MAJORITY VOTER MODEL FOR INFORMATION DIFFUSION

by

Rebecca Ryan

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2019

Contents

List of Figures

Abstract

A Majority Voter Model is a iterative process on graphs. Let G be a graph with a initial vertex colouring of n colours with the option of a vertex being uncoloured. The is a sequential process where once a vertex is coloured, it can no longer become uncoloured, and at each time step a vertex adopts the colour that occurs most frequently in its neighbourhood. We study two models that approach tie-breaking in a different way. In the event of a tie in the Conservative Majority Model, a vertex will conserve its current colour, and in the event of a tie in the Mixed Majority Model, a vertex will either conserve its colour if its current colour is in the majority or when its current colour is not present in the majority, the vertex will adopt the most preferred colour in the majority.

We classify the periodic configurations of the Conservative Model on paths, cycles, and toroidal grids. We also study the behaviour of uncoloured vertices. We introduce coalitions of colours that form under the model and determine some properties. We show that the removal of these coalitions does not affect the period of the conservative model. Additionally, when the initial configuration is random, we examine some threshold probabilities that ensure the survival of a colour on complete graphs and cycles. We find the period of both models and the length of their pre-period.

Acknowledgements

I would like to express my very great appreciation to my supervisor Dr. Jeannette Janssen. She has been incredibly kind, helpful, and patient throughout the writing process. This accomplishment would not have been possible without her guidance.

I would also like to thank my committee Dr. Jason Brown and Dr. Richard Nowakowski for taking the time to read this thesis and providing helpful feedback.

Lastly, I am so grateful for all the support my friends and family have given me.

Chapter 1

Introduction

Understanding how information spreads throughout networks is an important problem that is gaining popularity. This knowledge can be applied to a variety of fields such as economics, physics, sociology, and biology [9]. More commonly, with the rise of social networks, these studies are becoming increasingly important especially for tracking down viruses, or marketing [15] [10] [5] [22]. Such processes are referred to as dynamic processes in that the state of the information can change over time. From a political science perspective, these models also exhibit similar behaviour to voting [16] [8]. In biology these models can also help us understand cellular and evolutionary biology, epidemiology, and gene regulatory networks as seen in [3], [20], and [12].

Some of these dynamic processes are also important for the field of computer science. A popular topic in computing is automata. Automata can help with information processing, and are sometimes equivalent to Turing Machines which checks the limitations of mechanical computation of algorithms [19]. In some cases these processes can be self-organizing.

The focus of this thesis will be a majority rules process which is a model for social influence within a network. Each member of the network will adopt the opinion that occurs most frequently in their social neighbourhood. The hope is to be able to better understand this sort of information spread. In researching this model, as in other dynamic models, the objective is to discover the long term behaviour based on the initial state of the process.

1.1 Previous and Related Works

Many different dynamic processes have been previously studied. There exist both deterministic models and stochastic models. One example of a stochastic process is a random voter model in which there are voters in a network, and a voter assumes the opinion of one of its neighbours at random. Another example is a linear threshold model which has influence weights on each edge that change at random after each time step, and has two states: active and inactive [4]. A generalized linear threshold model with n states was studied in [14]. This is a linear threshold model where each node has an activation function. A multiple threshold model was studied in [22]. This model has multiple states for the vertices, and terminates when no new activation continues [22]. Meanwhile deterministic models have no element of probability when adopting a new state.

Other popular dynamic models are variations on the *prisoner's dilemma game*, which is also a model of social influence. This was originally a game used in classical game theory to study how humans cooperate with each other to best suit their own needs. Mathematicians started using the game to form a set of rules to create a deterministic process on graphs as seen in [7], [6], [21], and [11]. The model gives each node an option between cooperator and defector with an unbalanced scoring scheme. This means that being a cooperator will have a different score than being a defector. A node then looks to its neighbours and looks at the pay-off for either choice. A node makes a decision of either cooperator or defector in order to get the highest pay-off. In [6], an asynchronous updating model was studied. In [7] the model was studied on a random graph with different levels of connectivity. This sort of model can be used to study an evolutionary system as seen in [13].

Bootstrap percolation is another model that simulates the spread of information. The model starts on a graph with an initial set of active vertices. The model has a threshold r such that a node only becomes active if at least r of its neighbours are active. Once a node is activated, it remains activated. As stated in [9], this model can exhibit behaviours similar to rumour spread. Bootstrap percolation is a cellular automaton, and can also be a majority voter if r is large enough. In [9] they studied the model on the random graph $G(n, p)$ with results on the size of the final activation set, and the number of steps until the process stops.

Clearly there are many different models in existence, all attempting to model the behaviour of people based on social influence. Many of these papers sought out to find the best initial active set to achieve the most desirable final active set.

These sorts of updating models are quite popular in the field of economics as they help model consumer behaviour and aid in profit maximization. Numerous economic models as seen in [15], [10], [5],[23], [24],[22] provided many experimental results; however, there are few theoretical results. In [23] they used a cascade model to show the monotonicity of influence propagation for viral marketing and influence maximization. In [5] computational approaches were used to show product propagation through small-world networks. They studied a model in which products gain value based on how many adopters it has. It was discovered that a product is more likely to fail at becoming popular in random networks and more likely to thrive in highly clustered networks.

In [15] they created a model which looks at n states and considered product diffusion on small-world networks. The model is a majority voter model with random tie-breaking. Beforehand, most models only considered two states. They created two scenarios for the model, one involving perfect information, the other imperfect information. In perfect information all nodes are informed of the existence of each product. In imperfect information not every node is informed of each product and only discovers a new product through interaction with other nodes. Through many simulations they found that the perfect information model encourages faster diffusion in completely random networks, and the imperfect information model encouraged a winner-take-all scenario.

Others have looked at the influence certain communities of nodes have on the overall outcome of the final periodic state of the model as seen in [16]. This paper considers nodes that have formed coalitions, or communities that control the entire graph.

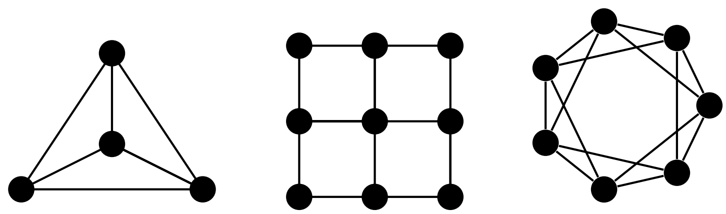
Perhaps the most influential paper for this thesis is [8]. They studied a biased majority model and a conservative majority model with two colours. They explored the models on cycles and toroidal grids. They showed that the period is 2 on both models with two colours, and found an upper bound of the pre-period of both models. They found a threshold probability that ensures a bichromatic configuration at the end of the process on cycles. Moreover, they showed that there are two threshold probabilities so that with high probability the final configuration will be monochromatic or bichromatic on the toroidal grid. From this paper we expanded upon the number of colours and added the possibility of being uncoloured.

1.2 Definitions

1.2.1 Graph Theory

We follow the definitions as given in [2]. A **graph** $G = (V, E)$ is a set of vertices V and edges E with a function ψ with associates each edge with an unordered pair of vertices in V. If we have $\psi(e) = uv$, then u and v the endpoints of e. We can also write the edge e as $e = uv$. We say the **order** of a graph is the number of vertices $|V|$ and the **size** of a graph is the number of edges $|E|$. A **loop** is a edge where the endpoints are identical. A graph is said to have **parallel edges** if there are at least two edges in E that have the same endpoints. A **simple graph** is a graph with no loops or parallel edges. A **multigraph** is a graph where loops and parallel edges are allowed. In this thesis unless otherwise stated all graphs are simple. We say two vertices u and v are **adjacent** if there exists an edge $e = uv$. We say an edge is **incident** to vertices u and v if u and v are the endpoints of the edge. The

Figure 1.1: Examples of families of graphs



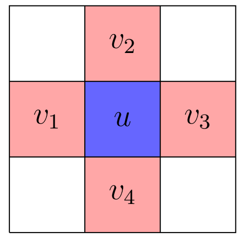
A complete graph A cartesian grid A 4-regular graph

neighbourhood of a vertex u, denoted $N(u)$, is the set of all vertices adjacent to u. The **degree** of vertex denoted $deg(v)$ is $|N(u)|$. A **cartesian product** of two graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$, denoted $G \Box H$, is the graph with vertex set $V(G) \times V(H)$ such that either $u_1u_2 \in E(G)$ and $v_1 = v_2$, or $v_1v_2 \in E(H)$ and $u_1 = u_2$. A **maximum matching** is a set of pairwise non-adjacent edges which covers as many vertices as possible. A **clique** is a set of mutually adjacent vertices.

The following are some families of graphs. A **complete graph** K_n is a graph of order n where any two vertices are adjacent. A **path** P_n is a graph of order n where the vertices can be listed sequentially where two vertices are adjacent if they are consecutive in the sequence, and are nonadjacent otherwise. A **cycle** C_n is a graph whose vertices are arranged in a cyclic sequence such that two vertices are adjacent if they are consecutive in the sequence and nonadjacent otherwise. A **cartesian grid** $G_{n,m}$ is the cartesian product of two paths $P_n \Box P_m$. A **toroidal** graph is a graph that can be placed on a torus such that no two edges intersect. A **toroidal grid** graph $T_{n,m}$ is the cartesian product of two cycles $C_n \Box C_m$. A **tree** is a graph which contains no cycles. A graph is **k-regular** if $d(v) = k$ for all $v \in V$. See Figures 1.1 for examples of some graphs. An **induced subgraph** of a graph $G = (V, E)$ is a set of vertices $X \subset V$ whose set of edges consists of all edges that have both endpoints in X.

Moreover, graph theory also has a connection to linear algebra. We can

Figure 1.2: von Neumann Neighbourhood



represent any graph as a matrix, an **adjacency matrix** of the graph. For a graph of order n, an adjacency matrix is a matrix A of size $n \times n$ where each entry (a_{uv}) is the number of edges joining vertex u to vertex v. For simple graphs, the entries a_{uv} will only either be a 1 or a 0. In the matrix, each row and column represent the neighbours of a vertex. These matrices are always symmetric. With this connection, it is now possible to use tools from linear algebra to discover facts about graphs.

Graphs can also be used to model finite deterministic automata. We will use the definition as defined in [19]. Let I be a finite alphabet, where an alphabet is a non-empty set where the elements are called letters. A **finite deterministic automata** over the set I is an ordered triple $A = (S, s_0, f)$, where S is a finite non-empty set of states, $s_0 \in S$ is the initial state, and f is a function mapping the Cartesian product $S \times I$ into the set S.

In graph theory we most commonly find automata in the form of cellular automata. This is an automaton that takes place on a Cartesian or toroidal grid. One of the more common types of neighbourhoods defined in cellular automata is the **von Neumann neighbourhood**. The von Neumann neighbourhood of a cell u consists of four neighbours orthogonally adjacent to u. We can see the von Neumann neighbourhood in Figure 1.2 where u is the blue cell, and its neighbourhood consists of the red cells labelled v_1, v_2, v_3, and v_4. There also exists one dimensional cellular automata where a cell u has two neighbours.

1.2.2 Probability Theory

To prove the existence of an event happening we use probability theory. We use the definitions as defined in [18]. First and foremost we look at the **probability space**. This contains three elements: the sample space Ω, which is the set of all possible outcomes, $\mathcal{F}$ which is the σ-field, a certain collection of subsets of Ω, and $\mathbb{P}$ the probability measure which is a function from $\mathcal{F}$ to $[0, 1]$. A simple example would be tossing a fair coin. Here, $\Omega = \{H, T\}$ and our σ-field is $\{\emptyset, H, T, \Omega\}$. Our probability measure is $P(\emptyset) = 0$, $P(H) = \frac{1}{2}$, $P(T) = \frac{1}{2}$, and $P(\Omega) = 1$. A **random variable** is a function X that maps elements from Ω to the real numbers $\mathbb{R}$. An **event** is a set of outcomes from Ω to which a probability is assigned. Let N be an event and let the random variable X_N be defined as:

$$X_N = \begin{cases} 1 & \text{if N occurs,} \\ 0 & \text{otherwise} \end{cases} \tag{1.1}$$

We call X_N the **indicator variable** of the event N. A **Bernoulli distribution** is a distribution of two outcomes, 0 or 1 (”failure” or ”success”), where the probability of a success is p and the probability of a failure is $1 - p$ for some $p \in [0, 1]$. A random variable with a Bernoulli distribution is called a **Bernoulli random variable**.

The **expected value** for a random variable X is denoted $E(X) = \sum_k P(X = k) \cdot k$. **Linearity of expectation** states that if we have $E(X)$ and $E(Y)$, then $E(X + Y)$ exists and $E(X + Y) = E(X) + E(Y)$. The **variance** of a random variable X is defined as $Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$. If we have two random variables X and Y then we can compute the **covariance**, $Cov(X, Y) = E(XY) - E(X)E(Y)$. If X and Y are independent then $Cov(X, Y) = 0$. If a random variable $X = \sum_{i=1}^{n} X_i$, then $Var(X) = \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n} \sum_{i \neq j} Cov(X_i, X_j)$. We say an event N happens **asymptotically almost surely** (a.a.s) if $Pr(N) \to 1$ as $n \to \infty$.

Theorem 1.1 (Markov's Inequality) *Let X is a positive random variable with a* *finite expected value* $E(X)$*. Then for every* $a > 0$

$$Pr(X \geq a) \leq \frac{E(X)}{a}.$$

Theorem 1.2 (Chebyshev's Inequality) *Let* X *be any random variable with finite expected value* $E(X)$*. Then if* $a > 0$ *is a real number, we have*

$$Pr(|X - E(X)| \geq a) \leq \frac{Var(X)}{a^2}.$$

Theorem 1.3 (Hoeffding's Inequality) *[1] Let* $X_1, X_2, ..., X_n$ *be a finite sequence of independent random variables. Assume that for all* $1 \leq k \leq n$*, one can find two constants* $a_k < b_k$ *such that* $a_k \leq X_k \leq b_k$ *almost surely. Denote* $S_n = X_1 + ... + X_n$*, then for any positive* x*,*

$$P(|S_n - E(S_n)| \geq x) \leq 2\exp\left(\frac{-2x^2}{\sum_{k=1}^{n}(b_k - a_k)^2}\right) \tag{1.2}$$

If the X_i are each indicator variables, then we can take $a_i = 0$ and $b_i = 1$, and Hoeffding's bound becomes $P(|S_n - E(S_n)| \geq x) \leq 2\exp(\frac{-2x^2}{n})$.

Also important in probability theory is understanding some basic asymptotic analysis. Let f and g be real valued functions.

- We say $f(x) = O(g(x))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} < \infty$.
- We say $f(x) = o(g(x))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.
- We say $f(x) = \Omega(g(x))$ if $\lim_{n\to\infty} \frac{g(n)}{f(n)} > 0$.
- We say $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$.
- We say f is much less than g, or $f \ll g$, if $\frac{f(n)}{g(n)} \to 0$ as $n \to \infty$.

1.3 The Model

In this section we describe four possible variants of the majority voter model. In each variant there is a different tie-breaking rule. These models originate from [8] and [15]. For each model we have a set of colours percolating through a graph. Each vertex is initially assigned a colour, and then changes its state based on the colour that occurs most often in the neighbourhood of the vertex. For each of these models, we will include an uncoloured state, which represents undecided. The vertices that are uncoloured will not have an effect on their neighbours. All four variants of the model depend on time. At time $t = 0$, a graph will be giving a random configuration of colours, and each vertex will update their colour for time $t = 1$. The process is then continued for $t \geq 0$.

