Spatial and Memory Effects in Social Systems

Steven Kenney

School of Mathematics and Physics
University of Portsmouth

The thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth
Doctor of Philosophy

June 2020

Declaration

‘Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award’.

This thesis consists of 52 figures and has a final word count of 41,900.

Steven Kenney

June 2020

Acknowledgements

Firstly I would like to express my unending gratitude to my supervisors Dr. James Burridge and Dr. Michal Gnacik for their continuous support, motivation, knowledge and understanding when things have been challenging. Without your constant guidance this would not have been possible and I consider myself incredibly fortunate to have worked with you both.

I would like to thank my parents for a lifetime of support but specifically for giving me somewhere to stay during my studies. I could not have finished this without everything you have done for me.

Finally I would like to thank my partner, Jess. There isn't enough paper in the world to list all the reasons. For everything.

Abstract

In this thesis we study the statistical physics of social dynamics and spreading phenomena. We first demonstrate how the formation of regional dialects in birdsong can be related to the physical process of magnetic domain formation and coarsening. We show that the death rate of the species is analogous to the thermodynamic temperature in physical systems, and compute the death rate at which dialects cannot form due to a transition to a disordered state. We then investigate interface motion in a two dimensional lattice model of opinion spread. In our model sites have memories for the states of their neighbours and change state when this memory reaches a threshold. We consider the impact of surface roughening on the rate of spread and construct multiple approximation techniques to understand how varying the configuration of site thresholds can control the speed of the interface. We then investigate the motion of an opinion wave spreading via threshold dynamics in a one dimensional random network. We present a method for enumerating site connections and understanding what causes the wave to become arrested, and how long we expect to wait for this to occur. Our method makes use of martingale stopping sequences, and provides upper bounds on the time until the wave stops. We also develop a coarse-grained cluster method for estimating the expected stopping time. This is analogous to the well known (n,m) cluster approximation, but may be systematically derived from our dynamics. Finally we claim that this system of one dimensional spread is analogous to a branching and coalescing random walk process and we formulate mean field equations to describe this.

Table of contents

List of figures

List of tables

Chapter 1

Brief Introduction

Over recent decades, models and techniques from statistical physics [9–12] have been adapted and applied to study interdisciplinary problems in biology and social systems [13–18]. In these applications the atomic constituents of the model are not physical particles, but human or animal agents, and we are interested in the macroscopic behaviour of very large communities. Statistical physics is an appropriate tool to use in this context, because macroscopic behaviour is often insensitive to fine details of the microscopic interactions which drive it.

Many of these social phenomena that we can observe in the real world involve copying, imitating or influencing people and these conformist effects lead to areas or regions in space where groups of people have aligned behaviours. Such behaviours where these effects are noticable include dialects, word choices and even political opinions. Where two differently aligned groups meet there is very often a well defined border with an observable and sharp transition between behaviours as you cross it. These borders between regions of alignment are called interfaces and examples are presented in figures (1.1) and (1.2). In the first example there is a clear interface across the middle of England and as you cross that boundary the way in which the local population pronounces the word 'butter' significantly changes and likewise in the second example there are clearly defined regions in the United States where one choice of word for 'soda' is preferred and sharp transitions between them.

What makes this subject so interesting is that these interfaces are not exclusive to humans. Figure (1.3) shows the formation of yellowhammer dialects in the Czech Republic and once again there are clear regions in which one song is incredibly dominant. As explained above, the fact that these interfaces arise in humans and animals suggests that there are some universal explanations that go beyond the microscopic behaviours and tendencies of individual agents and thus, a statistical physics approach is justified.

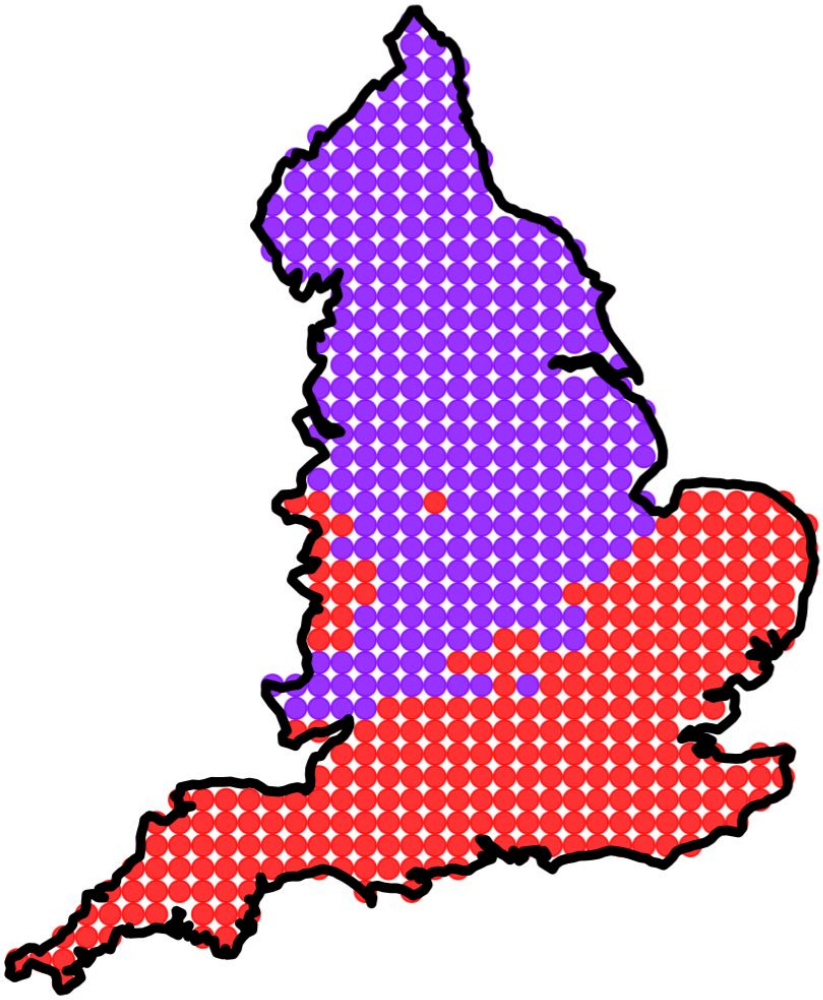


Fig. 1.1 A depiction of the local pronunciation for the word butter across England. Purple circles represent an ‘oo’ vowel sound such as the one that in the words hood and wood whereas red circles represent a defined ‘u’ sound such at the one in the words cut or bun.

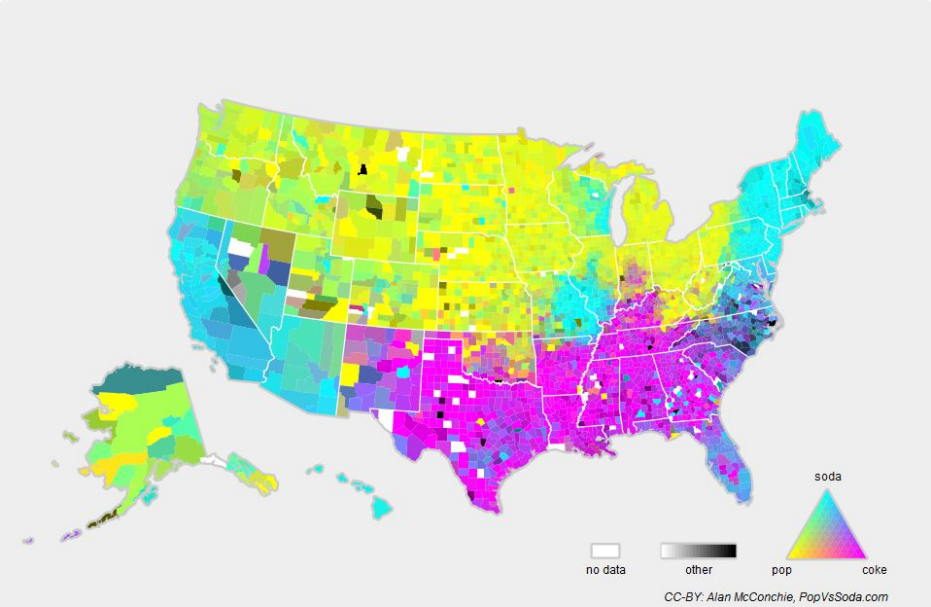


Fig. 1.2 A map of the United States of America highlighting the difference in word choices for the same item; in this case a carbonated beverage [1].

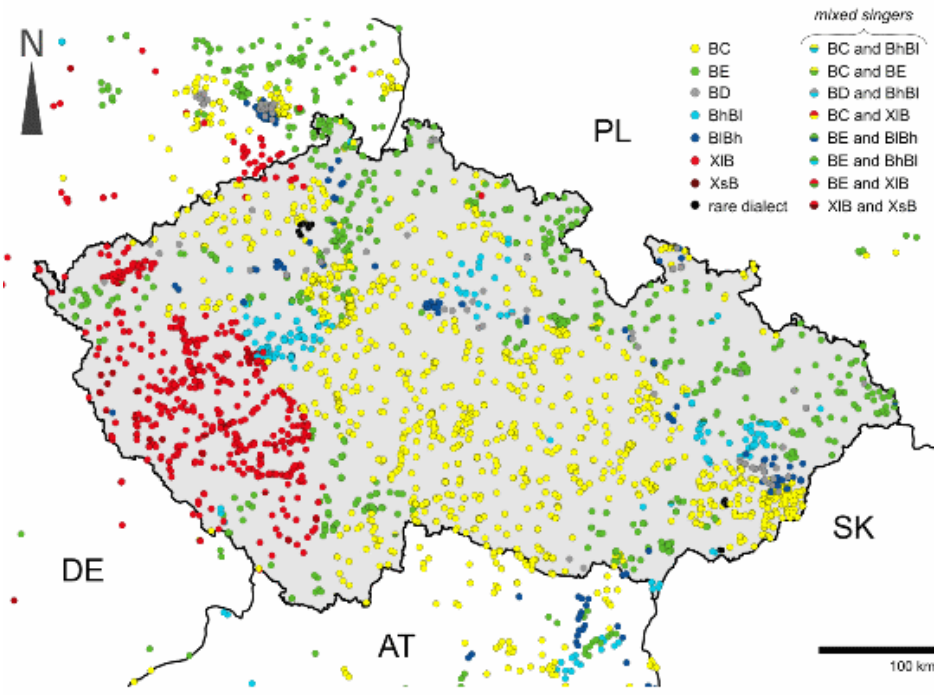


Fig. 1.3 A volunteer map for much of the Czech territory with different colour circles representing different dialects with some clear domain formations [2].

There are many questions that one can ask about these interfaces and the regions which they separate. These could include asking how interfaces form assuming an originally randomised setup, where the interfaces are likely to form in space and how these interfaces move over time. Throughout the body of this thesis we attempt to address these questions by utilising a variety of approaches and techniques.

Our work is divided into three parts, each of which studies the common theme of interfaces between domains in a different context.

Domain Walls and Birdsong

A major area of study in statistical physics is the existence of phase transitions from spatially disordered states, to stable and ordered domains [12, 13, 19]. Our first focus in this thesis is the existence and dynamics of domain walls in the social systems of birds, which are known to have regional dialects [4, 20, 21]. We make an analogy between dialect boundaries and domain walls, and show that the dialect patterns of a particular species - the Puget Sound White Crowned Sparrow [6] - are analogous to the formation of stripe states in high aspect ratio physical systems [22]. We achieve this by viewing the Sparrow's processes of song learning as analogous to spin alignment in the Ising model. By modelling the reproductive behaviour of the birds, along with the majority rule process of song selection, we can calculate analytical estimates for the critical death rate at which dialects can appear. We also show that the spatial structure of the dialect domains observed along the Pacific coast are predictable from the geometry of their habitat zone [23].

Motion of Interfaces

Our second focus is on the motion of interface fronts [24–26] in systems where an opinion or behaviour is spreading through a population of agents. These agents have fixed length memories for the opinions of their neighbours, and a threshold at which they change state [17, 27]. Interface fronts have been used when investigating spreading effects in multiple fields ranging from models in physics to sociology and biology. While our first focus was primarily regarding stability and stationary behaviour, in this second part of the project we are interested in the relationship between interface shape and speed, and the effects of disorder. We examine the impact of wave front shapes and surface roughening on the surface area of an interface, and investigate how this increases the spreading potential. We observe these effects on lattice structures with differing thresholds and resistances and provide a series of analytical approximations for the wave speed. Our most successful technique approximates

interface motion and shape on a disordered lattice by comparing it to the interface in a system formed from a single repeating tile.

Motion in One-Dimension

Our final focus is on opinion spread through random networks in one dimensional systems. This primarily regards information transfer between nodes and the core of our work revolves around calculations of expected travel distance. We adapt martingale techniques used to study stopping sequences in coin toss experiments to study the behaviour of network configurations. Using this martingale approach we calculate the expected distance an opinion will travel through the random network until connection density becomes too sparse for it to continue. These distances may also be viewed as the sizes of percolation clusters [9, 16]. We also derive a coarse grained clustering approach in order to reduce connection complexity and find that our method results in the (n,m) cluster approximation [28]. This model has some clear connections to percolation theory [9]; although the system does not strictly possess a percolation transition because it is one dimensional, it does exhibit percolation-like effects due to the quasi-two dimensional nature of the system when interactions become longer range.

Chapter 2

Preliminaries

In this chapter we introduce some fundamental models and concepts in statistical physics that we utilise later in our research. We begin by introducing the Ising model, a classic example system for understanding phase transitions which we refer back to in our research on the formation of bird song dialects in chapter (3). We outline the basics of percolation theory which describes the behaviour of connected clusters in a random graph which pertains to all chapters of our research. Lastly we introduce integrodifference equations, a technique used to model the spread of an invading wave when reproduction and dispersal occur at discrete intervals, which relates to our final work in chapter (5).

2.1 Ising Model

The Ising model is a mathematical model of ferromagnetism in statistical physics, created by Wilhelm Lenz and named after his student Ernst Ising, that uses discrete variables to represent magnetic dipoles of atomic spins in one of two states. In particular, it is worth mentioning that Ising has solved the model in one dimension in his thesis [29]. This Ising model is a classic statistical physical model due to its applications and comparisons to real world behaviour and is a well researched and known area [9, 30]. For a simple exposition on the Ising model we refer the reader to [10]. The spins are placed into a structure and allowed to interact with their neighbours. The two dimensional Ising model is one of the simplest models that shows a phase transition in the absence of an external field as the temperature of the system is raised or lowered around a critical point. In the Ising model we calculate the partition function, from which we can calculate other relevant quantities. Many of the early results we derive are not directly utilised in the main body of the thesis, however some of the results that they lead to are fundamental.

We begin by looking at a two dimensional lattice comprised of $N = L \times L$ sites, each with spin $s_i \in \{-1, 1\}$, with $i \in \{1, 2, \ldots, N\}$. In the case of a ferromagnet this would represent the magnetic states of atoms at locations on the crystal structure but can be used to represent many other things in different fields. The spin at any given site can be interpreted as an alignment direction in space. We are concerned with the relationship between spins at neighbouring positions i and j, with the pairing denoted as $\langle i, j \rangle$, that can interact with one another. An energy is then assigned to each pairing in the system, either $-J$ for a pair of the same spins or J for a pair of opposing spins, with $J > 0$, and thus the total internal energy in the system is

$$E_{\text{int}} = -J \sum_{\langle i,j \rangle} s_i s_j \tag{2.1}$$

with the sum running over all neighbouring pairs in the entire system.

The Ising model can also have an external field, H, applied to every spin in the system. Similarly to the internal interactions, a spin of opposing direction to the field has energy of $|H|$ and a spin of the same direction has energy $-|H|$. Therefore the energy for each spin due to the external field is $-Hs_i$ for $i \in \{1, 2, \ldots, N\}$ and the total energy contributed to the system from the external field is

$$E_{\text{ext}} = -H \sum_{i=1}^{N} s_i. \tag{2.2}$$

We can therefore combine the internal and external energy contributions to get the total energy for a system of N spins, that is, a Hamiltonian that we denote by

$$E_{\text{tot}} = \mathscr{H} = -H \sum_{i=1}^{N} s_i - J \sum_{\langle i,j \rangle} s_i s_j. \tag{2.3}$$

From these definitions of the model it is evident that for the system to lower the internal energy the neigbouring spins would need to align parallel. We discover the behaviour of the system as a function of temperature, $T > 0$ which defines an energy scale $k_B T$, where k_B is the Boltzmann constant. In a model with no external field, $H = 0$, and low temperature, $k_B T \ll J$, the neighbouring spin interactions are relatively strong and the system will align parallel to reduce the system's energy. Conversely in the case of high temperature, $k_B T \gg J$ the neighbouring spins are relatively weak, and thus the system effectively has random spin orientations.

We use the bold $\mathbf{s}$ to denote $\{s_1, s_2, \ldots, s_N\}$, the microscopic orientation of all sites at a given time. The temporal mean of an observable is a weighted ensemble average, $\langle A \rangle$, over all possible microstates. We let $p_{\mathbf{s}}$ denote the probability of the lattice being in a given state $\mathbf{s}$ with observable $A_{\mathbf{s}}$ having ensemble average

$$\langle A \rangle = \sum_{\mathbf{s}} p_{\mathbf{s}} A_{\mathbf{s}}, \tag{2.4}$$

where the sum runs over all 2^N configurations.

In the canonical ensemble, the system is in equilibrium with fixed volume and temperature. The probability, $p_{\mathbf{s}}$, to find the system in this configuration with total energy $E_{\mathbf{s}}$ is given by the Boltzmann distribution [31], namely

$$p_{\mathbf{s}} = \frac{\exp(-\beta E_{\mathbf{s}})}{\sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}})} \tag{2.5}$$

and thus the ensemble mean of observable $A_{\mathbf{s}}$ is

$$\langle A \rangle = \frac{1}{Z} \sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}}) A_{\mathbf{s}} \tag{2.6}$$

where the denominator is the partition function

$$Z(T, H, N) = \sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}}), \tag{2.7}$$

and $\beta = \frac{1}{k_B T}$ is the inverse temperature. The partition function is key to understanding the nature of the overall system as many fundamental properties and behaviours of the system can be calculated from it. The partition function acts as a weighted average over all microstates and links the macroscopic behaviour of the system to the microscopic behaviour of the sites.

Some examples of observable quantities are the total magnetism of the system

$$M_{\mathbf{s}} = \sum_{i=1}^{N} s_i \tag{2.8}$$

leading to the average total magnetism

$$\langle M \rangle = \frac{1}{Z} \sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}}) M_{\mathbf{s}}. \tag{2.9}$$

It follows that the average total energy is given by

$$\langle E \rangle = \frac{1}{Z} \sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}}) E_{\mathbf{s}}. \tag{2.10}$$

Thermodynamic entropy can be understood as a measure of the number of possible microstates of a system in thermodynamic equilibrium that lead to the same macrostate. That is different microscopic configurations that have the same average magnetism, average energy per spin and any other observables. The Gibbs entropy formula, developed by J Willard Gibbs [32], states that for a system in a canonical ensemble the Gibbs entropy S can be expressed as

$$S = -k_B \sum_r p_r \ln p_r \tag{2.11}$$

and using the Boltzmann distribution from equation (2.5) it follows that

$$S = k_B \ln Z + \langle E \rangle / T. \tag{2.12}$$

The fundamental thermodynamic relation states that the free energy of a system, F, is

$$F = \langle E \rangle - TS. \tag{2.13}$$

From equations (2.12) and (2.13) we can deduce that

$$F(T,H) = -k_B T \ln Z. \tag{2.14}$$

This total free energy is an extensive property of the system because it is proportional to the numbers of nodes in the lattice. We would prefer to work with intensive quantities, those that are unaffected by the scaling of the system, so we look at $f = F/N$ the free energy per spin, $m = \langle M \rangle / N$, the average magnetism per spin and $e = \langle E \rangle / N$, the average energy per spin. These intensive quantities can be calculated using only the free energy equation and the partition function which is why the partition function is so important.

As an example the average energy per spin can be shown to be

$$e(T,H) = f - T \left(\frac{\partial f}{\partial T} \right)_H, \tag{2.15}$$

where the subscript denotes that f is still a function of H and where the external field H is fixed. The proof of this is as follows

$$\begin{aligned} f - T\left(\frac{\partial f}{\partial T}\right)_H &= -\frac{1}{N}k_B T \ln Z + \frac{T}{N}\left(k_B \ln Z + k_B T \frac{1}{Z}\frac{\partial}{\partial T}Z\right) \\ &= \frac{k_B T^2}{NZ}\frac{\partial}{\partial T}Z \\ &= \frac{k_B T^2}{k_B T^2 NZ}\sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}}) E_{\mathbf{s}} \\ &= \frac{\langle E \rangle}{N}. \end{aligned} \tag{2.16}$$

Using the same method we can also show that

$$m(T,H) = -\left(\frac{\partial f}{\partial H}\right)_T, \tag{2.17}$$

with similar equations existing for entropy per spin and sensitivity per spin to both changes in external field and temperature.

In our defined system the free energy will be an analytical function and well-defined everywhere because N is finite and thus the partition function is made up of 2^N different terms, one for each possible permutation of the system. However in the thermodynamic limit as $N \to \infty$ there is no guarantee that the free energy will remain analytic and thus at certain temperatures and external fields the thermodynamic properties may be undefined and the system can undergo a phase transition. We will show by contradiction that in order to have a phase transition you must have neighbour interactions.

We will assume that our system has no nearest neighbour interactions, that is to say, that $J = 0$ in our energy expression, and the system is dependent on H the external field. For this system of non-interacting spins

$$E_{\text{tot}} = -H\sum_{i=1}^{N} s_i. \tag{2.18}$$

From equation (2.13) we see that average energy $\langle E \rangle$ and entropy S both contribute to the free energy of the system and the ratio of these contribution is $\langle E \rangle / TS \propto H\beta$, where $\propto$ means proportional to. From the second law of thermodynamics we get the principle of minimum energy, stating that the free energy of a system is minimised as it reaches equilibrium. We can then conclude that if $H\beta \to 0$ then the entropy term is dominant and entropy must be

maximised in order for free energy to be minimised, and thus the microstates consists of all the spins having random orientations with average energy and magnetisation being zero. Conversely if $H\beta \to \infty$ then the average energy term is dominant and in order for the free energy to be minimised we need the total energy in equation (2.18) to be minimised. This consists of two options, either all spins up or all spins down depending on the orientation of H. In this case $\langle M \rangle = \pm N$ and $\langle E \rangle = -N|H|$.

Since all of the thermodynamic quantities can be calculated using the partition function we begin by calculating

$$\begin{aligned} Z(T,H,N) &= \sum_{\mathbf{s}} \exp(-\beta E_{\mathbf{s}}) \\ &= \sum_{\mathbf{s}} \exp(\beta H \sum_{i=1}^{N} s_i) \\ &= \sum_{\mathbf{s}} \exp(\beta H s_1) \exp(\beta H s_2) \cdots \exp(\beta H s_N) \\ &= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} \exp(\beta H s_1) \exp(\beta H s_2) \cdots \exp(\beta H s_N) \\ &= \sum_{s_1=\pm 1} \exp(\beta H s_1) \sum_{s_2=\pm 1} \exp(\beta H s_2) \cdots \sum_{s_N=\pm 1} \exp(\beta H s_N) \\ &= (\exp(\beta H) + \exp(-\beta H))^N \\ &= (2\cosh \beta H)^N. \end{aligned} \tag{2.19}$$

This calculation is made simple by the absence of site interactions. Now that we have the partition function we can use equation (2.14) to calculate the free energy per spin in this system

$$\begin{aligned} F(T,H,N) = -\frac{1}{\beta} \ln(2\cosh \beta H)^N = -N \frac{1}{\beta} \ln(2\cosh \beta H) \\ f(T,H) = -\frac{1}{\beta} \ln(2\cosh \beta H) \end{aligned} \tag{2.20}$$

which is clearly analytic everywhere because $\cosh x \geq 1$ for all $x \in \mathbb{R}$. Due to this we can conclude that this model cannot undergo a phase transition for any values of H and T. Thus, by contradiction, in order to have a phase transition occur there must be interactions between spins, $J \neq 0$. For this non-interacting model the average magnetisation per spin can be

calculated using equation (2.17) as follows

$$m(T,H) = -\Big(\frac{\partial f}{\partial H}\Big)_T = \frac{1}{\beta}\frac{2\sinh\beta H}{2\cosh\beta H}\beta = \tanh\beta H. \tag{2.21}$$

From this we can see that for $T > 0$ the average magnetisation per spin can only be 0 at $H = 0$. As $T \to 0$ we know that $\beta \to \infty$ and the tanh function approaches a step function from -1 to 1. From this we can deduce there is no spontaneous magnetisation with no external field in a system of non-interacting spins. This outcome is intuitive from the basic description of the system but it's important to be able to prove in order to move on to more complicated energy models.

The next system to consider is one in which there is no external field applied and the only component of the system energy is the interaction of neighbouring spins. As before the system can be viewed as a competition to lower the free energy between thermal energy, leading to entropy and a disordered system, and the interaction energy, J, attempting to align neighbouring spins. Similarly to the system of non-interacting spins the behaviour is strongly dependent on the relative contribution of each term to the minimisation of the free energy equation, $\langle E\rangle/TS \propto J\beta$. Thus at high temperature, $J\beta \ll 1$, the free energy is minimised by maximising systemic entropy, leading to a disordered system of randomised spin orientations and conversely for low temperature systems $J\beta \gg 1$ and free energy is minimised by minimising the interaction energy in the system, causing the spins to align orientations with their neighbours. At low temperature this system will be in an ordered ferromagnetic configuration and at high temperature this system will be in a disordered paramagnetic state and it can be shown that there is a critical temperature at which the free energy contributions from the entropic term and the total energy term are relatively even and we undergo a sudden phase transition as we pass through this temperature [9, 12].

In order to investigate this critical temperature we introduce Glauber dynamics and the detailed balance condition. In the Glauber model, created by Roy J Glauber in 1963 [33], spins are selected one at a time and have a probability of changing their spin state dependent on the energy change this would cause. There are three different types of change; energy neutral, energy lowering and energy increasing and we need to fix the rates for these. For this we need to use the concept of detailed balance for a system in equilibrium which states that

$$p_{\mathbf{s}} w_j(\mathbf{s}) = p_{\mathbf{s}^j} w_j(\mathbf{s}^j) \tag{2.22}$$

where $\mathbf{s}^j$ is the state of the system derived from $\mathbf{s}$ with the spin at site j flipped and $w_j(\mathbf{s})$ is the transition rate from $\mathbf{s}$ to $\mathbf{s}^j$. These rates are continuous and non-zero as even the unfavourable state flips can occur. If we are considering a system of interacting spins with no external field then our Hamiltonian from (2.3) is

$$\mathscr{H} = -\sum_{\langle i,j \rangle} J_{i,j} s_i s_j \tag{2.23}$$

where the couplings of sites i, j can be ferromagnetic or antiferromagnetic, the strengths of the bonds can vary and neighbours can be different from site to site. Equation (2.22) gives us

$$\frac{w_i(\mathbf{s})}{w_i(\mathbf{s}^i)} = \frac{p_{\mathbf{s}^i}}{p_{\mathbf{s}}} = \frac{\exp(-\beta s_i \sum J_{i,j} s_j)}{\exp(\beta s_i \sum J_{i,j} s_j)} = \frac{1 - s_i \tanh\left(\beta \sum_{j \in \langle i \rangle} J_{i,j} s_j\right)}{1 + s_i \tanh\left(\beta \sum_{j \in \langle i \rangle} J_{i,j} s_j\right)} \tag{2.24}$$

where the sum is over the nearest neighbours of i , $\langle i \rangle$. The simplest set of rates satisfying this equation is

$$w_i(\mathbf{s}) = \frac{1}{2}\left[1 - s_i \tanh\left(\beta \sum_{j \in \langle i \rangle} J_{i,j} s_j\right)\right], \tag{2.25}$$

the rate proposed by Glauber. This is a continuous time Markov chain so it is a rate, however in our modelling we have to work with discrete them so we will interpret this as choosing a site at random and allowing it to flip orientation with probability

$$p_{\text{flip}} = \frac{1}{2}\left(\tanh \frac{\beta \Delta E}{2}\right), \tag{2.26}$$

where ΔE is the systemic energy change if that site were to flip. As expected energy minimising flips are favoured and energy neutral flips have a probability of exactly a half.

In order to investigate the critical temperature at which a phase transition occurs we introduce the mean field approximation which replaces the local interactions between spins with a representative field. This allows is to rewrite the Hamiltonian from equation (2.3) as

$$\mathscr{H} = -\sum_i h_i s_i \qquad \text{where} \qquad h_i = \frac{1}{2} J \sum_{j \in \langle i \rangle} s_j \tag{2.27}$$

so we can view the system as each spin being effected by a local inhomogenous field emanating from its neighbours, j. The approximation occurs when we replace the neighbouring spins by their mean value, the average local magnetisation m_{loc}, so the local field h_i can be

replaced with the homogenous mean field $h = Jzm_{\text{loc}}/2$ where z is the lattice coordination number, the number of neighbours a site has. This removes the direct interaction between linked neighbours and approximates it with a system of independent spins with the significant change that neighbouring spins are now uncorrelated. This mean field approach is only exact in case where every pair of sites in the system interact; in the case of the Ising model that is the complete graph of $N(N-1)/2$ links on N sites. We want the system energy to scale linearly with the system size so we choose the interaction strength to be inversely proportional to N giving us the Hamiltonian

$$\mathscr{H} = -\frac{1}{N}\sum_{i<j} s_i s_j \tag{2.28}$$

and the local field is the same as the magnetisation for all system sizes. The Glauber dynamics for this give us the site transition rate for a spin to flip of

$$w_i = \frac{1}{2}\left[1 - s_i \tanh\left(\frac{\beta}{N}\sum_{j\neq i} s_j\right)\right]. \tag{2.29}$$

When N is sufficiently large, we can replace s_i with the average spin S_i and the hyperbolic term becomes $\tanh\beta m$. Given that when a spin s_i flips it changes from 1 to -1 or vice versa, the sum change to the system is $-2s_i$ giving us the equation of motion for the average spin, $dS_i/dt = -2s_i w_i$. It follows that

$$\frac{dS_i}{dt} = -2s_i w_i = -2\left(\frac{1}{2}s_i - \frac{1}{2}s_i^2 \tanh\beta m\right) = -S_i + \tanh\beta m \tag{2.30}$$

and finally summing this equation over all sites we know that the average magnetisation must satisfy

$$\frac{dm}{dt} = -m + \tanh\beta m. \tag{2.31}$$

The equilibrium solution has $\frac{dm}{dt} = 0$ and as such there is an evident phase transition at the critical β value of $\beta_c = 1$. For $\beta < 1$ the only solution to this equation is $m = 0$, corresponding to the previous result that a system with high temperature leads to randomised spin orientations and zero average magnetisation. For $\beta \gg 1$ the equation has two solutions at $m = \pm 1$ coinciding with the result that a system at low temperature has spin alignment, either all up or all down. It is evident that a phase transition is occurring at this critical

temperature since as we transition through it we go from a randomised system to an aligned system.

For the two-dimensional Ising model with only nearest neighbour interactions and no external field the critical temperature is also a solved problem. First Rudolf Peierls showed in 1936 that there exists a phase transition by proving that at sufficiently low temperatures the amount of boundaries between up and down spins must be low enough that the majority of the system is aligned [34]. Then in 1944 Lars Onsager produced an analytical solution for the two-dimensional Ising model [35], showing the critical temperature to take exact value

$$\frac{k_B T_c}{J} = \frac{2}{\ln\left(1+\sqrt{2}\right)} \approx 2.269. \tag{2.32}$$

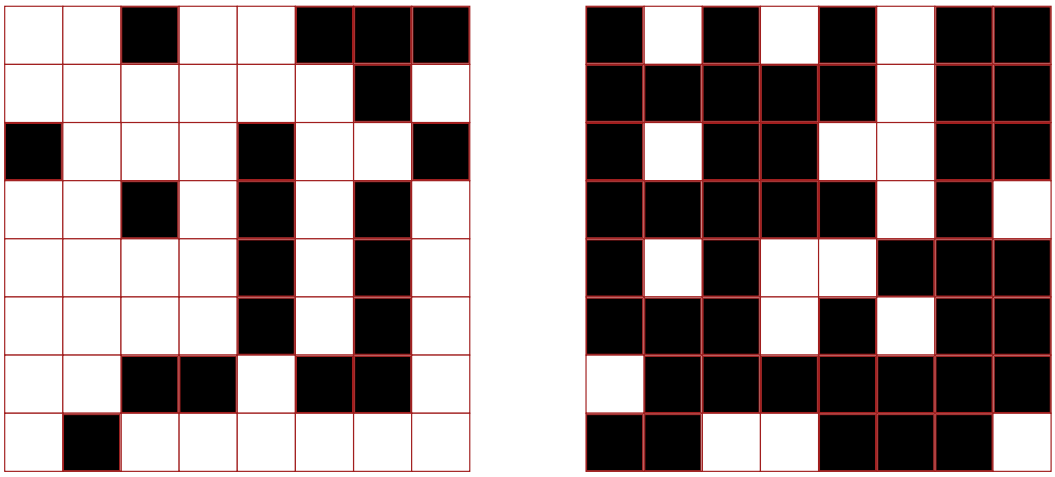
The existence of a critical temperature and phase transitions is an important feature that we use in later work.

2.2 Percolation

In this section we outline the concept of percolation and highlight some important definitions and results. Percolation is related to the existence of spanning clusters across a system and is a key idea in the research of spreading phenomena. Percolation is linked to the existence of stripe states in Ising model [22] systems which we observe in chapter (3) and our research in chapter (5) is similar to investigating percolation in a one dimensional system.

2.2.1 Percolation in One Dimension

In structured systems where sites can be in multiple states you often see the formation of clusters in which neighbouring sites are aligned. In many fields, such as epidemic spread [36] and networking and cascading effects [17], it is important to investigate these clusters and their geometrical properties including size, shape and long range connectivity. Percolation theory is the study of these connected clusters and their behaviours, allowing us to investigate many phenomena on random graphs. Percolation theory has been thoroughly investigated and there exists a plethora of literature on the subject [37, 38]. The first example we look at is site percolation on an $L \times L$ lattice where each site is either in the up state with probability p or in the down state with probability $q = 1 - p$. We define a cluster on this lattice as a group of nearest neighbour occupied sites in the same aligned state and let cluster size s be the number of sites in a given cluster. Intuitively with $p = 0$ the lattice has no sites in the up



(a) p = 0.3 (b) p = 0.7

Fig. 2.1 Percolation on a 2-D lattice with height and width of 8 units. For the subcritical p value we see a largest cluster size of 5 whereas we see a percolating cluster for the supercritical p value

state and as such there are no clusters of aligned up states and with $p = 1$ every site is up and there is one up cluster of size $s = L^2$. As expected the largest cluster in a given system increases with p as we go through the range $0 < p < 1$. A cluster is considered percolating if it has infinite size. Therefore in a confined $L \times L$ system only clusters that span from either left edge to right or top edge to bottom can possibly be percolating if the system was infinite in size. Therefore at $p = 1$ the sole cluster is percolating and at $p = 0$ there is no cluster, let alone a percolating one and it follows that there may exist a critical p, p_c, at which a percolating cluster begins to exist.

We begin considering percolation in one dimension where calculations can be made exactly. The first consideration in a one dimensional system of length L is that of the boundaries, as the sites on each end of the system have only one neighbour as opposed to the normal two. For an up cluster to exist in the middle of the system it requires a down site on both sides whereas at the boundary it only requires one down neighbour; however as we increase L the proportion of boundary sites to total sites tends to zero and their impact diminishes, eventually becoming non existent for an infinite system and we can ignore their impact. The first property we can compute is the probability that a given site i belongs to a cluster of size s, denoted by $p_{i \in C_s}$. The site must be part of s consecutive up sites, bound at both ends by down sites, and the site in question can be in any position in any of the s

positions within that cluster giving us

$$p_{i\in C_s} = s(1-p)^2 p^s. \tag{2.33}$$

Next we can calculate the frequency of clusters of a given size in a system. There are L sites in the system but with a cluster size of s we only have to apply the equation (2.33) on L/s sites to be certain of the frequency of s sized clusters and as such the frequency of s sized clusters is

$$N(s,p,L) = L(1-p)^2 p^s. \tag{2.34}$$

However this is an undesirable quantity because it is dependent on the system size and we want to consider the system becoming infinite in size. As such we can normalise by the number of sites in the system to instead get the cluster size density

$$n(s,p) = (1-p)^2 p^s. \tag{2.35}$$

In these equations $sn(s,p)$ is the probability that a given site belongs to an s sized cluster. From equation (2.35) it is clear that cluster size density $n(s,p)$ is strictly decreasing with increases in size s, thus for any fixed value of p larger clusters become rarer. As you can see from figure (2.2), for all differing values of p the curve begins decreasing slowly for a large range of s before dropping sharply for some characteristic value s_ξ which increases as p approaches 1. We rewrite $n(s,p)$ as

$$\begin{aligned} n(s,p) &= (1-p)^2 p^s \\ &= (1-p)^2 \exp(\ln p^s) \\ &= (1-p)^2 \exp(s \ln p) \\ &= (1-p)^2 \exp(-s/s_\xi), \end{aligned} \tag{2.36}$$

where we introduce the characteristic cluster size

$$s_\xi(p) = -\frac{1}{\ln p} \to \frac{1}{1-p} \text{ for } p \to 1^-. \tag{2.37}$$

The characteristic cluster size, shown in figure (2.3) is the typical largest cluster in a system and is an increasing function of p which diverges as p approaches 1.

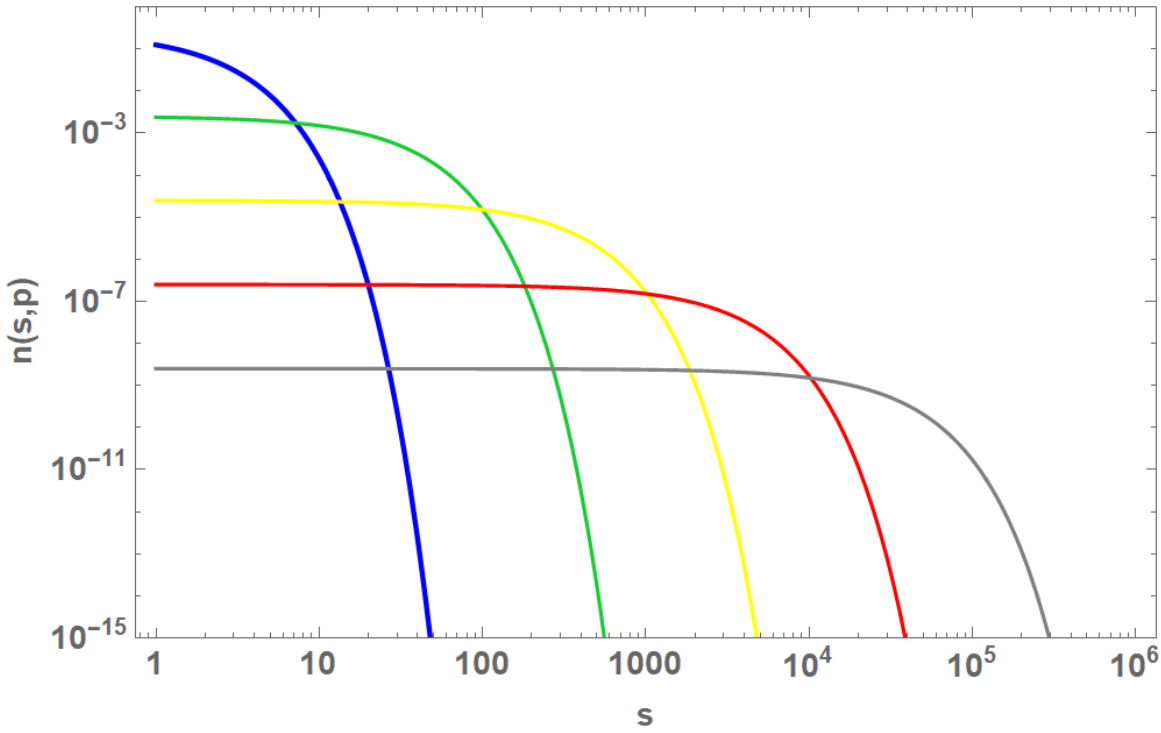


Fig. 2.2 A graph showing the cluster density against cluster size. The blue line represents $p = 0.5$, green represents $p = 0.95$, yellow represents $p = 0.995$, red represents $p = 0.9995$ and grey represents $p = 0.99995$

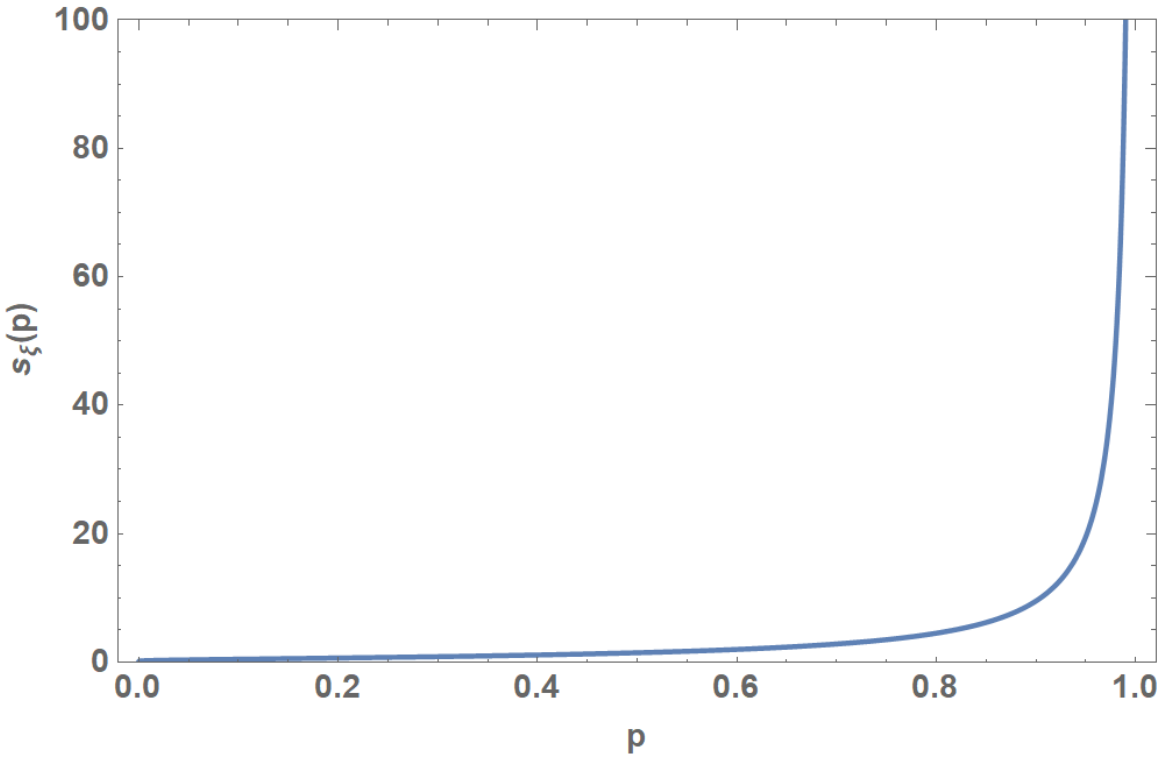


Fig. 2.3 A graph showing the characteristic cluster size against p. As p approaches 1 the characteristic cluster size diverges and as p approaches 0 the characteristic cluster size tends to 0.

