QUANTICOL

A Quantitative Approach to Management and Design of Collective and Adaptive Behaviours

quantiCOL

http://www.quanticol.eu

TR-QC-05-2016

An investigation of spatial moment closure

Revision: 1.0; September 2016

Author(s): Vashti Galpin (UEDIN)

Publication date: September 2016

Funding Scheme: Small or medium scale focused research project (STREP)
Topic: ICT-2011 9.10: FET-Proactive 'Fundamentals of Collective Adaptive Systems' (FOCAS)
Project number: 600708

Coordinator: Jane Hillston (UEDIN)
e-mail: Jane.Hillston@ed.ac.uk
Fax: +44 131 651 1426

Part. no.	Participant organisation name	Acronym	Country
1 (Coord.)	University of Edinburgh	UEDIN	UK
2	Consiglio Nazionale delle Ricerche – Istituto di Scienza e Tecnologie della Informazione "A. Faedo"	CNR	Italy
3	Ludwig-Maximilians-Universität München	LMU	Germany
4	Ecole Polytechnique Fédérale de Lausanne	EPFL	Switzerland
5	IMT Lucca	IMT	Italy
6	University of Southampton	SOTON	UK
7	Institut National de Recherche en Informatique et en Automatique	INRIA	France

COOPERATION

Abstract

This document investigates the use of spatial moment closure techniques on large grid-based spatial models. The basic SIR model is extended to a spatial SIR model by the addition of locations. Ordinary differential equations are derived for spatial moments and three different closure techniques are applied. Experimentation shows that (for the parameters under consideration) the fourth moment stochastic linearisation gives a better approximation than third moment stochastic linearisation or third moment log-normal approximation.

Contents

1 Introduction

In this document, we consider dynamic mathematical models defined over discrete space. We assume that discrete space is represented by an undirected graph with nodes representing locations and edges representing connections between the nodes. Movement and/or interaction between two nodes can only take place if there is an edge between those nodes. We consider spatial moment closure which is a group of approximation techniques applied to spatial moment ODEs (ordinary differential equations). These ODEs describe the dynamics of moments defined for variables of interest relating to space. Typically, these moment ODE systems are infinite, and moment closure techniques approximate higher order moments, thus reducing an infinite system of ODEs to a usable finite system of ODEs.

The term "spatial moment closure" has been applied in two different ways when considering discrete-space models. There are two main types of discrete-space models: individual and population. Individual discrete-space models are those where each node in the location graph is viewed either as a location which can take one of a finite number of discrete values, or alternatively, as a single individual at that location taking one of these values. Typically, the dynamics of such models are defined by transformation rules that define how the value at each node varies over time depending on the values of neighbouring nodes. In this context, spatial moment closure is also called pair approximation. This is because the variables of interest are those that count the number of patterns in the model over time. By patterns, we mean small subgraphs, starting with each possible combination of pairs of nodes joined by an edge (hence the term "pair approximation"). Moment ODEs are derived by considering the transformation rules that can affect each pair. The right-hand sides of these ODEs may contain references to three-node patterns. In general, an infinite system of moment ODEs are obtained and moment closure techniques can be applied [LD96, BP97, BP99, Gas15]. We do not consider spatial moment closure techniques for individual discrete space models further in this document.

Instead, we focus on population discrete-space models where multiple agents are located at each node in the graph. This is an extension and generalisation of a standard population continuous-time Markov model to include locations. We can derive ODEs for the spatial moments of the models. The moments are some quantity of interest that is summed across all locations and then divided by the number of locations. Again, we obtain an infinite systems of ODEs in most cases and moment closure techniques can be applied [PL97, MMRL02, MSH05].

The remainder of this document considers population discrete models, with a specific focus on epidemiological models of disease spread where space plays an important role. We start with general definitions before proceeding with examples.

2 Definitions and notation

Population CTMCs

A population continuous-time Markov chain (PCTMC) can be defined with or without locations. We give both definitions now.

Definition 1. *A* population continuous time Markov chain (PCTMC) *is a tuple* $(\mathbf{Y}, \mathcal{D}, \mathcal{T})$ *where*

- $\mathbf{Y} = (Y_1, \ldots, Y_n)$ *is a vector of* variables
- $\mathcal{D}$ *is a countable set of* states *defined as* $\mathcal{D} = \mathcal{D}_1 \times \ldots \times \mathcal{D}_n$ *where each* $\mathcal{D}_i \subseteq \mathbb{N}$ *represents the domain of* Y_i
- $\mathcal{T} = \{\tau_1, \ldots \tau_m\}$ *is the set of transitions of the form* $\tau_l = (\mathbf{v}, r)$ *where*
  - $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{Z}^n$ *is the* state change or update vector *where* v_i *describes the change in number of units of* Y_i *caused by transition* τ_l
  - $r : \mathcal{D} \to \mathbb{R}_{\geq 0}$ *is the* rate function *of transition* τ_l *with* $r(\mathbf{d}) = 0$ *whenever* $\mathbf{d} + \mathbf{v} \notin \mathcal{D}$.

As we wish to consider location indices in this document, we can introduce a slightly different definition that considers these. We assume n subpopulations as is the case for the standard CTMC definition above, and p locations. The notation $X_i^{(j)}$ denotes the size of the subpopulation i at location j.

Definition 2. *A* population continuous time Markov chain with locations (PLCTMC) *is a tuple* $(\mathbf{X}, \mathcal{D}, \mathcal{T})$ *where*

- $\mathbf{X} = (X_1^{(1)}, \ldots, X_p^{(1)}, \ldots, X_1^{(n)}, \ldots, X_p^{(n)})$ *is a vector of* variables
- $\mathcal{D}$ *is a countable set of* states *defined as* $\mathcal{D} = \mathcal{D}_1^{(1)} \times \ldots \times \mathcal{D}_1^{(p)} \times \ldots \times \mathcal{D}_n^{(1)} \times \ldots \times \mathcal{D}_n^{(p)}$ *where each* $\mathcal{D}_i^{(j)} \subseteq \mathbb{N}$ *represents the domain of* $X_i^{(j)}$
- $\mathcal{T} = \{\tau_1, \ldots \tau_m\}$ *is the set of transitions of the form* $\tau_l = (\mathbf{v}, r)$ *where*
  - $\mathbf{v} = (v_1^{(1)}, \ldots, v_1^{(p)}, \ldots, v_n^{(1)}, \ldots, v_n^{(p)}) \in \mathbb{Z}^{np}$ *is the* state change or update vector *where* v_i *describes the change in number of units of* $X_i^{(j)}$ *caused by transition* τ_l
  - $r : \mathcal{D} \to \mathbb{R}_{\geq 0}$ *is the* rate function *of transition* τ_l *with* $r(\mathbf{d}) = 0$ *whenever* $\mathbf{d} + \mathbf{v} \notin \mathcal{D}$.

The notation used in this document can be summarised as follows.

$$\begin{array}{llllll} \mathbf{X}_i & = & (X_i^{(1)}, \ldots, X_i^{(p)}) & X_i & = & \sum_{j=1}^{p} X_i^{(j)} \\ \mathbf{X}^{(j)} & = & (X_1^{(j)}, \ldots, X_n^{(j)}) & X^{(j)} & = & \sum_{i=1}^{n} X_i^{(j)} \\ \mathbf{X} & = & (\mathbf{X}_1, \ldots, \mathbf{X}_n) & X & = & \sum_{i=1}^{n}\sum_{j=1}^{p} X_i^{(j)} \quad = \quad \sum_{j=1}^{p}\sum_{i=1}^{n} X_i^{(j)} \end{array}$$

$$\langle X_i \rangle = 1/p \sum_{j=1}^{p} X_i^{(j)} \qquad \langle X_{i_1}^{q_1} \ldots X_{i_m}^{q_m} \rangle = 1/p \sum_{j=1}^{p} (X_{i_1}^{(j)})^{q_1} \ldots (X_{i_m}^{(j)})^{q_m}$$

We are interested in the expected values of spatial averages of single variables, $\mathbf{E}[\langle X_i \rangle]$ as well as expected values of spatial averages over products of variables $\mathbf{E}[\langle X_{i_1}^{q_1} \ldots X_{i_m}^{q_m} \rangle]$. This can be written more generally as $\mathbf{E}[\langle M(\mathbf{X}) \rangle]$ where $M(\cdot)$ is a polynomial defined over $X_1, \ldots, X_n$. However, since expectations are preserved by addition and by multiplication by a constant, we are only interested in those $M(\cdot)$ that are a product of powers of the variables $X_1, \ldots, X_n$. We will derive ODEs that define the behaviour of these expectations.

3 From non-spatial SIR to spatial SIR

We start by considering the SIR (Susceptible-Infected-Recovered) epidemiological model whose CMTC can be defined by the vector (S, I, R) and the domain $(\mathbb{R}_{\geq 0})^3$. The transitions between the states of the CTMC are

- $\tau_1 = ((-1, +1, 0), cSI)$ which represents infection and
- $\tau_2 = ((0, -1, +1), rI)$ which represents recovery.