Let $f : V \to \{0, 1, ..., n\}$ be a function that gives a configuration on a graph $G = (V, E)$. Let the set of vertices in the neighbourhood of v that have colour c at time t be denoted $N_c^t(v)$.

Let $v \in V$, then the majority of v is

$$\text{Maj}(f, v) = \text{argmax}\{|N_c^f(v)|\big|1 \leq c \leq n\}. \tag{1.3}$$

Thus, $\text{Maj}(f, v)$ is the colour that occurs the most often in the neighbourhood of v. A tie occurs at a vertex v if there exists at least 2 colours $c_1 \neq c_2$ such that

$$|N_{c_1}^t(v)| = |N_{c_2}^t(v)|, \tag{1.4}$$

and for all c, $1 \leq c \leq n$,

$$|N_c^t(v)| \leq |N_{c_2}^t(v)|. \tag{1.5}$$

This also includes the case where all neighbours of v are uncoloured since $|N_c^t(v)| = 0$ for all c. We let $\text{Maj}(f, v)$ be a colour, and $\{Maj(f, v)\}$ to be the set of all colours that occur most often in $N(v)$.

We will now define the variants of the majority voter model. Each colour is associated with an integer, and if a vertex v is uncoloured, then $f(v) = 0$.

Definition 1.1 (Conservative Majority Model) *In this model, vertices where a tie occurs retain their original colour. Precisely, for all $v \in V$, and all $t \geq 1$,*

$$f_t(v) = \begin{cases} f_{t-1}(v) & \textit{if a tie occurs at } v \\ Maj(f_{t-1}, v) & \textit{otherwise} \end{cases}$$

The Conservative model comes from [8].

Definition 1.2 (Biased Majority Model) *In this model, vertices where a tie occurs choose the best ranked colour. Assume that the colours $\{1, ..., c\}$ are ranked where colour 1 is the most desirable, colour 2 is second most desirable, and so on. Thus, in the event of a tie, the vertex will always choose the most desired colour in its neighbourhood. For all $v \in V$, and all $t \geq 1$,*

$$f_t(v) = \begin{cases} f_{t-1}(v) & \textit{if } \forall u \in N(v), f_{t-1}(u) = 0 \\ min\{Maj(f_{t-1}, v)\} & \textit{if there is a tie} \\ Maj(f_{t-1}, v) & \textit{otherwise} \end{cases}$$

The biased majority model comes from [8].

The following model uses elements from both the Conservative Majority Model and the Biased Majority Model.

Definition 1.3 (Mixed Majority Model) *Assume that the colours $\{1, ..., n\}$ are ranked where colour 1 is the most desirable, colour 2 is second most desirable, and so on. If the vertex's adopted colour is in the tie, then it will conserve its colour. In the event of a tie and the vertex's current colour is not in the tie, then the vertex will* *always choose the most desired colour in its neighbourhood. For all $v \in V$, and all $t \geq 1$,*

$$f_t(v) = \begin{cases} f_{t-1}(v) & \textit{if } \forall u \in N(v), f_{t-1}(u) = 0 \\ f_{t-1}(v) & \textit{if there is a tie and } f_{t-1}(v) \in Maj(f_{t-1}, v) \\ min\{Maj(f_{t-1}, v)\} & \textit{if there is a tie} \\ Maj(f_{t-1}, v) & \textit{otherwise} \end{cases}$$

For the next model we define the function $Rand(X)$. This function takes in a set X of elements, and outputs a random element from X.

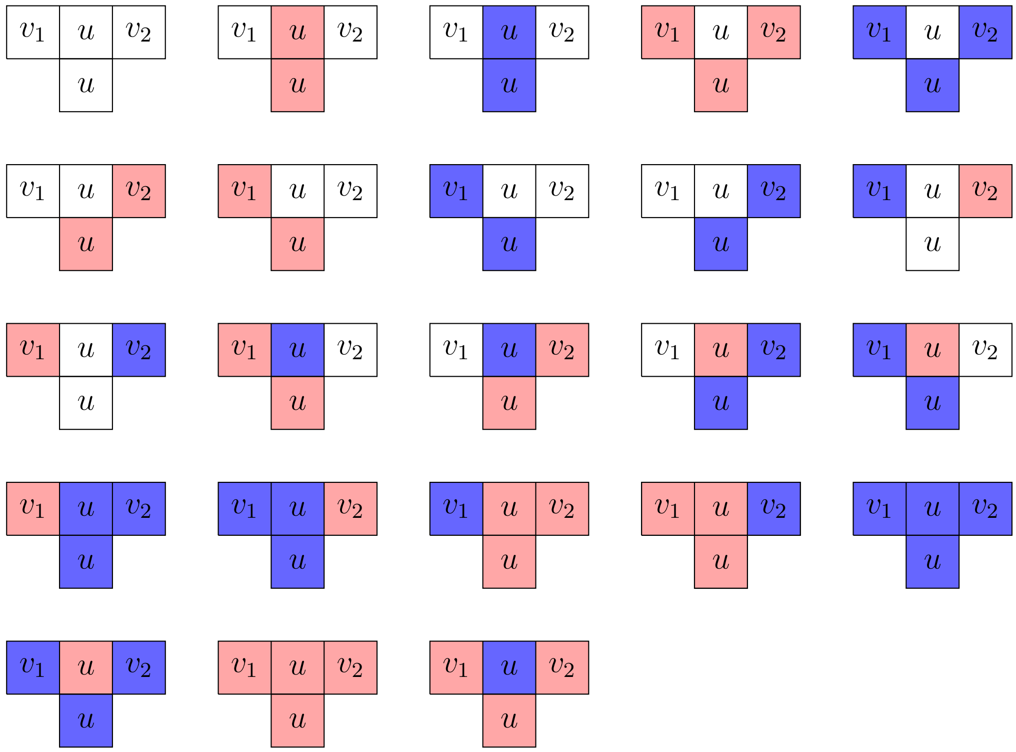
Definition 1.4 (Coin-Flip Majority Model) *This model has stochastic tie-breaking. In the event of a tie, a vertex will choose uniformly at random from the colours in $Maj(f_{t-1}, v)$.*

$$f_t(v) = \begin{cases} f_{t-1}(v) & \textit{if } \forall u \in N(v), f_{t-1}(u) = 0 \\ Rand(\{Maj(f_{t-1}, v)\}) & \textit{if there is a tie} \\ Maj(f_{t-1}, v) & \textit{otherwise} \end{cases}$$

The coin-flip model comes from [15].

A sequence $\{f_t\}_{t=0}^{\infty}$ is eventually periodic with period k if there exists some time t^* such that $f_{t^*}(v) = f_{t^*+k}(v)$ for all $t \geq t^*$. Let f be a configuration, and $\{f_t\}_{t=0}^{\infty}$ be a configuration sequence so that $f_0 = f$. We say that f is periodic with period k if $f_t = f_{t+k}$ for all $t \geq 0$. We say the **least period** of a function f is k if k is the smallest integer so that $f_t = f_{t+k}$. In this thesis when we say period or periodic, we always refer to the least period k. The Conservative, Biased, and Mixed models are inherently periodic as they have a finite number of configurations and are deterministic. However, the coin-flip model is stochastic, and whenever ties occur it will not be necessarily periodic.

Figure 1.3: One Dimensional Conservative Majority Cellular Automaton



The Conservative Majority Model will be the main focus of subsequent chapters. This is a finite deterministic cellular automata. In the Conservative Majority Model the rules are based on the neighbouring cells having a majority colour. The rules are defined in Figure 1.3 for the case when the graph is a path, and the corresponding automata is one dimensional. The top row represents the vertex u and its two neighbours at time t, and the bottom row is u at time $t+1$.

1.4 Main Results

This thesis covers a wide range of areas for the majority voter model with a primary focus on the Conservative Majority Model.

In Chapter 2, we classify all the possible period 1 and 2 configurations of the Conservative Majority Model on paths, cycles with n colours and uncoloured. We also classify the period 1 and 2 configurations on the toroidal grid with 2 colours and no uncoloured. We look at how an uncoloured vertex behaves in a periodic configuration.

In Chapter 3, we define persistent sets, which are sets of vertices in a graph such that at least half of their neighbours are in the set. We look at some basic properties of the sets and monochromatic persistent sets. We look at combining persistent sets, and how the period of a configuration is affected by persistent sets.

In Chapter 4 we look at the cycle and complete graph and the Conservative Majority Model with 2 colours and n colours, including uncoloured. We determine the possible initial configurations that determine whether or not a colour persists to the final configuration. In addition, with a random initial configuration, we look at the long term behaviour of the model on these graphs and the threshold probabilities that ensure the survival of a colour.

In Chapter 5, we prove that the period of the Mixed Majority Model with n colours is 2 and find the upper bound of its pre-period. Similarly, we find the upper bound of the pre-period for 2 colours and uncoloured on the Conservative Majority Model and prove that the period is 2. These proofs involve using a linear algebra framework to represent the model.

Chapter 2

Periodic Configurations

In this chapter we look at periodic configurations to give an idea of what to expect in the final stabilized state of the Conservative Majority Model. This narrows all the possibilities for periodic configurations, as for any graph $G = (V, E)$ on $|V|$ vertices with n colours and uncoloured, there are $|V|^{n+1}$ possible configurations.

For the Conservative Majority Model, we know that it is inherently periodic as the model is deterministic and has a finite number of configurations. In Chapter 5, we prove that the Conservative Majority Model with two colours and uncoloured has period 2. It has not yet been proven that the model with n colours has period 2; however, given experimental results, we believe that it has period 2. We classify paths, cycles and toroidal grids for period 1 and period 2 configurations. For the sections on paths and cycles, we will be under the assumption that the process has n colours and uncoloured. For the sections on toroidal grids, we will be under the assumption that the process has 2 colours and no uncoloured. Recall, each colour is associated with an integer and we let uncoloured be represented by 0.

Let f be a configuration, and $\{f_t\}_{t=0}^{\infty}$ be a configuration sequence so that $f_0 = f$. Recall that f is periodic with period k if $f_t = f_{t+k}$ for all $t \geq 0$. We let k be the least period. A vertex v in a configuration f has period k if $f_t(v) = f_{t+k}(v)$ for all $t \geq 0$. Given a graph $G = (V, E)$, a configuration f, and a vertex $v \in V$, we define the configuration f^{-v} on the graph $G - v$ as the restriction of f to $V - \{v\}$.

Lemma 2.1 *Let f be a configuration on a graph $G = (V, E)$, and let $v \in V$ be such that $f(v) = 0$, and let f^* be the configuration obtained by applying one step of the*

Conservative Majority Model. Let f^{-v} be the corresponding configuration on $G-v$, and f^{*-v} the configuration obtained by applying one step of the Conservative Majority Model to $G-v$ with configuration f^{-v}. Then, for all $u \in V-v$, $f^*(u) = f^{*-v}(u)$.

Proof: Since uncoloured vertices are not considered then $\text{Maj}(f,u)=\text{Maj}(g,u)$ for all $u \in V - \{v\}$. Moreover, a tie occurs at $u \in V - \{v\}$ in G if and only if a tie occurs at u in $G-v$. □

Theorem 2.1 *Suppose f is a periodic configuration for a graph $G = (V, E)$, and there exists some $v \in V$ so that $f(v) = 0$. Then f^{-v} is periodic for $G - v$.*

Proof: Let k be the period of f. Let $\{f_t\}_{t=0}^{\infty}$ be the configuration sequence so that $f_0 = f$. Configurations $f_0, f_1, ..., f_{k-1}$ represent all the configurations in the period. Vertices never can become uncoloured; therefore, during a period uncoloured vertices cannot become coloured. And so $f_i(v) = 0$ for $0 \le i \le k$. Lemma 2.1 implies that for $u \neq v$, $f_{i+1}(u) = f_{i+1}^{-v}(u)$ for all $0 \le i \le k$ and thus,

$$f_k^{-v}(u) = f_k(u) = f_0(u) = f_0^{-v}(u). \tag{2.1}$$

Therefore, f^{-v} is periodic with period k. □

In the following sections we will classify all period 1 and 2 configurations on paths, cycles, and toroidal grids.

2.1 Path and Cycle Configurations

In this section, we will classify the periodic configurations on paths, which can be useful in studying one dimensional cellular automata. Then, using the configurations of paths, we will classify all the configurations on cycles.

2.1.1 Path Configurations of Period 1

On a path we will define certain components of a period 1 configuration for each colour. We will number the vertices of a path P_n as $v_0, v_1, ..., v_{n-1}$. The **end vertices** of a path are v_0 and v_{n-1}. We first define the components for each colour, then the components for uncoloured.

A **separator vertex** is a vertex v_i, $0 < i < n-1$ so that v_{i-1}, v_i, v_{i+1} have three different colours, and v_{i-1}, v_i, v_{i+1} are not uncoloured and not end vertices. A **monochromatic path** is a subpath with vertices $v_i, v_{i+1}, ..., v_j$ so that $f(v_l) = c$ for $i \leq l \leq j$ and $f(v_{i-1}) \neq c$ or $i = 0$, and $f(v_{j+1}) \neq c$ or $j = n-1$. An **island vertex** is a vertex so that all neighbours are uncoloured. An **uncoloured separator vertex** is a single vertex v_i so that $f(v_i) = 0$ and $i \neq 0$, and v_{i-1} and v_{i+1} are different colours. See Figure 2.1 to see examples of monochromatic paths, island vertices, and uncoloured separator vertices.

Theorem 2.2 *A configuration f on a path for the Conservative Majority Model with n colours and uncoloured is period 1 if and only if for each colour $c \in \{1, ..., n\}$, the graph induced by colour c consists of monochromatic paths, separator vertices, or island vertices.*

Proof: Let f be a period 1 configuration on the path. Consider a subgraph induced by the colour c. Each component of the subgraph is either a single vertex or a path.

Suppose $v_i, ..., v_j$, $j - i \geq 2$ is a component. Then by definition $v_i, ..., v_j$ is a monochromatic path, and $f(v_{i-1}) \neq c$ and $f(v_{j+1}) \neq c$.

Next, suppose v_i is a component that is a vertex. Suppose first that v_i has an uncoloured neighbour. Then its other neighbour must also be uncoloured, otherwise v_i will change colour. Alternatively, v_i is an end vertex, with only one neighbour that is uncoloured, and will not change. If both neighbours are uncoloured, or v_i is an end vertex then v_i is an island vertex.

Assume that v_i has no uncoloured neighbour. If v_i is an end vertex, then its one neighbour must be the same colour as v_i, and it is part of a monochromatic path. Suppose that v_i is not an end vertex and has no uncoloured neighbour, then there must be a tie in its neighbourhood. Thus, v_i has two neighbours of different colours. Therefore, v_i is a separator vertex by definition. Similarly if there is an uncoloured vertex v, then it must be that there is a tie in its neighbourhood, making v_i an uncoloured separator vertex.

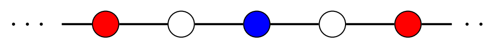
Suppose that f is not a period 1 configuration. Then there must be a vertex that changes colour. That is, there exists a single vertex v coloured c_1 with a strict majority in its neighbourhood. If v is an end vertex either coloured or uncoloured, then its neighbour must be another colour, and v changes. If there exists an path of uncoloured vertices, then one vertexin the path has a coloured neighbour and will change. If v is not an end vertex and is coloured, then both neighbours are a different colour giving a strict majority. If v is uncoloured, and not an uncoloured separator vertex, then it has a majority and changes colour.

As a vertex needs a strict majority to change colour, no vertex in a monochromatic path will change colour. Thus, v is a component that is one vertex. However, v is not a separator vertex or uncoloured separator vertex since it does not have two neighbours with different colours, and it is not an island vertex because it has one coloured neighbour. Therefore, the graph induced by c_1 has a component that is not a monochromatic path, separator vertex, uncoloured separator vertex or island vertex. □

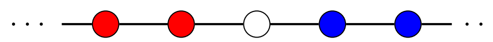
2.1.2 Cycle Configurations of Period 1

Now we classify the period 1 configurations on the cycle. The results from the classification from the path follow for the cycle; however, we add one more definition.