After calculating the characteristic cluster size the next quantity to consider is the average cluster size, which we can gauge by considering what is the expected cluster size that a random site with specific spin belongs to. The probability of choosing a specific site with a chosen spin state is $1/N_{\text{occ}}$ where N_{occ} is the number of sites with chosen spin state. The probability of a chosen site belonging to a specific s sized cluster is then s/N_{occ}. Using this we can weight the clusters based upon their size and deduce that the average size of a cluster that a random site of specific spin belongs to is

$$\chi(p) = \frac{1}{N_{\text{occ}}} \sum_{k=1}^{N_{\text{clu}}} s_k^2. \tag{2.38}$$

In order to work with this quantity we can relate it to the cluster number density, $n(s,p)$, defined previously. The quantity $\chi(p)$ is a sum running over all clusters in a system but in order to convert it in terms of cluster number density we instead want to view it as a sum over all cluster sizes and the frequency at which those cluster sizes occur as follows

$$\chi(p) = \frac{1}{N_{\text{occ}}} \sum_{k=1}^{N_{\text{clu}}} s_k^2 = \frac{1}{N_{\text{occ}}} \sum_{s=1}^{\infty} s^2 N(s,p,L). \tag{2.39}$$

The expected number of occupied sites within an L sized lattice is simply pL and the probability of an arbitrary site being part of any finite sized cluster is just p for all $p < 1$. Since $sn(s,p)$ is the probability that an arbitrary site belongs to a cluster of size s then the sum over all s values of $sn(s,p)$ must be equal to p. Using this fact and equations (2.34) and (2.35) we can show that

$$\begin{aligned} \chi(p) &= \frac{\sum_{s=1}^{\infty} s^2 N(s,p,L)}{pL} \\ &= \frac{\sum_{s=1}^{\infty} s^2 n(s,p)}{p} \\ &= \frac{\sum_{s=1}^{\infty} s^2 n(s,p)}{\sum_{s=1}^{\infty} s n(s,p)}. \end{aligned} \tag{2.40}$$

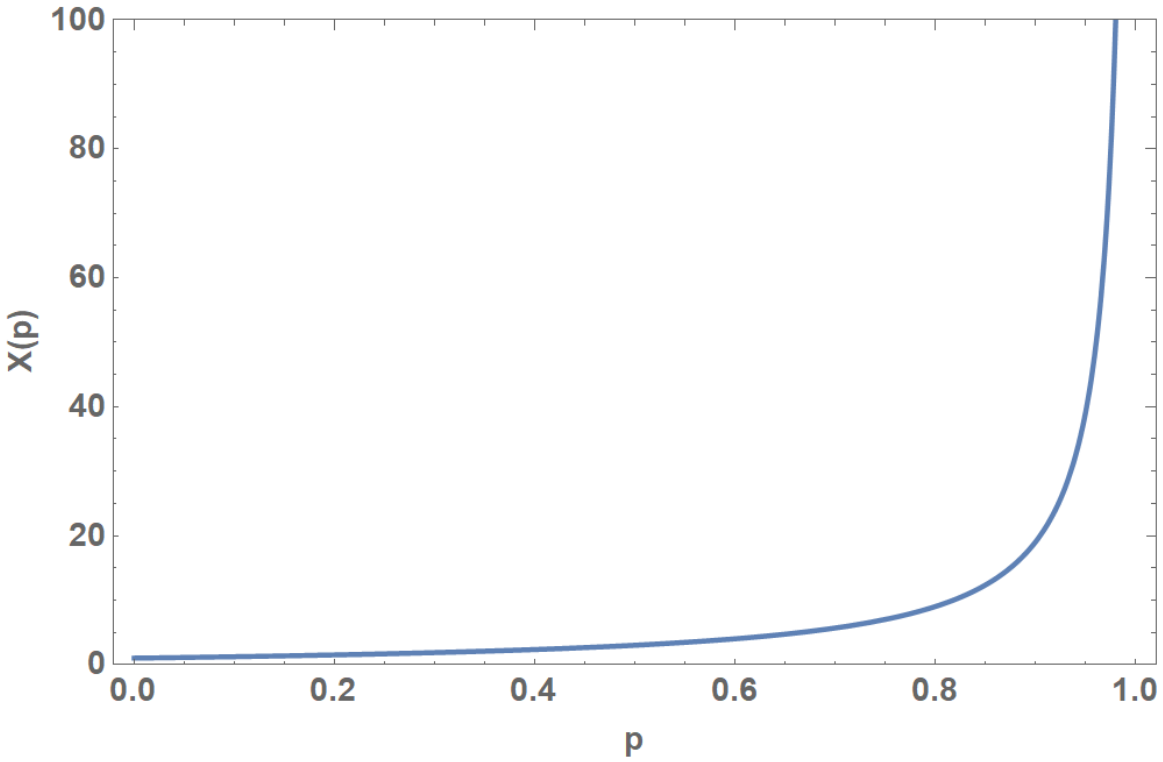


Fig. 2.4 A graph showing the average cluster size against p. As p approaches 1 the average cluster size diverges and as p approaches 0 the average cluster size tends to 0.

Finally by utilising the sum of a geometric series and (2.35) we can deduce that

$$\begin{aligned}\chi(p) &= \frac{1}{p}(1-p)^2 \sum_{s=1}^{\infty} s^2 p^s \\ &= \frac{1}{p}(1-p)^2 \left(p\frac{d}{dp}\right)\left(p\frac{d}{dp}\right)\left(\sum_{s=1}^{\infty} p^s\right) \\ &= \frac{1}{p}(1-p)^2 \left(p\frac{d}{dp}\right)\left(p\frac{d}{dp}\right)\left(\frac{p}{1-p}\right) \\ &= \frac{1+p}{1-p}. \end{aligned} \tag{2.41}$$

Clearly, $\chi(p)$ is an increasing function of p and this matches the physical behaviour of the system since the more occupied sites that exist the larger the average clusters become. Lastly in one dimension we introduce the correlation function, a quantity that measures the probability that two sites belong to the same finite cluster given that one of those sites is occupied. Let the site r_i in question be occupied. It follows that it can only be in the same finite cluster as site r_j if r_j is occupied along with all of the intermediate sites. Therefore the

site-site correlation function g takes the form

$$g(r_i, r_j) = p^{|r_i - r_j|} \tag{2.42}$$

where $r = |r_i - r_j|$ is the distance between the two named sites measured in lattice units. Evidently $g(r_i, r_i) = p^0 = 1$ and as the distance between chosen sites increases the correlation function is strictly decreasing, which can be expected from the definition and from the result in equation (2.35). Much like cluster sizes we can also have a characteristic correlation length, $\xi(p)$, given by

$$g(r_i, r_j) = \exp(\ln p^r) = \exp(r \ln p) = \exp(-r/\xi) \tag{2.43}$$

where

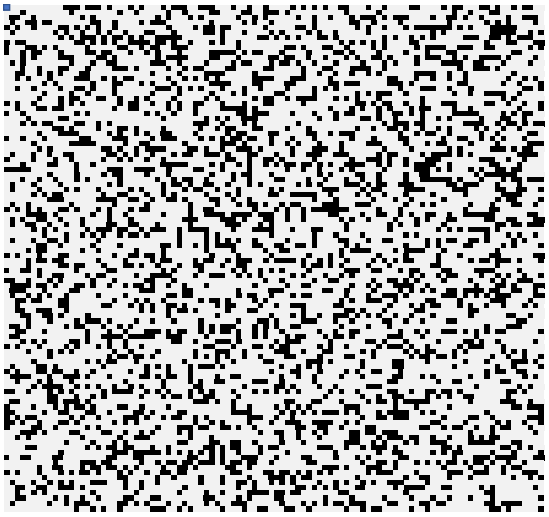
$$\xi(p) = -\frac{1}{\ln p} \propto \frac{1}{1-p} \text{ for } p \to 1^-. \tag{2.44}$$

This matches with the characteristic cluster size from equation (2.37) which is intuitive in one dimension since the characteristic correlation length represents the typical radius of the largest cluster. In higher dimensions these quantities will not match but are still related. Similarly there exists the following relationship between the sum over all correlation functions from site r_i and the average cluster size,

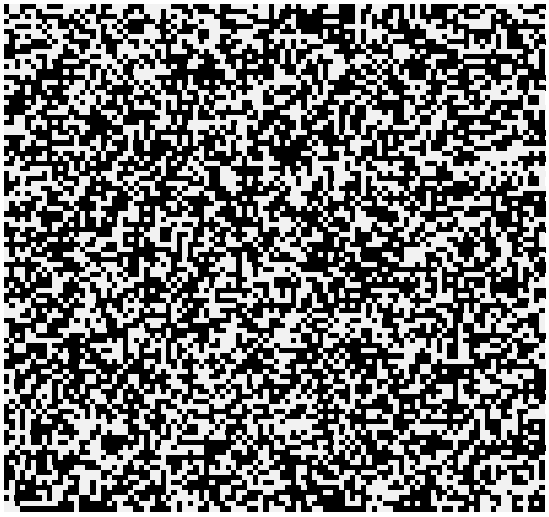
$$\begin{aligned}
\sum_{r_j} g(r_i, r_j) &= \sum_{r_j} p^{|r_i - r_j|} \\
&= \ldots + p^3 + p^2 + p + 1 + p + p^2 + p^3 + \ldots \\
&= 1 + 2p + 2p^2 + 2p^3 + \ldots \\
&= 1 + \frac{2p}{1-p} = \frac{1+p}{1-p} = \chi(p).
\end{aligned}$$

2.2.2 Percolation in Two Dimensions

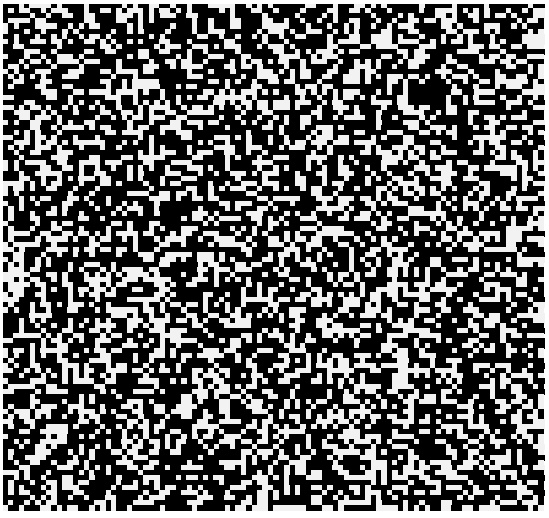
We now move onto looking at percolation in two dimensions and some necessary results we will utilise later. While much work has been done on site percolation in two dimensions there is no exact result for p_c on a square lattice. Simulation and numerical work has honed in on an approximation of around $p_c = 0.59275$ which is supported by the images in figure (2.5). There are still features of two dimensional systems that can be looked at however, particularly in their behaviours around the critical occupation probability.



(a) $p = 0.3$: Many spread, small finite clusters.



(b) $p = 0.55$: Large clusters forming, close to percolation.



(c) $p = 0.6$: A percolating cluster has formed spanning from top to bottom

(d) $p = 0.7$: The percolating cluster fills more space leaving less room for the smaller clusters

Fig. 2.5 Four lattices of size $L = 100$, each with a different occupation probability, to demonstrate the onset of a percolating cluster

Lattice Type	z	p_c
d = 1 line	2	1
d = 2 triangle	3	0.6971
d = 2 square	4	0.592746
d = 2 hexagonal	6	0.5
d = 3 simple cube	6	0.3116
d = 4 hypercubic	8	0.1969
Bethe	z	$1/(1-z)$

Table 2.1 Critical probabilities for the onset of a percolating cluster. z represents the number of nearest neighbours for each site in the structure [9].

One property we can investigate is the probability that an arbitrary site on the lattice belongs to a percolating cluster, $P_\infty(p)$. For finite lattice systems $P_\infty(p) > 0$ for $0 < p \leq p_c$ whereas in an infinite system $P_\infty(p) = 0$ for $0 < p \leq p_c$, a distinct difference brought on by the nature of a bounded system. As you increase the size of a finite system the increase in $P_\infty(p)$ as p approaches p_c becomes sharper and more rapid. The probability $P_\infty(p)$ is referred to as the order parameter to percolation and its behaviour close to p_c can be expressed as

$$P_\infty(p) \propto (p - p_c)^\beta \text{ for } p \to p_c^+. \tag{2.45}$$

Many values for β for varying lattice structures and dimensions have been calculated, including all two dimensional lattices where $\beta = 5/36$. Another quantity we can consider in two dimensions is the average cluster size, $\chi(p)$. Given that the average cluster size can, by definition, only include clusters of finite size we find that in two dimensions $\chi(p)$ is increasing for $p \to p_c^-$ and decreases as p increases above p_c because the onset of percolating clusters leaves reduced space for finite clusters to form. For a finite lattice, $\chi(p)$ can not diverge at p_c as the cluster size is limited by the size of the system, whereas in an infinite system $\chi(p)$ is divergent. This difference in behaviour between a finite and infinite system is the reason the infinite limit is required. This divergence for $p \to p_c$ is a power law with exponent γ in terms of the difference between p and p_c, that is

$$\chi(p) \propto |p - p_c|^{-\gamma} \text{ for } p \to p_c. \tag{2.46}$$

Much like for β this has been solved for multiple structures, including the two dimensional lattice where $\gamma = 43/18$.

Chapter 3

Birdsong Models

3.1 Sparrow Model

There are many real world systems which exhibit similar dynamics to the Ising model [11–13, 33], allowing us to use results from the Ising model research to explain the observable quantities and behaviours we see in reality. The original inspiration for the Ising model was ferromagnets but it has since been used in understanding tax evasion [39], human language uptake [40] and segregation patterns [41, 42]. One such example of a natural system we look at in this manner is the formation of birdsong dialects in different species [21, 43–45]; we will focus predominantly on the puget sound white-crowned sparrow utilising field observations by Nelson [4]. Almost all species of birds learn to sing songs [46, 47] as infants and their learning process is shown be down to a combination of genetic factors [48, 49] and learning from others [50, 51]. These learning processes can be of varying forms; some species learn from their parents while some learn from others in their surrounding population and some species can only learn as juveniles while others can learn and change at any point in their lives. The underlying dynamics for each species can vary based on a multitude of factors including social behaviours, learning styles, death rates and movement. However the fundamental balance of order, learning from and aligning with nearby agents, set against disorder, death of agents and dispersion, is consistent in many species meaning a comparison to ferromagnetic spin states can prove to be a useful method of understanding the formation of dialect territories and boundaries.

The observation that different birds have different songs and processes of social learning was made as far back as 1773 [52] but it was Thorpe's research into the song learning of chaffinches [53, 54] that ignited deeper investigations into this area. Research into dialect regions began in 1962 [55] and for the puget sound white-crowned sparrow Baptista [5]

noted distinct dialect regions on the Pacific coast of California, Oregon and Washington. The difference between dialects is classified by a variation in the terminal trill for which sample sonograms are shown in figure (3.1). The process of song learning in the white-crowned sparrow has been heavily researched including the required exposure time to adults and time period from birth that song learning is possible [20, 56]. Nelson compiled and outlined much of the puget sound white-crowned sparrow's learning processes, movements and death rates which form the foundation for our model [57]. During the breeding season the sparrows migrate to their breeding grounds having traveled between 500 and 1900 kilometres from their wintering grounds. Adult male sparrows are territorial and typically return back to the same locations after their migration and continue to sing the same songs as they had previously. The juveniles return later and fill in the vacant spaces in the territories left by the sparrows that have deceased. While at the breeding grounds the juveniles overproduce dialects, and on return to the nesting sites they sing multiple dialects. Field observation and playback experiments demonstrate that juveniles selectively conform to local majority, forgetting and removing the other dialects from their vocabularies. Once a yearling conforms to one particular song it keeps it for life and will continue to sing it for all subsequent year. It is this mechanism of selective learning that leads to the formation of distinct dialect domains and this forms the basis of our model.

From figure (3.3) we see there are many dialects along the entirety of the coast with each domain having a length of approximately 100 kilometres, whereas each bird only interacts within a range of approximately 110 metres [58]. As such any juvenile arriving on the coast will only feasibly encounter, at most, two variations of adult dialects, allowing us to only consider two dialects in our model. We approximate the positioning of the birds by considering a two dimensional lattice with one bird at each site and we can consider the dialect of each bird to be a binary spin state at that location.

In our model each time step represents the yearly return of the birds from the breeding grounds. With probability α a site will be replaced with a new juvenile and with probability $(1-\alpha)$ the site will remain the same, representing the same returning adult singing the same song. The probability α represents the death of the previous adult in that location with $\alpha \approx 0.4$ based upon the species life expectancy of around 2.5 years. The returning process is performed in two steps, firstly it is decided which adults survive and return, then secondly the gaps are filled in with new juveniles that make their song choice based upon their already returned neighbours, with this process sketched in figure (3.2).

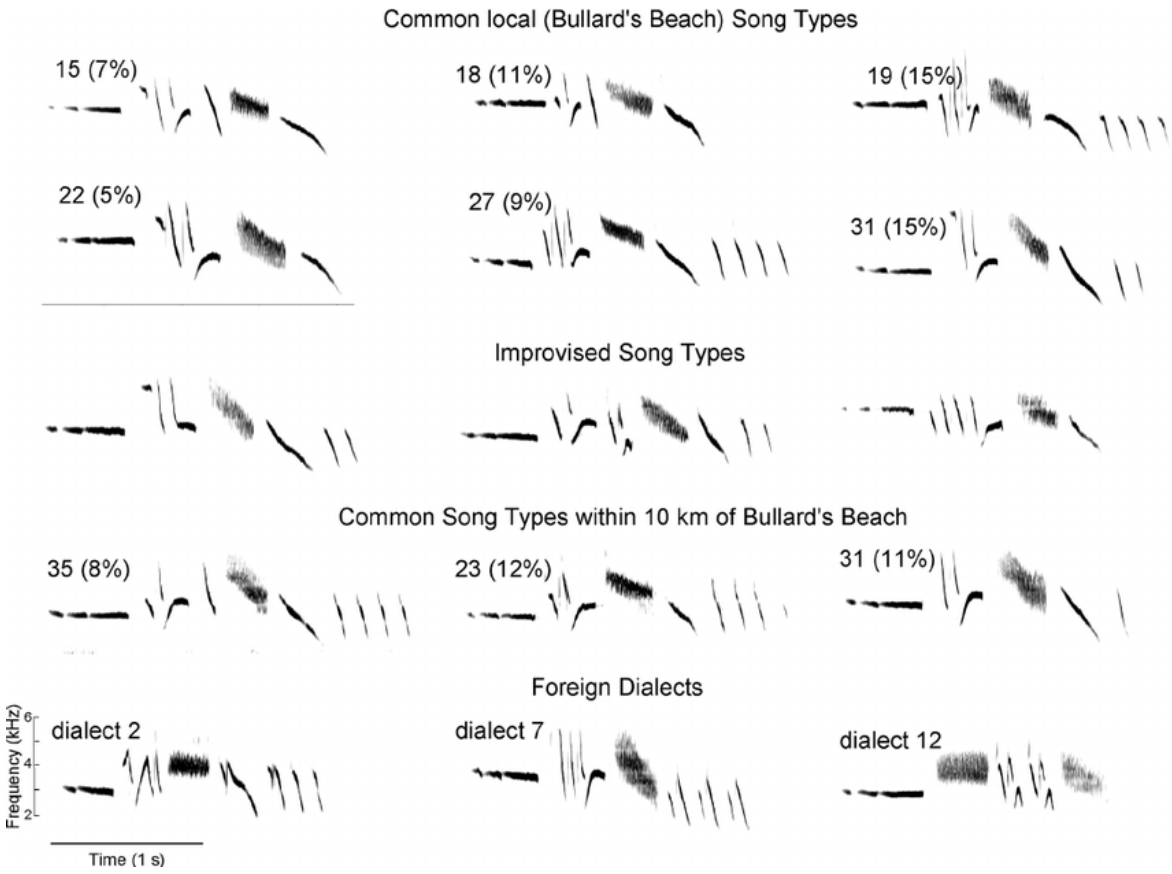


Fig. 3.1 Sound spectrograms of white-crowned sparrow songs. Shown are examples of the common song types in the study area at Bullard's Beach State Park, Oregon [3].

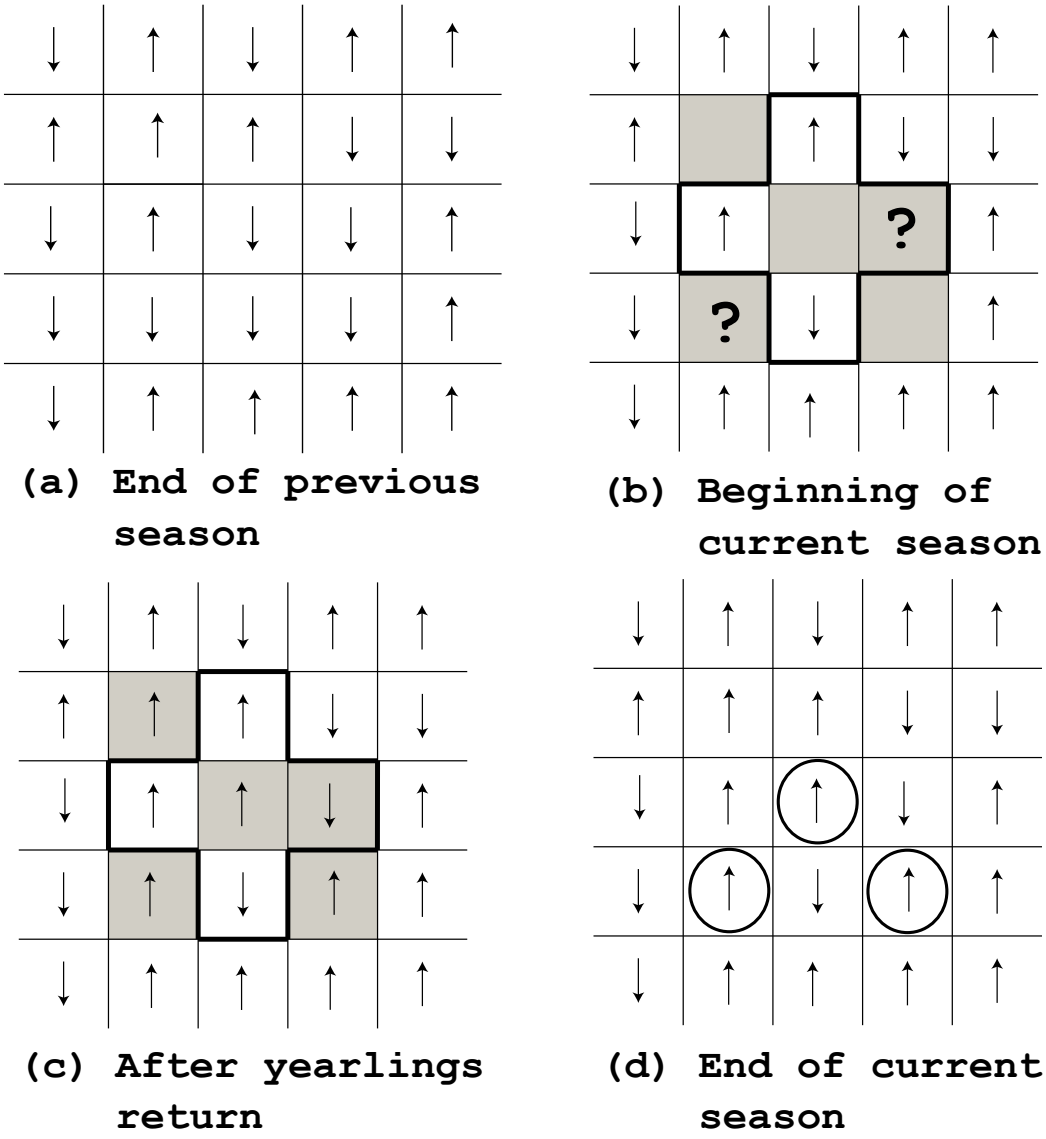


Fig. 3.2 Schematic representation of the lattice model with single phase reoccupation. Adult song states are either up or down. Grey squares represent territories vacated due to death of an adult bird. The heavy line traces the audible neighbourhood of the central site in the four neighbour model. Question marks represent equivocal information about local dialect coming from yearlings. States that change between time steps are circled.

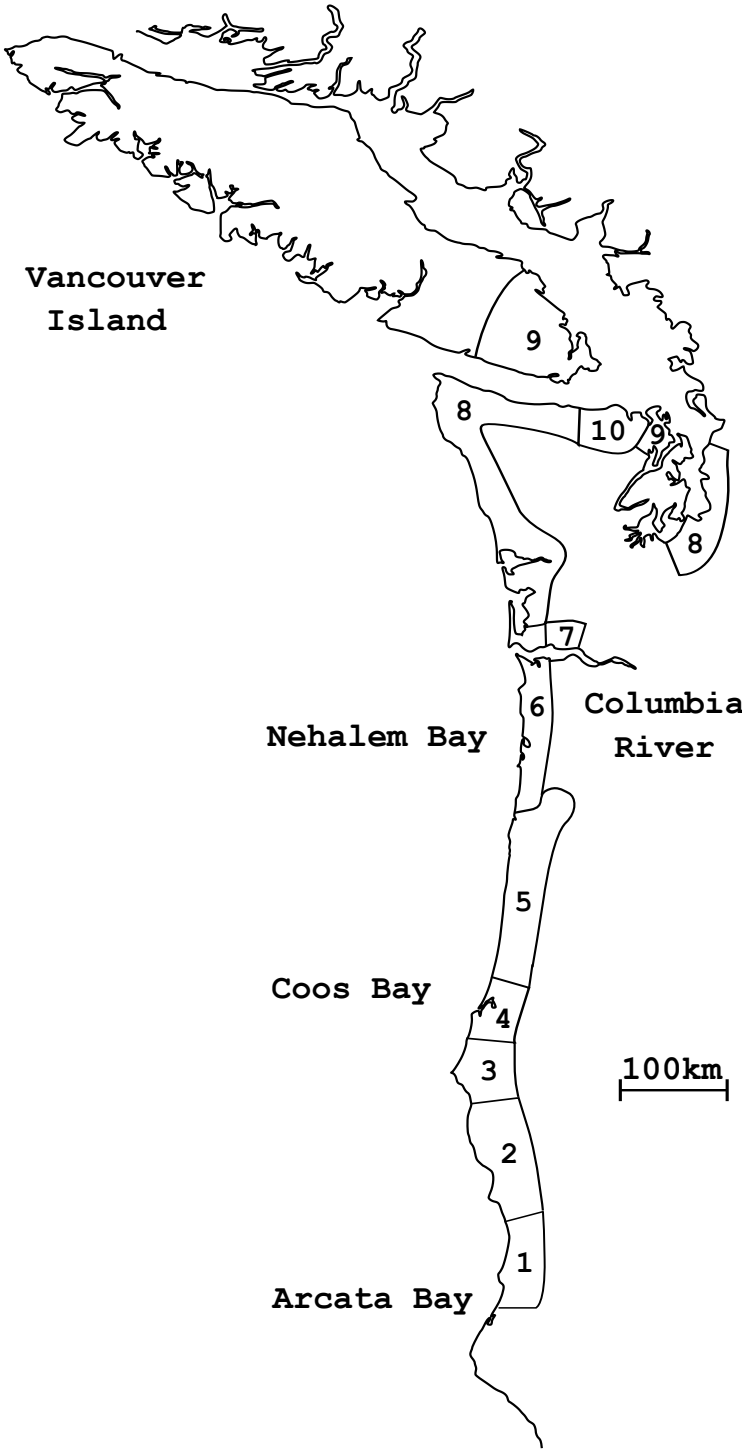


Fig. 3.3 Approximate geographical extent of song dialects of the puget sound white-crowned sparrow along the Pacific Northwest coast of the United States of America. Figure redrawn from original data in [4–7]

Fig. 3.4 A singing white-crowned sparrow in Stevens State Park, Oregon. Photo by Alex Lamoreaux [8]

3.1.1 Model Variations and Different Return Phase Structures

In the first version of the model for the puget sound white-crowned sparrow we look at the simplest version of events in which all of the juveniles return at the exact same time and only learn their songs from the previously existing adults in their immediate vicinity. In the event that multiple juveniles arrive in neighbouring sites they will not have an impact on each others song choice; the justification for this being that while the juveniles are adapting to the local dialect they will be singing many varied songs from their overlearned vocabulary and as such the net bias towards a specific song imparted on fellow juveniles should be negligible. In this example a returning juvenile can essentially only see and hear the local adults and, based on field observations, it is expected that they are highly likely to conform to the majority dialect. In the event that both dialects have equal influence on the juvenile they will make a choice with equal probability of $1/2$ for their final decision.

We later incorporate the information that following the arrival of adult birds the juveniles arrive throughout a window of approximately one month and that the median time to learn a permanent dialect by discarding overproduced songs is roughly twelve days. This arrival time difference allows for some early arriving juveniles to have an influence on the later territory holders. In the multiple return phase model we explicitly account for the sequential arrival of yearlings. We can divide the reoccupation portion of the model into n phases. During each phase a portion of the empty sites are filled and all site holders settle on their permanent dialect before the next phase arrives. Subsequent later phases are then able to make use of the songs that the earlier yearlings have chosen in order to make their own selection of adult song. The decision process is the exact same as the single return phase model; a yearling is influenced only by the birds in the local vicinity that have already chosen a permanent adult dialect but not from a yearling that arrives during the same phase as themselves. In the limit $n \to \infty$ the juveniles arrive one by one in a pure sequential model mimicking the physical process of cooperative random sequential adsorption [59].

3.1.2 Simulations and Results

We begin by looking at the single return phase model for which it is evident that if we let $\alpha = 0$ then no adults would be replaced by juveniles and the system would be in a permanent, fixed formation. Conversely if $\alpha = 1$ then at each time step every site would be replaced by juveniles and, with no adults to base their song choices on, the system would be an entirely random collection of spins. However for sufficiently low, positive values of α we see the formation of dialect domains and the formation of distinct boundaries. The cause

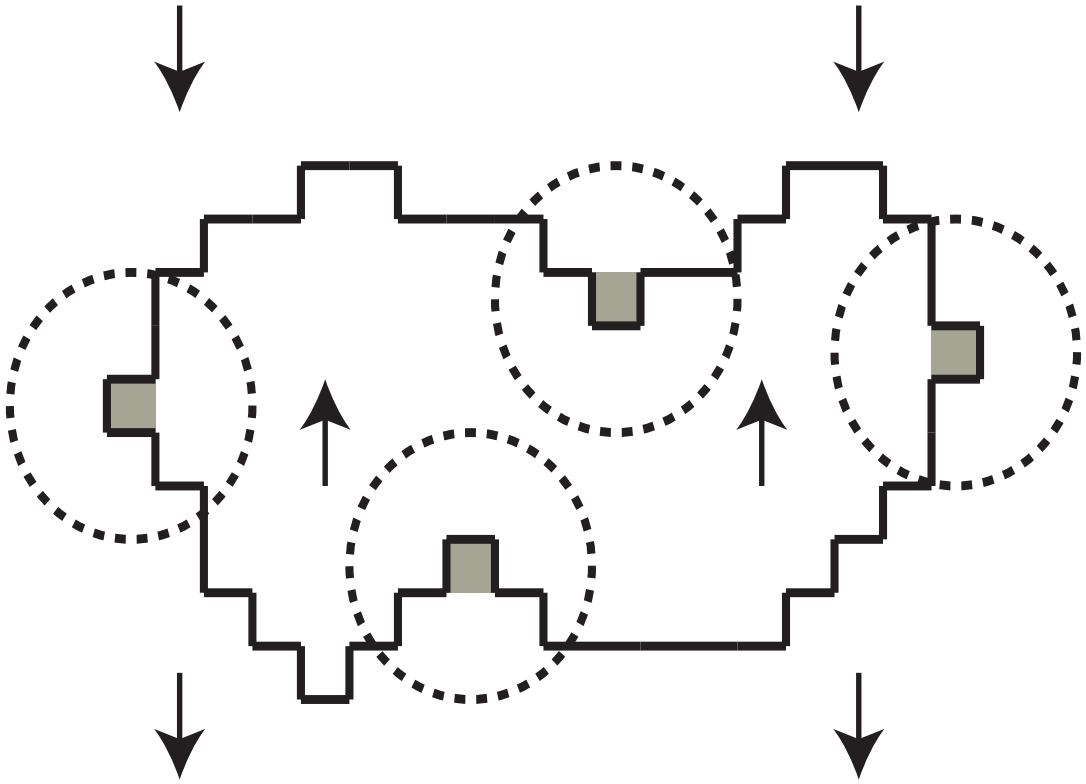


Fig. 3.5 Example of a dialect domain. The highlighted sites are particularly vulnerable to have a dialect switch after the next breeding season.

for the formation of distinct stripe states is highlighted in figure (3.5) and is a smoothing of dialect boundaries into smooth, straight lines due to what is effectively surface tension. The dynamics of the majority rule method is a well investigated topic [14, 27, 60, 61] and is a reasonable description of consensus formation in a network of interacting nodes. The majority rule method used here represents a nonlinear response to local dialects whereas for contrast a proportional response would be to align to a randomly chosen neighbour as is the case in the voter model [62]. However the voter model approach does not lead to the formation of domains and boundaries due to a lack of surface tension and the importance of a nonlinear learning rule has been identified in a non-spatial model of bird dialects [45].

Given that at low, positive values of α there are clear domain formations and at high α values there is randomness we aim to investigate the critical α, α_c, analogous to the Curie temperature [63] in ferromagnets, at which the transition between an ordered and disordered system occurs. To investigate this we introduce the correlation length, ξ, for a lattice of size L with periodic boundary conditions. We let $S(x,y) \in \{-1,1\}$ be the dialect of the bird at

site (x,y), allowing us to define the correlation length as

$$\xi = \mathbb{E}\left[\sum_{x'=1}^{L/2} S(x,y)S(x',y)\right], \tag{3.1}$$

where $\mathbb{E}[\cdot]$ denotes the expectation over the equilibrium probability distribution of system states. The correlation function is a measure of order between two sites of a given distance which decays over distance and equation (3.1) gives us a typical distance at which this correlation substantially drops. Due to translational invariance on a lattice with periodic boundary conditions ξ is independent of (x,y). It is clear from this definition that if the system is solely made up of one dialect then $\xi = L/2$, if the system is entirely random and there are no spatial correlations then $\xi = 0$ and if the system is made up of multiple finite domains then the value of ξ will lie somewhere between.

In order to estimate the equilibrium value of ξ we first create a very stable system with a high probability of being up, an initial condition of $p(\uparrow) = 0.95$, and a very low death rate of $\alpha = 0.01$. We allow the system to reach equilibrium and calculated ξ by computing the sum

$$\sum_{x'=1}^{L/2} S(x,y)S(x',y) \tag{3.2}$$

at all sites in the system over a sequence of time intervals until the spatial and time averages were fixed. Once the correlation length was confirmed we raise the death rate incrementally and repeat the process to recalculate ξ. Using this incremental system applied to an already stable system rather than reinitialising in a random state has two key benefits. Firstly α plays a role similar to thermodynamic temperature in magnetic systems and in systems that are quenched from an entirely random state you can have the formation of long lasting but non-equilibrium stripe states that can break translational invariance and reduce the accuracy in ξ. Simulations have suggested that in an $L \times L$ sub-critical system the average time to reach equilibrium is proportional to L^2 if no stripe state forms but this changes to $L^{3.4}$ if a stripe state forms [22]. We avoid this complication by starting in an already stable state. The second motivation is that this approach greatly reduces computation time since the system is already close to its equilibrium state as soon as α is raised. The results of this method are shown for the four neighbour case and eight neighbour case in figure (3.6) and these results match those for both systems with initially high α values that are lowered and also systems that are reinitialised for each new α value, provided no spontaneous stripe states are formed.

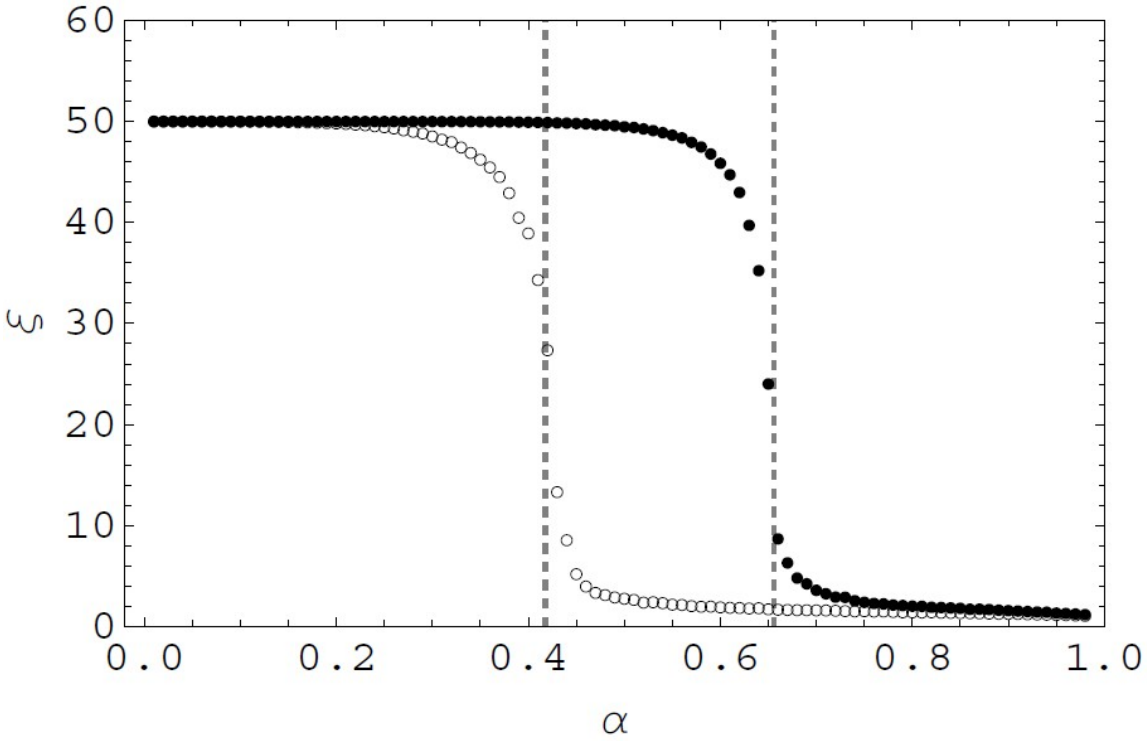


Fig. 3.6 Correlation length, ξ, shown as a function of death rate α in a 100 by 100 system with periodic boundary conditions. Open circles represent the four neighbour model and black dots represent the eight neighbour model. Dashed lines show analytical estimates of critical death rates $\alpha_c \approx 0.418$ (four neighbour model) and $\alpha_c \approx 0.657$ (eight neighbour model).

From the results in figure (3.6) we see that for sufficiently low death rates the system is made up of one single dialect domain. However as α increases we reach a critical value at which the correlation length plummets and the system fragments into disorder. The critical value is significantly higher in the eight neighbour model which is intuitive since you would expect to need a greater fraction of the population to die in order for a majority change around any given cell. Field observations indicate that in the case of the puget sound white-crowned sparrow the number of neighbours varies between one and four, suggesting the four neighbour model is a more accurate representation of the natural system. Interestingly from analysis we will find $\alpha_c \approx 0.4178$ in the four neighbour model which is remarkably close to the actual observed death rate which would make the existing dialect formations highly unstable since a slight increase in death rate due to a new invasive predator or disease could cause fragmentation. Given that real world death rates will fluctuate around a long term average and the multiple dialect territories along the coast have existed for a significant amount of time it implies that the observable critical death rate is significantly greater than the actual death rate.

This fact suggests that the extra information juveniles can receive in the multiple return phase model from their early arriving counterparts plays a significant role in the stabilisation of the system. In order to investigate this phenomenon we simulated the four neighbour model with multiple return phases; the cases $n = 2$ and $n \to \infty$. The results of these multiple phase models can be seen in figure (3.7) and it is evident that in both cases the critical death rate is shifted significantly above the previously seen 0.4178 and more importantly significantly ahead of the observed death rate showing that these systems would lead to stable domains.

Provided $\alpha < \alpha_c$ then, much like the Ising model with $T < T_c$, a system of one dialect is safe from the invasion of a differing song. From figure (3.5) we can see that if you introduce an invading cluster into a single dialect system the surface tension effects will cause the shape of the cluster to gradually approach circular, at which point the convexity causes boundary sites to switch to the original dialect and the cluster will shrink and disappear. Despite the system being unable to escape from having a single dialect state it can be shown that under the Ising model [12] systems can form stable states with the existence of multiple domains with opposing states. These stable configurations take on the form of stripe states [19], which because of their straight edges spanning across the system, are immune to the surface tension effects which cause domains with convex borders to shrink and as such these formations are extremely long lived. The probability of rectangular lattice systems forming into stripe states is strongly dependent on the aspect ratio of the lattice in question, with

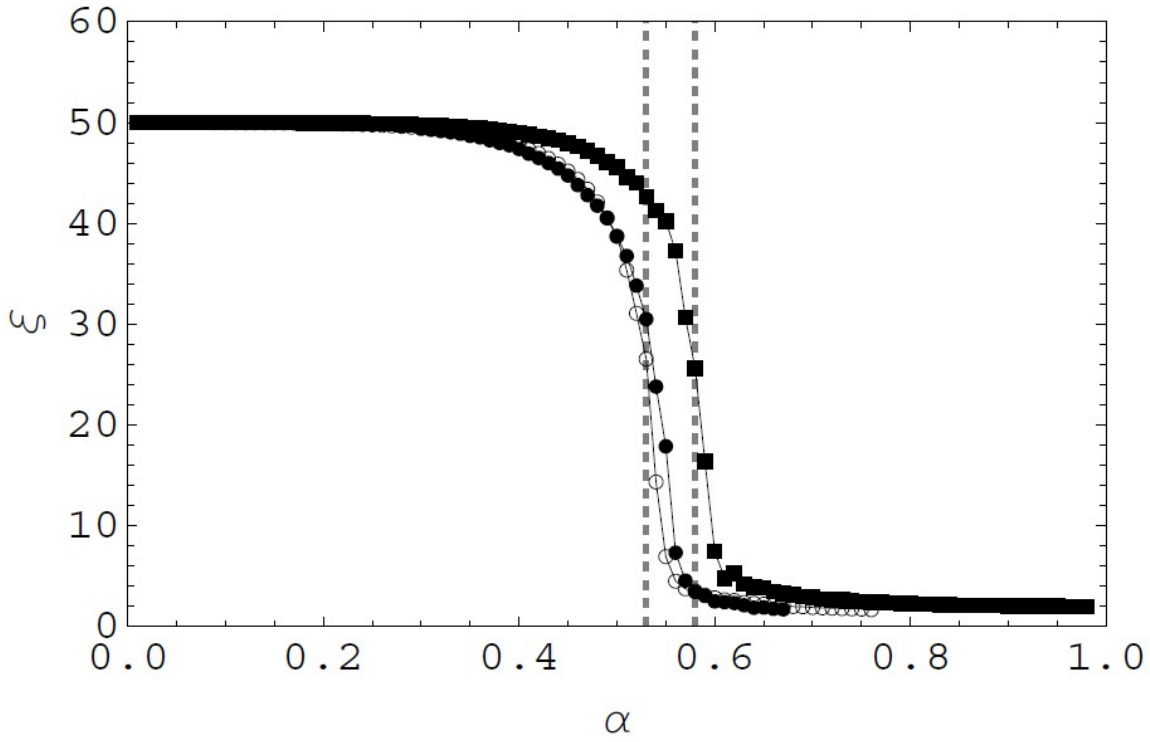


Fig. 3.7 Correlation length as a function of death rate α in a 100 by 100 system with periodic boundary conditions. Open circles correspond to the four neighbour, two-phase model, black dots correspond to the Ising model and black squares correspond to the sequential model. Dashed lines show estimated critical values from simulations, $\alpha_c \approx 0.53$ (two phase) and $\alpha_c \approx 0.58$ (fully sequential).

long and thin high aspect ratio systems being prone to producing many more stripes across the shorter dimension. From figure (3.3) we can see that the territories of the puget sound white-crowned sparrow are incredibly long and thin with high aspect ratios and this offers a plausible explanation for the striped formation of the dialect domains.