These transitions can also be written in the following form.

$$\begin{array}{rcll} S + I & \to & I + I & \text{at rate } cSI \\ I & \to & R & \text{at rate } rI \end{array}$$

Note that when S or I is zero, the transitions have zero rate and hence it is not possible for S or I to become negative. This is a nonspatial model since it is based on the assumption that all of the susceptibles, infecteds and recovereds are well-mixed in the space they occupy. In certain circumstances, we wish to move away from this assumption and instead make the different assumption that there are distinct locations in space, and each individual is situated at one of these locations, leading to

separate groups consisting of susceptibles, infecteds and recovereds. We then use superscripted indices in brackets to indicate which subpopulation we are referring to. The model can then be modified to a vector of the form $(S^{(1)}, I^{(1)}, R^{(1)}, \ldots, S^{(p)}, I^{(p)}, R^{(p)})$ assuming p locations, with domain $(\mathbb{R}_{\geq 0})^{3p}$. The two transition rules

$$\begin{array}{rcll} S^{(j)} + I^{(j)} & \rightarrow & I^{(j)} + I^{(j)} & \text{at rate } c^{(j)} S^{(j)} I^{(j)} \\ I^{(j)} & \rightarrow & R^{(j)} & \text{at rate } h^{(j)} I^{(j)} \end{array}$$

define the interactions that occur at each location j. However, since there is no interaction between locations, this a rather unsatisfactory model. We can add the following rules for movement.

$$\begin{array}{rcll} S^{(j)} & \rightarrow & S^{(k)} & \text{for } k \in \text{nb}(j) \quad \text{at rate } {m_S}^{(ij)} S^{(j)} \\ I^{(j)} & \rightarrow & I^{(k)} & \text{for } k \in \text{nb}(j) \quad \text{at rate } {m_I}^{(ij)} I^{(j)} \\ R^{(j)} & \rightarrow & R^{(k)} & \text{for } k \in \text{nb}(j) \quad \text{at rate } {m_R}^{(ij)} R^{(j)} \end{array}$$

Here, $\text{nb}(j)$ refers to the neighbourhood of location j, and includes every location k for which there is an edge (j, k) in the location graph. In the rules above, parameter homogeneity has not been assumed, and the infection, recovery and movement parameters can vary between locations.

For the spatial SIR model, the movement of recovereds have no effect on infection or recovery and hence the last rule can be omitted (and will be from now on). Furthermore, experimentation (not reported here) suggests that the movement of the susceptibles has little effect on infection, and could also be omitted. Reduction in the number of (applicable) rules can lead to faster simulation.

We can also derive a deterministic fluid approximation for the spatial SIR model as follows.

$$\begin{array}{rcl} \frac{dS^{(j)}}{dt} & = & -c^{(j)} S^{(j)} I^{(j)} - \sum_{k \in \text{nb}(j)} {m_S}^{(jk)} S^{(j)} + \sum_{k \in \text{nb}(j)} {m_S}^{(kj)} S^{(k)} \\ \frac{dI^{(j)}}{dt} & = & c^{(j)} S^{(j)} I^{(j)} - h^{(j)} I^{(j)} - \sum_{k \in \text{nb}(j)} {m_I}^{(jk)} I^{(j)} + \sum_{k \in \text{nb}(j)} {m_I}^{(kj)} I^{(k)} \\ \frac{dR^{(j)}}{dt} & = & h^{(j)} I^{(j)} \end{array}$$

We now consider some basic experimentation with these models. The different neighbourhoods that we consider are as follows.

No connectivity: The location graph has no edges, and a result of this, there is no movement between locations.

Full connectivity: The location graph has an edge between every pair of locations, and hence is a complete graph.

Von Neumann connectivity: This is the connectivity found in a grid where each location is connected to the locations to its north, south, east and west. This means that most locations have four neighbours. However, boundary locations have three neighbours or two if they are at the corners of the grid. Another solution would be to link the topmost locations with those at the bottom, and the leftmost with the rightmost thus defining a grid defined over a torus. However, this is unlikely to happen in reality so we work with the grid with boundaries.

The first example considers a 12x12 grid and an SIR model in which only infecteds move between locations[1]. Parameters and initial values are the same for each location. Figures 1, 2 and 3 show a single stochastic simulation of the spatial SIR model where only infecteds move, for the three different types of connectivity. Figure 4 illustrates how the results can differ once all infecteds have recovered and Figure 5 shows the trajectories for the spatial averages of the three species from simulations for each connectivity type. Figure 6 shows a deterministic simulation, and in contrast to the stochastic simulation, every location shows the same behaviour.

[1]For these grid-based models, it can be assumed that the location index j represents a pair defining a grid location but we elide the details.

The second example considers a 30x30 grid and an SIR model in which both susceptibles and infecteds move, and von Neumann connectivity is used. In the first figure, Figure 7, parameters and initial values do not vary between locations. In Figure 8, the initial values for locations vary. Most cells have no infecteds at the start of the simulation. The only locations that have infecteds are the outside border locations and the locations adjacent to those. By the end of the simulation, there is little indication of this starting configuration. In Figure 9, the bottom third of locations have a faster infection rate than those in the top two thirds, and it can be seen that this difference leads to a clear difference in the quantity of recovereds, with more recovereds in the bottom third region.

4 Motivation

The models we have considered are large with up to 900 locations, giving 2700 distinct subpopulations. The simulations using a Gillespie SSA algorithm are reasonably fast given the size but are at the limits of memory capability (on an ordinary laptop). An obvious question is then to consider whether there are usful approximations that can be used beyond the obvious deterministic approximation presented in the previous section. As Figure 5 shows, this deterministic approximation is qualitatively imprecise, since in the actual stochastic model, some susceptibles remain where as none remain in the deterministic approximation.

A technique that has been proposed in the literature is spatial moment closure. As described above, it involves the derivation of spatial moment ODEs and then applies moment closure techniques to these. These ODEs describe how the quantities

$$\mathbf{E}[\langle X_{i_1}^{q_1} \dots X_{i_m}^{q_m} \rangle] \qquad \text{where} \qquad \langle X_{i_1}^{q_1} \dots X_{i_m}^{q_m} \rangle = 1/p \sum_{j=1}^{p} (X_{i_1}^{(j)})^{q_1} \dots (X_{i_m}^{(j)})^{q_m}.$$

change over time. In the SIR model, we will be interested in moments such as $\mathbf{E}[\langle S \rangle]$, $\mathbf{E}[\langle S^2 \rangle]$, $\mathbf{E}[\langle SI \rangle]$ and $\mathbf{E}[\langle SI^2 \rangle]$ where $\mathbf{E}[\langle S \rangle]$ is the spatial average of S, $\mathbf{E}[\langle S^2 \rangle]$ is the spatial average of S^2, and so on. We can also consider central moments such as variance and covariance.

$$\begin{aligned} \mathbf{Var}[S] &= \mathbf{E}[\langle S^2 \rangle] - (\mathbf{E}[\langle S \rangle])^2 \\ \mathbf{Cov}[S, I] &= \mathbf{E}[\langle SI \rangle] - \mathbf{E}[\langle S \rangle]\mathbf{E}[\langle I \rangle] \end{aligned}$$

As we will see, including covariance terms improves the approximation, since these terms capture how the quantities of S and I vary with respect to each other. To ensure these terms are included, at least secind moments must be considered, and hence moment lcosure can only be applied to third moments or higher.

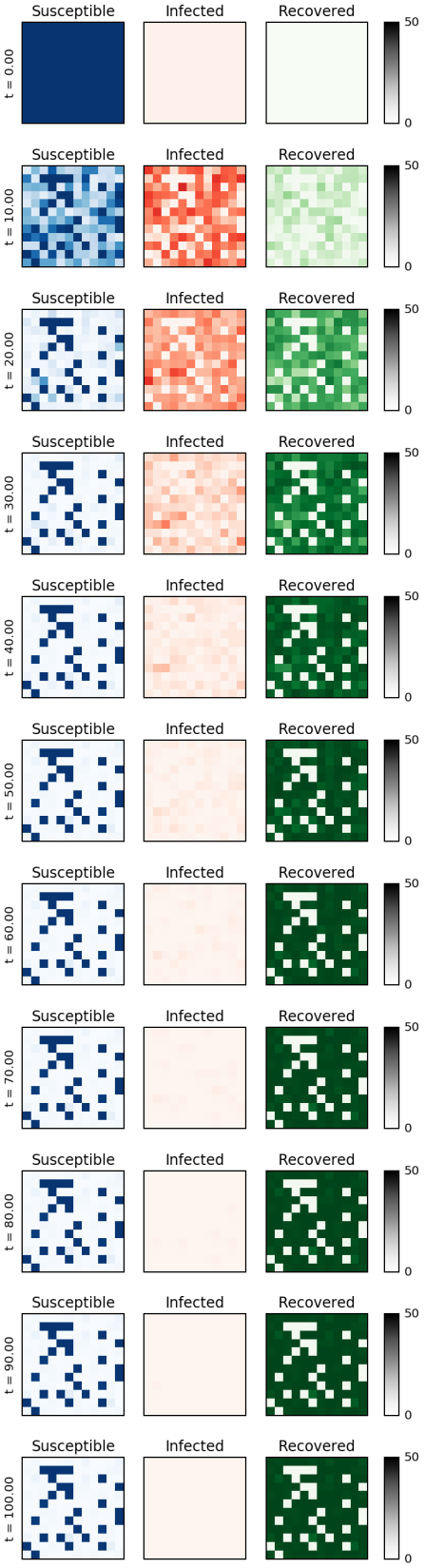


Figure 1: Single stochastic simulation of the SIR model on 12x12 grid with no connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, $m_I{}^{(ij)} = 0.00001$ (only infecteds move)