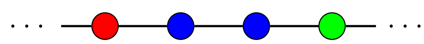
Figure 2.1: Configurations of period 1 on a path



An island vertex



A uncoloured separator vertex



A monochromatic path

We define a **monochromatic cycle** in a cycle C_n as $v_0, ..., v_{n-1}$ so that $f(v_i) = c$ for $0 \leq i \leq v_{n-1}$. We have the following corollary from Theorem 2.2.

Corollary 2.1 *Let f be a configuration on the cycle C_n for the Conservative Majority Mode, with n colours and uncoloured. The configuration f is period 1 if and only if for each colour $c \in \{1, ..., n\}$ and cycle $C = (V, E)$ f the graph induced by colour c consists of monochromatic paths, a monochromatic cycle, separator vertices, or island vertices.*

Proof: In the cycle there are no end vertices. However, in a configuration on a cycle it is possible to have monochromatic paths, separator vertices, and island vertices. The only other possible configuration is if the entire graph is one colour, in which case it is a monochromatic cycle and no vertex will change colour. Therefore, this result follows from Theorem 2.2. □

2.1.3 Path Configurations of Period 2

We classify period 2 configurations on the path. These patterns are more involved than period 1 configurations as period 2 configurations may have period 1 vertices as well as period 2 vertices.

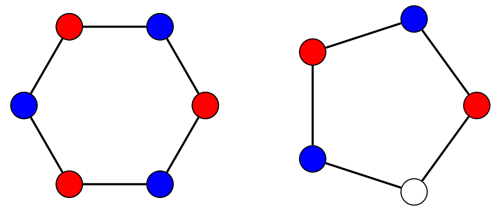
As before, in a period 2 configuration on a path, there may exist induced monochromatic paths, separator vertices, and island vertices. We use the definitions as introduced in 2.1.1. A **bichromatic path** is a path $v_i, ..., v_j$, where $f(v_i) = c_1 = f(v_{i+2}) = ... = c_2$, and $f(v_{i+1}) = f(v_{i+3}) = ... = c_2$ such that v_{i-1}, v_{j+1} are uncoloured separator vertices, or $i - 1 = 0$ or $j + 1 = n - 1$.

Theorem 2.3 *A periodic configuration f with n colours and uncoloured on a path has period 2 and not period 1 if and only if there exists a bichromatic path for some two colours c_i,$c_j \in \{c_1, ..., c_n\}$. Moreover, the graph induced by every pair of colours c_i and c_j consists of monochromatic/bichromatic paths, separator vertices, or island vertices.*

Proof: Suppose f does not have period 1. Then by Theorem 2.2 there is a colour c_i so that some component of the graph is not a monochromatic path, separator vertex, or island vertex. Thus, it must be a component that is a vertex v with 2 neighbours of the same colour c_j, $j \neq i$, v is an end vertex adjacent to a vertex of colour c_j, or v has an uncoloured neighbour and a neighbour of colour c_j. As f has period 2, then v will change colour, and so v has a majority of c_j, and becomes c_j in the next step. However, in order to become c_i again, it must have a majority of c_i again. This means that the neighbours change colour as well. And so, either each neighbour either has an uncoloured neighbour, another neighbour that is also c_i, or is an end vertex. Thus, this must be a bichromatic path by definition.

Now suppose that for some two colours c_i and c_j there is a bichromatic path. Then consider the graph induced by colour c_i. Each vertex in the bichromatic path is a component in the graph induced by c_i. Now, each vertex either has two neighbours coloured c_j, or one uncoloured neighbour, and is not a separator or island vertex. Therefore, by Theorem 2.2, this is not a period 1 configuration. Consider vertices in a bichromatic path. The end vertices v_k, v_l have one uncoloured neighbour, and one neighbour of the other colour. This gives a majority and so v_k, v_l will change to the

Figure 2.2: Examples of period 2 configurations



Bichromatic cycle Bichromatic path

other colour. All the vertices coloured c_j have two neighbours coloured c_i and the vertices coloured c_i have two neighbours coloured c_j. Thus, in the next step, every vertex coloured c_i will turn c_j and every vertex coloured c_j will turn c_i. If we apply the Conservative Majority process another step, then the bichromatic graph will have its original colouring. Therefore, this is a period 2 configuration. □

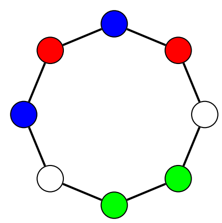
2.1.4 Cycle Configurations of Period 2

In this section we classify the possible period 2 configurations on the cycle. As before, all of the properties of path configurations also follow for cycles. See Figures 2.2 and, 2.3 for examples of period 2 configurations on the cycle. We define the **bichromatic cycle** on C_n for n even as $v_0, ..., v_{n-1}$, where $f(v_0) = c_1 = f(v_2) = ... = c_2$, and $f(v_1) = f(v_3) = ... = c_2$. We have the following corollary of Theorem 2.3 for period 2 configurations on cycles.

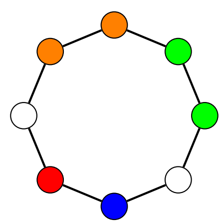
Theorem 2.4 *A periodic configuration f with n colours and uncoloured on the cycle has period 2 and not period 1 if and only if there exists a bichromatic path for some two colours c_i,$c_j \in \{c_1, ..., c_n\}$ or a bichromatic cycle. Moreover, the graph induced for every pair of colours c_i and c_j consists of monochromatic or bichromatic paths, a bichromatic cycle for n even, separator vertices, or island vertices.*

Proof: In the cycle there are no end vertices. However, in a configuration on a cycle it

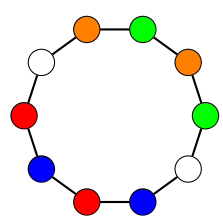
Figure 2.3: Configurations with both period 1 and 2 subgraphs on a cycle



Bichromatic and monochromatic paths on a cycle



Bichromatic and monochromatic paths on a cycle



Two bichromatic paths on a cycle

is possible to have monochromatic/ bichromatic paths, separator vertices, and island vertices. There are two other possible configurations. One configuration is if the entire graph is one colour, in which case it is a monochromatic cycle and is not period 2. If the cycle has an even number of vertices, then it is possible to have a bichromatic cycle in which case it is a period 2 configuration. Therefore this result follows from Theorem 2.3. $\square$

We have now classified all periodic configurations of paths and cycles. Next, we classify the periodic configurations on the toroidal grid.

2.2 Toroidal Grid Configurations

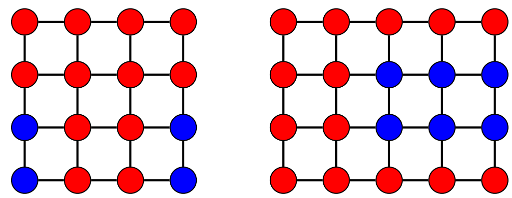
In this section we will look at how the Conservative Majority Model behaves on the toroidal grid. This can be helpful in further studies for two dimensional cellular automata, as we will know the final configurations. We classify all possible period 1 and 2 configurations for the toroidal grid with two colours and no uncoloured. We denote the graph by $T_{n,m}$.

Toroidal Grid Configurations of Period 1

Here we will classify the possible period 1 configurations on the toroidal grid. The toroidal grid is denoted $T_{n,m} = C_n \square C_m$, where we say this has width n and height m, and we will use the coordinate system to refer to a vertex on the graph. The height and width may be of any length. Each vertex will be labelled (x, y) where $1 \leq x \leq n$ and $1 \leq y \leq m$. We use colours r and b, and no uncoloured. Note that in the definitions below the roles of r and b may be reversed.

We define a **square** with four vertices, $(i, j), (i+1, j), (i, j+1), (i+1, j+1)$. We define a **polyomino region** as a connected subgraph that is a union of squares and paths such that both end vertices of the paths are in the union of squares. In

Figure 2.4: Period 1 Monochromatic Polyomino Region



such a region there exists no vertices with degree 1 in the subgraph. See Figure 2.4 for two examples of such a region. Note that the vertices in red represent the shape of the region.

We define a **monochromatic cycle** as an induced cycle with the property that each vertex is the same colour b. Each vertex in the cycle has two neighbours coloured b and so there will always be a majority b or a tie including b. See Figure 2.5 for an example of a monochromatic cycle. As we will see in the next chapter, this is an example of a persistent set.

We say a polyomino region is a **monochromatic polyomino region** if all vertices are the same colour b. Since each vertex in a polyomino region has at least two neighbours in a polyomino region, each vertex in a monochromatic polyomino region has a majority of b in their neighbourhood, and will retain its colour. Thus, all neighbours outside of the region are r. This would have the shape of a rectangle of width n and height m. However, these regions make take other shapes as seen in Figure 2.4.

Theorem 2.5 *A configuration f on the toroidal grid $T_{m,n}$ for the Conservative Majority Model with 2 colours and no uncoloured is period 1 if and only if for colours b and r, the graph induced by colour b or r consists of monochromatic polyomino regions, or cycles.*

Proof: Let f be a period 1 configuration. Consider a subgraph induced by the colour b. Let the other colour be r. Each component of the subgraph is either a path, cycle, single vertex, or some other subgraph.

Suppose there exists a component that is a single vertex (i, j), then all of its neighbours are r. And so, (i, j) will change colour and cannot be in a period 1 configuration. Suppose there is a component that forms an induced path. Then both end vertices will have a majority of r, and so it will change colour and cannot be in a period 1 configuration. Suppose now there is a component that forms an induced cycle, then by definition this is a monochromatic cycle, and is part of a period 1 configuration.

Suppose there is a connected component C with colour b such that it is not a path, cycle, or single vertex. Each vertex v in the component has at least two neighbours coloured b so that f is period 1. If these neighbours are and v part of a b-square, call this square S_v. Now, take the union of all S_v for each vertex v in C, call it R, and consider the component formed by $C - R$.

If a collection of disjoint paths remain, then it must be that at least one end vertex of each path is apart of R. Otherwise C would not be a connected component. Moreover, if only one end vertex is connected to R, then the other end vertex will have degree 1 in the component and must have a majority r, and will change colour. This is not period 1, so both end vertices must be connected to R. Similarly if a collection of non-disjoint paths remain, then for each path there exists two vertices with degree 1 in $C - R$. In order to not have a vertex of degree 1 in C, it must be that each degree 1 vertex is connected to R.

If a cycle that is not a square remains, then there must be a vertex v that has a neighbour in R. And so, v has three neighbours coloured b and must already be a part of R. Removing v gives a path in which both end vertices are a part of R. Every other vertex in the component has two neighbours coloured b, and so will not change colour. Thus, C is the union of squares S_v and paths where both end vertices are in C. Therefore, C must be a polyomino region and each induced subgraph in a period 1 configuration will either be a monochromatic polyomino region, or cycle.

Now suppose that f is not period 1. Then there must exist some vertex v with a strict majority in its neighbourhood so that it can change colour. That is, if v is coloured b, then it has either 3 or 4 neighbours coloured r. Thus, v cannot be a part of a monochromatic cycle, or polyomino region as there is not at least two neighbours coloured b. And so, if f is not period 1 there exists a component that is not a monochromatic cycle or polyomino region. □

2.2.1 Toroidal Configurations of Period 2

In this section we will classify all possible period 2 configurations on the toroidal grid. Again, period 2 configurations on the toroidal grid may have period 1 vertices as well as period 2 vertices. We use the colours b and r, and no uncoloured.

We say a polyomino region is a **bichromatic polyomino region** if each vertex with an odd y-coordinate and even x-coordinate will be coloured b, each vertex with an odd y-coordinate and odd x-coordinate will be coloured r, each vertex with an even y-coordinate and even x-coordinate will be coloured r, and each vertex with an even y-coordinate and an even x-coordinate will be coloured b. In addition, each vertex with two neighbours not in the subgraph, must have both neighbours coloured differently. Note that the colouring with r and b may be reversed.

We call an induced cycle a **bichromatic cycle** if $f((x_1, y_1)) = ... = f((x_{2k+1}, y_{2k+1})) = b$, $f((x_2, y_2)) = ... = f((x_{2k}, y_{2k})) = r$, and (x_1, y_1) is a neighbour of (x_p, y_p). For each coordinate (x_i, y_i), $1 \leq i, j \leq 2k$, the neighbours not in the cycle are coloured differently. See Figure 2.5 for an example of a bichromatic cycle.

Theorem 2.6 *A periodic configuration f on the Conservative Majority Model with 2 colours and no uncoloured on a toroidal grid has period 2 and not period 1 if and only if there exists at least one bichromatic polyomino region or bichromatic cycle.*

Moreover, the graph consists of monochromatic/bichromatic polyomino regions, and monochromatic/bichromatic cycles.

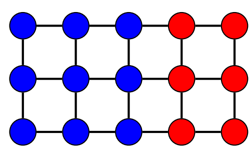
Proof: Suppose f does not have period 1, then there must exist some vertex with a strict majority colour in its neighbourhood. We will use colours b and r.

Consider the graph induced by all the period 2 vertices. If a period 2 vertex v has a majority of size 2 in the induced graph, then v is a part of a bichromatic cycle, or a path in a polyomino region. If v has a majority of size 3, then it must be a vertex in a polyomino region. Finally, if v has a majority of 4, then it is a vertex in a polyomino region. Note that in the entire graph, a vertex in a bichromatic cycle or a path in a bichromatic polyomino region will have a period 1 neighbour contributing to the majority. Every other vertex in the graph must have period 1, and by Theorem 2.5 this graph consists of monochromatic cycles, polyomino regions.

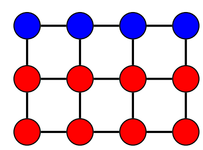
Suppose now that there exists period 2 vertices in f for colours b and r. Then the graph induced by b contains either a component that is a single vertex, and thus has a majority of size 4, or a vertex with a majority of 3. Then by Theorem 2.5 f does not have period 1.

Thus, all bichromatic components consist of bichromatic polyomino regions, and cycles, and all monochromatic components consist of monochromatic polyomino regions, and cycles.. □

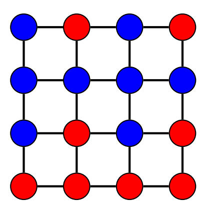
We have now classified all period 1 and 2 configurations of cycles, paths, and toroidal grids. This gives us a general idea of what to expect at the end of the Conservative Majority Model process. Moreover, some of the period 1 configurations are what we call monochromatic persistent sets which will be explored in the next chapter.



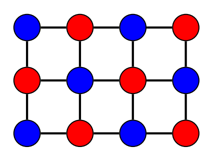
Period 1 polyomino region



Period 1 blue monochromatic cycle



Period 2 bichromatic cycles



Period 2 bichromatic polyomino region

Figure 2.5: Examples of Periodic Configurations on Toroidal Grid

Chapter 3

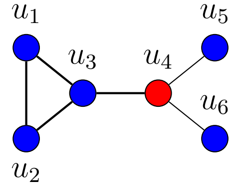
Persistent Sets

In this chapter we look at properties of persistent sets, which are sets of connected vertices that have at least half of their neighbours in the set. In terms of the spread of influence, these sets are particularly interesting because they can represent a group of people with an opinion that cannot be swayed. Moreover, with the existence of a persistent set in a configuration, it can be assured that a colour will not disappear.