In order to investigate the typical size of a stripe domain in such shaped systems we simulated the fully sequential model on high aspect ratio rectangles with the results shown in figure (3.8). In sufficiently large systems, when domains become large, the dynamics of their borders can be approximated as continuous curves and the specific length scale of individual sites becomes irrelevant. In such cases only the relative sizes, rather than the actual sizes, of the structures within the system are important. This scale free property allows us to compare the aspect ratios of stripe domains in our modelled system to those of the real dialect domains. We ran simulations on a 2000 by 50 system until it consisted of well-defined stripes. We coarse-grained this system into 10 by 10 cells, to simplify the measuring and counting, and found the domains had a mean aspect ratio, AR, of $\mu_s(\mathrm{AR}) = 1.4$ and standard deviation of $\sigma_s(\mathrm{AR}) = 1.1$. Using the maps of field observed dialects we have estimated the length of the real coastal domains displayed in figure (3.3) and distribution maps show that the breeding grounds for the puget sound white-crowned sparrow spread approximately 50 kilometres inland [7]. Combining these field observed results we can calculate that the real dialect domains have a mean aspect ratio $\mu(\mathrm{AR}) = 2.8$ and standard deviation $\sigma(\mathrm{AR}) = 1.9$, with the removal of dialect 7, since it's non coastal, and dialect 9, since it's disconnected on Vancouver Island, from these calculations. If we include dialect 9 then the calculation gives $\mu(\mathrm{AR}) = 2.5$ bringing it closer to the simulated results. We still note however that the standard error of the mean for the eight field observations is $\sigma(\mathrm{AR})/\sqrt{8} \approx 0.7$ so the simulated resulted is only differing by approximately one standard error. While the results don't match perfectly they are close enough to support that the simulated model is comparable to the real world and that the observed domains arise due to the processes modelled and fall into the classification of stripe states.

3.1.3 Links to the Kinetic Ising Model

We have made use of an analogy between the territories of dialects made up by song choices of birds and the domains of aligned spin states in ferromagnets but we now aim to make this connection explicit by deriving Ising models with transition rates matching those of our dialect system; allowing us to derive a relationship between the death rate α and

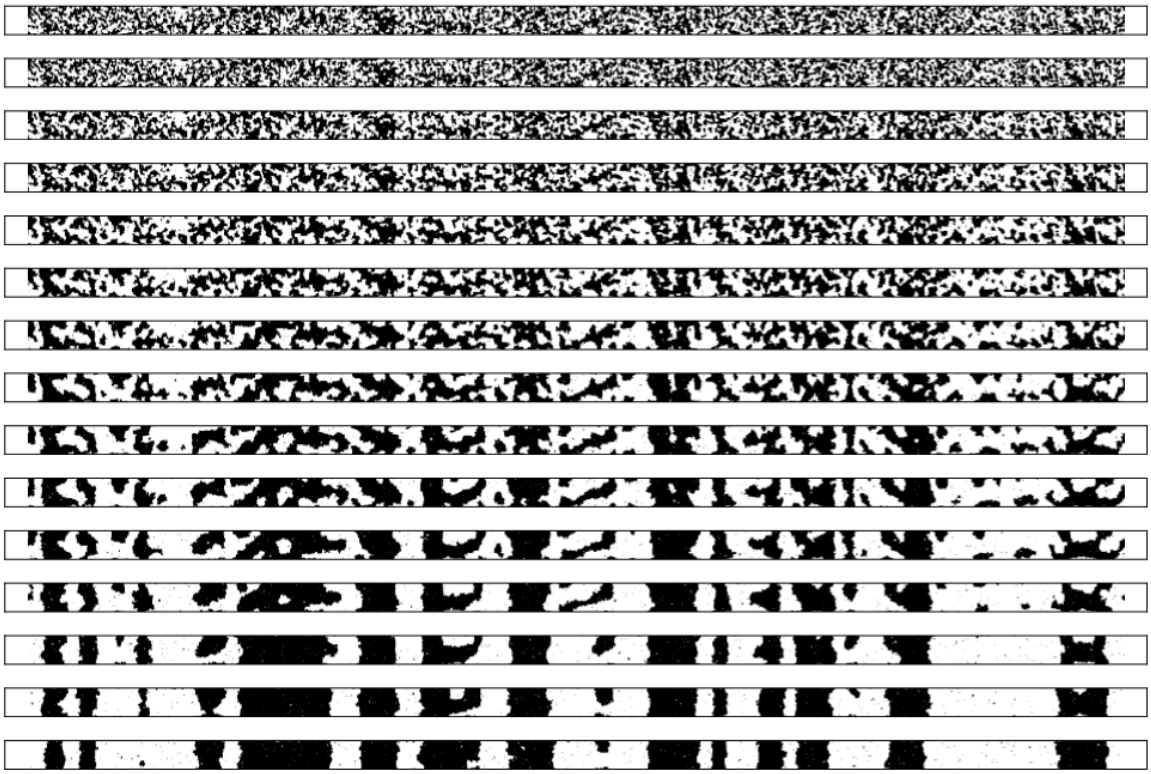


Fig. 3.8 Evolution of the sequential four neighbour model with $\alpha = 0.4$ on a 2000 by 50 lattice with wall boundary conditions along the horizontal sides (top and bottom) and periodic boundary conditions at the vertical sides. Black sites represent the up dialect, white sites represent the down dialect. The system is shown after $100 \times [2^k - 1]$ iterations for $k \in \{1, 2, \ldots, 9\}$.

thermodynamic temperature. We can also demonstrate multiple phase arrival of the yearlings can be viewed solely as extending the viewing range over which dialect choices are made.

Consider the four neighbour model, letting $\langle x,y\rangle$ denote the nearest neighbours of site (x,y). We can then define

$$\psi(x,y) = \sum_{\langle x,y\rangle} S(x',y'), \tag{3.3}$$

where $S(x',y') \in \{-1,1\}$ is the dialect at site (x',y') at the end of the breeding season. We now suppose that site (x,y) is vacant at the start of the next arrival phase due to the occupying agent in the model dying. The juvenile that occupies this site is unlikely to have access to $\psi(x,y)$ due to the non-return of some neighbours, instead only having access to a portion of that information. This disruption in accessing the complete information plays the same disordering role as temperature in the Ising model. We define an indicator function for the return of an adult at a given site, (x,y) after as

$$R(x,y) = \begin{cases} 1 & \text{with prob. } 1-\alpha \\ 0 & \text{with prob. } \alpha. \end{cases} \tag{3.4}$$

Using this indicator function in conjunction with (3.3) we can compute the sum of dialects a juvenile witnesses on its arrival at site (x,y) as

$$\overline{\psi}(x,y) = \sum_{\langle x,y\rangle} R(x',y')S(x',y'), \tag{3.5}$$

and the juvenile will choose his dialect in accordance to the sign of this potentially incomplete information term. We define the step function

$$f(z) = \begin{cases} 1 & \text{if } z > 0 \\ \frac{1}{2} & \text{if } z = 0 \\ 0 & \text{if } z < 1. \end{cases} \tag{3.6}$$

Given $\psi(x,y)$ the probability that the juvenile will choose the $\uparrow$ dialect is

$$\begin{aligned} p_{\uparrow}(\psi) &= \mathbb{P}(\overline{\psi} > 0) + \frac{1}{2}\mathbb{P}(\overline{\psi} = 0) \\ &= \mathbb{E}[f(\overline{\psi})], \end{aligned} \tag{3.7}$$

where the expectation is taken over all possible combinations of vacated sites. In order to compute this expectation as a function of α we first note that $N_\uparrow$ and $N_\downarrow$, the numbers of birds with the up and down dialects around site (x,y) *after* the adults have returned, meaning some site may be vacant, conditional on ψ, are binomially distributed with probabilities

$$\begin{aligned}&\mathbb{P}(N_\uparrow = u \cap N_\downarrow = v|\psi)\\ &\quad= \binom{2+\frac{\psi}{2}}{u}\binom{2-\frac{\psi}{2}}{v}\alpha^{4-u-v}(1-\alpha)^{u+v}.\end{aligned} \tag{3.8}$$

By noting that $\overline{\psi} = N_\uparrow - N_\downarrow$ we can compute the expectation from equation (3.7) as

$$\mathbb{E}[f(\overline{\psi})] = \sum_{u=0}^{2+\frac{\psi}{2}}\sum_{v=0}^{2-\frac{\psi}{2}}\mathbb{P}(N_\uparrow = u \cap N_\downarrow = v|\psi)f(u-v), \tag{3.9}$$

allowing us to exactly compute the probability for the dialect at site (x,y) after a breeding season:

$$p_\uparrow(\psi) = \begin{cases}\frac{\alpha^4}{2}, & \text{if } \psi = -4\\ \frac{3\alpha^2}{2} - 2\alpha^3 + \alpha^4, & \text{if } \psi = -2\\ \frac{1}{2}, & \text{if } \psi = 0\\ 1 - \frac{3\alpha^2}{2} + 2\alpha^3 - \alpha^4, & \text{if } \psi = 2\\ 1 - \frac{\alpha^4}{2}, & \text{if } \psi = 4.\end{cases}$$

From figure (3.9) we can see that $p_\uparrow(\psi)$ takes the form of a smoothed discrete version of $f(\psi)$.

We now make the link between this system of dialect spins at the end of a breeding season and a system of Ising spins, evolving under Glauber dynamics [33, 64] at β, the inverse thermodynamic temperature. In this Ising model case, sites on the lattice are selected one at a time with equal uniform probability and that site will be set to the dialect up state with probability

$$\omega_\uparrow(\psi) := \frac{1+\tanh\beta\psi}{2}, \tag{3.10}$$

which, like equation (3.10), is also a smoothed discrete version of $f(\psi)$. If we now consider a site (x,y) on a domain boundary then $\psi(x,y) \in \{-2,0,2\}$, meaning we can choose β such that $\omega_\uparrow(\psi)$ and $p_\uparrow(\psi)$ match for these values of ψ and the probability of those sites choosing

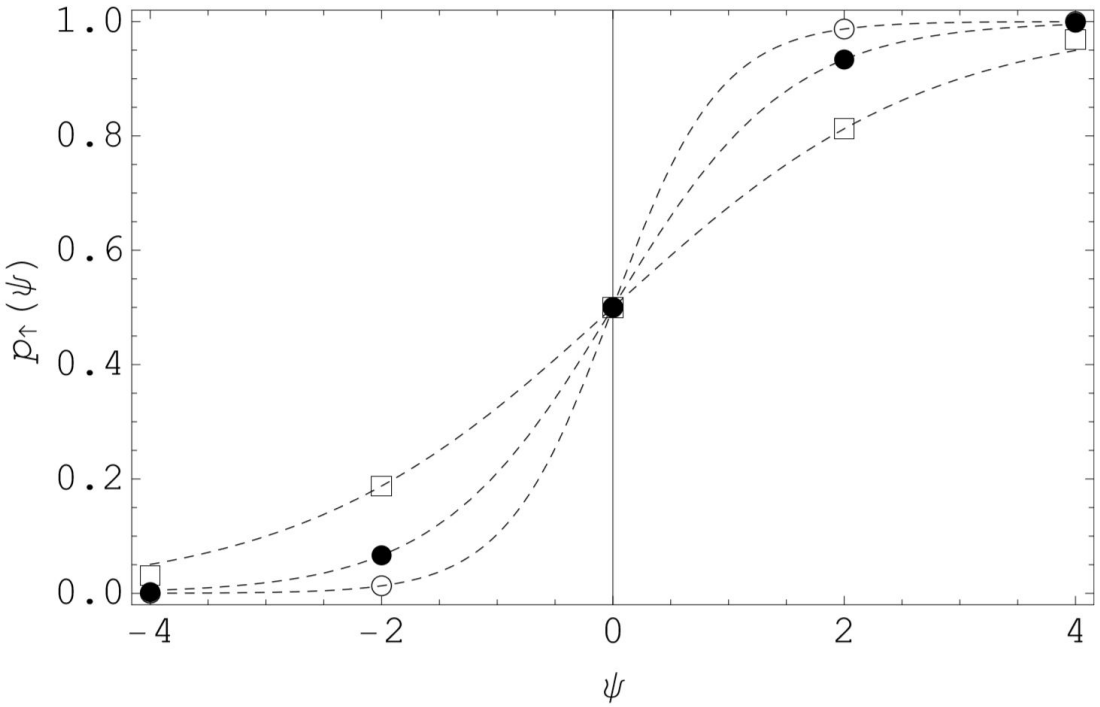


Fig. 3.9 The discrete function $p_\uparrow(\psi)$ for $\alpha \in \{0.1, 0.3, 0.6\}$ (open circles, filled circles, squares). Also shown as dashed lines are the continuous functions $w_\uparrow(\psi) = (1 + \tanh[\beta(\alpha)\psi])/2$.

dialects is identical across both models. Our justification for this is that the most important sites in determining system behaviour are those sites on or close to the boundaries, and those sites are likely to have either equal up and down neighbours or a minor imbalance. We can match our three target values with only one parameter because both expressions have rotational symmetry around a fixed value at 0. It is important however to note that in the dialect model the simultaneous updates across sites leads to a correlation between state flips that isn't present in Glauber dynamics, leading to a minor difference. If we match the values of $\omega_{\uparrow}(\psi)$ and $p_{\uparrow}(\psi)$, we find that the inverse temperature takes the form

$$\beta(\alpha) = \frac{1}{2}\tanh^{-1}(1-3\alpha^2+4\alpha^3-2\alpha^4). \tag{3.11}$$

We can now make use of this approximate analytical relationship between death rate and inverse temperature to estimate the critical death rate in the dialect model. Equating $\beta(\alpha)^{-1}$ to the exact critical temperature in the Ising model we deduce that

$$\frac{1}{\beta(\alpha)} = \frac{2}{\ln(1+\sqrt{2})} \tag{3.12}$$

and solving for α gives us the critical death rate

$$\alpha_c \approx 0.418 \tag{3.13}$$

to three significant figures. If we compare this to the previous simulations for correlation lengths, ξ in figure (3.6), we can acknowledge that this method of prediction is remarkably close in the four neighbour case.

We naturally extend this to the eight neighbour model by redefining $\psi(x,y)$ to be the sum of the eight nearest neighbours of site (x,y). In this eight neighbour case

$$\begin{aligned}\mathbb{P}(N_{\uparrow} = u \cap N_{\downarrow} = v|\psi) \\ = \binom{4+\frac{\psi}{2}}{u}\binom{4-\frac{\psi}{2}}{v}\alpha^{8-u-v}(1-\alpha)^{u+v}.\end{aligned} \tag{3.14}$$

We now let $p_{8\uparrow}(\psi)$ represent the probability that of a juvenile selecting the up dialect in the eight neighbour model and repeating the previous calculations we find that

$$p_{8\uparrow}(0) = \frac{1}{2} \tag{3.15}$$

$$p_{8\uparrow}(2) = 1 - 5\alpha^2 + 20\alpha^3 - 45\alpha^4 + 64\alpha^5 - \frac{115\alpha^6}{2} + 30\alpha^7 - 7\alpha^8 \tag{3.16}$$

$$p_{8\uparrow}(4) = 1 - \frac{15\alpha^4}{2} + 24\alpha^5 - 34\alpha^6 + 24\alpha^7 - 7\alpha^8 \tag{3.17}$$

$$p_{8\uparrow}(6) = 1 - \frac{7\alpha^6}{2} + 6\alpha^7 - 3\alpha^8 \tag{3.18}$$

$$p_{8\uparrow}(8) = 1 - \frac{\alpha^8}{2} \tag{3.19}$$

with $p_{8\uparrow}(\psi) = 1 - p_{8\uparrow}(-\psi)$ allowing us to compute the remaining four terms simply. As before we match up $p_{8\uparrow}(\psi)$ and $\omega_{\uparrow}(\psi)$ for $\psi \in \{-2, 0, 2\}$ and see

$$\beta_8(\alpha) = \frac{1}{2}\tanh^{-1}(2p_{8\uparrow}(2) - 1). \tag{3.20}$$

As before we equate this quantity with the known series estimate for the critical temperature in the eight neighbour Ising model, $T_c \approx 5.2599$, to obtain the approximate critical death rate of

$$\alpha_c \approx 0.657. \tag{3.21}$$

When we compare this to the approximate critical death rate of $\alpha_c \approx 0.651$ from the correlation simulations in figure (3.6) we see that this is an effective estimation method with a minor inaccuracy. We postulate that this discrepancy is derived from the fact that not all sites on domain boundaries have identical probabilities of changing dialect between the two models.

3.1.4 Multiple Return Phases

As previously shown in figure (3.7) the concept of multiple return phases offers both a greater comparison to the real world observed behaviours and also a more reasonable critical death rate given the long term stability of the dialect domains. As before we aim to form a link between the two phase, four neighbour model and the Ising model. In the case of the two phase model a juvenile has the potential to have its dialect decision be influenced by the dialects of a greater number of adults within the local vicinity. As seen in figure (3.10) if the

central site is occupied by a juvenile returning in the second phase then it is possible that one of the four nearest neighbours is occupied by a juvenile that returned in the first phase. In that case the agents in the outer sites will have a second hand influence on the decision of the central juvenile. With reference to figure (3.10) we let ψ_1, ψ_2 and ψ_3 be the sum of the dialects in, respectively, the sets of nearest neighbours, next nearest neighbours and next next nearest neighbour sites of a new central territory holder:

$$\begin{aligned} \psi_1 &:= S_1 + S_2 + S_3 + S_4, \\ \psi_2 &:= S_5 + S_6 + S_7 + S_8, \\ \psi_3 &:= S_9 + S_{10} + S_{11} + S_{12}. \end{aligned} \tag{3.22}$$

There is a notable difference between the sites of ψ_2 and the sites of ψ_3 because, while both terms can only have second hand influence on the central agent, the agents in the sites of ψ_2 have the potential to influence the central agent through two different immediate neighbours.

We refer to the sets of sites as the first, second and third rings. In the two phase model, after each breeding season, there are three different cases to consider at each site. Firstly an adult can return to a location and no more consideration needs to be made. Secondly the site can be occupied by a juvenile that returns during the first reoccupation phase; in which case the probability of this agent picking a given dialect is solely dependent on ψ_1 and is identical to the previously solved single phase model. It is in the event that a site is reoccupied by a juvenile during the second return phase that we must make these additional considerations of rings two and three. We define $p_{\uparrow}(\psi_1, \psi_2, \psi_3) = \mathbb{P}(\uparrow | \psi_1, \psi_2, \psi_3)$ to be the probability of the central agent choosing the up dialect. We can compute this probability if we assume that all dialect arrangements within each ring are equally probable. We adopt the site numbering used in figure (3.10) and let $\overrightarrow{S} = (S_1, S_2, \ldots, S_{12})$ represent the song states at the end of the previous breeding season, $\overrightarrow{R} = (R_1, R_2, \ldots, R_{12})$ be the indicators for the return of adults at the start of the next breeding season and $\overrightarrow{r} = (r_1, r_2, r_3, r_4)$ be the indicator for yearlings arriving in the first return phase in the four nearest neighbour sites. Lastly we let $\overrightarrow{B} = (B_1, B_2, B_3, B_4)$ be a vector of Bernoulli variables $B_i \in \{-1, 1\}$ which indicate the song choices of the agents in these sites in the event that these sites are occupied by first phase returners with equivocal dialect information. This vector of Bernoulli variables is only required for first phase returners in the nearest neighbour ring with equivocal information because those are the only sites that make a 'coinflip' decision that can also have a bearing

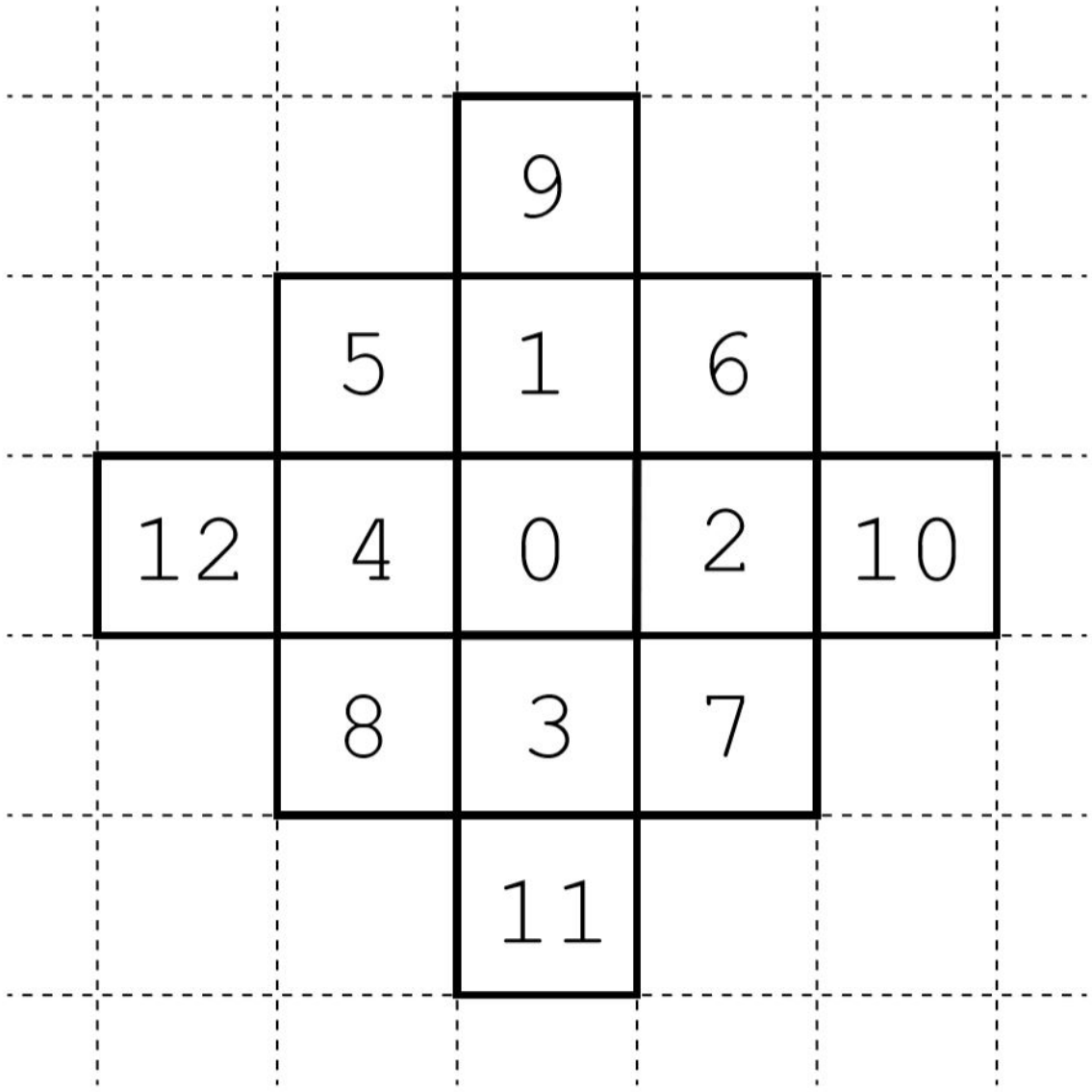


Fig. 3.10 Sites, outlined bold, that can influence the dialect of the central 0 site in the two phase return model.

on the central site's final decision. With these indicator functions we get

$$\begin{aligned}
\mathbb{P}(R_i = 0) &= \alpha \\
\mathbb{P}(R_i = 0 \cap r_i = 0) &= \frac{\alpha}{2} \\
\mathbb{P}(R_i = 0 \cap r_i = 1) &= \frac{\alpha}{2} \\
\mathbb{P}(R_i = 1) &= 1 - \alpha \\
\mathbb{P}(R_i = 1 \cap r_i = 0) &= 1 - \alpha \\
\mathbb{P}(R_i = 1 \cap r_i = 1) &= 0.
\end{aligned} \tag{3.23}$$

We also define

$$\begin{aligned}
\psi_1(\overrightarrow{S}) &:= \sum_{k=1}^{4} S_k \\
\psi_2(\overrightarrow{S}) &:= \sum_{k=5}^{8} S_k \\
\psi_3(\overrightarrow{S}) &:= \sum_{k=9}^{12} S_k.
\end{aligned} \tag{3.24}$$

In order to compute $p_{\uparrow}(\psi_1, \psi_2, \psi_3)$ we condition on the central site being empty and then consider the two cases corresponding to each phase in which the territory becomes occupied.

In the first case the central site is reoccupied during the first return phase and we define $p_1(\uparrow \mid \overrightarrow{S}, \overrightarrow{R})$ to be the probability of the central site choosing the up dialect conditional on $\overrightarrow{S}$ and $\overrightarrow{R}$. We define the influence this agent receives from existing nearest neighbour adults to be

$$X_0 = \sum_{k=1}^{4} S_k R_k \tag{3.25}$$

and it follows that

$$p_1(\uparrow \mid \overrightarrow{S}, \overrightarrow{R}) = I_{\{X_0 > 0\}} + \frac{1}{2} I_{\{X_0 = 0\}}, \tag{3.26}$$

where the indicator function I_A of the event A is defined as

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases} \tag{3.27}$$

These equations are a rewriting of equations (3.5) and (3.7) utilising the new notation that we need for the multiple return phase model.

In the more complex case, in which the central site is occupied by a juvenile during the second return phase we define $p_2(\uparrow \mid \overrightarrow{S}, \overrightarrow{R}, \overrightarrow{r}, \overrightarrow{B})$ to be the probability of the central agent choosing the up dialect conditional on $\overrightarrow{S}$, $\overrightarrow{R}$, $\overrightarrow{r}$ and $\overrightarrow{B}$. We then define

$$X_k = \sum_{i \in \langle k \rangle} S_i R_i \quad k \in \{1,2,3,4\}, \tag{3.28}$$

where $\langle k \rangle$ denotes the nearest neighbours around site k. X_k is a random variable representing the sum of all states around site k at the beginning of the current season. The crucial quantity that determines the final dialect choice of the central agent is the sum of states in sites $\{1,2,3,4\}$ at the end of the first phase. We let $\{s_k\}_{k=1}^4$ be these states where $s_k \in \{-1,0,1\}$. If $R_k = 1$ then site k is occupied by the returning adult and $s_k = S_k$. If $R_k = 0$ and $r_k = 1$ then the site is occupied by a first phase returner and bases its dialect on X_k. Lastly if $R_k = r_k = 0$ then the site is a fellow second phase returner and as such has no influence and $s_k = 0$. From these statements it follows that

$$s_k = R_k S_k + r_k [I_{\{X_k > 0\}} - [X_k < 0] + B_k I_{\{X_k = 0\}}]. \tag{3.29}$$

We can now define

$$x_0 = \sum_{i=1}^{4} s_i, \tag{3.30}$$

then

$$p_2(\uparrow \mid \overrightarrow{S}, \overrightarrow{R}, \overrightarrow{r}, \overrightarrow{B}) = I_{\{x_0 > 0\}} + \frac{1}{2} I_{\{x_0 = 0\}}. \tag{3.31}$$

Since a site with a non-returning adult is equally likely to be filled during the first phase as it is in the second phase, then conditional on $\overrightarrow{S}$, $\overrightarrow{R}$, $\overrightarrow{r}$ and $\overrightarrow{B}$ we have

$$\mathbb{P}(\uparrow | \overrightarrow{S}, \overrightarrow{R}, \overrightarrow{r}, \overrightarrow{B}) = \frac{1}{2}[p_1(\uparrow | \overrightarrow{S}, \overrightarrow{R}) + p_2(\uparrow | \overrightarrow{S}, \overrightarrow{R}, \overrightarrow{r}, \overrightarrow{B})]. \tag{3.32}$$

In order to compute this conditional probability we must first define

$$\begin{aligned} f_{Rr}(u,v) &= \mathbb{P}(R_i = u \cap r_i = v), \\ f_R(u) &= \mathbb{P}(R_i = u). \end{aligned} \tag{3.33}$$

The joint probability mass function of $\overrightarrow{R}$ and $\overrightarrow{r}$ is then

$$\begin{aligned} f(\overrightarrow{u}, \overrightarrow{v}) &:= \mathbb{P}(\overrightarrow{R} = \overrightarrow{u} \cap \overrightarrow{r} = \overrightarrow{v}) \\ &= \Pi_{k=1}^{4} f_{Rr}(u_k, v_k) \Pi_{k=5}^{12} f_R(u_k). \end{aligned} \tag{3.34}$$

We now make the assumption that conditional on the values of ψ_1, ψ_2, ψ_3 that all complying iterations of $\overrightarrow{S}$ are equally likely. We let $\mathbf{S}$ be the set of all values of $\overrightarrow{S}$ and define

$$A(y_1, y_2, y_3) = \{ \overrightarrow{S} \in \mathbf{S} | \psi_1(\overrightarrow{S}) = y_1, \psi_2(\overrightarrow{S}) = y_2, \psi_3(\overrightarrow{S}) = y_3 \}. \tag{3.35}$$

Since $\overrightarrow{B}$ is a set of four Bernoulli variables all sixteen values are equally probable and as such

$$p_{\uparrow}(y_1, y_2, y_3) = \frac{1}{16|A|} \sum_{\overrightarrow{S} \in A, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{B}} f(\overrightarrow{u}, \overrightarrow{v}) p(\uparrow | \overrightarrow{S}, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{B}). \tag{3.36}$$

The summation over A is required because, while we have assumed that all complying iterations of $\overrightarrow{S}$ are equally likely, they do not all lead to the same probability of the central site being up. To compute all permutations requires summing over 2^{32} combinations of song states $\overrightarrow{S}$, arrival times $(\overrightarrow{R}, \overrightarrow{r})$ and decision variables $\overrightarrow{B}$. These results were achieved using a simple Python program and the full results can be found in Appendix A with a sample result here,

$$p_{\uparrow}(0,2,4) = \frac{3}{256}\alpha^{11} - \frac{5}{128}\alpha^{10} + \frac{1}{256}\alpha^{9} - \frac{7}{128}\alpha^{7} \tag{3.37}$$

$$+ \frac{1}{4}\alpha^{6} - \frac{3}{8}\alpha^{5} + \frac{9}{64}\alpha^{4} - \frac{3}{16}\alpha^{2} + \frac{1}{4}\alpha + \frac{1}{2} \tag{3.38}$$

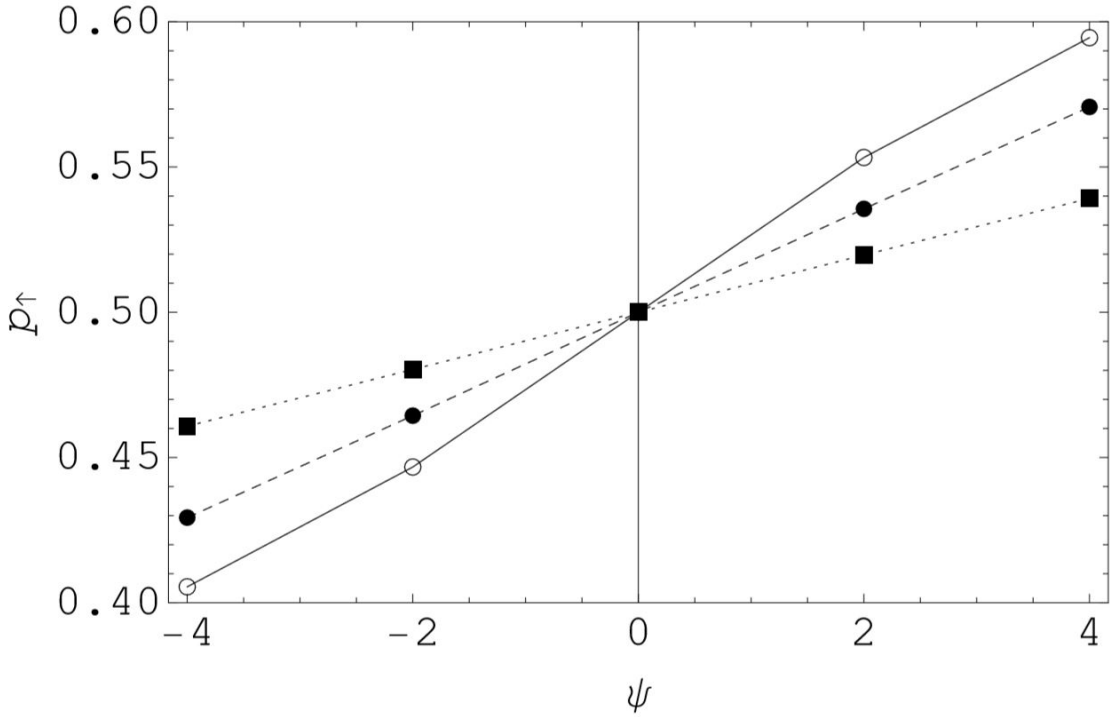


Fig. 3.11 The influence of second and third ring sites in the case where $\psi_1 = 0$ and $\alpha = 0.4$ the field observed death rate. Open circles: $p_\uparrow(0,\psi,\psi)$, closed circles: $p_\uparrow(0,\psi,0)$ and squares: $p_\uparrow(0,0,\psi)$.

These results have been verified by Monte Carlo simulation and for assurance the simple cases $\alpha = 1$ and $\alpha = 0$ equate exactly. For instance

$$p_\uparrow(0,2,4)|_{\alpha=0} = p_\uparrow(0,2,4)|_{\alpha=1} = \frac{1}{2}, \tag{3.39}$$

where for $\alpha = 1$ all neighbouring adults return and only ψ_1 matters, which in this case is 0 leading to each dialect being chosen with equal probability and for $\alpha = 0$ all adults perish, all information is lost and every site makes a random choice.

To quantify the importance of the outer rings we refer to figure (3.11), illustrating the effect the outer ring values can have on $p_\uparrow$ when the information ψ_1 from the inner ring is equivocal. We see that the influence provided from the outer rings can influence the final decision of the central agent by a factor of $\frac{3}{2}$.

By selecting sites at random from the lattice and updating them using the probabilities $p_\uparrow(\psi_1,\psi_2,\psi_3)$ which may be viewed as a form of Glauber dynamics with appropriately chosen Hamiltonian we obtain the correlation lengths shown in figure (3.7). This plot shows

us that, while their impact is weak, the outer rings in the two phase model create a significant increase in dialect domain stability that better aligns with the real world observations of long term stable domains in spite of fluctuations in adult death rate.

3.2 Birdsong Model Conclusion

Using observed data on the learning processes and nesting behaviours of the puget sound white-crowned sparrow [57], we have constructed a simple lattice model in order to explain and justify the existence of large dialect domains in their coastal breeding grounds. We have shown that the ordering effect of juvenile song conformity and the disordering effect of adult death can be compared to spin interactions and thermodynamic temperature in two dimensional ferromagnetic materials. In utilising this analogy we have shown how the distinct dialect domains and boundaries can be viewed as stripe states, calculated a critical death rate at which the existence of stable dialect regions would destabilise and we have shown how a variation in the return structure of juveniles can impact the long term stability due to an increase in the local region for which a site may consider in its majority rule calculations. Birdsong dialects are widely observed and take on a variety of forms [51]; we suggest that the analogy to ordering in physical systems may be usefully applied to other species and potentially related to the study of observed and historical human dialect domains [65].

The work in this chapter was a joint effort between myself and my supervisor, Dr James Burridge, and resulted in the publication of the paper ‘Birdsong dialect patterns explained using magnetic domains’[23]. The analysis and formulation of ideas were all produced in weekly meetings and simulations were constructed by both of us independently with comparison between results as confirmation.

Chapter 4

Interface Motion in Two Dimensions

In this chapter we investigate the spread of an interface on a two dimensional lattice structure. These models represent systems in which there is an invasive idea or opinion and the information spreads through the system. We look at models with nearest neighbour exchanges of information between fixed sites, resembling a system of face-to-face interactions between agents with low mobility. Concepts that exhibit this type of spreading include the adoption of new language and culture [66–68] and the erosion of surfaces [69, 70]. The intent is to gain a deeper quantitative understanding of how ideas spread and how the shape of the interface, determined by randomness in features of the agents, can impact wave speed.

4.1 Interface Shape and Speed

In this chapter we introduce the concept of memory into an invading opinion model on a two dimensional lattice. We investigate how individual sites in a network possessing a memory, and making their choices based upon it, impacts the velocity and shape of the invading wave. Memory has been shown to effect the behaviour of many mathematical models of social systems, including in games [71–75], language models [76, 77], spreading phenomena [78, 79], opinion dynamics [80] and it has also been studied in the context of the Ising Model [18]. In the work we present here, memory length may also be viewed as a threshold for a change of state. Threshold models [17, 81, 82] have applications in social dynamics [13], and also in physical settings such as rock weathering [69], where the threshold (or memory length) of a lattice site may be viewed as its strength, and the impact of wind or dust as the flow of social information. A fundamental concept to these models is that of surface roughening. We consider a smooth interface between two regions with different opinions, one of which is superior and invasive, one which is weaker and being usurped, and an exchange

of opinions across the interface. As the weaker opinion sites adopt the new opinion the interface roughens, surface area of the interface increases and the spread accelerates. The growing of a random surface can be described using the Kardar-Parisi-Zhang (KPZ) equation [25, 26]

We begin with simplified starting conditions and build upon that by looking at more complex arrangements, and the impact these changes have upon the dynamics of the system. We begin with an $m \times n$ lattice with each location occupied by one agent possessing a memory. After each time step the agents inspect the current opinions of their four nearest neighbours and remember them. They store that information for a limited memory length, with newer information replacing the oldest. This approach to site memory could be viewed as a threshold and is comparable to the dosage memory of a site in the generalised Watts model [83]. At the end of each time step the agents look over all opinions stored in their entire memory and make a conformist decision on what their opinion should be. This decision may be viewed as a modified form of majority (or minority) rule [14, 16, 60], but averaged over the memory of the decision maker. In the standard definition of the majority rule, agents inspect the current states of their neighbours, and make a decision based on these states. In our case we think of each agent as carrying a historical record of these states, and matching their state to the majority opinion in their historical record. Formally, we can define the model as follows. Letting $S_i(t) \in \{0,1\}$ be the state of lattice site i then the memory of this site is

$$m_i(t) = \frac{1}{4\tau_i} \sum_{s=0}^{\tau_i - 1} \sum_{j \in \langle i \rangle} S_j(t-s) \tag{4.1}$$

where τ_i is memory length of site i and $\langle i \rangle >$ denotes the four nearest neighbours of this site. This memory is an analogue of the dosage memory in the Watts model [17, 83]. The state of site i is then given by

$$S_i(t) = \begin{cases} 1 & \text{if } m_i(t) \geq \frac{1}{4} \\ 0 & \text{otherwise .} \end{cases} \tag{4.2}$$

The system has periodic boundary conditions at the vertical (left and right) edges with fixed boundaries along the horizontal edges. Initially the opinion setup of all agents is the down (‘0’) opinion except the agents on the bottom row which possess the invading up (‘1’) opinion. We initialise the system with all sites having a full memory of only their current state, that is to say that for sites in the bottom row $m_i(t) = 1$ and for all other sites $m_i(t) = 0$. This setup causes the first state flips to take some time rather than happen immediately at the first time step and makes a better comparison to a new idea invading a territory that has

only ever known the old idea. Due to the nature of the system's construction we need the agents to flip opinions when $1/4$ of their memory is filled with the invading up opinion or the wave would be incapable of moving. This is due to the fact that the first sites that flip from state 0 to 1 can at most only ever reach a value of $m = \frac{1}{4}$ as they have only have one up-state neighbour from below and three down-state neighbours. As a consequence of this it is impossible for an agent to flip away from the up opinion once adopted as there is no way for a new interaction memory to have less up-states in it than the oldest memory it replaces. Thus the invading wave is always connected and we can't have dispersed clusters of varying opinions or isolated agents changing opinion. If for each column we consider the highest agent of the up ('1') opinion to be at the front of the wave then we are interested in two main features: the velocity of the wave and the shape of the wave-front. Given that the agents all flip at the same memory proportion then the only variable that can influence these features is the individual agent's memory lengths.

4.1.1 Simplistic Model Construction

Throughout this section we will be looking at systems where individual agents have one of two memory lengths, the shorter α or the longer β. We begin by considering the simplest case of all; that is that all agents have the same memory length, say α. This model requires no simulation as we can simply deduce that after α time steps every agent in the lowest 'unflipped' row will convert to the invading opinion and after every subsequent α time steps the entirety of the next row will also flip. In this system it is simple to see that the velocity of the wavefront is the reciprocal of the memory length and that the invading wavefront is a constant flat line. In this simplest model every agent with the down opinion only ever receives influence of the new opinion from the agent directly behind them since, even though they observe the opinions of their neighbours to the side, they will be changing opinions at the exact same time step. It is this potential influence from the side neighbours that can lead to more interesting wave dynamics in systems with varied memory length patterns.

The next arrangement we consider is one in which the agents have a randomised mixture of both memory lengths α and β. Each agent is assigned memory length α with probability p or memory length β with probability $(1-p)$. We know the velocity and wave shape for the system when $p = 0$ and $p = 1$ but are interested in the wave dynamics as p varies between these values. We ran these simulations on a 200×200 lattice as it is large enough to dampen the effect of variance, but keeps the simulation time short, and set our memory lengths to $\alpha = 20$ and $\beta = 40$. As for the wave shape in this system we will be concerned with the

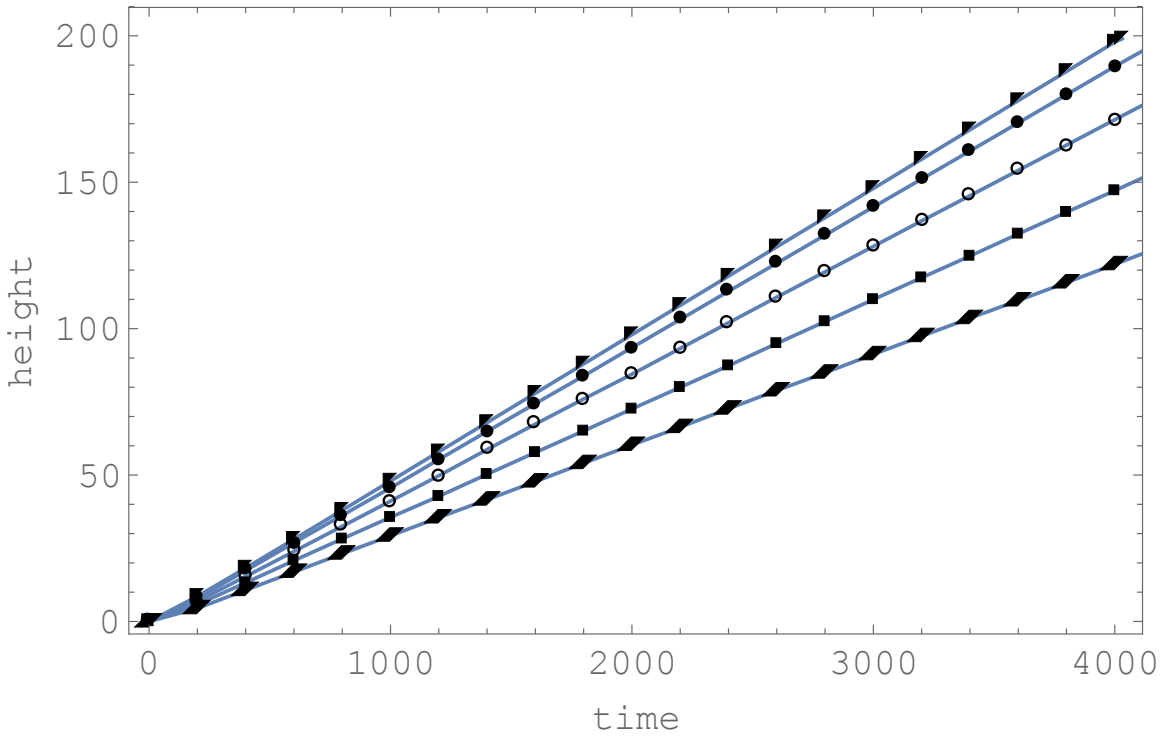


Fig. 4.1 Line plots showing the invasion progress over time for five randomised setups. The possible memory lengths are $\alpha = 20$ and $\beta = 40$ for which p is the probability of a site having memory length α and $(1-p)$ is the probability of a site having memory length β. Plotted systems are $p = 0.1$(diamonds), 0.3(squares), 0.5(empty circles), 0.7(filled circles) and 0.9(triangles). Higher p values correspond to faster invasion speeds .

roughness of the leading wave, calculated as the mean square distance between the front of each column and the mean wavefront. From the initial conditions it is clear that we can expect a flat line with some noise since there's no structured pattern to make the wave adopt a particular shape. The results are shown in plot (4.1).

