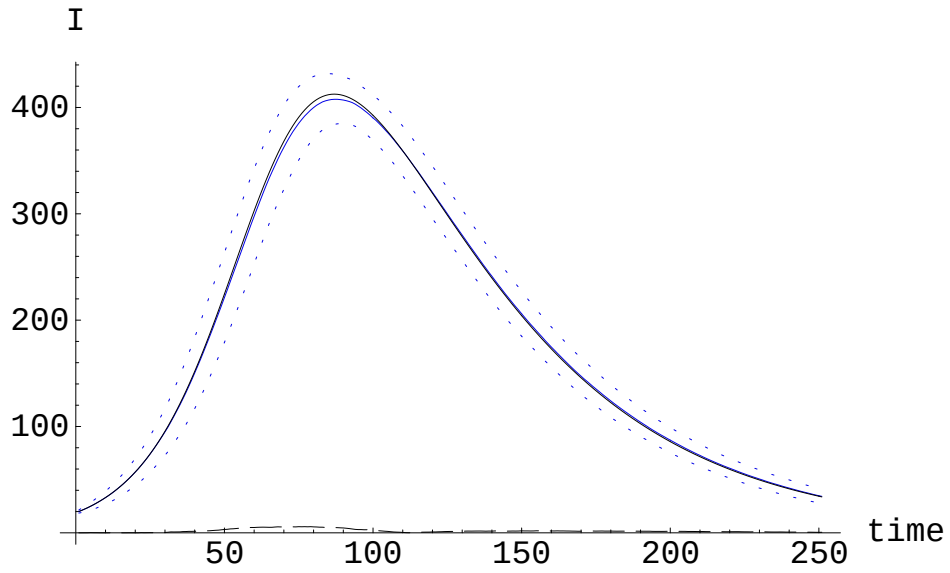

Figure 3.11: Infecteds (I) of Fig. 3.1 for $p_i = 0.08, p_r = 0.02$ and initial population $S1\{980\} \times I1\{20\} \times R1\{0\}$: — MFE Simulations — mean, ... mean±SD, --- |mean−MFE|

expressions for the steady states of the system, in terms of the parameters of the model. As an example we consider (3.13), the MFE for Fig. 3.1. We find the steady states by setting $S_{t+1} = S_t = S^*$, $I_{t+1} = I_t = I^*$ and $R_{t+1} = R_t = R^*$ and solving for S^*, I^* and R^* . Doing this we find that the steady state of (Fig. 3.13) is $(S^*, 0, R^*)$, which is a steady state for any values of S^* and R^* , including the special cases where $S^* = 0$ and $R^* = 0$.

It is further possible to analyse the stability of the steady states for small perturbations. For (3.13) we can rationalise about this without having to perform the full analysis. For small perturbations in S^* or R^* a new steady state will be reached, since any state where $I = 0$ is a steady state. Perturbations in I will cause the system to evolve to a new steady state with different values of S^* and R^* . The steady state $(S^*, 0, R^*)$ can therefore be thought of as stable since for any perturbation the system will evolve back to $(S^*, 0, R^*)$ although with the values of S^* and R^* changed. Alternatively any particular steady state (with specific values for S^* and R^*) is unstable since small perturbations will cause the system to evolve to a new state.

For (3.13) the only steady state is the disease free state $(S^*, 0, R^*)$. Steady states with $I^* \neq 0$ can exist in two different situations: either by recovered individuals losing immunity or by the population being of variable size, either due to birth and death or migration. The former situation could be added to the model in Fig. 3.1 by making the $R2$ agent

$$R2 \stackrel{\text{def}}{=} p_s.\surd : S1 + (1 - p_s).\surd : R1$$

where p_s is the probability that a recovered loses immunity. The latter situation is covered in Chapter 4, which introduces density dependent growth to disease models. We have not explicitly included migration in any of our

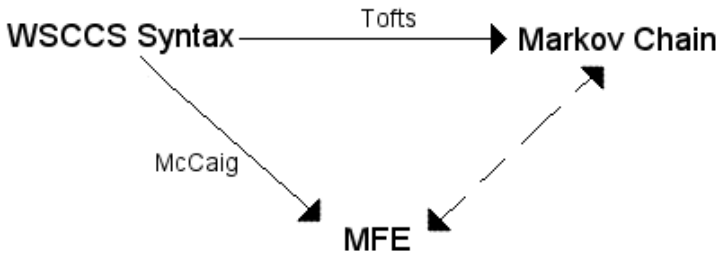

Figure 3.12: Relationship between MFE and Markov chain semantics

models; however, traditional mathematical models often feature single terms to describe the increase (due to births and migration into the population) and decrease (due to deaths and migration out of the population) in numbers of individuals that are like the terms we find by considering birth and death. In subsequent chapters we produce models that lead to MFE with non-trivial steady states, with the numbers of individuals in each group a function of the parameters of the model. Analysing the stability of such steady states would allow us to comment, for instance, on whether a disease can be expected to persist or die out over time. In this thesis we do not find and analyse these steady states, such analysis will be performed in future work.

3.4.2 Proof of correctness

Our algorithm offers an alternative semantics for WSCCS, which allows us to derive MFE directly from the WSCCS syntax (see Fig. 3.12). The standard WSCCS semantics give us the Markov chain for the system. In this section we are interested in rigorously relating the Markov chain and MFE semantics to show that at the limit, where the system consists of infinitely many agents, the mean of the Markov chain is equivalent to the MFE, the dashed line in Fig. 3.12.

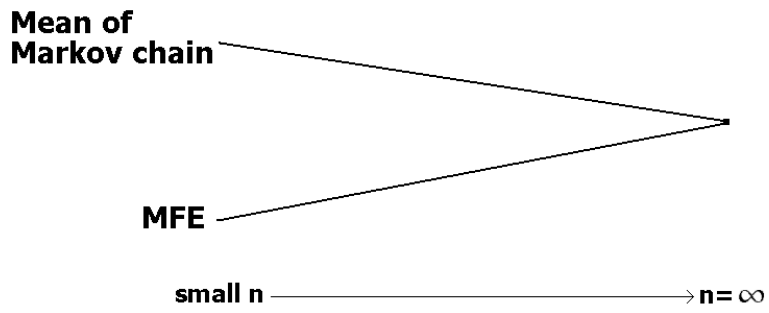

Figure 3.13: Convergence of MFE and mean of Markov chain as number of agents, $n \longrightarrow \infty$

Kurtz [58] presented limit theorems that relate the mean of a Markov chain to ordinary differential equations. For discrete time Markov chains, such as those that arise from WSCCS semantics, an intermediate stage derives equations for the change in the state of the system in a single step of time. By relating the conditions for the derivation of such terms to the process undertaken in our algorithm we demonstrate that in the limit, where a system consists of infinitely many agents, our mean field equations will be infinitesimally close to the mean of the Markov chain (Fig. 3.13).

The conditions that Kurtz set out for the limit theorem are:

- $X_n(k)$ is a sequence of discrete time Markov processes, with measurable state spaces, E_n , which is a subset of $\mathscr{B}^k$, the Borel sets [57] in $\mathbb{R}^k$

- processes are rescaled from $\{0, 1, ..., n\}$ to $[0, 1]$ by dividing through by n and letting $n \to \infty$ [22] – for our purposes n is the initial number of agents in the system

- the one step transition function is denoted by

$$\mu_n(x, \Gamma) = P\{X_n(k+1) \in \Gamma | X_n(k) = x\}$$

i.e. $\mu_n(x, \Gamma)$ is the probability of moving from x to a point in the set Γ in one timestep

- suppose there exist sequences of positive numbers α_n and ε_n such that

$$\lim_{n\to\infty} \alpha_n = \infty \qquad \text{and} \qquad \lim_{n\to\infty} \varepsilon_n = 0 \,,$$

$$\sup_n \sup_{x\in E_n} \alpha_n \int_{E_n} |z - x| \mu_n(x, dz) < \infty \tag{3.15}$$

and

$$\lim_{n\to\infty} \sup_{x\in E_n} \alpha_n \int_{|z-x|>\varepsilon_n} |z - x| \mu_n(x, dz) = 0 \,. \tag{3.16}$$

We now relate these conditions to our WSCCS models.

- The condition on the state space is true since all subsets of $\mathbb{R}^k$ are Borel sets and the states of WSCCS models are in $\mathbb{N}^k \subset \mathbb{R}^k$ where k is the number of types of agents in the model.

- The transition function $\mu_n(x, \Gamma)$, which denotes the probability of moving from state x to some state in Γ, is the same as the transition function of WSCCS.

- We think of z and x as being position vectors with a component representing each type of agent in the system.

- The term $|z - x|$, which appears in both (3.15) and (3.16), is the magnitude of the difference between the start state, x, and the destination state, z. This means that $|z - x|$ is the norm of the vector travelled in one timestep.

- As $n \to \infty$ the number of states that can be reached in one step becomes very large. Since we scale the process by dividing by n the

states z for which $\mu(x,z)$ is greatest will be close to x (such that $|z-x|$ is close to 0). For z where $|z-x|$ is larger, the probability of reaching z will be close to 0. This means that $\int_{|z-x|>\varepsilon_n} |z-x|\mu_n(x,dz)$ is infinitesimal and at the limit (where $n=\infty$) $\alpha_n = \infty$ $\alpha_n \int_{E_n} |z-x|\mu_n(x,dz) < \infty$ is true and (3.15) is satisfied.

- Similarly for (3.16), as $n \to \infty$ the proportion of $[0,1]$ that we are considering increases - since $\varepsilon_n \to 0$. At the limit the probability of reaching any point other than x (such that $|z-x| \neq 0$) is 0 so that (3.16) is satisfied.

Kurtz result then shows that for every $\delta > 0,\ t > 0$

$$\lim_{n\to\infty} \sup_{x\in E_n} P\Big\{ \sup_{k\leq \alpha_n t} |X_n(k) - X_n(0) - \sum_{l=0}^{k} \frac{1}{\alpha_n} F_n(X_n(l))| > \delta \\ \text{where } X_n(0) = x \Big\} = 0\ , \qquad (3.17)$$

where $F_n(x) = \alpha_n \int_{E_n} (z-x)\mu_n(x,dz)$.

Since the equations that are derived by our algorithm are one stage equations we note that applied to the behaviour of the process over only one timestep (3.17) becomes

$$\lim_{n\to\infty} \sup_{x\in E_n} P\Big\{ \sup_{k\leq \alpha_n t} |X_n(1) - X_n(0) - \int_{E_n} (z-x)\mu_n(x,dz)| > \delta \\ \text{where } X_n(0) = x \Big\} = 0\ . \qquad (3.18)$$

If we introduce a function $G(x) = \int_{E_n} (z-x)\mu_n(x,dz)$, (3.18) means that at the limit $n \to \infty$, the difference

$$X_n(1) - \{X_n(0) + G(X_n(0))\}\ ,$$

is infinitesimal; therefore, we assume that

$$X_n(1) = X_n(0) + G(X_n(0)) \ .$$

Since we are dealing with Markov processes, which have no memory of previous states, we can generalise further to find

$$X_n(k+1) = X_n(k) + G(X_n(k)) \ . \tag{3.19}$$

The form of $G(x) = \int_{E_n}(z-x)\mu_n(x, dz)$ is equivalent to the way in which we construct our MFE. We interpret the integral here as a summation. The integral, across the entire state space, of the product of the change of state and the probability of making that change, gives us the mean change of state. By adding this to the previous state of the models, (3.19), we obtain the MFE derived by our algorithm.

3.5 Summary

In this chapter we have presented an algorithm for deriving MFE from the WSCCS description of a model. These equations are amenable to a wide range of mathematical analyses that are used for traditional mathematical models of a system. A particular advantage of these MFE is that we can easily consider very large systems as well as studying the system using symbolic representation of the numbers of agents. For traditional process algebra analyses - Monte Carlo simulation or studying the Markov chain - studying very large models is computationally expensive.

Although previous studies derived MFE for WSCCS models [73, 82], they did so by intuitive reasoning. We now have a rigorous approach to the

derivation of MFE, which allows us to comment on the correctness of the MFE. In addition our algorithm makes it possible to derive equations for much bigger and more complex systems, provided they meet the conditions necessary for using the algorithm, where the mean behaviour may not be so obvious.

Throughout the following chapters we will make use of the algorithm, and will present the MFE for the models being considered, without comment on how this is done.

Chapter 4

Density Dependent Growth

In this chapter we present WSCCS models of population dynamics and compare the MFE that can be derived to the population level equations traditionally used to describe population dynamics. To do this we must capture births and deaths in our model and this is done by the same method used by Sumpter [82]. Birth is modelled by having individuals become a parallel agent, which includes one agent to represent the parent and one agent to represent the offspring. Death is captured by allowing individuals to become the null agent 0.

To capture realistic population growth, which has some upper limit determined by the environment, we require density dependent behaviour, with the likelihood of either giving birth or dying, dependent on the current size of the population. In the models presented here we add density dependence to both the birth and death rates in turn. Biologically there are many systems where the death rate will increase as the density increases. In this case food and shelter become scarce and individuals become weaker and are more likely to die. Alternatively this weakness may manifest itself as a reduced fecundity and a reduction in the birth rates. Mathematically, population

level models of disease spread put the density dependence into either the birth term or the death term and in some models in both the birth and death terms [33, 36, 42, 66]. If we are considering the dynamics of the entire population then it does not matter where the density dependence comes; however, when we add disease it does matter because we split the population into sub-groups. In the literature density dependence often comes in the birth terms in order to make the analysis easier.

Brännström and Sumpter [19] made use of a site-based framework to develop derivations of several different single species population models. They were able to derive several well known models (including the models proposed by Beverton and Holt [16], Hassell [40] and Ricker [78]) but notably not Verhulst's logistic equation [91], which is the most commonly used equation to describe population dynamics [6, 33, 95]. In this chapter we develop WSCCS models of population dynamics that introduce density dependence in different ways and compare the resulting MFE to models from the literature.

Previous WSCCS models of disease spread [73, 82] ignored birth and death of individuals. This is reasonable if we are considering a disease that has a short lived epidemic in comparison to the time scale of population growth or in a managed population like a farm. However, for many natural populations we need to consider births and deaths. In WSCCS density dependence can be introduced implicitly or by explicitly including agents that represent resources for which the population competes. In Sections 4.1-4.3 we present a number of models exploring different individual behaviour, and ways of representing that behaviour, exploring the resultant changes in overall population dynamics.

$$
\begin{aligned}
Rep &\stackrel{\text{def}}{=} N \times N \\
N &\stackrel{\text{def}}{=} p_d.\surd : 0 + p_b.\surd : Rep + (1 - p_d - p_b).\surd : N \\
\\
Population &\stackrel{\text{def}}{=} N\{n\}\lceil\{\surd\}
\end{aligned}
$$

Figure 4.1: Density dependence without food

4.1 Density dependence without resources

In the simplistic model given in Fig. 4.1 the N agents die with probability p_d, becoming the null agent 0, or give birth with probability p_b, becoming the agent *Rep*, which consists of two N agents in parallel. This model leads to a single MFE,

$$N_{t+1} = N_t(1 + p_b - p_d). \tag{4.1}$$

With fixed probabilities p_b and p_d the average behaviour of this model would be similar to that of the simple exponential growth model described by Malthus [62], $N_{t+1} = \lambda N_t$, with $0 \leq \lambda \leq 2$. With $p_b > p_d$ the population will become infinitely large; $p_b < p_d$ will lead to the population dying out, while $p_b = p_d$ will lead to an equilibrium state for any initial population size, $N_0 = n$. This model does not capture the reality of population growth but density dependent growth can be achieved by making use of the functional probabilities described in Chapter 2.

4.1.1 Density dependent birth

Density dependent birth can be added to the model in Fig. 4.1 by making the probability of giving birth functional with p_b inversely proportional to $\lfloor N \rfloor$. This is achieved by adding

$$p\_b \stackrel{\text{prob}}{=} min(max(0, p\_b_0 - k * \lfloor N \rfloor), pL) \ ,$$

where p_{b_0} is the probability of birth in the absence of crowding and k is a measure of the strength of the effect of crowding, $0 < k << 1$.

In the MFE we can now substitute

$$p_b = p_{b_0} - kN_t \ ,$$

and (4.1) then becomes

$$\begin{aligned} N_{t+1} &= N_t + (p_{b_0} - kN_t - p_d)N_t \\ &= N_t + (p_{b_0} - p_d)N_t\left(1 - \frac{kN_t}{p_{b_0} - p_d}\right) \\ &= N_t + rN_t\left(1 - \frac{N_t}{K}\right) \ , \end{aligned} \tag{4.2}$$

which is the discrete time version of Verhulst's logistic equation [91] with $r = (p_{b_0} - p_d)$ and carrying capacity $K = (p_{b_0} - p_d)/k$. The logistic equation is the most commonly used equation for describing population dynamics and is frequently included as a self limiting growth term in models of disease spread. We can see in this case that r is limited by $0 \leq r \leq 1$.

Towards chaos

The dynamics of the logistic equation have been well studied with a wide range of dynamics available for different parameter values [65]. For values of $r > 1$ oscillations begin to appear in the time series of the model with increased period oscillations and eventually chaos developing as r increases. In the equations derived from the WSCCS models such behaviour is never possible because the form of r $(p_{b_0} - p_d)$ involves subtracting a probability from another probability, which means that r can never be greater than 1. Chaos and cycles are widely observed in natural biological systems so it is desirable to be able to obtain such behaviours from our model.

If we consider instead the model in Fig. 4.2 we will see that such behaviour is possible for this model. The only difference in this model is that individuals can give birth to multiple offspring simultaneously, with $b \geq 1$ being the symbolic representation of the number of offspring. This model could be used to model a species that has an average litter size of b. This model leads to the MFE

$$\begin{aligned} N_{t+1} &= N_t + (bp_{b_0} - bkN_t - p_d)N_t \\ &= N_t + rN_t\left(1 - \frac{N_t}{K}\right), \end{aligned} \tag{4.3}$$

which is once again the logistic equation, this time with $r = bp_{b_0} - p_d$ and

$$K = \frac{bp_{b_0} - p_d}{bk}.$$

This means that oscillatory and chaotic behaviour is possible from this model since for $b \geq 2$ it is possible that $r > 1$.

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & min(max(0, p_{b0} - k * \lfloor N \rfloor), p_L) \\
 & & \\
Rep & \stackrel{\text{def}}{=} & N\{b+1\} \\
N & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_b.\surd : Rep + (1 - p_d - p_b).\surd : N \\
 & & \\
Population & \stackrel{\text{def}}{=} & N\{n\}\lceil\{\surd\}
\end{array}
$$

Figure 4.2: Density dependence without food - birth rate = b

4.1.2 Density dependent death

Density dependent death can similarly be added to Fig. 4.2 by replacing functional p_b with functional p_d directly proportional to N:

$$p_d \stackrel{\text{prob}}{=} min(max(0, p_{d_0} + k * \lfloor N \rfloor), p_L) \ ,$$

where p_d0 is the probability of death in the absence of crowding. This is added to the MFE by substituting for

$$p_d = p_{d_0} + kN_t \ ,$$

which gives us the MFE

$$
\begin{aligned}
N_{t+1} &= N_t + (bp_b - (p_{d_0} + kN_t))N_t \\
&= N_t + (bp_b - p_{d_0})N_t\Big(1 - \frac{kN_t}{bp_b - p_{d_0}}\Big) \\
&= N_t + rN_t\Big(1 - \frac{N_t}{K}\Big) \ . \qquad (4.4)
\end{aligned}
$$

This is once again the logistic equation, with $r = (bp_b - p_{d_0})$ and

$$K = \frac{bp_b - p_{d_0}}{k} \ .$$

4.2 Food as an explicit resource

The models in Section 4.1 assumed that we understand how the size of the population affects the growth of the population. The advantage of individual-based modelling techniques is that we can avoid such population level assumptions, with the population level behaviours arising from the individual interactions. To achieve this in models of population dynamics we use agents to represent a resource for which the individuals compete. Access to the resource can be used to determine the likelihood of either birth or death.

Sumpter [82] developed a mechanism for describing density dependent growth in a population, which made use of food as an agent. Individuals in the population compete for the available food resource, giving birth after eating, and die probabilistically. Eating is a prioritised activity, so if an individual can eat they must. This means that every member of the population will give birth at each step of time until the size of the population is larger than the number of food agents, after which the number of births will be equal to the number of food agents. By intuitive reasoning Sumpter derived the following MFE for his model:

$$N_{t+1} = (1 - p_d)N_t + \min[(1 - p_d)N_t, f] \,,$$

where p_d is the probability of death in any timestep and f is the number of food agents. Sumpter found that this MFE has a stable steady state of

$$N^* = \frac{f}{p_d} \,,$$

when $p_d \leq 0.5$.

In the models that follow we make use of this idea of food represented by agents. It should be noted that any other finite resource that a population requires and competes for (e.g. space) can be modelled in exactly the same way.

4.2.1 Density dependent birth

Using prioritised communication between the food agents and the agents representing members of the population forces all individuals to eat; however, in a population it is likely that some individuals, while foraging, may fail to find food that is present. By using non-prioritised communication between food agents and the members of the population we allow individuals to fail to eat even when food is present. This approach has the added effect of eliminating the min term from the MFE, making them more amenable to algebraic analysis.

The model given in Fig. 4.3 uses the same principles as Sumpter's model but features a non-prioritised *eat* action. This means that, even with sufficient food for all individuals to eat, agents may not eat. As well as removing priority this model also reverses the order of the communicative and probabilistic stages. Here we have communication (eating) followed by probabilistic choice while Sumpter's model featured choice followed by communication. We change order in this way to remain consistent with the disease models presented in Fig. 3.1 and in subsequent chapters, which have communication (transmission) followed by probabilistic choice.

The agents $N1$ and $N2$ represent the members of the population at the different stages of the model. The $N1$ agents can eat and become the parallel agent *Rep*, which consists of $b+1$ $N2$ agents and represents giving birth to b offspring. If they do not eat the $N1$ agents become a single $N2$ agent. In

$$
\begin{aligned}
N1 &\stackrel{\text{def}}{=} 1.eat : Rep + 1.\surd : N2 \\
F1 &\stackrel{\text{def}}{=} 1.\overline{eat} : F2 + 1.\surd : F2 \\
\\
Rep &\stackrel{\text{def}}{=} N2\{b+1\} \\
N2 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1-p_d).\surd : N1 \\
F2 &\stackrel{\text{def}}{=} 1.\surd : F1 \\
\\
Population &\stackrel{\text{def}}{=} N1\{n\} \times F1\{f\} \lceil \{\surd\}
\end{aligned}
$$

Figure 4.3: Density dependence on births with non-prioritised communication

the second stage of the model the $N2$ agents die or survive probabilistically. The total number of food agents is constant therefore the F agents $(F1, F2)$ should be thought of as units of food that the environment can produce in a time step rather than discrete portions of food that are consumed by the population. The fact that the number of food agents remains constant means that we do not have to derive an MFE for $F1$ and the number of food agents can be represented by the constant f in the MFE for $N1$.

The model in Fig. 4.3 leads to the MFE

$$N_{t+1} = (1-p_d)N_t + \frac{(1-p_d)bfN_t}{f+N_t} \, . \tag{4.5}$$

Here the term $(1-p_d)N_t$ represents the mean proportion of the existing population that survives the probabilistic death stage. The term $bfN_t/(f+N_t)$ represents the mean number of new births with the factor $(1-p_d)$ representing the proportion of new births that survive the probabilistic death stage, since offspring are able to probabilistically die immediately after birth.

We find the steady state of this model by setting $N_{t+1} = N_t = N^*$:

$$N^* = (1 - p_d)N^* + \frac{(1 - p_d)bfN^*}{f + N^*} .$$

Solving for N^* we get

$$N^* = \frac{(b - (b + 1)p_d)f}{p_d} .$$

With $b = 1$ we have $N^* = (1-2p_d)f/p_d$, and to ensure $N^* > 0$ we require $1/2 > p_d$. The steady state N^* is smaller than in Sumpter's model (which had $N^* = f/p_d$) if the same parameter values are used, since $1 - 2p_d < 1$. This is due to $N1$ agents being able to not eat even when there is food available, which leads to fewer births on average and a smaller population at equilibrium.

