SOLUTION OF URN MODELS BY GENERATING FUNCTIONS WITH APPLICATIONS TO SOCIAL, PHYSICAL, BIOLOGICAL, AND NETWORK SCIENCES

By

William Pickering

A Dissertation Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: MATHEMATICS

Examining Committee:

____

Chjan Lim, Dissertation Adviser

____

Boleslaw Szymanski, Member

____

Isom Herron, Member

____

David Isaacson, Member

Rensselaer Polytechnic Institute
Troy, New York

November 2016
(For Graduation December 2016)

© Copyright 2016
by
William Pickering
All Rights Reserved

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

Special thanks to my advisor, Chjan Lim, for his ongoing support and collaboration. We have collectively learned greatly from our research together. I also want to thank Boleslaw Szymanski, Isom Herron, and David Isaacson for also serving on my thesis committee. Each member has contributed a unique perspective that has molded this project.

I also want to thank the faculty of the RPI math department for their ongoing support and instruction. It was through these professionals that I began to fully understand the field of Applied Mathematics. I also want to thank Dawnmarie Robens, who has been of enormous assistance, even when I first stepped into the math department nearly a decade prior to the writing of this dissertation.

I finally thank my family who has done more to shape my trajectory that led to the creation of this thesis than anyone else. Without their love and support, these accomplishments would not be possible.

ABSTRACT

The primary subject of this thesis is the notion of an urn model and their applications to complex systems. It is demonstrated that several models of social, physical, and biological science are special cases of a large class of urn models that are then exactly solved. These models prescribe 2 or more urns with N balls distributed among them. Two balls are then drawn randomly and then redistributed among these urns. This redistribution is also stochastic, with distributions only depending on which urns the balls came from and the order in which they were drawn.

Examples of such an urn model date back to the Ehrenfest "dog-flea" model of molecular diffusion, created in 1907. In this model, one ball is chosen randomly to move to the opposite urn. The model was invented to describe the diffusion of particles within a closed container. This model was later diagonalized by Mark Kac in 1947 by using an innovative generating function method to calculate all eigenvalues and eigenvectors of the Markov transition matrix. We not only improve upon this method, but the models that we solve have a wider range of applications. In addition to the physical application of the Ehrenfest model, we have related such models to social interactions on networks, as well as genetic drift.

We improve upon the generating function method of Kac significantly to handle the more complex features of such urn models. The method of Kac reduces the eigenvalue problem to an ordinary differential equation that is solved exactly. Since our urn models draw two balls instead of one ball in the Ehrenfest model, we solve a partial differential equation instead. Interestingly, the method of solution given in this thesis easily solves the Ehrenfest model, where the original generating function method of Kac does not solve the general urn models easily.

Through the process of diagonalization, we have uncovered solutions that would otherwise be untenable by other methods. For instance, dynamics for small N, higher moments of consensus time, and the long time behavior of these urn models can be easily determined by these methods. Furthermore, an exact solution for all future probability distributions of some popular urn models had not been

found prior to the results given in this thesis. Due to the breadth of models that can be studied in this manner, this thesis represents a significant step forward in the subject of complex systems theory.

CHAPTER 1
Introduction

In this chapter, we discuss the basic premise of the thesis, its organization, and the applications of the models and the analysis.

1.1 Urn Models

In this section, we describe the various models that we study in this thesis. These urn models are stochastic systems that have N balls distributed among two or more urns. The urns are labeled (e.g. A and B), and can correspond to opinions, energy levels, alleles in an ecosystem, etc. The balls in the system can represent individual people, particles, organisms, etc. As time evolves, these individuals can change urns as balls are drawn and moved. Socially, this corresponds to people changing their opinions. Biologically, this represents a child that may have a different genetic structure than the parent it replaced.

While the stochastic properties of the models may change significantly, there are two things that are always consistent. The first property of these models is that N is conserved. That is, we assume that no balls are added or removed from the system. The second property is that each time step begins with two random balls being drawn from the urns. These balls are then redistributed among the urns stochastically, depending on which urns the balls came from and the order they were drawn.

1.2 Applications of Urn Models

Here we list some salient models that can be cast in the above urn framework. We discuss these models and their solution in more detail in the following chapters.

Portions of this chapter previously appeared as: W. Pickering, B. K. Szymanski, and C. Lim, "Analysis of the high dimensional naming game with committed minorities," *Phys. Rev. E*, vol. 93, no. 5, p. 052311, May 2016.

1.2.1 The Voter Model

The voter model has become an archetype for social influence in the literature [1–5]. The model is stated in a general network setting with N nodes. Each node also has one of two possible states (A or B). In a single discrete time step, one node is chosen randomly. This node then adopts the state of one of its neighbors, also chosen randomly. This process continues until every node has the same state. If the nodes have unanimous agreement in their state, we say that the system is in consensus.

A particularly interesting case of the network is the "mean-field" case: every node is connected to every other node. The random effects of the degree distribution and network process is averaged out in this case. This is the case in which the network is said to be the complete graph. As it happens, if the network is not the complete graph, the model may have similar features as the complete graph counterpart. For example, in chapters 2 and 5, we compare the voter model on the complete graph with the same model imposed on Erdős-Rényi random graphs [6]. We find that the solutions for the moments of consensus time and the number of states as a function of time (in the case where there are M states) are similar despite the network topology.

Such a model on the complete graph can be rewritten as an equivalent urn model for the macrostate of the system. In this new formulation, there are two urns, call them A and B. Two balls are drawn randomly from these urns, and then both of them are placed in urn that the second ball was drawn from. These models are equivalent in the sense that the probability distribution for the macrostates are equal to each other. That is, if $n_A(m)$ is the number of balls in urn A at time m, then this random variable has the same distribution as if $n_A(m)$ were the total number of nodes with state A on the complete graph. By recasting the model into an urn setting, we relate it to a special case of a general framework of urn models that can be solved by the generating function method we propose in this thesis.

1.2.2 The Naming Game

Another simple model of social influence is the naming game [7–14]. This model assumes that there are K opinions in the system, but an individual can adopt multiple opinions simultaneously. The opinions that an individual adopt is called its “word list”. By assuming this, each individual may adopt one of $2^K - 1$ possible states. The naming game model is defined as follows. In a single time step, choose one node in the network to be the speaker and another adjacent node to be the listener. The speaker chooses one opinion from its word list to send to the listener. The listener then changes its state according to the following rules:

1. If the spoken word is not in the listener’s word list, it is added to the list.

2. If the spoken word is in the listener’s word list, then the listener discards all words from its list except the spoken word. [8]

Again, this process continues until every individual adopts the same opinion. This model becomes very complex for large K. For fixed K, it is possible to express the macrostates of the system in terms of a large system of ordinary differential equations that described the mean population fraction for each word list [12]. However, as K becomes large, such as $K = O(N)$, this method fails. By treating the naming game as an urn model and solving it by the generating function method, we circumvent the exponential complexity of the differential equations and solve the problem for large K. This is possible because of the symmetries that are inherent in the naming game interactions.

In the naming game, we also want to consider the effect of committed minorities. These are individuals who never change their opinion regardless of the messages that they hear. We are interested in the influence of a small committed group of individuals on a majority population. In particular, we examine the number of zealots that it takes in order for the minority opinion to take over. As we show in chapter 6, there is a strong correlation between this critical fraction of zealots and the Shannon entropy of the social system.

We aim to establish that the naming game model can also account for dynamics of opinion spread in extreme initial conditions. Our motivating historical precedents are the dynamics of post-revolution opinion struggle. Often before revolution happens, the government identifies and suppresses the leading opposition minorities which are on the verge of achieving tipping fraction of support (e.g., Islamists before Iranian revolution of 1979 or Muslim Brothers before Egyptian revolution of 2011), so the revolution is conducted by a motley of opposition movements with different ideologies united only by opposition to the government. After the revolution, they remove suppression of such minorities allowing them to quickly win the majority of the population in agreement with the naming game model. However the case of the Russian revolution of February 1917 was different. The revolt was spontaneous, disorganized, and not led by any dominant minority exceeding tipping point fraction of the population as in the previous examples. Yet, in the midst of the disorder and dissent, a small Bolshevik party grasped the power and support of uncommitted individuals by November 1917, because their leader Lenin correctly diagnosed that the power laid on the streets. Here we study the case resembling such situations in the context of naming game, so there are committed minorities of multiple opinions. In [12], authors show that in such a case, a stalemate of opinion can more easily occur, in which no decision is reached. In contrast, we identify the new set of conditions for this case under which the loss of stability of a social system occurs. Under these conditions, instead of stagnation with no decision, a rapid change occurs in which a small minority quickly spreads their opinion to the uncommitted subpopulation. In addition, we show that in the presence of committed minorities, as opinion diversity of the uncommitted subpopulation increases, the size of the committed minority needed to the turn the uncommitted to the minority opinion decreases. In extreme cases, this critical committed minority is invariant of the system size. This suggests that too much dissent between individuals makes them susceptible to even a few zealots.

1.3 Motivation of the Solution

In every urn model that we consider, the system can be expressed as a Markov chain. For any Markov chain, there exists a transition matrix that describes the probability of moving from one state to another. The states in the Markov chain are the macrostates of the urn model, which are the total number of balls in each urn. The transition matrix describes how the probability distribution changes over time.