The simulations show us that rather than the velocity being linear in relation to the mean memory length it speeds up with the presence of mixed memory lengths and once the proportion of short memory agents is high enough the wave travels at the fastest velocity. This is due to the previously mentioned feature that once memory lengths are mixed you have agents on the same row flipping opinions at different times, with the early flipping agents passing their influence on to their side neighbours and accelerating their opinion change. Thus not only do the shorter memory length agents speed up the wave by themselves flipping sooner but they also cause the longer memory agents to flip sooner. Another result to note here is that the roughness of the wavefront takes time to stabilise but eventually remains fixed.

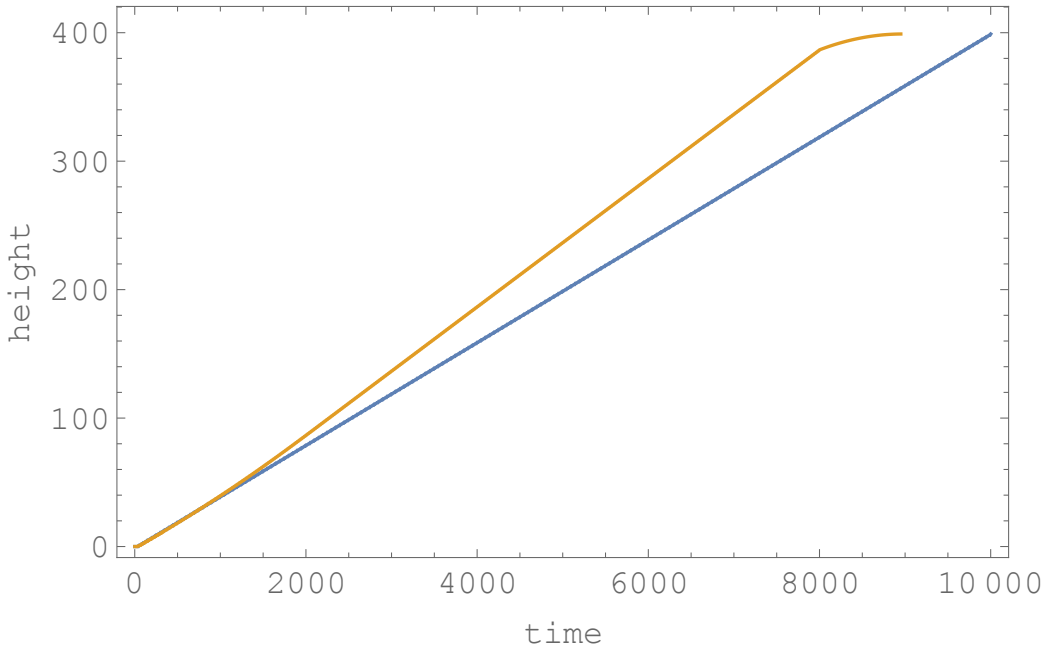


Fig. 4.2 Line plot showing the average height of the invading wave with respect to time. The orange steeper line corresponds to the stripe state system and the blue shallower line represents the system with alternating cells.

We next consider how we can manipulate the wave front velocity by choosing which agents receive the long and short memories, rather than a randomised assignment. It is apparent that the presence of short memory agents alongside their longer memory counterparts can accelerate the flipping process so we would expect different arrangements and configurations to produce different wave velocities and wavefront shapes. We begin with the two simplest arrangements possible, both with memory proportion $p = 0.5$, and with the same memory lengths as before, $\alpha = 20$ and $\beta = 40$. We will have a lattice with width of 200 and length of 400 for which system A will be a checkerboard pattern, where the memory length in every cell is different to its 4 nearest neighbours, and system B will be 2 distinct stripes in which all agents in columns 1-100 will have memory length α and all agents in columns 101-200 will have memory length β. You can see these results for the systems in plots (4.2) and (4.3).

The contrast between the two results show the importance of the lattice arrangement to the dynamics of the opinion wave. The two differing arrangements begin with the same velocity but once stabilised the stripe state accelerates to match the maximum potential velocity of $1/\alpha$. If we look at the shape of the invading wave in figure (4.4) we can garner a greater understanding of what is occurring at the wavefront boundary to cause this behaviour. Initially the wave invades faster through the short memory agents and slower through the

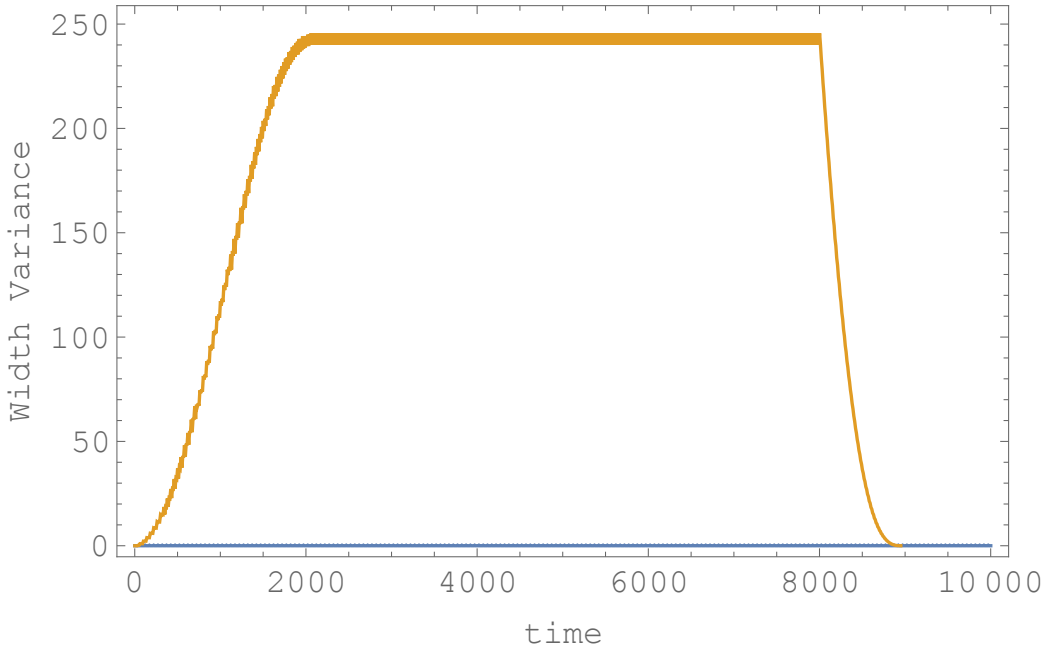


Fig. 4.3 Line plot showing the variance in invasion height across the width of the system with respect to time. The orange line represents the stripe state system and the blue line represents the alternating system. Both systems converge to a constant variance with the stripe state being significantly greater. The variance in the stripe state system eventually drops to zero as the wave reaches the end of the lattice.

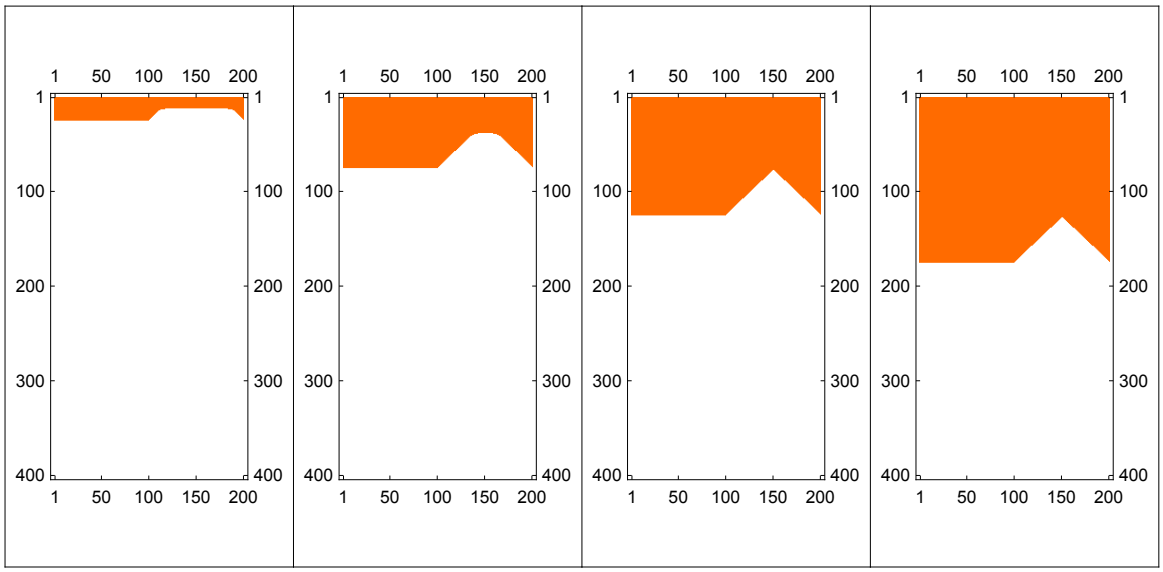


Fig. 4.4 Matrix plot of the invading wave travelling downwards on a striped system where orange sites are up and white sites are down. Columns 1 to 100 are short memory agents and columns 101 to 200 are long memory agents. These represent the system after 500, 1500, 2500 and 3500 time steps.

longer memory agents causing the wave variance to increase and the velocity to match the mean. As the system develops the long memory agents on the border of the stripe begin receiving twice as much influence as before, from below and the already flipped short memory agents to the side, causing the wave to pass through them at the same rate as the short memory agents. Once the two two slopes of up sites in (4.4) meet all of the long memory agents on the boundary are receiving twice the influence and the wave passes through every agent with maximum velocity.

We further investigate this effect by looking at the most extreme case; a similar system except this time with just one column of short memory agents with all other agents having the longer memory. We see in figures (4.5), (4.6) and (4.7) that this system exhibits the same behaviours. In this system the velocity begins as the lowest because the majority of the agents have a long memory and are only receiving the invading opinion from below, however once the system stabilises it reaches the same maximum velocity as the half striped system. Thus showing we can reach maximum velocity with a miniscule proportion of the population having the shorter memory length. These models highlight that, for systems with a repeated memory length pattern, eventually the roughness of the invading wave converges to a fixed point because the leading wavefront takes on a consistent shape. Since the wave shape converges to a fixed shape, and any column made of solely short agents must travel

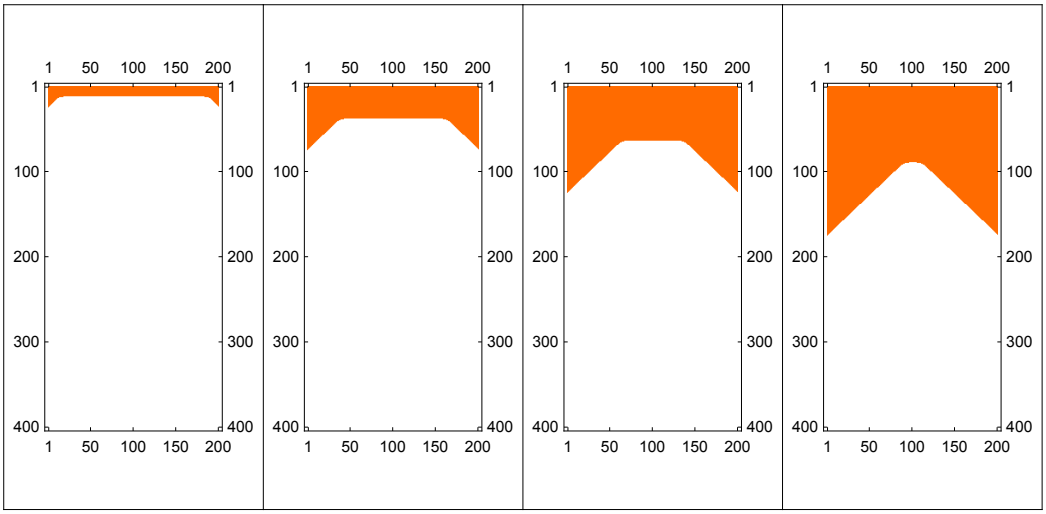


Fig. 4.5 Matrix plot of the invading wave on a slim striped system. Column 1 is made up of short memory agents and columns 2 to 200 are long memory agents. This shows the system after 500, 1500, 2500 and 3500 time steps.

at the maximum velocity, it is intuitive that once the system has been running long enough to become stable that the velocity of the wave will be at its maximum in all places. If this were not the case, because in some areas the wave travelled at a sub-maximal speed, then the wavefront shape would continue to deform.

In order to further investigate this phenomenon we take a closer look at the gradient of the invading wavefront on the boundary between long and short memory agents. From figures (4.4) and (4.5) it is clear in these systems that the gradient of the invading wave through the long memory agents is 1. Intuitively once the wavefront is stabilised the shape remains constant and makes a step up every α time steps, since it is travelling with maximum velocity. For a long memory agent on the border of the wavefront they will only receive influence from one side agent that flips before them and the agent behind them. Given that the wavefront has to move up every α time steps we know that every agent will receive exactly α influence in their stored memory from the agent behind them. Since they must get β encounters in their memory in order to flip it is apparent they must have $\beta - \alpha$ encounters with the already flipped agent to the side. This microscopic behaviour is illustrated in figure (4.8) and from this we can deduce that the slope gradient, the steepness of the boundary, in a striped system is $(\beta/\alpha - 1)$.

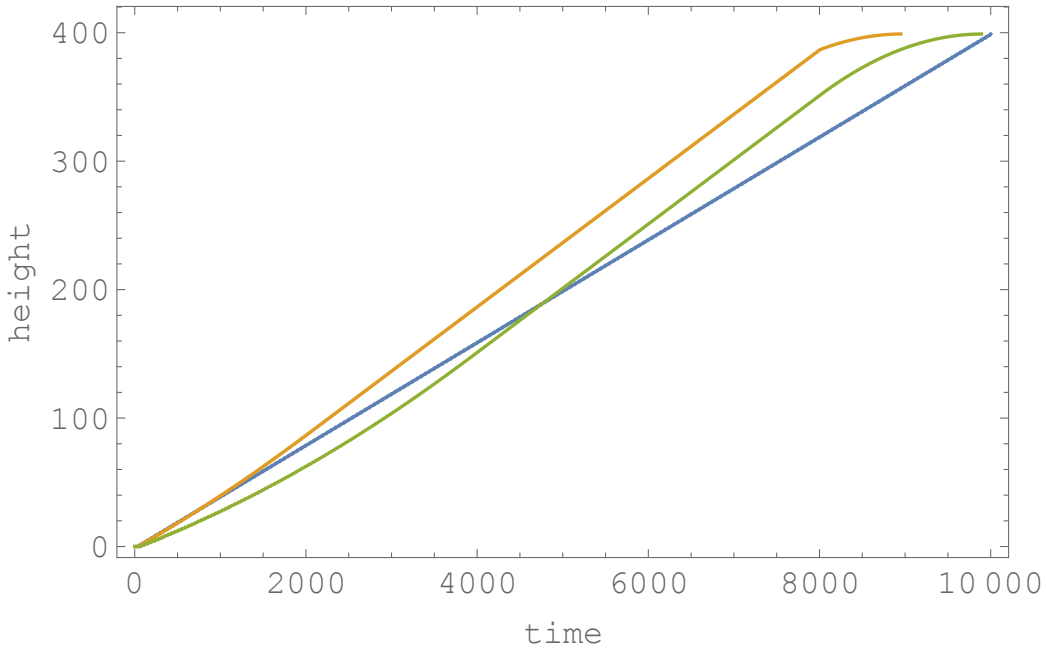


Fig. 4.6 Line plot showing the average height of the invading wave for three systems. The orange line represents a striped system with half short memory and half long memory, the green line represents a single column stripe setup and the blue line represents a system of alternating sites.

4.2 Tiled Systems

In order to further investigate the behaviour of a traveling wave through a large lattice system, still with two potential site thresholds, and randomised allocations we will begin by making the approximation that the system is made up of a smaller repeated square tile with width and height n. Intuitively as n grows large the number of tiles required to cover a specific large lattice reduces and the accuracy of this approximation increases until the point that the tile in consideration matches the original lattice and the tiled approximation will be exact. Due to the periodic boundary conditions on the side borders for this system we only ever need to consider the wave traveling through one column of tiles since all columns of tiles will behave identically. It also follows that once the wave travels through a small number of vertically stacked tiles it falls into a cyclical shape and velocity; only taking time to reach this state because the spread of the new idea at the bottom of the first tile is fixed and uniform whereas the influence spread on the transition between two tiles is dependent on the threshold allocations in the tile. Because the transfer of information into the first tile is different to the information transfer into subsequent tiles it is important to consider multiple tiles for the wave to reach a stable pattern.

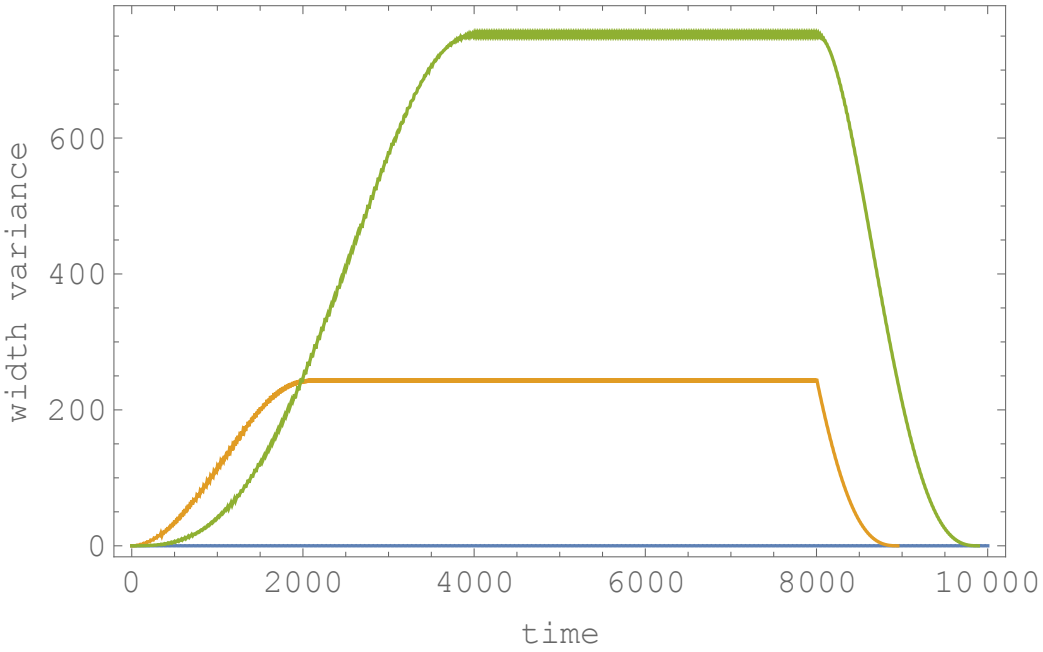


Fig. 4.7 Line plot showing the variance in invasion height across the width of three different systems. The orange line represents a striped system with half short memory and half long memory, the green line represents a single column stripe setup and the blue line represents a system of alternating sites.

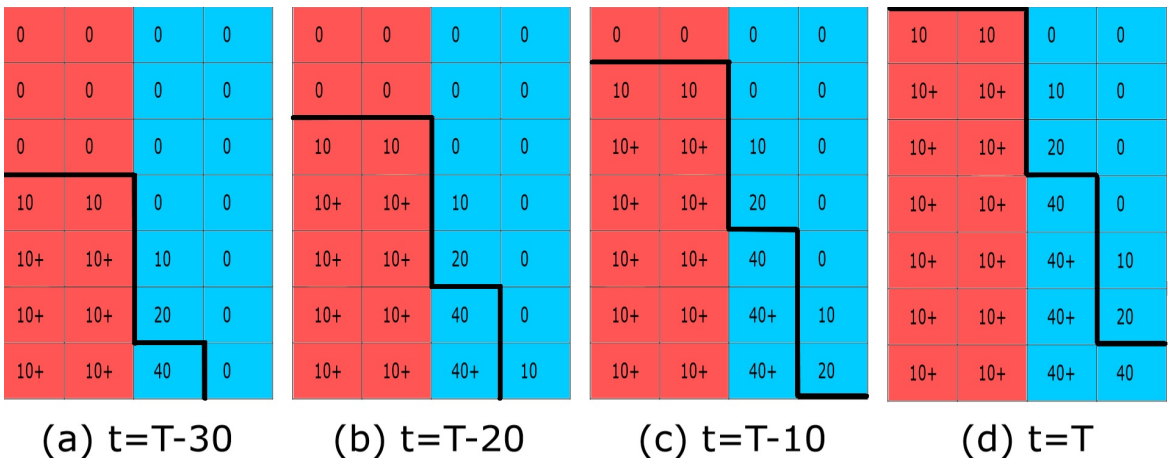


Fig. 4.8 Image depicting the wave boundary between long and short memory stripes of $\alpha = 10$ and $\beta = 40$. Blue cells are long memory agents, red cells are short memory agents, cells under the black line have converted and the value in the cell denotes the number of interactions with the invading opinion the agents have encountered.

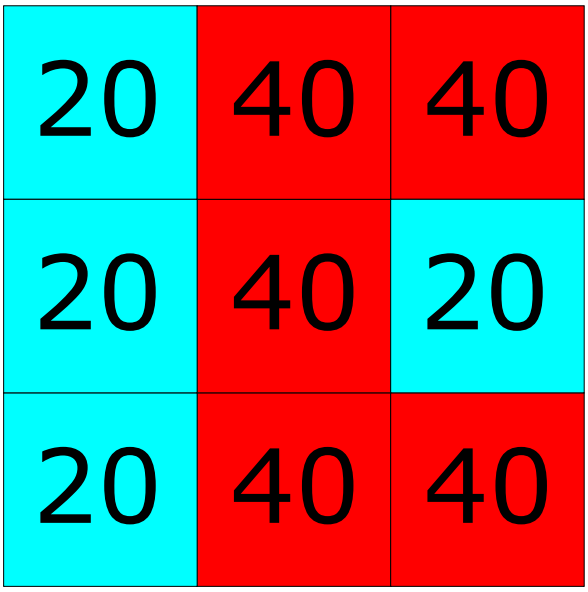


Fig. 4.9 An illustration of an example 3×3 tile. Example tile 1.

We begin with three example systems in which $n = 3$ with a fixed proportion of long and short memory sites and looking at the velocity of the wave in each case. For these examples the proportion of short memory agents is a fixed $\frac{4}{9}$. The tiles used for comparison are illustrated in figures (4.9), (4.10) and (4.11) and when running an invading wave through a vertical set of ten of each tile we get the graph (4.12) highlighting the average height of the wavefront against time.

As expected from the previous section the steady velocity of the system made up of the first tile is $\frac{1}{20}$ given that it will consist of one column made up entirely of short memory agents. As previously shown this column will travel at the highest possible velocity and the rest of the system will be pulled along to match it. This is an inherent issue with attempting to use tiles with smaller values of n to approximate the wave behaviour in a bigger system; with random tiles there are minimal patterns for the short and long memory agents to take and the disparity between different patterns is substantial. For comparison the velocity of the wave in systems comprised of tiles two and three is approximately $\frac{1}{27}$, a discrepancy of over 30% from the fastest tile. As we increase the value of n the likelihood of a tile having a column with only short memory agents drops significantly. Due to this we would expect this tiled system approach to be highly accurate and for the majority of randomised tiles to have very comparable results.

Using this tiled approach we do find a consistent trend; the wave velocity and shape enters a fixed pattern after traveling through the first two tiles. As previously mentioned the

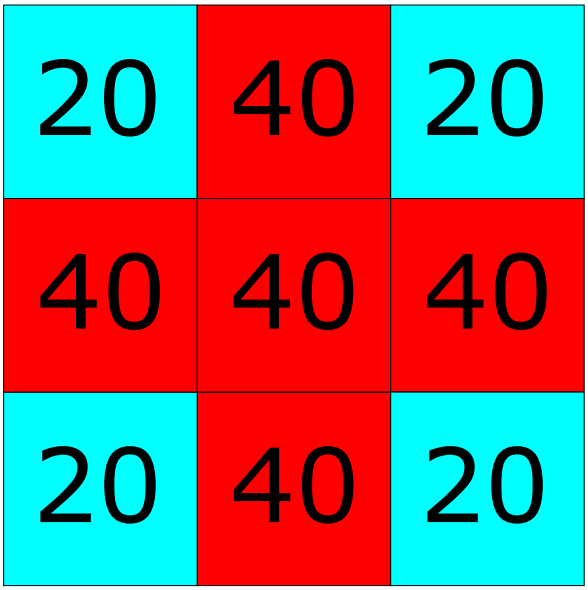


Fig. 4.10 An illustration of an example 3×3 tile. Example tile 2.

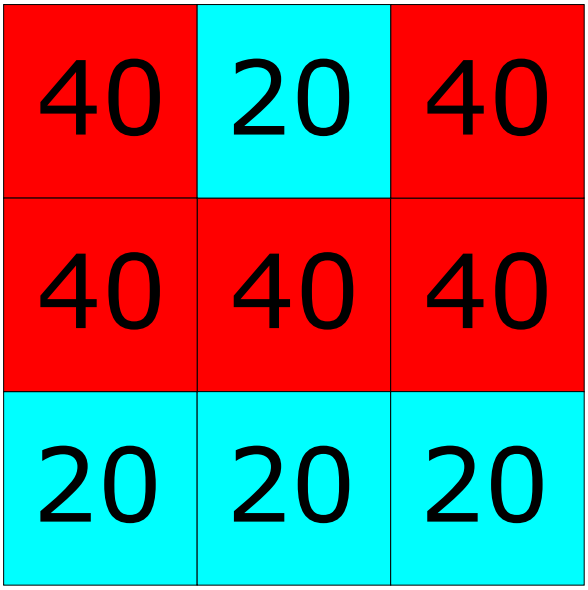


Fig. 4.11 An illustration of an example 3×3 tile. Example tile 3.

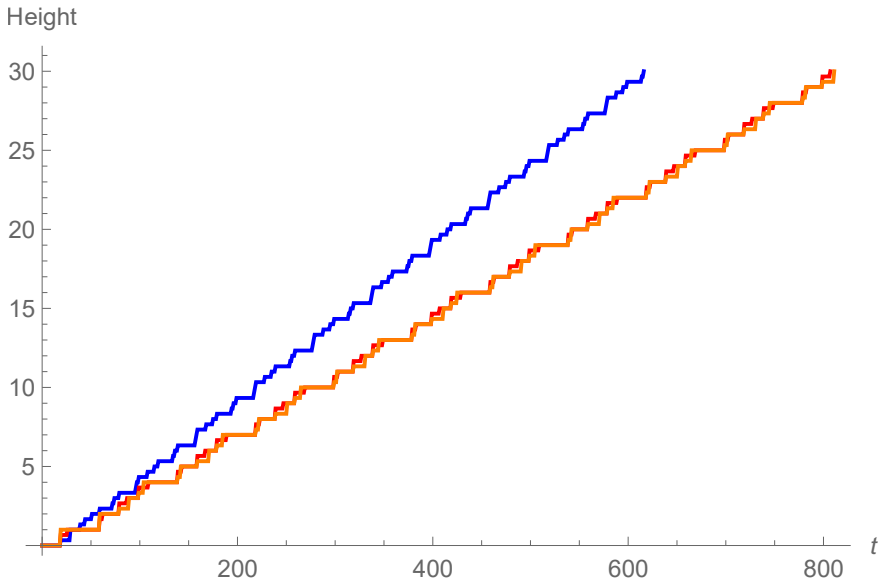


Fig. 4.12 A graph showing the average height of the invading wave through a vertical set of 10 of each example tile taken at every time step. Blue represents tile 1, red represents tile 2 and orange represents tile 3.

initial wave spread into the first tile is uniform whereas the information spread between the top of the first tile and the bottom of the second tile is dependent on the times at which the sites in the top row of tile one flipped. However in all observed cases this minor impact is eliminated before the wave reaches the top of the third tile. As such the difference in flip times between identical sites at the top of the second and third tiles can already give us the velocity of a system. An example of this is shown in figure(4.13) where the time differential between sites three rows behind has reached 80 and remains fixed at that value.

Given a randomised tile of sufficiently low n it is simple to calculate the exact flip times for the earlier rows by hand. Ideally we can find an optimal tile size for which calculating by hand is efficient and quick and the accuracy to an entirely randomised large lattice is high and sufficient. In order to compare the accuracy of tiled systems we first simulate a large lattice with height 1000 and width 100 and a probability of 0.4 for any given site to possess a short memory. In the previous examples with the 3 by 3 tiles the patterns were chosen specifically to have four short memory sites and five long memory states to get an understanding of the process. In this new construction each site has a probability of having a long or short memory so the exact number of long and short memory sites can vary. The height progression of this system is shown on the graph (4.14) and gives us an average wavefront velocity, to three significant figures, of 0.0399 rows per unit time. This system size is chosen as it is large

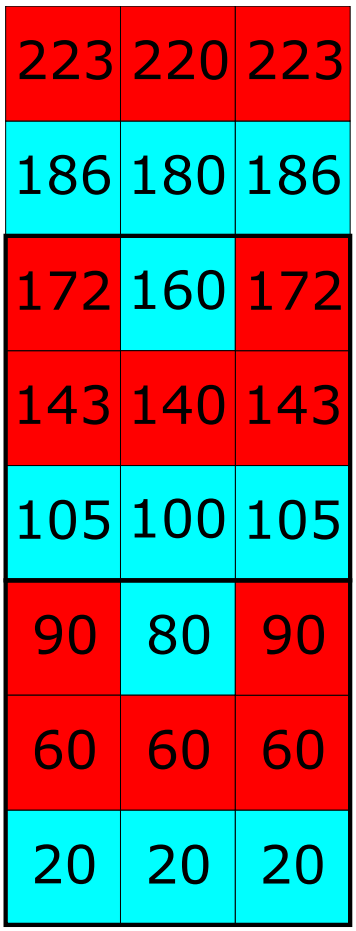


Fig. 4.13 An illustration of the exact flip times for sites from the bottom of a system made up of tile 2. In this example the wave has already reached its fixed velocity on the second row of the third tile.

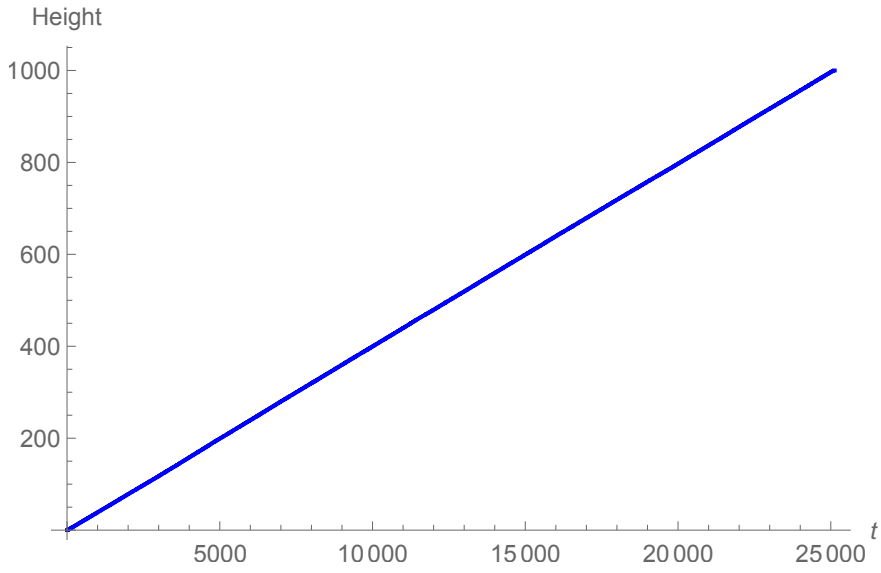


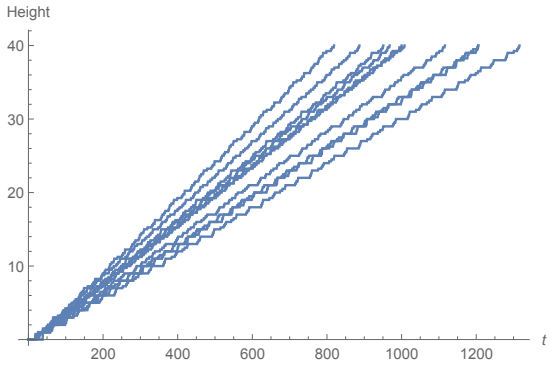
Fig. 4.14 A graph illustrating the height of an invading wave into a large lattice with variable thresholds. Sites can have short memory with probability $p = 0.4$ and long memory with probability $(1-p) = 0.6$. Short memory sites have threshold 20 and long memory sites have threshold 40.

n	Average Velocity	Standard Deviation
4	0.03931	0.006015
6	0.03884	0.003285
8	0.04047	0.002688
10	0.03991	0.002491

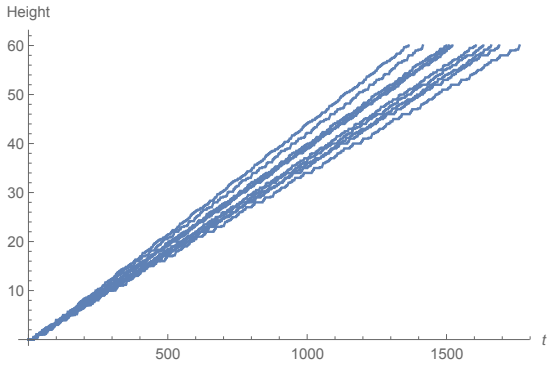
Table 4.1 Table showing the average wavefront velocity and standard deviations for ten simulated systems with varied tile size n. All values to 4 significant figures.

enough to remove any small number effects and repeated simulations at this system size generate similar results.

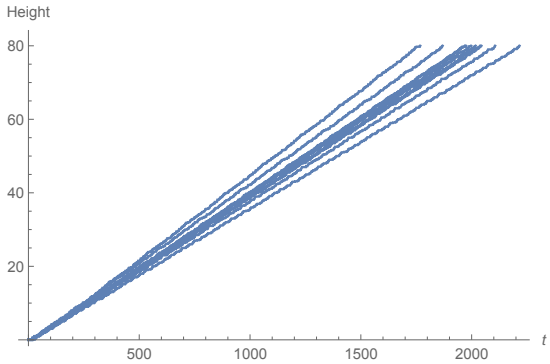
We want to compare how accurate the tiled approach is for differing values of n to this entirely randomised large scale lattice. In order to do this we created randomised tiles, still with $p = 0.4$ as the proportion of short memory sites, for $n = 4, n = 6, n = 8$ and $n = 10$. We aim to calculate the velocity of the wave through a sample of systems with each of these tile sizes and compare the behaviour of the velocities in each case. Graphs for height progression over time for 10 systems at each tile size can be seen in plots (4.15a), (4.15b), (4.15c) and (4.15d). Table (4.1) contains the information for the average velocity for systems of each tile size and the standard deviation across the 10 systems.



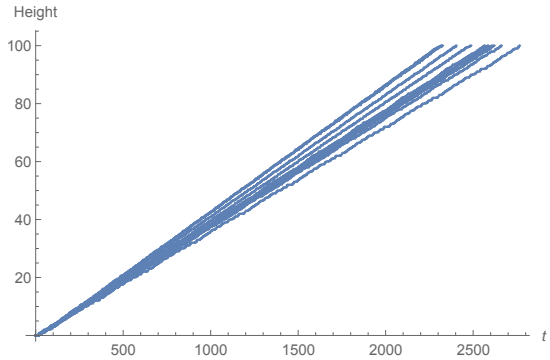
(a) Tiles of size $n = 4$.



(b) Tiles of size $n = 6$.



(c) Tiles of size $n = 8$



(d) Tiles of size $n = 10$

Fig. 4.15 Four graphs showing the wave front height over time of systems with randomised tiles for four different tile sizes n. In each case $p = 0.4$ is fixed, the proportion of sites with short memories.

These results highlight how effective the tiled approach is at generating an approximation of the wavefront velocity for a system of much larger size. As you average over a sufficient number of randomised tiles the approximate velocity is highly accurate even for tiles of size 4. Intuitively the standard deviation and variance of wavefront velocities is greater in smaller tiled systems since there is a distinct possibility of randomly creating tiles with entirely short agents making the wave rapid in comparison to the average. As such in order to use smaller tiles to approximate large scale behaviour you would ideally use a greater number of tiles. However given the simplicity of calculating the flip times by hand for small tiles this method is still incredibly efficient.

4.3 Interface Motion with Disorder

The previous systems highlight the impact of variable memory thresholds on the shape of the wavefront and the impact this has on the wave velocity. However all of those models looked at systems with only two potential memory lengths which is a significant restriction. If we are to consider the wave as an opinion traveling through a living population then it would be more reasonable to consider threshold length as a normally distributed attribute with a mean for the population and standard deviation allowing for variability between individuals. We will continue looking at the wave traveling through a two dimensional lattice with the same method and neighbour interactions as previous models.

In the construction of this model with equation (4.1) we declare each agent to have a memory length, τ, and as such they store 4τ pieces of information from observing each of the nearest neighbours. However due to the previous restriction of a site having to flip as soon as $1/4$ of its memory was filled with the new invading wave it is not possible for a site to replace a memory of the invading wave before it has flipped. That is to say that a site in the down state, that first encounters the invading wave at time t, will always flip at or before time $t+\tau_i$, before that first interaction is forgotten. As such a site with memory length τ with a required memory proportion to flip of $1/4$ is analogous to a site with an interaction threshold of τ. For simplicity we will continue the analysis by using the idea of a site threshold in which a down site will flip state once it encounters the new wave τ times.

As we are now considering sites to have their own thresholds, taken from a distribution rather than being assigned a binary choice of long or short, we will be more concerned with the exact thresholds at certain locations and the correlation between neighbours. As such we will begin using (i,j) notation for positioning to more accurately describe specific sites. We begin by letting $\tau_{i,j} = \lfloor z \rfloor$, where $z \sim \mathcal{N}(\mu, \sigma^2)$ be the individual threshold at location

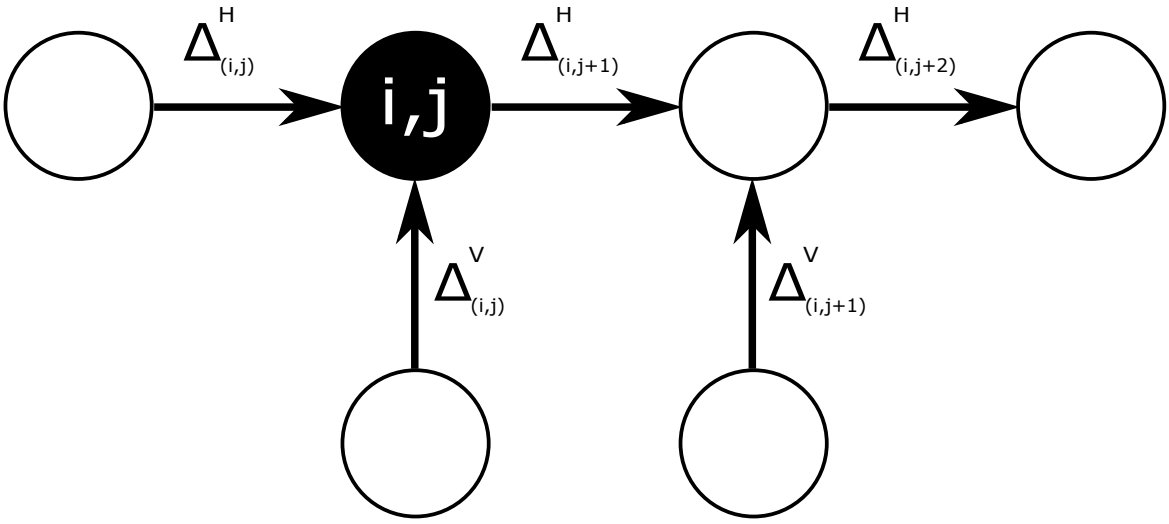


Fig. 4.16 Illustration of the time differential between sites flipping states and the directions representing the positive direction for the delta function and influence given.

(i, j) and $T_{i,j}$ be the flip time for site (i, j), the time step at which the opinion at the site flips to the dominant invading idea. By choosing μ to be substantially large we can approximate $\tau_{i,j}$ as a continuous normal distribution. Intuitively if $\sigma = 0$ then all sites have threshold μ and the wavefront is flat and the wave is always at minimal speed of $1/\mu$. By choosing low values of σ relative to μ we can be sure that the wave will be unable to bypass around low threshold sites and as such $T_{i+1,j} > T_{i,j}$ for all sites. For example if sites have a mean threshold of $1,000,000$ with a standard deviation of 50 then the possibility that a site can flip states before the one behind it is infinitesimally small.

We define the following delta terms to represent the flip time differentials between neighbouring sites, horizontally and vertically, as illustrated in (4.16).

$$\Delta^H_{i,j} = T_{i,j} - T_{i,j-1} \tag{4.3}$$

$$\Delta^V_{i,j} = T_{i,j} - T_{i-1,j}. \tag{4.4}$$

Given that the system has periodic boundary conditions on the edges it must be the case that the sum of all Δ^H terms must be 0 and as such $\mathbb{E}[\Delta^H] = 0$. Since Δ^V is the time differential between vertically aligned neighbours it is the case that $\mathbb{E}[\Delta^V] = 1/v$, where v is the velocity of the invading wave. We aim to calculate the velocity of the invading wave given the mean and variance of the original approximate threshold distribution, $\tau_{i,j} = \lfloor z \rfloor$, where $z \sim \mathcal{N}(\mu, \sigma^2)$.