Timestep length

The fact that each individual gives birth in each iteration of the model if they have access to the resource suggests that the timestep being captured is long in relation to the time between producing offspring (which can be a matter of days for some insects, several weeks for small mammals and more than a year for some large mammals). In addition we should note that in most ecological systems not all individuals can produce offspring (unless it is reasonable to assume that almost all of the population is female). This approach may, therefore, be useful in modelling asexual reproduction or cellular mitosis but in general for ecological models it is not reasonable. The model in Fig. 4.4 addresses these two factors by having individuals give birth with probability p_b if they eat.

$$
\begin{array}{rcl}
Rep & \stackrel{\text{def}}{=} & N1\{b+1\} \\
N1 & \stackrel{\text{def}}{=} & 1.eat : Nb2 + 1.\surd : N2 \\
F1 & \stackrel{\text{def}}{=} & 1.\overline{eat} : F2 + 1.\surd : F2 \\
 & & \\
N2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1-p_d).\surd : N1 \\
Nb2 & \stackrel{\text{def}}{=} & p_b.\surd : Rep + p_d.\surd : 0 + (1-p_d).\surd : N1 \\
F2 & \stackrel{\text{def}}{=} & 1.\surd : F1 \\
 & & \\
Population & \stackrel{\text{def}}{=} & N1\{n\} \times F1\{f\} \lceil \{\surd\}
\end{array}
$$

Figure 4.4: Density dependence on births with probability of reproducing after eating

The MFE derived from this model is

$$N_{t+1} = (1-p_d)N_t + \frac{bp_b f N_t}{f + N_t} \, ,$$

with resulting steady state

$$N^* = \frac{(bp_b - p_d)f}{p_d} \, .$$

Once again we must ensure that $N^* > 0$, which means we require $bp_b > p_d$.

4.2.2 Density dependent death

A similar model can be developed featuring density dependent death. In Fig. 4.5 the $N1$ agents can once again eat, becoming the agent $N2$, but here if they do not eat they die, becoming the null agent 0. The $N2$ agents then give birth probabilistically and to be realistic, $N2$ agents can also die probabilistically. This means that in each step of time a proportion of the population die, for instance, due to age and some die due to a lack of food.

$$
\begin{aligned}
Rep &\stackrel{\text{def}}{=} N1\{b+1\} \\
N1 &\stackrel{\text{def}}{=} 1.eat : N2 + 1.\surd : 0 \\
F1 &\stackrel{\text{def}}{=} 1.\overline{eat} : F2 + 1.\surd : F2 \\
\\
N2 &\stackrel{\text{def}}{=} p_b.\surd : Rep + p_d.\surd : 0 + (1 - p_b - p_d).\surd : N1 \\
F2 &\stackrel{\text{def}}{=} 1.\surd : F1 \\
\\
Population &\stackrel{\text{def}}{=} N1\{n\} \times F1\{f\} \lceil \{\surd\}
\end{aligned}
$$

Figure 4.5: Density dependence on deaths with non-prioritised communication

The MFE for this model is

$$N_{t+1} = (1 + bp_b - p_d)\frac{fN_t}{f + N_t}, \tag{4.6}$$

where $fN_t/(f + N_t)$ represents the proportion of the population that eat and therefore survive the competition for food, with the factor $(1 + bp_b - p_d)$ representing the increase in the population due to births and the decrease due to probabilistic death. Equation (4.6) can be rearranged to give

$$N_{t+1} = \frac{aN_t}{1 + cN_t}, \tag{4.7}$$

where $a = (1 + bp_b - p_d)$ and $c = 1/f$. Equation 4.7 is the Beverton-Holt model [16], which was originally proposed as a model of salmon populations displaying density dependent birth; however we have derived this equation from a model that features density dependent death. Although this model has previously been described it is not commonly used for describing density dependent populations so it is interesting that this term has naturally arisen from our WSCCS model that explicitly includes the population interacting

with a resource.

If we set $N_{t+1} = N_t = N^*$ in (4.6) and solve for N^* we find that the steady state of this model is

$$N^* = (bp_b - p_d)f \ .$$

In this case to ensure the steady state is positive we require $bp_b > p_d$.

Comparison of dynamics of MFE for implicit and explicit competition models

The logistic and Beverton-Holt models offer different ways to capture density dependent growth in a population, which leads us to ask the question: how do the dynamics of the two models compare? In Fig. 4.6 we plot (4.4) and (4.6) with the same values of $p_b = 0.2$, $b = 1$ and with p_{d_0} of (4.4) equal to p_d of (4.6) ($p_d = p_{d_0} = 0.15$). In addition the values of $k = 0.00005$ and $f = 20000$ were chosen to give $N^* = 1000$ for each model. We see that both models offer similar dynamics with the curves diverging only slightly as they approach the steady state. In Fig. 4.7 we produce graphs with $b = 12$ so that $r = (bp_b - p_{d_0}) > 1$ in the logistic model and oscillations are observed. In this case both models share a larger steady state,

$$N^* = \frac{bp_b - p_{d_0}}{k} = (bp_b - p_d)f = 45000 \ ,$$

and reach the steady state more quickly since more births take place during each timestep. The two curves remain close during the rapid growth phase and as they approach the steady state the Beverton-Holt curve smoothly settles to the steady state, while the logistic curve oscillates about the steady state as expected.

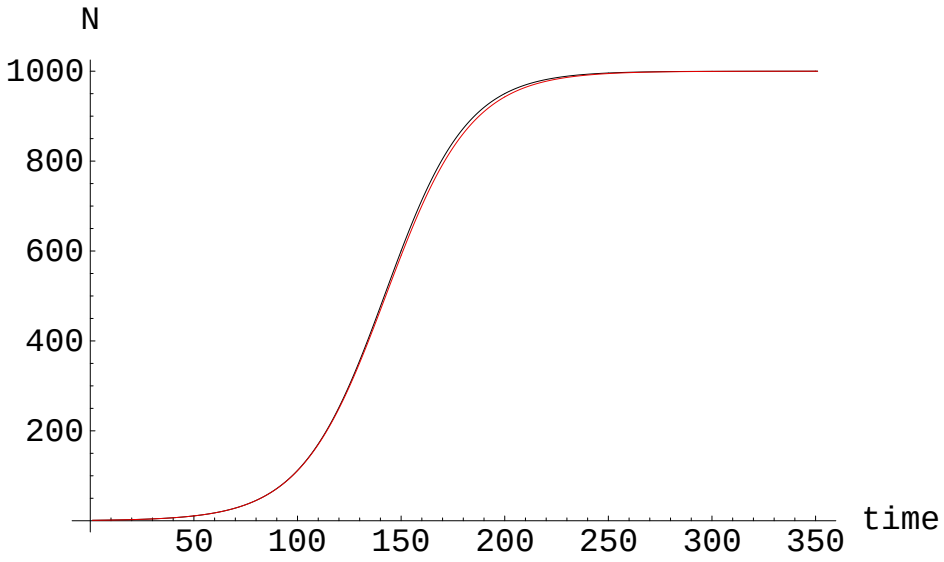

Figure 4.6: Logistic and Beverton-Holt models plotted with the same parameter values - $b = 1$ (— Logistic, — Beverton-Holt)

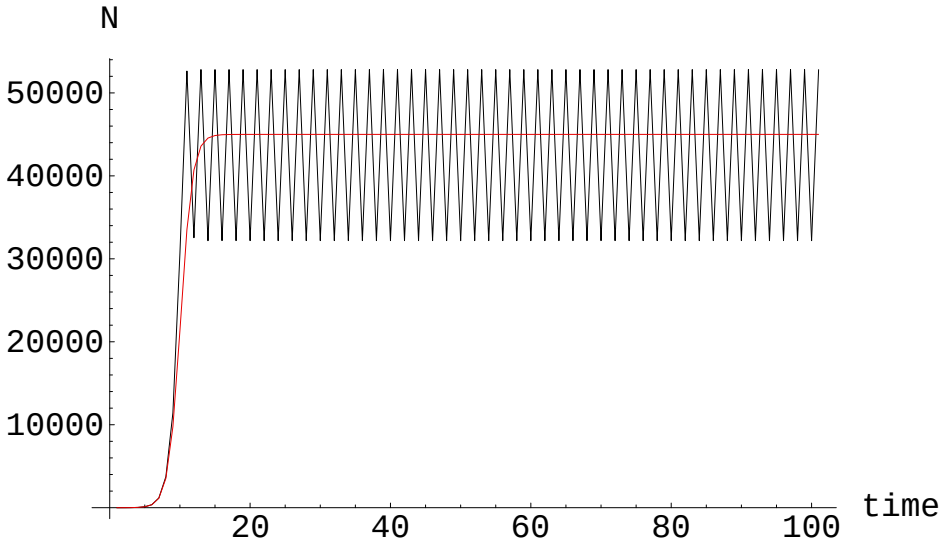

Figure 4.7: Logistic and Beverton-Holt models plotted with the same parameter values - $b = 12$ (— Logistic — Beverton-Holt)

4.3 Population dynamics and disease

While population dynamics are interesting in isolation they are also crucial in developing realistic models of other aspects of a population. In this section we present models of disease spread, based on the basic model proposed in 3.1, with the addition of population dynamics mechanisms from Sections 4.1 and 4.2. In all of these models we assume, for simplicity, that the disease in question does not cause death in the host species. While it is reasonable to say that for many disease systems the level of disease induced mortality is negligible we may also wish to consider situations where the disease has a significant effect on the population dynamics. Such systems could be modelled by having a different probability of death, $p\_dd \geq p\_d$, for the infected agents. In the resulting equations this would merely lead to p_{dd} replacing p_d in the equations for I_{t+1} .

4.3.1 SIR models with explicit competition

Density dependent death

The model in Fig. 4.8 adds the population dynamics of Fig. 4.5 to the model of infectious disease spread introduced in Fig. 3.1. The first stage in the model is the eating stage in which $S1, I1$ and $R1$ all compete for food and those that do not eat will die. The second stage is a contact stage in which infected ($Trans$) agents come into contact with the population and potentially pass the disease to susceptibles. The infected individuals are represented by parallel agents, with the $Trans$ agents passing on the disease and the $T2$ agents able to be contacted by a $Trans$ agent. Communication is prioritised so that all $Trans$ make contact. $S2$ that are contacted become $SI3$, which make a probabilistic choice to become infected or not, while

$$
\begin{aligned}
RepS &\stackrel{\text{def}}{=} S1\{b+1\} \\
S1 &\stackrel{\text{def}}{=} 1.eat : S2 + 1.\surd : t \\
RepI &\stackrel{\text{def}}{=} I1 \times S1\{b\} \\
I1 &\stackrel{\text{def}}{=} 1.eat : I2 + 1.\surd : 0 \\
RepR &\stackrel{\text{def}}{=} R1 \times S1\{b\} \\
R1 &\stackrel{\text{def}}{=} 1.eat : R2 + 1.\surd : 0 \\
Food1 &\stackrel{\text{def}}{=} 1.\overline{eat} : Food2 + 1.\surd : Food2 \\
\\
S2 &\stackrel{\text{def}}{=} \omega.infect : SI3 + 1.\surd : S3 \\
I2 &\stackrel{\text{def}}{=} T2 \times Trans \\
T2 &\stackrel{\text{def}}{=} \omega.infect : I3 + 1.\surd : I3 \\
Trans &\stackrel{\text{def}}{=} \omega.\overline{infect} : 0 + 1.\surd : 0 \\
R2 &\stackrel{\text{def}}{=} \omega.infect : R3 + 1.\surd : R3 \\
Food2 &\stackrel{\text{def}}{=} 1.\surd : Food3 \\
\\
S3 &\stackrel{\text{def}}{=} p_b.\surd : RepS + (1 - p_b - p_d).\surd : S1 + p_d.\surd : 0 \\
SI3 &\stackrel{\text{def}}{=} p_b.\surd : RepS + p_i.\surd : I1 + (1 - p_i - p_b - p_d).\surd : S1 \\
&\quad + p_d.\surd : 0 \\
I3 &\stackrel{\text{def}}{=} p_b.\surd : RepI + p_r.\surd : R1 + (1 - p_r - p_b - p_d).\surd : I1 \\
&\quad + p_d.\surd : 0 \\
R3 &\stackrel{\text{def}}{=} p_b.\surd : RepR + (1 - p_b - p_d).\surd : R1 + p_d.\surd : 0 \\
Food3 &\stackrel{\text{def}}{=} 1.\surd : Food1 \\
\\
Population &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times Food1\{f\} \lceil \{\surd\}
\end{aligned}
$$

Figure 4.8: Frequency dependent SIR model with density dependence in the deaths

$T2$ and $R2$ agents are not affected by contact since infected and recovered individuals cannot become infected again. After the contact stage the $Trans$ agents all become the null 0 so that the infected individuals are once again represented by a single agent. The final stage is the probabilistic stage in which all individuals can give birth to b susceptible individuals, with probability p_b, or die, with probability p_d. In addition the $SI3$ agents become infected with probability p_i and $I3$ agents can recover with probability p_r.

The system of MFE derived from this model is

$$\begin{aligned} S_{t+1} &= \frac{f}{f+N_t}\Big((1-p_d)S_t + bp_bN_t - \frac{p_iS_tI_t}{N_t}\Big)\,, \\ I_{t+1} &= \frac{f}{f+N_t}\Big((1-p_d-p_r)I_t + \frac{p_iS_tI_t}{N_t}\Big)\,, \\ R_{t+1} &= \frac{f}{f+N_t}\Big((1-p_d)R_t + p_rI_t\Big)\,, \end{aligned} \tag{4.8}$$

where $N_t = S_t+I_t+R_t$, the total population size. These are the standard SIR equations that have been found for WSCCS models [73] with an extra factor of $f/(f+N_t)$ on each equation, which is the proportion of the population that successfully eats. This is unconventional since in traditional models [6, 42] the transmission term (in this case $(p_iS_tI_t)/N_t$) is not affected by the density dependent birth or death term while here it is scaled by the death due to competition term $f/(f+N_t)$. This is not affected by the order in which the stages of the WSCCS model occur. For instance if the eating stage were to come last, following the contact and probabilistic stages, then the number of S agents that would become infected after the first two stages would be $p_aS_tI_t/N_t$, and the proportion of those that would survive competition for resources would be $f/f+N_t$. Such a model would therefore lead to the same system of equations (4.8) as for Fig. 4.8.

Once again we can say that the population dynamics in Fig. 4.8 come

directly from the competition for food rather than any assumptions that have been imposed on the model. We can therefore say that the system of equations (4.8) is a candidate for modelling population dynamics in disease systems with a constant resource, despite the differences to traditional models.

Density dependent birth

Now instead of having density dependent death we return to density dependent birth, see Fig. 4.9. The model given in Fig. 4.9 adds density dependence in birth to an SIR model. In this model all of the individuals in the population compete for the food that is available and if they do eat they give birth to b newborn individuals. These newborns do not make themselves available to be infected immediately but become susceptible individuals at the next step of time.

This model leads to the following system of MFE:

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t + \frac{bfN_t}{f+N_t} - \frac{p_i S_t I_t}{N_t} \,, \\ I_{t+1} &= (1-p_d-p_r)I_t + \frac{p_i S_t I_t}{N_t} \,, \\ R_{t+1} &= (1-p_d)R_t + p_r I_t \,. \end{aligned}$$

Alternatively, if we had designed the model with all individuals giving birth to susceptible individuals, which immediately make themselves available to

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} 1.eat : RepS + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} 1.eat : RepI + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} 1.eat : RepR + 1.\surd : R2 \\
Food1 &\stackrel{\text{def}}{=} 1.\overline{eat} : Food2 + 1.\surd : Food2 \\
\\
RepS &\stackrel{\text{def}}{=} S2 \times B2\{b\} \\
S2 &\stackrel{\text{def}}{=} \omega.infect : SI3 + 1.\surd : S3 \\
B2 &\stackrel{\text{def}}{=} 1.\surd : B3 \\
RepI &\stackrel{\text{def}}{=} I2 \times B2\{b\} \\
I2 &\stackrel{\text{def}}{=} T2 \times Trans \\
T2 &\stackrel{\text{def}}{=} \omega.infect : I3 + 1.\surd : I3 \\
Trans &\stackrel{\text{def}}{=} \omega.\overline{infect} : 0 + 1.\surd : 0 \\
RepR &\stackrel{\text{def}}{=} R2 \times B2\{b\} \\
R2 &\stackrel{\text{def}}{=} \omega.infect : R3 + 1.\surd : R3 \\
Food2 &\stackrel{\text{def}}{=} 1.\surd : Food3 \\
\\
S3 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1 - p_d).\surd : S1 \\
SI3 &\stackrel{\text{def}}{=} p_d.\surd : 0 + p_i.\surd : I1 + (1 - p_i - p_d).\surd : S1 \\
B3 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
I3 &\stackrel{\text{def}}{=} p_d.\surd : 0 + p_r.\surd : R1 + (1 - p_r - p_d).\surd : I1 \\
R3 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1 - p_d).\surd : R1 \\
Food3 &\stackrel{\text{def}}{=} 1.\surd : Food1 \\
\\
Population &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times Food1\{f\} \lceil \{\surd\}
\end{aligned}
$$

Figure 4.9: Frequency dependent SIR model with density dependence in the births

become infected and make the probabilistic choice to die, we would get a quite different transmission term in our MFE:

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t + (1-p_d)\frac{bfN_t}{f+N_t} - \frac{p_i I_t}{N_t(1+\frac{bf}{f+N_t})}\Big(S_t + \frac{bfN_t}{f+N_t}\Big) \,, \\ I_{t+1} &= (1-p_d-p_r)I_t + \frac{p_i I_t}{N_t(1+\frac{bf}{f+N_t})}\Big(S_t + \frac{bfN_t}{f+N_t}\Big) \,, \\ R_{t+1} &= (1-p_d)R_t + p_r I_t \,, \end{aligned}$$

featuring the term

$$\frac{p_i I_t}{N_t(1+\frac{bf}{f+N_t})}\Big(S_t + \frac{bfN_t}{f+N_t}\Big) \,.$$

Here the number of susceptible individuals with which the infecteds can make contact includes the births from the current iteration of the model, which introduces the factor

$$S_t + \frac{bfN_t}{f+N_t} \,.$$

Similarly the total number of individuals that can be contacted by the infecteds includes these births and the denominator in the transmission term becomes

$$N_t\left(1 + \frac{bf}{f+N_t}\right) \,.$$

The other change that arises in these equations is that the transmission term includes a factor capturing the fact that individuals can be born and die within a single iteration of the model leading to the term

$$(1-p_d)\frac{bfN_t}{f+N_t} \,.$$

This illustrates the importance of considering carefully the biological implications of choices within the model. In many cases it is reasonable to suggest that newborn individuals will not be available to be infected immediately and by capturing this in our model we derive simpler MFE.

4.3.2 SIR models with implicit competition

In the same way as for the models that make use of food as an agent we can add infectious disease dynamics to the models involving mechanisms for density dependence using functional probabilities outlined in Fig. 4.2.

Density dependent birth

Combining an SIR model with the density dependent birth model from Fig. 4.2 we get the model shown in Fig. 4.10. The first stage in the model features probabilistic birth and death. In the second stage newborn individuals $B2$ are not able to be infected and will become susceptible $S3$ agents at the next stage in the model.

Density dependent births is again achieved by defining $p_b = (p_{b_0} - kN_t)$ and the model leads to the following system of MFE:

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t - \frac{p_i(1-p_d)S_tI_t}{N_t} + bp_{b_0}\left(1 - \frac{kN_t}{p_{b_0}}\right)N_t \ , \\ I_{t+1} &= (1-p_r)(1-p_d)I_t + \frac{p_i(1-p_d)S_tI_t}{N_t} \ , \\ R_{t+1} &= (1-p_d)R_t + p_r(1-p_d)I_t \ . \end{aligned} \tag{4.9}$$

This is the typical sort of model we would write down directly, with each equation featuring a factor for the probability of surviving $(1 - pd)$, and a logistic type growth term in the susceptibles equation.

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & min(max(0, p_b 0 + k * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor)), pL) \\
 & & \\
S1 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + p_d.\surd : 0 + (1 - p_b - p_d).\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + p_d.\surd : 0 + (1 - p_b - p_d).\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + p_d.\surd : 0 + (1 - p_b - p_d).\surd : R2 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S2 \times B2\{b\} \\
S2 & \stackrel{\text{def}}{=} & \omega.infect : SI3 + 1.\surd : S3 \\
RepI & \stackrel{\text{def}}{=} & I2 \times B2\{b\} \\
I2 & \stackrel{\text{def}}{=} & T2 \times Trans \\
T2 & \stackrel{\text{def}}{=} & \omega.infect : I3 + 1.\surd : I3 \\
Trans & \stackrel{\text{def}}{=} & \omega.\overline{infect} : 0 + 1.\surd : 0 \\
RepR & \stackrel{\text{def}}{=} & R2 \times B2\{b\} \\
R2 & \stackrel{\text{def}}{=} & \omega.infect : R3 + 1.\surd : R3 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : S3 \\
 & & \\
S3 & \stackrel{\text{def}}{=} & 1.\surd : S1 \\
SI3 & \stackrel{\text{def}}{=} & p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
I3 & \stackrel{\text{def}}{=} & p_r.\surd : R1 + (1 - p_r).\surd : I1 \\
R3 & \stackrel{\text{def}}{=} & 1.\surd : R1 \\
 & & \\
Population & \stackrel{\text{def}}{=} & S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 4.10: Frequency dependent SIR model with density dependent probability of giving birth

Density dependent death

We can replace density dependent birth in Fig. 4.10 with density dependent death by replacing the functional p_b with functional p_d:

$$p\_d \stackrel{\text{prob}}{=} min(max(0, p_d0 + k * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor)), 1) \ .$$

The MFE that arise from Fig. 4.10 with density dependent death are

$$\begin{aligned} S_{t+1} &= (1 - p_{d_0} - kN_t)S_t - p_i(1 - p_{d_0} - kN_t)\frac{S_t I_t}{N_t} + bp_b N_t \ , \\ I_{t+1} &= (1 - p_{d_0} - kN_t)(1 - p_r)I_t + p_i(1 - p_{d_0} - kN_t)\frac{S_t I_t}{N_t} \ , \\ R_{t+1} &= (1 - p_{d_0} - kN_t)R_t + (1 - p_{d_0} - kN_t)p_r I_t \ . \end{aligned}$$

Mathematically these are more complex than for the density dependent birth case because the transmission term is scaled by the factor $(1 - p_{d_0} - kN_t)$ rather than $(1 - p_d)$. In both cases these terms are introduced because the individuals that die are not able to take part in transmission of the disease. In the current model the term is more complex because the probability of death is density dependent and therefore depends on the current population size. This means that, for systems where density dependent death is preferable, there is a trade off between the mathematical simplicity and increased tractability offered by models featuring density dependent birth, and the improved biological realism offered by considering density dependent death.

4.4 Summary

In this chapter we have introduced several different models that seek to capture realistic growth. The models in Figs. 4.1 and 4.2 make use of functional probabilities to introduce density dependent behaviour into the population.

From both of these models, for density dependent birth or density dependent death, the resulting MFE is a discrete time version of the logistic growth model.

The fact that we have achieved the logistic equation from these models may seem to contradict the findings of Brännström and Sumpter [19] who did not find the logistic equation for any of their models. It should be noted that we are effectively choosing our probabilities using the same assumptions that lead to the logistic equation in the traditional population level models by making the probabilities linearly proportional to the population size; however, Brännström and Sumpter did try to get the logistic equation and couldn't, so even though we have "forced it" it is still significant that we can get it. Other methods of choosing these probabilities, such as non-linear proportionality, would lead to different MFE. A major difference is that Brännström and Sumpter's site-based framework is inherently spatial whereas the models presented here assume random mixing of the population.

Alternative models were presented that made use of agents to represent food as a resource for the population. These mechanistic models more closely represent behaviour in a population with density dependence arising naturally from competition for resources. In the case where density dependent death is modelled we obtain the Beverton-Holt equation [16], which was proposed to describe the dynamics of fish populations. By adding this form of density dependence to a disease model we obtain the classical equations to describe disease spread [52] with each equation scaled by a factor representing survival after competition. This makes the equations more complex and less tractable algebraically. In traditional mathematical models this problem is addressed by making birth rather than death density dependent.

Although the models that explicitly capture competition for resources have a greater degree of biological realism, it also introduces a greater degree of mathematical complexity into the MFE. In subsequent chapters where births and deaths are included we make use of the functional probability method from Fig. 4.1 in order to simplify the equations and allow us to focus on the aspect of the population in which we are most interested.

Chapter 5

Modes of Transmission I: Frequency Dependent vs Density Dependent

Previous WSCCS models of disease spread [73] naturally lead to a frequency dependent transmission term of the form,

$$\frac{\beta S_t I_t}{N_t} ,$$

in the derived MFE whilst in traditional mathematical models the most commonly used term is the density dependent term proposed by Kermack and McKendrick [52]:

$$\beta SI .$$

Although frequency dependent transmission has been proposed as the more appropriate term to use by some authors [42], and other terms have been proposed [20, 48], the density dependent term is still the favoured option in many cases.