Generally speaking, let $\mathbf{a}^{(m)}$ be the probability distribution of the macrostates. If $\mathbf{T}$ is the Markov transition matrix given by the model, then the probability distribution at the next time step is given by

$$\mathbf{a}^{(m+1)} = \mathbf{T}\mathbf{a}^{(m)} \tag{1.1}$$

If we were to diagonalize $\mathbf{T}$, we could calculate $\mathbf{a}^{(m)}$ very easily. Specifically, let $\mathbf{T} = \mathbf{S}\Lambda\mathbf{S}^{-1}$ denote the diagonalization of $\mathbf{T}$. Therefore,

$$\mathbf{a}^{(m)} = \mathbf{S}\Lambda^{m}\mathbf{S}^{-1}\mathbf{a}^{(0)}. \tag{1.2}$$

So if we had access to the spectral decomposition of the transition matrix, then all future probability distributions would be known in closed form. Given this, the evolution of the model would be known in great detail, and exact solutions easily follow from these distributions. This motivates us to find all eigenvalues and eigenvectors of the transition matrix. To solve this, we rewrite the the eigenvalue problem as a partial differential equation for a generating function $G(x, y)$ for the eigenvector components and solve for eigenvalues λ and G. This method of diagonalization represents the key principle of this thesis – the properties of which are explored in detail in subsequent chapters.

CHAPTER 2
Exact Spectral Solution of the 2-State Voter Model on the Complete Graph

In this chapter, we will describe in detail the diagonalization of the 2-state voter model [1–5] by generating functions [15]. The voter model is an archetype in social opinion dynamics, and has been linked to population genetics and genetic algorithms [16–20]. The method of solution given here is also regarded as an archetype for other models and problems explored in this thesis. The solution of the eigenvalues and eigenvectors then allows us to solve for many other valuable quantities, such as the propagator, expected local time, expected consensus time.

2.1 The Voter Model as an Urn Model

The voter model is defined as follows. On a network of N nodes, each node has one of two states (A or B). In one discrete time step, a node is chosen randomly to adopt the state of one of its neighbors, also chosen randomly. This process will continue until all nodes have the same state. In the context of social systems, the states A and B are considered to be opinions. However, this model is also a special case of the Moran model of population genetics, in which A and B are alleles and mutation parameters are set to zero [16].

In this chapter, we assume that the network is a complete graph. That is, we assume every node is connected to all other nodes. This constitutes the mean field case often introduced to simplify the network topology [12, 15, 21–23]. More generalized networks are considered in chapter 4. It also happens that results found on the complete graph also hold for several sparse network topologies [12, 24].

We now proceed to recast this model on the complete graph as an equivalent urn-ball system. We do this by identifying a random walk that is equivalent to the random walk proposed by the voter model on the complete graph. The urn model

---

Portions of this chapter previously appeared as: W. Pickering and C. Lim, "Solution of the voter model by spectral analysis," *Phys. Rev. E*, vol. 91, no. 1, p. 012812, Jan. 2015.

that we propose is given as follows. There are two urns (call them A and B) with a total of N balls divided between the urns. In a discrete time step, choose two balls randomly, noting the order that they were drawn from the urns. Then, place both balls in the urn from which the first ball was drawn. Now, let $n_A(m)$ and $n_B(m)$ be the total number of balls in urn A and B respectively at discrete time m. Since the number of balls in the system is conserved, we have $n_A(m) + n_B(m) = N$ for all m. This allows us to characterize the system by a single variable, say n_A.

This urn process can be expressed as a random walk model. That is, we write,

$$n_A(m+1) = n_A(m) + \Delta n_A(m). \tag{2.1}$$

Since the total number of balls in an urn changes by at most 1, Δn_A only takes values from $\{-1, 0, 1\}$. From the definition of the urn model, the transition probabilities are given to be

$$Pr\{\Delta n_A(m) = 1 \mid n_A(m) = i\} = p_i, \tag{2.2}$$

$$Pr\{\Delta n_A(m) = -1 \mid n_A(m) = i\} = p_i, \tag{2.3}$$

$$Pr\{\Delta n_A(m) = 0 \mid n_A(m) = i\} = 1 - 2p_i. \tag{2.4}$$

Where,

$$p_i = \frac{i(N-i)}{N(N-1)}. \tag{2.5}$$

The transition probabilities given above exactly correspond to the transition probabilities for the network model. In the network setting, n_A is the total number of nodes that have state A. The relationship between urn models and network models is expanded upon in later chapters.

2.2 Markov Chain and Formulating the Propagator

The urn problem given in the previous section is a Markov Chain over the macrostate quantity n_A. With every (discrete) Markov chain, there exists a Markov transition matrix that describes the change in the probability distribution of the chain. Here, we establish the single step propagator and determine the structure of the transition matrix.

To represent the macrostate probability distribution, let

$$a_i^{(m)} = Pr\{n_A(m) = i\}. \tag{2.6}$$

This is the probability that the system is given by a particular macrostate $n_A = i$ at discrete time m. We can use generating functions to derive a recursion relation between future probability distributions $a_i^{(m+1)}$ conditioned on $a_i^{(m)}$. Let

$$R^{(m)}(x) = \sum_i a_i^{(m)} x^i. \tag{2.7}$$

We can now express the probability distribution for $\Delta n_A(m)$ as

$$D_i(x) = p_i x^{-1} + (1 - 2p_i) + p_i x, \tag{2.8}$$

where p_i is given by equation 2.5. Now, we utilize the following properties of generating functions of this type.

Product rule: If X and Y are integer random variables with probability generating functions $F(x)$ and $G(x)$ respectively, then the generating function of $X+Y$ is $F(x)G(x)$ [15, 25–27].

Sum rule: If the probability space is partitioned into N events, each with generating function $F_j(x)$, then the generating function for the entire space is $\sum_{j=1}^{N} F_j(x)$ [15, 25–27].

$$a_i^{(m+1)} = p_{i-1} a_{i-1}^{(m)} + (1 - 2p_i) a_i^{(m)} + p_{i+1} a_{i+1}^{(m)}. \tag{2.9}$$

From a numerical perspective, given the distribution $a_j^{(0)}$, one could iteratively apply equation 2.9 to calculate any $a_j^{(m)}$. However, Such a solution does not yield a detailed solution for the dynamics for various system sizes. Rather, we seek to find not only an efficient means of calculating $a_j^{(m)}$ for any j and m, we wish to find the manner in which it converges. To accomplish both of these tasks, we solve the for all of the eigenvalues and eigenvectors of the transition matrix. This transition matrix is found by rewriting equation 2.9 in matrix form:

$$\begin{bmatrix} \ddots & \ddots & \ddots & & & \\ & p_{i-1} & 1-2p_i & p_{i+1} & & \\ & & \ddots & \ddots & \ddots & \end{bmatrix} \begin{bmatrix} \vdots \\ a_i^{(m)} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ a_i^{(m+1)} \\ \vdots \end{bmatrix} \tag{2.10}$$

We express equation 2.10 more succinctly by $\mathbf{Ta}^{(m)} = \mathbf{a}^{(m+1)}$, where $\mathbf{T}$ is the transition matrix, and $\mathbf{a}^{(m)}$ is a vector that takes components $a_i^{(m)}$. Note that using repeated application of this formula gives

$$\mathbf{a}^{(m)} = \mathbf{T}^m \mathbf{a}^{(0)}. \tag{2.11}$$

In order to calculate arbitrary powers of a large matrix $\mathbf{T}$, we turn to the eigenvalues and eigenvectors of the matrix $\mathbf{T}$. To solve for these, we use similar generating function techniques as above.

2.3 Solution of the Spectral Problem by Generating Functions

In order to calculate arbitrary powers of a large matrix, we calculate all eigenvalues and eigenvectors exactly. To do this, we express the spectral problem in generating function form, and proceed to solve it explicitly. The eigenvalue problem can be stated as $\mathbf{Tv} = \lambda\mathbf{v}$, where the vector $\mathbf{v}$ takes components c_i. It is more convenient to express this problem in component form instead, which is given by

$$(\lambda - 1)c_i = p_{i-1}c_{i-1} - 2p_ic_i + p_{i+1}c_{i+1}. \tag{2.12}$$

Now we rewrite this in terms of a generating function $G(x, y)$ for c_i that we define by

$$G(x, y) = \sum_{i=0}^{N} c_i x^i y^{N-i}. \tag{2.13}$$

Now, with generating functions of this type, we make two observations:

1. If c_i is generated by G, then p_ic_i is generated by $\dfrac{xyG_{xy}}{N(N-1)}$

2. If c_i is generated by G, then c_{i-1} is generated by $\frac{x}{y}G$

3. If c_i is generated by G, then c_{i+1} is generated by $\frac{y}{x}G$

We use item 1 to generate $p_i c_i$ and then use items 2 and 3 to shift it appropriately to express equation 2.12 in terms of G. Doing so gives a very succinct and insightful expression for G:

$$N(N-1)(\lambda-1)G = (x-y)^2 G_{xy} \tag{2.14}$$

The form of this equation leads us to an natural choice for a change of variables. Let $u = x - y$ and $G(x,y) = H(u,y)$. We now show that H has the same form as G:

$$\begin{aligned}
H(u,y) &= \sum_{i=0}^{N} c_i (u+y)^i y^{N-i} &(2.15)\\
&= \sum_{i=0}^{N} \sum_{j=0}^{i} \binom{i}{j} c_i u^j y^{N-j} &(2.16)\\
&= \sum_{j=0}^{N} \sum_{i=j}^{N} \binom{i}{j} c_i u^j y^{N-j} &(2.17)\\
&= \sum_{i=0}^{N} \left[\sum_{j=i}^{N} \binom{j}{i} c_j \right] u^i y^{N-i} &(2.18)\\
&= \sum_{i=0}^{N} b_i u^i y^{N-i}. &(2.19)
\end{aligned}$$