A **persistent set** in a graph is a set of vertices I so that for each vertex $v \in I$, $|N(v) \cap I| \geq \frac{deg(v)}{2}$. If a configuration of a graph G is given, then we call a set a **monochromatic persistent set** if it is a persistent set and each vertex is the same colour. For example, as seen in Chapter 2, a persistent set of vertices in C_n could be two adjacent vertices. See Figure 3.1 for an example of a persistent set. The vertices labelled u_1, u_2, and u_3 form a persistent set and a monochromatic persistent set, while the vertices labelled u_5, and u_6 do not form a persistent set, but the vertices u_4, u_5, u_6 form a persistent set, but not a monochromatic persistent set.

Lemma 3.1 *Let $G = (V, E)$ be a graph, f a configuration, and I a monochromatic persistent set. Then for a sequence of configurations $\{f_t\}_{t=0}^{\infty}$ in the Conservative Majority Model, if $f_0 = f$, and for all $t \geq 0$, I is a monochromatic persistent set.*

Figure 3.1: A example of a persistent set



Proof: As I is a monochromatic persistent set, we have that for all $v \in I$, $|I \cap N(v)| \geq \frac{deg(v)}{2}$, and $f(v) = f_0(v) = c$, for some $c \neq u$, where u represents uncoloured. Now fix some $t \geq 0$ and assume that for all $v \in I$, $f_t(v) = c$. Then for each $v \in I$, $Maj(f_t, v) = c$ or $c \in \{\text{Maj}(f, v)\}$ in the event of a tie, and so $f_{t+1}(v) = c$. Thus, by definition I is a monochromatic persistent set in f_{t+1}. □

Recall that a vertex v has period k in f if for a sequence of configurations $\{f_t\}_{t=0}^{\infty}$, $f_0 = f$, we have that $f_t(v) = f_{t+k}(v)$ for all $t \geq 0$. We let k be the least period. In particular, a configuration f has period k if and only if all vertices in G have period k in f.

Corollary 3.1 *Let $G = (V, E)$ be a graph, f a configuration, and I a monochromatic persistent set. Then for all $v \in I$, v has period 1 in f.*

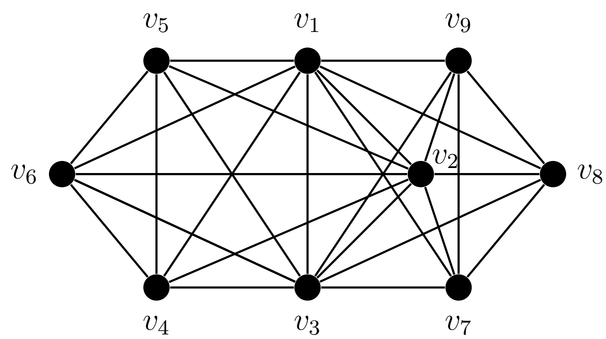
Clearly the union of two disjoint persistent sets is again a persistent set, therefore, we will only consider connected persistent sets.

Theorem 3.1 *The union of two persistent sets I and J is a persistent set.*

Proof: By definition each vertex v in I has $|I \cap N(v)| \geq \frac{deg(v)}{2}$, and each vertex v in J has $|J \cap N(v)| \geq \frac{deg(v)}{2}$. Therefore, each vertex v in $I \cup J$ has $|(I \cup J) \cap N(v)| \geq \frac{deg(v)}{2}$. □

On the other hand, the intersection of two persistent sets I and J is not always a persistent set. Let $G = (V, E)$ be a graph where $V = \{v_1, ..., v_9\}$, and $I = \{v_1, ..., v_6\}$ and $J = \{v_1, v_2, v_3, v_7, v_8, v_9\}$ each form a clique of size 6 such that $I \cap J = \{v_1, v_2, v_3\}$. It is easy to see that I and J are persistent sets; however, $I \cap J = \{v_1, v_2, v_3\}$ is not a persistent set because each vertex in $I \cap J$ has degree 8, but each vertex in the graph induced by $I \cap J$ has degree $2 < \frac{deg(v_i)}{2} = 4$, $i = 1, 2, 3$. See Figure 3.2.

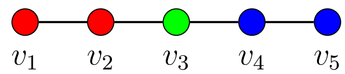
Figure 3.2: Counterexample for $I \cap J$ is not a persistent set



A **minimal monochromatic persistent** set in a configuration f is a monochromatic persistent set I so that no subset of I is a monochromatic persistent set. Likewise, a **maximal monochromatic persistent set** in a configuration f is a monochromatic persistent set J so that each persistent set J' containing J has at least one vertex of a different colour. Both maximal and minimal monochromatic persistent sets are not necessarily unique in a configuration.

Suppose you have a cycle C_n, and let f be a configuration for the Conservative Majority Model. An example of a minimal monochromatic persistent set I is two adjacent vertices u, v so that $f(u) = f(v)$. A single vertex cannot be a persistent set by definition, and so I forms a minimal monochromatic persistent set. However, for $n > 3$, there may be two different pairs of such monochromatic persistent sets; therefore, these sets are not unique. An example of a maximal monochromatic persistent J is three vertices v_0, v_1, v_2 in a K_5 so that $f(v_0) = f(v_1) = f(v_2)$, $f(v_4) \neq f(v_0)$, and $f(v_3) \neq f(v_0)$. Now, by definition the sets of vertices $\{v_0, v_1, v_2, v_3\}$, $\{v_0, v_1, v_2, v_4\}$, and $\{v_0, v_1, v_2, v_3, v_4\}$ all form persistent sets. However, each set of vertices is no longer monochromatic. Thus, any persistent set in K_5 containing J is not monochromatic, and so J is maximal. Moreover, by Theorem 3 the union of persistent sets is also a persistent set, and so the union of monochromatic persistent sets are persistent sets. Thus either there is a unique maximal monochromatic persistent set, or there is a maximal monochromatic persistent set for more than one colour.

Figure 3.3: A vertex of period 1 that is not a persistent set



There exist vertices of period 1 that are not in monochromatic persistent sets. For example, suppose you have P_5, with vertices $\{v_1, v_2, v_3, v_4, v_5\}$. Let $f(v_1) = f(v_2) = r$, $f(v_3) = g$, and $f(v_4) = f(v_5) = b$. Clearly, v_1, v_2 form a monochromatic persistent set, and v_4, v_5 form another monochromatic persistent set, but v_3 also has period 1, as there is always a tie in its neighbourhood, and it is not a persistent set. See Figure 3.3. However, as we will see in the upcoming Theorem, if a configuration for a graph G has period 2, and all period 1 vertices are in monochromatic persistent sets, then the removal of the persistent sets will not change the periodicity of the configuration.

Let I be a set of vertices, f be a configuration on a graph G and f^{-I} be the corresponding configuration on $G - I$.

Theorem 3.2 *Let $G = (V, E)$ be a graph, and let f be a period 2 configuration for G under the Conservative Majority Model with two colours and uncoloured. Let $\mathcal{P}$ be the set of all vertices that have period* 1*. Then $f^{-\mathcal{P}}$ is a period* 2 *configuration for $G - \mathcal{P}$.*

Proof: Let $G = (V, E)$ be a graph, and f a period 2 configuration on G. Let $\{f_t\}_{t=0}^{\infty}$ be a configuration sequence for G so that $f_0 = f$. Recall, a configuration has period 2 if all vertices in f have period 2. And so, all vertices either have period 1 or period 2.

Let $g = f^{-\mathcal{P}}$. Given the set of all period 1 vertices $\mathcal{P}$, we let $\{g_t\}_{t=0}^{\infty}$ be the corresponding configuration sequence obtained from applying the Conservative Majority Model to $G - \mathcal{P}$, with initial configuration g. We will now show that for all $t \geq 0$, $f_t|_{G-\mathcal{P}} = g_t$, and g_t has period 2.

Fix some t and assume $f_t(u) = b$ and $f_{t+1}(u) = r$ for some $u \in V$. Because f has period 2, we have $f_t = f_{t+2}$. Let R_2^t and B_2^t be the sets of red and blue neighbours of u respectively of period 2 at time t. Similarly, let R_1^t, and B_1^t be the sets of red and blue neighbours of u of period 1 at time t. Observe that both B_1^t and R_1^t are in the set $\mathcal{P}$.

As B_1^t and R_1^t are sets of period 1 vertices, then $B_1^t = B_1^{t+1}$ and $R_1^t = R_1^{t+1}$. Suppose $B_1^t \geq \frac{deg(u)}{2}$, then $\text{Maj}(f_t, u) = b$ or $\{b \in \text{Maj}(f_t, u)\}$, and $f_{t+1}(u) = b$, which is a contradiction of our assumption that $f_{t+1}(u) = r$. Now, suppose $R_1^t = R_1^{t+1} \geq \frac{deg(u)}{2}$, then $\text{Maj}(f_{t+1}, u) = r$ or $\{r \in \text{Maj}(f_t, u)\}$, and $f_{t+2}(u) = r \neq f_t(u)$, which is also a contradiction since we assumed that f had period 2. Therefore, both R_1^t and B_1^t are less than $\frac{deg(u)}{2}$.

As $f_t(u) = b$, $Maj(f_t, u) = r$, $f_{t+1}(u) = r$, and

$$R_2^t + R_1^t > B_2^t + B_1^t. \tag{3.1}$$

As $f_{t+1}(u) = r$, $Maj(f_{t+1}, u) = b$, $f_{t+2}(u) = f_t(u) = b$.

$$B_2^{t+1} + B_1^{t+1} > R_2^{t+1} + R_1^{t+1}. \tag{3.2}$$

As R_2^t is the set of period 2 neighbours, then at time $t+1$, each of the vertices in this set is now blue. Therefore, $R_2^t = B_2^{t+1}$. Similarly, $B_2^t = R_2^{t+1}$, and so

$$R_2^t + B_1^{t+1} > B_2^t + R_1^{t+1}. \tag{3.3}$$

Moreover combining (3.1) and (3.3), we have

$$2R_2^t + R_1^t + B_1^{t+1} > 2B_2^t + B_1^t + R_1^{t+1} \tag{3.4}$$

$$R_2^t > B_2^t. \tag{3.5}$$

As $R_2^t > B_2^t$ at time t, then $g_{t+1}(u) = r$, which also means that $B_2^{t+1} > R_2^{t+1}$, and so $g_{t+2}(u) = b = g_t(u)$, for each vertex u. Thus, as $g_0 = f_0$, we have that

$g_0(u) = f_0(u)$ for all $u \in G - \mathcal{P}$, and by the above calculations $g_1(u) = f_1(u)$ for all $u \in G - \mathcal{P}$. As both f and g have period 2 then $g_2 = g_0 = f_0 = f_2$, and so for all $t \geq 0$, $g_t = f_t$ for $G - \mathcal{P}$.

Thus, g is a period 2 configuration for $G - \mathcal{P}$. □

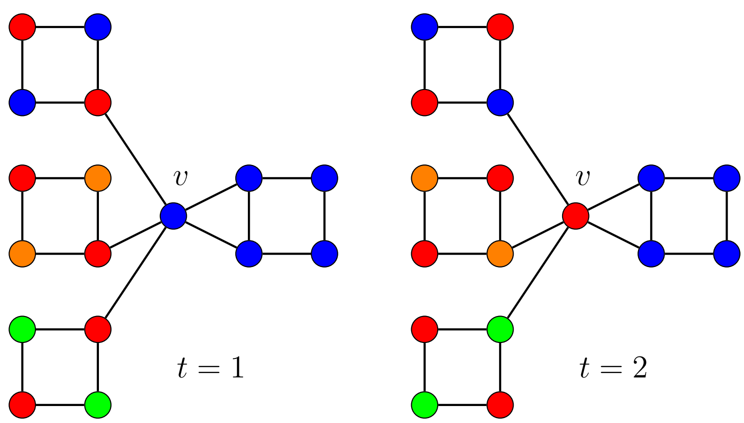
In general, from the above theorem, given that $f_0(u) = b$, the majority at time t for $t \geq 0$ at vertex u in $G - \mathcal{P}$ is determined by,

$$f_t^{-\mathcal{P}}(u) = \begin{cases} b & \text{if } t \text{ is even} \\ r & \text{if } t \text{ is odd} \end{cases} \tag{3.6}$$

The above theorem will only work for the existence of 2 colours. Suppose we let the Conservative Majority Model proceed with n colours and for some u, $f_t(u) = r$ and $f_{t+1}(u) = b$. With n colours a vertex does not necessarily need $\frac{deg(u)}{2} + 1$ neighbours of one colour to change. Now if each of the blue neighbours is an element of a monochromatic persistent set, then its removal will affect the periodicity of u. In fact the graph may no longer be periodic if the set is removed. Consider the following graph in Figure 3.4. The four adjacent blue vertices form a monochromatic persistent set. At $t = 1$, v has three red neighbours, and two blue neighbours. At $t = 2$, v has three blue neighbours, a green neighbour, and an orange neighbour. With the removal of the persistent set, at time $t = 2$, v will have a tie, and remain red. Thus, the removal of the persistent set has an effect on the period of vertex v. In fact, v now has period 1 instead of period 2.

We have explored some basic properties of persistent sets and their interaction with other colours. Persistent sets can be useful in determining final configurations of the models. In the following chapter we see how persistent sets of cycles can be used to see the persistence of a colour, and the long term behaviour of complete graphs.

Figure 3.4: Persistent sets with more than 2 colours



Chapter 4

Long Term Behaviour

In many of the majority voter model studies the eventual goal is to be able to predict the long term behaviour of the process. This is important for real world applications such as viral marketing.

We study the Conservative Majority Model with a random initial configuration. Precisely, given probabilities $P_0, P_1, ..., P_c$ so that $P_0 + P_1 + ... + P_c = 1$, we colour each vertex independently so that $Pr(f(v) = i) = P_i$. Then we set $f = f_0$ and run the model.

The goal is to be able to predict the periodic configurations that are reached in the long run. We call a configuration monochromatic if all the vertices are the same colour, and is multi-chromatic otherwise. In particular, we focus on complete graphs and cycles in this chapter. When we look at complete graphs, we show that with high probability if $P_b > P_i, i \neq b, 0$ then the graph will be monochromatic with colour b. Following this, we look at cycles and determine specific configurations that always ensure the survival of a colour. We then look at these configurations and determine the initial probability of a colour so that with high probability one of the specific configurations will be in the initial configuration and the colour will persist.

4.1 Complete Graphs

In this section we will look at the initial probabilities of each colour and their relation to the final configuration on the complete graph K_n. As each vertex in a complete graph has influence on every other vertex, one can easily see that if more than half of the vertices are coloured the same colour, then the whole graph will be that colour under any majority model and for any number of colours in the next step of the process. In general, as long as a colour is adopted by more vertices than any other colour, it will persist; it does not necessarily need to be half of the vertices. We look at the case with c colours and uncoloured. We consider the Conservative Majority Model with c colours, and uncoloured.

In this section the results are asymptotic based on the number of vertices n in a graph, and not the number of colours. The number of colours c is a constant. And so, our results happen asymptotically almost surely (a.a.s) as $n \to \infty$. See Section 1.2.2 for the definition of a.a.s.