From the definition of the system, site (i, j) will flip to the invading opinion once it encounters the invading opinion $\tau_{i,j}$ times from its nearest neighbours. A site will receive influence of the new opinion from all neighbours that flip at an earlier time step and, since the sites in a column flip in ascending order, a site can not receive influence from above. As such we need only consider potential influence from the agent below the two horizontal neighbours. We define the piecewise function J as

$$J(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases} \tag{4.5}$$

and therefore

$$\tau_{i,j} = \Delta^V_{i,j} + J(\Delta^H_{i,j}) + J(-\Delta^H_{i,j+1}). \tag{4.6}$$

Each $J(\Delta^H)$ represents the influence received from a horizontal neighbour and Δ^V represents the vertical flip time differential and as such the influence received from the site below. This is because at the time of flipping a site begins spreading influence at a rate of 1 per unit time. As such the time differential between neighbouring sites flipping is the influence that will be transferred between them. However this influence transfer is only in the direction of the earlier flipper passing influence on to the later flipper, hence the piecewise function J. Since the sites in a column will always flip in order, Δ^V will always be positive and J is not required there. Since the three terms on the right hand side of (4.6) represent the influence received from the relevant neighbours of (i, j) it follows that these must sum to an individual site's threshold in order for it to flip. In order to continue to analysis we make the reasonable assumptions that the delta terms are normally distributed with mean and variance as follows

$$\Delta^V \sim \mathcal{N}(1/v, \sigma_V^2) \tag{4.7}$$

$$\Delta^H \sim \mathcal{N}(0, \sigma_H^2). \tag{4.8}$$

We can utilise this expression for $J(\Delta^H)$ to take moments of the equation (4.6) and calculate some approximations for the velocity for the wave in terms σ and μ. Firstly we compute the first three moments of $J(\Delta^H)$ as follows

$$\mathbb{E}[J(\Delta^H)] = \int_0^\infty \frac{\Delta}{\sigma_H\sqrt{2\pi}} e^{-\frac{\Delta^2}{2\sigma_H^2}} d\Delta = \frac{\sigma_H}{\sqrt{2\pi}}, \tag{4.9}$$

$$\mathbb{E}[J^2(\Delta^H)] = \int_0^\infty \frac{\Delta^2}{\sigma_H\sqrt{2\pi}} e^{-\frac{\Delta^2}{2\sigma_H^2}} d\Delta = \frac{\sigma_H^2}{2}, \tag{4.10}$$

$$\mathbb{E}[J^3(\Delta^H)] = \int_0^\infty \frac{\Delta^3}{\sigma_H\sqrt{2\pi}} e^{-\frac{\Delta^2}{2\sigma_H^2}} d\Delta = \sqrt{\frac{2}{\pi}}\sigma_H^3. \tag{4.11}$$

In our first approximation for the velocity of the wave we will make the optimistic assumption that the delta terms in equation (4.6) are independent. Doing so allows us to take simple expectations as follows.

$$\begin{aligned} \mathbb{E}(\tau_{i,j}) &= \mathbb{E}[\Delta_{i,j}^V + J(\Delta_{i,j}^H) + J(-\Delta_{i,j+1}^H)] \\ \mu &= \frac{1}{v} + \frac{2\sigma_H}{\sqrt{2\pi}}, \end{aligned} \tag{4.12}$$

$$\begin{aligned} \mathbb{E}(\tau_{i,j}^2) &= \mathbb{E}[(\Delta_{i,j}^V + J(\Delta_{i,j}^H) + J(-\Delta_{i,j+1}^H))^2] \\ \mu^2 + \sigma^2 &= \frac{1}{v^2} + \sigma_V^2 + \frac{4\sigma_H}{v\sqrt{2\pi}} + (1+\frac{1}{\pi})\sigma_H^2, \end{aligned} \tag{4.13}$$

$$\begin{aligned} \mathbb{E}(\tau_{i,j}^3) &= \mathbb{E}[(\Delta_{i,j}^V + J(\Delta_{i,j}^H) + J(-\Delta_{i,j+1}^H))^3] \\ \mu^3 + 3\mu\sigma^2 &= \frac{1}{v^3} + \frac{3\sigma_V^2}{v} + 6(\sigma_V^2 + \frac{1}{v^2})\frac{\sigma_H}{\sqrt{2\pi}} + \frac{3}{v}(1+\frac{1}{\pi})\sigma_H^2 + \frac{7\sigma_H^3}{\sqrt{2\pi}}. \end{aligned} \tag{4.14}$$

If we solve these equations simultaneously for values of μ and σ we get a first estimate for the velocity of the wave. Unfortunately solving these equations for $\mu = 100$ and $\sigma = 10$ produces the solution that $\frac{1}{v} = 100$. This would imply that the variation in site memory thresholds, and the subsequent surface roughening, has no accelerating effect on the wavefront which we know to be incorrect. We first attempt to resolve this by removing the assumption that adjacent horizontal delta terms are uncorrelated, which seems intuitive, since we would expect that if a site has a low threshold and flips early then its corresponding deltas would be correlated.

Instead let us now assume that adjacent horizontal delta terms are part of a bivariate half normal distribution, $K(x,y)$ with probability density function

$$\frac{1}{2\pi\sigma_H^2\sqrt{1-\rho^2}}e^{-\frac{x^2+y2-2xy\rho}{2\sigma_H^2(1-\rho^2)}}, \tag{4.15}$$

for non-negative x and y and $0<\rho<1$. In this equation $x=J(\Delta_{i,j}^H)$, $y=J(-\Delta_{i,j+1}^H)$ and ρ is the correlation between the adjacent deltas. In order to take moments as before we need to calculate some additional expectations, A and B.

$$\begin{aligned}\mathbb{E}[K(x,y)] &= \int_0^\infty\int_0^\infty \frac{xy}{2\pi\sigma_H^2\sqrt{1-\rho^2}}e^{-\frac{x^2+y2-2xy\rho}{2\sigma_H^2(1-\rho^2)}}dxdy\\ A(\sigma_H,\rho) &= \frac{\rho\sigma_H^2\left(\pi+\sqrt{-1+\frac{1}{\rho^2}}-\arccos[\rho]\right)}{2\pi},\end{aligned} \tag{4.16}$$

$$\begin{aligned}\mathbb{E}[K(x^2,y)] &= \int_0^\infty\int_0^\infty \frac{x^2y}{2\pi\sigma_H^2\sqrt{1-\rho^2}}e^{-\frac{x^2+y2-2xy\rho}{2\sigma_H^2(1-\rho^2)}}dxdy\\ B(\sigma_H,\rho) &= \frac{(1+\rho)^2\sigma_H^3}{2\sqrt{2\pi}}.\end{aligned} \tag{4.17}$$

By utilising these bivariate expectations we can recalculate the moments from equations (4.12)-(4.14) this time under the assumption of correlated horizontal deltas.

$$\begin{aligned}\mathbb{E}(\tau_{i,j}) &= \mathbb{E}[\Delta_{i,j}^V + J(\Delta_{i,j}^H) + J(-\Delta_{i,j+1}^H)]\\ \mu &= \frac{1}{v}+\frac{2\sigma_H}{\sqrt{2\pi}},\end{aligned} \tag{4.18}$$

$$\begin{aligned}\mathbb{E}(\tau_{i,j}^2) &= \mathbb{E}[(\Delta_{i,j}^V + J(\Delta_{i,j}^H) + J(-\Delta_{i,j+1}^H))^2]\\ \mu^2+\sigma^2 &= \frac{1}{v^2}+\sigma_V^2+\sigma_H^2+\frac{4\sigma_H}{v\sqrt{2\pi}}+2A(\sigma_H,\rho),\end{aligned} \tag{4.19}$$

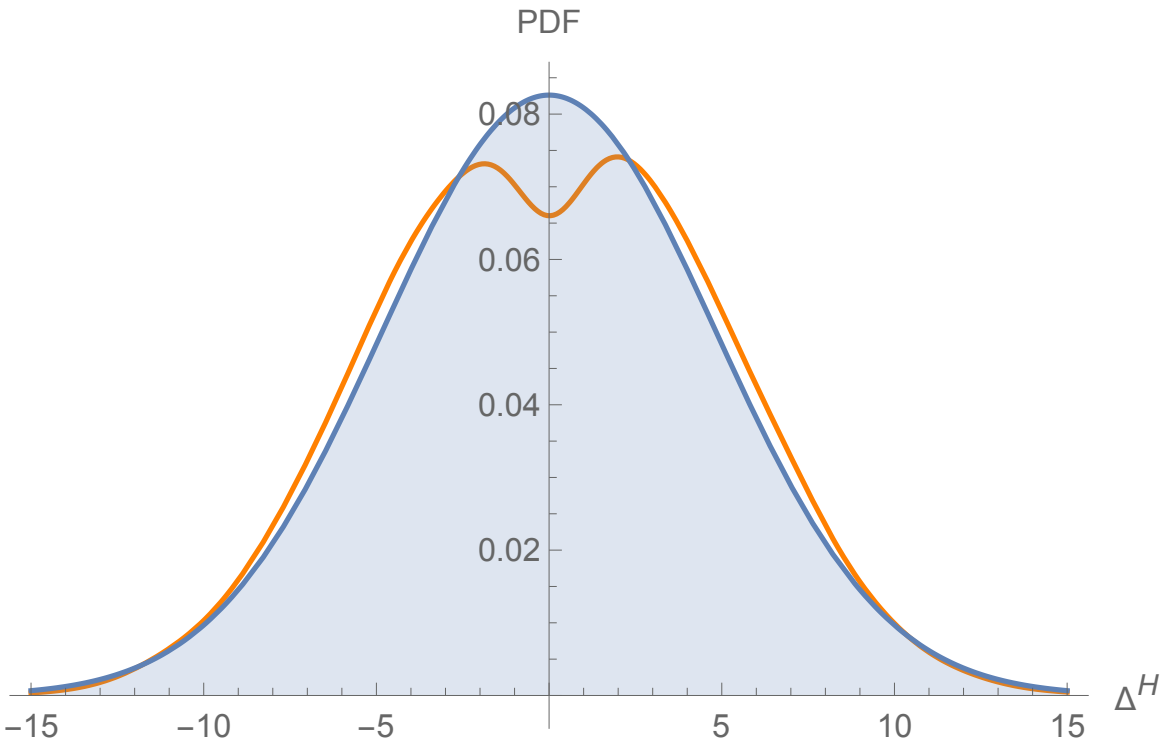


Fig. 4.17 A plot of the smoothed distribution for Δ^H against a Gaussian normal with the same mean and standard deviation. The orange plot represents Δ^H from simulated results and the blue plot represents the normal distribution.

$$\mathbb{E}(\tau_{i,j}^3) = \mathbb{E}[(\Delta_{i,j}^V + J(\Delta_{i,j}^H) + J(-\Delta_{i,j+1}^H))^3]$$
$$\mu^3 + 3\mu\sigma^2 = \frac{1}{v^3} + \frac{3\sigma_V^2}{v} + 6(\sigma_V^2 + \frac{1}{v^2})\frac{\sigma_H}{\sqrt{2\pi}} + \frac{3}{v}\sigma_H^2 + \frac{4\sigma_H^3}{\sqrt{2\pi}} + \frac{6}{v}A(\sigma_H,\rho) + 6B(\sigma_H,\rho). \tag{4.20}$$

Unfortunately in solving this new set of simultaneous equations for known values of μ and σ we still produce the only viable solution that $\frac{1}{v} = 100$, which we know to be incorrect. Given that this method should be correct it must be the case that some of our underlying assumptions are false. In order to test this we simulated a system with width of 50 and height of 1000 with $\mu = 100$ and $\sigma = 5$ to observe the actual distributions of Δ^H and Δ^V.

Evidently from plots (4.17) and (4.18) the assumption of normality, especially in the case of Δ^H is erroneous since it appears to be bi-modal, which would have a significant impact on the moments taken in all previous equations. The assumption of normality appears reasonable in the case Δ^V however. In order to use the moments technique above we would

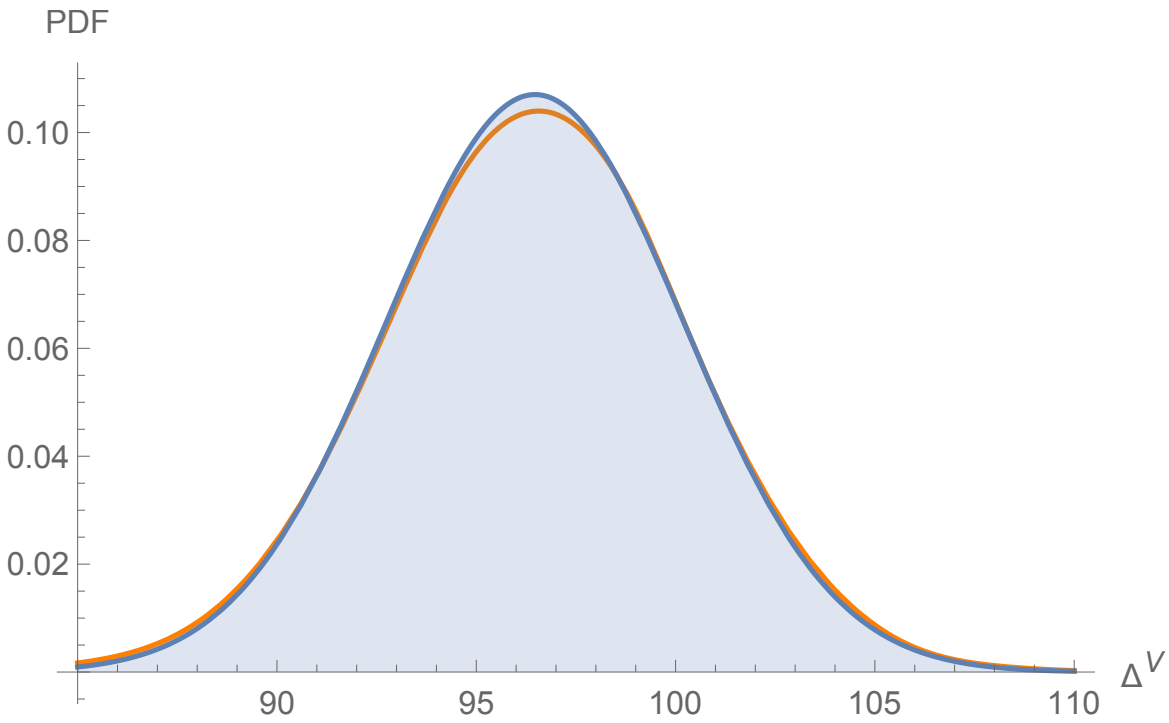


Fig. 4.18 A plot of the smoothed distribution for Δ^V against a Gaussian normal with the same mean and standard deviation. The orange plot represents Δ^V and the blue plot represents the normal distribution.

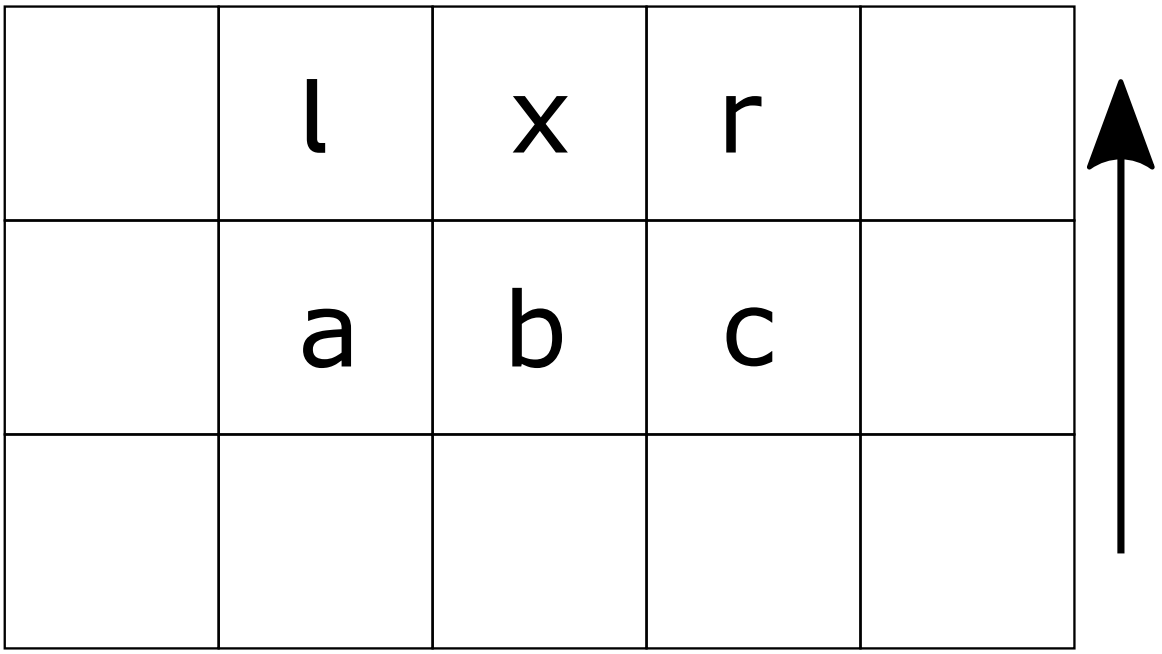


Fig. 4.19 An illustration of the tiles with an important impact on the time differential t_x.

need to ascertain the exact distribution for Δ^H, something beyond the current scope of this work.

We will instead look at an alternative method of approximating the velocity for this system based upon a mean field approach. Let us consider an arbitrary point in the middle of our system, site x, and let t_x be the time differential between site x flipping and the site behind flipping. The expected velocity of the wave is simply $\frac{1}{t_x}$. We begin by considering the simplest version of the system and then progressively introduce the variability in site thresholds that cause the complexity in calculating the velocity. From figure (4.19) we are interested in t_x and will begin with all named sites having fixed thresholds before introducing distributions for them. For each version of the system we attempt to compute t_x analytically allowing us to compare each approximation to the realised simulation results. We will begin with the assumption that fixed threshold sites have a threshold of $\tau = \mu$ and variable sites have threshold $\tau = \mu + b$ where b is taken from any distribution of choice. We again make the assumption that μ is substantially large and our discrete system can be treated as though it is continuous.

With this in mind let us again consider the simplest possible system in which every site (i, j) on the lattice has fixed threshold $\tau_{i,j} = \mu$. If we say that site x is on row n of the system then we know that all other agents on the same row flip at time $t = \mu n$, all sites on the line below flip at $t = \mu(n-1)$ and so on. Evidently in this system $t_x = \mu$ since, as

previously discussed, this system will exhibit no accelerated velocity and the wave will progress through the system as a uniform flat line. Thus in this version of the system, the zero order approximation, the velocity of the wave is exactly $1/\mu$.

Next let us consider τ_x being variable while all other sites remain with fixed thresholds. Given all the other sites have fixed thresholds of μ there are two cases to consider; $\tau_x < \mu$ and $\tau_x > \mu$. We define $I_x(t)$ to be the cumulative dose of the invasive opinion that site x has received at time t where

$$I_x(t) = \sum_{s=0}^{t} \sum_{j \in \langle i \rangle} S_j(t-s) \tag{4.21}$$

and we define T_x to be the exact time at which site x flips. Therefore $t_x = T_x - T_b$. If we let $I_x(t)$ be the information received at site x at time t then we know that once $I_x(t) = \tau_x$ site x flips. It $\tau_x < \mu$ we know that site x flips before its neighbours and as such it will only receive influence from the site directly below, which flipped at time $t = \mu(n-1)$. If $\tau_x > \mu$ then the site x will receive influence from the site behind until $t = \mu n$, at which point the neighbours will flip and the amount of influence being received is tripled. At the time of site x flipping

$$I_x(T_x) = \tau_x = \begin{cases} (T_x - \mu(n-1)) & \text{if } \tau_x \leq \mu \\ (T_x - \mu(n-1)) + 2(T_x - \mu n) & \text{if } \tau_x > \mu \end{cases} \tag{4.22}$$

and noting that $t_x = T_x - T_b$ where $T_b = \mu(n-1)$ allows us to rearrange to get

$$t_x = \begin{cases} \tau_x & \text{if } \tau_x \leq \mu \\ \mu + \frac{\tau_x - \mu}{3} & \text{if } \tau_x > \mu. \end{cases} \tag{4.23}$$

From our definition $\tau = \mu + b$ where b is from any distribution. If, for example, we allow b to be from a normal distribution with mean 0 and standard deviation σ then we can introduce a variable $p \sim \mathcal{N}(0, \sigma) = (\tau - \mu)$. Clearly in this case

$$t_x = \begin{cases} \mu + p & \text{if } \tau_x \leq \mu \\ \mu + \frac{p}{3} & \text{if } \tau_x > \mu \end{cases} \tag{4.24}$$

enabling us to perform the following integrations over the correct limits,

$$\int_{-\infty}^{0} \frac{pe^{\frac{-p^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dp = -\frac{\sigma}{\sqrt{2\pi}}$$
$$\int_{0}^{\infty} \frac{pe^{\frac{-p^2}{2\sigma^2}}}{3\sigma\sqrt{2\pi}} dp = \frac{\sigma}{3\sqrt{2\pi}}. \tag{4.25}$$

This allows us to make an analytical approximation for the time differential between site x and the site behind and with it a first approximation for the velocity of the wave using $v = \frac{1}{t_x}$,

$$v = \left(\mu - \frac{2\sigma}{3\sqrt{2\pi}}\right)^{-1}. \tag{4.26}$$

As an indicator of the accuracy of this approximation we make comparisons to the results from simulation. For a system with $\tau \sim \mathcal{N}(100,5)$ this approximation gives us $v = 1/98.7$ whereas the simulated result is $v = 1/96.5$ indicating that this approximation technique begins to predict an acceleration in wave speed but has a large room for improvement. In order to do this we increase the number of sites with variable thresholds.

We now extend on this by introducing threshold variability in more sites to create a second order approximation. We begin by looking at the case in which τ_x, τ_l and τ_r are variable with all other sites having mean threshold and attempt to calculate how accurate of an approximation this makes. In this case we immediately know $T_a = T_b = T_c = \mu(n-1)$ since there is no stochasticity in the system at this time. There are multiple cases that we need to consider in order to be exhaustive of which sites will be making an impact on t_x.

1. Site x flips before sites l and r and as such only receives influence from site b. This occurs when $\tau_x < \min(\tau_l, \tau_r)$.

2. Site x receives influence from site l which also received influence from its outer neighbour and behind. $\mu < \tau_l < \tau_x < \tau_r$.

3. Site x receives influence from site r which also received influence from its outer neighbour and behind. $\mu < \tau_r < \tau_x < \tau_l$.

4. Site x receives influence from site l which only received influence from behind. $\tau_l < \tau_x < \tau_r$ and $\tau_l < \mu$.

5. Site x receives influence from site r which only received influence from behind. $\tau_r < \tau_x < \tau_l$ and $\tau_r < \mu$.

6. Site x receives influence from sites l and r which both only received influence from behind. $\mu > \max(\tau_l, \tau_r)$ and $\tau_x > \max(\tau_l, \tau_r)$.

7. Site x receives influence from sites l and r which both received influence from outer neighbours and behind. $\mu < \min(\tau_l, \tau_r)$ and $\tau_x > \max(\tau_l, \tau_r)$.

8. Site x receives influence from sites l and r which both received influence from behind and only l received influence from the outer neighbour. $(\tau_r < \mu < \tau_l)$ and $\tau_x > \max(\tau_l, \tau_r)$.

9. Site x receives influence from sites l and r which both received influence from behind and only r received influence from the outer neighbour. $(\tau_l < \mu < \tau_r)$ and $\tau_x > \max(\tau_l, \tau_r)$.

Clearly there are many cases, but we can reduce them by acknowledging the horizontal symmetry which leads to multiple pairs of cases being identical. Those pairs are cases 2 and 3, cases 4 and 5 and cases 8 and 9. For each unique case we need to calculate t_x in order to get an overall approximation for the system velocity. We use $I_x(t)$ to denote the influence received by site x at time t.

Case 1: We can simply calculate t_x using the influence received by site x at time of flip is

$$\begin{aligned} I_x(T_x) = \tau_x &= T_x - T_b \\ &= t_x \end{aligned}$$

and so

$$t_x = \tau_x. \tag{4.27}$$

Case 2: In this case we use ll to denote the outer neighbour of site l. The influence received by site l at the time of its flip can be expressed as

$$\begin{aligned} I_l(Y_l) = \tau_l &= (T_l - T_a) + (T_l - T_{ll}) \\ &= T_l - \mu(n-1) + T_l - \mu n \end{aligned}$$

and so

$$T_l = \mu n + \frac{\tau_l - \mu}{2}. \tag{4.28}$$

With the flip time of site l we can now consider the influence received by site x at the moment it flips, shown to be

$$\begin{aligned} I_x(T_x) = \tau_x &= (T_x - T_b) + (T_x - T_l) \\ &= T_x - \mu(n-1) + T_x - (\mu n + \frac{\tau_l - \mu}{2}) \end{aligned}$$

and so utilising $t_x = T_x - \mu(n-1)$ we can rearrange for

$$t_x = \mu + \frac{\tau_x - \mu}{2} + \frac{\tau_l - \mu}{4}. \tag{4.29}$$

Case 4: We can use case 1 above to give us $T_l = \mu(n-1) + \tau_l$, from which it follows that, at the time of site x flipping,

$$\begin{aligned} I_x(T) = \tau_x &= (T_x - T_b) + (T_x - T_l) \\ &= T_x - \mu(n-1) + T_x - (\mu(n-1) + \tau_l) \end{aligned}$$

and we can rearrange to get

$$t_x = \mu + \frac{\tau_x - \mu}{2} + \frac{\tau_l - \mu}{2}. \tag{4.30}$$

Case 6: We again utilise case 1 to give us $T_l = \mu(n-1) + \tau_l$ and $T_r = \mu(n-1) + \tau_r$ which means that when site x flips

$$\begin{aligned} I_x(T_x) = \tau_x &= (T_x - T_b) + (T_x - T_l) + (T_x - T_r) \\ &= t_x + T_x - (\mu(n-1) + \tau_l) + T_x - (\mu(n-1) + \tau_r) \end{aligned}$$

and so

$$t_x = \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{3} + \frac{\tau_r - \mu}{3}. \tag{4.31}$$

Case 7: We can use case 2 to give us $T_l = \mu n + \frac{\tau_l - \mu}{2}$ and $T_r = \mu n + \frac{\tau_r - \mu}{2}$ and thus when site x flips

$$\begin{aligned} I_x(T_x) = \tau_x &= (T_x - T_b) + (T_x - T_l) + (T_x - T_r) \\ &= t_x + T_x - (\mu n + \frac{\tau_l - \mu}{2}) + T_x - (\mu n + \frac{\tau_r - \mu}{2}) \end{aligned}$$

and rearranging gives us

$$t_x = \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{6} + \frac{\tau_r - \mu}{6}. \tag{4.32}$$

Case 8: From previous cases $T_l = \mu n + \frac{\tau_l - \mu}{2}$ and $T_r = \mu n + (\tau_r - \mu)$. When site x flips

$$\begin{aligned} I_x(T_x) = \tau_x &= (T_x - T_b) + (T_x - T_l) + (T_x - T_r) \\ &= t_x + T_x - (\mu n + \frac{\tau_l - \mu}{2}) + T_x - (\mu n + (\tau_r - \mu)) \end{aligned}$$

and lastly rearranging gives us

$$t_x = \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{6} + \frac{\tau_r - \mu}{3}. \tag{4.33}$$

By listing and solving the cases exhaustively we can now define $t(x)$ as

$$t_x(\tau_x, \tau_l, \tau_r) = \begin{cases} \tau_x & \text{if } \tau_x < \min(\tau_l, \tau_r) \\ \mu + \frac{\tau_x - \mu}{2} + \frac{\tau_l - \mu}{4} & \text{if } \mu < \tau_l < \tau_x < \tau_r \\ \mu + \frac{\tau_x - \mu}{2} + \frac{\tau_r - \mu}{4} & \text{if } \mu < \tau_r < \tau_x < \tau_l \\ \mu + \frac{\tau_x - \mu}{2} + \frac{\tau_l - \mu}{2}. & \text{if } \tau_l < \tau_x < \tau_r \text{ and } \tau_l < \mu \\ \mu + \frac{\tau_x - \mu}{2} + \frac{\tau_r - \mu}{2}. & \text{if } \tau_r < \tau_x < \tau_l \text{ and } \tau_r < \mu \\ \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{3} + \frac{\tau_r - \mu}{3} & \text{if } \mu > \max(\tau_l, \tau_r) \text{ and } \tau_x > \max(\tau_l, \tau_r) \\ \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{6} + \frac{\tau_r - \mu}{6} & \text{if } \mu < \min(\tau_l, \tau_r) \text{ and } \tau_x > \max(\tau_l, \tau_r) \\ \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{6} + \frac{\tau_r - \mu}{3} & \text{if } (\tau_r < \mu < \tau_l) \text{ and } \tau_x > \max(\tau_l, \tau_r) \\ \mu + \frac{\tau_x - \mu}{3} + \frac{\tau_l - \mu}{3} + \frac{\tau_r - \mu}{6} & \text{if } (\tau_l < \mu < \tau_r) \text{ and } \tau_x > \max(\tau_l, \tau_r). \end{cases} \tag{4.34}$$

As in the first order evaluation we let b be from a normal distribution with mean 0 and standard deviation σ and let $p = (\tau_x - \mu), y = (\tau_l - \mu)$ and $z = (\tau_r - \mu)$. We now integrate each case in equation (4.34) over the given range to see how much t_x deviates from the fully

fixed zeroth order system as follows:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\min(y,z)} p\frac{e^{\frac{-p^2-y^2-z^2}{2\sigma^2}}}{(2\pi)^{\frac{3}{2}}\sigma^3}\,dp\,dy\,dz = -\frac{\sigma}{2\sqrt{\pi}} \tag{4.35}$$

$$\int_{0}^{\infty}\int_{y}^{\infty}\int_{p}^{\infty} (\frac{p}{2}+\frac{y}{4})\frac{e^{\frac{-p^2-y^2-z^2}{2\sigma^2}}}{(2\pi)^{\frac{3}{2}}\sigma^3}\,dz\,dp\,dy = \frac{\sigma\left(\pi - 4\sqrt{2}\arctan\frac{1}{\sqrt{2}} + 2\sqrt{2}\arctan\sqrt{2}\right)}{32\sqrt{2}\pi^{\frac{3}{2}}} \tag{4.36}$$

$$\int_{-\infty}^{0}\int_{y}^{\infty}\int_{p}^{\infty} (\frac{p}{2}+\frac{y}{2})\frac{e^{\frac{-p^2-y^2-z^2}{2\sigma^2}}}{(2\pi)^{\frac{3}{2}}\sigma^3}\,dz\,dp\,dy = \frac{-\pi\sigma^4 - 2\sqrt{2}\pi\sigma^4 + 2\sqrt{2}\sigma^4\arctan\frac{1}{\sqrt{2}}}{16\sqrt{2}\pi^{\frac{3}{2}}\sigma^3} \tag{4.37}$$

$$\int_{-\infty}^{0}\int_{-\infty}^{0}\int_{\max(y,z)}^{\infty} (\frac{p}{3}+\frac{y}{3}+\frac{z}{3})\frac{e^{\frac{-p^2-y^2-z^2}{2\sigma^2}}}{(2\pi)^{\frac{3}{2}}\sigma^3}\,dp\,dz\,dy = -\frac{\sigma}{6\sqrt{2\pi}} \tag{4.38}$$

$$\int_{0}^{\infty}\int_{0}^{\infty}\int_{\max(y,z)}^{\infty} (\frac{p}{3}+\frac{y}{6}+\frac{z}{6})\frac{e^{\frac{-p^2-y^2-z^2}{2\sigma^2}}}{(2\pi)^{\frac{3}{2}}\sigma^3}\,dp\,dz\,dy = \frac{\sigma\left((2+\sqrt{2})\pi - 4\arctan\sqrt{2}\right)}{48\pi^{\frac{3}{2}}} \tag{4.39}$$

$$\int_{0}^{\infty}\int_{-\infty}^{0}\int_{\max(y,z)}^{\infty} (\frac{p}{3}+\frac{y}{6}+\frac{z}{3})\frac{e^{\frac{-p^2-y^2-z^2}{2\sigma^2}}}{(2\pi)^{\frac{3}{2}}\sigma^3}\,dp\,dz\,dy = \frac{\sigma}{48\sqrt{\pi}}. \tag{4.40}$$

All of these integrals are solvable immediately except the integral from case 7 in which the final integration needs to be done by parts.

In order to calculate the velocity we must first define $f(\sigma)$, the effect of variable thresholds on flipping time, calculated by summing all possible cases in (4.35) as follows

$$\begin{aligned} f(\sigma) = &-\frac{\sigma}{2\sqrt{\pi}} - \frac{\sigma}{6\sqrt{2\pi}} + \frac{\sigma\left(\pi - 4\sqrt{2}\arctan\frac{1}{\sqrt{2}} + 2\sqrt{2}\arctan\sqrt{2}\right)}{16\sqrt{2}\pi^{\frac{3}{2}}} + \\ &\frac{-\pi\sigma^4 - 2\sqrt{2}\pi\sigma^4 + 2\sqrt{2}\sigma^4\arctan\frac{1}{\sqrt{2}}}{8\sqrt{2}\pi^{\frac{3}{2}}\sigma^3} + \frac{\sigma}{24\sqrt{\pi}} + \frac{\sigma\left((2+\sqrt{2})\pi - 4\arctan\sqrt{2}\right)}{48\pi^{\frac{3}{2}}}, \end{aligned} \tag{4.41}$$

and our approximation for the velocity in this system is simply $v = 1/(100 + f(\sigma))$. In order to assess the performance of the approximation we compare it to the realised results from simulation. We simulated a system with all sites having a variable threshold with $\tau \sim \mathcal{N}(100, \sigma)$ for a range of sigma values. A plot of this approximation against simulation results is shown in figure (4.20). As we can see from these results this method of approximation performs

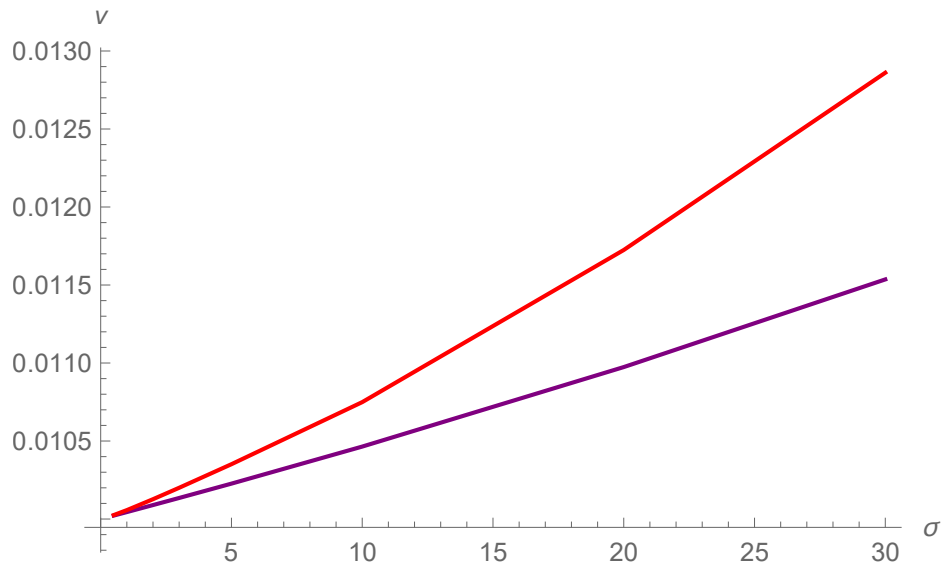


Fig. 4.20 A plot showing the velocity of the wave front in a system with mean threshold 100 and standard deviation σ. The red plot represents the simulation results and purple plot represents the τ_l, τ_x, τ_r variable approximation.

reasonably well for low σ values but we seek a more accurate approximation. We note that as σ gets large the velocity in the simulated systems increases at a much larger rate than the approximation, which is due to the wavefront shape being erratic when the thresholds have large relative differences. As such these approximations are mostly useful in low variance systems.

In this approximation we have made a simple attempt at a mean field approach; assuming some variable behavior among a sea of average behaviour. The next extension of this approach would be to consider sites a, b, c, l, r, x as variable, with all other sites having mean threshold, allowing us to include the accelerating impact of more variable thresholds in the local neighbourhood.

4.4 Tiled Approach to the Variable Threshold System.

In this final section we look at using the previous tiled approach in order to approximate the velocity of a wavefront in a system for which the sites have variable thresholds. The predominant advantage of approximating the system as being made up of a given tile repeated many times over is that it is possible to calculate an exact result while having the option to increase the tile size to improve accuracy. We begin by looking at the system as a repetition

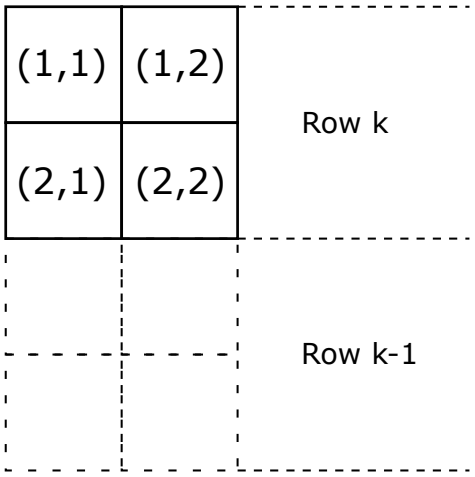


Fig. 4.21 An illustration of the tiled system with the labeling of the 4 sites with variable thresholds that are repeated through the whole system.

of a 2×2 tile as illustrated in figure (4.21) where the row number, k, represents the positional height of the tile in the system and the wave is traveling vertically upwards.

In this system we know the wavefront will have stable velocity v and we define $\gamma = \frac{2}{v}$ to be the time taken for the wave to travel through a tile. We define $T_{k,i,j}$, the flip time of site (i, j) in tile row k with expression

$$T_{k,i,j} = \gamma k + x_{i,j} \tag{4.42}$$

where $x_{i,j}$ is the additional time it takes site (i, j) to flip once the first site in a tile has flipped. We again utilise the J function defined in (4.5) to denote the influence received between sites based upon their flip time differences. We assume the thresholds are from distributions that are close to one another and with sufficiently low variances such that the sites in any given column all flip in ascending order and the wavefront behaviour is predictable. With this reasonable restraint we can generate equations for the thresholds in terms of the influence each site receives from its neighbours as follows

$$\begin{aligned}
\tau_{2,1} &= x_{2,1} - x_{1,1} + \gamma + 2J(x_{2,1} - x_{2,2}) \\
\tau_{2,2} &= x_{2,2} - x_{1,2} + \gamma + 2J(x_{2,2} - x_{2,1}) \\
\tau_{1,1} &= x_{1,1} - x_{2,1} + 2J(x_{1,1} - x_{1,2}) \\
\tau_{1,2} &= x_{1,2} - x_{2,2} + 2J(x_{1,2} - x_{1,1}).
\end{aligned} \tag{4.43}$$

Given the reflective symmetry of the tile structure and repetition in the structure of the system there are 4 tile cases we are concerned with,

1. $x_{1,2} < x_{1,1}$ and $x_{2,2} < x_{2,1}$ if $\tau_{1,2} < \tau_{1,1} + \frac{1}{3}(\tau_{2,1} - \tau_{2,2})$ and $\tau_{2,2} < \tau_{2,1}$

2. $x_{1,1} < x_{1,2}$ and $x_{2,1} < x_{2,2}$ if $\tau_{1,1} < \tau_{1,2} + \frac{1}{3}(\tau_{2,2} - \tau_{2,1})$ and $\tau_{2,1} < \tau_{2,2}$

3. $x_{1,2} > x_{1,1}$ and $x_{2,2} < x_{2,1}$ if $\tau_{1,2} > \tau_{1,1} + \frac{1}{3}(\tau_{2,1} - \tau_{2,2})$ and $\tau_{2,2} < \tau_{2,1}$

4. $x_{1,1} > x_{1,2}$ and $x_{2,1} < x_{2,2}$ if $\tau_{1,2} > \tau_{1,1} + \frac{1}{3}(\tau_{2,2} - \tau_{2,1})$ and $\tau_{2,1} < \tau_{2,2}$,

where the first 2 and last 2 cases have the same solution. The first 2 cases represents a tile in which the sites in one column always flip before the other column whereas the second case represents the first site to flip on each line coming from alternating columns. We begin by getting an expression for γ in each unique case.

For case 1

$$\begin{aligned}
\tau_{2,1} &= (x_{2,1} - x_{1,1}) + \gamma + 2(x_{2,1} - x_{2,2}) \\
\tau_{2,2} &= (x_{2,2} - x_{1,2}) + \gamma \\
\tau_{1,2} &= (x_{1,2} - x_{2,2}) \\
\tau_{1,1} &= (x_{1,1} - x_{2,1}) + 2(x_{1,1} - x_{1,2}) \\
x_{2,2} &= 0,
\end{aligned} \tag{4.44}$$

which we can solve for γ to get

$$\gamma_1 = \tau_{2,2} + \tau_{1,2} \tag{4.45}$$

which is intuitive since the time taken to travel through a tile in which one column always flips first should simply be the sum of said column's thresholds.

For case 3

$$\begin{aligned}
\tau_{2,1} &= (x_{2,1} - x_{1,1}) + \gamma + 2(x_{2,1} - x_{2,2}) \\
\tau_{2,2} &= (x_{2,2} - x_{1,2}) + \gamma \\
\tau_{1,2} &= (x_{1,2} - x_{2,2}) + 2(x_{1,2} - x_{1,1}) \\
\tau_{1,1} &= (x_{1,1} - x_{2,1}) \\
x_{2,2} &= 0,
\end{aligned} \tag{4.46}$$

which we again solve for γ to get

$$\gamma_2 = \frac{1}{4}(3\tau_{1,1} + 3\tau_{2,2} + \tau_{1,2} + \tau_{2,1}). \tag{4.47}$$

For an example approximation we will let all thresholds be from a Gaussian normal distribution with mean μ and standard deviation σ. We can then express each threshold as $\tau_{i,j} = \mu + u_{i,j}$ and we see that

$$\begin{aligned}\gamma_1 &= 2\mu + u_{2,2} + u_{1,2} \\ \gamma_2 &= 2\mu + \frac{1}{4}(3u_{1,1} + 3u_{2,2} + u_{1,2} + u_{2,1}).\end{aligned} \tag{4.48}$$

In order to calculate γ and the velocity of the wave in a given case we need to perform the integration

$$\int_D (2\mu + f(u_{1,1}, u_{1,2}, u_{2,1}, u_{2,2})) \frac{e^{-\frac{u_{1,1}^2+u_{1,2}^2+u_{2,1}^2+u_{2,2}^2}{2\sigma^2}}}{(2\pi)^2\sigma^4} du_{1,1}\, du_{1,2}\, du_{2,1}\, du_{2,2}, \tag{4.49}$$

where D is the domain over which the integration is performed and f is some linear combination of u terms.

In order to calculate the integral (4.49) we use the change of variable

$$z_{i,j} = \frac{u_{i,j}}{\sigma} \text{ with } du_{i,j} = \sigma\, dz_{i,j} \tag{4.50}$$

so that the original integration in equation (4.49) becomes

$$\int_{D'} (2\mu + \sigma f(z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2})) \frac{e^{-\frac{z_{1,1}^2+z_{1,2}^2+z_{2,1}^2+z_{2,2}^2}{2}}}{(2\pi)^2} dz_{1,1}\, dz_{1,2}\, dz_{2,1}\, dz_{2,2} \tag{4.51}$$

with D' being the new domain to integrate over.

This integration can now be separated into two parts. The first part is

$$\int_{D'} 2\mu \frac{e^{-\frac{z_{1,1}^2+z_{1,2}^2+z_{2,1}^2+z_{2,2}^2}{2}}}{(2\pi)^2} dz_{1,1}\, dz_{1,2}\, dz_{2,1}\, dz_{2,2}, \tag{4.52}$$

which when we sum over the whole domain by adding all 4 cases together, simply sums to 2μ. The second part

$$\sigma \int f(z_{1,1}, z_{1,2}, z_{2,1}, z_{2,2})) \frac{e^{-\frac{z_{1,1}^2+z_{1,2}^2+z_{2,1}^2+z_{2,2}^2}{2}}}{(2\pi)^2} dz_{1,1}\, dz_{1,2}\, dz_{2,1}\, dz_{2,2}, \tag{4.53}$$

is a linear term in σ where the integral is a constant K which we can solve for numerically in each case. We can numerically evaluate the integral for all cases with the domains defined by the τ inequalities in equation (4.4) and the function f from the equations in (4.48). We let K_1 be the solution for cases 1 and 2 and K_2 be the solution in cases 3 and 4 and through numerical integration we get the results

$$K_1 = -0.319459 \tag{4.54}$$

$$K_2 = -0.115127. \tag{4.55}$$

By combining the results for equations (4.52) and (4.53) we can get an expression for γ which is linear in μ and σ as follows

$$\begin{aligned} \gamma &= 2\mu + 2\sigma(K_1 + K_2) \\ &= 2\mu - 0.869172\sigma. \end{aligned} \tag{4.56}$$

We plot this approximation against realised simulation results in graph (4.22) where we see that for such a simple and small tile the approximation is reasonable. We conclude that, of all the approximation methods presented in this chapter, that this tiled approximation is the superior approach. It is simple to scale up by increasing the tile size, either to 3×3, or even just 3×2 and the number of cases required to go through to reach a solution is significantly fewer than the alternative approaches. The following work for this content in preparation for a publication would be to extend the number of variable locations in each approximation to see the effects of early flipping sites over a larger distance. An area of further investigation is also the shape of the wavefront within a tiled system as surface roughening is a key factor in wave acceleration.