Here, we define b_i to be the coefficients of H, which are given by

$$b_i = \sum_{j=i}^{N} \binom{j}{i} c_j. \tag{2.20}$$

Now, taking $u = x - y$ and rewriting equation 2.14 in terms of u and H gives

$$N(N-1)(\lambda-1)H = u^2(H_{uy} - H_{uu}). \tag{2.21}$$

Taking this and applying equation 2.19, we obtain a recursion relation for b_i:

$$N(N-1)(\lambda-1)b_i = (i-1)(N-i+1)b_{i-1} - i(i-1)b_i. \tag{2.22}$$

We can now express b_i explicitly in terms of b_{i-1} to create a recursion relation for b_i:

$$b_i = \frac{(i-1)(N-i+1)}{N(N-1)(\lambda-1)+i(i-1)} b_{i-1}. \tag{2.23}$$

We use this recursion relation to find the complete set of eigenvalues. Equation 2.19 defines a homogeneous polynomial in two variables (u, y). Because there are no negative powers, we require $b_i = 0$ when $i < 0$ regardless of λ. Using 2.23, we conclude therefore that $b_i = 0$ for every i, as long as there is no singularity in 2.23 for some $i = k$. This yields the trivial solution to the eigenvalue problem $H(u, y) = G(x, y) = 0$, which we discard. Therefore, to obtain non-trivial solutions to the eigenvalue problem, we demand that for some $i = k$, there exists a singularity in 2.23. This occurs when

$$N(N-1)(\lambda-1) + k(k-1) = 0. \tag{2.24}$$

Therefore, the complete set of eigenvalues of the transition matrix for the voter model on the complete graph are:

$$\lambda_k = 1 - \frac{k(k-1)}{N(N-1)}, \quad k = 0, \ldots, N \tag{2.25}$$

This implies that b_k can take any value. For convenience, set $b_k = 1$, and apply 2.23 to find every other b_i. Note that doing this will guarantee that $b_i = 0$ when $i > N$, which is also required for H to be of the form given in equation 2.19. This yields explicit solutions for b_i for a given λ from the set given in 2.25. Now, given an explicit solution for H, we obtain an explicit solution for G by substitution:

$$G(x,y) = \sum_{i=0}^{N} b_i (x-y)^i y^{N-i} \tag{2.26}$$

$$= \sum_{i=0}^{N} \sum_{j=0}^{i} \binom{i}{j} (-1)^{j-i} b_i x^j y^{N-j} \tag{2.27}$$

$$= \sum_{j=0}^{N} \sum_{i=j}^{N} \binom{i}{j} (-1)^{j-i} b_i x^j y^{N-j} \tag{2.28}$$

$$= \sum_{i=0}^{N} \left[\sum_{j=i}^{N} \binom{j}{i} (-1)^{i-j} b_j \right] x^i y^{N-i} \tag{2.29}$$

By the definition of G from equation 2.13, we have that

$$c_i = \sum_{j=i}^{N} \binom{j}{i} (-1)^{i-j} b_j \tag{2.30}$$

which gives the explicit solution for the corresponding eigenvector.

The solution for the eigenvectors given by 2.30 is exact, but the qualitative features of the eigenvectors is not obvious from this equation. To study the eigenvectors in more detail for large N, we turn to the differential eigenvalue problem, with the above solution in mind.

Note that if we divide equation 2.12 by $1/N^2$, we have

$$-k(k-1)c_i \sim \frac{\Delta_i^2 (p_i c_i)}{1/N^2} \tag{2.31}$$

where Δ_i^2 is the second centered difference operator. For large N, the second difference on the right side of equation 2.31 approaches the second derivative. Let $i/N \to x$, $\Delta x = 1/N$, and $u_k(x) \to c_i$. Then, equation 2.31 becomes

$$-k(k-1)u_k \sim \frac{d^2}{dx^2}[x(1-x)u_k]. \tag{2.32}$$

Expanding all derivatives gives the second order differential equation for u_k:

$$x(1-x)\frac{d^2u_k}{dx^2} + (2-4x)\frac{du_k}{dx} + (k(k-1)-2)u_k \sim 0. \tag{2.33}$$

This is a case of the hypergeometric differential equation [28, 29], whose solution is given in terms of the hypergeometric function ${}_2F_1$ as

$$u_k(x) \sim {}_2F_1(k+1, 2-k; 2; x). \tag{2.34}$$

It happens to be that the series for ${}_2F_1(k+1, 2-k; 2; x)$ terminates, giving a degree $k-2$ polynomial for the eigenvector components in the interior. It should be noted that for both $k=0$ and $k=1$, we have that $\lambda = 1$. These eigenvectors correspond to the consensus state in which all of the balls are in either urn A ($k=0$) or urn B ($k=1$). The eigenvalue 1 always corresponds to a stationary distribution, which is the case for both consensus states. More interestingly, for the second largest eigenvalue, λ_2, we have that the solution is $u_k \sim 1$. This implies that the uniform distribution is an eigenvector, and that for large time, the probability distribution approaches uniformity. In Fig. 2.1, we calculate the exact eigenvector components and compare them to the hypergeometric solution given in equation 2.34.

This solution not only allows us to efficiently calculate large powers of the propagator, it shows that the voter model on the complete graph is very diffusive in nature. The probability distribution tends to the uniform distribution for large time. Then at a rate of $1 - \frac{2}{N(N-1)}$, this uniform distribution tends to consensus. Now that eigenvalues and eigenvectors are known, we now use them to find other exact solutions.

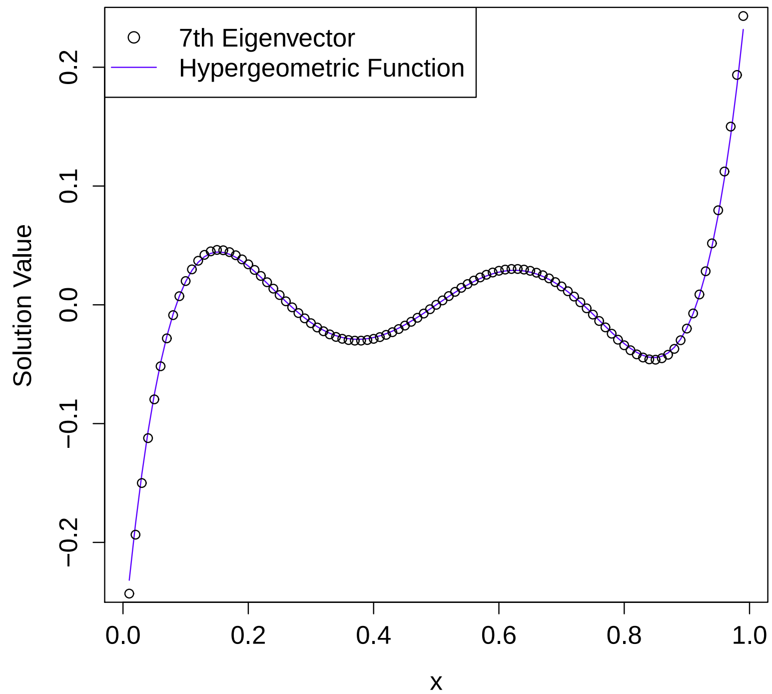

Figure 2.1: 7th Eigenvector of the discrete problem plotted with the exact solution for the limit as $N \to \infty$. The hypergeometric function in the figure is a fifth degree polynomial.

2.4 Exact Solutions of the Voter Model by Spectral Analysis

As described above, the entire urn process (or network model), is described in terms of a single macrostate quantity n_A that changes randomly with time. All future probability distributions depend on arbitrary powers of the single step Markov transition matrix. Not only does the solution to the spectral problem allow us to calculate these exactly, we can also find other quantities of interest in terms of the spectral solution. In particular, we will calculate all moments of consensus time (mean, variance, etc.), then we consider local times. The consensus time is the amount of scaled time (m/N) until all of the balls are in the same urn. The local times is the amount of scaled time spent at each macrostate prior to consensus.

2.4.1 Moments of Consensus Time

We wish to calculate not only the expected time until consensus is reached, but all moments thereof. In sociophysics [4], the consensus time is a quantity of great interest [9,11,15,21,22,30]. Other methods for calculating this rely on methods such as continuous time first step analysis and differential equations [11, 21]. For such methods to hold, it is assumed that N is large. However, these methods fail when N is small, can only solve for the first moment, and is only an asymptotic solution. The solution that is provided here is more robust in that every N is feasible, and every moment can be calculated.