Theorem 4.1 *Let $K_n = (V, E)$ be the complete graph of order n, and let c be the number of colours, and b an integer. If $P_b > P_i$ for $1 \leq i \leq c, i \neq b$, then asymptotically almost surely K_n will be monochromatic with colour b in the Conservative Majority Model.*

Proof: Let X be the random variable that counts the number of vertices that are coloured j. Fix colour j, $1 \leq j \leq c$, and let X_i be its indicator variable where for $v_i \in V$,

$$X_i = \begin{cases} 1 & \text{if } f(v_i) = j \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

For each X_i, we have that X_i is a Bernoulli random variable with probability P_j, $X = \sum_{i=1}^{n} X_i$, and $E(X) = P_j n$. Using Hoeffding's inequality as given in Section

1.2.2 and letting $0 = a_i$ and $1 = b_i$ for each X_i, we have that,

$$Pr(|X - E(X)| \geq \sqrt{n}\log n) \leq 2\exp\left(\frac{-2(\sqrt{n}\log n)^2}{\sum_{i=1}^{n}(1-0)^2}\right) \quad (4.2)$$

$$= 2\exp\left(\frac{-2n\log^2 n}{n}\right) \quad (4.3)$$

$$= 2\exp(-2log^2 n) \quad (4.4)$$

As $2\exp(-2\log^2 n)$ approaches zero as $n \to \infty$, then with high probability X will not deviate from its expected value any more than $\sqrt{n}\log n$. Next we show that with high probability, the number of other coloured vertices are each less than the number of vertices coloured b. Firstly, let $\epsilon < \frac{P_b - P}{2}$ where $P = max\{P_i, i \neq b\}$. By assumption, $\epsilon > 0$. Now, we show that with high probability, the number of vertices coloured blue is always more or at least as many than the number of vertices coloured any other colour. Let n_i be the number of vertices in K_n coloured i. We look at the event $E_b = \{n_b \geq (P_b - \epsilon)n\}$ and $E_i = \{n_i \leq (P_i + \epsilon)n\}$ for $1 \leq i \leq c, i \neq b$. We want to show that the event $E_b \cap \bigcap_{i \neq b} E_i$ happens with high probability. Therefore, we show $E_b^c \cup \bigcup_{i \neq b} E_i^c$ has a probability tending to 0 as n grows. And so, we use our results from the Hoeffding bound for each E_i^c:

$$P(E_i^c) = P(n_i > (P_i + \epsilon)n) \quad (4.5)$$

$$\leq Pr(|n_i - E(n_i)| \geq \epsilon n) \quad (4.6)$$

$$< P(|n_i - E(n_i)| \geq \sqrt{n}\log n) \quad (4.7)$$

$$\leq 2\exp(-2log^2 n) \quad (4.8)$$

And so, we use the union bound to bound $P(E_b^c \cup \bigcup_{i \neq u,b} E_i^c)$.

$$P(E_b^c \cup \bigcup_{i \neq u,b} E_i^c) \leq 2c\exp(-2log^2 n) \quad (4.9)$$

As $n \to \infty$, $2c\exp(-2\log^2 n) \to 0$. Therefore asymptotically almost surely the event $E_b \cap \bigcap_{i \neq b} E_i$ will happen. □

In the following section, we will look at the cycle graph with red, blue, and uncoloured, and find the initial probabilities that guarantee the persistence of blue.

4.2 Cycles

We will look at the initial configuration probabilities and their relation to the final configuration on the cycle. For the purpose of this chapter we will only consider the Conservative Majority Model with an initial configuration of only two colours: red, blue, and uncoloured. For the colours blue, red, and uncoloured, we let $P_b + P_r + P_u = 1$.

Consider the following configurations of induced subpaths of a cycle C_n. We assume the vertices of C_n are numbered $v_0, ..., v_{n-1}$. An induced subpath v_i, v_{i+1} of length 2 such that $f(v_i) = f(v_{i+1}) = b$ will be denoted BB. An induced subpath of length 3 v_j, v_{j+1}, v_{j+2} such that $f(v_j) = f(v_{j+2}) = u$ and $f(v_{j+1}) = b$ will be denoted UBU. The induced subpath $v_i, ..., v_j$ of length $2k$ for $2k \leq n-2$ such that $f(v_i) = f(v_j) = b$ and $f(v_{i+1}) = ... = f(v_{j-1}) = u$ is denoted as $BU...UB$. As $2k \leq n-2$, it is not possible that $v_i = v_j$, and there will be no overlap.

Theorem 4.2 *In the Conservative Majority Model on C_n, the colour blue will persist if at least one of the following occurs in the initial configuration:*

- *a BB induced subpath,*
- *a UBU induced subpath, or*
- *a $BU...UB$ induced subpath.*

Proof: If the first configuration BB appears, then the two vertices will always be blue as they form a monochromatic persistent set.

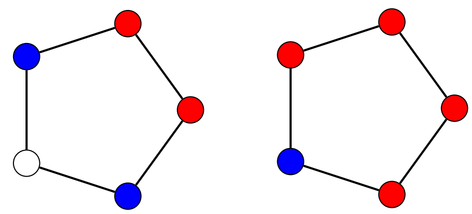
Assume that configuration UBU appears. Consider first that $n \geq 5$, there are six possibilities up to symmetry for the colours of the vertices adjacent to both uncoloured vertices. The possibilities are $UUBUU$, $BUBUU$, $BUBUB$, $BUBUR$, $UUBUR$ and $RUBUR$. For the first five cases, it is easy to see in the next time step that the configuration will contain the BB induced subpath, and the central vertex will remain blue. For the case of $RUBUR$, we note that at the vertices labelled U a tie occurs. Each uncoloured vertex has a red neighbour and a blue neighbour and so neither vertex will change in the next time step and the central vertex will remain blue. In following steps, for the configuration $RUBUR$, either the red vertices always remain red, or one red vertex turns blue. If they always remain red, then B persists. If a red vertex turns blue, then we have the configuration $RUBUB$, and in the next step we will have the configuration $RUBB$, and BB forms a persistent set. Therefore, blue persists.

If $n < 5$ and the path UBU appears, then $n = 3$ or $n = 4$. If $n = 3$, then UBU colours the entire graph, and so all vertices will turn blue in the next step. If $n = 4$, then both uncoloured vertices share a neighbour. If the neighbour is blue or uncoloured, then it is clear that all vertices will turn blue in the next steps. If the neighbour is red, then both uncoloured vertices have a tie in their neighbourhood, and do not change colour. Consequently, the blue vertex does not change either, and so blue persists.

Finally assume that the configuration $BU...UB$ appears. We use induction to show that when $BU...UB$ appears for any path of uncoloured vertices of length $2k$, blue will persists. For the base case let's consider a path of uncoloured vertices of length 0. With 0 uncoloured vertices, we have the configuration BB, which we know persists. Suppose that blue persists with a configuration $BU...UB$ with $2k$ uncoloured vertices. Consider a configuration $BU...UB$ with $2(k+1)$ uncoloured vertices. We label the vertices of the path as $v_1, ..., v_{2k+2}$. In the next step of the period vertices v_2 and v_{2k+2} will turn blue, which leaves us with a path of uncoloured vertices of length $2k$, which we assumed ensures the persistence of blue. Therefore, by induction we have that $BU...UB$ for any path of uncoloured vertices of length $2k$, blue will persist. $\square$

By symmetry, Theorem 4.2 also holds for the persistence of red. We note

Figure 4.1: Uncoloured vertex path of odd length



that for the configuration $BU...UB$, if the path is of odd length, then blue will not necessarily persist. Consider the configuration $RBUBR$ at time t on a cycle, at $t+1$ we have $RRBRR$ which then becomes the path R. See Figure 4.1.

Using the configurations BB, UBU, and $BU...UB$, we will find the initial probabilities so that with high probability b will persist. In [8], Gartner and Zehmakan proved that if both $P_b, P_r \gg \frac{1}{\sqrt{n}}$, then with high probability the final configuration will be bichromatic on the Conservative Majority Model and the Biased Majority Model. Gartner and Zehmakan only used two colours: red and blue, and there were no uncoloured vertices. See Section 1.2.2 for the definition of the notation $\gg$.

On a cycle, the configurations BB and RR are persistent sets. The authors considered a maximum matching of C_n into $\lfloor \frac{n}{2} \rfloor$ pairs. They let X_i for $1 \leq i \leq \lfloor \frac{n}{2} \rfloor$ be the event that a pair contains two blue vertices, and so $P(X_i = 1) = P_b^2$. Moreover, the variables X_i are independent. Letting $X = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} X_i$, $P(X = 0) \leq (1 - P_b^2)^{\lfloor \frac{n}{2} \rfloor} \leq (e^{-P_b^2})^{\lfloor \frac{n}{2} \rfloor}$, which approaches zero as n goes to infinity.

With the addition of uncoloured vertices, there is a relationship between the probability of being uncoloured P_u, and the probability of being coloured blue P_b. In fact, P_b may be less than $\frac{1}{\sqrt{n}}$, and still with high probability blue will persist provided that P_u is large enough.

Next, using the configurations from Theorem 4.2, we will now look at the initial probabilities required for a colour to persist to the final configuration. The following theorems all include uncoloured vertices.

Theorem 4.3 *If $P_b \gg \frac{1}{\sqrt{n}}$, then with high probability blue will persist for the Conservative Majority Model on the cycle C_n.*

Proof: Let X be the random variable which counts the number of times the pattern BB appears in the initial configuration and $X = \sum_{i=0}^{n-1} X_i$ where $X_i = 1$ if the pattern BB appears in positions i and $i+1$ calculated modulo n, and $X_i = 0$ if the pattern BB does not appear in positions i and $i+1$. We have that $E(X) = \sum_{i=0}^{n-1} E(X_i)$. The probability that $X_i = 1$ is P_b^2 and so, $\sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} P(X_i = 1) = nP_b^2$.

Next we use Chebyshev's inequality.

$$Pr(X = 0) \leq Pr(|X - E(X)| \geq E(X)) \leq \frac{Var(X)}{E(X)^2} \tag{4.10}$$

Now by definition,

$$Var(X) = \sum_{i}^{n} Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j). \tag{4.11}$$

In calculating the covariance, we need to consider the cases in which X_i and X_j overlap in 1 or 0 vertices. As defined in Section 1.2.2 for random variable X the variance is $Var(X) = E(X^2) - (E(X))^2$. The expected value of X_i is P_b^2, and so $Var(X_i) = P_b^2 - P_b^4$. Similarly, for two random variables X, Y, the covariance is $Cov(X, Y) = E(XY) - E(X)E(Y)$. If X_i and X_j overlap in one vertex, then there will be a path of three blue vertices, meaning $E(X_iX_j) = P_b^3 - P_b^4$, and $E(X_i)E(X_j)$ is simply P_b^4. If X_i and X_j overlap in two vertices, then $X_i = X_j$ and so the covariance is 0.

$$Var(X_i) = P_b^2 - P_b^4 \tag{4.12}$$

$$Cov(X_i, X_j) = \begin{cases} {P_b}^3 - {P_b}^4 & \text{if } |j - i| = 1 \\ 0 & \text{if } |j - i| \geq 2 \end{cases} \tag{4.13}$$

Now using (4.12) and (4.13) we can calculate the upper bound of Chebyshev's inequality.

$$Var(X) = n(P_b^2 - P_b^4 + 2(P_b^3 - P_b^4)) \tag{4.14}$$

$$= nP_b^2(1 + 2P_b - 3P_b^2) \tag{4.15}$$

This gives us the following,

$$\frac{Var(X)}{E(X)^2} = \frac{nP_b^2(1 + 2P_b - 3P_b^2)}{n^2P_b^4} \tag{4.16}$$

$$= \frac{1 + 2P_b - 3P_b^2}{nP_b^2} \tag{4.17}$$

$$\leq \frac{3}{nP_b^2}. \tag{4.18}$$

As $n \to \infty$, we have that (4.18) approaches zero if and only if nP_b^2 approaches infinity. So if $P_b \gg \frac{1}{\sqrt{n}}$. This implies that $P(X = 0)$ goes to zero. Thus, with high probability the initial configuration will contain BB, and the final configuration will contain the colour blue. □

Theorem 4.4 *If $P_b {P_u}^2 \gg \frac{1}{n}$, then with high probability blue will persist for the Conservative Majority Model on the cycle C_n.*

Proof: Let X be the random variable which counts the number of times UBU appears in the initial configuration and $X = \sum_{i=0}^{n-1} X_i$ where $X_i = 1$ if the pattern UBU appears in positions $i, i+1$, and $i+2$ calculated modulo n, and $X_i = 0$ if the pattern does not appear in these positions. We have that $E(X) = \sum_{i=0}^{n-1} E(X_i)$ by linearity of expectation. The probability that $X_i = 1$ is $P_b P_u^2$ and so, $\sum_{i=0}^{n-1} E(X_i) = \sum_{i=0}^{n-1} P(X_i = 1) = nP_b P_u^2$.

We use Chebyshev's inequality and (4.11). In calculating the covariance, we need to consider three possible cases for events X_i and X_j: $|j-i|=2$, $|j-i|=1$, and $|j-i|\geq 3$. These represent the cases when the patterns overlap in 1, 2, or 0 vertices. The expected value of X_i is $P_bP_u^2$, and so $Var(X_i) = P_bP_u^2 - P_b^2P_u^4$. If X_i and X_j overlap in one vertex, then there will be a path $UBUBU$, and $Cov(X_i,X_j)$ is simply $P_b^2P_u^3 - P_b^2P_u^4$. If X_i and X_j overlap in two vertices, then the path is $UBBU$ which does not contain UBU. Thus, they have a negative covariance and $Cov(X_i,X_j) = -P_b^2P_u^4$. When X_i, X_j overlap in three vertices then $X_i = X_j$ and there is zero covariance.

$$Var(X_i) = P_b{P_u}^2 - {P_b}^2{P_u}^4 \tag{4.19}$$

$$Cov(X_i,X_j) = \begin{cases} {P_b}^2{P_u}^3 - {P_b}^2{P_u}^4 & \text{if } |j-i|=2 \\ -P_b^2P_u^4 & \text{if } |j-i|=1 \\ 0 & \text{if } |j-i|\geq 3 \end{cases} \tag{4.20}$$

Using (4.19), and (4.20) the variance is calculated.

$$Var(X) = n(P_b{P_u}^2 - P_b^2P_u^4 + 2(P_b^2P_u^3 - P_b^2P_u^4 - P_b^2P_u^4)) \tag{4.21}$$

$$= nP_bP_u^2(1 - 5P_bP_u^2 + 2P_bP_u) \tag{4.22}$$

We have the following,

$$\frac{Var(X)}{E(X)^2} \leq \frac{nP_bP_u^2(1-5P_bP_u^2+2P_bP_u)}{n^2P_b^2P_u^4} \tag{4.23}$$

$$= \frac{1-5P_bP_u^2+2P_bP_u}{nP_bP_u^2} \tag{4.24}$$

$$\leq \frac{3}{nP_bP_u^2} \tag{4.25}$$

As $n\to\infty$, we have that (4.25) approaches zero if and only if $nP_bP_u^2$ approaches infinity. So if $P_bP_u^2 \gg \frac{1}{n}$, then $P(X=0)$ goes to zero. Thus, with high probability the initial configuration will contain UBU, and thus the final configuration will contain blue. □

By symmetry, the above theorems also hold for the persistence of red in the final configuration.

In looking at the configuration $BUUB$, we note that from Theorem 15 it is implied that with high probability, if $P_b^2 {P_u}^2 \gg \frac{1}{n}$, then blue persists. If the configuration $BUUB$ occurs with high probability, then it is very likely that the configuration UBU also appears, since if $P_b^2 P_u^2 \gg \frac{1}{n}$, then $P_b P_u^2 \gg \frac{1}{n}$ as $P_b P_u^2 \geq P_b^2 P_u^2$. For the general case of $BU...UB$ we have that if $P_b^2 P_u^{2k} \gg \frac{1}{n}$ then $P_b^2 P_u^2 \gg \frac{1}{n}$ and $P_b P_u^2 \gg \frac{1}{2}$, and the patterns $BUUB$ and UBU will appear. Thus, again the bound for $P_b P_u^{2k}$ is already implied from the previous case.

With n colours and uncoloured, we can still use the same methods to determine whether or not a final configuration will contain a specific colour on the cycle. In the next chapter we look at how long the process of the Conservative Majority Model takes to reach its final configuration and that it has period 2, and similarly for a similar majority model.