4.5 Interface Spreading Conclusion

By utilising a memory and threshold model on a two dimensional lattice with nearest neighbour interactions we created multiple approximation methods to calculate the velocity

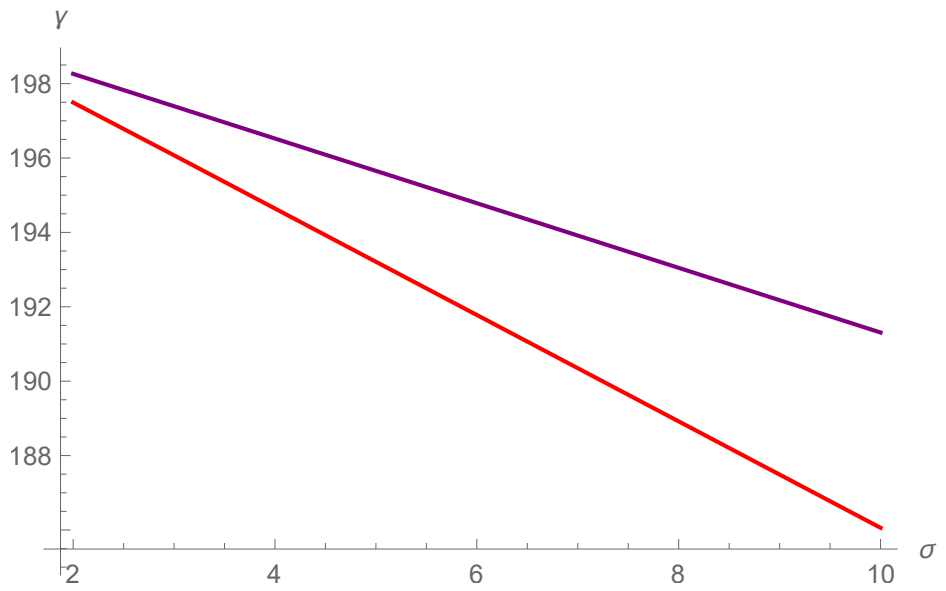


Fig. 4.22 A line plot for the γ of a system against σ for both a simulated system and the tile approximation. The red line represents the simulation result and the purple line represents the tile approximation.

of the wave. We explored the basic behaviour of the traveling wave through simple site threshold setups noting that velocity is not proportional to the mean system threshold and is instead more closely linked to the fastest possible route. We highlighted how some low threshold sites can accelerate the wave by increasing the surface area through which influence is spread and how this relates to the gradient of the wavefront. We have shown that the behaviour of a large system can be accurately approximated using repetition of a small tile where the exact velocity of the wave is much simpler to compute while demonstrating how very small tiles can exhibit high variance in this method.

We then proceeded to look at systems with variable memories or thresholds and utilising the time differences between flips of neighbouring sites to approximate the wave velocity. In doing so, we established the non-normality of the distribution from which the flip time differentials originate. We used mean field approximations of progressively higher order to estimate the velocity of the wavefront, acknowledging the results were promising but higher order systems, in which more sites are considered to have variable thresholds, may be required to reach a more accurate result. Finally we combined the previous tiled system analysis with the variable threshold system and calculated analytical approximations for the wave travel time based upon the time to travel through one tile in a self replicated system.

Chapter 5

Threshold Spreading in One Dimension

The investigation of how ideas spread through communication networks is a heavily researched area due to the impact of social networks on modern society. The understanding of how people can be influenced and who or where to target efforts [84] offers huge potential benefits to those wishing to effectively spread ideas; ranging from advertisers being able to get their products positively seen by more people [85, 86] to the spread of political ideas [87, 88] and 'fake news' [89]. While this field of research often involves complex networks in this chapter we will look at the spread of an invasive idea in much simpler systems to investigate their behaviour with the possibility of the results and techniques being able to upscale to different systems. We focus on one dimensional systems in which agents communicate with others within a given range with the main consideration being how far an invasive idea is expected to spread and what connectivity features can stop it. Spreading effects on networks can have many varied contagion methods [90]. Our system is a susceptible infectious (SI) model [91, 92], meaning all sites are in one of two states, either susceptible or infectious. The similar but distinctly different alternative models would be the susceptible, infectious, susceptible (SIS) model [93] in which infected sites can revert back to susceptible or the susceptible, infectious, recovered (SIR) model [94] where infected sites can recover or be refractory and have an immunity. We choose the SI model as it exhibits the features we are concerned with and lacks unnecessary complexity.

We start by looking at a simple toy model to introduce the setup, notation and basic concepts and move on to calculate the range of site connectivity, firstly using a martingale approach and secondly a clustered system approach. We conclude by considering the system to be analogous to a mean-field random walk and investigating the distribution of connections which allow the invading wave to propagate indefinitely.

5.1 One-Dimensional Fixed Neighbour Threshold Model

The first model we will look at is very simple but it has some important features and requires certain techniques to solve that will be useful in more complex situations. We take a line with agents at sites $i, i \in \mathbb{Z}_{\geq 0}$ and define a connection between two sites to be a reciprocal communication in which they can pass on influence of an idea. We define a site i as either being in the 'up' opinion state, $S_i = 1$, denoting that they have adopted the new invasive idea or in the 'down' opinion state $S_i = 0$ in which they follow the original standard idea. At the beginning of the system all sites are in the 'down' opinion state except a small cluster at the start of the line that have adopted the new invasive idea, and all sites have a threshold, τ, a measure of their resistance to adopt the new idea. Once a down site interacts with τ up sites they will flip their opinion state and begin to spread influence of the new idea themselves. The key distinction between the systems in this chapter and the previous chapter is that these models require a site to communicate to τ unique sites in order to flip, whereas in the previous chapter a site could flip after enough time of communicating to just one neighbour with the invasive idea. There is a system parameter r, the link range, which indicates the number of sites in the neighbourhood of a site that it can and does communicate with. The application of threshold rules as a simplistic comparison to social influence is an established area of research with many models having been created including the Watts model [17] and the Centola-Macy model [95]. Our system is highly comparable to these models in which sites have a fixed and permanent threshold taken from some distribution and each node can be in a binary state at a given time. Sites then update their state based upon the number of connected neighbours exceeding or remaining below the threshold. The key distinction of this work is introducing new methods to approach the model which can then be applied to other systems.

In this first model each site interacts with a cluster of r sites in its neighbourhood with the possibility of looking slightly forwards or backwards. This system allows us to enumerate the connectivity of a site depending on which of the sites in its neighbourhood it is connected to. We define the connection state s_i of site i as the difference between the number of forward connections and the number of backward connections. Examples of connection states are shown in figure (5.1). Note the distinction between the opinion state of i, S_i and the connection state of i, s_i. Given that the links between sites in this system are reciprocal, and the connection blocks are dense, it means that the connectivity state of site i has a direct impact on the connectivity states of all the sites within range r and in fact further as we will show.

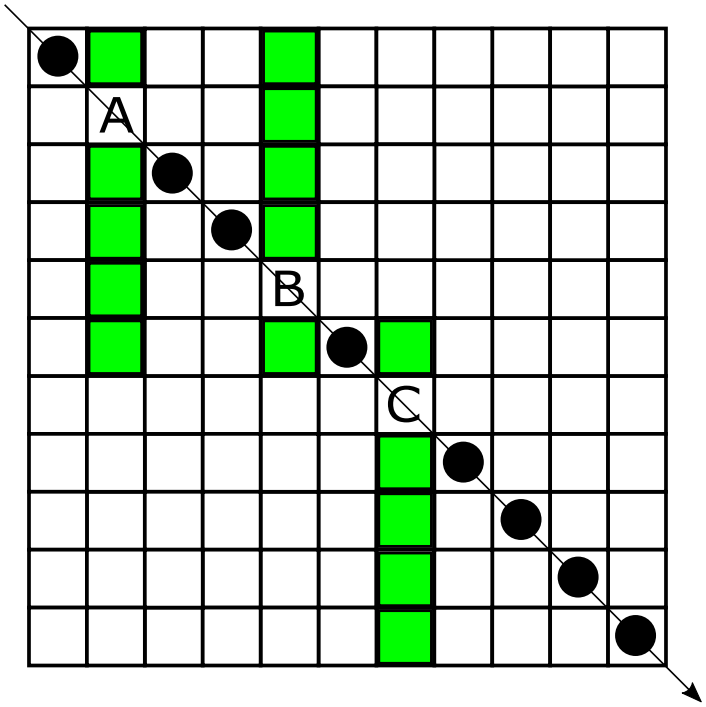


Fig. 5.1 An example of the restricted connectivity of the initial model. The leading diagonal represents sites on the one dimensional line and green squares represents reciprocal connections. Only 3 sites in the system are shown as an example. In this case $r = 5, s_A = 3, s_B = -3$, and $s_C = 3$.

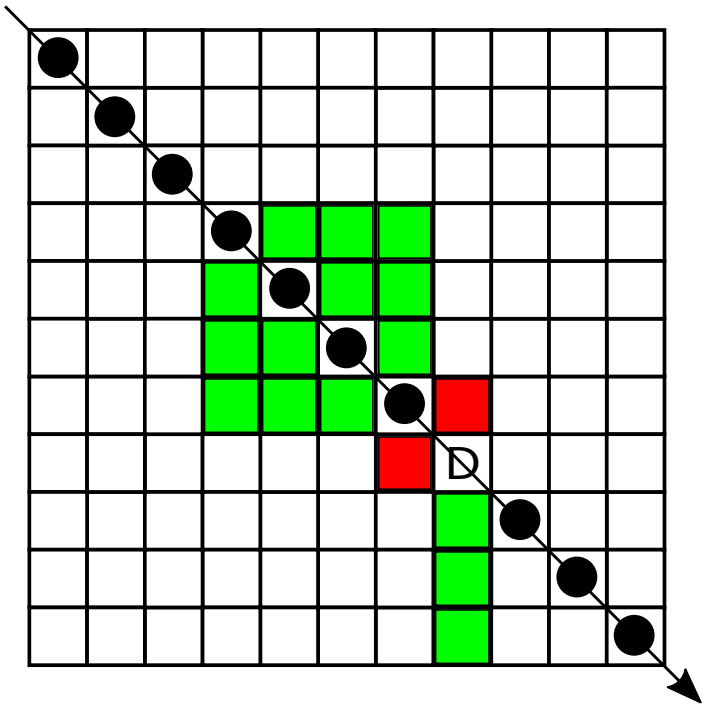


Fig. 5.2 A system in which $r = 3$ and $s_{D-4} = r = 3$. The red square highlights a site which can't represent a connection due to this, forcing the connection states $S_D = 3$.

Evidently for a system with a link range r there are $r+1$ possible connection states for a site to be in, ranging from only communicating forwards to only communicating backwards. We begin by considering a site i in the system in connection state $s_i = r$, that is that they communicate only with agents in front of them. Due to the reciprocal nature of links we know that site $(i+r)$ must communicate r sites backwards and as such their only possible connection state is $s_{i+r} = -r$. It follows that if $s_{i+r} = -r$ then site $(i+r+1)$ can not be communicating to any backwards sites and thus $s_{i+r+1} = r$. This idea is illustrated in figure (5.2). From this position it is clear that the system is now in a series of disjoint clusters and any invading ideas will never be able to exceed the threshold at site $(i+r+1)$ giving us our first rule in this model; $(i+r)$ is an upper bound for the travel distance of the wave where site i is the first site in the system in connection state r.

In the case where $r = 3$ therefore we have confirmed that if we ever see a site i for which $s_i = 3$ then the future nature of the system is fully decided, likewise if $s_i = -3$. Given that the only two other states possible are 1 and -1 we shall consider a system limited to only those two options. For simplicity let $s_i = 1$ and in only considering the agents in front, we know site i has reciprocal links to sites $(i+1)$ and $(i+2)$. We now consider the state at site $(i+2)$; we have already determined that they must connect to site i and since in this case we

have disallowed the extreme connection states the only option becomes $s_{i+2} = -1$. If now we next consider site $(i+1)$ we have two options and both are valid, however whichever state $(i+1)$ is in must be the opposite of state $(i+3)$. If $s_{i+1} = 1$ then, using the same justification we used to explain $s_i = -s_{i+2}$, we can say $s_{i+1} = -s_{i+3}$. If $s_{i+1} = -1$ then it does not have a link to site $(i+3)$, but $(i+3)$ does link to $(i+2)$ and thus $s_{i+3} = 1$. For the system with extreme cases disallowed it is true that $s_i = -s_{i+2}$ for all $i \in \mathbb{Z}_{\geq 0}$.

This leads to two more rules. Firstly if we draw out the system in the way shown in figure (5.1) and figure (5.2) then the system has an intuitive line of reflection across the leading diagonal due to the reciprocal nature of the links. This graphically demonstrates the fact that $\sum_{k=i}^{\infty} s_k = 0$ as every connection forward for one site is a connection backwards for another. In fact we can expand on this rule and assert $\sum_{k=i}^{i+r+1} s_k = 0$ showing that for all possible systems $s_i = s_{i+r+1}$ and the system will always enter cycles of length $r+1$. Secondly because the connection blocks are reciprocal and dense it is not possible for site $(i+1)$ to communicate further back than site i communicates. Thus in order for sites to exceed their threshold and change opinion to the new invading idea they must always do so in order. Due to this fact, and because in order to flip ideas you must interact with τ sites with the new opinion, every single site must interact with at least τ agents from behind or the invading wave breaks down. Combining these rules gives us a very strict pattern of acceptable cycles that can allow a wave to continue indefinitely.

If we take for instance $r = 4$ and $\tau = 2$ we know from the above rules that $\sum_{k=i}^{i+r+1} s_k = 0$, the states are in a cycle of 4 and that all sites must link to at least 2 sites from behind. Thus the only viable configuration for this wave to continue unrestricted is $(0,0,0,0,\ldots)$. While this model is rather trivial and solvable it introduces some important concepts and acts as the prelude to much more interesting models.

5.2 One-Dimensional Ranged Interaction Model

The greatest restriction of the previous model is that each site interacts with a fixed cluster of neighbours and because of this the states of ahead agents are predetermined and the system is easily understood. We now expand on that system by allowing sites the possibility of connecting or not connecting to a site within a fixed radius around them but keeping the property of a site threshold. In this system the links are still reciprocated in the sense that if site i passes influence to site j then site j passes influence to site i, but because sites don't have the fixed cluster of connections there isn't a forced cycle of site states. This difference in the system leads to some fundamental behaviour changes for the traveling wave, most

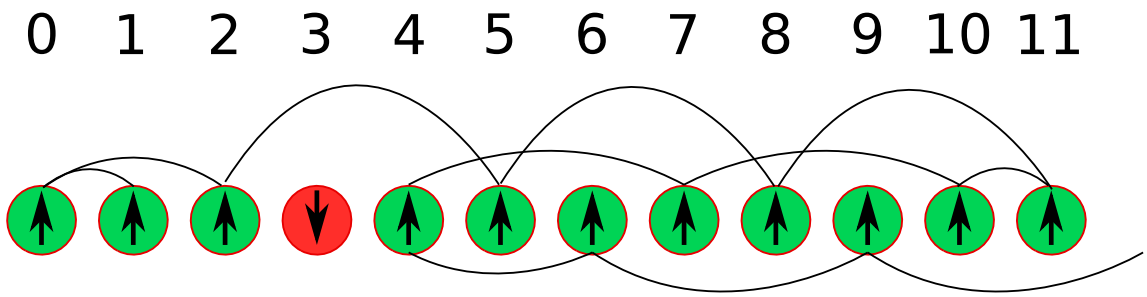


Fig. 5.3 A sample system with link range $r = 3$ and threshold $\tau = 1$. This example exhibits two important features of this model; firstly that the wave is still traveling despite site 3 not having flipped and secondly that the wave has traveled forward, backwards and then forwards again. Since sites $0, 1$ and 2 start in the up state the order of sites flipping is $5, 8, 11, 10, 7, 4, 6$ and lastly 9.

notably that for site i to take on the new invasive idea it is not required for site $(i-1)$ to have already switched. In fact it is entirely possible for a site to never switch opinions but for the wave to travel past them or for the wave to travel past but come back and influence them at a later time step. It is because of these traits that the analysis becomes more complicated.

We begin by looking at a system with link range $r = 3$, threshold $\tau = 1$ and probability of making a link, p, the simplest system that exhibits some of these behaviours. We begin all systems with the first r sites having the new invasive idea and all other sites having the standard weaker idea, this way the invasive opinion begins at its strongest since it has the most possibilities to pass on. The question we attempt to answer is how far will the wave travel before no longer influencing any new sites to flip? As shown in figure (5.3) the wave can travel forwards, then influence sites behind and then travel forwards again and, although the probability of these long sequences occurring becomes very low, they are possible. Therefore in calculating the expectation that an arbitrary site will eventually adopt the new invasive idea you have to look at the connections in both directions for a large distance. This makes any approach involving calculating the expected opinion of sites incredibly complicated.

For these reasons we need to use a new approach beginning with adjusting the enumeration of the site states. Due to the reciprocal nature of links between sites if we start at site 0 they can only connect to the r sites in front of them with 2^r permutations. If we then consider site 1 it is already determined whether it connects to site 0 and as such we need again only consider the r sites in front. We can repeat this process and notice that all configurations of network connections can be created by only determining the forward connections of each site. In order to enumerate this we simply allow a sites state to be the binary expansion of its

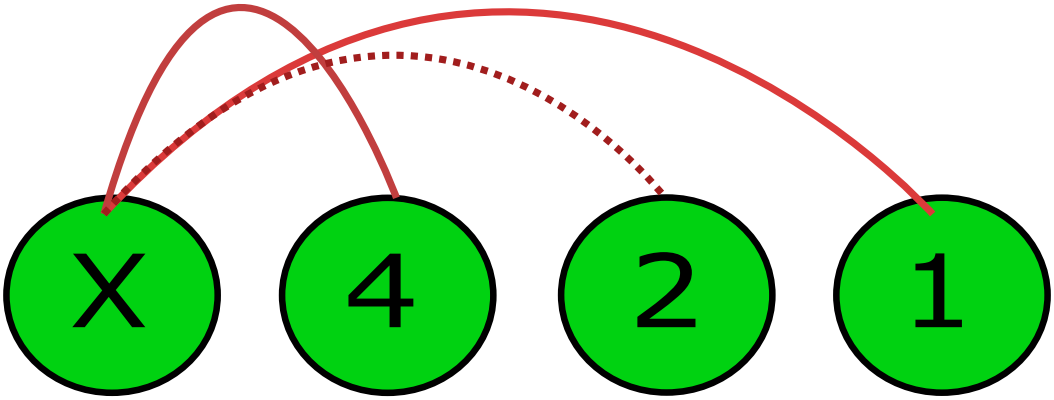


Fig. 5.4 An example of the forward connection enumeration on a system with link range r. In this example the dotted line represents no link and as such node X has link state 5.

forward connections as shown in figure (5.4). As such all possible network configurations can be expressed as a series of integers between 0 and $2^r - 1$ and as an example if we look at figure (5.3) we can say $(s_0, s_1, \ldots, s_{10}, s_{11}) = (6,0,1,0,3,1,1,1,1,1,4,0)$.

With this structure in place, instead of looking directly at how far a wave travels, we can consider which sequences of network connections (site states) cause the invading wave to get pinned.

5.2.1 Stopping Sequences

We begin by looking at the case where $p = 0.5$ and as such all possible states for a single nodes connections are equally likely. At time of creation our system has $S_0 = S_1 = S_2 = 1$ and $S_k = 0$ for all $k > 2$. As the system evolves over time the wave propagates through the system and some of the down sites will become up, with the opposite switch being impossible. Consider a site in the system $i, i \neq 0$ for which sites $S_i = S_{i+1} = S_{i+2} = 1$. In our consideration of sequences that cause the wave to stop we can truncate off all of the sites before site i since the system is essentially re-originated at this point. A second consideration in this model is that for the wave to stop there must be a furthest forward up site j such that $S_j = 1$ and $S_k = 0$ for all $k > j$. The state of all sites above j become irrelevant to the wave stopping since the invading opinions never reaches them or goes beyond. This gives us our definition of a stopping sequence.

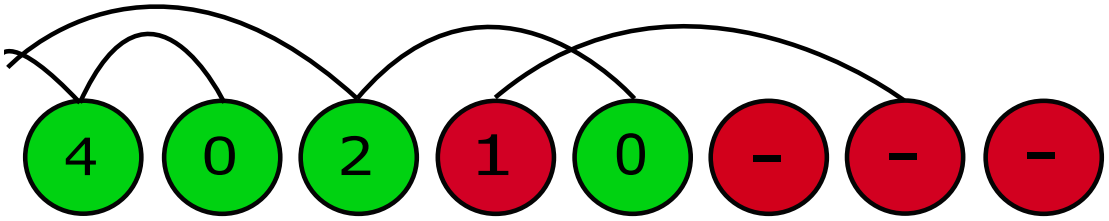


Fig. 5.5 Stopping sequence (4, 0, 2, 1, 0). Whenever this sequence of forward links occurs the wave is guaranteed to stop.

Definition. *Stopping Sequence : A stopping sequence for a system is the sequence of site states* $(s_i, s_{i+i}, \ldots, s_{i+r}, \ldots, s_{j-1}, s_j)$ *in which site i is the greatest site in the system for which* $S_i = S_{i+1} = \ldots = S_{i+r} = 1$ *and* $S_j = 1$, $S_k = 0$ *for all* $k > j$.

In our system with $r = 3, \tau = 1$ and $p = 0.5$ the simplest stopping sequence to consider is (0,0,0), three sites in a row that make no forward links and as such the wave can never reach past them. For this system there happens to be 8 stopping sequences of length 3, 72 sequences of length 5 and an exponentially growing number of longer sequences. Shorter sequences are universally more likely to occur and cause a wave to stop than longer sequences, however not all stopping sequences of the same length are equally probable even though all site states have equal probability of $1/8$. An example of this are the two stopping sequences shown in figures (5.5) and (5.6); in which both sequences contain the same number of links yet the sequence (4 1 0 2 0) is almost twice as likely as (4 0 2 1 0) to be the sequence to stop a wave, with respective probabilities of 0.01871 and 0.0096.

The reason for there being no stopping sequence of size four is due to the definition of a stopping sequence. The final site j of a stopping sequence has to be in the up state, that is $S_j = 1$, and therefore $s_j = 0$. We then consider site $j-1$; if site $j-1$ does not connect to sites $j+1$ or $j+2$ then it doesn't matter which value S_{j-1} takes. If site $j-1$ does connect past site j then $S_{j-1} = 0$ which puts restrictions on the connection states s_{j-4}, s_{j-3} and s_{j-2}. Likewise if site $j-2$ does not connect to site $j+1$ then S_{j-2} can take either value and if site $j-2$ does connect to site $j+1$ then $S_{j-2} = 0$ and the connection states s_{j-5}, s_{j-4} and s_{j-3} matter. In short the stopping sequences of length three occur when sites $j-1$ and $j-2$ make no connections beyond site j. In systems where sites $j-1$ and $j-2$ do make connections past site j then the stopping sequence must go as far back as at least site $j-4$, making the sequence have length of at least five.

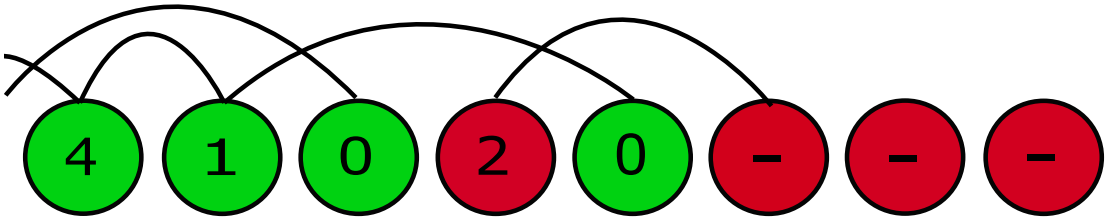


Fig. 5.6 Stopping sequence (4, 1, 0, 2, 0). Whenever this sequence of forward links occurs the wave is guaranteed to stop.

Table 5.1 Probability that three given sequences are first finished at each time step given no current flips.

Time step	1	2	3	4	5	6
Current flip sequence	-	-	-	-	-	-
Prob. of HTHT first ending	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{3}{64}$
Prob. of THTT first ending	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
Prob. of TTTT first ending	0	0	0	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$

Our explanation for the difference in probability of occurrence as a stopping sequence, for different sequences of the same length, comes from Penney's game [96], named after its inventor Walter Penney, in which a fair coin is tossed and you gamble on the occurrence of sequences. The game is non-transitive because for three sequences you often find that betting on sequence A occurring before sequence B is profitable, betting on sequence B occurring before sequence C is profitable and betting on sequence C occurring before sequence A is profitable. This non-transitivity has some interesting results that follow over to our work.

5.2.2 Penney's Game

In order to understand the difference in probabilities for sequences of the same length we begin by considering sequences of coin flips in which there are only two possible results at each site rather than the eight of the network model. Before any coin is flipped it is true that for all sequences there is a $1/16$ chance of the sequence finishing at the fourth coin flip. However as the coin flip results start coming in the probability of sequences being completed at given time steps can vary wildly as shown in tables (5.1), (5.2), and (5.3).

Table 5.2 Probability that three given sequences are first finished at each time step given heads was the result on the first flip.

Time step	1	2	3	4	5	6
Current flip sequence	H	-	-	-	-	-
Prob. of HTHT first ending	0	0	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
Prob. of THTT first ending	0	0	0	0	$\frac{1}{16}$	$\frac{1}{16}$
Prob. of TTTT first ending	0	0	0	0	$\frac{1}{16}$	$\frac{1}{32}$

Table 5.3 Probability that three given sequence are first finished at each time step given the first coin flip was a head and the second was a tails

Time step	1	2	3	4	5	6	7
Current flip sequence	H	T	-	-	-	-	-
Prob. of HTHT first ending	0	0	0	$\frac{1}{4}$	0	0	$\frac{1}{16}$
Prob. of THTT first ending	0	0	0	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
Prob. of TTTT first ending	0	0	0	0	$\frac{1}{8}$	0	$\frac{1}{32}$

From these examples it is evident that the expected wait time for a given sequence is not consistent for each sequence and fluctuates as the results come in. In order to calculate the expected wait time until we encounter a given sequence we introduce a thought experiment. Imagine a fair gambling game, run by a casino in which we repeatedly toss a coin until we hit a required sequence at which point the game ends. At every coin toss a new gambler arrives and bets £1 on the first result of a sequence; if he loses he leaves disappointed and if he wins he remains and bets everything on the next result. For instance if the sequence we are considering is HTTT every gambler will bet on H for their first bet and if they win they will bet on T for the second flip etc. figure (5.7) is a graphic representation for the total winnings for all gamblers that are still involved when the game ends for the sequences HTHT and THTT.

As we can see in the case of HTHT the gamblers have a combined final winnings the gamblers have a combined winning of £20 and in the case of THTT the gamblers have final winnings of £18, remembering that since this is the first time the sequence has occurred any gamblers that played before this point must have lost. Given that the game is fair, by definition the expected net gain to the casino must be 0. Since the gamblers bet a new £1 at every time step the expected number of flips to encounter THTT must be 18 and the expected number of flips to encounter HTHT must be 20. However it can be shown that the

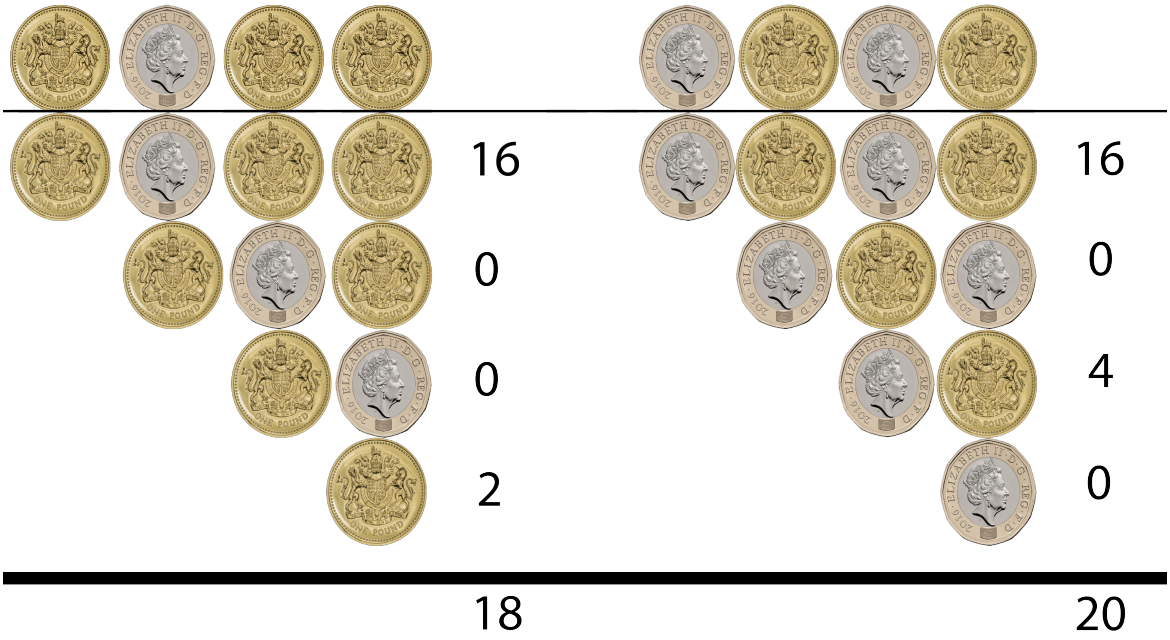


Fig. 5.7 Illustration of the gamblers fortunes when the sequences THTT and HTHT first occur and the game ends. Each row represents a gambler and the bets they make if they are still in the game, while the number at the end of each row is the gamblers fortune when the game finishes.

odds are 9 to 5 in favour of seeing HTHT before THTT. In the original version of Penney's game we have two players A and B competing to guess a sequence that will occur first. Player A chooses a sequence to bet on, reveals it and then player B chooses their sequence. Provided the sequences are of length three or greater it is always possible for player B to pick a sequence that gives them favourable odds of winning with this being shown in figure (5.8)

An intuitive explanation for these results is that unless player A wins at the earliest possible chance, on the third coin flip, then at any point in time at which they could win on the next coin flip there is a high chance that player B has just won. This is because the player B intentionally chooses a sequence that ends with the first two flips in the sequence chosen by A. This is the most apparent in the case where player A chooses TTT (or HHH); if their sequence is completed on the third coin flip then they immediately win and their choice was unbeatable. If their sequence is to complete at any time, $t > 3$, then it must be true that the sequence HTT (or THH) was completed at time $t-1$. Thus if player A chooses TTT they have a $1/8$ chance of still winning against HTT by winning on flip three, but after that they can never win. It is this behaviour that contributes to the HTHT sequence being a 9 to 5 favourite to be seen before THTT despite having a longer expected wait time. It is also worth noting that this game is non-transitive in the sense that if strategy A beats strategy B and strategy B beats strategy C then strategy A doesn't necessarily beat strategy C, meaning that, unless you have prior information to your opponents choice, there isn't an optimal choice that guarantees you are favourite to win.

In order to calculate the expected wait time to encounter any sequence from a list and also calculate the probability for any given sequence to be the first to occur we refer to a paper by Shuo-Yen Robert Li [97] in which a method is outlined to exactly calculate such values. We let Z be a discrete random variable (the theory works for an arbitrary discrete distribution so one may pick their favourite), Σ be the set of all possible values of Z and let $Z_1, Z_2, \ldots$ be a sequence of independent random variables with the same distribution as Z. Let $A = (a_1, a_2, \ldots, a_m)$ be a sequence over Σ and $B = (b_1, b_2, \ldots, b_n$ be a sequence over Σ that is not a connected sub-sequence of $(a_1, a_2, \ldots, a_{m-1})$. We denote N_B to be the waiting time until the sequence B occurs in a run of the process $Z_1, Z_2, \ldots$. We previously calculated N_B in the coin system for sequence $B = (H, T, H, T)$ given a new system with no current gamblers. Now imagine we arrive to the game late and we already see the sequence $A = (H, T)$ is on the table; this changes the expected time to now see sequence B. We define $\mathbb{E}[N_{AB}]$ to be the expected wait time to see sequence B given the sequence A as a starting point. Clearly in this case when we arrive at the table the gambler that began playing two steps before currently has a fortune of £4 and all other gamblers have lost. Given we know the gamblers have

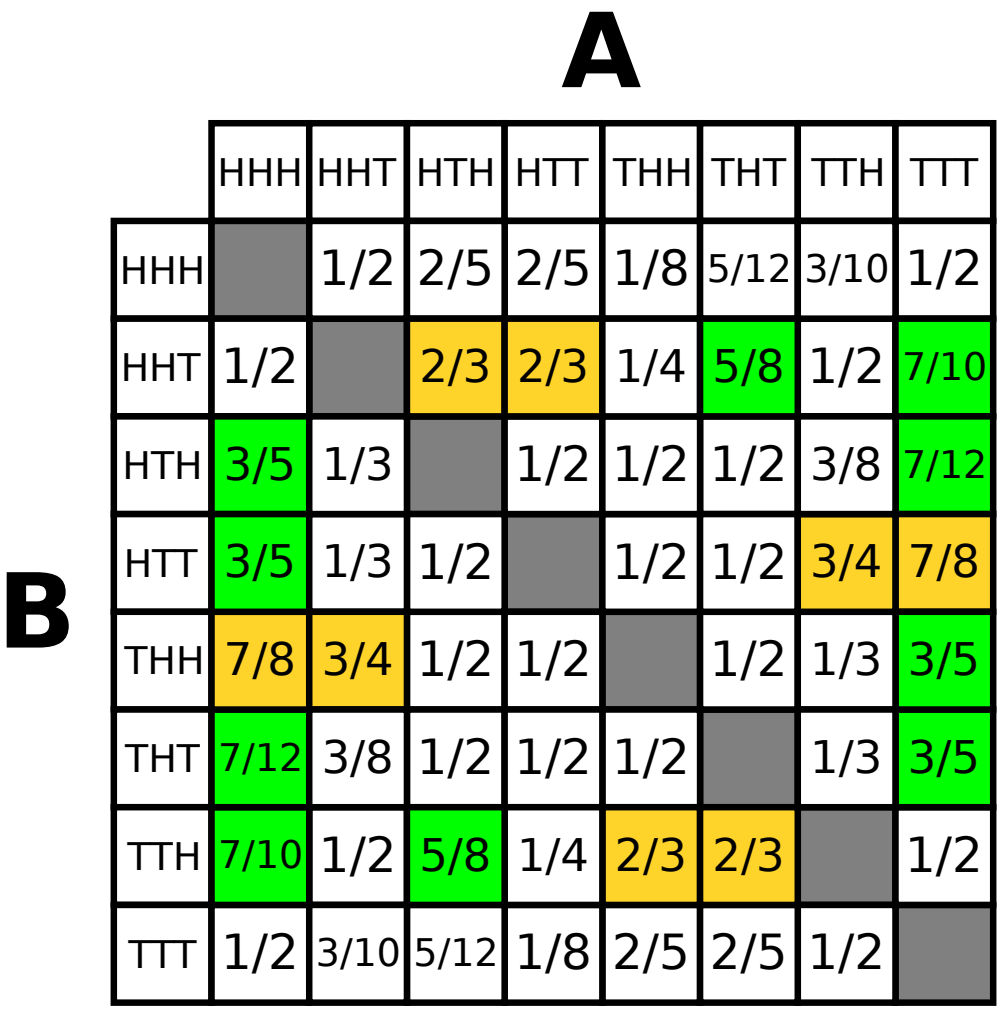


Fig. 5.8 Table showing the probability of B winning given the players choose the given sequences. Cells highlighted in gold correspond to the optimal choice for B, given the choice made by A, and cells highlighted green are alternate winning selections. These numbers are derived from Penney's game.

a fortune of £20 at the game's conclusion we expect them to gain another £16. Since the game is fair, by definition, if we expect the gamblers to win £16 more then we also expect them to gamble £16 more and since £1 is added at each time step we can deduce that for $B = (H,T,H,T)$ and $A = (H,T)$ we have $\mathbb{E}[N_{AB}] = 16$.

Let $A = (a_1, a_2, \ldots, a_m)$ and $B = (b_1, b_2, \ldots, b_n$ be two sequences over Σ. For every pair of integers (i, j), let

$$\delta_{i,j} = \begin{cases} P(Z = b_j)^{-1} & \text{if } 1 \leq i \leq m, 1 \leq j \leq n, \text{ and } a_i = b_j \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

We then define

$$A * B = \sum_{k=1}^{m} \prod_{j=0}^{m-k} \delta_{k+j,j+1} = \delta_{1,1}\delta_{2,2}\cdots\delta_{m,m} + \delta_{2,1}\delta_{3,2}\cdots\delta_{m,m-1} + \ldots + \delta_{m,1} \tag{5.2}$$

It follows from our previous example that $\mathbb{E}[N_B] = B * B$, which gives us the expected wait time for sequence B to occur given no starting sequence, and $\mathbb{E}[N_{AB}] = B * B - A * B$ which gives us the expected wait time for sequence B given starting sequence A with B not being a connected sub-sequence of B. From this we can re-derive the main result from Li's paper [97] which we will utilise in our network problem.

Let $A_1, A_2, \ldots, A_N$ be sequences over Σ. For each sequence A_i we want to calculate the probability that A_i is realised in the sequence $Z_1, Z_2, \ldots$ before the other $n-1$ sequences. As before we consider the situation where a sequence A is given to begin the process and write N_i for N_{AA}. We then let N be the minimum stopping time among all the stopping times $N_1, N_2, \ldots, N_n$. We want to compute $P(N_j = N)$ for each j. To avoid the case of ties we assume that no sequence contains another as a sub-sequence. Given the starting sequence A, let p_i be the probability that A_i precedes all other sequences.

From this description we have

$$\begin{aligned} \mathbb{E}[N_i] &= \mathbb{E}[N] + \mathbb{E}[N_i - N] \\ &= \mathbb{E}[N] + \sum_{j=1}^{n} p_j \mathbb{E}[N_i - N | N = N_j] \end{aligned} \tag{5.3}$$

and given that $\mathbb{E}[N_i] = A_i * A_i - A * A_i$ and $\mathbb{E}[N_i - N | N = N_j] = A_i * A_i - A_j * A_i$ from previous work we can substitute these into equation (5.3) to conclude that

$$\begin{aligned} A_i * A_i - A * A_i &= \mathbb{E}[N] + \sum_{j=1}^{n} p_j (A_i * A_i - A_j * A_i) \\ &= \mathbb{E}[N] + A_i * A_i - \sum_{j=1}^{n} p_j (A_j * A_i) \end{aligned} \tag{5.4}$$

and thus

$$A * A_i = \sum_{j=1}^{n} p_j (A_j * A_i) - \mathbb{E}[N]. \tag{5.5}$$

In matrix form we can write this as

$$\begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ -1 & A_1 * A_1 & A_2 * A_1 & \dots & A_n * A_1 \\ -1 & A_1 * A_2 & A_2 * A_2 & \dots & A_n * A_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & A_1 * A_n & A_2 * A_n & \dots & A_n * A_n \end{bmatrix} \begin{bmatrix} \mathbb{E}[N] \\ p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} 1 \\ A * A_1 \\ A * A_2 \\ \vdots \\ A * A_n \end{bmatrix},$$

and, provided that we can calculate all the terms in the coefficient matrix M, we can solve for the values of $\mathbb{E}[N]$ and p_i for all i.

Now let us relate all of this work back to our network model with $r = 3$ and connection probability p. The probability of a site being in a given state is a function of its forward connections with

$$\begin{aligned} p(0) &= (1-p)^3 \\ p(1) = p(2) = p(4) &= p(1-p)^2 \\ p(3) = p(5) = p(6) &= p^2(1-p) \\ p(7) &= p^3. \end{aligned} \tag{5.6}$$

For simplicity in the current calculations we will use $p = 0.5$ so all states have equal probability of $1/8$ and as such the 'gamblers payoff' is 8 times the stake if they are to guess the correct site state. The shortest possible stopping sequence is of length three for obvious reasons in this system and beyond that they can be of any length except four. However while there are infinitely many potential stopping sequences of ever increasing length the

Table 5.4 Results of a Monte Carlo simulation of 5×10^7 iterations for a system in which $r = 3, \tau = 1$ and $p = 0.5$ showing all stopping sequences of length three.

Stopping Sequence	Count	Approximate probability
(0,0,0)	3004361	0.0600872
(0,4,0)	5301726	0.106035
(2,0,0)	5771157	0.115423
(2,4,0)	6083482	0.12167
(4,0,0)	3059938	0.0611988
(4,4,0)	6080914	0.121618
(6,0,0)	6081648	0.121633
(6,4,0)	6081213	0.121624

probability of the longer sequences occurring becomes infinitesimally low. In order to generate an extensive list of potential stopping sequences one can simply set up a system in its initial format with the first sites being a sequence to test and run the wave through the connected system for all possible configurations. In table (5.4) you see the results of a Monte Carlo simulation of 5×10^7 iterations for a system in which $r = 3, \tau = 1$ and $p = 0.5$ showing all stopping sequences of length three. Over those 5×10^7 iterations there were 26425 distinct stopping sequences with the longest sequence being 23 sites long; however the eight sequences of length three account for 82.9% of all results and all sequences of length three and five account for 95.4%. Importantly from our Monte Carlo simulation we deduce that the expected waiting time, given no initial sequence, for a stopping sequence to occur is 62.281.

In order to utilise the method outlined in the paper by Li we must calculate the coefficient matrix using the previously described product method and then solve given there is no starting sequence. This process is computationally demanding since as you increase the stopping sequence length in consideration the number of potential sequences grows exponentially and the number of calculations involved in producing the coefficient matrix grows quadratically. In this work we computed the full list of results for all sequences up to length 10 which produces an incredibly close approximation to the Monte Carlo results as shown in table (5.5). The full list of results for the Monte Carlo and the matrix calculations have been produced but are too large to include in this work.

The matrix calculation method produces an expected wait time of 62.3005 to encounter a stopping sequence of length ten or lower. Given that this is an approximate method, considering only a truncated set of possible stopping sequences, the result is satisfyingly

Table 5.5 Table showing probabilities of length three stopping sequences occurring. The matrix calculation is performed using all sequences of length ten or lower.

Stopping Sequence	Monte Carlo	Matrix Calculation
(0,0,0)	0.0600872	0.060112
(0,4,0)	0.106035	0.106056
(2,0,0)	0.115423	0.115498
(2,4,0)	0.12167	0.121681
(4,0,0)	0.0611988	0.0611718
(4,4,0)	0.121618	0.121681
(6,0,0)	0.121633	0.121681
(6,4,0)	0.121624	0.121681

accurate. Running a Monte Carlo simulation for a high number of iterations can be very time inefficient as is computing a full list of stop sequences up to a given length to then perform the matrix calculation method. A better solution to generating a fast approximation is a hybrid of both methods; run a short Monte Carlo simulation to generate a list of the most common stopping sequences and then perform the matrix calculation on that shorter list.

It is important to note that using the matrix calculation technique will always produce an upper bound for the expected wait time. We let $B_n = \{A_1, A_2, \ldots, A_n\}$ be the set of all stopping sequences considered in our matrix calculation, B_∞ be the infinite set of all possible stopping sequences and $C = B_\infty \backslash B_n$. Let $Z_1, Z_2, \ldots$ be a sequence of independent identically distributed random variables as before and let N_B, N_C and N_∞ be the waiting time until a sequence from those sets occurs in the process $Z_1, Z_2, \ldots$. If $N_B < N_C$ then $N_B = N_\infty$ since $B_n \subset B_\infty$. If $N_C < N_B$ then $N_B > N_\infty$. Since C is non-empty and all sequences can occur first it must be true that $\mathbb{E}[N_\infty] \leq \mathbb{E}[N_B]$.