The consensus time, τ, is the amount of scaled time (m/N) until all of the balls are in the same urn. We use the series expression for the *pth* moment of consensus time to calculate this. Let q_m be the probability that the system achieves consensus at discrete time m. Then, the *pth* moment of the consensus time is given by

$$E[\tau^p] = \sum_{m=1}^{\infty} q_m \left(\frac{m}{N}\right)^p \tag{2.35}$$

Now, q_m is given by

$$q_m = \frac{1}{N}(a_1^{(m-1)} + a_{N-1}^{(m-1)}). \tag{2.36}$$

We use the solution to the spectral problem to find $a_1^{(m-1)}$ and $a_{N-1}^{(m-1)}$. By spectral decomposition, we can write the initial probability distribution $\mathbf{a}^{(0)}$ as

$$\mathbf{a}^{(0)} = \sum_{k=2}^{N} d_k \mathbf{v}_k \tag{2.37}$$

for eigenvector $\mathbf{v}_k$ corresponding to eigenvalue λ_k. We take the sum starting from $k = 2$ because the eigenvectors that correspond to $k = 0, 1$ are the consensus states. We assume that the system does not start in consensus, so $d_0 = d_1 = 0$. The constants d_k depend on the initial distribution, and a comprehensive means of calculating these is given in chapter 3. Now, since the future probability distributions are simply powers of the transition matrix multiplied into $\mathbf{a}^(0)$, we have

$$\mathbf{a}^{(m-1)} = \sum_{k=2}^{N} d_k \lambda_k^{m-1} \mathbf{v}_k \tag{2.38}$$

Therefore, if we let $s_k = d_k([\mathbf{v}_k]_1 + [\mathbf{v}_k]_{N-1})$, we have

$$a_1^{(m-1)} + a_{N-1}^{(m-1)} = \sum_{k=2}^{N} s_k \lambda_k^{m-1} \tag{2.39}$$

Now we have that the *pth* moment of consensus time is given by

$$\begin{aligned} E[\tau^p] &= \sum_{m=1}^{\infty} \frac{1}{N} \sum_{k=2}^{N} s_k \lambda_k^{m-1} \left(\frac{m}{N}\right)^p && (2.40) \\ &\sim \sum_{k=2}^{N} \frac{1}{N^{p+1}} s_k \frac{p!}{(1-\lambda_k)^{p+1}} && (2.41) \\ &= p!(N-1)^{p+1} \sum_{k=2}^{N} \frac{s_k}{[k(k-1)]^{p+1}}. && (2.42) \end{aligned}$$

This is the exact solution for the moments of consensus time. We can estimate this fairly easily by noting that the probability distribution is dominated by the second largest eigenvalue, whose eigenvector is a constant. Let $S(m)$ be the probability that the system has not reached consensus by time m. Now, estimating this by the second largest eigenvalue, we have that $S(m) = O(\lambda_2^m)$. We now use this to estimate q_m. Note that the probability of reaching consensus at time m is the probability that the system is in consensus $(1 - S(m))$ minus the probability that the system was already in consensus $(1 - S(m-1))$. Therefore, we have

$$q_m = S(m-1) - S(m) = O\left(\lambda_2^{m-1}(1-\lambda_2)\right). \tag{2.43}$$

Using this, the moments of consensus time can be bounded as follows:

$$E[\tau^p] = \sum_{m=0}^{\infty} q_m \left(\frac{m}{N}\right)^p \tag{2.44}$$

$$= \frac{1-\lambda_2}{\lambda_2 N^p} \sum_{m=0}^{\infty} O(\lambda_2^m) m^p \tag{2.45}$$

$$\sim \frac{1-\lambda_2}{\lambda_2 N^p} O\left(\frac{p!\lambda_2^{p+1}}{(1-\lambda_2)^{p+1}}\right) \tag{2.46}$$

$$= O\left(p!\left[\frac{\lambda_2}{(1-\lambda_2)N}\right]^p\right). \tag{2.47}$$

This shows that the expected time to consensus ($p = 1$) is $O(N)$, the second moment of consensus time ($p = 2$) is $O(N^2)$, etc. The expected time to consensus is consistent with previous research [11, 21, 31], but the higher moments are new developments.

2.4.2 Expected Local Times

The local time of a macrostate $n_A = i$ is the total amount of time that the system spends at this macrostate prior to reaching consensus. Since the local time may be different for various macrostates, we arrange the local times into a vector, $\mathbf{M}$, whose *ith* component is the local time for $n_A = i$. We use the solution to the spectral problem to calculate the expected value of $\mathbf{M}$.

We begin by letting $M_i(m)$ be the total amount of time spent at macrostate $n_A = i$ by time m. Also, let $M_i = \lim_{m\to\infty} M_i(m)$. Let $\Delta M_i(m) = M_i(m+1) - M_i(m)$ denote the change in the local time in a single step. Now, the local time for $n_A = i$ will increase by $1/N$ if the system is in this macrostate at time $m+1$. This occurs with probability $a_i^{(m+1)}$. Otherwise, the local time for $n_A = i$ does not change. Therefore, $E[\Delta M_i(m)] = a_i^{(m+1)}/N$. If we sum $E[\Delta M_i(m)]$ from $m = 0$ to $m \to \infty$, the series is telescopic, so we have

$$E[M_i] - E[M_i(0)] = \frac{1}{N} \sum_{m=0}^{\infty} a_i^{(m+1)}, \tag{2.48}$$

which implies that

$$E[M_i] = \frac{1}{N} \sum_{m=0}^{\infty} a_i^{(m)} \tag{2.49}$$

Since we have control over $a_i^{(m)}$ by the solution of the spectral problem, we can find closed solution to this infinite series expression. We express the probability distribution $\mathbf{a}$ in terms of the spectral decomposition given in equation 2.38, and obtain the solution for the expected local time as

$$E[\mathbf{M}] = \frac{1}{N} \sum_{m=0}^{\infty} \sum_{k=2}^{N} d_k \lambda_k^m \mathbf{v}_k \tag{2.50}$$

$$= \frac{1}{N} \sum_{k=2}^{N} \frac{d_k}{1-\lambda_k} \mathbf{v}_k \tag{2.51}$$

$$= (N-1) \sum_{k=2}^{N} \frac{d_k}{k(k-1)} \mathbf{v}_k \tag{2.52}$$

The first and last components of $E[\mathbf{M}]$ should be discarded, since they correspond to consensus states, and we only concern ourselves with macrostate that are not consensus states. Every other component of $E[\mathbf{M}]$ corresponds to the exact solution to the expected local time. This solution is exact for any initial distribution $\mathbf{a}^{(0)}$.

Here, we examine a continuous time treatment of the local time problem in order to find the behavior for large N. To do this, let $\rho_i = i/N$, $t_m = m/N$, and $u(\rho_i, t_m) = N a_i^{(m)}$. The quantity ρ_i is interpreted as the density of urn A, and u is a probability density function. In this notation, as $N \to \infty$, we have that the Fokker-Plank equation is

$$\frac{\partial u}{\partial t} \sim \frac{1}{N} \frac{\partial^2}{\partial \rho^2} [\rho(1-\rho)u]. \tag{2.53}$$

We now take the continuous time limit of the infinite series for local time given in equation 2.49 to obtain

$$M(\rho) = \frac{1}{N} \sum_{m=0}^{\infty} a_i^{(m)} \to \frac{1}{N} \int_0^{\infty} u(\rho, t) dt. \tag{2.54}$$

Just as we calculated the infinite sum in the discrete problem for all discrete times

m, we now integrate the Fokker-Plank equation given by equation 2.53 to obtain

$$-u(\rho, 0) = \frac{d^2}{d\rho^2}[\rho(1-\rho)M]. \tag{2.55}$$

Since the system reaches consensus in finite time, and the probability distribution $a_i^{(m)} \to 0$ as $m \to \infty$, the upper limit of the integrand on the left side vanishes. Now, let $f(\rho) = u(\rho, 0)$ be the given initial distribution of the system, and let $T(\rho) = \rho(1-\rho)M(\rho)$. We now wish to solve

$$-f(\rho) = \frac{d^2T}{d\rho^2}. \tag{2.56}$$

From the definition of T, we have the boundary conditions $T(0) = T(1) = 0$. We solve this by applying the Green's function of this problem [28], which is given by

$$\tilde{g}(\rho, \xi) = \begin{cases} \rho(1-\xi), & \rho < \xi \\ \xi(1-\rho), & \rho > \xi \end{cases}. \tag{2.57}$$

Therefore, we can write the solution for T as

$$T(\rho) = \int_0^1 f(\xi)\tilde{g}(\rho, \xi)d\xi. \tag{2.58}$$

This yields the solution for $M(\rho)$ as

$$M(\rho) = \int_0^1 f(\xi)g(\rho, \xi)d\xi \tag{2.59}$$

where

$$g(\rho, \xi) = \begin{cases} \dfrac{1-\xi}{1-\rho}, & \rho < \xi \\ \\ \dfrac{\xi}{\rho}, & \rho > \xi \end{cases}. \tag{2.60}$$

There are a few noteworthy properties of this solution. Firstly, if $f(\rho) = \delta(\rho - \rho_0)$ for some ρ_0, this distribution corresponds to $a_i = 1$ for $i/N = \rho_0$ and $a_i = 0$ elsewhere. That is, the system always starts with $n_A = N\rho_0$. If we assume this distribution, then the solution for the local times is simply

$$M(\rho) = g(\rho, \rho_0). \tag{2.61}$$

Another interesting initial condition is the uniform probability distribution function. This is $f(\rho) = 1$. Recall that this distribution is an eigenvector of the model, which implies that in the discrete solution, $d_2 = \frac{1}{N+1}$ when $\mathbf{v}_2 = \mathbf{1}$, and $d_k = 0$ otherwise. So, according to the discrete solution, the expected local time is also uniform, with time $\frac{N-1}{2(N+1)} \sim \frac{1}{2}$. Now, taking $f(\rho) = 1$ in the Green's function solution, we have that $\mathbf{M}(\rho) = 1/2$, which is also uniform, and the asymptotic value of the discrete solution. Figure 2.2 compares simulations of the voter model to the exact solution for the expected local times.