Chapter 5

Periods and Pre-Periods

Knowing whether or not a model is periodic gives insight to recurring trends. It is also important as it will make the model easier to understand, and knowing the length of the period allows us to classify periodic configurations. Moreover, we study the length of the pre-period, which tells us how long it will take for a majority model process to stabilize.

In this chapter we will give an upper bound on the pre-period for the Conservative Majority Model with two colours and uncoloured, and prove that the model has period 2. That is, for every starting configuration, there exists a time t^* so that $f_{t^*} = f_{t^*+2}$. We then look at a mixed majority model, and prove that its model has least period 2, and find an upper bound of its pre-period.

5.1 The Linear Algebra Framework for two Colours and Uncoloured

The objective of this section is to find an upper bound of the pre-period and the length of the period of the Conservative Majority Model with two colours and uncoloured. As previously mentioned, the addition of the uncoloured vertex makes the model slightly more complicated on graphs due to the fact that it acts differently from other colours. Firstly, a vertex cannot adopt being uncoloured, and once a vertex becomes coloured, it will never be uncoloured again. This property has the potential to add extra steps to the pre-period. An uncoloured vertex can have a tie in its neighbourhood for several steps before having a majority and turning colour. There have been results on pre-periods for both the Conservative and Biased Majority Models in [8] and [17].

We find an upper bound of the pre-period for the Conservative Majority Model with two colours, red and blue, and uncoloured. In order to do so, we represent the model in a linear algebra framework. This framework was first described in [17] with a general linear algebra function. The model given in [17] has two colours and no uncoloured vertices. It is described as follows.

Let G be a graph of order n, and let A be the adjacency matrix of G, and x be a vector in $\{-1, 1\}^n$, where

$$x(v) = \begin{cases} -1 & \text{if } f(v) = red \\ 1 & \text{if } f(v) = blue. \end{cases} \tag{5.1}$$

The product $(Ax)(v)$ gives some integer k which represents the difference of the blue and red neighbours of a vertex v. If there are more blue neighbours then $(Ax)(v) > 0$, and if there are more red neighbours, $(Ax)(v) < 0$. If there is a tie, then $(Ax)(v) = 0$.

Given any vector x of size n we have the sign function $sgn(x) : \mathbb{R}^n \rightarrow \{-1, 1\}^n$, where

$$sgn(x(v)) = \begin{cases} 1 & \text{if } x(v) \geq 0 \\ -1 & \text{if } x(v) < 0. \end{cases} \tag{5.2}$$

For the general linear algebra framework of the Conservative Majority Model, we have an initial opinion vector $x_0 \in \{-1, 1\}^n$, and for $t \geq 0$, $x_{t+1} = sgn(Ax_t)$.

For the Conservative Model with red, blue and uncoloured, we change the domain of sgn to $\{-1, 0, 1\}$, where $sgn(x(v)) = 0$ if $x(v) = 0$. We modify A to add a memory process to the framework, as in the event of a tie we may have $(Ax)(v) = 0$, but $x(v) \neq 0$. Let $0 < \epsilon < 1$. Define $B = A + \epsilon I_n$. Then if there is a tie and $x(v) = 1$, then $(Bx)(v) = \epsilon$, and if there is a tie and $x(v) = -1$, then $(Bx)(v) = -\epsilon$. Thus, for the linear algebra framework with two colours and uncoloured, we have an initial opinion vector $x_0 \in \{-1, 0, 1\}^n$, and for $t \geq 0$, $x_{t+1} = sgn(Bx_t)$.

Theorem 5.1 *Let $\{f_t\}_{t\geq 0}$ be the configuration sequence obtained by applying the rules of the Conservative Majority Model. Let $\{x_t\}_{t=0}^{\infty}$ be the sequence of vectors given by $x_{t+1} = sgn(Bx_t)$. Let $f_0 = x_0$, then for all $t \geq 0$, $f_t = x_t$.*

Proof: We consider the Conservative Majority Model with the colours red, blue, and uncoloured, which are represented in the vector x as $-1, 1, 0$ respectively.

We use induction on t. For our base case we have $f_0 = x_0$. Let $t > 0$ be some time and assume that $f_t = x_t$. For some vertex v, let $b_t = |N_b^t(v)|$, and $r_t = |N_r^t(v)|$ be the number of blue neighbours of v at time t and the number of red neighbours at time t. We look at the cases where $b_t > r_t$, and $b_t = r_t$.

If $b_t > r_t$, then $f_{t+1}(v) = b$. Moreover, when $b_t > r_t$, then $(Bx_t)(v) > 0$, and so $x_{t+1}(v) = sgn((Bx_t)(v)) = 1$. Therefore, $f_{t+1}(v) = x_{t+1}(v)$.

Suppose v is blue at time t. If $b_t = r_t$, then $f_{t+1}(v) = b$. When $b_t = r_t$, then $(Bx)(v) = b_t - r_t + \epsilon > 0$, and $x_{t+1}(v) = sgn((Bx_t)(v)) = 1$. And so, $f_{t+1}(v) = x_{t+1}$.

Suppose now that v is uncoloured at time t. If $b_t = r_t$, then $f_{t+1}(v) = u$. When $b_t = r_t$, then $(Bx)(v) = b_t - r_t + 0\epsilon = 0$, and $x_{t+1}(v) = sgn((Bx_t)(v)) = 0$. And so, $f_{t+1}(v) = x_{t+1}$.

The argument is symmetric for $r_t > b_t$, and when v is red at time t.

For each vertex v, we have $f_{t+1}(v) = x_{t+1}(v)$, which means $f_{t+1} = x_{t+1}$. Therefore, for every $t \geq 0$, we have $f_t = x_t$. □

Now, we will introduce some notation that will be used. Let x be a vector of size n, and let $||x|| = \sum_{i=1}^{n} |x(i)|$. Let A be a matrix of size $n \times n$ then $||A|| = \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|$.

To find an upper bound of the pre-period of the Conservative Majority Model with two colours and uncoloured, we define a potential function for the linear algebra framework:

$$V(t) = x_{t+1}^T B x_t - u_t, \tag{5.3}$$

where u_t is the number of uncoloured vertices at time t.

We now prove some properties of this potential function and framework.

Lemma 5.1 *Let $\{x_t\}_{t=0}^{\infty}$ be a vector sequence representing the Conservative Majority Model. Then $x_{t+1}^T B x_t = ||Bx_t||$.*

Proof: We have that $||Bx_t|| = \sum_{i=1}^{n} |(Bx_t)(i)|$. Now, we know that $x_{t+1}(v) = sgn((Bx_t)(v))$. In general, $sgn(x)x = ||x||$ for any vector x. Here we have $x_{t+1}^T Bx_t = (sgn(Bx_t))^T Bx_t$. $\square$

Lemma 5.2 *Let $G = (V, E)$ be a graph, and t^* be the minimum t so that x_t is a periodic configuration. Then for all t, $0 \le t \le t^*$, $V(t) + 2 - \delta \le V(t+1)$, when $\epsilon = 1 - \delta$ and $\delta = \frac{1}{(|V||E|)^2}$.*

Proof: By Lemma 5.1, we know that $x_{t+2}^T Bx_{t+1} = ||B_{t+1}||$. As we are in the pre-period then $x_t \neq x_{t+2}$, and so there must be some vertex v where $x_t(v)$ differs from $x_{t+2}(v) = sgn(B_{t+1}(v))$. For each vertex that doesn't change from t to $t+2$, the difference of the matrix multiplication is $x_{t+2}(v)(Bx_{t+1})(v) - x_t(v)(Bx_{t+1})(v) = 0$, as $x_{t+2}(v) = x_t(v)$.

Suppose $sgn(x_t(v)) = -1$, $sgn(x_{t+1}(v)) = -1$ and $sgn(x_{t+2}(v)) = 1$. Then we have that at time $t+1$, there must be a blue majority so that $x_{t+2}(v) = 1$. If that is the case, then $(Bx_{t+1})(v) \ge 1$, and $x_{t+2}(v)(Bx_{t+1})(v) \ge 1$. Moreover, as $x_t(v) = -1$, then $x_t(v)(Bx_{t+1})(v) \le -1$. And so $x_{t+2}(v)(Bx_{t+1})(v) - x_{t-1}(v)(Bx_{t+1})(v) \ge 2$. If no uncoloured vertices change, $u_t - u_{t+1} = 0$, otherwise $u_t - u_{t+1} > 0$. Combining with all the other vertices we have $V(t+1) - V(t) \geq 2 - \delta$.

Let $sgn(x_t(v)) = -1$, $sgn(x_{t+1}(v)) = 1$ and $sgn(x_{t+2}(v)) = 1$. Then, if there is not a majority of blue, it is the case that there is a tie at $t+1$ and v conserves its colour. If this is the case then $(Bx_{t+1})(v) = \epsilon$. And so $x_{t+2}(v)(Bx_{t+1})(v) = \epsilon$, $x_t(v)(Bx_{t+1})(v) = -\epsilon$, and $x_{t+2}(v)(Bx_{t+1})(v) - x_t(v)(Bx_{t+1})(v) = 2\epsilon = 2 - \delta$. As before $u_{t-1} - u_t = 0$, and so $V(t+1) - V(t) \geq 2 - \delta$.

Now, let $sgn(x_t(v)) = 0$, $sgn(x_{t+1}(v)) = -1$, and $sgn(x_{t+2}(v)) = 1$. Then $u_t > u_{t+1}$, and $u_t - u_{t+1} = 1$. As there is a change in colour from time $t+1$ to $t+2$, there must a blue majority, and so $(Bx_{t+1})(v) \geq 1$. This means that $x_{t+2}(v)(Bx_{t+1})(v) \geq 1$. As $sgn(x_t(v)) = 0$, then $x_t(v)(Bx_{t+1})(v) = 0$, and the difference at least 1. Therefore again, $V(t+1) - V(t) \geq 2 - \delta$.

Let $sgn(x_t(v)) = 0$, $sgn(x_{t+1}(v)) = 1$, and $sgn(x_{t+2}(v)) = 1$. Then, if there is not a majority of blue, it is the case that there is a tie at $t+1$ and v conserves its colour. If this is the case then $(Bx_{t+1})(v) = \epsilon$. And so $x_{t+2}(v)(Bx_{t+1})(v) = \epsilon$, $x_t(v)(Bx_{t+1})(v) = 0$, and $x_{t+2}(v)(Bx_{t+1})(v) - x_t(v)(Bx_{t+1})(v) = \epsilon$. As before $u_{t-1} - u_t = 1$, and so combining these two equalities, $V(t+1) - V(t) \geq 2 - \delta$.

Suppose now that $sgn(x_t(v)) = 0$, $sgn(x_{t+1}(v)) = 0$, and $sgn(x_{t+2}(v)) = 1$. We have that $u_t - u_{t+1} = 0$, and there is a tie at v at time t. Let B_{t+1} and R_{t+1} be the number of blue and red neighbours of v at time $t+1$. If $R_{t+1} + B_{t+1}$ is odd, and $u_t = u_{t+1}$, then at time t we must have $R_t + B_t$ is odd. However, in the event of a tie we have that $(Bx_t)(v) = \pm\epsilon$ which cannot happen if $R_t + B_t$ is odd as $|R_{t+1} - B_{t+1}| \geq 1$.

Suppose that $R_{t+1} + B_{t+1}$ is even. As $sgn(x_t(v)) = 0$, then $x_t(v)(Bx_{t+1})(v) = 0$. At time $t+1$ at v, $(Bx_{t+1})(v) \geq 1$, as there is assumed to be a blue majority. We know $R_t = B_t$, and $B_{t+1} > R_{t+1}$. Now,

$$B_{t+1} - B_t \geq 1 \tag{5.4}$$

$$R_t - R_{t+1} \geq 1 \tag{5.5}$$

$$B_{t+1} - R_{t+1} + R_t - B_t \geq 2 \tag{5.6}$$

$$B_{t+1} - R_{t+1} \geq 2. \tag{5.7}$$

Therefore, $(Bx_{t+1})(v) \geq 2$, which means $x_{t+2}^T Bx_{t+1} - x_t^T Bx_{t+1} \geq 2 - \delta$, and $V(t+1) - V(t) \geq 2 - \delta$. □

Corollary 5.1 *The potential function $V(t)$ either increases or $x_t = x_{t+2}$.*

Proof: From Lemma 5.2, it follows that the function $V(t)$ does not decrease. So letting k be the length of the period we have,

$$V(0) + (2 - \delta) \leq V(1) + (2 - \delta) \leq ... \leq V(k) + (2 - \delta) = V(0) + (2 - \delta) \tag{5.8}$$

$$V(0) = V(1) = ... = V(k-1). \tag{5.9}$$

Clearly in the period no uncoloured vertices change so $u_0 = ... = u_{k-1}$, and so each time in the period $x_{t-1}Bx_t = x_{t+1}Bx_t$, which is only true when $x_{t-1} = x_{t+1}$. Therefore $x_t = x_{t+2}$ □

This corollary shows that the Conservative Majority Model has period 2.

Theorem 5.2 *Let $G = (V, E)$ be a graph of order n. The Conservative Majority Model with two colours and uncoloured has a pre-period with at most $\left\lfloor |E| + \frac{n+u_0}{2} \right\rfloor$ steps, where u_0 is the number of uncoloured vertices at time 0.*

Proof: We use the linear algebra framework of the Conservative Majority Model, where A is the adjacency matrix of G, $B = A + \epsilon I_n$, and $\delta = \frac{1}{(|V||E|)^2}$ and

$\epsilon = 1 - \delta$. Let t^* be the pre-period, and let $x_0 \in \{-1, 0, 1\}^n$ be the initial vector of colours at time $t = 0$ and $x_{t+1} = sgn((Bx_t))$ for $t \geq 0$.

Consider the sum $V(t)$ from $t = 0$ to $t^* - 1$, we have the following:

$$\sum_{t=0}^{t^*-1}(V(t+1) - V(t)) = (x_{t^*+1}^T B x_{t^*} - x_0^T B x_1) + (u_0 - u_{t^*}). \tag{5.10}$$

By Lemma 5.2, we know that our potential function for the model $V(t)$ is increasing by at least $2 - \delta$ for each time step; therefore, we have:

$$(x_{t^*+1}^T B x_{t^*} - x_0^T B x_1) + (u_0 - u_{t^*}) \geq (2 - \delta)t^*. \tag{5.11}$$

Next, we find bounds for $x_{t^*+1}^T B x_{t^*}$, $x_0^T B x_1$, u_0, and u_{t^*}. To find an upper bound of $x_{t^*+1}^T B x_{t^*}$, we note that all entries in x_{t^*} are at most 1, and every entry in B is non-negative. Since we know $x_{t^*+1}^T B x_{t^*} = ||Bx_{t^*}||$, then $x_{t^*+1}^T B x_{t^*} \leq ||B\mathbf{e}|| = ||B||$, where $\mathbf{e}$ is the all ones vector, and so $x_{t^*}^T B x_{t^*+1} \leq ||B||$. A lower bound for $x_0 B x_1$ is simply 0. Finally, $u_{t^*} \geq 0$ always as in the period it is possible for there to be zero uncoloured, and we leave u_0 as is.