This technique is usable for systems with greater link ranges and thresholds by simply adjusting the stopping sequences considered and recalculating the coefficient matrix. Likewise for adjustment in the connection probability, p, the 'gamblers payoff' will vary so the values in the coefficient matrix calculation will change but the process remains consistent.

5.3 Cluster Model

While the previous technique deals with many of the complications of a one dimensional network it has a significant drawback of requiring large computational time. Once you scale the system up in size the ability to calculate accurate approximations becomes more and

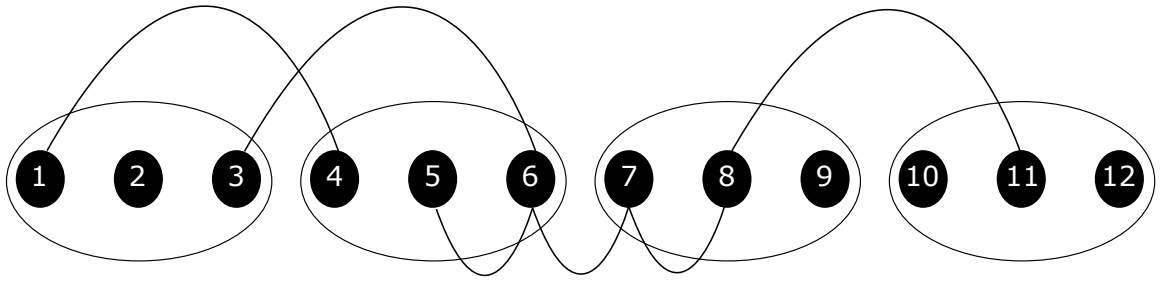


Fig. 5.9 A sample set of connections demonstrating the coarse grained cluster approach. $S_{0,2} = \{3,5,6,7,8\}$, $S_{1,3} = \{5,6,7,8,11\}$ and thus the overlap is non-empty and $B_{0,3}$ has occurred.

more unfeasible due to time constraints. In order to create a more robust model we used the concept of clustering agents into groups and instead of considering the progress of the wave through individual sites we can instead look at the wave progression through coarse-grained clusters. The benefit to clustering agents together into clusters of size r, where r is the range of communication, is that you bypass the previously stated complication of the wave passing by a site only to return to it later. It would not be possible for a cluster C_n to be influenced by an invading opinion without cluster C_{n-1} having been influenced prior.

As with the previous model we will use $r = 3$ and $\tau = 1$ for simplicity. We let C_i denote a given cluster and let $B_{i,i+k}$ denote the event that a given cluster C_i is connected to cluster C_{i+k} through all intermediary clusters. In order to approximate the traveling of an opinion wave over large systems we need to consider how the connections pass through multiple B terms with common connections. For example if $B_{0,2}$ occurs and $B_{1,3}$ occurs then those B terms both include the connections between C_1 and C_2 and we need to know what this says about the probability of $B_{0,3}$ occurring. Let $S_{i,i+2}$ be the set of all sites in clusters C_i, C_{i+1} and C_{i+2} involved in or connected to the sites involved in $B_{i,i+2}$. If the intersection of $S_{i,i+2}$ and $S_{i+1,i+3}$ is non empty then there is an overlap and we know that $B_{0,3}$ has occurred. This is highlighted in figure (5.9) We define the overlap term, O, as follows.

$$O^c_{i,j} = B_{i,j} \cap B_{j-1,j+1} \cap B^c_{i,j+1} \tag{5.7}$$

$$O_{i,j} = B^c_{i,j} \cup B^c_{j-1,j+1} \cup B_{i,j+1}. \tag{5.8}$$

We first define the complement of an overlap from the understanding of what it means for an overlap to not occur, to which we can apply De Morgans law to get the overlap term. By using a combination of B and O terms chained together we can make approximations

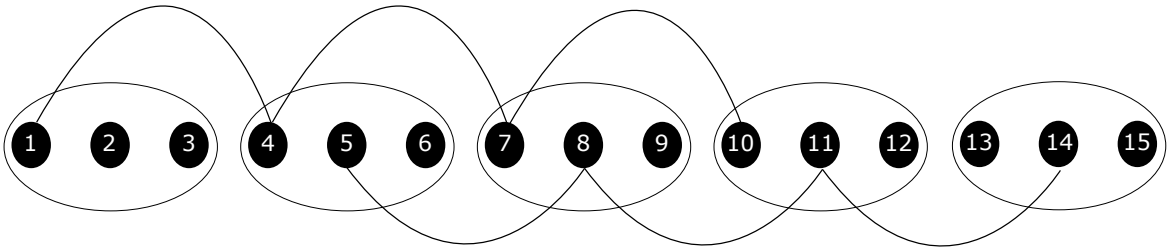


Fig. 5.10 An example system in which $B_{0,2}, B_{1,3}, B_{2,4}, O_{0,2}$ and $O_{1,3}$ all hold but $B_{0,4}$ does not occur.

for the connectivity of much longer systems. This approximation allows us to calculate the probability of much longer range connections by repeatedly using small range probabilities which are easily calculable.

$$\begin{aligned}\mathbb{P}(B_{i,j+1}) &= \mathbb{P}(B_{i,j} \cap O_{i,j} \cap B_{j-1,j+1}) \\ &= \mathbb{P}(B_{j-1,j+1} \cap O_{i,j} | B_{i,j}) \mathbb{P}(B_{i,j}) \\ &\approx \mathbb{P}(B_{j-1,j+1} \cap O_{j-2,j} | B_{i,j}) \mathbb{P}(B_{i,j}) \end{aligned} \tag{5.9}$$

The reason this method is only an approximation is because it is possible for all of the B and O terms in a chain to be satisfied but the wave would be stopped if it wasn't for the coarse grained approach as shown in figure (5.10).

The error that leads to this method being an approximation is due to having multiple routes through the same clusters and the overlaps connecting to the 'wrong' path through. As shown in the previous section however attempting to calculating these waves without using an approximation is not feasible since even calculating $B_{0,9}$ with $r = 3$ requires running through approximately 2×10^{24} permutations. As later results show waves can travel through thousands of clusters making exact calculations completely impossible.

We can rewrite the approximation in equation (5.9) using the definition of the overlap term in equation (5.7).

$$\mathbb{P}(B_{i,j+1}) \approx \mathbb{P}(B_{j-1,j+1} \cap O_{j-2,j} | B_{j-2,j}) \mathbb{P}(B_{i,j}) \tag{5.10}$$

$$= \frac{\mathbb{P}(B_{j-1,j+1} \cap B_{j-2,j} \cap (B^{\mathsf{c}}_{j-2,j} \cup B^{\mathsf{c}}_{j-1,j+1} \cup B_{j-2,j+1}))}{\mathbb{P}(B_{j-2,j})} \mathbb{P}(B_{i,j}) \tag{5.11}$$

$$= \frac{\mathbb{P}(B_{j-1,j+1} \cap B_{j-2,j} \cap B_{j-2,j+1}) \mathbb{P}(B_{i,j})}{\mathbb{P}(B_{j-2,j})} \tag{5.12}$$

$$= \frac{\mathbb{P}(B_{j-2,j+1}) \mathbb{P}(B_{i,j})}{\mathbb{P}(B_{j-2,j})} \tag{5.13}$$

Repeating the process for the $\mathbb{P}(B_{i,j})$ term until it reduces to $\mathbb{P}(B_{i,i+2})$, and assuming spatial homogeneity, we can show that for $n \geq 3$

$$\mathbb{P}(B_{0,n}) = \frac{\mathbb{P}(B_{i,i+3})^{n-2}}{\mathbb{P}(B_{i,i+2})^{n-3}} \tag{5.14}$$

This is a rigorous and new derivation for the usage of overlapping clusters to describe the kinetics of a lattice model outlined by ben-Avraham [28]. This gives us an approximation for the probability of connecting from C_i to C_{j+1}. In order to calculate the expectation for how far a wave travels we also need to include the probability that a wave stops at a certain cluster, say C_{j+1}. In order for that to occur we need to first reach C_{j+1} and then one of two situations. Either to not have the connection $B_{j,j+2}$ or have that connection but not have the overlap term $O_{j-1,j+1}$.

$$\begin{aligned}(B_{j,j+2} \cap O^{\mathsf{c}}_{j-1,j+1}) \cup B^{\mathsf{c}}_{j,j+2} =& (B_{j,j+2} \cup B^{\mathsf{c}}_{j,j+2}) \\ & \cap (O^{\mathsf{c}}_{j-1,j+1} \cup B^{\mathsf{c}}_{j,j+2}) \\ =& B^{\mathsf{c}}_{j,j+2} \cup O^{\mathsf{c}}_{j-1,j+1} \\ =& (B_{j,j+2} \cap O_{j-1,j+1})^{\mathsf{c}}\end{aligned} \tag{5.15}$$

We can define L to be a random variable corresponding to the furthest forward cluster a wave connects to, starting from C_0. We may deduce the expectation of L by calculating $\mathbb{P}(L = n)$ for all $n \geq 0$

$$\mathbb{P}(L = n) = \begin{cases} \mathbb{P}(B^{\mathsf{c}}_{i,i+1}) & n = 0 \\ \mathbb{P}(B_{i,i+1} \cap B^{\mathsf{c}}_{i,i+2}) & n = 1 \\ \mathbb{P}(B_{i,i+2})\mathbb{P}((B_{i+1,i+3} \cap O_{i,i+2})^{\mathsf{c}} | B_{i,i+2}) & n = 2 \\ \frac{\mathbb{P}(B_{i,i+3})^{n-2}}{\mathbb{P}(B_{i,i+2})^{n-3}} \mathbb{P}((B_{i+1,i+3} \cap O_{i,i+2})^{\mathsf{c}} | B_{i,i+2}) & n \geq 3. \end{cases} \tag{5.16}$$

It follows that

$$\begin{aligned}\mathbb{E}(L) =& \mathbb{P}(B_{i,i+1} \cap B^{\mathsf{c}}_{i,i+2}) + \\ & 2\mathbb{P}(B_{i,i+2})\mathbb{P}((B_{i+1,i+3} \cap O_{i,i+2})^{\mathsf{c}} | B_{i,i+2}) + \\ & \mathbb{P}((B_{i+1,i+3} \cap O_{i,i+2})^{\mathsf{c}} | B_{i,i+2}) \sum_{n=3}^{\infty} \frac{n\mathbb{P}(B_{i,i+3})^{n-2}}{\mathbb{P}(B_{i,i+2})^{n-3}}\end{aligned} \tag{5.17}$$

These expressions are all exactly calculable in terms of p for all combinations of link range, r, and threshold, τ. For a system of four clusters, in which the states of only the first three clusters are important and with our example values of $r = 3, \tau = 1$, we can simulate the system for all 8^{27} permutations and get the wave progression counts shown in table (B.1) in the appendices. The probability of each B event can be calculated exactly using said table where each expression is the sum over the relevant column where each individual term has the number of permutations as a coefficient multiplied by p raised to the power of connections made and $(1-p)$ raised to the power of unmade connections. As an example, from the column titled 'no progression', we get

$$\begin{aligned}\mathbb{P}(B^c_{i,i+1}) =& 1p^0(1-p)^{27} + 21p^1(1-p)^{26} + 210p^2(1-p)^{25} + 1330p^3(1-p)^{24} + \\ & 5985p^4(1-p)^{23} + \ldots + 210p^{19}(1-p)^8 + 21p^{20}(1-p)^7 + 1p^{21}(1-p)^6.\end{aligned} \tag{5.18}$$

When we apply this process to each B event and simplify each expression we get the following,

$$\mathbb{P}(B^c_{i,i+1}) = (1-p)^6, \tag{5.19}$$

$$\begin{aligned}\mathbb{P}(B_{i,i+1} \cap B^c_{i,i+2}) =& (1-p)^6 p(6 + 11p - 20p^2 - 58p^3 + 185p^4 - 218p^5 + \\ & 135p^6 - 44p^7 + 6p^8),\end{aligned} \tag{5.20}$$

$$\begin{aligned}\mathbb{P}(B_{i,i+2}) =& p^2(10 + 16p - 122p^2 - 97p^3 + 1668p^4 - 5018p^5 + 8563p^6 - 9758p^7 + \\ & 7834p^8 - 4488p^9 + 1808p^{10} - 489p^{11} + 80p^{12} - 6p^{13}),\end{aligned} \tag{5.21}$$

$$\begin{aligned}\mathbb{P}(B_{i,i+3}) =& -p^3(-15 - 78p + 62p^2 + 1356p^3 + 282p^4 - 21654p^5 + 28502p^6 + \\ & 236937p^7 - 1317827p^8 + 3636910p^9 - 6789919p^{10} + 9406120p^{11} - \\ & 10071638p^{12} + 499938p^{13} - 5696602p^{14} + 3027774p^{15} - 1264000p^{16} + \\ & 406453p^{17} - 97327p^{18} + 16365p^{19} - 1726p^{20} + 86p^{21}),\end{aligned} \tag{5.22}$$

Table 5.6 Table showing the expected furthest forward location for an invading opinion wave using different approximation methods in comparison to a Monte Carlo simulation.

p	Matrix Calculation	Monte Carlo	Cluster Model	Monte Carlo
0.1	4.24570	4.246450 ± 0.00092082	1.59889	1.599207 ± 0.000376435
0.2	6.42674	6.424702 ± 0.00195299	2.41172	2.410474 ± 0.00069223
0.3	11.0823	11.06076 ± 0.00419497	4.0036	3.995358 ± 0.00141711
0.4	23.1918	23.14034 ± 0.0101898	8.05982	8.038073 ± 0.00340359
0.5	62.3897	62.2819 ± 0.0297835	21.144	21.09976 ± 0.00993008
0.6	229.760	229.7294 ± 0.185438	76.9642	76.90930 ± 0.0618138
0.7	1299.9	1299.941 ± 1.05747	433.722	433.6470 ± 0.35249
0.8	15154.8	15153.84 ± 21.3473	5052.09	5051.613 ± 7.11576
0.9	991109	989261.3 ± 1400.51	330469	329754.1 ± 466.838

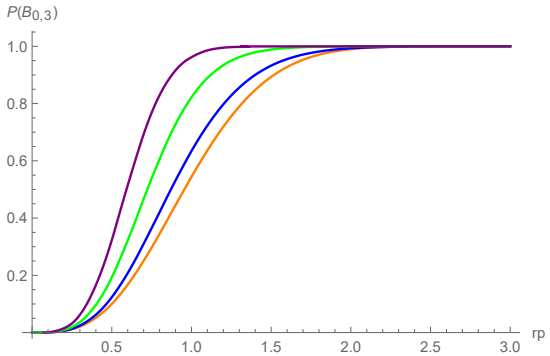
$$\begin{aligned}\mathbb{P}(B_{i,i+2} \cap (B_{i,i+3} \cap O_{i,i+2})^c) =& (-1+p)^6 p^2 (10 + 61p + 16p^2 - 654p^3 - 70p^4 + \\ & 4119p^5 - 291p^6 - 36342p^7 + 110574p^8 - 181112p^9 + \\ & 195745p^{10} - 148513p^{11} + 80336p^{12} - 30567p^{13} + \\ & 7815p^{14} - 1210p^{15} + 86p^{16}) \end{aligned} \tag{5.23}$$

and lastly using the definition of conditional probability

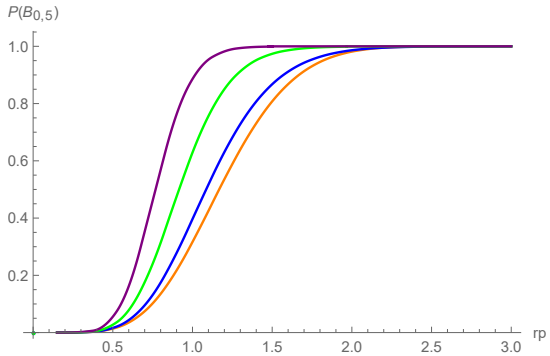
$$\mathbb{P}((B_{i,i+3} \cap O_{i,i+2})^c) | B_{i,i+2}) = \frac{\mathbb{P}(B_{i,i+2} \cap (B_{i,i+3} \cap O_{i,i+2})^c)}{\mathbb{P}(B_{i,i+2})}. \tag{5.24}$$

By substituting equations (5.19), (5.20), (5.21), (5.22), (5.23) and (5.24) into equation (5.17) we can compute the expected furthest forward cluster that the invading wave will reach before becoming pinned. We include these results alongside Monte Carlo simulations and the matrix calculation method from the previous section for comparison, see table (5.6).

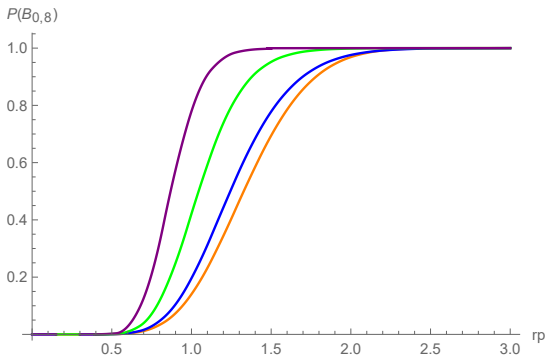
As we see from table (5.6) the results for both methods have a very high level of accuracy in the systems chosen. Each approximation has a level of inaccuracy but the source of inaccuracy in each case is different allowing a choice of approximation depending on the requirements. The matrix calculation approximation has inaccuracy due to the inability to be exhaustive over all cases, however the user can choose how accurate they want to make the method depending on how long they are willing to wait and how much computational power they have access to. The cluster model approach has inherent inaccuracy involved in the overlap term which is unavoidable, however the time taken to calculate an approximation is very low which makes it a useful method to get a fast estimate.



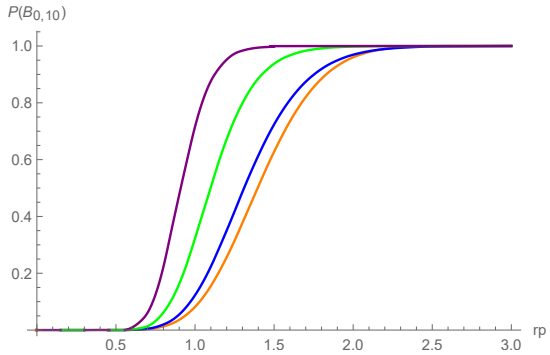
(a) $B_{0,3}$: The probability of connecting three clusters forward.



(b) $B_{0,5}$: The probability of connecting five clusters forward.



(c) $B_{0,8}$: The probability of connecting eight clusters forward.



(d) $B_{0,10}$: The probability of connecting ten clusters forward.

Fig. 5.11 The results of four, one million iteration, Monte Carlo simulations. Plots show the probability of connecting a set number of clusters forward against rp, the average links forward that each agent makes. Purple corresponds to a system with $r = 20$, green represents $r = 10$, blue represents $r = 5$ and orange represents $r = 3$.

Finally in this section we acknowledge the impact that link range and connection probability have on the long range connectivity of the network. With link range r and connection probability p we know that rp is the expected number of forward sites a connection makes. Intuitively as the cluster size increases there are more possible connections between clusters and thus the probability of clusters being connected increases. We would expect that if rp remains fixed then, as cluster size increases, the wave will travel through more clusters and likewise if rp increases the wave will travel through more clusters. The connectivity over a range of clusters for different cluster sizes is displayed in the set of graphs (5.11).

From these plots there is an early indication that as cluster size, or link range, increases there is a critical value of rp for traveling a specific distance. With a value of rp lower than this critical point the probability of reaching the distance is miniscule whereas with a greater

rp you are almost certain to reach the given distance. Interestingly this criticality is very similar to a percolation effect something which is impossible in one dimension, suggesting that the cluster approximation approach is in many ways comparable to a higher dimensional problem due to the ability for the wave to bypass isolated resistant sites.

5.4 Mean Field Random Walk Approach

In this section we outline a different approach to analysing the behaviour of the wave when the cluster size becomes significantly large. As indicated from the cluster model results we would anticipate that for a given link range or cluster size n and some arbitrary distance D there is a critical connection probability p at which point the wave becomes highly likely to travel said distance. However in the cluster model, unless $p = 1$, there is always a chance that the wave is halted at any point due to specific sequences of states. In this final method we take a mean field approach, which bypasses this chance that a progressing wave gets 'unlucky'.

In order to use this approach we first show that the model can be defined exactly using a random walk approach. In the previous definition of the model when the sites have threshold $\tau = 1$ if the wave spreads to a new site x at time t then we know that if x is connected to a non-flipped site y then y flips at time $t + 1$. Given that the link connections between sites in the system remain fixed from initiation if site x has no connected 'down' sites then that part of the wave essentially stops there. We can picture a freshly flipped site as having a random walker at that location. This walker can split and move to any number of 'unflipped' sites within range n at the next time step, with each split having equal probability, p, or the walker dies in this location. If two random walkers in the system land on the same site simultaneously they coalesce into one. The behaviour of coalescing random walks has been investigated previously [98, 99] but with focus on time taken to coalesce rather than the distance and velocity of node exploration.

We next consider the sites themselves and the states that they can be in. For the aforementioned $\tau = 1$ case each site in the system can be in one 3 states at a given time t. The potential states for a site are

1. No random walker has visited this site. Site is 'unflipped'.

2. A random walker is currently at this site. Site has 'flipped'.

3. A random walker has been at this site but is no longer here. Site has 'flipped'.

Let the state of a site x at time t be defined as $S_x(t) \in \{1,2,3\}$. The state of the entire system can be denoted as $\underline{S}(t) = (S_1(t), S_2(t), S_3(t), ...)$.

As n becomes large we are interested in $p \ll \frac{1}{n}$ since we anticipate the critical p value will be very small. We define the indicator function $I(A)$,

$$I(A) = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false.} \end{cases} \tag{5.25}$$

With these stated definitions and the understanding of the magnitude of p we can define the probabilities of a site being in a given stats at time $t = 1$ as

$$\mathbb{P}(S_x(t+1) = 3|\underline{S}(t)) = I(S_x(t) = 2) + I(S_x(t) = 3) \tag{5.26}$$

$$\begin{aligned} \mathbb{P}(S_x(t+1) = 2|\underline{S}(t)) &= I(S_x(t) = 1)(1 - (1-p)^{\sum_{y=x-N}^{y=x+N} I(S_y(t)=2)}) \\ &\approx pI(S_x(t) = 1) \sum_{y=x-N}^{y=x+N} I(S_y(t) = 2) \text{ as } p \to 0 \end{aligned} \tag{5.27}$$

$$\begin{aligned} \mathbb{P}(S_x(t+1) = 1|\underline{S}(t)) &= I(S_x(t) = 1)(1-p)^{\sum_{y=x-N}^{y=x+N} I(S_y(t)=2)} \\ &\approx I(S_x(t) = 1)(1 - p \sum_{y=x-N}^{y=x+N} I(S_y(t) = 2)) \text{ as } p \to 0, \end{aligned} \tag{5.28}$$

where we use the approximation that $(1-p)^l \approx 1 - pl + \frac{p^2}{2}O(l^2)$ for $\frac{1}{p} \ll l$ which holds since l is of order n.

These equations represent a system for which the connection kernel is uniform, that is that the probability of connection between any sites in range is a fixed p. We can however replace that kernel with any probability distribution of our choosing. Later we will consider the kernel to be Gaussian which maintains the connection symmetry but makes closer connections more likely than distant ones. The overall behaviour of the system with these differing kernels should be similar just with different exact solutions.

For systems with any distribution kernel we let p_k = probability of a jump with displacement $k \in \mathbb{Z}$ then we can rewrite the probability of a site not being connected to a site with a random walker (state 2) as

$$\mathbb{P}(\text{not being connected to a 2}) = \prod_{y=-\infty}^{\infty} (1 - p_{x-y})^{I(S_y(t)=2)} \tag{5.29}$$

We can now use this to begin looking at how the system changes over time as the wave progresses. First we note

$$\begin{aligned}\exp\left(\ln\left[\prod_{y=-\infty}^{\infty}(1-p_{x-y})^{I(S_y(t)=2)}\right]\right) &= \exp\left(\sum_{y=-\infty}^{\infty}\ln(1-p_{x-y})I(S_y(t)=2)\right)\\ &= \exp\left(-\sum_{y=-\infty}^{\infty}p_{x-y}I(S_y(t)=2)\right)\\ &\approx 1-\sum_{y=-\infty}^{\infty}p_{x-y}I(S_y(t)=2) \qquad (5.30)\end{aligned}$$

using the approximation that $e^{-x} \approx (1-x)$ for sufficiently small values of x.

We then define

$$\psi_k(x,t) = \mathbb{E}[I(S_x(t)=k)] = \mathbb{P}(S_x(t)=k) \qquad (5.31)$$

for which the discrete time derivative is

$$\begin{aligned}\dot{\psi}_k(x,t) &= \psi_k(x,t+1) - \psi_k(x,t)\\ &= \mathbb{P}(S_x(t+1)=k) - \mathbb{P}(S_x(t)=k)\\ &= \mathbb{E}\Big[\mathbb{E}[I(S_x(t+1)=k|\underline{S}(t)]\Big] - \psi_k(x,t) \qquad (5.32)\end{aligned}$$

Now, for the different k values we have, neglecting correlations between sites,

$$\begin{aligned}\dot{\psi}_1(x,t) &= \psi_1(x,t)(1-\sum_{y=x-n}^{y=x+n}p_{x-y}\psi_2(y,t)) - \psi_1(x,t)\\ &= -\psi_1(x,t)\sum_{y=x-n}^{y=x+n}p_{x-y}\psi_2(y,t) \qquad (5.33)\\ \dot{\psi}_2(x,t) &= \psi_1(x,t)\sum_{y=x-n}^{y=x+n}p_{x-y}\psi_2(y,t) - \psi_2(x,t) \qquad (5.34)\\ \dot{\psi}_3(x,t) &= \psi_2(x,t). \qquad (5.35)\end{aligned}$$

These results are observable from the definition of the system. The rate of change between sites in state 1 between time steps is simply the amount of sites in state 1 that change to state 2. The change in sites in state 2 is caused by all previous state 2 sites transitioning into state 3 with some state 1 sites transitioning into state 2. The limits in equations (5.33) and (5.34) are truncated versions of the limits in (5.30) because p_k is only non-zero for sites in range.

We now promote space and time to continuous variables to get the integral equations

$$\dot{\psi}_1(x,t) = -\psi_1(x,t)\int_{-\infty}^{\infty} p_{x-y}\psi_2(y,t)dy \tag{5.36}$$

$$\dot{\psi}_2(x,t) = \psi_1(x,t)\int_{-\infty}^{\infty} p_{x-y}\psi_2(y,t)dy - \psi_2(x,t) \tag{5.37}$$

omitting the equation for $\psi_3(x,t)$ due to it having no impact on the wave behaviour. We now replace the p notation in equations (5.36) and (5.37) with ϕ representing any probability kernel and perform the integration

$$\begin{aligned}\int_{-\infty}^{\infty}\phi(x-y)\psi_2(y,t)dy &= \int_{-\infty}^{\infty}\phi(u)\psi_2(x-u,t)du \\ &\approx \int_{-\infty}^{\infty}\phi(u)\left[\psi_2(x,t)+\frac{u^2}{2}\psi_2''(x,t)\right]du \\ &= np\psi_2(x,t)+\frac{\psi_2''(x,t)}{2}\int_{-\infty}^{\infty}\phi(u)u^2du,\end{aligned} \tag{5.38}$$

where we use the substitution $u=(x-y)$, and used the Taylor series for $\psi_2(x-u)$ about $x=u$ with the assumption that higher derivatives of ψ_2 are sufficiently small. Therefore if ϕ is a Gaussian kernel the integral has solution

$$\int_{-\infty}^{\infty}\phi(x-y)\psi_2(y,t)dy = np\psi_2(x,t)+\frac{pn^3}{2}\psi_2''(x,t) \tag{5.39}$$

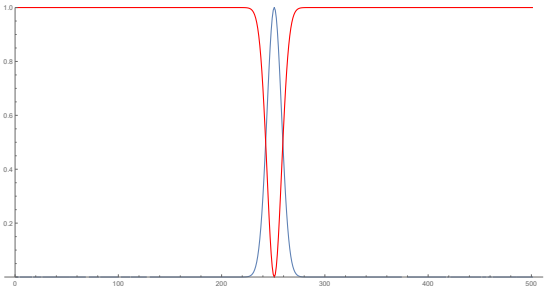
and we get the pair of derivatives

$$\dot{\psi}_1 = -np\psi_1\left[\psi_2+\frac{n^2}{2}\psi_2''\right] \tag{5.40}$$

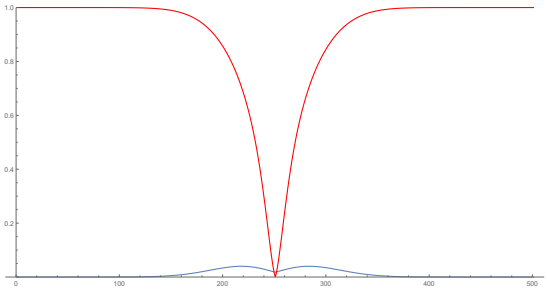
$$\dot{\psi}_2 = np\psi_1\left[\psi_2+\frac{n^2}{2}\psi_2''\right]-\psi_2. \tag{5.41}$$

We can solve these simultaneous equations using the method of lines [100, 101] to see how the proportions of each state in the system progress over time. We omit state 3 for these results since it is easily derived from the other results. We begin with a cluster of sites in state 2 at the centre of the system with all other sites in state 1 and allow the wave to spread through the system. The results for systems with $n=20$ and $p=0.03$ and 0.06 are shown in figures (5.12) and (5.13).

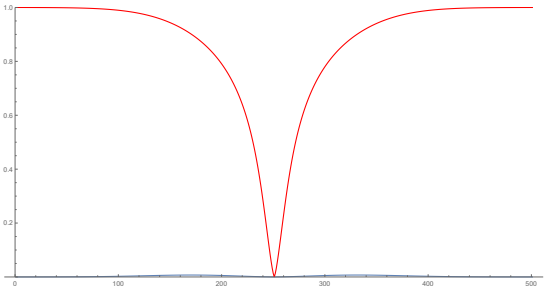
In the case $p=0.03$ we see the wave begins spreading but the size of the propagating cluster is progressively whittled down until the wave stops moving. In the case of $p=0.06$ we



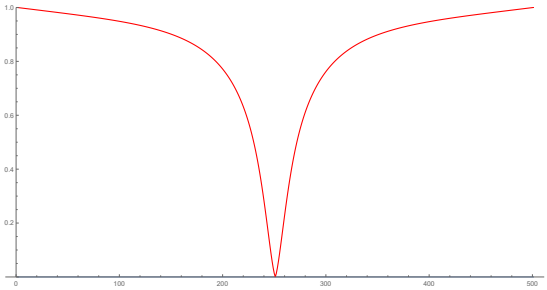
(a) State proportions at site initiation.



(b) Site proportions after one time step.

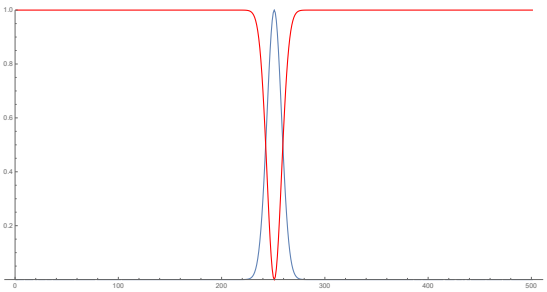


(c) Site proportions after two time steps.

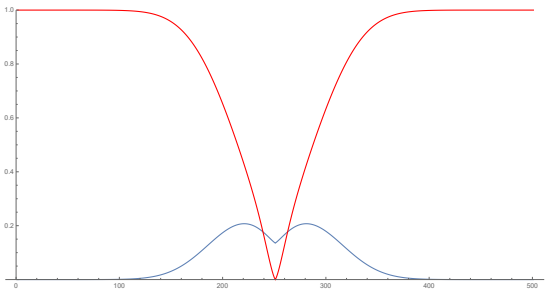


(d) Site proportions after three time steps.

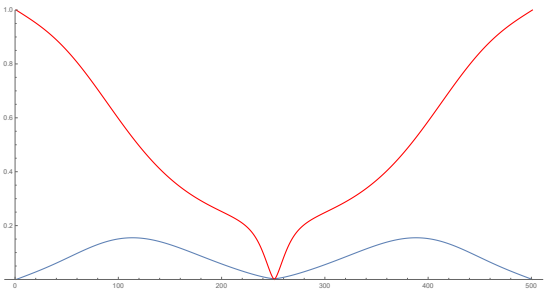
Fig. 5.12 The results for the method of lines on a system with $n = 20$ and $p = 0.03$. The blue line represents the proportion of the system in that state 2, the red line represents the proportion of the system in state 1.



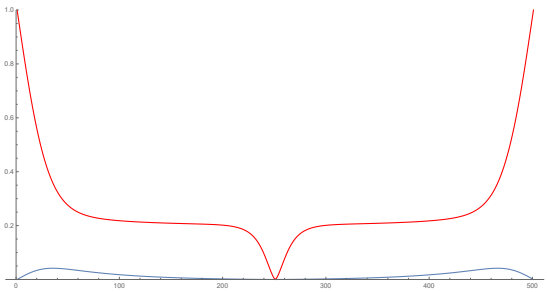
(a) State proportions at site initiation.



(b) Site proportions after one time step.



(c) Site proportions after two time steps.



(d) Site proportions after three time steps

Fig. 5.13 The results for the method of lines on a system with $n = 20$ and $p = 0.06$. The blue line represents the proportion of the system in that state 2, the red line represents the proportion of the system in state 1.

see the propagating cluster converges to a shape which it then maintains forever as the wave will never stop. These results give us a strong indicator that for this set of equations there is a critical np value close to 1 at which point the wave is expected to continue forever. We acknowledge however that this approach ignored correlation between sites and the equations of motion were predicated on small number approximations at the beginning.

5.5 One-Dimensional Network Threshold Model Conclusion

In this chapter we investigated the spread and size of a propagating spread on a one dimensional network. We introduced a toy model to demonstrate the enumeration of site connection states on a simplistic network. We also introduced a close range reciprocal network in one dimension and presented how all network configurations can be expressed as a sequence of integers, which each number representing the binary expansion of each site's forward connections. We demonsrated the idea of stopping sequences, consecutive arrangements of site connection states that guarantee a disconnect in the systems connectivity and thus the ending of an idea spread. We utilised Penney's game [96] and a martingale approach to the study of sequence patterns [97] to compute the expected time for a wave to stop and compared this to simulation results. We then took an alternative approach to estimating the size of a spreading wave by considering a coarse-grained clustering approach. We acknowledge the benefits of this clustered system in terms of computational efficiency and simplicity in how the wave only travels in one direction but also the inherent inaccuracy of the system due to a required approximation. We continued by demonstrating how both approaches are highly accurate in comparison to a Monte Carlo simulation. Finally we considered how this system is analogous to a mean field random walk in one dimension and established a set of site states and transition rates and used numerical methods to show how the proportion of sites in each state varies over time as the wave spreads.

Chapter 6

Conclusion

The aim of this thesis has been to apply models and techniques from Statistical Physics to the spatial ordering and spreading of opinions and exchange of information [12, 13]. We began by modelling the song learning of birds and the way in which they formed distinct dialect boundaries, with particular focus on the Puget Sound White-crowned sparrow [4–6, 57]. Through the use of two dimensional lattice models with localised information exchange we were able to make direct comparisons between the observed field results and the established Ising model [29]. We demonstrated how the boundaries of dialect regions form in the same manner as magnetic domain walls, and how the formation of distinct vertical dialect regions is analogous to the emergence of stripe states in ferro-magnets [19]. We were able to examine the stability of the dialect regions against the destabilising impacts of death by making a direct relationship between the thermodynamic temperature and the death rate of the sparrow. The approach of analysing the song learning and behaviour of birds in comparison to the Ising model of thermodynamics is novel and has potential applications in other species [51].

We investigated interface motion [25, 92] on a two dimensional lattice with site specific memory and nearest neighbour interactions [17, 82]. Our intent was to gain a deeper understanding of the velocity of interface spreading, in particular the impact that variation in site specific memories could have. We attempted to highlight how the structure and pattern of longer and shorter memory agents can have a substantial effect on the the interface shape, roughness and velocity. We demonstrated how approximating the behaviour of a large system can be achieved by considering the system to be a repetition of a much smaller tile for which you can find the velocity analytically. We analysed the system using a collection of low order approximations, giving promising early results. These methods can be extended to look at progressively more sites in order to increase accuracy. Treating this system as the repetition

of smaller tile was a new approach and yielded promising approximations with the potential to be improved.

We created models of opinion spread through a one dimensional random network and analysed the spreading process [13, 16]. We utilised the enumeration of sites in our network in order to generate lists of sequences that were certain to cause a wave to stop and then used martingale methods to calculate the expected stopping time [97]. This method produces a rigorous upper bound on the stopping time which is a close approximation to the simulated Monte Carlo. Lastly we aimed to investigate a cluster approach to approximating expected travel time in our one dimensional random network [28]. As part of our calculations we exactly calculated the probability of connecting between neighbouring clusters and identified an approximation for longer range connections. We tested this cluster approximation for the distance travelled in a random one dimensional network against the previous martingale approach and Monte Carlo simulation. All results were close and suggest both approximation methods are of high accuracy at moderate link ranges.

The overarching theme that is consistent through all chapters of the thesis is that of interfaces. Interfaces are found in nearly all systems in which there is a transfer of information or conformity of ideas, either through local interactions in space or on a network. The shape of the interface in all of these systems controls the rate and locations in which information is exchanged and it is the understanding of how this interface moves, adjusts over time and eventually gets pinned that allows us to predict the long term system behaviours. Chapter 3, the birdsong chapter, has a heavy focus on the original formation of these interfaces, the smoothing of the interface through effects similar to surface tension and the locations of these interfaces within a system. The threshold models in Chapter 4 allow us to begin looking at the speed of a travelling interface through a system with a particular focus on the accelerating effect of surface roughening, in this case driven by individual site variability. In Chapter 5 we begin to work on the understanding of the interface motion on a simple network with a particular focus on pinning and restraints that cause an interface to get stuck.

The results we have found have differing levels of generality ranging from applicable in many areas to very specific to the case we have focused on. The formation of dialect domains in the puget sound white-crowned sparrow and understanding of this through comparison to the Ising model is a technique that can be utilised in many different areas. Since the publication of the paper that Chapter 3 is based upon [23] similar techniques have already been used on human dialects. It is reasonable to consider that this comparison between learnt behaviours through close range interactions and the Ising model would be applicable in many areas but that would need further investigation. Certainly the smoothing of interfaces

through surface tension and the pinning of these interfaces in geographical locations is a generalised phenomenon. The results in Chapter 4 are likely the most general due to the model construction being very broad and unspecific. Being able to calculate the velocity of a travelling interface as some function of a variable localised resistance would have applications in a vast amount of physical systems ranging from the aforementioned dialect spread but also physical effects such as acid erosion. The work in Chapter 5 would likely be the least applicable in other areas as the results are very specific to that specific system and its constraints. The largest drawback to applying that work generally is simply the computational power required to construct an approximation; however potentially with faster processing speeds the enumeration and stopping sequence techniques could be expanded to much more complex systems eventually.

In order to build upon the work in this thesis one would predominantly look at expanding on the mean-field random walk approach in the final chapter and the tiled approach to interface velocity in Chapter 4. The birdsong and Ising model work is interesting but it feels like there are less places in which to expand upon it without just applying the same technique to different behaviours and species. The mean-field random walk approach to explain the information transfer and site states on a one-dimensional network should enable some understanding of long term interface behaviours in much larger systems than the earlier techniques in that chapter can handle. While Chapter 4 has the least conclusive results the tiled approach to approximating interface motion is certainly the broadest topic and the most applicable in a variety of real world applications. Expanding the approximation techniques by looking at progressively larger tiles and perhaps constructing an algorithm to compute the wave velocity, as opposed to going through case by case by hand, could yield some significant results.

References

[1] A. McConchie. The pop vs. soda page. *http://popvssoda.com/*, 2020.

[2] P. Pipek and A. Petrusek. Yellowhammers tweet with regional hashtags. *https://www.bou.org.uk/blog-pipek-petrusek-yellowhammer-dialects*, 2018.

[3] D. A. Nelson and A. Poesel. Tutor choice and imitation accuracy during song learning in a wild population of the puget sound white-crowned sparrow. *Behavioral Ecology and Sociobiology*, 68:1741–1752, 2014.

[4] D. A. Nelson. Song overproduction, selective attrition and song dialects in the white-crowned sparrow. *Animal Behaviour*, 60:887–898, 2000.

[5] L. F. Baptista. Geographic variation in song and dialects of the puget sound white-crowned sparrow. *The Condor*, 79:356–370, 1977.

[6] D. A. Nelson and J. A. Soha. Perception of geographical variation in song by male puget sound white-crowned sparrows, zonotrichia leucophrys pugetensis. *Animal Behaviour*, 68:395–405, 2004.

[7] J. Grinnell and A. H. Miller. *The Distribution of the Birds of California*. Cooper Ornithological Club, 1944.

[8] A. Lamoreaux. Puget sound white-crowned sparrows. *https://nemesisbird.com/birding/identification/puget-sound-white-crowned-sparrows*, 2014.

[9] K. Christensen. *Complexity and Criticality*. Imperial College Press, 2005.

[10] B. A. Cipra. An introduction to the ising model. *The American Mathematical Monthly*, 94(10):937–959, 1987.

[11] H. A. Kramers and G. H. Wannier. Statistics of the two dimensional ferromagnet. *Phys. Rev.*, 60:252, 1941.

[12] P. L. Krapivsky, S. Redner, and E. Ben-Naim. *A Kinetic View of Statistical Physics*. Cambridge University Press, 2010.

[13] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81:591, 2009.

[14] P. Chen and S. Redner. Majority rule dynamics in finite dimensions. *Phys. Rev. E*, 71:036101, 2005.

[15] T. Galla. Intrinsic noise in game dynamical learning. *Phys. Rev. Letters*, 103:198702, 2009.

[16] S. Galam. Minority opinion spreading in random geometry. *Eur. Phys. J. B*, 25:403, 2002.

[17] D. J. Watts. A simple model of global cascades on random networks. *PNAS*, 99(9):5766–5771, 2002.

[18] F. Cacciolo, S. Franz, and M. Marsili. Ising model with memory: coarsening and persistence properties. *J. Stat. Mech. Theor. Exp.*, 2008:07006, 2008.

[19] V. Spirin, P. L. Krapivsky, and S. Redner. Fate of zero-temperature ising ferromagnets. *Phys. Rev. E*, 63:036118, 2001.

[20] M. A. Cunningham and M. C. Baker. Vocal learning in white-crowned sparrows:sensitive phase and song dialects. *Behavioral Ecology and Sociobiology*, 13:259–269, 1983.

[21] J. R. Krebs. Bird song dialects. *Nature*, 257:181–182, 1975.

[22] K. Barros, P. L. Krapivsky, and S. Redner. Freezing into stripe states in two-dimensional ferromagnets and crossing probabilities in critical percolation. *Phys. Rev. E*, 80:040101(R), 2009.

[23] J. Burridge and S. Kenney. Birdsong dialect patterns explained using magnetic domains. *Phys. Rev. E*, 93(6):062402, 2016.

[24] J. Medlock and M. Kot. Spreading disease: integro-differential equation old and new. *Mathematical Biosciences*, 184:201–222, 2003.

[25] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamics scaling of growing interfaces. *Phys. Rev. Letters*, 56:889, 1986.

[26] A. L. Barabási and H. E. Stanley. *Fractal Concepts in Surface Growth*. Cambridge University Press, 1995.

[27] S. Galam. Majority rule hierarchical structures and democratic totalitarism: A statistical approach. *Journal of Mathematical Psychology*, 30(4):426–434, 1986.

[28] D. ben Avraham and J.Kohler. Mean-field (n,m)-cluster approximation for lattice models. *Phys. Rev. A*, 45(12):8358–8370, 1992.

[29] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift Für Physik*, 31(1):253–258, 1924.

[30] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press, 1982.