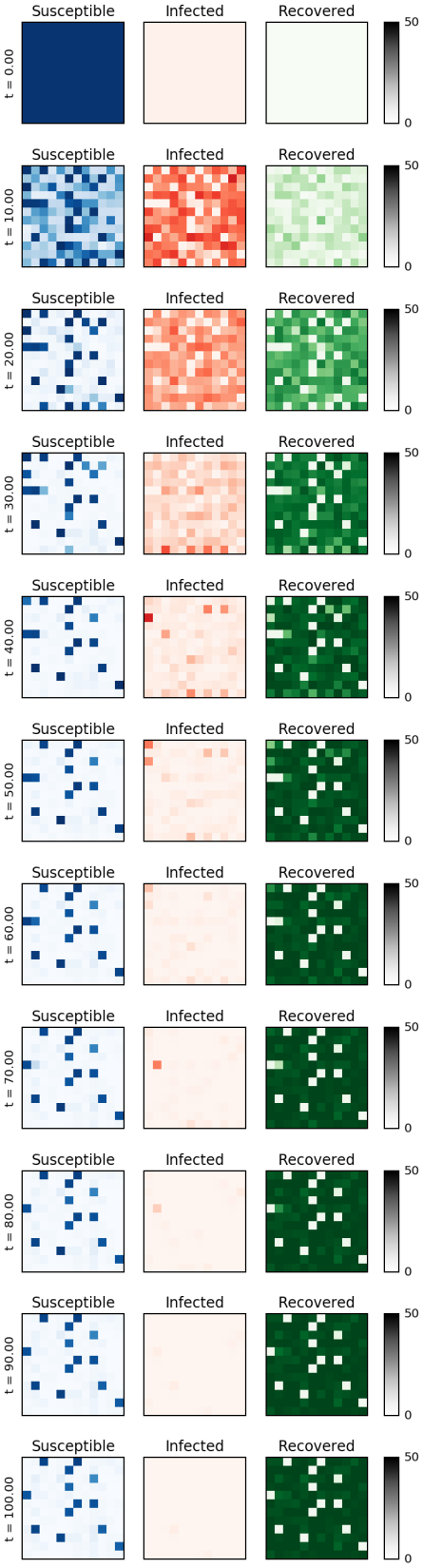


Figure 2: Single stochastic simulation of the SIR model on 12x12 grid with full connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, $m_I{}^{(ij)} = 0.00001$ (only infecteds move)

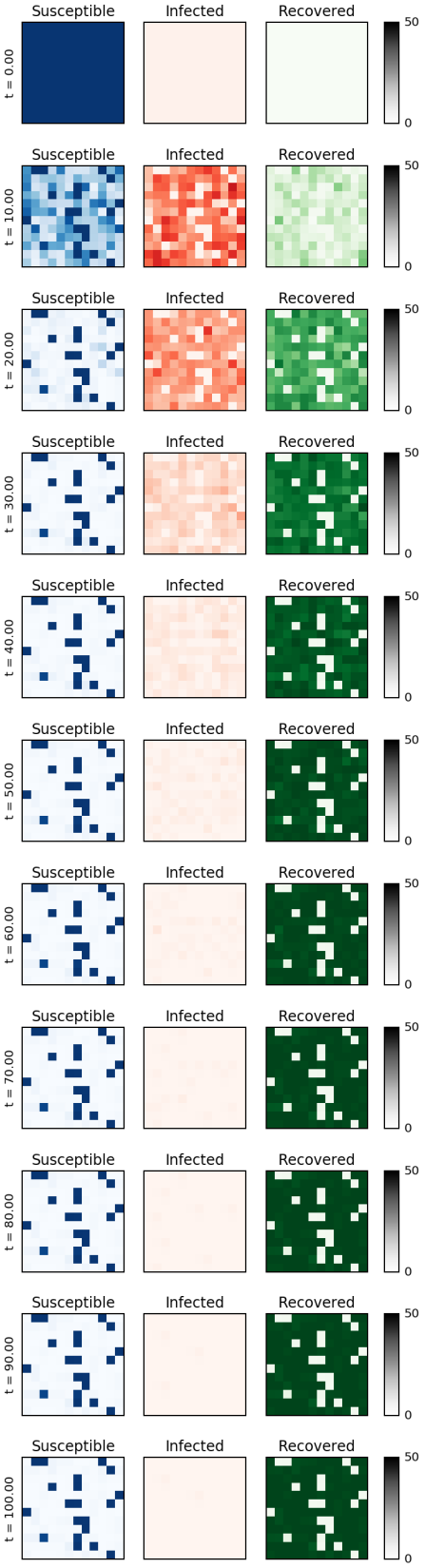


Figure 3: Single stochastic simulation of the SIR model on 12x12 grid with von Neumann connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, ${m_I}^{(ij)} = 0.00001$ (only infecteds move)

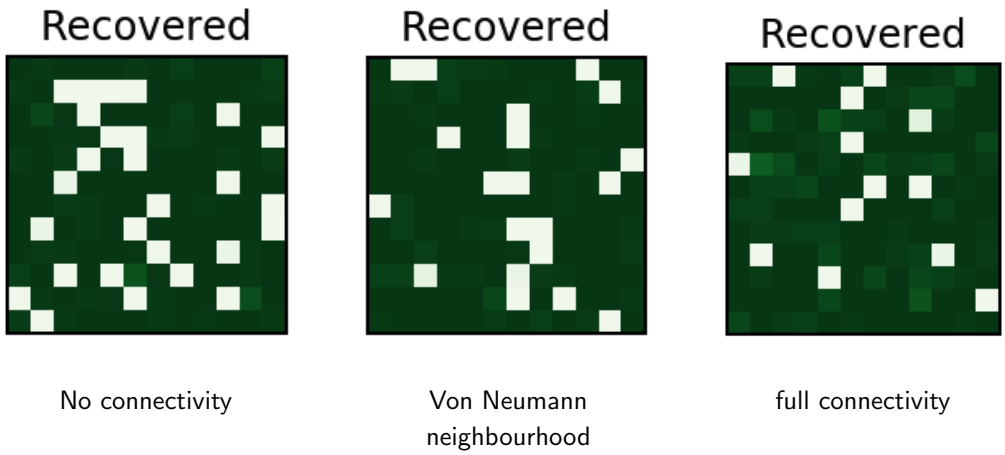


Figure 4: Single stochastic simulation at time $t = 100$ of the SIR model on 12x12 grid for various connectivities with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, $m_I{}^{(ij)} = 0.00001$ (only infecteds move)

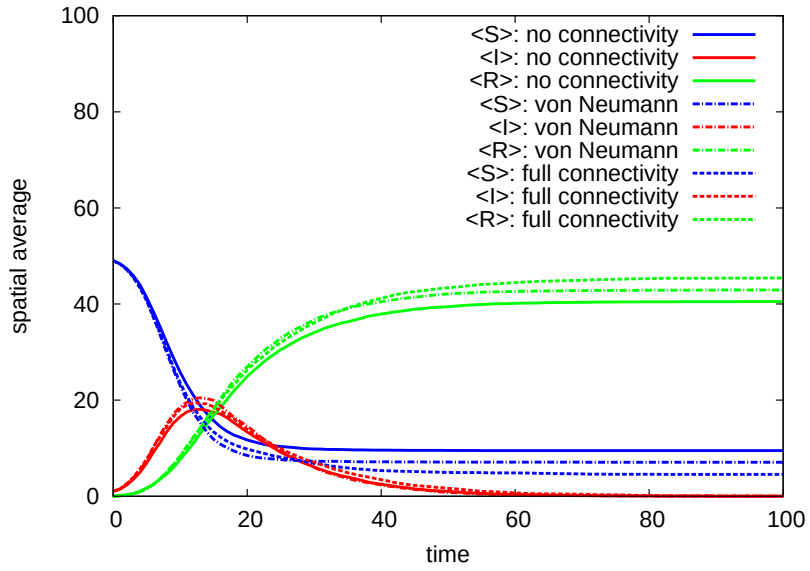


Figure 5: Spatial averages for the SIR model on 12x12 grid for various connectivities with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, $m_I{}^{(ij)} = 0.00001$ (only infecteds move)

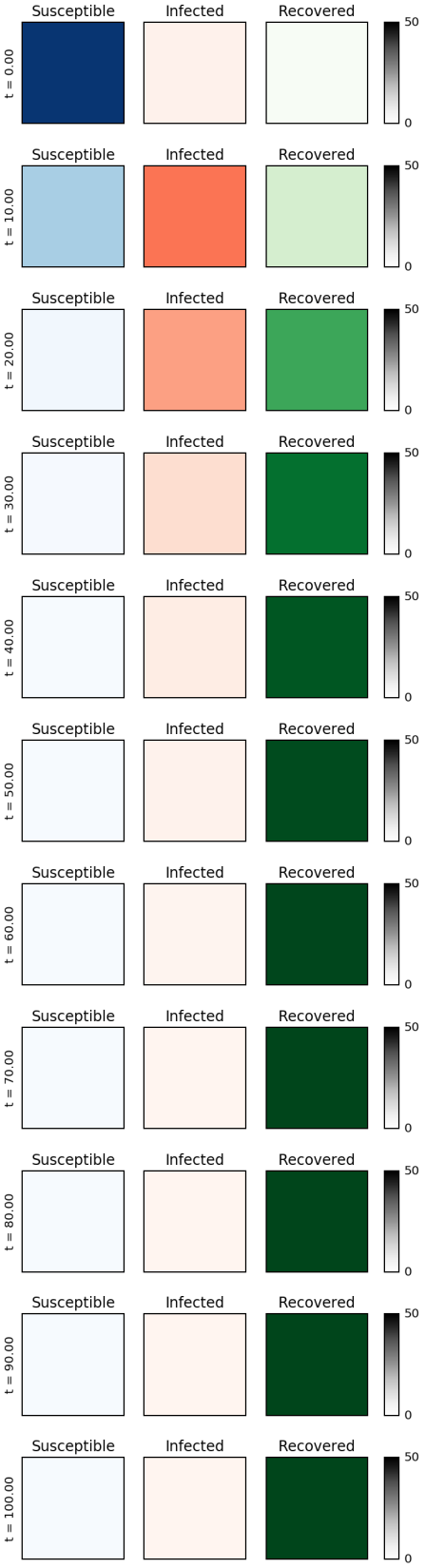


Figure 6: Deterministic simulation of the SIR model on 12x12 grid with full connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, ${m_I}^{(ij)} = 0.00001$ (only infecteds move)