Now, combining all the bounds we have:

$$(x_{t^*+1}^T B x_{t^*} - x_0^T B x_1) + (u_0 - u_{t^*}) \leq (||B|| - 0) + (u_0 - 0) \tag{5.12}$$

$$= ||B|| + u_0, \tag{5.13}$$

which implies that $(2 - \delta)t^* \leq ||B|| + u_0$. We have

$$t^* \leq \frac{||B|| + u_0}{2 - \delta} \tag{5.14}$$

$$= \frac{||A|| + n\epsilon + u_0}{2 - \delta} \tag{5.15}$$

$$\leq \frac{2|E| + n + u_0}{2 - \delta}. \tag{5.16}$$

Since t^* is an integer, and $\delta, \epsilon < 0$, we then get that our pre-period is at most $\left\lfloor |E| + \frac{n+u_0}{2} \right\rfloor$. $\square$

As mentioned earlier, Poljak and Turzik [17] studied the Conservative Majority Model using the linear algebra framework with just two colours and no uncoloured. To find the upper bound, they modified the matrix A for their framework to deal with ties in their proof. Using the adjacency matrix A they created a larger matrix C of size $n+1$ where $c_{ij} = a_{ij}$ for $i, j \leq n$ and,

$$c_{n+1,i} = c_{i,n+1} = \begin{cases} 1 & \text{if } \sum_j a_{ij} \text{ even} \\ 0 & \text{if } \sum_j a_{ij} \text{ odd} \end{cases} \tag{5.17}$$

$$c_{n+1,n+1} = ||A|| + 1 \tag{5.18}$$

With the addition of the extra row and column, they have $\sum_{j=1}^{n} c_{ij}$ is odd for $i = 1, ..., n$. With only two colours, this means that there will never be a tie in the neighbourhood of any vertex, and so for all $v \in V$, $(Cx)(v) \neq 0$. Thus in the event of a tie and $(Ax)(v) = 0$, then $(Cx)(v) = \pm 1$ depending the current colour. The function used is sgn as described earlier. Now, with the matrix C, the function becomes $h : \mathbb{R}^{n+1} \to \{-1, 1\}^{n+1}$, where $h(x_t(1), ..., x_t(n), x_t(n+1)) = (sgn(x_t(1)), ..., sgn(x_t(n)), 1)$.

Our proof followed a similar procedure to this paper. Their potential function was $V'(t) = x_{t+1}^T C x_t$, and they had that $V(t+1) - V(t) \geq 2$. And so, after summing over the pre-period they had the value $x_{t^*+1}^T C x_{t^*} - x_0^T C x_1 \geq 2t^*$. In the paper the upper bound of $x_{t^*+1}^T C x_{t^*}$ was $4|E| + 2s + 1$, where s is the number of vertices with even degree in G. The lower bound of $x_0^T C x_1$ was $2|E| - s + n + 1$. Combining the bounds, the upper bound of the pre-period was at most $|E| + 3s - n$.

In our model we included uncoloured vertices. When uncoloured vertices are included, dealing with ties becomes more difficult. For example, a vertex can have odd degree and still have a tie in its neighbourhood. As uncoloured vertices have no effect on the decisions of their neighbours, it is fitting to let $u = 0$ as it will not contribute to the overall sum in $x_{t+1}^T B x_t$. Thus, we have split the number line to the signs of the numbers. However, if we tried this method with more than 2 colours, and uncoloured then it would not work, as there is no natural way to divide the number line into more than 3 parts for this process.

We will now compare the bounds from our model and the model in [17]. Our model has a bound of $\lfloor |E| + \frac{n+u_0}{2} \rfloor$ and the other model has a bound of $|E| + 3s - n$. Our bound for two colours and uncoloured is most closely reached by graphs with fewer edges, such at trees, cycles, and paths. The upper bound of the pre-period of a C_n is $\lfloor |E| + \frac{n+u_0}{2} \rfloor = \frac{3n+u_0}{2}$. Now suppose we have C_4 and let the vertices be ordered cyclically, and let $x_0 = (0, -1, -1, 1)$. The upper bound of our pre-period will be 6. At time 1, we have $x_1 = (0, -1, -1, -1)$, and finally at time 2 we have $x_2 = (-1, -1, -1, -1)$. As we can see here, it took 3 steps to reach the period, where our bound predicted at most 6 steps. Therefore, the upper bound can be reached within at least 3 steps with cycles. Moreover, a graph in which the bound is reached is P_2 with initial configuration $x_0 = (1, 0)$. The predicted length of the pre-period is 2, and clearly $x_1 = (1, 1)$ in the next step.

With Poljak and Turzik's bound $|E| + 3s - n$, a bound on cycles with no uncoloured vertices is $n + 3n - n = 3n$ which is double our bound for cycles when there are no uncoloured vertices. Also, in graphs with a large number of even degree vertices, our bound will be much better. For example k-regular graphs for even k. Our predicted bound is $\frac{(k+1)n}{2}$, and their bound is $\frac{(k+4)n}{2}$ which is larger. However, in the case of graphs with no even degree vertices, we have that the bound in [17] is better. For example, our bound for a K_4 graph with no uncoloured vertices is 8, whereas their paper has an upper bound of only 2.

Figure 5.1: P_2 with 2 colours matrices A and C

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} \epsilon & 0 & 1 & 0 \\ 0 & \epsilon & 0 & 1 \\ 1 & 0 & \epsilon & 0 \\ 0 & 1 & 0 & \epsilon \end{pmatrix}$$

5.2 The Linear Algebra Framework for n Colours

Often, there are more than just two colours, states, or opinions. Thus, it would be more realistic to observe the Conservative Majority Model with n colours with uncoloured. In [17], Poljak and Turzik were able to find an upper bound of the pre-period and the period of a majority voter model with n colours. We will modify the framework of their model to find the period and the pre-period of the Mixed Majority Model with n colours. Recall Definition 1.3 from Chapter 1, this model contains properties of both the biased and conservative majority models, where for all $v \in V$, if there is a tie in v's neighbourhood and v's colour is in the tie, v conserves its colour for the next step. Otherwise, v will pick the most preferred colour that is in the tie.

Let $G = (V, E)$ be a graph with adjacency matrix A, and let there be n colours. We modify A to adapt to the process. Each non-zero entry in A is replaced with the identity matrix I_n, each entry A_{ii} is replaced with $\mathcal{E}_n = \epsilon I_n$, and each zero entry of A is replaced with 0_n. Let B be the modified matrix of size $|V|n \times |V|n$. See Figure 5.1 for an example of the constructed matrix B for a path P_2 with 2 colours. Let x be the opinion vector of size $|V|n$. Here the vector x is divided into $|V|$ blocks of size n. Each block $x(v)$ represents the colouring of v. Each row of $x(v)$ represents one of the n colours. An entry in $x(v)$ will be 1 if that colour is adopted by v, and 0 otherwise. For example, if there are 5 colours, and $x(v)$ is the vector $[0, 0, 1, 0, 0]$, then vertex v has chosen colour 3. Let $B(i, j)$ denote the $n \times n$ matrix, which replaced the entry A_{ij}.

For any vector x, and standard bases vectors e_i, let

$$Max(x) = max\{x \cdot e_i, 1 \leq i \leq n\} \tag{5.19}$$

and,

$$min(x) = e_i, \text{ where } i = min\{j : x \cdot e_j = Max(x)\}\}. \tag{5.20}$$

Given a vector x of size mn, we define $g : \mathbb{R}^{mn} \rightarrow \{0,1\}^{mn}$ as

$$g(x(m)) = min(x(m)) \tag{5.21}$$

In this framework of the Mixed Majority Model, we have an initial opinion vector $x_0 \in \{0,1\}^{mn}$, and for $t \geq 0$, $x_{t+1} = g(Bx_t)$.

Theorem 5.3 *Let* $\{f_t\}_{t=0}^{\infty}$ *be the configurations sequence obtained by applying the Mixed Majority Model with* n *colours. Let* $\{x_t\}_{t=0}^{\infty}$ *be the sequence of vectors given by* $x_{t+1} = g(Bx_t)$*. Let* $f_0 = x_0$*, then for all* $t \geq 0$*,* $f_t = x_t$*.*

Proof: First we show that $(Bx)(v)$ counts the number of neighbours of each colour of v. In $(Bx)(v)$, each entry i is determined by $B(v,1)x(1) + ... + B(v,|V|)x(|V|)$. For $1 \leq j \leq |V|$, if $B(v,j)x(j) = I_n$ then $B(v,j)x(j) = x(j) = e_i$ where i is the colour of the vertex j. Note that $B(v,j)$ is I_n if and only if j is a neighbour of v. Therefore, $(Bx)(v)$ gives the number of neighbours of each colour. Moreover, $B(v,v)x(v) = \epsilon e_i$, where i is the current colour, which creates a memory process for this linear algebra framework.

We use induction on t. We have already $f_0 = x_0$. Let $t > 0$ and assume that $f_t = x_t$.

Suppose that $f_t(v) = b$ and $Maj(v,f) = i$ and i is the only colour in the majority. The set $\{j : x(v) \cdot e_j = Max((Bx_t)(v)), 1 \leq j \leq n\}$ has one element i. This is due to the fact that $(Bx_t)(v)$ counts the number of neighbours of each colour, and so there will be only one element i such that $e_i = Max((Bx_t)(v))$. Thus, in both cases $f_{t+1}(v) = i$ and $x_{t+1}(v) = g((Bx_t)(v)) = e_i$. And so, $f_{t+1}(v) = x_{t+1}(v)$.

Suppose again that $f_t(v) = b$ and $i_1, ..., i_k$ are the colours in the majority, such that i_1 is the most preferred colour and i_k is the least preferred colour. Then $f_{t+1}(v) = i_1$ as colour i_1 is the more preferred colour. Then colours $i_1, ..., i_k \in \{j : x(v) \cdot e_j = Max(x), 1 \leq j \leq n\}$, and so $x_{t+1}(v) = g((Bx_t)(v)) = e_{i_1}$. And so, $f_{t+1}(v) = x_{t+1}(v)$.

Finally suppose that $f_t(v) = b$ so that b is one of the colours in the majority. Let $m = Max(x_t(v))$. If i is a majority colour not equal to b, then row i of $(Bx_t)(v)$ equals m. The value of the rows of $(B(x_t))(v)$ of each majority colour are all the integer m. Because of the addition of $\mathcal{E}_n$ along the diagonal of B, we have that row b of $(Bx_t)(v)$ is $m+\epsilon$, and so b is in the majority. Therefore, $x_{t+1}(v) = g((Bx_t)(v)) = e_b$. And so, $f_{t+1}(v) = x_{t+1}(v)$.

Thus, for each vertex v in G, we have $f_{t+1}(v) = x_{t+1}(v)$, which means that $f_{t+1} = x_{t+1}$. Therefore, for all $t \geq 0$, we have $f_t = x_t$. □

We define the potential function of the Mixed Majority Model as

$$V(t) = x_{t+1}^T B x_t + \frac{1}{2} \sum_{v \in V} \delta(v_{t+1}, v_t), \tag{5.22}$$

where v_t represents the colour of v at time t, and

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise.} \end{cases} \tag{5.23}$$

Lemma 5.3 *For all $t \geq 0$, if $x_t \neq x_{t+2}$ then $V(t) + \frac{1}{2} - \epsilon \leq V(t+1)$.*

Proof: If we are in the pre-period then $x_t \neq x_{t+2}$, and so there must be at least one vertex v where $x_t(v)$ differs from $x_{t+2}(v) = g((Bx_{t+1})(v))$.

Let $V(t) = \sum_v V(v,t)$ where $V(v,t) = x_{t+1}(v)(Bx_t)(v) + \frac{1}{2}\delta(v_{t+1}, v_t)$. Then we look at the change of $V(v,t+1) - V(v,t)$. This will contribute to the overall sum of $V(t+1) - V(t)$.

Suppose that $x_t(v) = x_{t+1}(v) = j$ and $x_{t+2}(v) = i \neq j$. Recall the vector $(Bx_{t+1})(v)$ represents the number of each colour present in the neighbourhood of v. Since v's colour at time $t+2$ is i, the majority colour in v's neighbourhood at time $t+1$ is i. Since v changed colour from time $t+1$ to $t+2$, j is not a majority colour of v at time $t+1$. Thus, the component of $(Bx_{t+1})(v)$ corresponding to i denoted b_i is the largest, and the component corresponding to j denoted b_j is strictly smaller. So $b_j + 1 - \epsilon \leq b_i$, where the ϵ comes from the contribution of the colour of v itself at time $t+1$.

Since $x_t(v) = e_j$, we have that $x_t(v)(Bx_{t+1})(v)$ equals b_j, and similarly, $x_{t+2}(v)(Bx_{t+1})(v) = b_i$. Then,

$$x_{t+2}(v)(Bx_{t+1})(v) - x_t(v)(Bx_{t+1})(v) + \frac{1}{2}(\delta(v_{t+1}, v_{t+2}) - \delta(v_t, v_{t+1})) \tag{5.24}$$

$$\geq 1 - \epsilon + \frac{1}{2}(\delta(v_{t+1}, v_{t+2}) - \delta(v_t, v_{t+1})) \tag{5.25}$$

Since $\delta(v_{t+1}, v_{t+2}) = 0$ and $\delta(v_t, v_{t+1}) = 1$, the contribution to $V(v,t+1) - V(v,t)$ is at least $1/2 - \epsilon$.

Suppose that $x_{t+2}(v) = x_{t+1}(v) = i$ and $x_t(v) = j \neq i$. Since v's colour at time $t+2$ is i, the majority colour in v's neighbourhood at time $t+1$ is i. Since v remained the same colour from time $t+1$ to $t+2$, b_i is the largest component and there is an epsilon term as v's colour contributes $+\epsilon$ to b_i. Thus, b_j is strictly smaller than b_i and $b_j + \epsilon \leq b_i$.

Since $x_t(v) = e_j$, we have that $x_t(v)(Bx_{t+1})(v)$ equals b_j, and similarly, $x_{t+2}(v)(Bx_{t+1})(v) = b_i$. Then,

$$x_{t+2}(v)(Bx_{t+1})(v) - x_t(v)(Bx_{t+1})(v) + \frac{1}{2}(\delta(v_{t+1}, v_{t+2}) - \delta(v_t, v_{t+1})) \tag{5.26}$$

$$\geq \epsilon + \frac{1}{2}(\delta(v_{t+1}, v_{t+2}) - \delta(v_t, v_{t+1})) \tag{5.27}$$

Since $\delta(v_{t+1}, v_{t+2}) = 1$ and $\delta(v_t, v_{t+1}) = 0$, the contribution to $V(v, t+1) - V(v, t)$ due to vertex v is at least $1/2 + \epsilon$.

Suppose that $x_t(v) = l$, $x_{t+1}(v) = j \neq l$, and $x_{t+2}(v) = i \neq j, l$. Since v's colour at time $t+2$ is i, the majority colour in v's neighbourhood at time $t+1$ is i. As v changed colour from time $t+1$ to $t+2$, j is not a majority colour of v at time $t+1$. Thus, b_i is the largest value, and b_j is strictly smaller. So $b_j + 1 \leq b_i$.

Since $x_t(v) = e_i$, we have that $x_t(v)(Bx_{t+1})(v)$ equals b_j, and similarly, $x_{t+2}(v)(Bx_{t+1})(v) = b_i$. Then,

$$x_{t+2}(v)(Bx_{t+1})(v) - x_t(v)(Bx_{t+1})(v) + \frac{1}{2}(\delta(v_{t+1}, v_{t+2}) - \delta(v_t, v_{t+1})) \tag{5.28}$$

$$\geq 1 + \frac{1}{2}(\delta(v_{t+1}, v_{t+2}) - \delta(v_t, v_{t+1})) \tag{5.29}$$

Since $\delta(v_{t+1}, v_{t+2}) = 0$ and $\delta(v_t, v_{t+1}) = 0$, v's contribution to $V(v, t+1) - V(v, t)$ is at least 1.