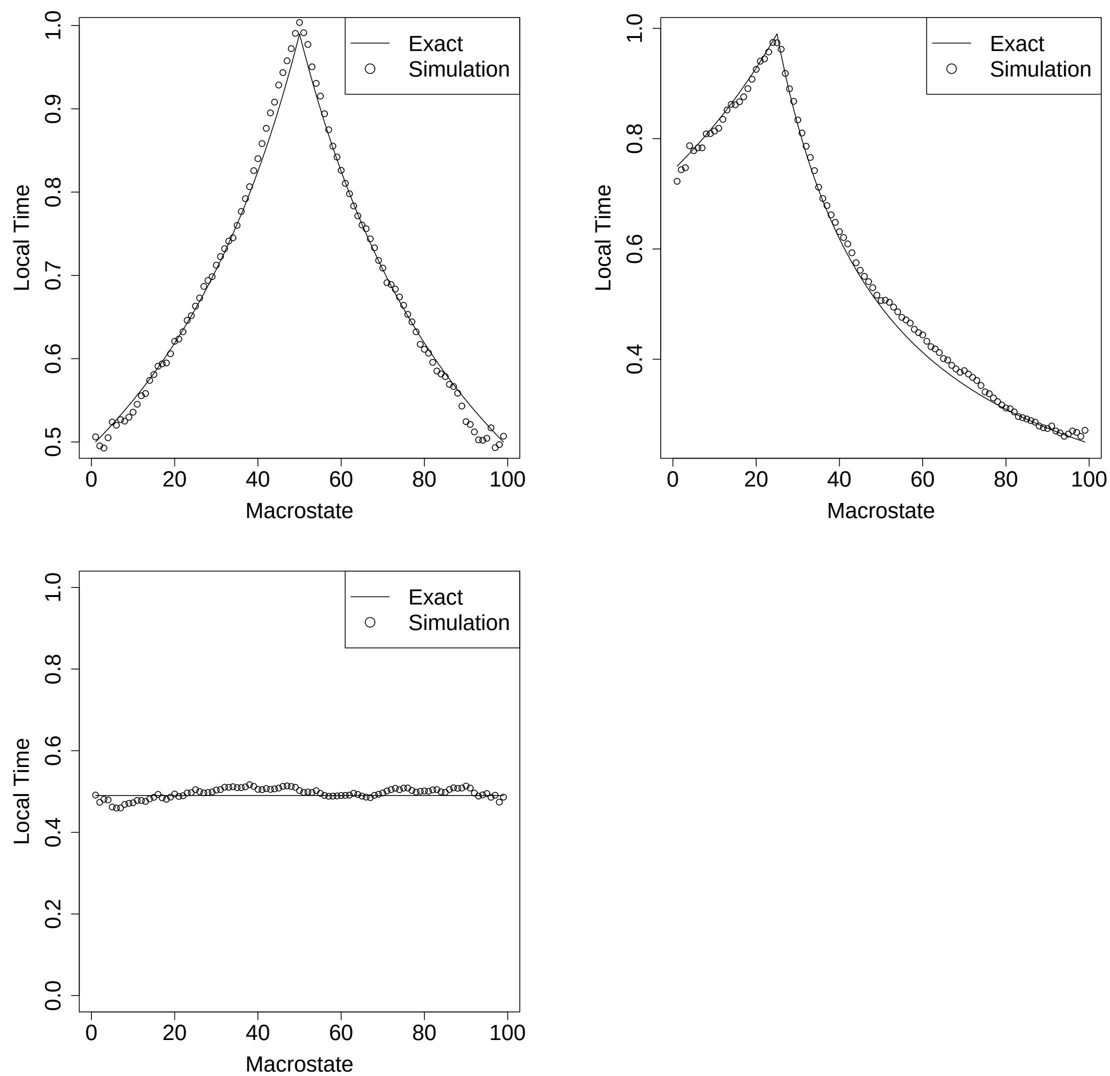

Figure 2.2: Local times are examined when $n_A(0) = N/2$ (top), $n_A(0) = N/4$ (middle), and the uniform initial distribution (bottom). The exact solution is compared with Monte Carlo simulations over 3000 runs of the Voter model for $N = 100$. Note the similarities to the Green's function solution for these cases.

CHAPTER 3
Exact Solution of the 2-Urn Models

Here we define the 2-Urn model and proceed to solve it exactly by similar generating function methods. The method of solution is a generalization of the generating function method of Mark Kac [32] to solve the Ehrenfest urn model [33]. This solution reconciles notions of increasing thermodynamic entropy with time reversibility by exactly diagonalizing the transition matrix. The Ehrenfest model chooses a random ball to change urns in a single step. On the other hand, the 2-urn models that we discuss here are more general than this, since we examine the case when two particles interact with each other. Interestingly enough, even though we choose two random balls from the urn, the Ehrenfest model exists as a given parameter configuration of the 2-urn models. In addition, the voter model and the Moran model with mutation are both cases of the 2-urn process. The method of solution is similar to the solution to the voter model by generating functions, given in chapter 2. Also, the 2-urn models have a natural network extension that we explore in chapter 4.

3.1 Linear 2-Urn Models and Solution

In this section, we will define the 2-urn models and then proceed to solve them by generating functions. The general 2-urn model is specified by 6 parameters that represent transition probabilities. However, the generating function solution for the eigenvalues and eigenvectors only holds when one equality constraint is satisfied. We call this constraint the *linearity constraint*, and the subclass of urn models in which this constraint holds are called *linear urn models*. For the linear urn models, we can diagonalize them exactly to calculate the m-step propagator as well as other solutions. For a six parameter system, with one equality constraint, there are effectively five free parameters that define a linear urn model. The *nonlinear* cases are studied in later sections.

We define the model as follows. For any 2-urn model, there are two urns (A and B) with N balls distributed between them (as in the voter model). In a

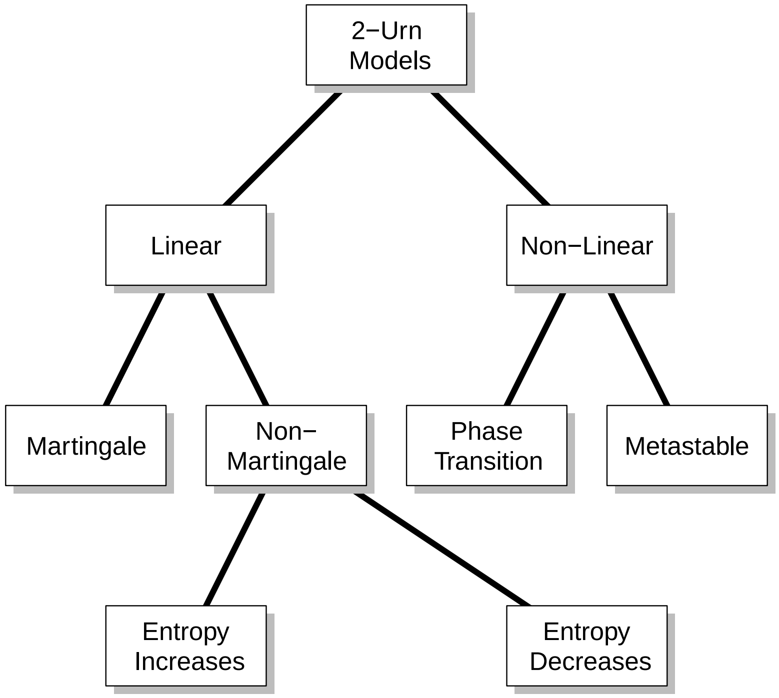

Figure 3.1: Classification tree for the 2-Urn problems that shows all relevant subclasses. Among the linear cases $(2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0)$**, the martingales** $(\alpha_1 = \beta_1, \alpha_2 = \beta_2 = \gamma_1 = \gamma_2 = 0)$ **are equivalent to the voter model and the non-martingales constitute a much larger class of models. We show that an entropy increase in the non-martingale cases is equivalent to macroscopic symmetry of the model given by equation** (3.41)**. In the nonlinear models, there exists a phase transition over a single parameter that separates metastable convergence from logarithmic convergence.**

discrete time step, two balls are drawn randomly from the urn, one after the other. The balls are then redistributed between the urns stochastically, with probabilities that only depend on which urns the balls were drawn from and the order that they were drawn.

The definition of these models is very simple, yet they offer many interesting and rich properties. A categorization of the relevant cases that we discuss is given in figure 3.1.

Let $n_A(m)$ be the number of balls in urn A at discrete time m. Each model is characterized through six rate parameters, which we denote by $\{\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2\}$. These parameters represent the redistribution probabilities after the balls were selected. These parameters are given to be

$$\alpha_1 = Pr\{\Delta n_A = 1|AB\} + Pr\{\Delta n_A = 1|BA\} \tag{3.1}$$

$$\alpha_2 = Pr\{\Delta n_A = 1|BB\} \tag{3.2}$$

$$\beta_1 = Pr\{\Delta n_A = -1|AB\} + Pr\{\Delta n_A = -1|BA\} \tag{3.3}$$

$$\beta_2 = Pr\{\Delta n_A = -1|AA\} \tag{3.4}$$

$$\gamma_1 = Pr\{\Delta n_A = 2|BB\} \tag{3.5}$$

$$\gamma_2 = Pr\{\Delta n_A = -2|AA\}. \tag{3.6}$$

Since these parameters correspond to probabilities, all parameters must be non-negative, $\alpha_1 + \beta_1 \leq 2$, $\alpha_2 + \gamma_1 \leq 1$, and $\beta_2 + \gamma_2 \leq 1$. If one were to use negative parameters (negative probabilities), the resulting urn model would not be physically viable. However, interestingly enough, one could still use these techniques to diagonalize the formal transition matrix that it would imply.