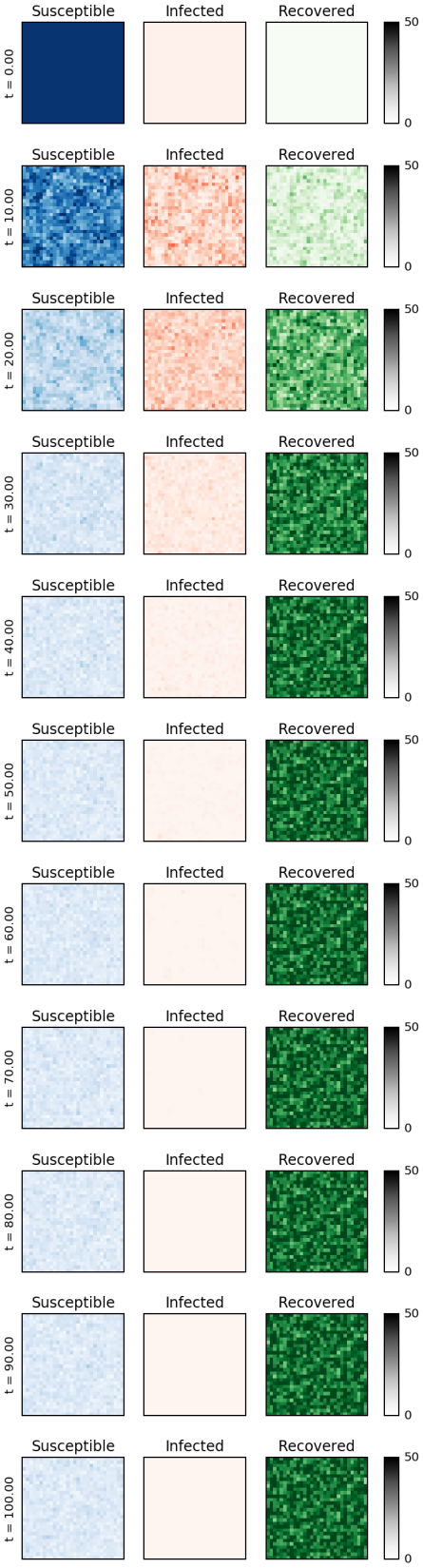


Figure 7: Single stochastic simulation of the SIR model on 30x30 grid with von Neumann connectivity with initial values $S^{(i)}(0) = 30, I^{(i)}(0) = 10, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.001$, $h^{(i)} = 0.2$, ${m_S}^{(ij)} = {m_I}^{(ij)} = 0.05$ (only susceptibles and infecteds move)

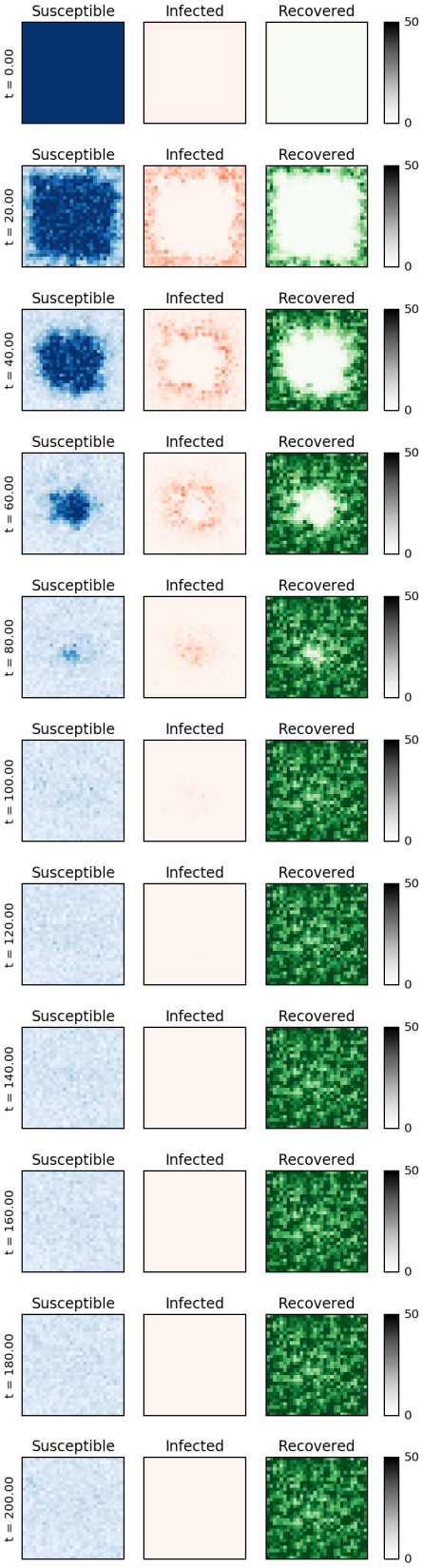


Figure 8: Single stochastic simulation of the SIR model on 30x30 grid with von Neumann connectivity with initial values $S^{(ij)}(0) = 30$, $I^{(ij)}(0) = 0$ except $I^{(ij)}(0) = 10$ for two cell border, $R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.001$, $h^{(i)} = 0.2$, ${m_S}^{(ij)} = {m_I}^{(ij)} = 0.05$

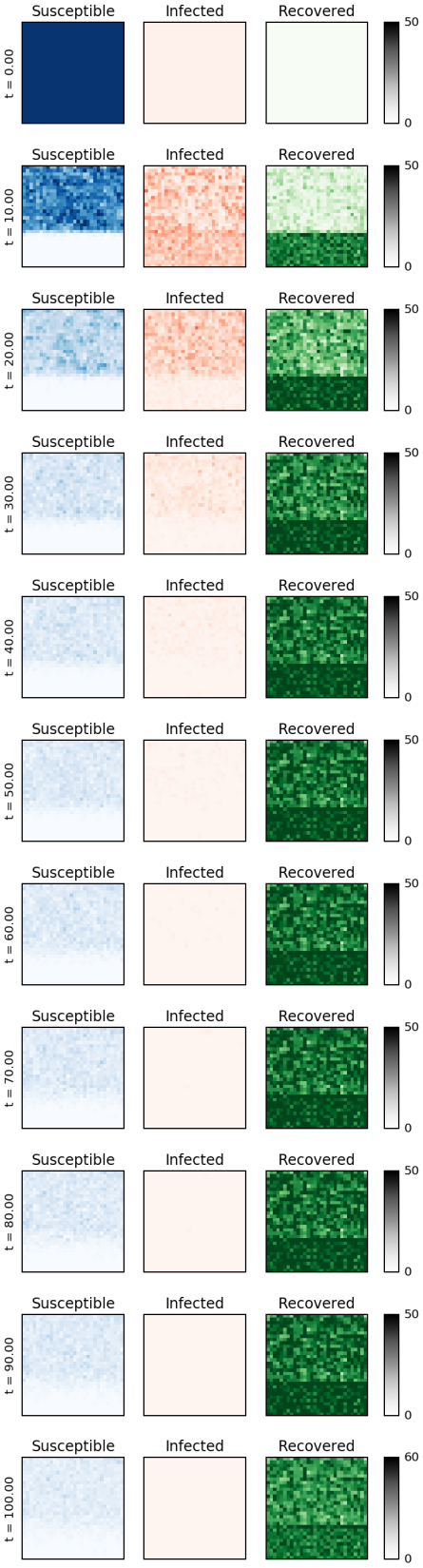


Figure 9: Single stochastic simulation of the SIR model on 30x30 grid with von Neumann connectivity with initial values $S^{(ij)}(0) = 30$, $I^{(ij)}(0) = 10$, $R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.001$ except bottom third of grid where $c^{(i)} = 0.005$, $h^{(i)} = 0.2$, ${m_S}^{(ij)} = {m_I}^{(ij)} = 0.05$