Now, summing over all v, we have $\sum_v V(v, t+1) - \sum_v V(v, t) = V(t+1) - V(t) \geq \frac{1}{2} - \epsilon$. Therefore, each vertex changing colour from time t to time $t+2$ contributes at least $\frac{1}{2} - \epsilon$ to $V(t+1) - V(t)$. For any vertex v so that $v_{t+2} = v_t$, we have that $\delta(v_{t+2}, v_{t+1}) = \delta(v_t, v_{t+1})$, and $x_{t+1}(v)(Bx_{t+2})(v) - x_t(v)(Bx_{t+1})(v) = (x_{t+2}(v) - x_t(v))(Bx_{t+1})(v) = 0$, and thus these vertices do not contribute positively or negatively to $V(t+1) - V(t)$. Thus, we have that $V(t)$ is increasing, and if $x_{t+2} \neq x_t$, then $V(t+1) - V(t) < 1/2 - \epsilon$. Thus, when the process has become periodic, $t > t^*$, $V(t)$ must remain constant since for a period k, we have

$$V(0) \leq V(1) \leq ... \leq V(k-1) \leq V(k) = V(0). \tag{5.30}$$

Thus, if $V(t)$ is not increasing, then $V(t) = V(t+1)$, and $x_t = x_{t+2}$. □

Theorem 5.4 *Let $G = (V, E)$ be a graph, and n be the number of colours. Then the Mixed Majority Model has period 2 and pre-period at most $4|E| - |V|$.*

Proof: We use the linear algebra framework of the Mixed Majority Model. Let t^* be the pre-period of the Mixed Majority Model with n colours and uncoloured, and let $x_0 \in \{0,1\}^{|V|n}$ be the initial vector of colours at time $t = 0$, and $x_{t+1}(v) = g((Bx_t)(v))$ for $t \geq 0$.

From Lemma 5.3 we have $V(t) < V(t+1)$ when $x_t \neq x_{t+2}$ and that in the period $V(t) = V(t+1)$. Now, letting k be the period and $t > t^*$, we have $V(t) = V(t+k)$. And so, V must be constant for all $t > t^*$. However, if the period is larger than 2, then $x_t \neq x_{t+2}$. This means that $V(t)$ is still increasing, and so $V(t) \neq V(t+1)$ is not in the period. Thus, $x_t = x_{t+2}$ and the period is 2.

Consider the increase of $V(t)$ from $t = 0$ to $t^* - 1$,

$$\sum_{t=0}^{t^*-1}(V(t+1) - V(t)) = (x_{t^*+1}Bx_{t^*} - x_0Bx_1) + \frac{|V|}{2}(\delta(u_{t^*+1}, u_{t^*}) - \delta(u_1, u_0)) \quad (5.31)$$

By Lemma 5.3, we know that $V(t)$ increases by $\frac{1}{2} - \epsilon$ at each time step. Thus, summing over all time steps of the pre-period we have that the lower bound of the above sum is $(\frac{1}{2} - \epsilon)t^*$. Therefore, we have

$$(x_{t^*+1}Bx_{t^*} - x_0Bx_1) + \frac{|V|}{2}(\delta(u_{t^*+1}, u_{t^*}) - \delta(u_1, u_0)) \geq \left(\frac{1}{2} - \epsilon\right)t^*. \quad (5.32)$$

Next, we find an upper bound to $(x_{t^*+1}Bx_{t^*} - x_0Bx_1) + \frac{|V|}{2}(\delta(u_{t^*+1}, u_{t^*}) - \delta(u_1, u_0))$.

The sum $\delta(u_{t^*+1}, u_{t^*}) - \delta(u_1, u_0)$ can always be bounded above by 1, and so multiplying by $\frac{|V|}{2}$, we get an upper bound of $\frac{|V|}{2}$.

The upper bound of $x_{t^*+1}Bx_{t^*}$ can be given by $||A|| + |V|\epsilon$. From Lemma 5.2 we know $(Bx_t)(v)$ counts the number of neighbours of each colour for v, the highest value $(Bx_t)(v)$ can have is $deg(v)$. If that is the case, then if every neighbour of v is coloured some colour i and $x_{t+1}(v) = e_i$, then $x_{t+1}(v)(Bx_t)(v) = deg(v)$. Thus, for any vertex $x_{t+1}(v)(Bx_t)(v) \leq deg(v)$. Moreover, along the diagonal we have $\mathcal{E}_n$ and so any multiplication of Bx_t we have an additional $|V|\epsilon$. Therefore, the upper bound is $\sum_{v \in V} deg(v) = ||A|| + |V|\epsilon$, where $||A|| = 2|E|$.

The value $x_1(v)(Bx_0)(v)$ is never zero since if v changes colour, then it must have a majority. And so it must have at least one neighbour of the majority colour. This means there exists some colour i so the $(Bx_0)(v) \geq 1$ and $x_1(v)(Bx_0)(v) \geq 1$. Thus, the next possible lowest bound for each vertex is one. Therefore, a lower bound to $x_0Bx_1 = x_1Bx_0$ is $|V| + |V|\epsilon$.

The upper bound to $(x_{t^*+1}Bx_{t^*} - x_0Bx_1) + \frac{|V|}{2}(\delta(u_{t^*+1}, u_{t^*}) - \delta(u_1, u_0))$ is

$$2|E| + \epsilon|V| - (|V| + \epsilon|V|) + \frac{|V|}{2}. \tag{5.33}$$

We choose $0 < \epsilon < 1$ so that $\epsilon t^* < 1$. Since t^* is an integer, we may remove the ϵ term. Therefore we have,

$$2|E| + \epsilon|V| - (|V| + \epsilon|V|) + \frac{|V|}{2} = 2|E| - \frac{|V|}{2} \tag{5.34}$$

$$\geq \frac{t^*}{2}. \tag{5.35}$$

Thus, $4|E| - |V| \geq t^*$ and that is the upper bound of the pre-period of the Mixed Majority Model. □

Our bound here is much lower than the bound given in [17] which was $2M^2||B||n(4n+1) + 2nr$ where $M = max\{||g(x(v))|| | v = 1, ..., |V|, x(v) \in \mathbb{Z}^n\}$, where n is the number of colours, and r is the size of the matrix. With our framework, $||g((Bx)(v))||$ will always equal 1, which is not the case in Poljak and Turzik's work so that already lowers the bound with our model.

For example, the upper bound of the pre-period for the cycle under our framework is $4|V| - |V| = 3|V|$; however, the upper bound of the pre-period from the other paper is $2(1)^2(4|E|)n(4n+1) + 2n(|V|n) = 8|E|(4n^2 + n) + 2n^2|V|$, which is clearly much larger. In fact, our bound will always be better because of the $8|E|$ term.

The graph K_2 with any configuration almost reaches the expected upper bound $4|E| - |V| = 2$. On P_3, the calculated upper bound is $4(2) - 3 = 5$; however for a configuration with 3 colours, we have $x_0 = \{(1,0,0),(0,1,0),(0,0,1)\}$ and $x_1 = \{(0,1,0),(0,0,1),(0,1,0)\}$, which is periodic with only 2 steps. The upper bound calculated on the complete graph K_n is $4|E| - |V| = 2(n(n-1)) - n = 2n^2 - 3n$ which is $\mathcal{O}(n^2)$. However, most complete graphs become stable within 1 or 3 time steps. Thus, although our bound is better than previous work, there is most likely a better bound for all graphs.

Having this result is important, as it allows us to be able to classify periodic configurations. Moreover, with such a small period it can be easier to determine the final proportions of colours, and makes the model easier to understand.

5.2.1 The Problem with the Conservative Majority Model

The initial goal was to find the length of the pre-period and the period with n colours of the Conservative Majority Model. However, under the linear algebra framework, it was not a easy construction. In order to prove the length of the pre-period and period, we needed a symmetric matrix B so that we could say $x_{t+1}Bx_t = x_tBx_{t+1}$, and a function so that $x_{t+1}Bx_t - x_tBx_{t-1} = f(x_t)Bx_t - f(x_{t-1})Bx_t \geq 0$ always. However under this linear algebra framework, we had the problem that $(g(Bx_t) - g(Bx_{t-1}))^TBx_t$ could be less than zero in the event of a tie.

Suppose a vertex v has a tie in its neighbourhood, but none of its neighbours have adopted the same colour as v, then we would have $x_{t+1}(v)(Bx_t)(v) = 0$. However, suppose that $x_{t-1}(v)$ is equal to a colour that some of the neighbours of v have adopted at time t, then $x_{t-1}(v)(Bx_t)(v) > 0$. To avoid a negative sum, we must account for this possibility. To do so, it requires modifying the value $x_{t+1}(v)(Bx_t)(v)$, so that $x_{t+1}(v)(Bx_t)(v) - x_{t-1}(v)(Bx_t)(v) \geq 0$. This requires, giving an advantage to only $x_{t+1}(v)(Bx_t)(v)$ so that $x_{t+1}(v)(Bx_t)(v) \geq x_{t-1}(v)(Bx_t)(v)$ always. However, this is no longer a symmetric system, as the same advantage would not be given to $x_t(v)(Bx_{t+1})(v)$. Alternatively, one could modify B so that the function is symmetric.

In doing so, it is required to create a non-symmetric matrix, and so $x_{t+1}Bx_t \neq x_tBx_{t+1}$.

As our proof methods require matrices and functions with which we can utilize the symmetry, it is difficult to find a suitable linear algebra framework for the Conservative Majority Model.

This concludes Chapter 5. We have created a new linear algebra framework for some majority models, found the upper bounds of pre-periods with up to n colours and showed that the period is 2 in both models.

Chapter 6

Conclusion

In this thesis we found several properties of the Conservative Majority Model. We also obtained theoretical results of the behaviour of the model with more than 2 colours with the addition of the uncoloured vertex.

In Chapter 2 we classified all period 1 and 2 configurations of the Conservative Majority Model on paths, cycles, and toroidal grids for n colours and uncoloured, and 2 colours and no uncoloured respectively. Further exploration here might be to classify the configurations on other families of graphs. Other directions could be to study the stability of the network when an extra edge or vertex is added.

In Chapter 3, we defined the notion of persistent sets in graphs, and monochromatic persistent sets. These sets could be useful in determining the final proportions of colours in the final configuration with high probability.

In Chapter 4, we studied some threshold probabilities that ensure the persistence of a colour. Results were found on the cycle and the complete graph where uncoloured vertices were included. Other possible directions are to find the threshold values for other families of graphs, in particular for the toroidal grid as it can be related to 2-dimensional cellular automata, and we have classified the periodic configurations.

Finally in Chapter 5 we proved that a model similar to the Conservative Majority Model called the Mixed Model is eventually periodic with period 2 and found the upper bound of its pre-period. We also showed that the upper bound of the pre-period of the Conservative Majority Model with 2 colours and uncoloured vertices is at most $\lfloor |E| + \frac{n+u_0}{2} \rfloor$, where u_0 is the number of uncoloured vertices at the initial configuration, and that it has period 2. This upper bound is almost reached by cycles, and paths. It would be interesting to find out if there exists a family of graphs with a certain initial configuration with the exact upper bound. We also proved that with n colours, the Mixed Model's pre-period has an upper bound of $4|E| - |V|$. Once again, further research can be done on lowering this upper bound, as in most cases, the upper bound is not reached. However, we have improved the bound from previous works.

Based on experimental results, the Conservative Majority Model appears to have period 2. Further work to be done is to find a way to create a symmetric model, or another representation of the process to see if it is indeed period 2.

In some cases there are cellular automata with deterministic updating rules that are in fact equivalent to a Turing Machine provided the automaton has external memory. This is known as Turing-Complete. An intriguing direction for this model is to study it in the world of computing and automata.

Question 6.1 *Is the Conservative or Biased Majority Voter Model Turing Complete?*

As the study of the Majority Voter Model with more that 2 colours and uncoloured is relatively new with few theoretical results, there are many directions for future research.

Bibliography

[1] Delyon B. Rio E. Bercu, B. *Concentration Inequalities for Sums and Martingales*. 2191-8198. SpringerBriefs in Mathematics, Switzerland, 1 edition, 2015.

[2] J.A. Bondy and U.S.R Murty. *Graph Theory*. 244. Springer, London, 1 edition, 2008.

[3] Claudine Chaouiya, Ouerdia Ourrad, and Ricardo Lima. Majority rules with random tie-breaking in boolean gene regulatory networks. *PLOS ONE*, 8(7):1–14, 07 2013.

[4] Wei Chen, Laks V. S. Lakshmanan, and Carlos Castillo. *Information and Influence Propagation in Social Networks*. Morgan & Claypool Publishers, 2013.

[5] Hanool Choi, Sang-Hoon Kim, and Jeho Lee. Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management*, 39(1):170–177, 2010.

[6] Christopher Duffy and Jeannette Janssen. The spread of cooperative strategies on grids with random asynchronous updating. *Internet Mathematics*, 2017.

[7] O. Durán and R. Mulet. Evolutionary prisoner's dilemma in random graphs. *Physica D: Nonlinear Phenomena*, 208(3):257 – 265, 2005.

[8] Bernd Gärtner and Ahad N. Zehmakan. Biased majority rule cellular automata. *CoRR*, abs/1711.10920, 2017.

[9] Svante Janson, Tomasz Łuczak, Tatyana Turova, and Thomas Vallier. Bootstrap percolation on the random graph $g_{n,p}$. *Ann. Appl. Probab.*, 22(5):1989–2047, 10 2012.

[10] Eocman Lee, Jeho Lee, and Jongseok Lee. Reconsideration of the winner-take-all hypothesis: Complex networks and local bias. *Management Science*, 52(12):1838–1848, December 2006.

[11] Angsheng Li and Xi Yong. Emergence of super cooperation of prisoner's dilemma games on scale-free networks. *PLoS ONE*, 10(2), 2015.

[12] V. Nicosia, F. Bagnoli, and V. Latora. Impact of network structure on a model of diffusion and competitive interaction. *EPL (Europhysics Letters)*, 94(6), June 2011.

[13] Martin Nowak and Robert May. The spatial dilemma of evolution. 3, 02 1993.

[14] Nishith Pathak, Arindam Banerjee, and Jaideep Srivastava. A generalized linear threshold model for multiple cascades. pages 965–970. IEEE Publishing, December 2010.

[15] Giovanni Pegoretti, Francesco Rentocchini, and Giuseppe Vittucci Marzetti. An agent-based model of innovation diffusion: network structure and coexistence under different information regimes. *Journal of Economic Interaction and Coordination*, 7(2):145–165, October 2012.

[16] D. Peleg. *Immunity against Local Influence*, pages 168–179. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[17] Svatopluk Poljak and Daniel Turzík. On pre-periods of discrete influence systems. *Discrete Applied Mathematics*, 13(1):33 – 39, 1986.

[18] Jeffrey S. Rosenthal. *A first look at rigorous probability theory*. 10. World Scientific, Singapore, 1 edition, 2000.

[19] Arto Salomaa. *Theory of Automata*. Pergamon Press, Hungary, 1 edition, 1969.

[20] B. Schönfish and A.M. de Roos. Synchronous and asynchronous updating in cellular automata. *Biosystems*, 51:123–143, 1999.

[21] A. Szolnoki, M. Perc, and G. Szabó. Diversity of reproduction rate supports cooperation in the prisoner's dilemma game on complex networks. *The European Physical Journal B*, 61(4):505–509, February 2008.

[22] H. Zhang, H. Zhang, A. Kuhnle, and M. T. Thai. Profit maximization for multiple products in online social networks. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9, April 2016.

[23] Y. Zhu, D. Li, H. Guo, and R. Pamula. New competitive influence propagation models in social networks. In *2014 10th International Conference on Mobile Ad-hoc and Sensor Networks*, pages 257–262, Dec 2014.

[24] Y. Zhu, D. Li, and Z. Zhang. Minimum cost seed set for competitive social influence. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9, April 2016.