Turner et al. [90] developed cellular automata models that implemented frequency dependent and density dependent transmission at the level of the individual. Cellular automata do not have a rigorous method for deriving equations, such as the method described in Chapter 3 for WSCCS, so the population level behaviour was found by fitting equations to the results obtained from the cellular automata. They found, counterintuitively, that irrespective of the individual-based behaviour implemented the frequency dependent transmission term was found to most accurately describe behaviour at the population level. This seems to imply that whatever rules we have at the individual level, we will always get frequency dependent transmission at the population level.

In this chapter we investigate whether this is also true for WSCCS or whether it is possible to develop WSCCS models that will lead to MFE featuring density dependent transmission. This is done by incorporating the assumptions that Begon et al. [10] used to produce derivations for the density dependent term, namely that the number of contacts made by an infected individual is directly proportional to the size of the population. Our approach differs slightly from that of Begon et al. since they derive transmission from a general term $Scpv$ where S is the number of susceptible individuals, c is the rate of contacts made by susceptibles, p is the probability that contact is with an infected individual and v is the probability that a susceptible becomes infected after contact with an infected individual. This assumes that susceptible individuals are 'active' in seeking out contacts and the individuals they contact are 'passive'. In WSCCS it is natural to write models such that infected individuals make contact and the rate of transmission is governed by the probability that contact is with a susceptible, as well as the probability that susceptibles become infected after contact.

Models are presented here that make use of either prioritised or non-prioritised communication to investigate the transmission terms that arise in the MFE. We are particularly interested in whether describing density dependent contact at the individual level leads to the expected density dependent transmission term in the resulting MFE. All models in this chapter make use of the logistic growth mechanism with density dependent birth described in Section 4.1.1. For simplicity the number of births is one in all of these models but we saw in Chapter 4 how we could use a symbolic representation (b) for the number of births.

5.1 Frequency dependent transmission

Norman and Shankland [73] found the frequency dependent transmission term

$$\frac{p_a S_t I_t}{N_t}$$

arose naturally from their models. However it was pointed out that the overall population size was constant, $N_t = N$, since births and deaths were not included in the model, and therefore this term could be written as $\beta S_t I_t$ with

$$\beta = \frac{p_a}{N} \, .$$

In Chapter 4 (Fig. 4.10) we added a mechanism for self limiting growth to a disease model. This gave MFE that once again feature a frequency dependent transmission term, now with $\beta = p_i(1 - p_d)$ but with N_t no longer constant.

Norman and Shankland also found that the choice of prioritised or non-prioritised communication need not affect the MFE derived from the system. This is true again here and the model in Fig. 5.1 features non-prioritised

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & min(max(0, p_b 0 + k * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor)), pL) \\
 & & \\
S1 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + p_d.\surd : 0 + (1 - p_b - p_d).\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + p_d.\surd : 0 + (1 - p_b - p_d).\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + p_d.\surd : 0 + (1 - p_b - p_d).\surd : R2 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S2 \times B2 \\
S2 & \stackrel{\text{def}}{=} & 1.infect : SI3 + 1.\surd : S3 \\
RepI & \stackrel{\text{def}}{=} & I2 \times B2 \\
I2 & \stackrel{\text{def}}{=} & T21.\overline{infect} : I3 + 1.\surd : I3 \\
RepR & \stackrel{\text{def}}{=} & R2 \times B2 \\
R2 & \stackrel{\text{def}}{=} & 1.infect : R3 + 1.\surd : R3 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : S3 \\
 & & \\
S3 & \stackrel{\text{def}}{=} & 1.\surd : S1 \\
SI3 & \stackrel{\text{def}}{=} & p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
I3 & \stackrel{\text{def}}{=} & p_r.\surd : R1 + (1 - p_r).\surd : I1 \\
R3 & \stackrel{\text{def}}{=} & 1.\surd : R1 \\
 & & \\
Population & \stackrel{\text{def}}{=} & S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 5.1: SIR model with frequency dependent transmission and non-prioritised communication

communication yet leads to the same MFE (4.9) as Fig. 4.10, which made use of prioritised communication. In Sec. 5.2 we will see that this is not always true as models are presented differing only by the choice to use prioritised communication or not, and can lead to very different MFE.

5.2 Density dependent transmission

To model density dependent transmission at the individual level the infected individuals must be able to make multiple contacts within each iteration of the model. This can be done in several ways to achieve a contact rate of c: i) infected individuals are modelled by a parallel agent, which contains c agents that can transmit the disease ii) infected individuals can perform up to c instances of the *infect* action in parallel iii) and infected individuals can make contact in c consecutive stages of the model. Here we consider each of these approaches in turn. We may expect the three approaches to lead to the same population level behaviour but we will see that this is not always true.

5.2.1 Non-prioritised communication

Parallel agents

The model featured in Fig. 5.2 features a parallel infected agent, $I1$, consisting of c transmitting agents, $Trans$, which all become the null agent 0, and a single $Ih1$ agent that does not take part in a communicating action but becomes the agent $I2$ at the next stage. The maximum number of contacts that an infected agent can make, c, is a proportion of the total population

$$
\begin{array}{rcl}
c & \stackrel{\text{int}}{=} & min(Round(j * (\lfloor S0 \rfloor + \lfloor I0 \rfloor + \lfloor R0 \rfloor)), Cmax) \\
p_b & \stackrel{\text{prob}}{=} & p_b 0 - k(\lfloor S0 \rfloor + \lfloor I0 \rfloor + \lfloor R0 \rfloor) \\
 & & \\
S0 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + (1 - p_b).\surd : S1 \\
I0 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + (1 - p_b).\surd : I1 \\
R0 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + (1 - p_b).\surd : R1 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S1 \times B1 \\
RepI & \stackrel{\text{def}}{=} & I1 \times B1 \\
RepR & \stackrel{\text{def}}{=} & R1 \times B1 \\
S1 & \stackrel{\text{def}}{=} & 1.infect : SI2 + 1.\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & Ih1 \times Trans\{c\} \\
Trans & \stackrel{\text{def}}{=} & 1.\overline{infect} : 0 + 1.\surd : 0 \\
Ih1 & \stackrel{\text{def}}{=} & 1.\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & 1.infect : R2 + 1.\surd : R2 \\
B1 & \stackrel{\text{def}}{=} & 1.\surd : B2 \\
 & & \\
S2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1 - p_d).\surd : S0 \\
SI2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_i.\surd : I0 + (1 - p_d - p_i).\surd : S0 \\
I2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_r.\surd : R0 + (1 - p_d - p_r).\surd : I0 \\
R2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1 - p_d).\surd : R0 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : S0 \\
 & & \\
Population & \stackrel{\text{def}}{=} & S0\{s\} \times I0\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 5.2: Parallel infected agent with non-prioritised communication

size, $c = jN_t$. The model leads to the MFE,

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t - p_i(1-p_d)\frac{jN_tS_tI_t}{S_t+jN_tI_t+R_t} + p_{b_0}\left(1-\frac{kN_t}{p_{b_0}}\right)N_t \ , \\ I_{t+1} &= (1-p_d-p_r)I_t + p_i(1-p_d)\frac{jN_tS_tI_t}{S_t+jN_tI_t+R_t} \ , \\ R_{t+1} &= (1-p_d)R_t + p_rI_t \ . \end{aligned} \tag{5.1}$$

The transmission term in this system of MFE cannot be simplified, other than gathering constants to give

$$\frac{\beta N_tS_tI_t}{S_t+jN_tI_t+R_t} \ ,$$

with $\beta = p_i(1-p_d)j$, and therefore we do not obtain the traditional density dependent transmission term.

Parallel actions

An alternative method to introduce density dependent contact would be to use an infected agent with parallel actions,

$$I1b1.\overline{infect}^c : I2 + 1.\overline{infect}^{(c-1)} : I2... + 1.\overline{infect} : I2 + 1.\surd : I2 \ .$$

The model in Fig. 5.3 implements this for $c = 3$ with $I1b$ replacing $I1, Ih1$ and $Trans$ in Fig. 5.2. This model leads to MFE of the same form as Fig. 5.1 with the exception of the transmission term, which becomes

$$p_iS_t\frac{\sum_{n3=0}^{I_t}\sum_{n2=0}^{I_t-n3}\sum_{n1=0}^{I_t-(n2+n3)} M\binom{S_t+R_t-1}{\sum_{m=1}^{3}(m\times n_m)-1}}{\sum_{n3=0}^{I_t}\sum_{n2=0}^{I_t-n3}\sum_{n1=0}^{I_t-(n2+n3)} M\binom{S_t+R_t}{\sum_{m=1}^{3}(m\times n_m)}} \ , \tag{5.2}$$

where M is the multinomial coefficient

$$M = \frac{I_t!}{(\prod_{k=1}^{3} n_k!)(I_t - \sum_{l=1}^{3} n_l)!} \, .$$

Using the general form of $I1b$ given above, which can perform up to c instances of the *infect* output action, the transmission term in the MFE is generalised and becomes

$$p_i S_t \frac{\sum_{n_c=0}^{I_t} \sum_{n_{c-1}=0}^{I_t - n_c} \cdots \sum_{n_1=0}^{I_t - \sum_{j=2}^{c} n_j} M\binom{S_t + R_t - 1}{\sum_{m=1}^{c}(m \times n_m) - 1}}{\sum_{n_c=0}^{I_t} \sum_{n_{c-1}=0}^{I_t - n_c} \cdots \sum_{n_1=0}^{I_t - \sum_{j=2}^{c} n_j} M\binom{S_t + R_t}{\sum_{m=1}^{c} m \times n_m}} \, , \tag{5.3}$$

where M is the multinomial coefficient

$$M = \frac{I_t!}{(\prod_{k=1}^{c} n_k!)(I_t - \sum_{l=1}^{c} n_l)!} \, .$$

This transmission term is algebraically intractable because it does not simplify any further and, therefore, these MFE are not amenable to the usual range of algebraic analysis. In Chapter 3 it was shown that any model featuring non-prioritised communication that makes use of parallel actions will lead to MFE with a term of this sort, based on the general term (3.10). In addition to such MFE being intractable, calculating the time series will be more computationally expensive than for the terms that arise from any of the general terms for other types of communication, (3.6, 3.7, 3.9). We avoid writing models that use communication of this form because the resulting MFE do not offer the advantages that can generally be obtained using MFE as a method of analysis.

The problems we have when using non-prioritised communication come about because of the interpretation of the parameter c. To capture the assumptions of Begon et al. we require that the number of contacts be

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & p_b 0 - k(\lfloor S0 \rfloor + \lfloor I0 \rfloor + \lfloor R0 \rfloor) \\
 & & \\
S0 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + (1-p_b).\surd : S1 \\
I0 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + (1-p_b).\surd : I1 \\
R0 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + (1-p_b).\surd : R1 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S1 \times B1 \\
RepI & \stackrel{\text{def}}{=} & I1 \times B1 \\
RepR & \stackrel{\text{def}}{=} & R1 \times B1 \\
S1 & \stackrel{\text{def}}{=} & 1.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1b & \stackrel{\text{def}}{=} & 1.\overline{\mathit{infect}}^{3} : I2 + 1.\overline{\mathit{infect}}^{2} : I2 \\
 & & +1.\overline{\mathit{infect}}^{1} : I2 + 1.\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & 1.\mathit{infect} : R2 + 1.\surd : R2 \\
B1 & \stackrel{\text{def}}{=} & 1.\surd : B2 \\
 & & \\
S2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1-p_d).\surd : S0 \\
SI2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_i.\surd : I0 + (1-p_d-p_i).\surd : S0 \\
I2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_r.\surd : R0 + (1-p_d-p_r).\surd : I0 \\
R2 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1-p_d).\surd : R0 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : S0 \\
 & & \\
\mathit{Population} & \stackrel{\text{def}}{=} & S0\{s\} \times I0\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 5.3: Parallel infect action with non-prioritised communication

directly proportional to the overall population size. However, using non-prioritised communication, the agents do not necessarily make c contacts, c is merely the maximum number of contacts that each infected individual can make. The mean number of contacts made by infected individuals is a function not only of the population size but also of the mix of different types of agent in the population, so therefore we are not really capturing density dependent contact. In addition, if we were to genuinely capture functional behaviour that makes use of parallel actions (in the example considered here c is fixed) we would have to extend the functional parameters notation by defining functional parallel actions. Here we do not capture the desired behaviour, in terms of contact rate, so there is no advantage to be gained by defining functional parallel actions notation. To overcome this problem, in Sec. 5.2.2 models are presented that make use of prioritised communication so that agents must communicate where possible.

Consecutive contact actions

The third method that was proposed to implement multiple contacts was consecutive contact stages. This would involve the contact stage being replaced by c stages, in each of which the infected agents can perform the output action. Non-prioritised communication would once again result in such a model not capturing the desired behaviour, with the mean contact rate being dependent on the mix of the population as well as the value of c.

5.2.2 Prioritised communication

Parallel agents

The model in Fig. 5.4 differs from the model in Fig. 5.2 because all of the communicating agents are prioritised so that they communicate where

$$
\begin{array}{rcl}
c & \stackrel{\text{int}}{=} & min(Round(j * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor)), Cmax) \\
p_b & \stackrel{\text{prob}}{=} & p_b 0 - k(\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor) \\
 & & \\
S1 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + (1 - p_b).\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + (1 - p_b).\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + (1 - p_b).\surd : R2 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S2 \times B2 \\
RepI & \stackrel{\text{def}}{=} & I2 \times B2 \\
RepR & \stackrel{\text{def}}{=} & R2 \times B2 \\
S2 & \stackrel{\text{def}}{=} & \omega.infect : SI3 + 1.\surd : S3 \\
I2 & \stackrel{\text{def}}{=} & T2 \times Trans\{c\} \\
Trans & \stackrel{\text{def}}{=} & \omega.\overline{infect} : 0 + 1.\surd : 0 \\
T2 & \stackrel{\text{def}}{=} & \omega.infect : I3 + 1.\surd : I3 \\
R2 & \stackrel{\text{def}}{=} & \omega.infect : R3 + 1.\surd : R3 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : B3 \\
 & & \\
S3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1 - p_d).\surd : S1 \\
SI3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_i.\surd : I1 + (1 - p_d - p_i).\surd : S1 \\
I3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_r.\surd : R1 + (1 - p_d - p_r).\surd : I1 \\
R3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1 - p_d).\surd : R1 \\
B3 & \stackrel{\text{def}}{=} & 1.\surd : S1 \\
 & & \\
Population & \stackrel{\text{def}}{=} & S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 5.4: Parallel infected agent with prioritised communication

possible. In addition, the infected $I2$ agent is a parallel agent that includes a $T2$ agent, which can perform the *infect* input action, thereby explicitly modelling the situation where an infected individual makes an infectious contact with another infected individual. Once again the number of contacts that infected individuals can make is a proportion of the population size, $c = jN_t$. This model leads to the following system of mean field equations:

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t - \min\left(p_i S_t, \frac{jp_i(1-p_d)N_t S_t I_t}{N_t}\right) + p_{b_0}\left(1 - \frac{kN_t}{p_{b_0}}\right)N_t \ , \\ I_{t+1} &= (1-p_d-p_r)I_t + \min\left(p_i S_t, \frac{jp_i(1-p_d)N_t S_t I_t}{N_t}\right) \ , \\ R_{t+1} &= (1-p_d)R_t + p_r I_t \ . \end{aligned} \tag{5.4}$$

The transmission term here can be simplified to

$$\min\left(p_i S_t, \beta S_t I_t\right) \ ,$$

where $\beta = jp_i(1-p_d)$. Here we have density dependent transmission except where $jN_t I_t \geq N_t$ $(jI_t \geq 1)$, when all susceptible individuals will be contacted at every time step. Whether $jI_t \geq 1$ will ever be true depends on the chosen parameter values.

Parallel actions

As in the non-prioritised case it is also possible to write a model that utilises parallel actions, rather than a parallel $I2$ agent, to allow infected individuals to make c contacts. This is done by making use of a $Trans$ agent of the form

$$Transb \stackrel{\text{def}}{=} \omega^c.\overline{infect}^c : I3 + \omega^{c-1}.\overline{infect}^{c-1} : I3 + ... + \omega.\overline{infect} : I3 + 1.\surd : I3 \ .$$

The model in Fig. 5.5 implements this for $c = 3$ and leads to the following system of MFE:

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t - \min\left(p_i S_t, \frac{3S_t I_t}{N_t}\right) + p_{b_0}\left(1 - \frac{kN_t}{p_{b_0}}\right)N_t \ , \\ I_{t+1} &= (1-p_d-p_r)I_t + \min\left(p_i S_t, \frac{3S_t I_t}{N_t}\right) \ , \\ R_{t+1} &= (1-p_d)R_t + p_r I_t \ . \end{aligned} \tag{5.5}$$

If notation were defined to allow us to make the number of parallel instances of the infect action functional, with c taking the same form as in Fig. 5.4, this model would lead to (5.4), the MFE for the parallel agent method. This is because the behaviours being captured are identical with each infected individual making exactly c contacts when possible and all agents are contacted when $jI_t \geq 1$. Notation could be defined to do this but at present there seems to be no advantage since the behaviour it would capture can already be described using functional parallel agents.

Consecutive contact actions

In addition to the parallel action and parallel agent approach described above it is possible to allow the infected individuals to make contact on c consecutive contact stages before the probabilistic choice stage occurs. The model in Figs. 5.6 and 5.7 implements this approach for the situation where $c = 3$. The first stage of the model is a probabilistic stage, in which births and deaths happen. In the subsequent three stages the $Trans$ agents ($Trans2, Trans3, Trans4$) perform the *infect* output action and after the three stages become the null agent, 0. All of the types of agents that can perform the *infect* input action (susceptible, infected and recovered) can perform at most one instance of the action. The susceptible agents at each

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & p_b 0 - k(\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor) \\
 & & \\
S1 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + (1-p_b).\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + (1-p_b).\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + (1-p_b).\surd : R2 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S2 \times B2 \\
RepI & \stackrel{\text{def}}{=} & I2 \times B2 \\
RepR & \stackrel{\text{def}}{=} & R2 \times B2 \\
S2 & \stackrel{\text{def}}{=} & \omega.infect : SI3 + 1.\surd : S3 \\
Ib2 & \stackrel{\text{def}}{=} & T2 \times Transb \\
Transb & \stackrel{\text{def}}{=} & \omega^3.\overline{infect}^3 : I3 + \omega^2.\overline{infect}^2 : I3 \\
 & & +\omega.\overline{infect} : I3 + 1.\surd : I3 \\
T2 & \stackrel{\text{def}}{=} & \omega.infect : I3 + 1.\surd : I3 \\
R2 & \stackrel{\text{def}}{=} & \omega.infect : R3 + 1.\surd : R3 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : B3 \\
 & & \\
S3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1-p_d).\surd : S1 \\
SI3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_i.\surd : I1 + (1-p_d-p_i).\surd : S1 \\
I3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + p_r.\surd : R1 + (1-p_d-p_r).\surd : I1 \\
R3 & \stackrel{\text{def}}{=} & p_d.\surd : 0 + (1-p_d).\surd : R1 \\
B3 & \stackrel{\text{def}}{=} & 1.\surd : S1 \\
 & & \\
Population & \stackrel{\text{def}}{=} & S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 5.5: Parallel infect action with prioritised communication

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & p_b 0 - k(\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor R1 \rfloor) \\
 & & \\
S1 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + (1-p_b).\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + (1-p_b).\surd : I2 \\
R1 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + (1-p_b).\surd : R2 \\
RepS & \stackrel{\text{def}}{=} & S2 \times B2 \\
RepI & \stackrel{\text{def}}{=} & I2 \times B2 \\
RepR & \stackrel{\text{def}}{=} & R2 \times B2 \\
S2 & \stackrel{\text{def}}{=} & \omega.infect : SI3 + 1.\surd : S3 \\
I2 & \stackrel{\text{def}}{=} & T2 \times Trans2 \\
Trans2 & \stackrel{\text{def}}{=} & \omega.\overline{infect} : Trans3 + 1.\surd : Trans3 \\
T2 & \stackrel{\text{def}}{=} & \omega.infect : Ic3 + 1.\surd : 03 \\
R2 & \stackrel{\text{def}}{=} & \omega.infect : Rc3 + 1.\surd : R3 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : B3 \\
 & & \\
S3 & \stackrel{\text{def}}{=} & \omega.infect : SI4 + 1.\surd : S4 \\
SI3 & \stackrel{\text{def}}{=} & 1.\surd : SI4 \\
Ic3 & \stackrel{\text{def}}{=} & 1.\surd : Ic4 \\
Trans3 & \stackrel{\text{def}}{=} & \omega.\overline{infect} : Trans4 + 1.\surd : Trans4 \\
T3 & \stackrel{\text{def}}{=} & \omega.infect : Ic4 + 1.\surd : 04 \\
Rc3 & \stackrel{\text{def}}{=} & 1.\surd : Rc4 \\
R3 & \stackrel{\text{def}}{=} & \omega.infect : Rc4 + 1.\surd : R4 \\
B3 & \stackrel{\text{def}}{=} & 1.\surd : B4
\end{array}
$$

Figure 5.6: Prioritised contact on successive contacts. Part 1

$$
\begin{aligned}
S4 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI5 + 1.\surd : S5 \\
SI4 &\stackrel{\text{def}}{=} 1.\surd : SI5 \\
Ic4 &\stackrel{\text{def}}{=} 1.\surd : I5 \\
Trans4 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : 0 + 1.\surd : 0 \\
T4 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I5 + 1.\surd : I5 \\
Rc4 &\stackrel{\text{def}}{=} 1.\surd : R5 \\
R4 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R5 + 1.\surd : R5 \\
B4 &\stackrel{\text{def}}{=} 1.\surd : B5 \\
\\
S5 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1 - p_d).\surd : S1 \\
SI5 &\stackrel{\text{def}}{=} p_d.\surd : 0 + p_i.\surd : I1 + (1 - p_d - p_i).\surd : S1 \\
I5 &\stackrel{\text{def}}{=} p_d.\surd : 0 + p_r.\surd : R1 + (1 - p_d - p_r).\surd : I1 \\
R5 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1 - p_d).\surd : R1 \\
B5 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 5.7: Prioritised contact on successive contacts. Part 2

of these stages ($S2, S3$ or $S4$) can perform the infect input action, however if an agent communicates at the first or second stage it becomes an SI agent ($SI3$ or $SI4$), which does not communicate at subsequent communicative stages. At the final stage in the model any susceptible agent that has been contacted by an infected agent will be an $SI5$, which makes the probabilistic choice to become infected or not. In the case of the infected and recovered agents, if they make contact at the first or second contact stage they respectively become an Ic or Rc agent, which cannot communicate further in the later stages of the model. In the final probabilistic stage the infected and recovered agents are respectively $I5$ or $R5$, irrespective of whether they have communicated at any of the three stages. Despite the fact that this

model features more stages, and different agent behaviour at each stage, the overall mean behaviour is the same as that of the model in Fig. 5.4 with the constant contact rate $c = 3$. The MFE for this model are therefore

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t - \min\left(p_i S_t, \frac{3p_i S_t I_t}{N_t}\right) + p_{b_0}\left(1 - \frac{kN_t}{p_{b_0}}\right)N_t \ , \\ I_{t+1} &= (1-p_d-p_r)I_t + \min\left(p_i S_t, \frac{3p_i S_t I_t}{N_t}\right) \ , \\ R_{t+1} &= (1-p_d)R_t + p_r I_t \ . \end{aligned} \tag{5.6}$$

By having the number of stages in which the infected individuals make contact, and therefore the contact rate, a function that depends on the size of the population this model would lead to the same MFE, (5.4), as the parallel actions and parallel agents formulations. Again notation could be defined to allow this but at present there seems to be no advantage since we can already capture the desired behaviour with functional parallel agents.

5.3 Summary

In this Chapter we have presented several models that attempt to capture the assumptions used by Begon et al. [10] in their derivation of the traditional density dependent transmission term βSI . Using non-prioritised communication we found that density dependent transmission does not arise because we cannot correctly capture the assumptions of Begon et al. Using non-prioritised communication the average number of contacts made by each infected individual is not set merely by the density dependent parameter c but is also influenced by the mix of different agents in the population.

In contrast using prioritised communication we considered three different methods of implementing the contact rate c, which all lead to the

transmission term $\min(p_i S_t, \beta S_t I_t)$. The min term is a limitation of the discrete time nature of WSCCS. In continuous time ODE models, with fixed rates, the change in the population is continuously recalculated over infinitesimal steps of time, which prevents any of the groups in the population becoming negative. Other than occasions where the min term takes effect we have the traditional density dependent transmission term.