$$
\begin{aligned}
\frac{d\mathbf{E}[\langle S\rangle]}{dt} &= -c\mathbf{E}[\langle SI\rangle] \\
\frac{d\mathbf{E}[\langle I\rangle]}{dt} &= c\mathbf{E}[\langle SI\rangle] - h\mathbf{E}[\langle I\rangle] \\
\frac{d\mathbf{E}[\langle R\rangle]}{dt} &= h\mathbf{E}[\langle I\rangle] \\
\frac{d\mathbf{E}[\langle SI\rangle]}{dt} &\approx c(\mathbf{E}[\langle S^2I\rangle] - \mathbf{E}[\langle SI^2\rangle]) - (c+h)\mathbf{E}[\langle SI\rangle] - p(m_I + m_S)\mathbf{Cov}[S,I] \\
\frac{d\mathbf{E}[\langle S^2\rangle]}{dt} &\approx -c(2\mathbf{E}[\langle S^2I\rangle] - \mathbf{E}[\langle SI\rangle]) - 2pm_S(\mathbf{Var}[S] - \mathbf{E}[\langle S\rangle]) \\
\frac{d\mathbf{E}[\langle I^2\rangle]}{dt} &\approx c(2\mathbf{E}[\langle SI^2\rangle] + \mathbf{E}[\langle SI\rangle]) - h(2\mathbf{E}[\langle I^2\rangle] - \mathbf{E}[\langle I\rangle]) - 2pm_I(\mathbf{Var}[I] - \mathbf{E}[\langle I\rangle]) \\
\frac{d\mathbf{E}[\langle S^2I\rangle]}{dt} &\approx -c(2\mathbf{E}[\langle S^2I^2\rangle] - \mathbf{E}[\langle S^3I\rangle] - \mathbf{E}[\langle SI^2\rangle] - c\mathbf{E}[\langle SI\rangle] + (c-h)\mathbf{E}[\langle S^2I\rangle] \\
&\quad + pm_S\mathbf{E}[\langle SI\rangle] + pm_S\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - 2pm_S\mathbf{Cov}[S,SI] - pm_I\mathbf{Cov}[S^2,I] \\
\frac{d\mathbf{E}[\langle SI^2\rangle]}{dt} &\approx c(2\mathbf{E}[\langle S^2I^2\rangle] + \mathbf{E}[\langle SI^3\rangle] + \mathbf{E}[\langle S^2I\rangle]) - 2(c+h)\mathbf{E}[\langle SI^2\rangle] + (c+h)\mathbf{E}[\langle SI\rangle] \\
&\quad + pm_I\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - 2pm_I\mathbf{Cov}[SI,I] - pm_S\mathbf{Cov}[S,I^2]
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Var}[X] &= \mathbf{E}[\langle X^2\rangle] - (\mathbf{E}[\langle X\rangle])^2 \\
\mathbf{Cov}[X,Y] &= \mathbf{E}[\langle XY\rangle] - \mathbf{E}[\langle X\rangle]\mathbf{E}[\langle Y\rangle]
\end{aligned}
$$

Figure 10: Spatial moment ODEs for the spatial SIR model with movement of both susceptibles and infecteds

5 Spatial moment closure in the case of full connectivity

Spatial moment closure has been proposed as a solution to this size problem, allowing for global reasoning and we now illustrate how the spatial moment ODEs can be derived. Spatial moments are the expected values of products of variable averaged over all locations. For convenience, we will consider parameter homogeneity and parameters will not vary between locations, hence $c^{(i)} = c$, $h^{(i)} = h$, ${m_I}^{(ij)} = m_I$ and ${m_S}^{(ij)} = m_S$ for all Thus for the spatial SIR model presented above, the spatial moment ODEs are given in Figure 10. We assume full connectivity as it makes the ODEs more straightforward to derive. As can be seen from the derivation of $\mathbf{E}[\langle SI\rangle]$ below, this introduces p, the number of locations into the ODEs[2]. Note that some ODEs are approximated and the reasons for this are described below.

We first discuss the form these ODEs take before giving some example derivations of the ODEs. All the ODEs for the first moments and second moments have been listed. For the third moments, only the ODEs of the two moments that appear in the RHS of second joint moment ODE are given. Both of these include references to fourth moments: $\mathbf{E}[\langle S^3I\rangle]$, $\mathbf{E}[\langle S^3I\rangle]$ and $\mathbf{E}[\langle SI^3\rangle]$. Deriving these fourth moments ODEs would include fifth moments on the RHS of the equations and so on. Spatial moment closure approximation has not been applied to the ODEs in Figure 10 but an approximation (which is likely to be accurate for a system with a large total population [MMRL02]) has been applied whereby $\mathbf{E}[\langle M_1(\mathbf{X})\rangle\langle M_2(\mathbf{X})\rangle]$ is approximated by $\mathbf{E}[\langle M_1(\mathbf{X})\rangle]\mathbf{E}[\langle M_2(\mathbf{X})\rangle]$. The approximation of $\mathbf{E}[\langle X\rangle]\mathbf{E}[\langle Y\rangle]$ by $\mathbf{E}[\langle X\rangle\langle Y\rangle]$ has much less impact as an approximation when compared to replacing $\mathbf{E}[\langle XY\rangle]$ $\mathbf{E}[\langle X\rangle]\mathbf{E}[\langle Y\rangle]$ because two spatial averages are involved both terms rather than the product

[2]To approximate von Neumann connectivity, this could possibly be changed to four, to capture the number of neighbours that most locations have. This requires further investigation. In previous work, the full connectivity spatial moments have been used to approximate von Neumann connectivity [MSH05].

of two spatial averages and the spatial average of a product. This approximation is necessary for all spatial moments of more than one variable.

We now work from the stochastic definition of the SIR model to illustrate how to dervie ODEsm, starting with

$$S^{(j)}(t+\delta t) - S^{(j)}(t) \;=\; \delta S^{(j)}$$

and hence it is necessary to determine $\delta S^{(j)}$. This can be done by considering all transitions that affect $S^{(j)}$ and then multiplying them by the rate and the time period δt. So to derive the ODEs for $\langle S\rangle$ we need to consider all of the transitions which reduce or increase S_j.

$$\begin{aligned}
\delta S^{(j)} &= ((S^{(j)}-1)(t) - S^{(j)}(t))(cS^{(j)}I^{(j)}(t))\delta t \\
&\quad + \sum_{k=1}^{p}((S^{(j)}-1)(t) - S^{(j)}(t))(m_S S^{(j)}(t))\delta t \\
&\quad + \sum_{k=1}^{p}((S^{(j)}+1)(t) - S^{(j)}(t))(m_S S^{(j)}(t))\delta t \\
&= [-cS^{(j)}I^{(j)}(t) - \sum_{k=1}^{p} m_S S^{(j)}(t) + \sum_{k=1}^{p} m_S S^{(k)}(t)]\delta t
\end{aligned}$$

Taking spatial averages gives the following.

$$\begin{aligned}
& 1/p\sum_{j=1}^{p} S^{(j)}(t+\delta t) - 1/p\sum_{j=1}^{p} S^{(j)}(t) \\
=\; & [-c/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t) - m_S/p\sum_{j=1}^{p}\sum_{k=1}^{p} S^{(j)}(t) + m_S/p\sum_{j=1}^{p}\sum_{k=1}^{p} S^{(k)}(t)]\delta t \\
=\; & [-c\langle SI\rangle - m_S\langle S\rangle\sum_{k=1}^{p} 1 + m_S\langle S\rangle\sum_{j=1}^{p} 1]\delta t \\
=\; & -c\langle SI\rangle\delta t
\end{aligned}$$

Rewriting this in terms of the spatial averages notation and taking expectations gives

$$\mathbf{E}[\langle S\rangle](t+\delta t) - \mathbf{E}[\langle S\rangle](t) \;=\; -c\mathbf{E}[\langle SI\rangle]\delta t$$

and then dividing by δt and and taking the limits as $\delta t \to 0$ results in

$$\frac{d\mathbf{E}[\langle S\rangle]}{dt} \;=\; -c\mathbf{E}[\langle SI\rangle]$$

The ODEs for $\mathbf{E}[\langle I\rangle]$ and $\mathbf{E}[\langle R\rangle]$ are derived in a similar fashion. The ODEs for second and higher moments are similar in form but more complex as the effect of each transition on the variables in the moment must be considered.

We next derive $\mathbf{E}[\langle SI\rangle]$ to show how this is done.

$$\begin{aligned}
& S^{(j)}I^{(j)}(t+\delta t) - S^{(j)}I^{(j)}(t) \\
= \; & \big((S^{(j)}-1)(I^{(j)}+1)(t) - S^{(j)}I^{(j)}(t)\big)(cS^{(j)}I^{(j)}(t))\delta t \\
& + \big(S^{(j)}(I^{(j)}-1)(t) - S^{(j)}I^{(j)}(t)\big)(hI^{(j)}(t))\delta t \\
& - \textstyle\sum_{k=1}^{p}\big(S^{(j)}(I^{(j)}-1)(t) - S^{(j)}I^{(j)}(t)\big)(m_I I^{(j)}(t))\delta t \\
& + \textstyle\sum_{k=1}^{p}\big(S^{(j)}(I^{(j)}+1)(t) - S^{(j)}I^{(j)}(t)\big)(m_I I^{(k)}(t))\delta t \\
& + \textstyle\sum_{k=1}^{p}\big((S^{(j)}-1)I^{(j)})(t) - S^{(j)}I^{(j)}(t)\big)(m_S I^{(j)}(t))\delta t \\
& + \textstyle\sum_{k=1}^{p}\big((S^{(j)}+1)I^{(j)}(t) - S^{(j)}I^{(j)}(t)\big)(m_S I^{(k)}(t))\delta t \\
= \; & [c(S^{(j)})^2 I^{(j)}(t) - cS^{(j)}(I^{(j)})^2(t) - cS^{(j)}I^{(j)}(t) - hS^{(j)}I^{(j)}(t) \\
& - m_I S^{(j)}(t)I^{(j)}(t)\textstyle\sum_{k=1}^{p} 1 + m_I S^{(j)}(t)\sum_{k=1}^{p} I^{(k)}(t) \\
& - m_S S^{(j)}(t)I^{(j)}(t)\textstyle\sum_{k=1}^{p} 1 + m_S I^{(j)}(t)\sum_{k=1}^{p} S^{(k)}(t)]\delta t \\
= \; & [c(S^{(j)})^2 I^{(j)}(t) - cS^{(j)}(I^{(j)})^2(t) - cS^{(j)}I^{(j)}(t)) - hS^{(j)}I^{(j)}(t) \\
& - pm_I S^{(j)}I^{(j)}(t) + pm_I S^{(j)}\langle I\rangle - pm_S S^{(j)}I^{(j)}(t) + pm_S I^{(j)}\langle S\rangle]\delta t
\end{aligned}$$