The parameters of the urn model may be interpreted as the influence of people in social settings. Values of α_1 and β_1 correspond to the impact a person has on another with the opposite opinion. An individual may respond to peer-pressure, or be influenced by a coherent argument from the other individual. This is the archetype for the voter model, which has parameter configuration $\{1, 0, 1, 0, 0, 0\}$. The other parameters, $\alpha_2, \beta_2, \gamma_1$, and γ_2 can correspond to mutation and competition between individuals. Also, these parameters can represent push-pull factors to Lee's model of migration [34], and quadratic transition probabilities reflect the assumptions made in Gravity models of migration and trade [35, 36]. Existing models with explicit parameter configurations also include the Moran model of genetic drift, which has parameters $\{1 - \mu_1, \mu_2, 1 - \mu_2, \mu_1, 0, 0\}$, where μ_1 and μ_2 are mutation probabilities [16, 17]. For these models, the population size, N, is not always large and thus the discrete stochastic treatment we provide is necessary.

With these parameters, we can express the transition probabilities for the corresponding random walk by

$$p_i^{(1)} = \alpha_1 \frac{i(N-i)}{N(N-1)} + \alpha_2 \frac{(N-i)(N-i-1)}{N(N-1)} \tag{3.7}$$

$$p_i^{(2)} = \gamma_1 \frac{(N-i)(N-i-1)}{N(N-1)} \tag{3.8}$$

$$q_i^{(1)} = \beta_1 \frac{i(N-i)}{N(N-1)} + \beta_2 \frac{i(i-1)}{N(N-1)} \tag{3.9}$$

$$q_i^{(2)} = \gamma_2 \frac{i(i-1)}{N(N-1)}. \tag{3.10}$$

Here, we define $p_i^{(k)} = Pr\{\Delta n_A = k | n_A = i\}$ and $q_i^{(k)} = Pr\{\Delta n_A = -k | n_A = i\}$. Note that if we consider the parameter configuration $\{1,1,1,1,0,0\}$, the model exactly simplifies to the Ehrenfest urn model.

Now we express the single step propagator of the 2-urn models. From this, we will define the Markov transition matrix, and then use generating functions to diagonalize it. Let $a_i^{(m)} = Pr\{n_A(m) = i\}$. Also let Δ_{ki} be defined such that $\Delta_{ki}[\phi_i] = \phi_{i+k} - \phi_i$ for a given grid function ϕ_i. Using this, we form the single step propagator for the 2-urn models:

$$\begin{aligned} a_i^{(m+1)} - a_i^{(m)} = \Delta_{-1i}[p_i^{(1)} a_i^{(m)}] + \Delta_{-2i}[p_i^{(2)} a_i^{(m)}] \\ + \Delta_{+1i}[q_i^{(1)} a_i^{(m)}] + \Delta_{+2i}[q_i^{(2)} a_i^{(m)}]. \end{aligned} \tag{3.11}$$

3.1.1 Generating Function Solution of the Spectral Problem

The transition probabilities correspond to elements in a pentadiagonal Markov transition matrix. We solve for all eigenvalues and eigenvectors of this model by extending the procedure chapter 2. For eigenvalue λ and eigenvector $\mathbf{v}$ with components c_i, let $G(x,y) = \sum_i c_i x^i y^{N-i}$ be the generating function for the eigenvectors. We rewrite the spectral problem for the single step propagator given in equation 3.11 as a partial differential equation for G using the differentiation and shift properties of G [15, 25–27]. Following the methods in chapter 2, the PDE for G is

$$
\begin{aligned}
N(N-1)(\lambda-1)G = \gamma_1(x^2-y^2)G_{yy} + \alpha_1 x(x-y)G_{xy} \\
+ \alpha_2 y(x-y)G_{yy} - \gamma_2(x^2-y^2)G_{xx} - \beta_1 y(x-y)G_{xy} \\
- \beta_2 x(x-y)G_{xx}. \quad (3.12)
\end{aligned}
$$

Just with the voter model, we let $u = x - y$ and $H(u, y) = G(x, y)$. As was shown in chapter 2, this change of variables allows us to define $H(u, y) = \sum_i b_i u^i y^{N-i}$. We make this change of variables because we can solve for b_i and λ exactly when

$$2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0, \quad (3.13)$$

which we call the *linearity constraint*. If this constraint does not hold, then there exist uyH_{uu} terms in the resulting transformed PDE, which produces super-diagonal terms. Only a sub-diagonal terms are allowed for the method to generate all eigenvalues and eigenvectors explicitly. The relationship between this constraint and linear systems is shown below. Now, with this change of variables, and assuming that the linearity constraint holds, the PDE for H is given as:

$$
\begin{aligned}
N(N-1)(\lambda-1)H = [\gamma_1 u^2 + (2\gamma_1 + \alpha_2)uy]H_{yy} \\
+ [(\alpha_1 - 2\gamma_1)u^2 + (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)uy]H_{uy} \\
+ (\gamma_1 - \alpha_1 - \gamma_2 - \beta_2)u^2 H_{uu}. \quad (3.14)
\end{aligned}
$$

This is equivalent to a recursion relation for b_i, which is given to be

$$b_i = \frac{A_{i-2}b_{i-2} + B_{i-1}b_{i-1}}{C_i} \quad (3.15)$$

where

$$A_i = \gamma_1(N-i)(N-i-1) \quad (3.16)$$

$$B_i = [(-2\gamma_1 + \alpha_1)i + (2\gamma_1 + \alpha_2)(N-i-1)](N-i) \quad (3.17)$$

and

$$C_i = N(N-1)(\lambda - 1) - (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)i(N-i) \\ - (\gamma_1 - \alpha_1 - \gamma_2 - \beta_2)i(i-1) \quad (3.18)$$

This allows us to find all eigenvalues exactly. Similar to the argument in chapter 2, we require $b_i = 0$ for $i < 0$ and $i > N$ as well. The difference equation given in 3.15 therefore implies that every b_i=0 unless there is a singularity at $i = k$. Since this corresponds to the trivial solution to the problem, we set the denominator of equation 3.15 to zero when $i = k$. Solving for λ shows that the eigenvalues are

$$\lambda_k = 1 - \frac{(-\alpha_1 + 4\gamma_1 + 2\alpha_2 + \beta_1)k(N-k)}{N(N-1)} - \frac{(-\gamma_1 + \alpha_1 + \gamma_2 + \beta_2)k(k-1)}{N(N-1)} \quad (3.19)$$

for $k = 0 \ldots N$. Since there is a singularity at $i = k$, b_k can take any value. This represents the fact that any constant multiplied by an eigenvector is still an eigenvector. We now apply equation 3.15 iteratively to find all b_i for $i > k$. This provides the means for finding b_i, and therefore, an explicit representation for $H(u, y)$. We obtain the generating function G by substitution to find

$$G(x, y) = \sum_{i=0}^{N} \sum_{j=i}^{N} (-1)^{j-i} \binom{j}{i} b_j x^i y^{N-i}. \quad (3.20)$$

Therefore, all of the components of the eigenvectors can be explicitly calculate by these means, so long as $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0$. This constraint is related to the continuous time representation of the 2-urn model. Let $\bar{\rho} = E[n_A/N]$. Using the transition probabilities, we find that the expected change in the macrostate is given by

$$E[\Delta n_A | n_A = i] = 2p_i^{(2)} + p_i^{(1)} - q_i^{(1)} - 2q_i^{(2)}. \quad (3.21)$$

For $\Delta t = 1/N$, we have that $\Delta n_A = \Delta\rho/\Delta t$, and a differential equation for the expected value of ρ can be written as

$$\frac{d\bar{\rho}}{dt} = 2\gamma_1 + \alpha_2 + (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)\bar{\rho} \\ + (2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2)\bar{\rho}^2. \quad (3.22)$$

Notice that this ODE is linear if and only if $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0$, which is the solvability condition of the generating function method. If this condition is not satisfied, there will be a b_{i+1} term in 3.15, and the method fails. Furthermore, 3.13 is not satisfied, then 3.22 is not a linear ODE, and we call the urn model *nonlinear*. We leave as a conjecture that the nonlinear cases cannon be solved by these generating functions for a different change of variables $(x, y) \to (u, v)$. This is the first major categorical distinction shown in Fig. 3.1.

3.1.2 Generating Function Method as a Similarity Transform

Since we use a linear change of variables on a polynomial generating function, method is equivalent to a well constructed similarity transformation. The generating function method proves that the similarity transform yields the spectral decomposition of the original transition matrix. Let $\mathbf{T}$ be the transition matrix implied by equation 3.11. So, the eigenvalue problem is given as $\mathbf{Tv} = \lambda\mathbf{v}$. The vector $\mathbf{v}$ takes components c_i. Now the change of variables is represented by the substitution $\mathbf{w} = \mathbf{Pv}$ for some given invertible matrix $\mathbf{P}$. Making this substitution shows that the eigenvalue problem for $\mathbf{w}$ is given by $\mathbf{PTP^{-1}w} = \lambda\mathbf{w}$. We let the components of $\mathbf{w}$ be b_i, which are the same b_i when using equation 3.15. The generating function method prescribes the matrix $\mathbf{P}$ so that the new matrix $\mathbf{L} = \mathbf{PTP^{-1}}$ is lower triangular with a bandwidth of at most two. As shown by the calculations in chapter 2, equations 2.20 and 2.30, The components of the transformation matrices that perform this triangularization are determined to be

$$[\mathbf{P}]_{ij} = \binom{j}{i} \quad (3.23)$$

$$[\mathbf{P}^{-1}]_{ij} = (-1)^{j-i}\binom{j}{i}. \quad (3.24)$$

We use the convention that $\binom{j}{i} = 0$ when $i > j$, which suggest that $\mathbf{P}$ and $\mathbf{P}^{-1}$ are upper triangular. The matrix $\mathbf{P}$ is called the upper Pascal matrix, since the columns of $\mathbf{P}$ are the rows of Pascal's Triangle [37]. Since $\mathbf{T} = \mathbf{P}^{-1}\mathbf{L}\mathbf{P}$, we have shown that the lower triangular matrix $\mathbf{L}$ expressed in a new basis given by the rows of Pascal's triangle yields the original Markov transition matrix.