This result is in contrast to the results of Turner et al. [90] who found that irrespective of whether frequency dependent or density dependent behaviour was implemented at the individual level the frequency dependent term best captured behaviour at the population level, by fitting equations to simulation results. Turner et al. used cellular automata models that were inherently spatial and naturally feature clustering of infection. The WSCCS models presented here assume that the population is well mixed and do not allow for clustering. This is a key difference between the models that explains why different results were found.

Since we have three methods that lead to the traditional density dependent term in the MFE it is clear that we cannot guarantee finding a unique model that leads to a given system of MFE. The three approaches that achieve density dependent transmission were found by trial and error and the question of whether it is possible to obtain a WSCCS model that leads to a given system of difference equations remains open. In this chapter we have demonstrated that it is possible to define individual behaviour in our models that leads to the transmission term $\beta S_t I_t$. In addition Norman and Shankland [73] found that the term $\beta S_t I_t / N_t$ came about from their models. Although these two common terms can be obtained from our individual based modelling approach, it may not be possible to find WSCCS models that lead to some of the other terms that have been proposed [20, 48].

Chapter 6

Modes of Transmission II: Direct vs Indirect

Many different methods of disease transmission have been identified though no consensus exists about classification of these different methods. Here we make use of a description published online by Mount Sinai Hospital, Toronto [29], which defines the following six methods by which transmission can occur:

- Direct contact: requires physical contact between susceptible and infected individuals (includes sexual transmission).
- Fomite transmission: susceptible individual is infected by contact with an infected touch-surface (fomite). Some organisms can survive on a fomite for lengthy periods (e.g. Norwalk virus).
- Droplet contact: susceptible individuals can become infected by contact with infected droplets that are produced when infected individuals cough or sneeze. Droplets cannot persist in the environment for long and therefore require close proximity between infected and

susceptible individuals. Diseases that can be transmitted in this way include measles and SARS.

- Airborne transmission: droplet nuclei (evaporated droplets) or infected dust particles can remain suspended in air. Infection can persist in this way for longer periods and can travel over greater distances. Diseases that can be transmitted in this way include influenza, tuberculosis and foot and mouth disease.

- Fecal-oral transmission: susceptible individuals come into contact with infected fecal matter through consumption of contaminated food or water. This can occur over large distances, for example by contaminated water supply. Usually associated with microorganisms that infect the digestive system.

- Vector-borne transmission: susceptible individuals come into contact with an infected vector (e.g. mosquitoes, ticks, rats). Vectors are mobile so that transmission of this form can happen over large distances.

For the purposes of developing WSCCS models we consider all of these modes of transmission to be represented by either direct or indirect transmission, as outlined in Table 6.1. The models presented in the preceding chapters all represent direct transmission since they require communication directly between susceptible and infected individuals. In order to capture the full range of disease behaviour it is necessary to be able to develop models of indirect transmission.

The traditional ODE models of indirect spread [6, 9] make use of an additional equation, which describes the amount of free living infection in the environment. In this chapter we present WSCCS models that capture indirect transmission by introducing agents to represent the environment.

Direct Transmission	Indirect Transmission
Direct contact	Fomite transmission
Droplet contact	Airborne transmission
	Fecal-oral transmission
	Vector-borne transmission

Table 6.1: Modes of transmission classified as direct or indirect

When an environment agent becomes infected it retains the infection and can pass it on to susceptible individuals at a later time. This is similar to an idea developed by Bradley et al. [17] to model the spread of internet worms using PEPA. In that case agents represent computers and routers on the network. Infected computers can pass the virus on to a router, which passes the infection on to another computer. This introduces a delay in the spread of the virus between computers since there is a time cost associated with both the transmission to the router and transmission from the router on to an uninfected computer.

Models have been developed that make use of a finite number of environment agents, or allow for unlimited quantities of infected environment agents. Since these models do not feature any density dependent behaviour we consider a simplified situation that does not include any births or deaths. This simplification is reasonable for many disease systems where the infection spreads rapidly through the population, so that the probability of any individual giving birth or dying during the course of an epidemic is so low as to make such events insignificant, and that have negligible levels of disease induced mortality.

6.1 Finite environment

In this section we consider models that feature a finite number of environment agents and are three stage models. At the first stage any infected environment agents that exist contact the population, potentially passing on the infection to susceptible individuals (this step uses prioritised communication - this means that we are not implicitly modelling a situation where an environment agent can contact another environment agent). The second stage involves the infected individuals passing the infection on to the environment (uses non-prioritised communication - this means that not all of the environment will automatically become infected when there are sufficiently many infected individuals) and the final tick involves probabilistic choices. If the order of the two contact stages were reversed - transmission from the population to the environment followed by transmission from the environment to the population - the MFE would not reflect the time delay that is inherent in traditional models of indirect transmission, although the transmission terms in the MFE would be considerably more complicated. We use frequency dependent transmission to avoid the added complexity associated with density dependent transmission in WSCCS and so that we focus on the effects of indirect transmission. If we wished to capture density dependent transmission we could implement this in the same way as in Chapter 5, with the number of contacts made with the infected environment agents proportional to the size of the population. We consider three models that feature finite environment and differ in the way in which infection in the environment decays: i) infection decays probabilistically ii) infection decays after a fixed period, persisting for one iteration of the model iii) infection decays after a fixed period, persisting for more than one iteration of the model.

6.1.1 Infection decays probabilistically

The model in Fig. 6.1 features environment agents that lose their infectivity probabilistically. During the probabilistic stage the infected environment agents ($Ei3$) become uninfected $E1$ agents with probability p_o, representing infection decaying. The agents that became infected during the current iteration of the model ($Ei03$) do not make the choice to become uninfected at this stage and are all able to pass on the infection during the next iteration. This model leads to the following system of five equations:

$$\begin{aligned}
S_{t+1} &= S_t - \frac{p_i Ei_t S_t}{S_t + I_t + R_t} \,, \\
I_{t+1} &= (1-p_r) I_t + \frac{p_i Ei_t S_t}{S_t + I_t + R_t} \,, \\
R_{t+1} &= R_t + p_r I_t \,, \\
E_{t+1} &= E_t + p_o Ei_t - \frac{E_t I_t}{E_t + Ei_t + I_t} \,, \\
Ei_{t+1} &= (1-p_o) Ei_t + \frac{E_t I_t}{E_t + Ei_t + I_t} \,.
\end{aligned}$$

By noting that the total number of agents representing the environment is fixed, $E_t + Ei_t = E_{t+1} + Ei_{t+1} = C$ for some constant C, we can rewrite the equations for the environment as

$$\begin{aligned}
Ei_{t+1} &= (1-p_o) Ei_t + \frac{E_t I_t}{C + I_t} \,, \\
E_{t+1} &= C - Ei_{t+1} \,.
\end{aligned}$$

This greatly simplifies the equation for E_t but we would not achieve further simplification by substituting for Ei_t in the equations for S_t and I_t since the equation for Ei_{t+1} is expressed in terms of E_t and Ei_t .

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
E1 &\stackrel{\text{def}}{=} 1.\surd : E2 \\
Ei1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : Ei2 + 1.\surd : Ei2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S3 \\
I2 &\stackrel{\text{def}}{=} 1.\overline{\mathit{environ}} : I3 + 1.\surd : I3 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R3 \\
SI2 &\stackrel{\text{def}}{=} 1.\surd : SI3 \\
E2 &\stackrel{\text{def}}{=} 1.\mathit{environ} : Ei03 + E3 \\
Ei2 &\stackrel{\text{def}}{=} 1.\mathit{environ} : Ei3 + Ei3 \\
\\
S3 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
I3 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1 - p_r).\surd : I1 \\
R3 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
Ei3 &\stackrel{\text{def}}{=} (1 - p_o).\surd : Ei1 + p_o.\surd : E1 \\
Ei03 &\stackrel{\text{def}}{=} 1.\surd : Ei1 \\
E3 &\stackrel{\text{def}}{=} 1.\surd : E1 \\
SI3 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \times E1\{c\} \lceil \{\surd\}
\end{aligned}
$$

Figure 6.1: Indirect transmission. Finite environment, infection decays probabilistically.

6.1.2 Infection persists in environment for fixed time

Infection persists for one iteration

The model in Fig. 6.2 features infected environment agents that only retain their infectivity for one iteration. After the transmission stage, when the disease can be passed on to the susceptible individuals, the infected $Ei1$ agents all become uninfected $E2$ agents, which can be infected by the infected $I2$ agents. The model in Fig. 6.2 leads to the following system of five equations:

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i Ei_t S_t}{S_t + I_t + R_t} \,, \\ I_{t+1} &= (1-p_r) I_t + \frac{p_i Ei_t S_t}{S_t + I_t + R_t} \,, \\ R_{t+1} &= R_t + p_r I_t \,, \\ E_{t+1} &= E_t + Ei_t - \frac{(E_t + Ei_t) I_t}{E_t + Ei_t + I_t} \,, \\ Ei_{t+1} &= \frac{(E_t + Ei_t) I_t}{E_t + Ei_t + I_t} \,. \end{aligned} \tag{6.1}$$

By once again noting that there are a fixed number of environment agents, $E_t + Ei_t = E_{t+1} + Ei_{t+1} = C$, we can rewrite the equations for Ei_{t+1} and E_{t+1} as

$$\begin{aligned} Ei_{t+1} &= \frac{C \times I_t}{C + I_t} \,, \\ E_{t+1} &= C - Ei_{t+1} \,. \end{aligned}$$

We can now substitute for Ei_t ,

$$Ei_t = \frac{C \times I_{t-1}}{C + I_{t-1}} \,,$$

in the equations for S_{t+1} and I_{t+1} and, since Ei_t is expressed merely in terms

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
E1 &\stackrel{\text{def}}{=} 1.\surd : E2 \\
Ei1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : E2 + 1.\surd : E2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S3 \\
SI2 &\stackrel{\text{def}}{=} 1.\surd : SI3 \\
I2 &\stackrel{\text{def}}{=} 1.\overline{\mathit{environ}} : I3 + 1.\surd : I3 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R3 \\
E2 &\stackrel{\text{def}}{=} 1.\mathit{environ} : Ei3 + 1.\surd : E3 \\
\\
S3 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
I3 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1 - p_r).\surd : I1 \\
R3 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
Ei3 &\stackrel{\text{def}}{=} 1.\surd : Ei1 \\
E3 &\stackrel{\text{def}}{=} 1.\surd : E1 \\
SI3 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \times E1\{e\} \lceil \{\surd\}
\end{aligned}
$$

Figure 6.2: Indirect transmission. Finite environment, infection persists for only one iteration.

of C and I_{t-1}, represent the mean of the model by the following system of three equations:

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i C \times I_{t-1} S_t}{(C + I_{t-1})(S_t + I_t + R_t)} , \\ I_{t+1} &= (1 - p_r) I_t + \frac{p_i C \times I_{t-1} S_t}{(C + I_{t-1})(S_t + I_t + R_t)} , \\ R_{t+1} &= R_t + p_r I_t . \end{aligned} \tag{6.2}$$

This means that, for this model, we can describe the population by a system of three second order difference equations, if we are not interested in the number of infected environment agents at any given time. However if we do wish to know what proportion of the environment is infected at any given time we must consider the system of five first order equations (6.1).

Another interesting question is what happens when we consider very large C, suggesting that the environment is plentiful. Considering the limiting case where $C \to \infty$ gives

$$\frac{C \times I_{t-1}}{C + I_{t-1}} \to I_{t-1} ,$$

and at the limit the transmission term in (6.2) becomes

$$\frac{p_i S_t I_{t-1}}{S_t + I_t + R_t} . \tag{6.3}$$

This is the familiar frequency dependent transmission term with a delay since the term depends not only on the state of the system at time t but also on the number of infected agents at time $(t-1)$. In Section 6.2 we present models that explicitly feature unlimited environments and we will compare the terms arising from those models to (6.3).

Infection persists for two iterations

Rather than persisting for a single iteration we can consider the situation where infection persists in the environment for more than one iteration. We present a model in which infection persists for two iterations, Fig. 6.3, and then generalise to find MFE for the general case where the infection persists for n iterations, for some integer n.

After the contact phase of Fig. 6.3 the $Ei1$ infected environment agents all become $Eib2$ agents, which can also be contacted by the infected agents when the disease is passed on to the environment, and if an $Eib2$ agent does come into contact with a $I2$ agent they become the infected environment agent $Ei3$, which can pass on the infection for in the following two iterations. In the next iteration these agents will be $Eib1$ agents, which can once again pass the infection to the population, and at this stage lose the infection becoming $E2$ uninfected environment agents.

This model leads to the following system of six equations to describe the evolution of the model:

$$\begin{aligned}
S_{t+1} &= S_t - \frac{p_i(Ei_t + Eib_t)S_t}{S_t + I_t + R_t} \,, \\
I_{t+1} &= (1-p_r)I_t + \frac{p_i(Ei_t + Eib_t)S_t}{S_t + I_t + R_t} \,, \\
R_{t+1} &= R_t + p_r I_t \,, \\
E_{t+1} &= E_t + Eib_t - \frac{(E_t + Eib_t)I_t}{E_t + Ei_t + Eib_t + I_t} \,, \\
Ei_{t+1} &= \frac{(E_t + Ei_t + Eib_t)I_t}{E_t + Ei_t + Eib_t + I_t} \,, \\
Eib_{t+1} &= Ei_t - \frac{Ei_t I_t}{E_t + Ei_t + Eib_t + I_t} \,.
\end{aligned} \tag{6.4}$$

$$
\begin{aligned}
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
E1 &\stackrel{\text{def}}{=} 1.\surd : E2 \\
Ei1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : Eib2 + 1.\surd : Eib2 \\
Eib1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : E2 + 1.\surd : E2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S3 \\
SI2 &\stackrel{\text{def}}{=} 1.\surd : SI3 \\
I2 &\stackrel{\text{def}}{=} 1.\overline{\mathit{environ}} : I3 + 1.\surd : I3 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R3 \\
E2 &\stackrel{\text{def}}{=} 1.\mathit{environ} : Ei3 + 1.\surd : E3 \\
Eib2 &\stackrel{\text{def}}{=} 1.\mathit{environ} : Ei3 + 1.\surd : Eib3 \\
\\
S3 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
SI3 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1-p_i).\surd : S1 \\
I3 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1-p_r).\surd : I1 \\
R3 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
E3 &\stackrel{\text{def}}{=} 1.\surd : E1 \\
Ei3 &\stackrel{\text{def}}{=} 1.\surd : Ei1 \\
Eib3 &\stackrel{\text{def}}{=} 1.\surd : Eib1 \\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \times E1\{e\} \lceil \{\surd\}
\end{aligned}
$$

Figure 6.3: Indirect transmission. Finite environment, infection persists for only two iterations.

We can note that the environment is once again constant, this time with $C = E_t + Ei_t + Eib_t$. This means the equations for the environment can be written as

$$\begin{aligned} Ei_{t+1} &= \frac{C \times I_t}{C + I_t} , \\ Eib_{t+1} &= Ei_t - \frac{Ei_t I_t}{C + I_t} , \\ E_{t+1} &= C - Ei_{t+1} - Eib_{t+1} . \end{aligned}$$

Unlike the situation where the infection persists in the environment for one iteration we cannot simplify the the equations by substituting for Ei_t and Eib_t in the equations for S_{t+1} and I_{t+1} and to study the population we must consider the full system of six equations (6.4). However if we once again consider what happens when $C \to \infty$ we find

$$\begin{aligned} Ei_{t+1} &= I_t , \\ Eib_{t+1} &= Ei_t = I_{t-1} , \end{aligned}$$

and we can substitute for Ei_t and Eib_t in the equations for S_{t+1} and I_{t+1} . This allows us to describe the population by a system of three third order equations:

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i(I_{t-1} + I_{t-2})S_t}{S_t + I_t + R_t} , \\ I_{t+1} &= (1 - p_r)I_t + \frac{p_i(I_{t-1} + I_{t-2})S_t}{S_t + I_t + R_t} , \\ R_{t+1} &= R_t + p_r I_t . \end{aligned} \tag{6.5}$$

If the infection persists for more than two iterations the MFE would become more complicated in two ways. In a general case, where the infection

persists for n iterations, the transmission term would take the form

$$\frac{p_i(Eia_t + Eib_t + Eic_t + ... + Ein_t)S_t}{S_t + I_t + R_t} ,$$

and there would be $n + 1$ equations that describe the environment. Again considering $C \to \infty$ the transmission term takes the form

$$\frac{p_i(I_{t-1} + I_{t-2} + I_{t-3} + ... + I_{t-n})S_t}{S_t + I_t + R_t} .$$

6.1.3 Summary of MFE

The terms associated with transmission of the infection from infected individuals to the environment take the form

$$\frac{E_t I_t}{E_t + Ei_t + I_t} . \tag{6.6}$$

The I_t in the denominator of (6.6) suggests that the infected individuals, as well as passing the infection to the environment, are also absorbing the infectious contacts made by other infecteds. If we were to directly write down equations to describe indirect transmission this term would instead take the form

$$\frac{E_t I_t}{E_t + Ei_t} ,$$

since we would expect the rate at which new infected portions of environment are created to depend on the number of infecteds (I_t), which spread the infection, and the probability that a particular portion of environment is uninfected $E_t/(E_t + Ei_t)$.

We could model this by making use of prioritised communication between infected individuals and the environment, which would lead to terms of the form

$$\frac{E_t I_t}{E_t + Ei_t} .$$

Such an approach, however, would mean that when there are sufficiently many infected individuals in the population ($I_t > C$) all of the environment would become infected and the term should correctly be written as

$$min\left(E_t, \frac{E_t I_t}{E_t + Ei_t}\right) .$$

This is a consequence of the fact that we are considering a fixed quantity of environment; however, it may be more realistic to consider an unlimited environment since the area that contains the environment may be very small such that the number of distinct portions of environment is so large as to not be a limiting factor. In Section 6.2 we consider models for the situation where there is no upper limit on the number of environment agents.

The other point about these equations that is unusual is the terms used to describe transmission from the environment to the population. These take the general form

$$\frac{\beta S_t Ei_t}{S_t + I_t + R_t} .$$

Mathematical models [6] more commonly assume density dependent transmission between the environment and the population. We could implement this in WSCCS using the mechanisms described in Chapter 5 but chose not to do so here so that the focus of our models is the indirect transmission and to ensure that this does not become confused by the greater complexity required by density dependent transmission.

6.2 Unlimited environment

In this section we present models in which the infected individuals probabilistically spawn infected environment agents. The infected environment agents can, in future iterations, pass the infection to susceptible individuals. In this case, when a unit of environment is no longer infectious, it becomes the null agent 0. This means that there is no artificial upper limit on the number of infected environment agents that can be created. We would expect models developed in this way to lead to the same equations as those derived for the fixed number of environment agents when we considered $C \to \infty$.

Once again models were developed with the infection either decaying probabilistically or persisting for a fixed period. In both of these models it is no longer necessary to have a communication phase where infection of the environment can occur so these are two stage models. At the first stage communication occurs with the population potentially making contact with the infected environment agents (again using prioritised communication) while the second stage involves probabilistic choice, including the choice of the infected individuals to produce an infected environment agent.

6.2.1 Infection decays probabilistically

The model in Fig. 6.4 has the environment agents probabilistically being removed from the system. With probability p_o the infected environment agent $E2$ becomes the null agent 0.

$$
\begin{aligned}
IE &\stackrel{\text{def}}{=} I1 \times E1 \\
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
E1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : E2 + 1.\surd : E2 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
I2 &\stackrel{\text{def}}{=} p_e.\surd : IE + p_r.\surd : R1 + (1 - p_e - p_r).\surd : I1 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
SI2 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
E2 &\stackrel{\text{def}}{=} (1 - p_o).\surd : E1 + p_o.\surd : 0 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 6.4: Indirect transmission. Unlimited environment, infection decays probabilistically.

The system of MFE for this model is

$$
\begin{aligned}
S_{t+1} &= S_t - \frac{p_i E_t S_t}{S_t + I_t + R_t} \,, \\
I_{t+1} &= (1 - p_r) I_t + \frac{p_i E_t S_t}{S_t + I_t + R_t} \,, \\
R_{t+1} &= R_t + p_r I_t \,, \\
E_{t+1} &= (1 - p_o) E_t + p_e I_t \,.
\end{aligned}
$$

These equations are simpler than the corresponding MFE for the case where there is a fixed number of environment agents, since we have only one equation for the environment. Here there is no advantage in substituting for E_t in the equations for S_{t+1} and I_{t+1} since E_{t+1} is expressed in terms of E_t.

$$
\begin{aligned}
IE &\stackrel{\text{def}}{=} I1 \times E1 \\
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
E1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : 0 + 1.\surd : 0 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
I2 &\stackrel{\text{def}}{=} p_e.\surd : IE + p_r.\surd : R1 + (1 - p_e - p_r).\surd : I1 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
SI2 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 6.5: Indirect transmission. Unlimited environment, infection persists for only one iteration.

6.2.2 Infection persists in environment for fixed time

Infection persists for one iteration

Fig. 6.5 features a model with environment agents that become the null agent after the transmission stage irrespective of whether they make contact or not. The MFE for this model are

$$
\begin{aligned}
S_{t+1} &= S_t - \frac{p_i E_t S_t}{S_t + I_t + R_t} \,, \\
I_{t+1} &= (1 - p_r) I_t + \frac{p_i E_t S_t}{S_t + I_t + R_t} \,, \\
R_{t+1} &= R_t + p_r I_t \,, \\
E_{t+1} &= p_e I_t \,.
\end{aligned}
$$

Substituting for E_t ($E_t = p_e I_{t-1}$) in the transmission term allows us to represent the model by a system of three equations since we do not care about the amount of infection in the environment, only the infection in the

population. This gives us

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i p_e I_{t-1} S_t}{S_t + I_t + R_t} \,, \\ I_{t+1} &= (1-p_r) I_t + \frac{p_i p_e I_{t-1} S_t}{S_t + I_t + R_t} \,, \\ R_{t+1} &= R_t + p_r I_t \,. \end{aligned}$$

These equations feature a transmission term,

$$\frac{p_i p_e I_{t-1} S_t}{S_t + I_t + R_t} \,,$$

of the same form as (6.3), the transmission term that was obtained in Section 6.1.2 for the limit of the fixed number of environment agents, $C \to \infty$.

Infection persists for two iterations

The model in Fig. 6.6 has the infection persisting for two iterations. After the contact phase the $E1$ infected environment agents all become $Eb2$ agents, which then become $Eb1$ agents and can pass on the infection for a second time before becoming the null agent 0.

This model leads to the following system of five equations:

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i (E_t + Eb_t) S_t}{S_t + I_t + R_t} \,, \\ I_{t+1} &= (1-p_r) I_t + \frac{p_i (E_t + Eb_t) S_t}{S_t + I_t + R_t} \,, \\ R_{t+1} &= R_t + p_r I_t \,, \\ E_{t+1} &= p_e I_t \,, \\ Eb_{t+1} &= E_t = p_e I_{t-1} \,. \end{aligned}$$

We can now substitute for E_t and Eb_t in the equations for S_{t+1} and I_{t+1}.

$$
\begin{aligned}
IE &\stackrel{\text{def}}{=} I1 \times E1 \\
S1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : SI2 + 1.\surd : S2 \\
I1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : I2 + 1.\surd : I2 \\
R1 &\stackrel{\text{def}}{=} \omega.\mathit{infect} : R2 + 1.\surd : R2 \\
E1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : E2 + 1.\surd : E2 \\
Eb1 &\stackrel{\text{def}}{=} \omega.\overline{\mathit{infect}} : 0 + 1.\surd : 0 \\
\\
S2 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
I2 &\stackrel{\text{def}}{=} p_e.\surd : IE + p_r.\surd : R1 + (1 - p_e - p_r).\surd : I1 \\
R2 &\stackrel{\text{def}}{=} 1.\surd : R1 \\
SI2 &\stackrel{\text{def}}{=} p_i.\surd : I1 + (1 - p_i).\surd : S1 \\
E2 &\stackrel{\text{def}}{=} 1.\surd : Eb1 \\
\\
\mathit{Population} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 6.6: Indirect transmission. Unlimited environment, infection persists for only two iterations.

This allows us to describe the model by a third order system of three equations, provided we are not interested in the number of infected environment agents:

$$\begin{aligned} S_{t+1} &= S_t - \frac{p_i p_e (I_{t-1} + I_{t-2}) S_t}{S_t + I_t + R_t} \ , \\ I_{t+1} &= (1 - p_r) I_t + \frac{p_i p_e (I_{t-1} + I_{t-2}) S_t}{S_t + I_t + R_t} \ , \\ R_{t+1} &= R_t + p_r I_t \ . \end{aligned}$$

The transmission term here,

$$\frac{p_i p_e (I_{t-1} + I_{t-2}) S_t}{S_t + I_t + R_t} \ ,$$

is once again of a similar form to (6.3). The main difference here is the transmission term, which depends not only on I_{t-1} but also on I_{t-2} .