Therefore, expanding and taking spatial averages gives the following.

$$\begin{aligned}
& 1/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t+\delta t) - 1/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t) \\
= \; & [c/p\sum_{j=1}^{p}(S^{(j)})^2 I^{(j)}(t) - c/p\sum_{j=1}^{p} S^{(j)}(I^{(j)})^2(t) - c/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t) - h/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t) \\
& - pm_I/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t) + pm_I/p\langle I\rangle\sum_{j=1}^{p} S^{(j)}(t) \\
& - pm_S/p\sum_{j=1}^{p} S^{(j)}I^{(j)}(t) + pm_S/p\langle S\rangle\sum_{j=1}^{p} I^{(j)}(t)]\delta t
\end{aligned}$$

This can be rewritten as

$$\begin{aligned}
\langle SI\rangle(t+\delta t) - \langle SI\rangle(t) \; = \; & [c\langle S^2 I\rangle(t) - c\langle SI^2\rangle(t) - c\langle SI\rangle(t) - h\langle SI\rangle(t) \\
& - pm_I\langle SI\rangle(t) + pm_I\langle S\rangle(t)\langle I\rangle(t) - pm_S\langle SI\rangle(t) + pm_S\langle S\rangle(t)\langle I\rangle(t)]\delta t
\end{aligned}$$

and the following ODE can be derived in the same way as before.

$$\begin{aligned}
\frac{d\mathbf{E}[\langle SI\rangle]}{dt} \; = \; & c\mathbf{E}[\langle S^2 I\rangle] - c\mathbf{E}[\langle SI^2\rangle] - c\mathbf{E}[\langle SI\rangle] - h\mathbf{E}[\langle SI\rangle] \\
& - pm_I\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle S\rangle\langle I\rangle] - pm_S\mathbf{E}[\langle SI\rangle] + pm_S\mathbf{E}[\langle S\rangle\langle I\rangle]
\end{aligned}$$

As mentioned above, we will approximate the term $\mathbf{E}[\langle S\rangle\langle I\rangle]$ with $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle]$.

$$\begin{aligned}
\frac{d\mathbf{E}[\langle SI\rangle]}{dt} \; \approx \; & c\mathbf{E}[\langle S^2 I\rangle] - c\mathbf{E}[\langle SI^2\rangle] - c\mathbf{E}[\langle SI\rangle] - h\mathbf{E}[\langle SI\rangle] \\
& - pm_I\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - pm_S\mathbf{E}[\langle SI\rangle] + pm_S\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle]
\end{aligned}$$

Gathering terms together appropriately gives the following ODE.

$$\frac{d\mathbf{E}[\langle SI\rangle]}{dt} \; \approx \; c(\mathbf{E}[\langle S^2 I\rangle] - \mathbf{E}[\langle SI^2\rangle]) - (c+h)\mathbf{E}[\langle SI\rangle] - p(m_I + m_S)\mathbf{Cov}[S, I]$$

This equation contains two terms of an order higher than $\mathbf{E}[\langle SI\rangle]$. They are $\mathbf{E}[\langle S^2I\rangle]$ and $\mathbf{E}[\langle SI^2\rangle]$. Similarly, if we derive ODEs for these terms, they will contain fourth moments on the right hand side[3]

We will now use two different spatial moment closure techniques. The first is stochastic linearization which will be applied separately to third and fourth moments, giving two different approximations. The second is to use the third moments for the log-normal distribution, thus making the assumption that the real moments of the data are similar to those of the log-normal distributions. The log-normal distribution is chosen because it has positive support, and hence, is suitable for modelling distributions of populations [MSH05].

Third moment stochastic linearisation: Approximate

- $\mathbf{E}[\langle S^2I\rangle]$ with $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle SI\rangle]$ or $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle I\rangle]$, and
- $\mathbf{E}[\langle SI^2\rangle]$ with $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I^2\rangle]$ or $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I\rangle]$.

Fourth moment stochastic linearisation: Approximate

- $\mathbf{E}[\langle S^3I\rangle]$ with $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle SI\rangle]$ or $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle S^2I\rangle]$,
- $\mathbf{E}[\langle SI^3\rangle]$ with $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ or $\mathbf{E}[\langle SI^2\rangle]\mathbf{E}[\langle I\rangle]$, and
- $\mathbf{E}[\langle S^2I^2\rangle]$ with $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle I^2\rangle]$, $\mathbf{E}[\langle SI\rangle]^2$, $\mathbf{E}[\langle S^2I\rangle]\mathbf{E}[\langle I\rangle]$ or $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle SI^2\rangle]$.

Log-normal distribution for third order Use the following approximations [MMRL02].

$$\mathbf{E}[\langle S^2I\rangle] \approx \frac{\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle SI\rangle]^2}{\mathbf{E}[\langle S\rangle]^2\mathbf{E}[\langle I\rangle]} \qquad \mathbf{E}[\langle SI^2\rangle] \approx \frac{\mathbf{E}[\langle I^2\rangle]\mathbf{E}[\langle SI\rangle]^2}{\mathbf{E}[\langle I\rangle]^2\mathbf{E}[\langle S\rangle]}$$

The results of these approaches are now shown. Figure 12(a) compares the results of a stochastic simulation of the spatial averages with the trajectory from the deterministic model. As can be seen, the results are both quantitatively and qualitatively differently. The deterministic trajectory shows that all susceptibles become infected and recover, whereas the stochastic simulation indicates that significant numbers remain susceptible. Such different outcomes could have very different implications for what to do if there is the possibility of a further outbreak caused by an infected agent arriving at one of the locations.

The results of the third moment stochastic linearisation are shown in Figure 12(b). This trajectory was obtained for the approximation using $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle I\rangle]$ and $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I\rangle]$. The choice of $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle SI\rangle]$ and $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I\rangle]$ provide similar results. The other two choices resulted in simulation failure due to the minimum possible step size being reached, and so these approximations both seem to be poorly behaved. The results of the third moment log-normal distribution are shown in Figure 12(c). Both of these results are disappointing as they are much closer to the deterministic trajectory that the stochastic trajectory.

The fourth moment stochastic linearisation was affected by numerical problems. Only two choices of approximations gave reasonable (but very different) results and three others gave wrong approximations. The results for the approximation using $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle SI\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle SI\rangle]$ are given in Figure 13(a) and those of the approximation using $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle SI\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle S^2I\rangle]$ are given in Figure 13(b). The results of the approximation using $\mathbf{E}[\langle S^2I\rangle]\mathbf{E}[\langle I\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle S^2I\rangle]$ are given in Figure 13(c), and the choice of using $\mathbf{E}[\langle S^2I\rangle]\mathbf{E}[\langle I\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle SI\rangle]$ gives similar results. On the other hand, the choice $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle SI^2\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle S^2I\rangle]$ gives very incorrect results with negative populations. All other choices for approximation gave a minimum step size error.

The results for the (well-behaved) fourth moment stochastic linearisation clearly gives better results that either closure technique applied to third moments. Figure 11 shows the covariance of S and I for the different techniques.

[3]Compact derivation of these ODEs are presented in the appendix of this document. Note that these derivations provide different results to those that appear in the slides at `http://www.sti.uniurb.it/events/sfm16quanticol/programme.html` and give better approximations.

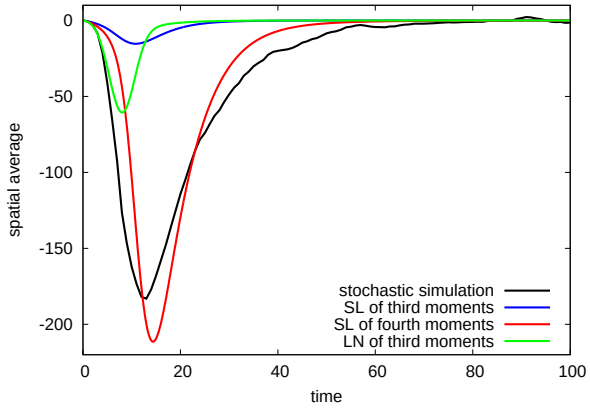
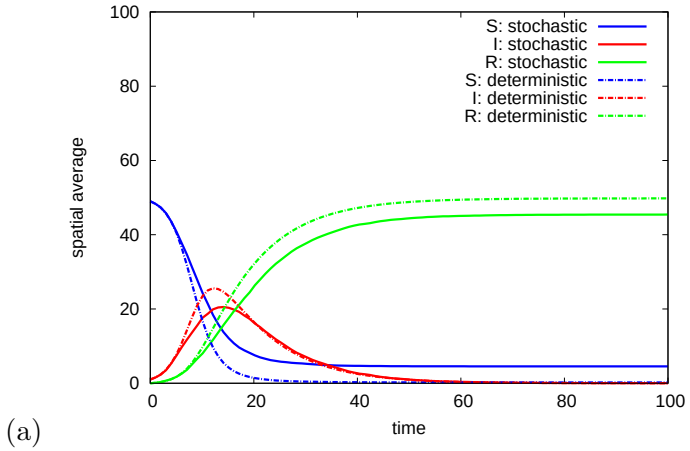