[31] L. Boltzmann. Studien über das gleichgewicht der lebendigen kraft zwischen bewegten materiellen punkten. *Wiener Berichte*, 58:517–560, 1868.

[32] J. W. Gibbs. *Elementary Principles in Statistical Mechanics*. Charles Scribner's Sons, 1902.

[33] R. J. Glauber. Time-dependent statistics of the ising model. *J. Math. Phys*, 4:294, 1963.

[34] R. Peierls. On ising's model of ferromagnetism. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(3):477, 1936.

[35] L. Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65(3-4):117–149, 1944.

[36] J. C. Miller. Percolation and epidemics in random clustered networks. *Phys. Rev. E*, 80:020901, 2009.

[37] D. Stauffer and A. Aharony. *Introduction To Percolation Theory: Revised Second Edition*. Taylor and Francis, 1994.

[38] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, 2006.

[39] G. Zaklan, F. Westerhoof, and D. Stauffer. Analysing tax evasion dynamics via the ising model. *Journal of Economic Interaction and Coordination*, 4(1), 2009.

[40] D. Stauffer. Social applications of two-dimensional ising models. *American Journal of Physics*, 76:470–473, 2008.

[41] T. C. Shelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2):143–186, 1971.

[42] D. Stauffer and C. Schulze. Urban and scientific segregation: The schelling-ising model. *arXiv:0710.5237*, 2007.

[43] D. A. Nelson and P. Marler. Selection based learning in bird song development. *Proc. Natl. Acad. Sci.*, 91:10498, 1994.

[44] J. Podos and P. S. Warren. The evolution of geographic variation in birdsong. *Adv. Stud. Behav.*, 37:403, 2007.

[45] R. Planque, N. Britton, and H. Slabbekoorn. On the maintenance of bird song dialects. *J. Math. Biol.*, 68:505, 2014.

[46] T. F. Wright, C. R. Dahlin, and A. Salinas-Melgoza. Stability and change in vocal dialects of the yellow-naped amazon. *Animal Behaviour*, 76:1017–1027, 2008.

[47] X. J. Yand, F. M. Lei, G. Wang, and A. J. Jesse. Syllable sharing and inter-individual syllable variation in anna's hummingbird calypte anna songs in san francisco california. *Folia Zoologica*, 56(3):307–318, 2007.

[48] C. Scharff and I. Adam. Neurogenetics of birdsong. *Current Opiniion in Neurobiology*, 23:29–36, 2013.

[49] S. Moorman, C. V. Mello, and J. J. Bolhuis. From songs to synapses: Molecular mechanisms of birdsong memory. *BioEssays*, 33(5):377–385, 2011.

[50] K. N. Laland. Social learning strategies. *Animal Learning and Behaviour*, 32(1):4–14, 2004.

[51] P. Marler and H. Slabbekoorn. *Nature's Music: The Science of Birdsong*. Elsevier, London, 2004.

[52] D. Barrington. Experiments and observations on the singing of birds. *Philosophical Transactions of the Royal Society of London*, 63:249–291, 1773.

[53] W. Thorpe. The process of song-learning in the chaffinch as studied by means of the sound spectrograph. *Nature*, 173:465–469, 1954.

[54] W. Thorpe. The learning of song patterns by birds, with especial reference to the song of the chaffinch fringilla coelebs. *International Journal of Avian Science*, 100(4):535–570, 1958.

[55] P. Marler and M. Tamura. Song dialects in three populations of white-crowned sparrows. *The Condor*, 64:368–377, 1962.

[56] P. Marler. A comparative approach to vocal learning:song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology*, 71:1–25, 1970.

[57] D. A. Nelson. Song overproduction and selective attrition lead to song sharing in the field sparrow (spizella pusilla). *Behavioral Ecology and Sociobiology*, 30:415–424, 1992.

[58] D. A. Nelson. private communication, 2015.

[59] J. W. Evans. Random and cooperative sequential adsorption. *Rev. Mod. Phys.*, 65:1281, 1993.

[60] P. L. Krapivsky and S. Redner. Dynamics of majority rule in two-state interacting spin systems. *Phys. Rev. Letters*, 90:238701, 2003.

[61] S. Galam. Social paradoxes of majority rule voting and renormalization group. *Journal of Statistical Physics*, 61(3-4):943–951, 1990.

[62] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen. Critical coarsening without surface tension: The universality class of the voter model. *Phys. Rev. Letters*, 87:045701–1, 2001.

[63] P. Curie. *Lois expérimentales du magnétisme. Propriétés magnétiques des corps à diverses températures*. Annales de chimie et de physique, 1895.

[64] Z. R. Yang. Glauber dynamics of the kinetic ising model. *Phys. Rev. B*, 46(18):578–584, 1992.

[65] J. Burridge. Spatial evolution of human dialects. *Phys. Rev. X*, 7:031008, 2017.

[66] J. Burridge. Unifying models of dialect spread and extinction using surface tension dynamics. *Royal Society Open Science*, 5:1–21, 2018.

[67] O. García. Language spread and its study in the twenty-first century. *The Oxford Handbook of Applied Linguistics*, 27:398–418, 2010.

[68] J. Fort. Demic and cultural diffusion propagated the neolithic transition across different regions of europe. *Journal of the Royal Society Interface*, 12:0166, 2015.

[69] J. Burridge and R. Inkpen. Formation and arrangement of pits by a corrosive gas. *Phys. Rev. E*, 91:022403, 2015.

[70] G. E. Mustoe. The origin of honeycomb weathering. *Geological Society of America Bulletin*, 93:108–115, 1982.

[71] J. Nash. Non-cooperative games. *Annals of Mathematics*, 54:286, 1951.

[72] M. Milinski and C. Wedekind. Working memory constrains human cooperation in the prisoner's dilema. *Proc. Natl. Acad. Sci. USA*, 95:13755–13758, 1998.

[73] D. Pais, C. H. Caicedo-Nunez, and N. E. Leonard. Hopf bifurcations and limit cycles in evolutionary network dynamics. *SIAM J. Appl. Dyn. Syst.*, 11:1754, 2012.

[74] E. Wesson and R. Rand. Hopf bifurcations in delayed rock-paper-scissors replicator dynamics. *Dy. Games. Appl.*, 2015. online first (DOI) 10.1007/s11071-015-2169-z.

[75] J. Burridge, Y. Gao, and Y. Mao. Forgetfulness can help you win games. *Phys. Rev. E*, 92:032119, 2015.

[76] G. J. Baxter, R. A. Blythe, W. Croft, and A. J. McKane. Utterance selection model of language change. *Phys. Rev. E*, 73(4):046118, 2006.

[77] G. Baxter and W. Croft. Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change*, 28:129–173, 2016.

[78] J. Burridge and M. Gnacik. Infrequent social interaction can accelerate the spread of a persuasive idea. *Phys. Rev. E*, 94:062319, 2016.

[79] J. P. Gleeson, K. P. O'Sullivan, R. A. Banos, and Y. Moreno. Effects of network structure, competition and memory time on social spreading phenomena. *Phys. Rev. X*, 6:021019, 2016.

[80] L. Guzman-Vargas and R. Hernandez-Perez. Small-world topolgy and memory effects on decision time in opinion dynamics. *Physica A*, 372:326, 2006.

[81] P. Dodds and D. J. Watts. Threshold models of social influence. *The Oxford Handbook of Analytical Sociology*, pages 1–23, 2011.

[82] P. S. Dodds and D. J. Watts. Universal behavior in a generalized model of contagion. *Phys. Rev. Letters*, 92(21):8701–1, 2004.

[83] P. Dodds and D. J. Watts. A generalized model of social and biological contagion. *Journal of Theoretical Biology*, 232(4):587–684, 2005.

[84] D. Kempe, J. M. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4):105–147, 2015.

[85] E. Wright, N. M. Khanfar, C. Harrington, and L. E. Kizer. The lasting effects of social media trends on advertising. *Journal of Business and Economics Research*, 8(11):73–80, 2010.

[86] S. C. Chu. Viral advertising in social media. *Journal of Interactive Advertising*, 12(1):30–43, 2011.

[87] S. Engesser, N. Ernst, F. Esser, and F. Büchel. Populism and social media: how politicians spread a fragmented ideology. *Information, Communication and Society*, 20(8):1109–1126, 2017.

[88] J. Ratkiewicz, M. D. Conover, M. Meiss, B. Goncalves, A. Flammini, and F. Menczer. Detecting and tracking political abuse in social media. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[89] S. Vosoughi, D. Roy, and S. Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[90] M. Porter and J. Gleeson. *Dynamical Systems on Networks*. Springer, 2016.

[91] T. Zhou, J. G Liu, W. J. Bai, G. Chen, and B. H. Wang. Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity. *Phys. Rev. E*, 75(05):056109, 2006.

[92] M. Barthe´lemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Phys. Rev. Letters*, 92:178701, 2004.

[93] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Letters*, 86:3200, 2001.

[94] M. E. J. Newman. The spread of epidemic disease on networks. *Phys. Rev. E*, 66:016128, 2002.

[95] D. Centola and M. Macy. Complex contagions and the weakness of long ties. *American Journey of Sociology*, 113(3):702–734, 2007.

[96] W. Penney. Problem 95: Penney-ante. *Journal of Recreational Mathematics*, 2:241, 1969.

[97] S. Y. R. Li. A martingale approach to the study of occurrence of sequence patterns in repeated experiments. *The Annals of Probability*, 8(6):1171–1176, 1980.

[98] R. I. Oliveira. Mean field conditions for coalescing random walks. *Annals of Probability*, 41:3420–3461, 2013.

[99] C. Cooper, A. Frieze, and T. Radzik. Multiple random walks in random regular graphs. *Society of Industrial and Applied Mathematics Journal of Discrete Mathematics*, 23:3420–3461, 2009.

[100] W. E. Schiesser. *The numerical method of lines: Integration of partial differential equations*. Academic Press, 1991.

[101] Z. Kamont. Numerical method of lines. *Hyperbolic Functional Differential Inequalities and Applications*, pages 181–204, 1999.

Appendix A

Dialect selection probabilities

$$p_{\uparrow}(4,4,4) = \frac{3}{256}\alpha^{11} + \frac{5}{256}\alpha^{10} + \frac{1}{128}\alpha^{9} - \frac{3}{64}\alpha^{8} - \frac{1}{64}\alpha^{7} - \frac{3}{32}\alpha^{6} - \frac{17}{128}\alpha^{4} + \frac{1}{2}$$

$$p_{\uparrow}(4,4,2) = \frac{9}{512}\alpha^{10} - \frac{3}{512}\alpha^{9} + \frac{3}{256}\alpha^{7} - \frac{17}{128}\alpha^{6} + \frac{5}{128}\alpha^{5} - \frac{23}{128}\alpha^{4} + \frac{1}{2}$$

$$\begin{aligned} p_{\uparrow}(4,4,0) = &- \frac{3}{512}\alpha^{11} + \frac{3}{128}\alpha^{10} - \frac{31}{768}\alpha^{9} + \frac{1}{16}\alpha^{8} - \frac{19}{1536}\alpha^{7} - \frac{53}{384}\alpha^{6} + \frac{29}{256}\alpha^{5} \ldots \\ &- \frac{95}{384}\alpha^{4} - \frac{1}{192}\alpha^{3} + \frac{1}{2} \end{aligned}$$

$$\begin{aligned} p_{\uparrow}(4,4,-2) = &- \frac{3}{512}\alpha^{11} + \frac{13}{512}\alpha^{10} - \frac{41}{512}\alpha^{9} + \frac{5}{32}\alpha^{8} - \frac{57}{512}\alpha^{7} - \frac{29}{256}\alpha^{6} + \frac{59}{256}\alpha^{5} \ldots \\ &- \frac{43}{128}\alpha^{4} - \frac{1}{64}\alpha^{3} + \frac{1}{2} \end{aligned}$$

$$p_{\uparrow}(4,4,-4) = \frac{3}{256}\alpha^{10} - \frac{7}{64}\alpha^{9} + \frac{19}{64}\alpha^{8} - \frac{79}{256}\alpha^{7} - \frac{1}{16}\alpha^{6} + \frac{51}{128}\alpha^{5} - \frac{57}{128}\alpha^{4} - \frac{1}{32}\alpha^{3} + \frac{1}{2}$$

$$\begin{aligned} p_{\uparrow}(4,2,4) = &\frac{3}{512}\alpha^{11} + \frac{1}{256}\alpha^{10} - \frac{15}{512}\alpha^{9} + \frac{5}{128}\alpha^{8} - \frac{1}{256}\alpha^{7} - \frac{13}{128}\alpha^{6} + \frac{3}{32}\alpha^{5} - \frac{33}{128}\alpha^{4} \ldots \\ &+ \frac{1}{2} \end{aligned}$$

$$\begin{aligned} p_{\uparrow}(4,2,2) = &- \frac{9}{2048}\alpha^{11} + \frac{3}{512}\alpha^{10} - \frac{21}{1024}\alpha^{9} + \frac{51}{1024}\alpha^{8} - \frac{19}{2048}\alpha^{7} - \frac{99}{1024}\alpha^{6} \ldots \\ &+ \frac{101}{1024}\alpha^{5} - \frac{59}{256}\alpha^{4} - \frac{11}{256}\alpha^{3} + \frac{1}{2} \end{aligned}$$

$$p_{\uparrow}(4,2,0) = -\frac{5}{1024}\alpha^{11} + \frac{11}{3072}\alpha^{10} - \frac{19}{768}\alpha^{9} + \frac{13}{256}\alpha^{8} + \frac{15}{1024}\alpha^{7} - \frac{343}{3072}\alpha^{6} + \frac{155}{1536}\alpha^{5} \ldots - \frac{35}{192}\alpha^{4} - \frac{35}{384}\alpha^{3} - \frac{1}{192}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,2,-2) = \frac{3}{2048}\alpha^{11} - \frac{7}{512}\alpha^{10} + \frac{5}{1024}\alpha^{9} - \frac{11}{1024}\alpha^{8} + \frac{181}{2048}\alpha^{7} - \frac{141}{1024}\alpha^{6} + \frac{85}{1024}\alpha^{5} \ldots - \frac{27}{256}\alpha^{4} - \frac{37}{256}\alpha^{3} - \frac{1}{64}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,2,-4) = \frac{3}{256}\alpha^{11} - \frac{29}{512}\alpha^{10} + \frac{59}{512}\alpha^{9} - \frac{3}{16}\alpha^{8} + \frac{119}{512}\alpha^{7} - \frac{85}{512}\alpha^{6} + \frac{7}{256}\alpha^{5} + \frac{1}{128}\alpha^{4} \ldots - \frac{13}{64}\alpha^{3} - \frac{1}{32}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,0,4) = -\frac{1}{512}\alpha^{11} - \frac{7}{768}\alpha^{10} - \frac{7}{768}\alpha^{9} + \frac{1}{12}\alpha^{8} - \frac{127}{1536}\alpha^{7} - \frac{31}{384}\alpha^{6} + \frac{155}{768}\alpha^{5} \ldots - \frac{35}{128}\alpha^{4} - \frac{5}{64}\alpha^{3} + \frac{1}{2}$$

$$p_{\uparrow}(4,0,2) = -\frac{7}{1024}\alpha^{11} + \frac{5}{1536}\alpha^{10} - \frac{19}{1536}\alpha^{9} + \frac{1}{64}\alpha^{8} + \frac{57}{1024}\alpha^{7} - \frac{35}{256}\alpha^{6} + \frac{205}{1536}\alpha^{5} \ldots - \frac{59}{384}\alpha^{4} - \frac{53}{384}\alpha^{3} - \frac{1}{96}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,0,0) = -\frac{5}{384}\alpha^{10} + \frac{23}{1152}\alpha^{9} - \frac{65}{1152}\alpha^{8} + \frac{1}{9}\alpha^{7} - \frac{23}{288}\alpha^{6} + \frac{1}{48}\alpha^{5} - \frac{59}{1152}\alpha^{4} - \frac{1}{6}\alpha^{3} \ldots - \frac{5}{144}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,0,-2) = \frac{7}{1024}\alpha^{11} - \frac{29}{1536}\alpha^{10} + \frac{9}{512}\alpha^{9} - \frac{17}{384}\alpha^{8} + \frac{85}{3072}\alpha^{7} + \frac{65}{768}\alpha^{6} - \frac{55}{512}\alpha^{5} \ldots + \frac{1}{64}\alpha^{4} - \frac{61}{384}\alpha^{3} - \frac{7}{96}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,0,-4) = \frac{1}{512}\alpha^{11} + \frac{19}{768}\alpha^{10} - \frac{23}{256}\alpha^{9} + \frac{9}{64}\alpha^{8} - \frac{385}{1536}\alpha^{7} + \frac{45}{128}\alpha^{6} - \frac{57}{256}\alpha^{5} + \frac{11}{384}\alpha^{4} \ldots - \frac{7}{64}\alpha^{3} - \frac{1}{8}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-2,4) = -\frac{3}{256}\alpha^{11} + \frac{5}{512}\alpha^{10} + \frac{21}{512}\alpha^{9} - \frac{5}{32}\alpha^{8} + \frac{129}{512}\alpha^{7} - \frac{99}{512}\alpha^{6} + \frac{9}{256}\alpha^{5} \ldots + \frac{1}{128}\alpha^{4} - \frac{13}{64}\alpha^{3} - \frac{1}{32}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-2,2) = -\frac{3}{2048}\alpha^{11} - \frac{5}{512}\alpha^{10} + \frac{27}{1024}\alpha^{9} - \frac{65}{1024}\alpha^{8} + \frac{155}{2048}\alpha^{7} - \frac{11}{1024}\alpha^{6} \ldots - \frac{29}{1024}\alpha^{5} + \frac{1}{128}\alpha^{4} - \frac{47}{256}\alpha^{3} - \frac{1}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-2,0) = \frac{5}{1024}\alpha^{11} - \frac{11}{3072}\alpha^{10} - \frac{17}{768}\alpha^{9} + \frac{9}{256}\alpha^{8} - \frac{95}{1024}\alpha^{7} + \frac{503}{3072}\alpha^{6} - \frac{139}{1536}\alpha^{5} \ldots + \frac{7}{384}\alpha^{4} - \frac{21}{128}\alpha^{3} - \frac{19}{192}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-2,-2) = \frac{9}{2048}\alpha^{11} + \frac{9}{512}\alpha^{10} - \frac{59}{1024}\alpha^{9} + \frac{89}{1024}\alpha^{8} - \frac{477}{2048}\alpha^{7} + \frac{347}{1024}\alpha^{6} - \frac{173}{1024}\alpha^{5} \ldots + \frac{3}{64}\alpha^{4} - \frac{37}{256}\alpha^{3} - \frac{9}{64}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-2,-4) = -\frac{3}{512}\alpha^{11} + \frac{11}{256}\alpha^{10} - \frac{17}{512}\alpha^{9} + \frac{5}{128}\alpha^{8} - \frac{83}{256}\alpha^{7} + \frac{67}{128}\alpha^{6} - \frac{9}{32}\alpha^{5} \ldots + \frac{13}{128}\alpha^{4} - \frac{1}{8}\alpha^{3} - \frac{3}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-4,4) = \frac{9}{256}\alpha^{10} - \frac{3}{16}\alpha^{9} + \frac{27}{64}\alpha^{8} - \frac{169}{256}\alpha^{7} + \frac{25}{32}\alpha^{6} - \frac{67}{128}\alpha^{5} + \frac{5}{128}\alpha^{4} + \frac{1}{32}\alpha^{3} \ldots - \frac{3}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-4,2) = \frac{3}{512}\alpha^{11} + \frac{11}{512}\alpha^{10} - \frac{63}{512}\alpha^{9} + \frac{1}{4}\alpha^{8} - \frac{231}{512}\alpha^{7} + \frac{157}{256}\alpha^{6} - \frac{107}{256}\alpha^{5} + \frac{11}{128}\alpha^{4} \ldots - \frac{3}{64}\alpha^{3} - \frac{3}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-4,0) = \frac{3}{512}\alpha^{11} + \frac{3}{128}\alpha^{10} - \frac{53}{768}\alpha^{9} + \frac{7}{64}\alpha^{8} - \frac{159}{512}\alpha^{7} + \frac{59}{128}\alpha^{6} - \frac{69}{256}\alpha^{5} + \frac{15}{128}\alpha^{4} \ldots - \frac{25}{192}\alpha^{3} - \frac{3}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-4,-2) = \frac{15}{512}\alpha^{10} - \frac{5}{512}\alpha^{9} + \frac{1}{64}\alpha^{8} - \frac{67}{256}\alpha^{7} + \frac{41}{128}\alpha^{6} - \frac{9}{128}\alpha^{5} + \frac{17}{128}\alpha^{4} - \frac{7}{32}\alpha^{3} \ldots - \frac{3}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(4,-4,-4) = -\frac{3}{256}\alpha^{11} + \frac{7}{256}\alpha^{10} + \frac{9}{128}\alpha^{9} - \frac{1}{64}\alpha^{8} - \frac{21}{64}\alpha^{7} + \frac{3}{16}\alpha^{6} + \frac{3}{16}\alpha^{5} + \frac{17}{128}\alpha^{4} \ldots - \frac{5}{16}\alpha^{3} - \frac{3}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,4,4) = \frac{3}{256}\alpha^{11} + \frac{1}{128}\alpha^{10} - \frac{3}{256}\alpha^{9} - \frac{1}{32}\alpha^{8} - \frac{9}{128}\alpha^{7} + \frac{9}{128}\alpha^{6} - \frac{47}{128}\alpha^{4} + \frac{39}{64}\alpha^{3} \ldots - \frac{15}{32}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,4,2) = \frac{3}{512}\alpha^{10} - \frac{1}{512}\alpha^{9} - \frac{1}{256}\alpha^{8} - \frac{17}{256}\alpha^{7} + \frac{5}{64}\alpha^{6} - \frac{9}{128}\alpha^{5} - \frac{85}{256}\alpha^{4} + \frac{21}{32}\alpha^{3} \ldots - \frac{33}{64}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,4,0) = -\frac{3}{512}\alpha^{11} + \frac{3}{256}\alpha^{10} - \frac{5}{384}\alpha^{9} + \frac{5}{256}\alpha^{8} - \frac{37}{512}\alpha^{7} + \frac{7}{64}\alpha^{6} - \frac{33}{256}\alpha^{5} - \frac{79}{256}\alpha^{4} \ldots + \frac{269}{384}\alpha^{3} - \frac{9}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,4,-2) = -\frac{3}{512}\alpha^{11} + \frac{7}{512}\alpha^{10} - \frac{15}{512}\alpha^{9} + \frac{7}{128}\alpha^{8} - \frac{57}{512}\alpha^{7} + \frac{41}{256}\alpha^{6} - \frac{43}{256}\alpha^{5} - \frac{19}{64}\alpha^{4} \ldots + \frac{95}{128}\alpha^{3} - \frac{39}{64}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,4,-4) = -\frac{9}{256}\alpha^{9} + \frac{15}{128}\alpha^{8} - \frac{53}{256}\alpha^{7} + \frac{29}{128}\alpha^{6} - \frac{23}{128}\alpha^{5} - \frac{19}{64}\alpha^{4} + \frac{25}{32}\alpha^{3} - \frac{21}{32}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,2,4) = \frac{3}{512}\alpha^{11} - \frac{1}{128}\alpha^{10} - \frac{7}{512}\alpha^{9} + \frac{5}{256}\alpha^{8} - \frac{1}{16}\alpha^{7} + \frac{31}{256}\alpha^{6} - \frac{9}{64}\alpha^{5} - \frac{5}{16}\alpha^{4} \ldots + \frac{45}{64}\alpha^{3} - \frac{9}{16}\alpha^{2} + \frac{1}{2}$$

$$p_{\uparrow}(2,2,2) = -\frac{9}{2048}\alpha^{11} - \frac{1}{1024}\alpha^{9} + \frac{1}{64}\alpha^{8} - \frac{111}{2048}\alpha^{7} + \frac{91}{1024}\alpha^{6} - \frac{97}{1024}\alpha^{5} - \frac{333}{1024}\alpha^{4} \ldots + \frac{11}{16}\alpha^{3} - \frac{69}{128}\alpha^{2} - \frac{3}{128}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,2,0) = -\frac{5}{1024}\alpha^{11} + \frac{11}{3072}\alpha^{10} - \frac{5}{384}\alpha^{9} + \frac{15}{512}\alpha^{8} - \frac{37}{1024}\alpha^{7} + \frac{121}{3072}\alpha^{6} - \frac{65}{1536}\alpha^{5} \ldots - \frac{505}{1536}\alpha^{4} + \frac{171}{256}\alpha^{3} - \frac{199}{384}\alpha^{2} - \frac{3}{64}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,2,-2) = \frac{3}{2048}\alpha^{11} - \frac{1}{128}\alpha^{10} - \frac{3}{1024}\alpha^{9} + \frac{1}{128}\alpha^{8} + \frac{25}{2048}\alpha^{7} - \frac{19}{1024}\alpha^{6} - \frac{1}{1024}\alpha^{5} \ldots - \frac{323}{3024}\alpha^{4} + \frac{165}{256}\alpha^{3} - \frac{1}{2}\alpha^{2} - \frac{9}{128}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,2,-4) = \frac{3}{256}\alpha^{11} - \frac{23}{512}\alpha^{10} + \frac{39}{512}\alpha^{9} - \frac{13}{128}\alpha^{8} + \frac{57}{512}\alpha^{7} - \frac{39}{512}\alpha^{6} + \frac{3}{256}\alpha^{5} \ldots - \frac{71}{256}\alpha^{4} + \frac{79}{128}\alpha^{3} - \frac{31}{64}\alpha^{2} - \frac{3}{32}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,0,4) = -\frac{1}{512}\alpha^{11} - \frac{5}{384}\alpha^{10} + \frac{1}{64}\alpha^{9} + \frac{7}{256}\alpha^{8} - \frac{143}{1536}\alpha^{7} + \frac{15}{128}\alpha^{6} + \frac{23}{256}\alpha^{5} \ldots - \frac{247}{768}\alpha^{4} + \frac{45}{64}\alpha^{3} - \frac{9}{16}\alpha^{2} - \frac{1}{32}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,0,2) = -\frac{7}{1024}\alpha^{11} + \frac{11}{1536}\alpha^{10} - \frac{7}{512}\alpha^{9} + \frac{35}{1536}\alpha^{8} - \frac{37}{3072}\alpha^{7} + \frac{7}{768}\alpha^{6} - \frac{11}{512}\alpha^{5} \ldots - \frac{161}{512}\alpha^{4} + \frac{509}{768}\alpha^{3} - \frac{25}{48}\alpha^{2} - \frac{1}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,0,-2) = \frac{7}{1024}\alpha^{11} - \frac{23}{1536}\alpha^{10} + \frac{25}{1536}\alpha^{9} - \frac{19}{512}\alpha^{8} + \frac{55}{1024}\alpha^{7} - \frac{9}{256}\alpha^{6} + \frac{53}{1536}\alpha^{5} \ldots - \frac{487}{1536}\alpha^{4} + \frac{481}{768}\alpha^{3} - \frac{11}{24}\alpha^{2} - \frac{1}{8}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,0,-4) = \frac{1}{512}\alpha^{11} + \frac{1}{48}\alpha^{10} - \frac{25}{384}\alpha^{9} + \frac{65}{768}\alpha^{8} - \frac{113}{1536}\alpha^{7} + \frac{7}{384}\alpha^{6} + \frac{61}{768}\alpha^{5} \ldots - \frac{93}{256}\alpha^{4} + \frac{41}{64}\alpha^{3} - \frac{7}{16}\alpha^{2} - \frac{5}{32}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-2,4) = -\frac{3}{256}\alpha^{11} + \frac{11}{512}\alpha^{10} + \frac{1}{512}\alpha^{9} - \frac{9}{128}\alpha^{8} + \frac{67}{512}\alpha^{7} - \frac{53}{512}\alpha^{6} + \frac{5}{256}\alpha^{5} \ldots - \frac{71}{256}\alpha^{4} + \frac{79}{128}\alpha^{3} - \frac{31}{64}\alpha^{2} - \frac{3}{32}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-2,2) = -\frac{3}{2048}\alpha^{11} - \frac{1}{256}\alpha^{10} + \frac{19}{1024}\alpha^{9} - \frac{23}{512}\alpha^{8} + \frac{143}{2048}\alpha^{7} - \frac{57}{1024}\alpha^{6} \ldots + \frac{29}{1024}\alpha^{5} - \frac{303}{1024}\alpha^{4} \ldots + \frac{161}{256}\alpha^{3} - \frac{61}{128}\alpha^{2} - \frac{15}{128}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-2,0) = \frac{5}{1024}\alpha^{11} - \frac{11}{3072}\alpha^{10} - \frac{1}{96}\alpha^{9} + \frac{7}{512}\alpha^{8} - \frac{3}{1024}\alpha^{7} - \frac{41}{3072}\alpha^{6} + \frac{73}{1536}\alpha^{5} \ldots - \frac{485}{1536}\alpha^{4} + \frac{493}{768}\alpha^{3} - \frac{181}{384}\alpha^{2} - \frac{9}{64}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-2,-2) = \frac{9}{2048}\alpha^{11} + \frac{3}{256}\alpha^{10} - \frac{39}{1024}\alpha^{9} + \frac{27}{512}\alpha^{8} - \frac{137}{2048}\alpha^{7} + \frac{33}{1024}\alpha^{6} \ldots + \frac{61}{1024}\alpha^{5} - \frac{337}{1024}\alpha^{4} + \frac{21}{32}\alpha^{3} - \frac{15}{32}\alpha^{2} - \frac{21}{128}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-2,-4) = -\frac{3}{512}\alpha^{11} + \frac{1}{32}\alpha^{10} - \frac{9}{512}\alpha^{9} + \frac{5}{256}\alpha^{8} - \frac{13}{128}\alpha^{7} + \frac{23}{256}\alpha^{6} + \frac{3}{64}\alpha^{5} - \frac{21}{64}\alpha^{4} \ldots + \frac{43}{64}\alpha^{3} - \frac{15}{32}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-4,4) = \frac{3}{128}\alpha^{10} - \frac{29}{356}\alpha^{9} + \frac{31}{128}\alpha^{8} - \frac{71}{256}\alpha^{7} + \frac{17}{128}\alpha^{6} + \frac{15}{128}\alpha^{5} - \frac{15}{32}\alpha^{4} + \frac{21}{32}\alpha^{3} \ldots - \frac{3}{8}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-4,2) = \frac{3}{512}\alpha^{11} + \frac{5}{512}\alpha^{10} - \frac{37}{512}\alpha^{9} + \frac{19}{128}\alpha^{8} - \frac{87}{512}\alpha^{7} + \frac{23}{256}\alpha^{6} + \frac{19}{256}\alpha^{5} - \frac{25}{64}\alpha^{4} \ldots + \frac{85}{128}\alpha^{3} - \frac{27}{64}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-4,0) = \frac{3}{512}\alpha^{11} + \frac{3}{256}\alpha^{10} - \frac{1}{24}\alpha^{9} + \frac{17}{256}\alpha^{8} - \frac{137}{1536}\alpha^{7} + \frac{5}{96}\alpha^{6} + \frac{13}{256}\alpha^{5} - \frac{245}{768}\alpha^{4} \ldots + \frac{257}{384}\alpha^{3} - \frac{15}{32}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-4,-2) = \frac{9}{512}\alpha^{10} - \frac{3}{512}\alpha^{9} + \frac{3}{256}\alpha^{8} - \frac{15}{256}\alpha^{7} + \frac{1}{64}\alpha^{6} + \frac{7}{128}\alpha^{5} - \frac{65}{256}\alpha^{4} + \frac{43}{64}\alpha^{3} \ldots - \frac{33}{64}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(2,-4,-4) = -\frac{3}{256}\alpha^{11} + \frac{1}{64}\alpha^{10} + \frac{13}{256}\alpha^{9} - \frac{13}{129}\alpha^{7} - \frac{3}{128}\alpha^{6} + \frac{3}{32}\alpha^{5} - \frac{25}{128}\alpha^{4} + \frac{43}{64}\alpha^{3} \ldots - \frac{9}{16}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,4,4) = \frac{3}{128}\alpha^{11} - \frac{1}{128}\alpha^{10} - \frac{1}{16}\alpha^{9} - \frac{1}{32}\alpha^{8} - \frac{1}{16}\alpha^{7} + \frac{7}{32}\alpha^{6} - \frac{1}{16}\alpha^{5} - \frac{9}{64}\alpha^{4} \ldots - \frac{3}{16}\alpha^{3} + \frac{1}{16}\alpha^{2} + \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,4,2) = -\frac{3}{256}\alpha^{10} + \frac{1}{256}\alpha^{9} - \frac{1}{64}\alpha^{8} - \frac{13}{128}\alpha^{7} + \frac{15}{64}\alpha^{6} - \frac{13}{64}\alpha^{5} - \frac{3}{32}\alpha^{3} \ldots$$
$$- \frac{1}{16}\alpha^{2} + \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,4,0) = \frac{3}{256}\alpha^{11} + \frac{11}{384}\alpha^{9} - \frac{3}{64}\alpha^{8} - \frac{59}{768}\alpha^{7} + \frac{53}{192}\alpha^{6} - \frac{47}{128}\alpha^{5} + \frac{13}{96}\alpha^{4} \ldots$$
$$- \frac{3}{16}\alpha^{2} + \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,4,-2) = -\frac{3}{256}\alpha^{11} + \frac{1}{256}\alpha^{10} + \frac{11}{256}\alpha^{9} - \frac{3}{32}\alpha^{8} - \frac{9}{256}\alpha^{7} + \frac{43}{128}\alpha^{6} - \frac{69}{128}\alpha^{5} + \frac{17}{64}\alpha^{4} \ldots$$
$$+ \frac{3}{32}\alpha^{3} - \frac{5}{16}\alpha^{2} + \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,2,-4) = -\frac{3}{128}\alpha^{10} + \frac{5}{64}\alpha^{9} - \frac{1}{8}\alpha^{8} - \frac{3}{128}\alpha^{7} + \frac{13}{32}\alpha^{6} - \frac{45}{64}\alpha^{5} + \frac{25}{64}\alpha^{4} + \frac{3}{16}\alpha^{3} - \frac{7}{16}\alpha^{2} \ldots$$
$$+ \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,2,4) = \frac{3}{256}\alpha^{11} - \frac{5}{128}\alpha^{10} + \frac{1}{256}\alpha^{9} - \frac{7}{128}\alpha^{7} + \frac{1}{4}\alpha^{6} - \frac{3}{8}\alpha^{5} + \frac{9}{64}\alpha^{4} - \frac{3}{16}\alpha^{2} + \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,2,2) = -\frac{9}{1024}\alpha^{11} - \frac{3}{256}\alpha^{10} + \frac{19}{512}\alpha^{9} - \frac{19}{512}\alpha^{8} - \frac{59}{1024}\alpha^{7} + \frac{113}{512}\alpha^{6} - \frac{151}{256}\alpha^{5} \ldots$$
$$+ \frac{25}{256}\alpha^{4} + \frac{1}{128}\alpha^{3} - \frac{9}{64}\alpha^{2} + \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,2,-2) = \frac{3}{1024}\alpha^{11} - \frac{1}{256}\alpha^{10} - \frac{11}{512}\alpha^{9} + \frac{27}{512}\alpha^{8} - \frac{83}{1024}\alpha^{7} + \frac{47}{512}\alpha^{6} - \frac{39}{512}\alpha^{5} \ldots$$
$$+ \frac{3}{256}\alpha^{4} + \frac{1}{128}\alpha^{3} - \frac{3}{64}\alpha^{2} + \frac{1}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,2,-4) = \frac{3}{128}\alpha^{11} - \frac{17}{256}\alpha^{10} + \frac{19}{256}\alpha^{9} - \frac{1}{32}\alpha^{8} - \frac{5}{256}\alpha^{7} + \frac{7}{256}\alpha^{6} - \frac{1}{128}\alpha^{5} + \frac{1}{2}$$

$$p_{\uparrow}(0,0,4) = -\frac{1}{256}\alpha^{11} - \frac{13}{384}\alpha^{10} + \frac{31}{384}\alpha^{9} - \frac{11}{192}\alpha^{8} - \frac{21}{156}\alpha^{7} + \frac{53}{192}\alpha^{6} - \frac{47}{128}\alpha^{5} \ldots$$
$$+ \frac{5}{32}\alpha^{4} + \frac{7}{96}\alpha^{3} - \frac{5}{24}\alpha^{2} + \frac{1}{6}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,0,2) = -\frac{7}{512}\alpha^{11} + \frac{17}{768}\alpha^{10} - \frac{23}{768}\alpha^{9} + \frac{23}{384}\alpha^{8} - \frac{149}{1536}\alpha^{7} + \frac{17}{128}\alpha^{6} - \frac{119}{768}\alpha^{5} \ldots + \frac{23}{384}\alpha^{4} + \frac{1}{24}\alpha^{3} - \frac{5}{48}\alpha^{2} + \frac{1}{12}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,0,-2) = \frac{7}{512}\alpha^{11} - \frac{17}{768}\alpha^{10} + \frac{23}{768}\alpha^{9} - \frac{23}{384}\alpha^{8} + \frac{149}{1536}\alpha^{7} - \frac{17}{128}\alpha^{6} + \frac{119}{768}\alpha^{5} \ldots - \frac{23}{384}\alpha^{4} - \frac{1}{24}\alpha^{3} + \frac{5}{48}\alpha^{2} - \frac{1}{12}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,0,-4) = \frac{1}{256}\alpha^{11} + \frac{13}{384}\alpha^{10} - \frac{31}{384}\alpha^{9} + \frac{11}{192}\alpha^{8} + \frac{21}{256}\alpha^{7} - \frac{53}{192}\alpha^{6} + \frac{47}{128}\alpha^{5} - \frac{5}{32}\alpha^{4} \ldots - \frac{7}{96}\alpha^{3} + \frac{5}{24}\alpha^{2} - \frac{1}{6}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-2,4) = -\frac{3}{128}\alpha^{11} + \frac{17}{256}\alpha^{10} - \frac{19}{256}\alpha^{9} + \frac{1}{32}\alpha^{8} + \frac{5}{256}\alpha^{7} - \frac{7}{256}\alpha^{6} + \frac{1}{128}\alpha^{5} + \frac{1}{2}$$

$$p_{\uparrow}(0,-2,2) = -\frac{3}{1024}\alpha^{11} + \frac{1}{256}\alpha^{10} + \frac{11}{512}\alpha^{9} - \frac{27}{512}\alpha^{8} + \frac{83}{1024}\alpha^{7} - \frac{47}{512}\alpha^{6} + \frac{39}{512}\alpha^{5} \ldots - \frac{3}{256}\alpha^{4} - \frac{1}{128}\alpha^{3} + \frac{3}{64}\alpha^{2} - \frac{1}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-2,0) = \frac{5}{512}\alpha^{11} - \frac{11}{1536}\alpha^{10} + \frac{1}{384}\alpha^{9} - \frac{1}{64}\alpha^{8} + \frac{41}{512}\alpha^{7} - \frac{83}{512}\alpha^{6} + \frac{47}{256}\alpha^{5} - \frac{19}{384}\alpha^{4} \ldots - \frac{1}{96}\alpha^{3} + \frac{3}{32}\alpha^{2} - \frac{1}{8}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-2,-2) = \frac{9}{1024}\alpha^{11} + \frac{3}{256}\alpha^{10} - \frac{19}{512}\alpha^{9} + \frac{19}{512}\alpha^{8} + \frac{59}{1024}\alpha^{7} - \frac{113}{512}\alpha^{6} + \frac{151}{512}\alpha^{5} \ldots - \frac{25}{256}\alpha^{4} - \frac{1}{128}\alpha^{3} + \frac{9}{64}\alpha^{2} - \frac{3}{16}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-2,-4) = -\frac{3}{256}\alpha^{11} + \frac{5}{128}\alpha^{10} - \frac{1}{256}\alpha^{9} + \frac{7}{128}\alpha^{7} - \frac{1}{4}\alpha^{6} + \frac{3}{8}\alpha^{5} - \frac{9}{64}\alpha^{4} + \frac{3}{16}\alpha^{2} \ldots - \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-4,4) = \frac{3}{128}\alpha^{10} - \frac{5}{64}\alpha^{9} + \frac{1}{8}\alpha^{8} + \frac{3}{128}\alpha^{7} - \frac{13}{32}\alpha^{6} + \frac{45}{64}\alpha^{5} - \frac{25}{64}\alpha^{4} - \frac{3}{16}\alpha^{3} + \frac{7}{16}\alpha^{2} \ldots - \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-4,2) = \frac{3}{256}\alpha^{11} - \frac{1}{256}\alpha^{10} - \frac{11}{256}\alpha^{9} + \frac{3}{32}\alpha^{8} + \frac{9}{256}\alpha^{7} - \frac{43}{128}\alpha^{6} + \frac{69}{128}\alpha^{5} - \frac{17}{64}\alpha^{4}\ldots \\ - \frac{3}{32}\alpha^{3} + \frac{5}{16}\alpha^{2} - \frac{1}{4}\alpha + \frac{1}{2} \tag{A.1}$$

$$p_{\uparrow}(0,-4,0) = \frac{3}{256}\alpha^{11} - \frac{11}{384}\alpha^{9} + \frac{3}{64}\alpha^{8} + \frac{59}{768}\alpha^{7} - \frac{53}{192}\alpha^{6} + \frac{47}{128}\alpha^{5} - \frac{13}{96}\alpha^{4} + \frac{3}{16}\alpha^{2}\ldots \\ - \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-4,-2) = \frac{3}{256}\alpha^{10} - \frac{1}{256}\alpha^{9} + \frac{1}{64}\alpha^{8} + \frac{13}{128}\alpha^{7} - \frac{15}{64}\alpha^{6} + \frac{13}{64}\alpha^{5} + \frac{3}{32}\alpha^{3} + \frac{1}{16}\alpha^{2}\ldots \\ - \frac{1}{4}\alpha + \frac{1}{2}$$

$$p_{\uparrow}(0,-4,-4) = -\frac{3}{128}\alpha^{11} + \frac{1}{128}\alpha^{10} + \frac{1}{16}\alpha^{9} + \frac{1}{32}\alpha^{8} + \frac{1}{16}\alpha^{7} - \frac{7}{32}\alpha^{6} + \frac{1}{16}\alpha^{5} + \frac{9}{64}\alpha^{4}\ldots \\ + \frac{3}{16}\alpha^{3} - \frac{1}{16}\alpha^{2} - \frac{1}{4}\alpha + \frac{1}{2}$$

Appendix B

Connections Made	Connections Not Made	No progression	Stops in C_1	Stops in C_2	Reaches C_2	Reaches C_3
0	27	1	0	0	0	0
1	26	21	6	0	0	0
2	25	210	131	10	10	0
3	24	1330	1329	251	266	15
4	23	5985	8303	2824	3262	438
5	22	20349	35868	18641	24513	5872
6	21	54264	114194	80178	127552	47374
7	20	116280	278893	238409	492857	254448
8	19	203490	536728	515833	1479857	964024
9	18	293930	829133	847866	3563762	2715896
10	17	352716	1041164	1093966	7042405	5948439
11	16	352716	1071323	1133267	11613856	10480589
12	15	293930	906978	955578	16182952	15227374
13	14	203490	631949	660000	19222861	18562861
14	13	116280	361062	373322	19580958	19207636
15	12	54264	167759	171820	17161837	16990017
16	11	20349	62488	63486	12955058	12891572
17	10	5985	18241	18414	8412059	8393645
18	9	1330	4025	4044	4681470	4677426
19	8	210	632	633	2219233	2218600
20	7	21	63	63	887946	887883
21	6	1	3	3	296006	296003
22	5	0	0	0	80730	80730
23	4	0	0	0	17550	17550
24	3	0	0	0	2925	2925
25	2	0	0	0	351	351
26	1	0	0	0	27	27
27	0	0	0	0	1	1

Table B.1 Table showing the counts for the final wave condition for the 8^{27} different configurations of site state connectivity.