If similarly the infection persists in the environment for a fixed period of more than two iterations the equations for the environment can be eliminated. If n is the number of iterations for which the infection persists the transmission term would then take the form

$$\frac{p_i p_e (I_{t-1} + I_{t-2} + ... + I_{t-n}) S_t}{S_t + I_t + R_t} \ .$$

In addition if we were interested in the number of infected environment agents at a given time we would have to consider n separate equations, each of the form

$$Ei_{t+1} = p_e I_{t-i} \ ,$$

for $1 \leq i \leq n$.

The MFE found here are of the same form as those found in Section

6.1 when we considered the fixed number of environment agent $C \to \infty$: however, here we have captured this behaviour explicitly rather than relying on manipulation of the derived MFE.

6.3 Summary

In this chapter we have presented models for the indirect spread of disease, making use of agents to represent portions of environment that can become infected and pass the disease back to the population. Models were presented that feature either a fixed number, or unlimited environment agents as well as models where the infection decays in the environment in different ways. From a biological realism perspective the models that have the infection decaying probabilistically (Figs. 6.1 and 6.4) are preferable since all events in the real world are fundamentally stochastic. For example infection in the environment can decay more or less quickly depending on exposure to sunlight or temperature. In addition the advantage of modelling in WSCCS is that we capture the underlying behaviour of individuals from which the population level behaviour emerges. However, for the purposes of performing algebraic analysis on the MFE the more simple terms that come about from the models in which infection persists for one iteration (Figs. 6.2 and 6.5) are preferable. These models have the further advantage that the equations for the environment can be eliminated and the mean behaviour of the population can be described by only three equations.

Some diseases persist in the environment for only short periods of time, particularly those that cannot survive in a dried state, e.g. measles [29], and for these diseases the probability of the infection in the environment decaying within the duration of an iteration of the model may be close to 1. It would be a reasonable assumption to model these diseases as persisting

in the environment for one iteration of the model. However, the models in which the disease deterministically persists for two iterations of the model (Figs. 6.3 and 6.6), or more than two iterations, are unrealistic since if we expect all of the infection to decay after n iterations we would expect a proportion to have decayed after $n-1$ iterations. Therefore we can say that if a model is to be developed in which the infection persists for a fixed period it should be for only one iteration of the model.

The other consideration about these models is whether we should have a fixed or unlimited number of environment agents. It should be noted that the area contaminated by an infected individual may be small with an adjacent portion of environment remaining uncontaminated. For this reason the number of units of environment will be very large and it will be reasonable to assume that there is no limit on the quantity of infection contained by the environment.

Chapter 7

Superspreaders

Superspreaders have been identified as being important in the spread of many diseases [60]. In such cases it has been observed that a small proportion of the infected individuals are responsible for the majority of new infections. Two mechanisms have been proposed to explain superspreaders and this chapter presents models of each mechanism. In the following superspreader models the infected portion of the population consists of two distinct groups: standard infected individuals (I) and superspreaders (U).

The first mechanism that leads to superspreaders is increased infectiousness (also known as supershedders). If a susceptible individual is contacted by a supershedder the probability of becoming infected (p_{iu}) is higher than the probability of becoming infected having been contacted by an infected individual (p_i), i.e. $p_{iu} = \alpha p_i$ for some constant $\alpha > 1$. The second mechanism that leads to superspreaders is an increased contact rate. For contact superspreaders the contact rate (c_u) is higher than the contact rate of the infected individuals (c_i), i.e. $c_u = \alpha c_i$.

Supershedders may arise because of a compromised immune system (meaning, for instance, that more virus is present in their body so the quantity

shed is greater) or because of some genetic predisposition that causes them to shed a greater quantity of the disease. Contact superspreaders are more gregarious or well travelled than the average and therefore make more contacts than other individuals in the population.

7.1 Supershedders

To capture the behaviour of supershedders we must be able to differentiate between whether a susceptible individual has been contacted by a superspreader or by a standard infected individual. To do this we have the infected and superspreader individuals perform different actions (*infect* and *infectU*) with the contacted susceptible individual making the choice to become infected using the relevant probability depending on the action performed. The standard infected and superspreader individuals have the same contact rate: fixed at one here for simplicity. If necessary an increased contact rate could be used, but it should be the same for the two groups if we are to distinguish between supershedders and contact superspreaders. The models in this section make use of the method for implementing births and deaths from Fig. 4.1, which leads to a logistic growth term in the susceptibles equation (by choosing the probability of giving birth to be inversely proportional to the population size, $p_b = p_{b_0} - kN_t$). The other probabilities are as in previous chapters i.e. p_r - probability of recovery; p_d - probability of death due to natural causes; and p_{dd} - probability of death due to the disease. Transmission is frequency dependent, with infected individuals and superspreaders able to make one contact per time step, but existing methods of implementing density dependent transmission could be introduced and would change only the transmission term in the derived equations. The first stage in the model involves probabilistic births with all individuals giving

birth to a newborn individual, which will go on to become a susceptible individual at the first stage of the next iteration of the model.

7.1.1 Single contact stage

Prioritised communication

The first model considered here, shown in Fig. 7.1, makes use of a different form of communication than the models presented in previous chapters. To distinguish between contact with standard infecteds and supershedders the susceptible $S2$ agents must perform different actions when communicating with $Trans2$ and $TransU2$ agents. To do this the $S2$ agents,

$$S2 \stackrel{\text{def}}{=} \omega.infect : SI3 + \omega.infectU : SU3 + 1.\surd : S3 \ ,$$

choose between two different communicating actions (*infect* and *infectU*), as well as free action $\surd$. The $S2$ agent is prioritised to perform these actions at the same priority level and with the same weight so the mean number of $S2$ agents that perform each action depends only on the mix of different agents in the population. In addition the $T2$ and $R2$ agents communicate in the same way to model the situations where an infected or superspreader individual makes contact with an individual that already has the disease, or has previously had the disease and is now immune to further infection. At the next stage the $SI3$ and $SU3$ agents (susceptibles that have been contacted by standard infecteds and supershedders respectively) have different probabilities of infection (p_i and $p_{iu} = \alpha p_i$ respectively).

Although the algorithm described in Chapter 3 does not cover communication of this form we can find generalised MFE, with labels SI_{new} and SU_{new} denoting the numbers of susceptibles that have communicated with

$$
\begin{aligned}
p_b &\overset{\text{prob}}{=} = p_b 0 - k * (\lfloor S1 \rfloor + [\lfloor I1 \rfloor + \lfloor U1 \rfloor + \lfloor R1 \rfloor) \\
\\
S1 &\overset{\text{def}}{=} p_b.\surd : RepS + (1-p_b).\surd : S2 \\
I1 &\overset{\text{def}}{=} p_b.\surd : RepI + (1-p_b).\surd : I2 \\
U1 &\overset{\text{def}}{=} p_b.\surd : RepU + (1-p_b).\surd : U2 \\
R1 &\overset{\text{def}}{=} p_b.\surd : RepR + (1-p_b).\surd : R2 \\
\\
RepS &\overset{\text{def}}{=} S2 \times B2 \\
RepI &\overset{\text{def}}{=} I2 \times B2 \\
RepU &\overset{\text{def}}{=} U2 \times B2 \\
RepR &\overset{\text{def}}{=} R2 \times B2 \\
S2 &\overset{\text{def}}{=} \omega.infect : SI3 + \omega.infectU : SU3 + 1.\surd : S3 \\
I2 &\overset{\text{def}}{=} T2 \times Trans2 \\
Trans2 &\overset{\text{def}}{=} \omega.\overline{infect} : I3 + 1.\surd : I3 \\
T2 &\overset{\text{def}}{=} \omega.infect : 0 + \omega.infectU : 0 + 1.\surd : 0 \\
U2 &\overset{\text{def}}{=} T2 \times TransU2 \\
TransU2 &\overset{\text{def}}{=} \omega.\overline{infectU} : U3 + 1.\surd : U3 \\
R2 &\overset{\text{def}}{=} \omega.infect : R3 + \omega.infectU : R3 + 1.\surd : R3 \\
B2 &\overset{\text{def}}{=} 1.\surd : B3 \\
\\
S3 &\overset{\text{def}}{=} (1-p_d).\surd : S1 + p_d.\surd : 0 \\
SI3 &\overset{\text{def}}{=} (p_i * (1-p_s)).\surd : I1 + (p_i * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i).\surd : S1 + p_d.\surd : 0 \\
SU3 &\overset{\text{def}}{=} (p_i u * (1-p_s)).\surd : I1 + (p_i u * p_s).\surd : U1 \\
&\quad +(1-p_d- p_i u).\surd : S1 + p_d.\surd : 0 \\
I3 &\overset{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : I1 + (p_d+p_{dd}).\surd : 0 \\
U3 &\overset{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : U1 + (p_d+p_{dd}).\surd : 0 \\
R3 &\overset{\text{def}}{=} p_d.\surd : 0 + (1-p_d).\surd : R1 \\
B3 &\overset{\text{def}}{=} 1.\surd : S1 \\
\\
Popn &\overset{\text{def}}{=} S1\{s\} \times I1\{i\} \times U1\{u\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 7.1: Supershedder model with density dependent probability of giving birth - contact in one stage - prioritised communication

standard infecteds and superspreaders respectively. The MFE found in this way are

$$\begin{aligned} S_{t+1} &= (1-p_d)S_t - p_i SI_{new} - \alpha p_i SU_{new} + N_t(p_{b_0} - kN_t) \ , \\ I_{t+1} &= (1-p_d-p_{dd}-p_r)I_t + p_i(1-p_s)SI_{new} + \alpha p_i(1-p_s)SU_{new} \ , \\ U_{t+1} &= (1-p_d-p_{dd}-p_r)U_t + p_i p_s SI_{new} + \alpha p_i p_s SU_{new} \ , \\ R_{t+1} &= (1-p_d)R_t + p_r(I_t+U_t) \ , \end{aligned} \tag{7.1}$$

where $p_{iu} = \alpha p_i$ for some $\alpha > 1$.

Despite not being able to automatically obtain expressions for SI_{new} and SU_{new} it is possible to obtain such expressions by careful consideration of the system. In doing this we make use of multinomial coefficients of the form

$$\binom{X}{X_{infect}, X_{infectU}, X_t} = \frac{X!}{X_{infect}!X_{infectU}!X_t!} \ ,$$

where X is the number of agents of type X in the population and X_{infect}, $X_{infectU}$ and X_t are the numbers of the X agents that perform the actions $infect, infectU$ and $\surd$ respectively. The weight associated with a particular change in the population over this communicating stage is

$$\binom{S2}{S2_{infect}, S2_{infectU}, S2_t}\binom{T2}{T2_{infect}, T2_{infectU}, T2_t}\binom{R2}{R2_{infect}, R2_{infectU}, R2_t} \ .$$

Note we do not have to consider binomial coefficients for the number of the $Trans2$ and $TransU2$ agents that communicate since there will always be

sufficient $S2, T2$ and $R2$ for them all to communicate. By considering the average of all the ways in which the population can evolve we get

$$SI_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} a\binom{S2}{a,b,S2-a-b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c,d,T2-c-d}\binom{R2}{I2-a-c,U2-b-d,R2-I2-U2+a+b+c+d}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a,b,S2-a-b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c,d,T2-c-d}\binom{R2}{I2-a-c,U2-b-d,R2-I2-U2+a+b+c+d}}$$

and

$$SU_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} b\binom{S2}{a,b,S2-a-b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c,d,T2-c-d}\binom{R2}{I2-a-c,U2-b-d,R2-I2-U2+a+b+c+d}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a,b,S2-a-b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c,d,T2-c-d}\binom{R2}{I2-a-c,U2-b-d,R2-I2-U2+a+b+c+d}},$$

since $Trans2 = I2$ and $TransU2 = U2$.

In this form these terms are intractable but they can be simplified to give tractable terms. Firstly we note that these multinomial coefficients can be rewritten as the product of two binomial coefficients, for instance

$$\begin{aligned}\binom{S2}{a,b,S2-a-b} &= \frac{S2!}{a!b!(S2-a-b)!} \\ &= \frac{S2!(S2-a)!}{a!b!(S2-a-b)!(S2-a)!} \\ &= \frac{S2!}{a!(S2-a)!}\times\frac{(S2-a)!}{b!(S2-a-b)!} \\ &= \binom{S2}{a}\binom{S2-a}{b}.\end{aligned}$$

Similarly we find

$$\binom{T2}{c,d,T2-c-d} = \binom{T2}{c}\binom{T2-c}{d}$$

and

$$\binom{R2}{I2-a-c, U2-b-d, R2-I2-U2+a+b+c+d} = \binom{R2}{I2-a-c}\binom{R2-I2+a+c}{U2-b-d}.$$

Using these binomial coefficient representations SI_{new} and SU_{new} become

$$SI_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} a\binom{S2}{a}\binom{S2-a}{b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c}\binom{T2-c}{d}\binom{R2}{I2-a-c}\binom{R2-I2+a+c}{U2-b-d}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a}\binom{S2-a}{b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c}\binom{T2-c}{d}\binom{R2}{I2-a-c}\binom{R2-I2+a+c}{U2-b-d}}$$

and

$$SU_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} b\binom{S2}{a}\binom{S2-a}{b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c}\binom{T2-c}{d}\binom{R2}{I2-a-c}\binom{R2-I2+a+c}{U2-b-d}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a}\binom{S2-a}{b}\sum_{c=0}^{I2-a}\sum_{d=0}^{U2-b}\binom{T2}{c}\binom{T2-c}{d}\binom{R2}{I2-a-c}\binom{R2-I2+a+c}{U2-b-d}}.$$

Using Vandermonde's convolution [35],

$$\sum_{k}\binom{j}{m+k}\binom{s}{i-k} = \binom{j+s}{m+i},$$

these terms become

$$SI_{new} = \frac{S2_t I2_t}{N2_t}$$

and

$$SU_{new} = \frac{S2_t U2_t}{N2_t},$$

where $N2_t = S2_t + I2_t + U2_t + R2_t = S2_t + T2_t + R2_t$. By substituting for

SI_{new} and SU_{new} the MFE for this model, (7.1), become

$$\begin{aligned}
S_{t+1} &= (1-p_d)S_t - p_i\frac{S_tI_t}{N_t} - \alpha p_i\frac{S_tU_t}{N_t} + N_t(p_{b_0} - kN_t) \ , \\
I_{t+1} &= (1-p_d-p_{dd}-p_r)I_t + p_i(1-p_s)\frac{S_tI_t}{N_t} + \alpha p_i(1-p_s)\frac{S_tU_t}{N_t} \ , \\
U_{t+1} &= (1-p_d-p_{dd}-p_r)U_t + p_ip_s\frac{S_tI_t}{N_t} + \alpha p_ip_s\frac{S_tU_t}{N_t} \ , \\
R_{t+1} &= (1-p_d)R_t + p_r(I_t+U_t) \ . \qquad (7.2)
\end{aligned}$$

To allow comparison to subsequent models, which have different forms of communication, we consider the one stage behaviour of a small population consisting of three $S2$, one $I2$ (consisting of one $T2$ and one $Trans2$), and one $U2$ (consisting of one $T2$ and one $TransU2$) over the communicating stage. Since this is the one stage behaviour over the communicating stage it depends solely on the numbers of agents of different types present and not on any of the probabilities in the model. We find that the mean numbers of susceptibles that communicate in this case are $SI_{new} = 0.6$ and $SU_{new} = 0.6$. Note that the total mean number of susceptible individuals that make contact with infected or superspreader individuals is $SI_{new} + SU_{new} = 0.6 + 0.6 = 1.2$, which is the same as we would have if we were considering a standard SIR model (such as that described in Fig. 3.1) with a population consisting three susceptible and two infected agents. Subsequent models lead to different terms for the communication stage so we will carry out this calculation to allow us to compare the models. Here we have the same mean number of contacts as we would expect without superspreaders: however, after the subsequent stage we would expect to find an increased number of new infections, for the same value of p_i . It would however be possible to choose a different value of p_i that would give the same overall mean behaviour. For instance considering a superspreader with $p_iu = \alpha p\_i$, p_s

will be the proportion of infected individuals that are superspreaders and it would be possible to design a model without superspreaders with probability of infection $p_i' = (1-p_s)p_i + \alpha p_i p_s$ that would have the same mean behaviour.

Non-prioritised communication

Norman and Shankland [73] demonstrated that for their models the choice of whether to use prioritised communication or not does not affect the mean behaviour. In Chapter 5 we demonstrated that this is not always true and here we investigate this question for models of the form of Fig. 7.1. The model in Fig. 7.2 differs from Fig. 7.1 only by the decision to use non-prioritised communication, with the $S2$ and $R2$ agents able to perform the actions *infect*, *infectU* or $\surd$ with equal weight. The mean numbers of these agents that perform the respective actions once again depend on the mix of different agents in the population. As in previous non-prioritised models we do not use parallel infected agents since infectious agents have the option not to communicate. In the models of Norman and Shankland this option not to communicate has the same effect numerically as explicitly modelling the situation where an infected individual communicates with another infected individual in the prioritised communication case.

This model once more leads to the equations (7.1) and again we cannot automatically derive expressions for SI_{new} and SU_{new} . However, by once again considering this specific model we can obtain expressions for the mean numbers of $S2$ agents that communicate with $I2$ and $U2$ agents. For this model we note that the weight with which any population change happens is

$$\binom{S2}{S2_{infect}, S2_{infectU}, S2_t}\binom{R2}{R2_{infect}, R2_{infectU}, R2_t}\binom{I2}{I2_{infect}}\binom{U2}{U2_{infect}},$$

$$
\begin{aligned}
p_b &\overset{\text{prob}}{=} p_b 0 - k * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor U1 \rfloor + \lfloor R1 \rfloor) \\
\\
S1 &\overset{\text{def}}{=} p_b.\surd : RepS + (1-p_b).\surd : S2 \\
I1 &\overset{\text{def}}{=} p_b.\surd : RepI + (1-p_b).\surd : I2 \\
U1 &\overset{\text{def}}{=} p_b.\surd : RepU + (1-p_b).\surd : U2 \\
R1 &\overset{\text{def}}{=} p_b.\surd : RepR + (1-p_b).\surd : R2 \\
\\
RepS &\overset{\text{def}}{=} S2 \times B2 \\
RepI &\overset{\text{def}}{=} I2 \times B2 \\
RepU &\overset{\text{def}}{=} U2 \times B2 \\
RepR &\overset{\text{def}}{=} R2 \times B2 \\
S2 &\overset{\text{def}}{=} 1.infect : SI3 + 1.infectU : SU3 + 1.\surd : S3 \\
I2 &\overset{\text{def}}{=} 1.\overline{infect} : I3 + 1.\surd : I3 \\
U2 &\overset{\text{def}}{=} 1.\overline{infectU} : U3 + 1.\surd : U3 \\
R2 &\overset{\text{def}}{=} 1.infect : R3 + 1.infectU : R3 + 1.\surd : R3 \\
B2 &\overset{\text{def}}{=} 1.\surd : B3 \\
\\
S3 &\overset{\text{def}}{=} (1-p_d).\surd : S1 + p_d.\surd : 0 \\
SI3 &\overset{\text{def}}{=} (p_i * (1-p_s)).\surd : I1 + (p_i * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i).\surd : S1 + p_d.\surd : 0 \\
SU3 &\overset{\text{def}}{=} (p_i u * (1-p_s)).\surd : I1 + (p_i u * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i u).\surd : S1 + p_d.\surd : 0 \\
I3 &\overset{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : I1 + (p_d+p_{dd}).\surd : 0 \\
U3 &\overset{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : U1 + (p_d+p_{dd}).\surd : 0 \\
R3 &\overset{\text{def}}{=} p_d.\surd : 0 + (1-p_d).\surd : R1 \\
B3 &\overset{\text{def}}{=} 1.\surd : S1 \\
\\
Popn &\overset{\text{def}}{=} S1\{s\} \times I1\{i\} \times U1\{u\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 7.2: Supershedder model with density dependent probability of giving birth - contact in one stage - non-prioritised communication

which considers the numbers of $I2$ and $U2$ that communicate, since they can choose not to, as well as the numbers of $S2$ and $R2$ that perform the various actions. The multinomial coefficients can once again be replaced by the product of two binomial coefficients and the weight with which population changes happen is

$$\binom{S2}{S2_{infect}}\binom{S2-S2_{infect}}{S2_{infectU}}\binom{R2}{R2_{infect}}\binom{R2-R2_{infect}}{S2_{infectU}}\binom{I2}{I2_{infect}}\binom{U2}{U2_{infect}}.$$

By considering the mean of all possible outcomes for the system we find

$$SI_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} a\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\sum_{d=b}^{U2}\binom{I2}{c}\binom{U2}{d}\binom{R2}{c-a}\binom{R2-c+a}{d-b}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\sum_{d=b}^{U2}\binom{I2}{c}\binom{U2}{d}\binom{R2}{c-a}\binom{R2-c+a}{d-b}}$$

and

$$SU_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} b\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\sum_{d=b}^{U2}\binom{I2}{c}\binom{U2}{d}\binom{R2}{c-a}\binom{R2-c+a}{d-b}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\sum_{d=b}^{U2}\binom{I2}{c}\binom{U2}{d}\binom{R2}{c-a}\binom{R2-c+a}{d-b}}.$$

The term

$$\sum_{c=a}^{I2}\sum_{d=b}^{U2}\binom{I2}{c}\binom{U2}{d}\binom{R2}{c-a}\binom{R2-c+a}{d-b},$$

which appears in the numerator and denominator of both terms, can be simplified to

$$\sum_{c=a}^{I2}\binom{I2}{c}\binom{R2}{c-a}\binom{R2+U2-c+a}{R2-c+a+b}$$

by noting that

$$\binom{R2-c+a}{d-b} = \binom{R2-c+a}{R2-c+a-d+b}.$$

This gives us the terms

$$SI_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} a\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\binom{I2}{c}\binom{R2}{c-a}\binom{R2+U2-c+a}{R2-c+a+b}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\binom{I2}{c}\binom{R2}{c-a}\binom{R2+U2-c+a}{R2-c+a+b}}$$

and

$$SU_{new} = \frac{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a} b\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\binom{I2}{c}\binom{R2}{c-a}\binom{R2+U2-c+a}{R2-c+a+b}}{\sum_{a=0}^{S2}\sum_{b=0}^{S2-a}\binom{S2}{a}\binom{S2-a}{b}\sum_{c=a}^{I2}\binom{I2}{c}\binom{R2}{c-a}\binom{R2+U2-c+a}{R2-c+a+b}} ,$$

which cannot be simplified any further.

As for the previous model, we consider the mean one stage behaviour for a population consisting of three $S2$, one $I2$ and one $U2$ and find the mean numbers of $S2$ agents that communicate with $I2$ and $U2$ are $SI_{new} = 0.692$ and $SU_{new} = 0.692$. Here the total number of infectious contacts (either with standard infecteds or superspreaders) is 1.384, which is more than we find in a standard SIR model with three susceptible and two infected agents. The point of the supershedders is that they are somehow more infectious but here we also find that the model featuring supershedders leads to more infectious contacts. This is not the behaviour that we wish to capture in the model so that this form of communication is not suitable for describing biological systems featuring superspreaders. This problem arises because of the form of the $S2$ and $R2$ agents. In the models of Norman and Shankland [73] the susceptible individuals were of the form

$$S2 \stackrel{\text{def}}{=} 1.\mathit{infect} : I1 + 1.\surd : S1 ,$$

which means that agents are weighted 1 to communicate and 1 to perform the free action $\surd$. However, in Fig. 7.2 the agents are effectively weighted 2 to communicate (1 to perform *infect* and 1 to perform *infectU*) and 1 to

perform the free action $\surd$. This makes the agents more likely to communicate, although the mean numbers that do so are still dependent on the mix of individuals present in the system.

Future work The models in Figs. 7.1 and 7.2 raise interesting questions for future work, namely

- what would the general terms be for the mean outcome of communication of this sort and can they be simplified?
- what effect do weights other than 1 have on communication and what would the general terms be if we did not impose the restriction on the weights of communicating agents?

By answering these questions it would be possible to extend the scope of the algorithm described in Chapter 3.

7.1.2 Consecutive contact stages

In Chapter 5 we demonstrated that it was possible to communicate on successive stages and make the same mean number of contacts as we find by communicating in parallel. Here we investigate the use of this form of communication to implement a model of superspreaders with the standard infected agents (I) communicating on the first stage and superspreaders communicating on the second stage. This choice of order does not affect the overall mean number of contacts that are made. This form of communication is interesting because we can obtain MFE directly using our algorithm.