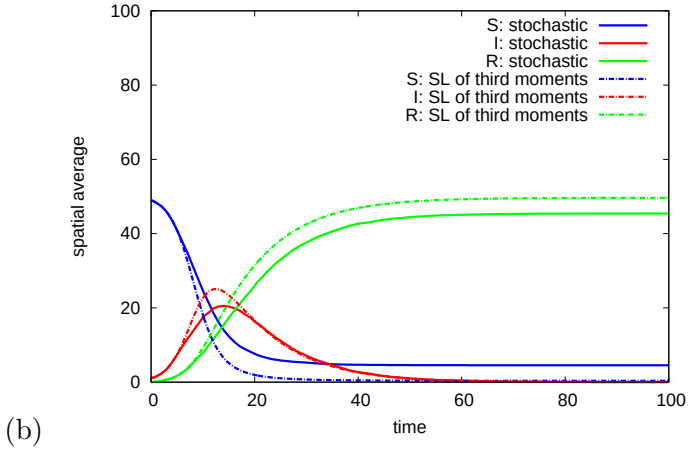
Figure 11: Covariance of S and I for different techniques: simulation of the SIR model on 12x12 grid with full connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, ${m_S}^{(ij)} = {m_I}^{(ij)} = 0.00001$

6 Conclusion

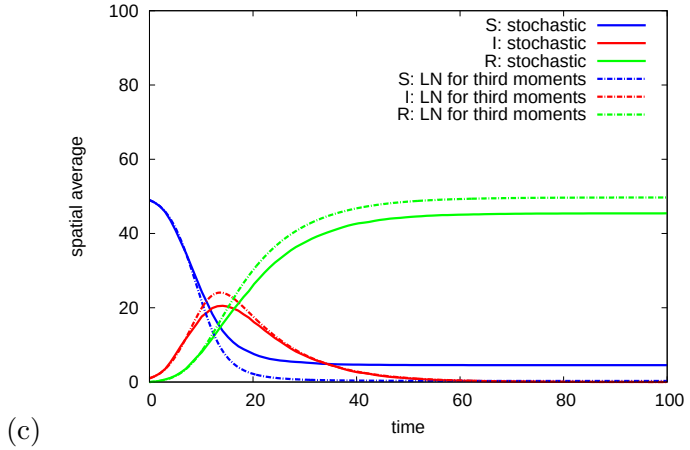
As was seen from the results, neither the deterministic approximation, the third moment stochastic linearisation approximation or the third moment log normal approximation (shown in Figure 12) give a good qualitative approximation, in the sense that the stochastic model indicates that not all susceptibles will succumb to the disease, where as these three approximations indicate the opposite. Whether susceptibles remain in the population has implications in the case of new infecteds entering the population and it would be preferable that an approximation could show this. In the case of the fourth moment stochastic linearisation (shown in Figure 13), of the various possibilities for factoring moments, the successful ones are demonstrated. Additionally, Figure 11 illustrates how well the covariance of S and I is approximated. Hence, we can conclude that the fourth moment stochastic linearisation provides the best results, though there is a dependency on the results on the specific way of factoring moments. This raises the question of how to choose which approximation to use when it is not possible to compare it with the results of the stochastic model and this requires further investigation.



(a)

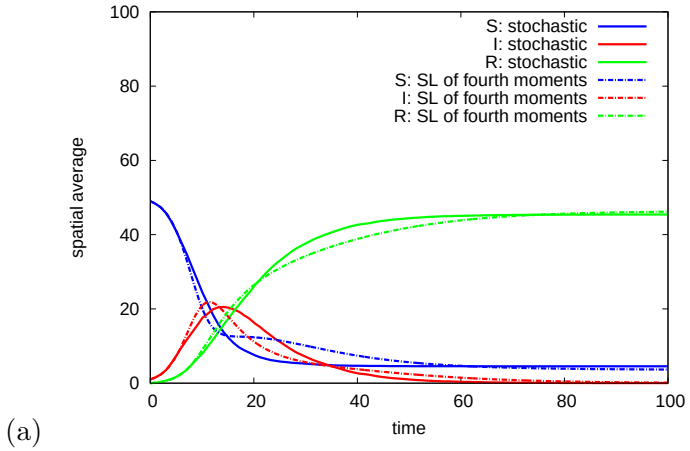


(b)

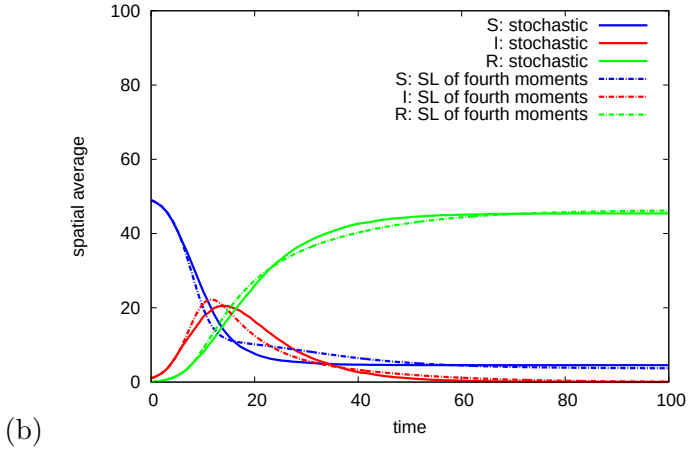


(c)

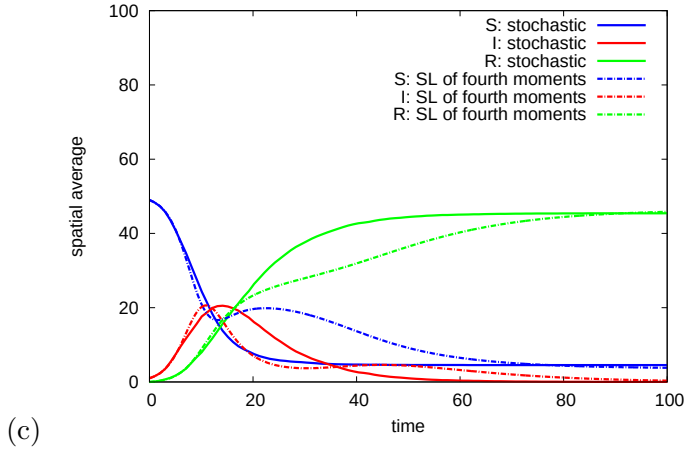
Figure 12: Stochastic versus (a) deterministic (b) third moment stochastic linearisation (c) third moment log-normal approximation: simulation of the SIR model on 12x12 grid with full connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, ${m_S}^{(ij)} = {m_I}^{(ij)} = 0.00001$



(a)



(b)



(c)

Figure 13: Stochastic versus three different fourth moment stochastic linearisation approximations using (a) $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle SI\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle SI\rangle]$ (b) $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle SI\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle S^2I\rangle]$ (c) using $\mathbf{E}[\langle S^2I\rangle]\mathbf{E}[\langle I\rangle]$, $\mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I^2\rangle]$ and $\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle S^2I\rangle]$: simulation of the SIR model on 12x12 grid with full connectivity with initial values $S^{(i)}(0) = 49, I^{(i)}(0) = 1, R^{(i)}(0) = 0$, and parameters $c^{(i)} = 0.011$, $h^{(i)} = 0.1$, $m_S{}^{(ij)} = m_I{}^{(ij)} = 0.00001$

References

[BP97] B. Bolker and S.W. Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, 52(3):179–197, 1997.

[BP99] B.M. Bolker and S.W. Pacala. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *The American Naturalist*, 153(6):575–602, 1999.

[Eng06] S. Engblom. Computing the moments of high dimensional solutions of the master equation. *Applied Mathematics and Computation*, 180(2):498–515, 2006.

[Gas15] N. Gast. The power of two choices on graphs: The pair-approximation is accurate? *SIGMETRICS Performance Evaluation Review*, 43(2):69–71, 2015.

[LD96] S.A. Levin and R. Durrett. From individuals to epidemics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1347):1615–1621, 1996.

[MMRL02] G. Marion, X. Mao, E. Renshaw, and J. Liu. Spatial heterogeneity and the stability of reaction states in autocatalysis. *Physical Review E*, 66(5):051915, 2002.

[MSH05] G. Marion, D.L. Swain, and M.R. Hutchings. Understanding foraging behaviour in spatially heterogeneous environments. *Journal of Theoretical Biology*, 232(1):127–142, 2005.

[PL97] S.W. Pacala and S.A. Levin. Biologically generated spatial pattern and the coexistence of competing species. In D. Tilman and P.M. Kareiva, editors, *Spatial ecology: the role of space in population dynamics and interspecific interactions*, pages 204–232. Princeton University Press, 1997.