The explicit solution for the spectral decomposition can be found by the generating function method. We do this by diagonalizing the matrix $\mathbf{L} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{-1}$. Here, $\boldsymbol{\Lambda} = diag(\lambda_0, \ldots, \lambda_N)$ given above and $\mathbf{W}$ are the eigenvectors of $\mathbf{L}$. The kth column of $\mathbf{W}$ are the b_i that are calculated using 3.15 corresponding to λ_k. Since $b_i = 0$ for $i < j$, $\mathbf{W}$ is lower triangular. Therefore, since $\mathbf{W}$ is a triangular matrix, $\mathbf{W}^{-1}$ can be found explicitly via forward substitution. This gives an explicit result for $\mathbf{L} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{-1}$. Now, since $\mathbf{T} = \mathbf{P}^{-1}\mathbf{L}\mathbf{P}$, we have

$$\mathbf{T} = \mathbf{P}^{-1}\mathbf{W}\Lambda\mathbf{W}^{-1}\mathbf{P}. \tag{3.25}$$

Which yields the spectral decomposition of the matrix $\mathbf{T}$, given by

$$\mathbf{T} = \mathbf{S}\Lambda\mathbf{S}^{-1} \tag{3.26}$$

where the matrix $\mathbf{S}$ contain the eigenvectors of $\mathbf{T}$. Both $\mathbf{S}$ and $\mathbf{S}^{-1}$ are given in terms of $\mathbf{P}$, $\mathbf{W}$, and their inverses found above by

$$\mathbf{S} = \mathbf{P}^{-1}\mathbf{W} \tag{3.27}$$

$$\mathbf{S}^{-1} = \mathbf{W}^{-1}\mathbf{P} \tag{3.28}$$

So, in order to find the initial distribution expressed in the eigenbasis of $\mathbf{T}$, we would calculate

$$\mathbf{d} = \mathbf{W}^{-1}\mathbf{P}\mathbf{a}^{(0)}. \tag{3.29}$$

The components of the vector $\mathbf{d}$ are the values of d_k that are used throughout this thesis when calculating various solutions.

Table 3.1: Summary of Exact Solutions

Quantity	Discrete Solution
Macro-state probability	$a_i^{(m)} = \sum_{k=0}^{N} d_k [\mathbf{v}_k]_i \lambda_k^m$
Consensus time	$E[\tau^p] \sim \sum_{k=1}^{N} \frac{d_k p!}{[N(1-\lambda_k)]^{p+1}} \times \left\{\beta_1 [\mathbf{v}_k]_1 + \frac{2\gamma_2}{N-1}[\mathbf{v}_k]_2\right\}$
Local time	$E[\mathbf{M}] \sim \frac{1}{N} \sum_{\substack{k=0 \\ \lambda_k \neq 1}}^{N} \frac{d_k}{1-\lambda_k} \mathbf{v}_k$
Shannon entropy	$S = -\bar{\rho}\log_2 \bar{\rho} - (1-\bar{\rho})\log_2(1-\bar{\rho})$ $\bar{\rho} = \sum_{i=0}^{N} \frac{i}{N} a_i^{(m)}$

3.2 Propagator, Consensus Time, Local Time, and Shannon Entropy

The primary goal of finding the spectral decomposition of the transition is to efficiently calculate large powers of the matrix. These large matrix powers are used to calculate future probability distributions. Using the decomposition given in 3.1.2, we have that the future probability distributions are expressed as

$$\mathbf{a}^{(m)} = \mathbf{S}\Lambda^m \mathbf{S}^{-1}\mathbf{a}^{(0)} \tag{3.30}$$

Let d_k be the initial distribution expressed in the eigenbasis given by equation 3.29. With d_k known, we can alternatively express the macrostate probability distribution as

$$a_i^{(m)} = \sum_{k=0}^{N} d_k \lambda_k^m [\mathbf{v}_k]_i. \tag{3.31}$$

As with the voter model, this is key to finding other interesting solutions, such as the expected time to consensus, the local time, and changes in entropy. The solutions themselves are discussed in table 3.1, with additional details below.

3.2.1 Consensus Time for 2-Urn Models

The consensus time is the amount of scaled time spent before the system reaches an absorbing state in which all of the balls are in a single urn. Here, we only consider linear 2-Urn models in which consensus of B is an absorbing state, but consensus of A is non absorbing. If both consensus of A and B were absorbing states, then we would require $\alpha_2 = \beta_2 = \gamma_1 = \gamma_2 = 0$. For the linearity condition to hold, $\alpha_1 = \beta_1$, which is equivalent to the voter model on the complete graph with a single rate parameter that scales time. Voter models of this type are the only cases of the 2-urn models that are *martingale*, which we have discussed in chapter 2. In this chapter, we only consider the *non-martingale* cases in which consensus of one urn is absorbing. The non-martingale cases of the 2-Urn model have very different behavior compared to the martingale cases described in chapter 2. Now, if we require $n_A = 0$ to be an absorbing state in the Markov chain, we require $Pr\{\Delta n_A = 0 | n_A = 0\} = 1$. For this to be true, we take $\alpha_2 = \gamma_1 = 0$.

We can apply similar means of calculating the moments of consensus time as with the voter model. Let s_m be the probability that the system reaches consensus at discrete time m. Therefore, by definition, the *pth* moment of the consensus time τ is given by

$$E[\tau^p] = \sum_{m=1}^{\infty} s_m (m/N)^p \tag{3.32}$$

The probability of reaching consensus, s_m, can be expressed as the probability that the system is one step from consensus times the probability that the system moves into the consensus state. Therefore, $s_m = q_1^{(1)} a_1^{(m-1)} + q_2^{(2)} a_2^{(m-1)}$. Since we can calculate $a_i^{(m)}$ explicitly, we can write s_m explicitly as

$$s_m = \sum_{k=1}^{N} d_k \lambda_k^{m-1} \left\{ \frac{\beta_1}{N} [\mathbf{v}_k]_1 + \frac{2\gamma_2}{N-1} [\mathbf{v}_k]_2 \right\}. \tag{3.33}$$

Now, we substitute this into equation 3.32 and evaluate the resulting infinite geometric series. This yields

$$E[\tau^p] \sim \sum_{k=1}^{N} \frac{d_k p!}{[N(1-\lambda_k)]^{p+1}} \left\{ \frac{\beta_1}{N}[\mathbf{v}_k]_1 + \frac{2\gamma_2}{N-1}[\mathbf{v}_k]_2 \right\}, \tag{3.34}$$

which is given in Table 3.1. Note that only the first few terms of this series can yield a good approximation to the moments of consensus time.

Even though this solution is exact, the form of equation 3.34 does not immediately reveal the asymptotic dependence on N. If we do not use the full diagonalization of the matrix, we can still obtain an approximation of the moments of consensus time. One means of approximating the consensus time is to find the time until the survival probability is $1/N$ [9, 11, 38, 39]. This implies that less than one ball would be in urn A at that time. Methods such as this are utilized in stochastic systems that have significant drift. Now, the survival probability is estimated by powers of the second largest eigenvalue of the system, given by λ_1^m. Now, using the above criterion, and converting to scaled time $\tau = m/N$ gives

$$\lambda_1^{\tau N} = \frac{1}{N}. \tag{3.35}$$

Using the exact solution for the eigenvalues given by equation 3.19, we have that $\lambda_1 = 1 - (\beta_1 - \alpha_1)/N$. The consensus time is found by solving for τ and using that $\ln \lambda_1 \sim -(1-\lambda_1)$, we find that

$$E[\tau] \sim \frac{\ln N}{\beta_1 - \alpha_1} + O(1). \tag{3.36}$$

Equation 3.36 gives the first term in the asymptotic expansion, and shows that there is an additional constant term. Figure 3.2 compares simulation data against this estimate and shows that there is good agreement between them.

3.2.2 Local Time for 2-Urn Models

The local time is the amount of scaled time (m/N) that the system spends at each macrostate prior to consensus. As with the consensus time, we assume that only consensus of B is an absorbing state, since if both states were absorbing, the

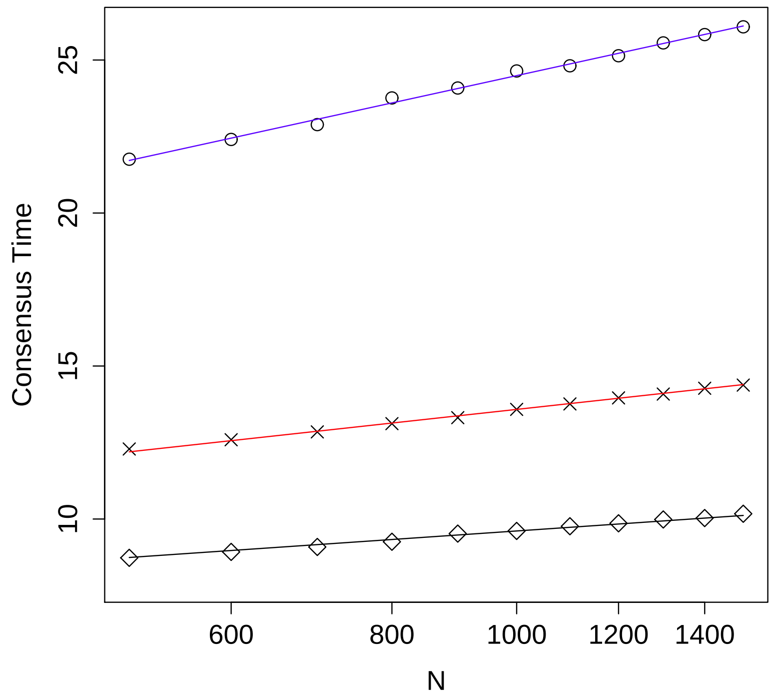

Figure 3.2: Consensus times for three linear models plotted on a logarithmic scale in N. These models have parameter configurations $\{1/4, 0, 1, 1/4, 0, 1/4\}$ ($\circ$), $\{1/2, 0, 1, 1/6, 0, 1/6\}$ ($\times$), and $\{3/4,\ 1, 1/2, 0, 1/12\}$ ($\diamond$). Solid lines are the estimates given by equation 3.36 with a fitted additive constant. The constant is dominated by the logarithm for large N.