Prioritised communication

The model in Figs. 7.3 and 7.4 has contact by infected individuals and by superspreaders occurring in different stages of the model. At the second

$$
\begin{aligned}
p_b &\overset{\text{prob}}{=} p_b 0 - k * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor U1 \rfloor + \lfloor R1 \rfloor) \\
\\
S1 &\overset{\text{def}}{=} p_b.\surd : RepS + (1 - p_b).\surd : S2 \\
I1 &\overset{\text{def}}{=} p_b.\surd : RepI + (1 - p_b).\surd : I2 \\
U1 &\overset{\text{def}}{=} p_b.\surd : RepU + (1 - p_b).\surd : U2 \\
R1 &\overset{\text{def}}{=} p_b.\surd : RepR + (1 - p_b).\surd : R2 \\
\\
RepS &\overset{\text{def}}{=} S2 \times B2 \\
RepI &\overset{\text{def}}{=} I2 \times B2 \\
RepU &\overset{\text{def}}{=} U2 \times B2 \\
RepR &\overset{\text{def}}{=} R2 \times B2 \\
S2 &\overset{\text{def}}{=} \omega.\mathit{infect} : SI3 + 1.\surd : S3 \\
I2 &\overset{\text{def}}{=} T2 \times Trans2 \\
Trans2 &\overset{\text{def}}{=} \omega.\overline{\mathit{infect}} : I3 + 1.\surd : I3 \\
T2 &\overset{\text{def}}{=} \omega.\mathit{infect} : 0 + 1.\surd : 03 \\
U2 &\overset{\text{def}}{=} T2 \times TransU2 \\
TransU2 &\overset{\text{def}}{=} 1.\surd : TransU3 \\
R2 &\overset{\text{def}}{=} \omega.\mathit{infect} : R3c + 1.\surd : R3 \\
B2 &\overset{\text{def}}{=} 1.\surd : B3
\end{aligned}
$$

Figure 7.3: Supershedder model with density dependent probability of giving birth, Part 1

$$
\begin{aligned}
S3 &\stackrel{\text{def}}{=} \omega.infectU : SU4 + 1.\surd : S4 \\
SI3 &\stackrel{\text{def}}{=} 1.\surd : SI4 \\
I3 &\stackrel{\text{def}}{=} 1.\surd : I4 \\
TransU3 &\stackrel{\text{def}}{=} \omega.\overline{infectU} : U4 + 1.\surd : U4 \\
T3 &\stackrel{\text{def}}{=} \omega.infectU : 0 + 1.\surd : 0 \\
R3 &\stackrel{\text{def}}{=} \omega.infectU : R4 + 1.\surd : R4 \\
R3c &\stackrel{\text{def}}{=} 1.\surd : R4 \\
B3 &\stackrel{\text{def}}{=} 1.\surd : B4 \\
\\
S4 &\stackrel{\text{def}}{=} (1-p_d).\surd : S1 + p_d.\surd : 0 \\
SI4 &\stackrel{\text{def}}{=} (p_i * (1-p_s)).\surd : I1 + (p_i * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i).\surd : S1 + p_d.\surd : 0 \\
SU4 &\stackrel{\text{def}}{=} (p_i u * (1-p_s)).\surd : I1 + (p_i u * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i u).\surd : S1 + p_d.\surd : 0 \\
I4 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : I1 + (p_d+p_{dd}).\surd : 0 \\
U4 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : U1 + (p_d+p_{dd}).\surd : 0 \\
R4 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1-p_d).\surd : R1 \\
B4 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
\\
Popn &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 7.4: Supershedder model with density dependent probability of giving birth, Part 2

stage in the model susceptible, infected, superspreader and recovered individuals are able to be contacted by the infected individuals and at the third stage those which have not already been contacted are able to be contacted by the superspreaders.

The fourth stage in the model involves individuals making probabilistic choices. Susceptible individuals that have not been contacted by an infected individual or a superspreader, and recovered individuals, die (with probability p_d) or survive (with probability $1 - p_d$). The susceptible individuals that have been contacted by an infected individual die (with probability p_d), become infected (with probability $p_i(1-p_s)$), become a superspreader (with probability $p_i \times p_s$) or remain susceptible (with probability $1-p_d-p_i$). Similarly susceptible individuals that have been contacted by a superspreader die (with probability p_d), become infected (with probability $p_{iu}(1 - p_s)$), become a superspreader (with probability $p_{iu} \times p_s$) or remain susceptible (with probability $1-p_d-p_{iu}$). Infected and superspreader individuals recover (with probability p_r), die (with probability p_d+p_{dd}, p_{dd} being the probability of dying due to the disease) or survive (with probability $1 - p_d - p_{dd} - p_r$).

The mean behaviour of this model is represented by the following system of equations:

$$\begin{aligned}
S_{t+1} &= (1 - p_d)S_t - \frac{S_t(p_i I_t + \alpha p_i U_t)}{N_t} + N_t(p_{b_0} - kN_t) \\
I_{t+1} &= (1 - p_d - p_{dd} - p_r)I_t + \frac{(1 - p_s)S_t(p_i I_t + \alpha p_i U_t)}{N_t}, \\
U_{t+1} &= (1 - p_d - p_{dd} - p_r)U_t + \frac{p_s S_t(p_i I_t + \alpha p_i U_t)}{S_t + I_t + U_t + R_t}, \\
R_{t+1} &= (1 - p_d)R_t + p_r(I_t + U_t). \qquad (7.3)
\end{aligned}$$

We can see that these equations take the same general form as (7.1), the equations for Figs. 7.1 and 7.2, but here we have the following expressions

for the numbers of susceptible individuals contacted,

$$SI_{new} = \frac{S_t I_t}{N_t} \qquad \text{and} \qquad SU_{new} = \frac{S_t U_t}{N_t} \ ,$$

which we can obtain because the form of communication used is covered by the algorithm of Chapter 3.

To compare the behaviour of this model to the behaviour of the previous models over the communication stages we once again consider the mean behaviour of a population consisting of three $S2$, one $I2$ and one $U2$. For the model in Fig. 7.3 communication happens over two stages and the mean behaviour can be found using the following two systems of one stage equations, which are produced by the algorithm while deriving (7.3):

$$\begin{aligned}
T2_t &= I2_t + U2_t \ , \\
SI3_{t+1} &= \frac{S2_t I2_t}{S2_t + T2_t + R2_t} \ , \\
S3_{t+1} &= S2_t - \frac{S2_t I2_t}{S2_t + T2_t + R2_t} \ , \\
I3_{t+1} &= I2_t \ , \\
T3_{t+1} &= T2_t - \frac{T2_t I2_t}{S2_t + T2_t + R2_t} \ , \\
TransU3_{t+1} &= U2_t \ , \\
R3c_{t+1} &= \frac{R2_t I2_t}{S2_t + T2_t + R2_t} \ , \\
R3_{t+1} &= R2_t - \frac{R2_t I2_t}{S2_t + T2_t + R2_t} \ ,
\end{aligned}$$

and

$$
\begin{aligned}
SU4_{t+1} &= \frac{S3_t TransU3_t}{S3_t + T3_t + R3_t} \ , \\
S4_{t+1} &= S3_t - \frac{S3_t TransU3_t}{S3_t + T3_t + R3_t} \ , \\
SI4_{t+1} &= SI3_t \ , \\
I4_{t+1} &= I3_t \ , \\
U4_{t+1} &= TransU3_t \ , \\
R4_{t+1} &= R3_t + R3c_t \ .
\end{aligned}
$$

Using these equations we find that the mean population after these two stages consists of 1.8 $S4$, 0.6 $SI4$, 0.6 $SU4$, 1 $I4$ and 1 $U4$. This is the same population that we found after the communication stage of Fig. 7.1 and therefore we have the same number of susceptibles contacted as we would for a standard SIR model.

Non-prioritised communication

The model in Figs. 7.5 and 7.6 replaces the prioritised communication in Fig. 7.3 and 7.4 with non-prioritised communication in the usual way. We once again wish to investigate whether the choice of non-prioritised communication

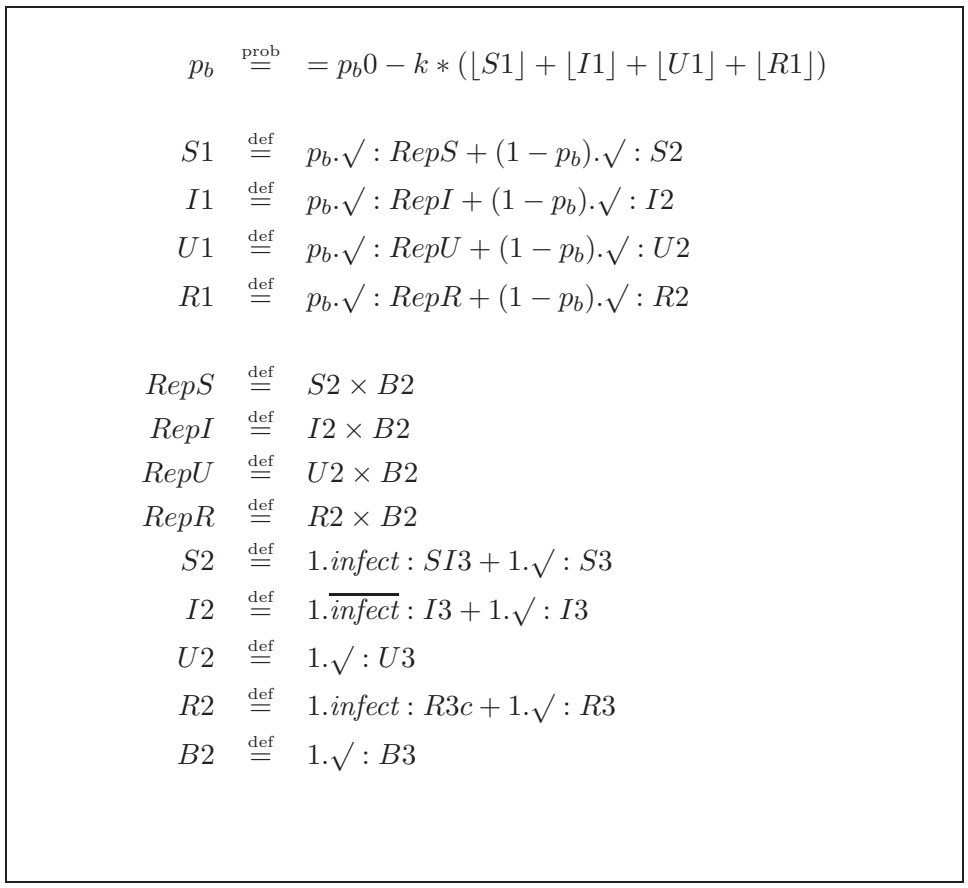

Figure 7.5: Supershedder model with density dependent probability of giving birth - non-prioritised communication, Part1

$$
\begin{aligned}
S3 &\stackrel{\text{def}}{=} 1.\mathit{infectU} : SU4 + 1.\surd : S4 \\
SI3 &\stackrel{\text{def}}{=} 1.\surd : SI4 \\
I3 &\stackrel{\text{def}}{=} 1.\surd : I4 \\
U3 &\stackrel{\text{def}}{=} 1.\overline{\mathit{infectU}} : U4 + 1.\surd : U4 \\
R3 &\stackrel{\text{def}}{=} 1.\mathit{infectU} : R4 + 1.\surd : R4 \\
R3c &\stackrel{\text{def}}{=} 1.\surd : R4 \\
B3 &\stackrel{\text{def}}{=} 1.\surd : B4 \\
\\
S4 &\stackrel{\text{def}}{=} (1-p_d).\surd : S1 + p_d.\surd : 0 \\
SI4 &\stackrel{\text{def}}{=} (p_i * (1-p_s)).\surd : I1 + (p_i * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i).\surd : S1 + p_d.\surd : 0 \\
SU4 &\stackrel{\text{def}}{=} (p_i u * (1-p_s)).\surd : I1 + (p_i u * p_s).\surd : U1 \\
&\quad +(1-p_d-p_i u).\surd : S1 + p_d.\surd : 0 \\
I4 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : I1 + (p_d+p_{dd}).\surd : 0 \\
U4 &\stackrel{\text{def}}{=} p_r.\surd : R1 + (1-p_d-p_{dd}-p_r).\surd : U1 + (p_d+p_{dd}).\surd : 0 \\
R4 &\stackrel{\text{def}}{=} p_d.\surd : 0 + (1-p_d).\surd : R1 \\
B4 &\stackrel{\text{def}}{=} 1.\surd : S1 \\
\\
\mathit{Popn} &\stackrel{\text{def}}{=} S1\{s\} \times I1\{i\} \times U1\{u\} \times R1\{r\} \lceil \{\surd\}
\end{aligned}
$$

Figure 7.6: Supershedder model with density dependent probability of giving birth - non-prioritised communication, Part 2

affects the mean behaviour of the system, which is done by obtaining the MFE:

$$\begin{aligned}
S_{t+1} &= (1-p_d)S_t + N_t(p_{b_0} - kN_t) - \\
&\quad S_t\left(\frac{p_i I_t}{S_t+I_t+R_t} + \frac{\alpha p_i U_t R_t}{(S_t+R_t)^2 + (S_t+I_t+R_t)U_t}\right) \\
I_{t+1} &= (1-p_d-p_{dd}-p_r)I_t + \\
&\quad (1-p_s)S_t\left(\frac{p_i I_t}{S_t+I_t+R_t} + \frac{\alpha p_i U_t R_t}{(S_t+R_t)^2 + (S_t+I_t+R_t)U_t}\right), \\
U_{t+1} &= (1-p_d-p_{dd}-p_r)U_t + \\
&\quad p_s S_t\left(\frac{p_i I_t}{S_t+I_t+R_t} + \frac{\alpha p_i U_t R_t}{(S_t+R_t)^2 + (S_t+I_t+R_t)U_t}\right), \\
R_{t+1} &= (1-p_d)R_t + p_r(I_t+U_t).
\end{aligned}$$

These differ from the MFE for the case where we used prioritised communication, (7.3), because we are once again not capturing the desired behaviour. The agents that can perform the output actions miss an opportunity to pass on the infection by performing $\surd$. For this model to have the same mean behaviour as Fig. 7.3 we need the choice to do $\surd$ to have the same effect as communicating with a $T2$ or $T3$ in Fig. 7.3. However the $I2$ and $U3$ agents should be able to interact with agents performing the output action on both communicative stages. This is not possible without introducing the possibility that they can be responsible for absorbing two infectious contacts. For example if $U2$ and $I3$ have the forms

$$\begin{aligned}
U2 &\stackrel{\text{def}}{=} 1.\mathit{infect} : U3 + 1.\surd : U3 \\
I3 &\stackrel{\text{def}}{=} 1.\mathit{infect} : I4 + 1.\surd : I4
\end{aligned}$$

the infected and superspreader individuals could explicitly communicate with an agent performing an output action as well as failing to perform the

output action, which in simple models [73] has the same numerical effect.

We again use the two sets of one stage equations to determine how a small population behaves over the communication stages:

$$\begin{aligned}
S3_{t+1} &= S2_t - \frac{S2_t I2_t}{S2_t + I2_t + R2_t} \ , \\
SI3_{t+1} &= \frac{S2_t I2_t}{S2_t + I2_t + R2_t} \ , \\
I3_{t+1} &= I2_t \ , \\
U3_{t+1} &= U2_t \ , \\
R3c_{t+1} &= \frac{R2_t I2_t}{S2_t + I2_t + R2_t} \ , \\
R3_{t+1} &= R2_t - \frac{R2_t I2_t}{S2_t + I2_t + R2_t} \ ,
\end{aligned}$$

and

$$\begin{aligned}
S4_{t+1} &= S3_t - \frac{S3_t U3_t}{S3_t + U3_t + R3_t} \ , \\
SI4_{t+1} &= SI3_t \ , \\
SU4_{t+1} &= \frac{S3_t U3_t}{S3_t + U3_t + R3_t} \ , \\
I4_{t+1} &= I3_t \ , \\
U4_{t+1} &= U3_t \ , \\
R4_{t+1} &= R3_t + R3c_t \ .
\end{aligned}$$

Considering again a population consisting of three $S2$, one $I2$ and one $U2$ we find that the mean population after communication consists of 1.5577 $S4$, 0.75 $SI4$, 0.6923 $SU4$, 1 $I4$ and 1 $U4$. This means that the total number of communicative contacts $(0.75 + 0.6923 = 1.4423)$ is greater than for any of the previous models and furthermore the standard infecteds have made more contacts on average than the superspreaders. This is not the desired

behaviour, since there are equal numbers of $I2$ and $U2$ and they should be equally likely to communicate. Also this behaviour would not be maintained if the order of infected and superspreader contact was reversed. Changing the order would lead to the numbers of $SI4$ and $SU4$ being switched.

7.1.3 Time series

We study the behaviour of the model from Fig. 7.3 by considering the time series of the MFE, (7.3), and also the time series of the mean of 1000 simulations of the model. This was done for a wide range of parameter values and in most cases the MFE fitted the mean of simulations well, as we have found for other models. However for this superspreader model we found that for some parameter values the two time series diverge and it is such a case we consider here. The graph in Fig. 7.7 considers the total number of infected individuals in the population $(I + U)$ for an initial population of $S1\{200\} \times I1\{40\} \times U1\{10\}$ with $p_s = 0.2$, $p_i = 0.02$, $\alpha = 16$, $p_r = 0.02$, $p_d = 0.01$, $p_{dd} = 0.005$, $p_{b_0} = 0.2$ and $k = 0.0008$. These values of p_s and α mean that on average we expect 20% of the infected individuals to be responsible for 80% of new infections, which has been proposed as the proportions that arise in superspreader systems [60, 94]. We see in Fig. 7.7 that the MFE and the simulations match well during the initial peak of the infection but over time the MFE settles to a steady state while the mean of the simulations gradually tends towards extinction of the disease.

In Section 7.1.1 we discussed the idea that it is possible to choose a different value of p_i such that the mean behaviour of a system without superspreaders will be the same as for the system with superspreaders. This is done for Fig. 7.3 by setting $\alpha = 1$, so that the U agents have the same behaviour as the I, and by setting $p_i = 0.08$ (to satisfy $0.8 \times 0.02 + 0.2 \times$

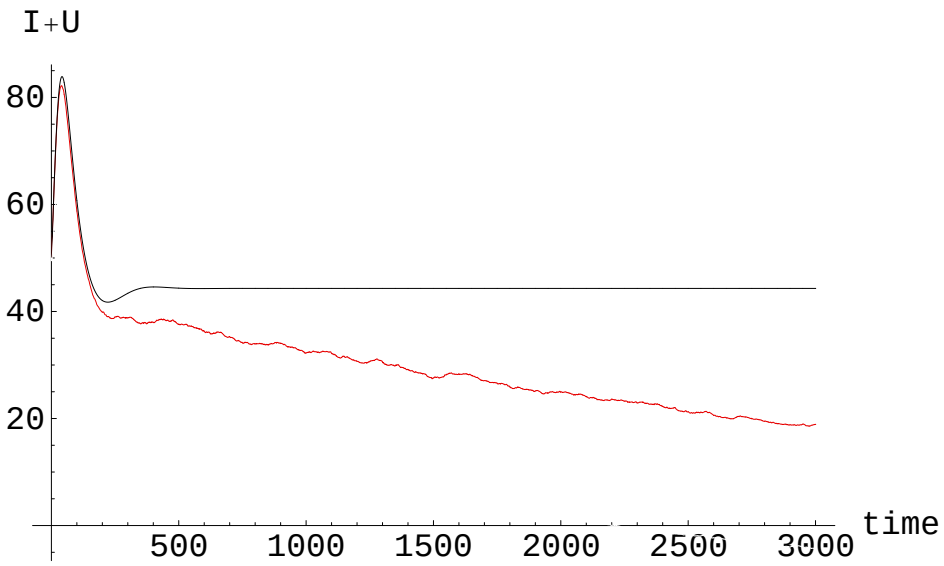

Figure 7.7: Total infecteds $(I+U)$ with supershedders: — MFE; Simulations — mean.

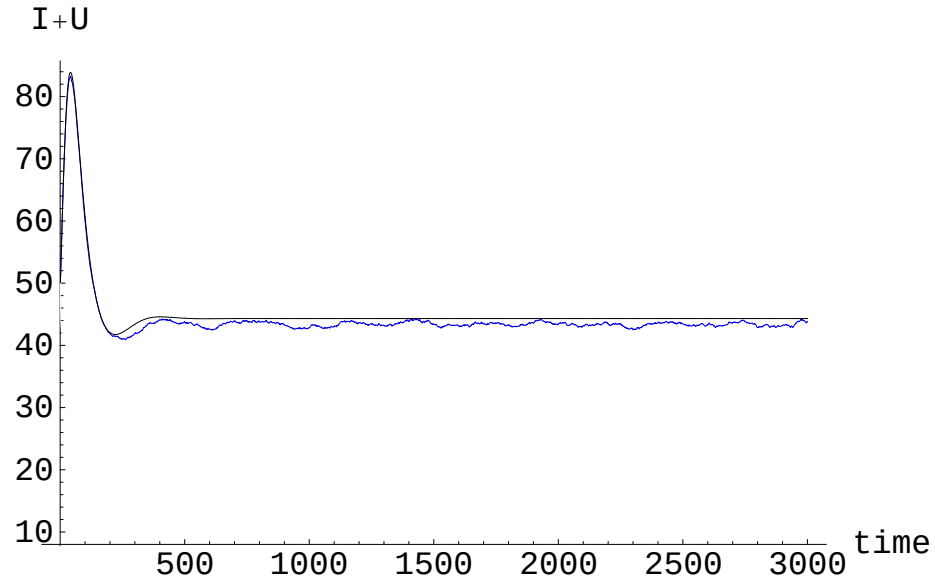

Figure 7.8: Total infecteds $(I+U)$ without supershedders, $p_i = 0.08$: — MFE; Simulations — mean.

$16 \times 0.02 = 0.08$) - all other parameter values are kept the same as for Fig. 7.7. This case is shown in Fig. 7.8 where we can see that the MFE have the same time series but the simulations now closely match the MFE for the entire duration being considered. This leads us to ask why the mean of the simulations should have such different behaviour with superspreaders present when the MFE have the same time series.

Matthews and Woolhouse [64] suggested that the presence of superspreaders could be expected to increase the variability of a system. We investigate this for our system by considering, in Fig. 7.9, the two cases (with or without superspreaders) in a single graph for a shorter period to examine the variability before the means diverge significantly. For the superspreader model we found that the distribution of our simulations was skewed so we plot the median and quartiles for both models. We can see that during the initial peak of infection the means remain close but at an early stage there is a difference between the quartiles and by the end of the period considered the model with superspreaders has significantly greater inter-quartile range and the median is markedly diverging from the MFE. If we consider a single simulation (Fig. 7.10) we find that the infection dies out, with the time at which the infection dies out varying stochastically between separate runs of the simulation. This explains the stochastic fade-out witnessed when we consider the mean of many simulations. Over time the number of simulations where $I_t + U_t = 0$ increases so that the mean of many simulations tends to 0. We do not find this behaviour in the case without superspreaders as the smaller variability means that individual simulations do not experience extinction of the disease.

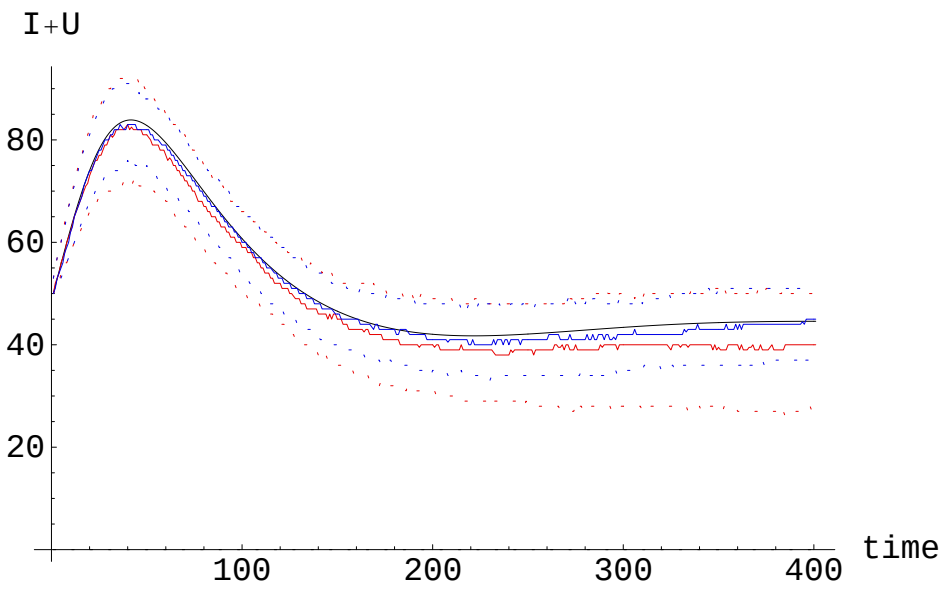

Figure 7.9: Total infection $(I+U)$: — MFE; Simulations with supershedders — median, ... upper and lower quartiles; Simulations without supershedders — median, ... upper and lower quartiles.