Appendix: derivations for $\mathbf{E}[\langle S^2 I\rangle]$ and $\mathbf{E}[\langle SI^2\rangle]$

$$
\begin{aligned}
\delta(S_j^2 I_j) &= \big((S_j-1)^2(I_j+1) - S_j^2 I_j\big)(cS_jI_j) + \big(S_j^2(I_j-1) - S_j^2 I_j\big)(hI_j) + \\
&\quad \textstyle\sum_{k=1}^p \big(S_j^2(I_j-1) - S_j^2 I_j\big)(m_I I_j) + \sum_{k=1}^p \big(S_j^2(I_j+1) - S_j^2 I_j\big)(m_I I_k) + \\
&\quad \textstyle\sum_{k=1}^p \big((S_j-1)^2 I_j - S_j^2 I_j\big)(m_S S_j) + \sum_{k=1}^p \big((S_j+1)^2 I_j - S_j^2 I_j\big)(m_S S_k) \\
&= (-2S_jI_j + I_j + S_j^2 - 2S_j + 1)(cS_jI_j) + (-S_j^2)(hI_j) + \\
&\quad m_I \textstyle\sum_{k=1}^p (-S_j^2) I_j + m_I \sum_{k=1}^p (S_j^2) I_k + \\
&\quad m_S \textstyle\sum_{k=1}^p (-2S_jI_j + I_j) S_j + m_S \sum_{k=1}^p (2S_jI_j + I_j) S_k \\
&= -2cS_j^2I_j^2 + cS_jI_j^2 + cS_j^3I_j - 2cS_j^2I_j + cS_jI_j - hS_j^2I_j - \\
&\quad pm_IS_j^2I_j + pm_I\langle I\rangle S_j^2 - 2pm_SS_j^2I_j + pm_SS_jI_j + 2pm_S\langle S\rangle S_jI_j + pm_S\langle S\rangle I_j \\
&\Rightarrow\ 1/p \textstyle\sum_{j=1}^p (-2cS_j^2I_j^2 + cS_jI_j^2 + cS_j^3I_j - 2cS_j^2I_j + cS_jI_j - hS_j^2I_j - \\
&\quad 2pm_SS_j^2I_j - pm_IS_j^2I_j + pm_I\langle I\rangle S_j^2 + 2pm_S\langle S\rangle S_jI_j + pm_SS_jI_j + pm_S\langle S\rangle I_j) \\
&= -2c\langle S^2I^2\rangle + c\langle SI^2\rangle + c\langle S^3I\rangle - 2c\langle S^2I\rangle + c\langle SI\rangle - h\langle S^2I\rangle - \\
&\quad 2pm_S\langle S^2I\rangle - pm_I\langle S^2I\rangle + pm_I\langle S^2\rangle\langle I\rangle + 2pm_S\langle S\rangle\langle SI\rangle + pm_S\langle SI\rangle + pm_S\langle S\rangle\langle I\rangle \\
d\mathbf{E}[\langle S^2I\rangle]/dt &= -2c\mathbf{E}[\langle S^2I^2\rangle] + c\mathbf{E}[\langle SI^2\rangle] + c\mathbf{E}[\langle S^3I\rangle] - 2c\mathbf{E}[\langle S^2I\rangle] + c\mathbf{E}[\langle SI\rangle] - \\
&\quad h\mathbf{E}[\langle S^2I\rangle] + pm_S\mathbf{E}[\langle SI\rangle] + pm_S\mathbf{E}[\langle S\rangle\langle I\rangle] - \\
&\quad 2pm_S\mathbf{E}[\langle S^2I\rangle] - pm_I\mathbf{E}[\langle S^2I\rangle] + pm_I\mathbf{E}[\langle S^2\rangle\langle I\rangle] + 2pm_S\mathbf{E}[\langle S\rangle\langle SI\rangle] \\
&\approx -2c\mathbf{E}[\langle S^2I^2\rangle] + c\mathbf{E}[\langle SI^2\rangle] + c\mathbf{E}[\langle S^3I\rangle] - 2c\mathbf{E}[\langle S^2I\rangle] + c\mathbf{E}[\langle SI\rangle] - \\
&\quad h\mathbf{E}[\langle S^2I\rangle] + pm_S\mathbf{E}[\langle SI\rangle] + pm_S\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - \\
&\quad 2pm_S(\mathbf{E}[\langle S^2I\rangle] - \mathbf{E}[\langle S\rangle]\mathbf{E}[\langle SI\rangle]) - pm_I(\mathbf{E}[\langle S^2I\rangle] - \mathbf{E}[\langle S^2\rangle]\mathbf{E}[\langle I\rangle]) \\
&= -2c\mathbf{E}[\langle S^2I^2\rangle] + c\mathbf{E}[\langle SI^2\rangle] + c\mathbf{E}[\langle S^3I\rangle] - 2c\mathbf{E}[\langle S^2I\rangle] + c\mathbf{E}[\langle SI\rangle] - \\
&\quad h\mathbf{E}[\langle S^2I\rangle] + pm_S\mathbf{E}[\langle SI\rangle] + pm_S\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - \\
&\quad 2pm_S\mathbf{Cov}[S, SI] - pm_I\mathbf{Cov}[SS, I]
\end{aligned}
$$

$$
\begin{aligned}
\delta(S_jI_j^2) &= \big((S_j-1)(I_j+1)^2 - S_jI_j^2\big)(cS_jI_j) + \big(S_j(I_j-1)^2 - S_jI_j^2\big)(hI_j) + \\
&\quad \textstyle\sum_{k=1}^{p}\big(S_j(I_j-1)^2 - S_jI_j^2\big)(m_II_j) + \sum_{k=1}^{p}\big(S_j(I_j+1)^2 - S_jI_j^2\big)(m_II_k) + \\
&\quad \textstyle\sum_{k=1}^{p}\big((S_j-1)I_j^2 - S_jI_j^2\big)(m_SS_j) + \sum_{k=1}^{p}\big((S_j+1)I_j^2 - S_jI_j^2\big)(m_SS_k) \\
&= (2S_jI_j + S_j + I_j^2 - 2I_j - 1)(cS_jI_j) + (-2S_jI_j + S_j)(hI_j) + \\
&\quad m_I\textstyle\sum_{k=1}^{p}(-2S_jI_j + S_j)I_j + m_I\sum_{k=1}^{p}(2S_jI_j + S_j)I_k + \\
&\quad m_S\textstyle\sum_{k=1}^{p}(-I_j^2)S_j + m_S\sum_{k=1}^{p}(I_j^2)S_k \\
&= 2cS_j^2I_j^2 + cS_j^2I_j + cS_jI_j^3 - 2cS_jI_j^2 - cS_jI_j - 2hS_jI_j^2 + hS_jI_j - \\
&\quad 2pm_IS_jI_j^2 + pm_IS_jI_j + 2pm_I\langle I\rangle S_jI_j + pm_I\langle I\rangle S_j - pm_SS_jI_j^2 + pm_S\langle S\rangle I_j^2 \\
&\Rightarrow\ 1/p\textstyle\sum_{j=1}^{p}(2cS_j^2I_j^2 + cS_j^2I_j + cS_jI_j^3 - 2cS_jI_j^2 - cS_jI_j - 2hS_jI_j^2 + hS_jI_j - \\
&\quad 2pm_IS_jI_j^2 - pm_SS_jI_j^2 + pm_S\langle S\rangle I_j^2 + 2pm_I\langle I\rangle S_jI_j + pm_IS_jI_j + pm_I\langle I\rangle S_j) \\
&= 2c\langle S^2I^2\rangle + c\langle S^2I\rangle + c\langle SI^3\rangle - 2c\langle SI^2\rangle - c\langle SI\rangle - 2h\langle SI^2\rangle + h\langle SI\rangle - \\
&\quad 2pm_I\langle SI^2\rangle - pm_S\langle SI^2\rangle + pm_S\langle S\rangle\langle I^2\rangle + 2pm_I\langle SI\rangle\langle I\rangle + pm_I\langle SI\rangle + pm_I\langle S\rangle\langle I\rangle \\
d\mathbf{E}[\langle SI^2\rangle]/dt &= 2c\mathbf{E}[\langle S^2I^2\rangle] + c\mathbf{E}[\langle S^2I\rangle] + c\mathbf{E}[\langle SI^3\rangle] - 2c\mathbf{E}[\langle SI^2\rangle] - c\mathbf{E}[\langle SI\rangle] - \\
&\quad 2h\mathbf{E}[\langle SI^2\rangle] + h\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle S\rangle\langle I\rangle] - \\
&\quad 2pm_I\mathbf{E}[\langle SI^2\rangle] - pm_S\mathbf{E}[\langle SI^2\rangle] + pm_S\mathbf{E}[\langle S\rangle\langle I^2\rangle] + 2pm_I\mathbf{E}[\langle SI\rangle\langle I\rangle] \\
&\approx 2c\mathbf{E}[\langle S^2I^2\rangle] + c\mathbf{E}[\langle S^2I\rangle] + c\mathbf{E}[\langle SI^3\rangle] - 2c\mathbf{E}[\langle SI^2\rangle] - c\mathbf{E}[\langle SI\rangle] - \\
&\quad 2h\mathbf{E}[\langle SI^2\rangle] + h\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - \\
&\quad 2pm_I(\mathbf{E}[\langle SI^2\rangle] - \mathbf{E}[\langle SI\rangle]\mathbf{E}[\langle I\rangle]) - pm_S(\mathbf{E}[\langle SI^2\rangle] - \mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I^2\rangle]) \\
&= 2c\mathbf{E}[\langle S^2I^2\rangle] + c\mathbf{E}[\langle S^2I\rangle] + c\mathbf{E}[\langle SI^3\rangle] - 2c\mathbf{E}[\langle SI^2\rangle] - c\mathbf{E}[\langle SI\rangle] - \\
&\quad 2h\mathbf{E}[\langle SI^2\rangle] + h\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle SI\rangle] + pm_I\mathbf{E}[\langle S\rangle]\mathbf{E}[\langle I\rangle] - \\
&\quad 2pm_I\mathbf{Cov}[SI, I] - pm_S\mathbf{Cov}[S, I^2]
\end{aligned}
$$