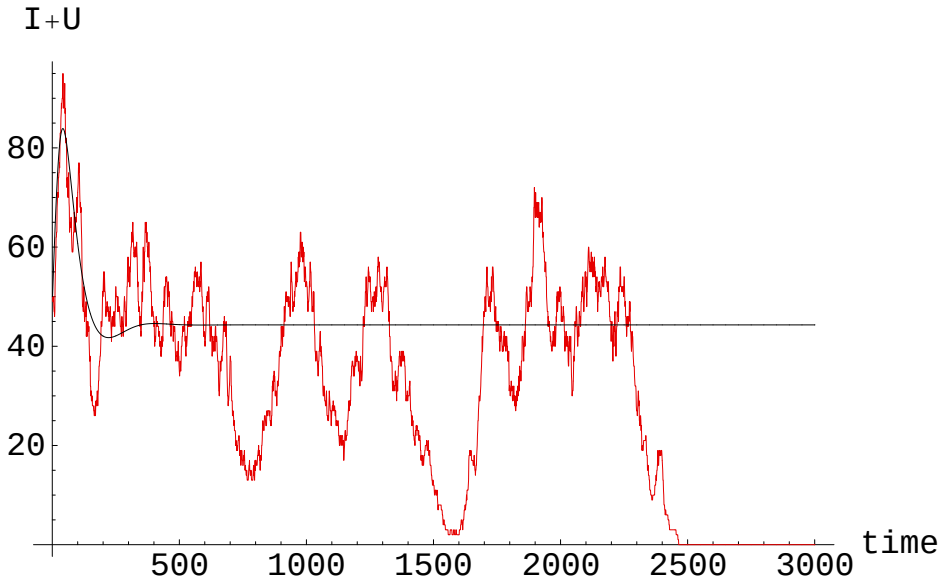

Figure 7.10: Total infecteds $(I+U)$ with supershedders - single simulation: — MFE; Simulation —.

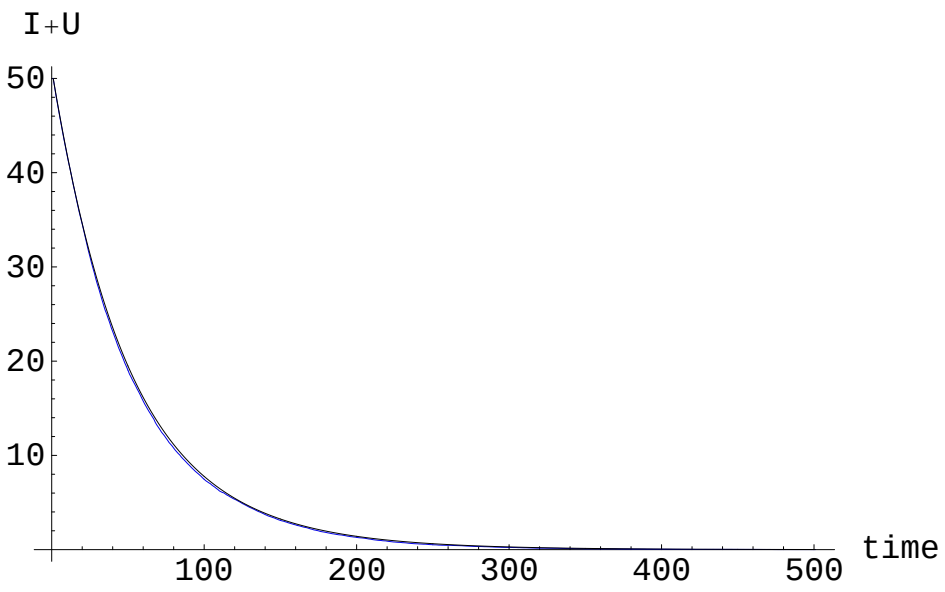

Figure 7.11: Total infecteds $(I + U)$ without supershedders, $p_i = 0.02$: — MFE; Simulations — mean.

It should also be noted that for the probability of infection used in the superspreader case ($p_i = 0.02$) the disease cannot persist without a portion of the infecteds being superspreaders that give a greater probability of infection. This is demonstrated in Fig. 7.11, which uses $p_i = 0.02, \alpha = 1$ and all other parameter values as before. In this case the disease dies without an initial epidemic occurring and the mean of the simulations match the MFE well over the entire period for which the disease survives.

7.2 Contact Superspreaders

The model in Fig. 7.12 features superspreaders that have a higher contact rate than the standard infected individuals. Density dependence in the population is introduced by the same mechanism as in the infectiousness models of section 7.1. The increased contact rate for the superspreaders

$$
\begin{array}{rcl}
p_b & \stackrel{\text{prob}}{=} & p_b 0 - k * (\lfloor S1 \rfloor + \lfloor I1 \rfloor + \lfloor U1 \rfloor + \lfloor R1 \rfloor) \\
 & & \\
S1 & \stackrel{\text{def}}{=} & p_b.\surd : RepS + (1-p_b).\surd : S2 \\
I1 & \stackrel{\text{def}}{=} & p_b.\surd : RepI + (1-p_b).\surd : I2 \\
U1 & \stackrel{\text{def}}{=} & p_b.\surd : RepU + (1-p_b).\surd : U2 \\
R1 & \stackrel{\text{def}}{=} & p_b.\surd : RepR + (1-p_b).\surd : R2 \\
 & & \\
RepS & \stackrel{\text{def}}{=} & S2 \times B2 \\
RepI & \stackrel{\text{def}}{=} & I2 \times B2 \\
RepU & \stackrel{\text{def}}{=} & U2 \times B2 \\
RepR & \stackrel{\text{def}}{=} & R2 \times B2 \\
S2 & \stackrel{\text{def}}{=} & \omega.infect : SI3 + 1.\surd : S3 \\
I2 & \stackrel{\text{def}}{=} & T2 \times Trans\{ci\} \\
Trans & \stackrel{\text{def}}{=} & \omega.\overline{infect} : 0 + 1.\surd : 0 \\
T2 & \stackrel{\text{def}}{=} & \omega.infect : I3 + 1.\surd : I3 \\
U2 & \stackrel{\text{def}}{=} & TU2 \times Trans\{cu\} \\
TU2 & \stackrel{\text{def}}{=} & \omega.infect : U3 + 1.\surd : U3 \\
R2 & \stackrel{\text{def}}{=} & \omega.infect : R3 + 1.\surd : R3 \\
B2 & \stackrel{\text{def}}{=} & 1.\surd : B3 \\
 & & \\
S3 & \stackrel{\text{def}}{=} & (1-p_d).\surd : S1 + p_d.\surd : 0 \\
SI3 & \stackrel{\text{def}}{=} & (p_i * (1-p_s)).\surd : I1 + (p_i * p_s).\surd : U1 \\
 & & +(1-p_i-p_d).\surd : S1 + p_d.\surd : 0 \\
I3 & \stackrel{\text{def}}{=} & p_r.\surd : R1 + (1-p_r-p_d-p_{dd}).\surd : I1 + (p_d+p_{dd}).\surd : 0 \\
U3 & \stackrel{\text{def}}{=} & p_r.\surd : R1 + (1-p_r-p_d-p_{dd}).\surd : U1 + (p_d+p_{dd}).\surd : 0 \\
R3 & \stackrel{\text{def}}{=} & (1-p_d).\surd : R1 + p_d.\surd : 0 \\
B3 & \stackrel{\text{def}}{=} & 1.\surd : S1 \\
 & & \\
Popn & \stackrel{\text{def}}{=} & S1\{s\} \times I1\{i\} \times R1\{r\} \lceil \{\surd\}
\end{array}
$$

Figure 7.12: Contact superspreader model with density dependent probability of giving birth

is achieved by having the superspreader parallel agent ($U2$) feature more $Trans$ agents (which can pass on the infection) than the infected parallel agent ($I2$) i.e. $c_u = \alpha c_i$ with $\alpha > 1$.

The mean behaviour of this model is given by the following mean field equations (which are derived directly by our algorithm):

$$\begin{aligned}
S_{t+1} &= (1-p_d)S_t - \min\left[p_i S_t, \frac{p_i S_t(c_i I_t + \alpha c_i U_t)}{N_t}\right] \\
&\quad + N_t(p_{b_0} - kN_t), \\
I_{t+1} &= (1-p_d-p_{dd}-p_r)I_t + (1-p_s)\min\left[p_i S_t, \frac{p_i S_t(c_i I_t + \alpha c_i U_t)}{N_t}\right], \\
U_{t+1} &= (1-p_d-p_{dd}-p_r)U_t + p_s \min\left[p_i S_t, \frac{p_i S_t(c_i I_t + \alpha c_i U_t)}{N_t}\right], \\
R_{t+1} &= (1-p_d)R_t + p_r(I_t + U_t). \qquad (7.4)
\end{aligned}$$

This system of equations can be analysed by considering the two options within the minimum term. This means that the behaviour of the model is described by either

$$\begin{aligned}
S_{t+1} &= (1-p_d)S_t - p_i S_t \ , \\
&\quad + N_t(p_{b_0} - kN_t) \ , \\
I_{t+1} &= (1-p_d-p_{dd}-p_r)I_t + (1-p_s)p_i S_t \ , \\
U_{t+1} &= (1-p_d-p_{dd}-p_r)U_t + p_s p_i S_t \ , \\
R_{t+1} &= (1-p_d)R_t + p_r(I_t + U_t) \ , \qquad (7.5)
\end{aligned}$$

or

$$\begin{aligned}
S_{t+1} &= (1-p_d)S_t - \frac{p_i S_t(c_i I_t + \alpha c_i U_t)}{N_t} \,, \\
&\quad + N_t(p_{b_0} - kN_t) \,, \\
I_{t+1} &= (1-p_d-p_{dd}-p_r)I_t + \frac{(1-p_s)p_i S_t(c_i I_t + \alpha c_i U_t)}{N_t} \,, \\
U_{t+1} &= (1-p_d-p_{dd}-p_r)U_t + \frac{p_s p_i S_t(c_i I_t + \alpha c_i U_t)}{N_t} \,, \\
R_{t+1} &= (1-p_d)R_t + p_r(I_t + U_t) \,. \qquad (7.6)
\end{aligned}$$

We are most interested in (7.6) since this applies in general, with (7.5) applying only when much of the population is already infected. With $c_i = 1$ (7.6) is the same as the system of equations for the supershedder model given in (7.3) and where $c_i > 1$ we have (7.3) with p_i rescaled by a factor of c_i. This means that for the contact superspreaders model we have the same mean behaviour as for supershedders, except when a large proportion of the population is already infected.

7.3 Summary

In this chapter we have presented several models of disease spread featuring superspreaders. In Section 7.1 we considered models where contact with a superspreader made the susceptible individuals more likely to go on to become infected, than if contact was with a standard infected individual. To do this, the superspreader and standard infected agents perform different actions with the susceptible agents going into a different state depending on the action performed, with differing probabilities of becoming infected. This can be implemented in two ways in WSCCS: either the agents performing the input actions can make the choice between two communicative actions

in a single stage or contact with standard infecteds and superspreaders can happen on consecutive stages.

We found that for both forms of contact non-prioritised communication led to a greater mean number of contacts by the susceptibles than for a model without superspreaders. This is not the behaviour we wish to capture, since the superspreaders should only have the effect of making infection more likely after contact than the standard infecteds.

For the consecutive contact, prioritised communication model we studied the time series of the resulting MFE and also the time series arising from the mean of many simulations of the model. We found that the variability of the stochastic simulations is greater when superspreaders are present than for a corresponding model with the same mean behaviour but no superspreaders. This agrees with result found by Mathews and Woolhouse [64] who found that superspreaders had the effect of increasing the variability in the outcomes of the system. For the particular parameter values being considered we found that this increased variability had the effect of allowing the mean infection in the simulations to die out. Meanwhile the MFE settle to a steady state that features fixed numbers of standard infected and superspreader individuals. This is an example of a model for which the MFE do not always offer a good approximation to the long term average behaviour of the model, although the MFE are a very good approximation to the simulations during the crucial early epidemic phase. The divergence of the MFE from the mean of the simulations is caused by the increased variability introduced by the superspreaders, although for many choices of the parameter values the MFE still offer a good approximation to the mean of the system. So far we have found the variability by calculating the standard deviation, or quartiles, of a large number of simulations, which is computationally ex-

pensive. A preferable approach, which could be addressed in future work, would derive equations for the variability in the system directly from the WSCCS syntax. This would allow us to see how variability changes with key parameters such as α.

In Section 7.2 we presented a model for a system featuring contact superspreaders. This behaviour was captured by having both types of infected individual represented by parallel agents with the superspreaders including more agents that perform the output action. The model led to MFE that feature a min term but for cases where the majority of the population has not been infected the MFE match those for the infectiousness superspreader models.

In Section 7.1 we found that models featuring prioritised communication were able to capture the desired behaviour while those featuring non-prioritised communication were not. This leads us to conclude that prioritised communication is preferable since it can more accurately model the desired systems. We further found that the same mean population level behaviour was found for prioritised models featuring either a single contact stage (with agents able to choose between two input actions) or consecutive contact stages (with different actions performed on different stages). The consecutive contact model has the advantage that MFE can be derived directly using our algorithm while the single contact stage model offers a more intuitive way of capturing the desired behaviour. At present there is a trade off between the simplicity of deriving equations automatically and more intuitively describing the system. This could be overcome by expanding the algorithm to cover models that feature communication of the form used in Figs. 7.1.

Chapter 8

Conclusions & Future Work

8.1 Conclusions

In this thesis we have investigated the use of process algebra as a tool to model and analyse the spread of infectious disease. We have presented an algorithm to formalise the process of deriving MFE from process algebra models. This algorithm relates the behaviour of the population to the behaviour of, and interaction between, the individuals that make up the population. We went on to develop models to address specific biological questions and using this algorithm we derived MFE that describe the average behaviour of the system. The questions considered were:

- Changing scale - how can we rigorously move from individual level to population level description of a system?
- Population growth - what individual level behaviours lead to different equations?
- Density dependent transmission - can individual behaviour be defined that leads to density dependent transmission MFE?

- Indirect transmission - how can we capture indirect transmission in individual level models?
- Superspreaders - what effect do superspreaders have on the variability of a system?

8.1.1 Modelling disease in WSCCS

Suitability of WSCCS

WSCCS has previously been used to produce basic disease models [73, 82] for which MFE could be obtained. The models presented here have sought to introduce greater biological realism and are therefore more complicated than the previous models. The attraction of individual-based modelling is that populations can most easily be studied in terms of individual behaviour and interactions. In disease systems these individual interactions are fundamentally important to the spread of the disease.

The main advantage of WSCCS is that we have a formal framework for describing individual behaviour and a range of analytical techniques with which to study the resulting populations, most notably by deriving MFE - see below. The formal nature of WSCCS means that we can obtain equations describing the population level behaviour and have confidence that they arise from the defined individual level behaviour.

In the course of our work we have found several restrictions on how best to describe disease systems in WSCCS. For instance Norman and Shankland found that the choice to use prioritised or non-prioritised communication did not affect the overall behaviour of their system, or the resulting MFE. However, in Chapters 5 and 7 we found that this is not always true and prioritised communication offers the most intuitive relationship between

individual-level and population-level behaviour. This may seem undesirable, since priority can be thought of as "forcing" disease transmission, however many diseases force infected individuals to pass on the disease. For instance a fox with rabies becomes aggressive and tends to fight, thereby potentially transmitting the disease [5].

Another important factor that affects models is the order in which different behaviours happen. We may naively assume that changing the order of the stages in the model will have no effect on the overall behaviour described by the MFE since all of the same behaviour will happen within one timestep of the MFE. However, in Chapter 3 we found that merely switching the order of stages can have a significant effect on the MFE. This is because switching the order changes the underlying biological assumptions of the model. We must therefore consider carefully the biological implications of the order in which we choose to present different behaviours in the model.

Advantages of MFE

By developing an algorithm to derive MFE rigorously from the WSCCS description of the model we have addressed the state explosion problem of process algebra. The MFE offer a simple way to produce the time series for the mean behaviour of the system. Traditional process algebra techniques - simulations or Markov chain analysis - are computationally expensive for large systems or are restricted in the size of the system that can be investigated.

In addition the MFE are analogous to the traditional mathematical equations used to model biological systems. A wide range of analyses are available for these mathematical models and the MFE that arise from our algorithm are amenable to these analyses. By developing mathematical models in this

way we can be sure that the population level equations are a direct consequence of the individual behaviour described in the WSCCS model. This is in contrast to the traditional approach to developing mathematical models in which assumptions are made about the population level behaviour, although it is the individual level behaviour that can most easily be observed.

In general we have seen that the MFE offer a very good approximation to the mean of the system but there are limitations. By relating our algorithm to the proof offered by Kurtz [58] we see that we would expect the MFE to match exactly the mean of the system in the limiting case where the system being considered is infinitely large. In general the MFE offer a better approximation as the size of the system increases but most important is the number of infected individuals in the initial population. By increasing the initial number of infected individuals, with a complementary reduction in the number of susceptible individuals so that the overall population size is unchanged, we find the MFE more closely match the mean behaviour of the system, which matches the results of West and Thompson [93] who found that changing the initial number of infected individuals had the greatest effect on the convergence of their stochastic and deterministic models.

An important question related to the derivation of MFE is whether it is possible to define WSCCS behaviour that will lead to desired MFE. This was achieved in Chapters 4 and 5, however in Chapters 5 and 7 we found that multiple models could be defined that lead to the same MFE. In addition it is clear that WSCCS models cannot be defined that will lead directly to equations featuring exponential or logarithmic terms [40, 78].

8.1.2 Population growth

In Chapter 4 we developed models that sought to capture realistic population dynamics, in which there is a limit on the size of the population. Many different mathematical models of population dynamics exist and our models led to two of these. For models including explicit competition for resources, with separate agents representing food, the MFE that we found was the Beverton-Holt model [16]. Although this model has previously been proposed to describe population growth it is not widely used. One of the disadvantages of the Beverton-Holt model is an increased mathematical complexity, compared to the more commonly used logistic model [91], but the fact that it has arisen from our simple models that capture competition for resources suggests that it is a good candidate for describing population dynamics.

Models where competition was implicitly included, with the probabilities of either birth or death dependent on the population size, led to the logistic model [91]. This is in contrast to the findings of Brännström and Sumpter [19] who presented a site based individual level model of population dynamics, making assumptions that led to several different mathematical models but most notably they did not find the logistic model. While Brännström and Sumpter's range of assumptions were limited by the framework within which their models were developed, in our models we were free to choose density dependence of births or deaths to be in any form we wish. The linear proportionality we implemented effectively incorporates the population level assumptions on which the logistic equation is based, however it is still interesting that we found the logistic equation since Brännström and Sumpter tried, and failed, to do so.

8.1.3 Disease transmission

Density dependent vs frequency dependent

The question of whether the density dependent transmission term

$$\beta SI$$

or the frequency dependent term

$$\frac{\beta SI}{N}$$

is most appropriate for capturing disease transmission has been of interest [10]. The frequency dependent term most naturally arises from WSCCS models but in Chapter 5 we showed that it was possible to describe individual level behaviour that would lead to density dependent transmission. This is in contrast to the results of Turner et al. [90] who found that whatever individual level behaviour was described it could best be approximated at the population level by the density dependent transmission term. The difference is likely to be due to the different treatment of spatial information in the different models. Turner et al.'s cellular automata models were inherently spatial, which led to clustering of infection, while our WSCCS models assume a randomly mixed population. We are now in a position to look at transmission of specific diseases and investigate what transmission term should be used to describe them.

Indirect transmission

In Chapter 6 we developed models to capture indirect disease transmission that utilised separate WSCCS agents to represent the environment.

Different models featured either an unlimited number of environment agents, with infected agents probabilistically spawning infected environment, or a fixed number of environment agents, which become infected by coming into contact with infected individuals. Another variable in the model was the way in which infection decays within the environment. If infection persists for a fixed period of time then in the unlimited environment case the mean of the system can be represented by equations only for the groups in the population (S, I and R) with a delay in the transmission term. In all other cases it was necessary to consider equations to describe the environment as well as the population. In any environment containing a population the infected portions of environment will be small compared to the environment as a whole so it is reasonable to consider the quantity of environment that can become infected to be unlimited.

8.1.4 Spatial information

All of the models considered in this thesis assume random mixing of individuals. In Chapters 4 and 5 we found differences between our results and those of other individual-based modelling approaches [19, 90] that incorporate spatial information about individuals in the system. These other studies found counterintuitive results at the population level while we are able to describe individual behaviour that intuitively leads to desired population level behaviour. The major difference with our models is that we have removed spatial information although we could explicitly add space later.

8.1.5 Superspreader models

In Chapter 7 we presented models of systems featuring superspreaders. These models led to MFE that were similar to the model proposed by

Kemper [51] for a superspreader system. By performing simulations of the system we found that, for some parameter values, the MFE did not offer a good fit to the mean of the simulations: the MFE settle to a steady state while the simulations show the system displaying stochastic fade out of the disease. This was shown to be due to the increased variability introduced to the system by the presence of superspreaders. This result was predicted by Matthews and Woolhouse [64] who suggested that superspreaders could be expected to increase the variability of the system.

8.2 Future work

In this section we mention some potential directions to extend the work presented in this thesis.

8.2.1 Variability

Although we have demonstrated that MFE generally offer a very good approximation to the mean behaviour of the system, the average behaviour they describe only tells us part of the story: the variability of the system, and the range of possible outcomes, are also important. In Chapter 7 we saw that the presence of superspreaders leads to greater variability in the potential outcomes of the model. With large enough variability we saw that this could affect the suitability of the MFE as an approximation to the mean behaviour of the system. To calculate the standard deviation we must perform many simulations of the system, which is computationally expensive. This detracts from the advantage offered by MFE since we must still perform simulations of the system.

The underlying probabilistic rules of WSCCS would make it possible to derive equations, in a similar way as for the MFE, which describe the

standard deviation of the system over time. With such equations it would be possible to comment on whether stochastic fade out may occur for particular parameter values. For models that feature only probabilistic agents this is straightforward, drawing on well known results for the binomial distribution from probability theory: however, for models featuring communication it is not so clear how these equations would look, with the corresponding results for the hypergeometric distribution (which is the relevant distribution for probabilistic events of the form of communication in WSCCS) resulting in equations that do not match the standard deviation of the simulations.

8.2.2 ODEs from WSCCS

Although many discrete time difference equation models have been proposed to describe biological systems it is more common to use ordinary differential equations to model a system in continuous time. For this approach to be attractive to mathematical biologists as a method for developing models it would be advantageous if we could obtain ODEs. One way of obtaining ODEs for our models would be by thinking of our MFE as the result of applying Euler's method to a system of ODEs. This would mean, for instance, that we could consider the system of MFE that arose from our initial example model in Chapter 3, (3.13), to be the result of applying Euler's method to the following system of ODEs:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{p_i SI}{N} \,, \\ \frac{dI}{dt} &= -p_r I + \frac{p_i SI}{N} \,, \\ \frac{dR}{dt} &= p_r I \,. \end{aligned}$$

A preferable approach would derive ODEs directly from the WSCCS description of the system. As mentioned in Chapter 3 the MFE arise from an intermediate stage of Kurtz's [58] derivation of ODEs for Markov processes. It would therefore be possible to amend the algorithm presented in Chapter 3 so that we derive ODEs rather than discrete difference equations. By deriving such equations we could more directly compare the equations that arise from WSCCS models to existing mathematical models of biological systems.

8.2.3 Spatial information

All of the models presented in the previous chapters are based on the idea that the population is well mixed so that all individuals can interact with any other. This assumption is used widely when developing models of biological systems, however, in reality biological systems typically feature some sort of spatial heterogeneity. By incorporating spatial information into our models we could more realistically describe the behaviour of the system.

A natural first step to incorporating spatial heterogeneity to our models would be to consider a model of disease spread in metapopulations. This could be done by describing a system consisting of several subpopulations, each containing susceptibles, infecteds and recovereds (S_A, I_A, R_A; S_B, I_B, R_B;...; S_X, I_X, R_X). Interaction would happen only within these subpopulations (e.g. S_A become infected by interacting with I_A) and individuals can migrate between subpopulations probabilistically (e.g. I_A can become I_B with probability $p_{m_{ab}}$).

8.2.4 Extending the scope of the algorithm

In Chapter 7 we presented models for which we could not use our algorithm to derive MFE, although we did obtain MFE by carefully considering the system. In this case the mean behaviour captured by these MFE is the same as for another model for which we could obtain MFE using our algorithm, so the restrictions of the algorithm do not limit the system that can be described. Nevertheless we may wish in future to use features of WSCCS that are not currently allowed by the algorithm. Further work on how we capture the mean behaviour should make it possible to extend the algorithm to allow it to be used with a wider range of WSCCS models than is currently possible.

Appendix A

WSCCS

In this Appendix we summarise the formal syntax and semantics of WSCCS. The information given here is summarised from [88].

A.1 Syntax

A.1.1 Actions

Action names, $a \in Act$, are chosen from an arbitrary set and as such it is useful to choose action names that are suggestive of the system being described. The inverse of the action a (typically input) is $\overline{a}$ (typically output) and the identity action is denoted by $\surd$. When actions must occur in parallel we denote the multiplication by $\#$ such that $a\#\overline{a} = \surd$.

A.1.2 Relative frequency expressions (RFE)

RFE, e, are defined by the following syntax with x ranging over a set of variable names and c ranging over a fixed field (e.g. $\mathscr{N}$ or $\mathscr{R}$):

$$e ::= x|c|e + e|e \times e \ .$$

In these expressions we have commutative and associative multiplication and addition, with multiplication distributing over addition.

A.1.3 Weights

The set of WSCCS weights $\mathscr{W}$, denoted by w_i, are of the form $e\omega^k$, $e = e\omega^0$. In such weights e is the relative frequency with which this choice should be taken and k is the priority of this choice with ω an infinite object, $\omega > e \ \forall \ e$. The following multiplication and addition rules apply with $k \geq k'$:

$$\begin{aligned} e\omega^k + f\omega^{k'} &= e\omega^k = f\omega^{k'} + e\omega^k \ , \\ e\omega^k + f\omega^k &= (e+f)\omega^k = f\omega^k + e\omega^k \ , \\ e\omega^k * f\omega^{k'} &= (ef)\omega^{k+k'} = f\omega^{k'} * e\omega^k \ . \end{aligned}$$

A.1.4 Grammar

The possible WSCCS expressions are given by the following BNF grammar:

$$E ::= X | a : E | \Sigma\{wi.E_i | i \in I\} | E \times E | E\lceil A | \Theta(E) | E[S] | X \stackrel{\text{def}}{=} E \ .$$

Here $X \in Var$, a set of process variables; $a \in Act$, an action group; $w_i \in \mathscr{W}$, a set of weights; S is a set of renaming functions, $S : Act \rightarrow Act$ such that $S(\surd) = \surd$ and $\overline{S(a)} = S(\overline{a})$; action subsets $A \subseteq Act$ with $\surd \in A$; and arbitrary indexing sets I . The informal interpretation of the operators is as follows

- 0 a process that cannot proceed, representing deadlock;
- X the process bound to the variable X ;
- $a : E$ a process that can perform a becoming E ;

- $\Sigma\{w_i.E_i | i \in I\}$ the weighted choice between processes E_i , the weight of E_i being w_i . Considering a large number of repeated experiments of this process, we expect to see E_i chosen with relative frequency $w_i/\Sigma_{i\in I} w_i$. The binary plus operator can be used in place of the indexed sum i.e. writing $\Sigma\{1_1.a : 0, 2_2.b : 0 | i \in \{1, 2\}\}$ as $1.a + 2.b$;

- $E \times F$ the synchronous parallel composition of E and F . At each stage each process must perform an action with the composition performing the composition of the individual actions;

- $E \lceil A$ a process that can only perform actions in the group A . This operator is used to enforce communication on actions $b \notin A$;

- $\Theta(E)$ represents taking the prioritised parts of the process E only ;

- $E[S]$ represents E relabelled by the function S;

- $X \stackrel{\text{def}}{=} E$ represents binding the process variable X to the expression E .

A.2 Semantics

The semantics of WSCCS is transition based, defining the actions that a process can perform and the weight with which a state can be reached. The operational rules of WSCCS, presented in Table A.1, follow the informal description of the operators given above. In particular note the two different arrows that feature in the table: $\stackrel{a}{\rightarrow}$ represents a transition, associated with the action a ; and $\stackrel{w}{\longmapsto}$ represents a transition associated with a weight w .We may specify multiple ways to choose the same process with the same weight

$\overline{a{:}E \xrightarrow{a} E}$	$\overline{\sum\{w_i.E_i \mid i\in I\} \overset{w_i}{\longmapsto} E_i}$
$\dfrac{E \xrightarrow{a} E' \quad F \xrightarrow{b} F'}{E\times F \xrightarrow{a\#b} E'\times F'}$	$\dfrac{E \overset{w}{\longmapsto} E' \quad F \overset{v}{\longmapsto} F'}{E\times F \overset{wv}{\longmapsto} E'\times F'}$
$\dfrac{E \xrightarrow{a} E' \quad F \overset{w}{\longmapsto} F'}{E\times F \overset{w}{\longmapsto} E\times F'}$	$\dfrac{E \overset{w}{\longmapsto} E' \quad F \xrightarrow{a} F'}{E\times F \overset{w}{\longmapsto} E'\times F}$
$\dfrac{E \xrightarrow{a} E' \quad a\in A}{does_A(E)}$	$\dfrac{E \overset{w}{\longmapsto} E' \quad does_A(E')}{does_A(E)}$
$\dfrac{E \xrightarrow{a} E' \quad a\in A}{E\lceil A \xrightarrow{a} E'\lceil A}$	$\dfrac{E \overset{w}{\longmapsto} E' \quad does_A(E')}{E\lceil A \overset{w}{\longmapsto} E'\lceil A}$
$\dfrac{E \xrightarrow{a} E'}{E[S] \xrightarrow{S(a)} E'[S]}$	$\dfrac{E \overset{w}{\longmapsto} E'}{E[S] \overset{w}{\longmapsto} E'[S]}$
$\dfrac{E \xrightarrow{a} E' \quad X \overset{def}{=} E}{X \xrightarrow{a} E'}$	$\dfrac{E \overset{w}{\longmapsto} E' \quad X \overset{def}{=} E}{X \overset{w}{\longmapsto} E'}$
$\dfrac{E \xrightarrow{a} E'}{\Theta(E) \xrightarrow{a} \Theta(E')}$	$\dfrac{E \overset{n\omega^k}{\longmapsto} E' \not\exists (k'>k).E \overset{m\omega^{k'}}{\longmapsto}}{\Theta(E) \overset{n}{\longmapsto} \Theta(E')}$

Table A.1: Operational rules for WSCCS

and therefore the processes are multi-related by weight, e.g.

$$1.P + 1.P + 1.Q \tag{A.1}$$

can evolve to P with cumulative weight 2, so we must retain both evolutions. The auxiliary predicate $does_A(E)$, which denotes the ability of E to perform A after zero or more probabilistic actions, is well defined since only finitely branching choice expressions are allowed.

A.3 Equational rules

Table A.2 features equational rules that form a sound and complete equational system for WSCCS.

(Σ_1) $\Sigma_{i\in I} w_i.Ei = \Sigma_{j\in J} v_j.E_j$ $\begin{cases} \text{there is a surjection } f : I \longmapsto J \text{ with} \\ v_j = \sum\{w_i | i \in I \wedge f(i) = j\}\ , \\ \text{and for all } i \text{ with } f(i) = j \text{ then } E_i = E_j\ . \end{cases}$

(Exp_1) $a : E \times b : F = ab : (E \times F)$

(Exp_2) $a : E \times \Sigma_{j\in J} v_j.F_j = \Sigma_{j\in J} v_j.(a : E \times F_j)$

(Exp_3) $(\Sigma_{i\in I} w_i.E_i) \times (\Sigma_{j\in J} v_j.F_j) = \Sigma_{(i,j)\in(I,J)} v_i w_j.(E_i \times F_j)$

(Res_1) $(a : E)\lceil A = \begin{cases} a : (E\lceil A) \text{ if } a \in A \\ 0 \text{ otherwise}\ . \end{cases}$

(Res_2) $(\Sigma_{i\in I} w_i.E_i)\lceil A = \Sigma_{j\in J} w_j.(E_j\lceil A)$ where $J = \{i \in I | d_A(E_i)\}$

(Θ_1) $\Theta(a.E) = a.\Theta(E)$

(Θ_2) $\Theta(\Sigma_{i\in I} w_i.E_i) = \Sigma_{j\in J} \mathcal{N}(w_j).\Theta(E_j)$ where $J = \{i \in I | w_i = e\omega^{max_\omega(\{w_i\})}\}$

(Ren) $\Sigma_{i\in I} w_i.E_i = \Sigma_{i\in I} e w_i.E_i$ where e is an EVF

Table A.2: Equational rules for WSCCS

Bibliography

[1] M. Abadi and A.D. Gordon. A calculus for cryptographic protocols: The Spi Calculus. In *Fourth ACM Conference on Computer and Communications Security*, pages 36–47, 1997.

[2] E. Ahmed and A.S. Elgazzar. On some applications of cellular automata. *Physica A-Statistical Mechanics and its Applications*, 296:529–538, 2001.

[3] L.J. Allen, T. Lewis, C.F. Martin, M.A. Jones, C.K. Lo, M. Stamp, G. Mundel, and A.B. Way. Analysis of a measles epidemic. *Statistics in Medicine*, 12:229–239, 1993.

[4] R.M. Anderson, B.T. Greenfell, and R.M. May. Oscillatory fluctuations in the incidence of infectious-disease and the impact of vaccination - time-series analysis. *Journal of Hygiene*, 93:587–608, 1984.

[5] R.M. Anderson, H.C. Jackson, R.M. May, and A.M. Smith. Population-dynamics of fox rabies in europe. *Nature*, 289:765–771, 1981.

[6] R.M. Anderson and R.M. May. The population-dynamics of micro-parasites and their invertebrate hosts. *Philosophical transactions of the Royal Society of London Series B*, 291:451–524, 1981.

[7] J. Antonovics, Y. Iwasa, and M.P. Hassell. A generalized-model of parasitoid, venereal, and vector-based transmission processes. *American Naturalist*, 145:661–675, 1995.

[8] J.C.M. Baeten. A brief history of process algebra. *Theoretical Computer Science*, 335(2/3):131–146, 2005.

[9] N.D. Barlow and J.M. Kean. Simple models for the impact of rabbit calicivirus disease (rcd) on australasian rabbits. *Ecological Modelling*, 109:225–241, 1998.

[10] M. Begon, M. Bennet, R.G. Bowers, N.P. French, S.M. Hazel, and J. Turner. A clarification of transmission terms in host-microparasite models: numbers, densities and areas. *Epidemiology and infection*, 129:147–153, 2002.

[11] J.A. Bergstra and J.W. Klop. Algebra of Communicating Processes with Abstraction. *Theoretical Computer Science*, 37:77–121, 1985.

[12] J.A. Bergstra and C.A. Middelburg. Process algebra for hybrid systems. *Theoretical Computer Science*, 335:215–280, 2005.

[13] D. Bernoulli. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prevenir. *Histoire de l'Acad. Roy. Sci. (Paris) avec Mém. des Math. et Phys. and Mém*, pages 1–45, 1760.

[14] D. Bernoulli and S. Blower. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Reviews in Medical Virology*, 14:275–288, 2004.

[15] K. Berthier, M. Langlais, P. Auger, and D. Pontier. Dynamics of a feline virus with two transmission modes with exponentially growing

host populations. *Proceedings of the Royal Society of London Series B*, 267:2049–2056, 2000.

[16] R.J.H. Beverton and S.J. Holt. *On the dynamics of exploited fish populations*, volume 19 of *Fisheries Investigations, Series 2*. H.M.S.O., 1957.

[17] J.T. Bradley, S.T. Gilmore, and J. Hillston. Analysing distributed internet worm attacks using continuous state-space approximation of process algebra models. *Journal of Computer and System Sciences*, 2007. In Press - doi:10.1016/j.jcss.2007.07.005.

[18] J.T. Bradley, S.T. Gilmore, and N. Thomas. Performance analysis of stochastic process algebra models through stochastic simulation. *Proceedings of the 5th International Workshop on Performance Modeling, Evaluation, and Optimization of Parallel and Distributed Systems*, pages 121–130, 2006.

[19] A. Brännström and D.J.T. Sumpter. The role of competition and clustering in population dynamics. *Proceedings of the Royal Society of London Series B*, 272:2065–2072, 2005.

[20] C.J. Briggs and H.C.J. Godfray. The dynamics of insect-pathogen interactions in stage-structured environments. *American Naturalist*, 145:855–887, 1995.

[21] L. Brodo, P. Degano, and C. Priami. A tool for quantitative analysis of calculus processes. In *ICALP Satellite Workshops*, pages 535–550, 2000.

[22] B.J. Cairns, J.V. Ross, and T. Taimre. A comparison of models for predicting population persistence. *Ecological Modelling*, 201:19–26, 2007.

[23] M. Calder, S. Gilmore, and J. Hillston. Automatically deriving odes from process algebra models of signalling pathways. In *Proceedings of Computational Methods in Systems Biology (CMSB 2005)*, pages 204–215, 2005.

[24] M. Calder, S. Gilmore, and J. Hillston. Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA. *Transactions on Computational Systems Biology VII*, 4230:1–23, 2006.

[25] L. Cardelli. From processes to ODEs by chemistry. 2007. Unpublished draft. Downloaded from http://lucacardelli.name/.

[26] L. Cardelli. On process rate semantics. 2007. To appear in Theoretical Computer Science, Elsevier, 2007. Downloaded from http://lucacardelli.name/.

[27] C. Castillo-Chavez and A.A. Yakubu. Dispersal, disease and life-history evolution. *Mathematical Biosciences*, 173:35–53, 2001.

[28] F. Corradini and W. Vogler. Measuring the performance of asynchronous systems with pafas. *Theoretical Computer Science*, 335:187–213, 2005.

[29] Mount Sinai Hospital Department of Microbiology. Faq: Methods of disease transmission. Webpage: retrieved September 2007 http://microbiology.mtsinai.on.ca/faq/transmission.shtml.

[30] Z.L. Feng and J.X. Velasco-Hernandez. Competitive exclusion in a vector-host model for the dengue fever. *Journal of Mathematical Biology*, 35:523–544, 1997.

[31] A. Fenton, J.P. Fairbairn, R.A. Norman, and P.J. Hudson. Parasite transmission: reconciling theory and reality. *Journal of Animal Ecology*, 71:893–905, 2001.

[32] R. Fujie and T. Odagaki. Effects of superspreaders in spread of epidemic. *Physica A-Statistical Mechanics and its Applications*, 374:843–852, 2007.

[33] L.Q. Gao and H.W. Hethcote. Disease transmission models with density-dependent demographics. *Journal of Mathematical Biology*, 30:717–731, 1992.

[34] B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical transactions of the Royal Society of London Series B*, 115:513–585, 1825.

[35] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics: A foundation for computer science*. Addison-Wesley, 1989.

[36] D. Greenhalgh. An epidemic model with a density-dependent death rate. *IMA Journal of Mathematics Applied in Medicine and Biology*, 7:1–26, 1990.

[37] D. Greenhalgh. Some results for an SEIR epidemic model with density dependence in the death rate. *IMA Journal of Mathematics Applied in Medicine and Biology*, 9:67–106, 1992.

[38] G.R. Grimmett and D.R. Stirzaker. *Probability and Random Processes*. Oxford University Press, 1992.

[39] C.M. Guldberg and P. Waage. Concerning chemical affinity. *Erdmann's Journal fr Practische Chemie*, 127:69–114, 1879.

[40] M.P. Hassell. Density-dependence in single-species populations. *Journal of Animal Ecology*, 45:283–296, 1975.

[41] M. J. Hatcher and C. Tofts. The evolution of polygenic sex determination with potential for environmental manipulation. Technical Report UMCS-95-4-2, Department of Computer Science, University of Manchester, 1995.

[42] H.W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42:599–653, 2000.

[43] H.W. Hethcote and P. van den Driessche. An SIS epidemic with variable population and a delay. *Journal of Mathematical Biology*, 34:177–194, 1995.

[44] J. Hillston. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.

[45] J. Hillston. Fluid Flow Approximation of PEPA models. In *QEST'05, Proceedings of the 2nd International Conference on Quantitative Evaluation of Systems*, pages 33–42. IEEE Computer Society Press, Torino, September 2005.

[46] J. Hillston and N. Thomas. Product form solution for a class of PEPA models. *Performance Evaluation*, 35:171–192, 1999.

[47] C.A.R. Hoare. Communicating sequential processes. *Communications of the ACM*, 21:666–677, 1978.

[48] M.E. Hochberg. Non-linear transmission rates and the dynamics of infectious disease. *Journal of Theoretical Biology*, 153:301–321, 1991.

[49] Wolfram Research Inc. Mathematica 5.2, Wolfram Research. http://www.wolfram.com/products/mathematica/index.en.html (Accessed October 2007).

[50] J. Joo and J.L. Lebowitz. Pair approximation of the stochastic susceptible-infected-recovered-susceptible epidemic model on the hypercubic lattice. *Physical Review E*, 70, 2004.

[51] J.T. Kemper. Identification of superspreaders for infectious-disease. *Mathematical Biosciences*, 48:111–127, 1980.

[52] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics i. *Proceedings of the Royal Society of London Series A*, 115:700–721, 1927.

[53] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics ii - the problem of endemicity. *Proceedings of the Royal Society of London Series A*, 138:55–83, 1932.

[54] W.O. Kermack and A.G. McKendrick. Contributions to the mathematical theory of epidemics iii - further studies of the problem of endemicity. *Proceedings of the Royal Society of London Series A*, 141:94–122, 1933.

[55] R.J. Knell, M. Begon, and D.J. Thompson. Transmission of *Plodia interpunctella* granulosis virus does not conform to the mass action model. *Journal of Animal Ecology*, 67:592–599, 1998.

[56] C.M. Kribs-Zaleta and J.X. Velasco-Hernandez. A simple vaccination model with multiple endemic states. *Mathematical Biosciences*, 164:183–201, 2000.

[57] K. Kuratowski. *Topology: Volume 1.* Polish Scientific Publishers, 1966.

[58] T.G. Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability*, 7:49–58, 1970.

[59] M. Liljenstam, D.M. Nicol, V.H. Berk, and R.S. Gray. Simulating realistic network worm traffic for worm warning system design and testing. In *Proceedings of the 2003 ACM Workshop on Rapid Malcode (WORM)*. ACM Press, 2003.

[60] J.O. Lloyd-Smith, S.J. Schreiber, P.E. Kopp, and W.M. Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438:355–359, 2005.

[61] A.J. Lotka. *Elements of physical biology*. Williams and Wilkins, 1925.

[62] T. Malthus. An essay on the principle of population, 1798.

[63] M. Marusic, Z. Bajzer, J.P. Freyer, and S. Vukpavlovic. Analysis of growth of multicellular tumor spheroids by mathematical models. *Cell Proliferation*, 27:73–94, 1994.

[64] L. Matthews and M. Woolhouse. New approaches to quantifying the spread of infection. *Nature Reviews Microbiology*, 3:529–536, 2005.

[65] R.M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.

[66] J. Menalorca and H.W. Hethcote. Dynamic models of infectious-diseases as regulators of population sizes. *Journal of Mathematical Biology*, 30:693–716, 1992.

[67] R. Milner. *A Calculus of Communicating Systems*, volume 92 of *Lecture Notes in Computer Science*. Springer-Verlag, 1980.

[68] R. Milner. *Communication and Concurrency*. Prentice Hall, 1989.

[69] R. Milner. *Communicating and Mobile Systems: the π-Calculus*. Cambridge University Press, 1999.

[70] R. Milner, J. Parrow, and D. Walker. A calculus of mobile processes .1. *Information and Computation*, 100:1–40, 1992.

[71] J. Munyandorero. Length-based Beverton and Holt spawning stock biomass per-recruit-models, with application to the lates stappersii (boulenger) stock in lake tanganyika. *Fisheries Management and Ecology*, 8:1–14, 2001.

[72] E.A.C Newton and P. Reiter. A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics. *American Journal of Tropical Medicine and Hygiene*, 47:709–720, 1992.

[73] R. Norman and C. Shankland. Developing the use of process algebra in the derivation and analysis of mathematical models of infectious disease. In *Computer Aided Systems Theory - EUROCAST 2003*, volume 2809 of *Lecture Notes in Computer Science*, pages 404–414. Springer-Verlag, 2003.

[74] C.A. Petri. *Kommunikation mit automaten*. PhD thesis, Institut fuer Instrumentelle Mathematik, Bonn, 1962.

[75] P. Phillipe. Chaos, population biology, and epidemiology - some research implications. *Human Biology*, 65:525–546, 1993.

[76] D.A. Rand, M. Keeling, and H.B. Wilson. Stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies. *Proceedings of the Royal Society of London Series B*, 259:55–63, 2004.

[77] A. Regev, E.M. Panina, W. Silverman, L. Cardelli, and E. Shapiro. Bioambients: an abstraction for biological compartments. *Theoretical Computer Science*, 325:141–167, 2004.

[78] W.E. Ricker. Stock and recruitment. *Journal of the Fisheries Research Board of Canada*, 11:559–623, 1954.

[79] B.J. Ross. A process algebra for stochastic music composition. In *Proceedings 1995 International Computer Music Conference*, pages 448–451, 1995.

[80] J.J. Ryder, M.R. Miller, A. White, R.J. Knell, and M. Boots. Host-parasite population dynamics under combined frequency- and density-dependent transmission. *Oikos*, 116:2017–2026, DEC 2007.

[81] K. Sato, H. Matsuda, and A. Sasaki. Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology*, 32:251–268, 1994.

[82] D. Sumpter. *From Bee to Society: an agent based investigation of honeybee colonies*. PhD thesis, UMIST, 2000.

[83] D.J.T. Sumpter, G.B. Blanchard, and D.S. Broomhead. Ants and agents: a process algebra approach to modelling ant colony behaviour. *Bulletin of Mathematical Biology*, 63:951–980, 2001.

[84] S.Y. Tang, R.A. Cheke, and Y.N. Xiao. Optimal impulsive harvesting on non-autonomous Beverton-Holt difference equations. *Nonlinear Analysis - Theory and Applications*, 65:2311–2341, 2006.

[85] P.H. Thrall, J. Antonovics, and D.W. Hall. Host and pathogen coexistence in sexually-transmitted and vector-borne diseases characterized

by frequency-dependent transmission. *American Naturalist*, 142:543–552, 1993.

[86] C. Tofts. Using process algebra to describe social insect behaviour. *Transactions of the Society for Computer Simulation*, 9:227–283, 1993.

[87] C. Tofts. Processes with probabilities, priority and time. *Formal Aspects of Computing*, 6:536–564, 1994.

[88] C. Tofts. Exact, analytic, and locally approximate solutions to discrete event-simulation problems. *Simulation Practice and Theory*, 6:721–759, 1998.

[89] C. Tofts. Exploiting strong attractors to slaughter monsters - taming 10^{1500} states and beyond. Technical Report HPL-2006-121, HP Laboratories, Bristol, 2006.

[90] J. Turner, M. Begon, and R.G. Bowers. Modelling pathogen transmission: the interrelationship between local and global approaches. *Proceedings of the Royal Society of London Series B*, 270:105–112, 2002.

[91] P.F. Verhulst. Notice sur la loi que la population suit dans son accroissement. *Corr. Math. et Phys.*, 10:113–121, 1838.

[92] V. Volterra. Variations and fluctuations of the number of individuals in animal species living together. In *Animal Ecology*, pages 409–448. McGraw-Hill, 1931.

[93] R.W. West and J.R. Thompson. Models for the simple epidemic. *Mathematical Biosciences*, 141:29–39, 1997.

[94] M.E.J. Woolhouse, C. Dye, J.F. Etard, T. Smith, J.D. Charlwood, G.P. Garnett, P. Hagan, J.L.K. Hii, P.D. Ndhlovu, R.J. Quinnell, C.H.

Watts, S.K. Chandiwana, and R.M. Anderson. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proceedings of the National Academy of Sciences of the United States of America*, 94:338–342, 1997.

[95] J.S Zhou and H.W. Hethcote. Population-size dependent incidence in models for diseases without immunity. *Journal of Mathematical Biology*, 32:809–834, 1994.