system reduces to the voter model previously studied in chapter 2. We organize the local times into a vector $\mathbf{M}$, whose components M_i correspond to macrostate $n_A = i$. Let $M_i(m)$ be the total time spent at macrostate $n_A = i$ by discrete time m. Similar to the voter model, we consider $E[\Delta M_i(m)]$, where $\Delta M_i(m) = M_i(m+1) - M_i(m)$. The random variable $\Delta M_i(m)$ takes value $1/N$ if the system is in macrostate $n_A = i$ at time $m+1$, and otherwise takes value zero. So, $E[\Delta M_i(m)] = a_i^{(m+1)}$. Therefore, the expected local time is given by

$$E[M_i(\infty)] - E[M_i(0)] = \sum_{m=1}^{\infty} E[\Delta M_i(m)] = \frac{1}{N} \sum_{m=1}^{\infty} a_i^{(m+1)}. \tag{3.37}$$

Now, $E[M_i(0)]$ is given by the initial probability distribution of the model: $a_i^{(0)}/N$. We now apply the diagonalization of linear urn models given by equation 3.31 to find $E[M_i(\infty)]$ exactly. We now make this substitution are organize the result into a vector to find

$$E[\mathbf{M}] = \frac{1}{N}\sum_{k=1}^{N}\frac{d_k\mathbf{v_k}}{1-\lambda_k}. \tag{3.38}$$

The $k = 0$ term is not included because $d_0 = 0$. This is because $k = 0$ corresponds to the consensus state, and we stop measuring the local time once consensus has been attained. This result is also stated in Table 3.1.

3.2.3 Shannon Entropy in Social Systems

In social systems in particular, we wish to analyze the effect of disagreement in the population. If there is high dissent in the population, the less stable it may be. To measure this, we consider the Shannon entropy of the messages that the people as the measurement of uncertainty of the population [40]. In a social setting, a 2-urn process would correspond to a message from a speaker (first ball) to a neighboring listener (second ball). We consider these messages as information content, and if there is information, one can define entropy. If the population is in a consensus state, then only a single opinion will be transmitted from person to person. In this case the Shannon Entropy will be zero, which is consistent with zero dissent. Similarly, the dissent is maximum if each opinion is represented in equal amounts.

If a ball is drawn from urn A, we say that message A is transmitted. This occurs with probability n_A/N. Since n_A is a random variable at future time m, the probability that the message is A is $E[\rho(m)]$, which we denote by $\bar{\rho}(m)$. This expected value is given by

$$\bar{\rho}(m) = \sum_{i=0}^{N}\frac{i}{N}a_i^{(m)}. \tag{3.39}$$

Since there are only two urns (A or B), only one of two messages can be expressed at a given time. Since the probability of message A is $\bar{\rho}$, the probability of message B is $1-\bar{\rho}$. So, the Shannon entropy, $\mathcal{H}$, is

$$\mathcal{H} = -\bar{\rho}\ln(\bar{\rho}) - (1-\bar{\rho})\ln(1-\bar{\rho}). \tag{3.40}$$

We take the natural logarithm here out of convenience, since the consensus time is closely related to entropy for 2-Urn processes. In fact, the consensus time for the voter model is exactly the entropy scaled by N [21, 41, 42]. Entropy is maximized for $\bar{\rho} = 1/2$, and minimized at consensus states ($\bar{\rho} = 0$ or $\bar{\rho} = 1$), as stated above.

It may be useful to think of entropy in a thermodynamic sense. For information theory presented here, there is no explicit second law of thermodynamics. In fact, many social systems cause entropy to decrease, and individuals converge on a consensus opinion. As such, the second law of thermodynamics does not hold for all 2-urn processes, though there is a subclass in which entropy always increases (such as the Ehrenfest model). It becomes insightful to know under which conditions the second law of thermodynamics holds or not.

Here we will show that a non-martingale linear urn process has a principle of maximum entropy if and only if the *macroscopic symmetry property* holds, which we define below. Let U_1 and U_2 be the urns from which the first and second balls were drawn respectively. Also let U^c denote the urn that is complementary to U. For example, if $U_1 = A$, then $U_1^c = B$. A 2-urn process is *macroscopically symmetric* if for all U_1 and U_2,

$$E[n_A|U_1, U_2] = E[n_B|U_1^c, U_2^c]. \tag{3.41}$$

Physically, this means that the macroscopic behavior of the system is unchanged if the urns were interchanged. This implies a form of symmetry between urn A and B.

First we show that if macroscopic symmetry holds for linear models, then there is a principle of maximum entropy. It is sufficient to show that the equilibrium density must be $\bar{\rho} = 1/2$ to guarantee an increase in entropy. If we assume macroscopic symmetry, then we require that

$$2\gamma_1 - 2\gamma_2 + \alpha_2 - \beta_2 = 0, \tag{3.42}$$

$$\alpha_1 = \beta_1. \tag{3.43}$$

Notice that if these constraints hold, then the 2-urn model satisfies the linearity constraint given in equation 3.13. There are no nonlinear 2-urn problems that are macroscopically symmetric. Now we consider the mean density of the system for large N given in equation 3.22 with these parameter constraints, and find that

$$\frac{d\bar{\rho}}{dt} = (2\gamma_1 + \alpha_2)(1 - 2\bar{\rho}), \tag{3.44}$$

which has a monotonic attraction to $\bar{\rho} = 1/2$ for all initial conditions. This shows that entropy increases for non-martingale 2-urn models. That is, we wish to show that $2\gamma_1 - 2\gamma_2 + \alpha_2 - \beta_2 = 0$ and $\alpha_1 = \beta_1$.

To show the converse, we assume that entropy increases for non-martingale linear models and show that macroscopic symmetry holds. Since we assume that equation 3.22 is linear, we get

$$\frac{d\bar{\rho}}{dt} = 2\gamma_1 + \alpha_2 + (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)\bar{\rho}. \tag{3.45}$$

We also have that $\bar{\rho} = 1/2$ must be an equilibrium of the system. Setting 3.45 to zero for $\bar{\rho} = 1/2$ gives $\alpha_1 = \beta_1$. Since the model is linear, we substitute $\alpha_1 = \beta_1$ into equation 3.13 and find that $2\gamma_1 - 2\gamma_2 + \alpha_2 - \beta_2 = 0$. This shows that the system is macroscopically symmetric.

This result shows that the principle of increasing entropy is fundamentally related to symmetry and linearity. Since physical models obey the second law and have symmetries associated with them, then this result justifies the study of linear models in particular.

3.3 Nonlinear Models

For linear models, the propagator can be solved exactly, which guarantees an enormous control over the solution. In this section, we study more general models for which $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 \neq 0$. Because the linearity condition does not hold, we cannot solve for all eigenvalues and eigenvectors exactly. Qualitatively, nonlinear models have such a wide range of unusual behavior that we would not expect them all to be solved so easily. To focus our study, we consider nonlinear models such that only consensus of B is an absorbing state. Because the system is nonlinear, we show that there exists a phase transition across $\alpha_1 = \beta_1$, which would otherwise not exist for linear models.

3.3.1 Phase Transition for Nonlinear Models

Here we consider similar stable consensus type models as in the linear case where if all of the balls are in B, all balls will remain in B with probability 1. Consensus of B is an absorbing state in the corresponding Markov chain. By assuming this, we take $\gamma_1 = \alpha_2 = 0$. We make these substitutions in equation 3.22 to find

$$\frac{d\bar{\rho}}{dt} = (-2\gamma_2 - \alpha_1 + \beta_1 - \beta_2)\bar{\rho}\left(\frac{\alpha_1 - \beta_1}{-2\gamma_2 - \alpha_1 + \beta_1 - \beta_2} + \bar{\rho}\right). \tag{3.46}$$

Clearly, $\bar{\rho}_1 = 0$ is stationary. This is expected, because consensus of B is an absorbing state by assumption. What might seem unexpected, however, is that this root is not always stable, and that there exists a phase transition when $\alpha_1 = \beta_1$. This can be seen by considering equation 3.46 for $\bar{\rho} << 1$.

The second root, $\bar{\rho}_2$, can be expressed by a single parameter, ω. We let

$$\omega = \frac{\alpha_1 - \beta_1}{2\gamma_2 + \beta_2} \tag{3.47}$$

and observe that

$$\bar{\rho}_2 = \frac{\omega}{1+\omega}. \tag{3.48}$$

The stability of these roots are characterized completely by ω, because for $\omega > 0$, the root at the origin is unstable and $\bar{\rho}_2$ becomes stable. The root ρ_2 also exists within the physical domain ($0 \leq \bar{\rho} \leq 1$), which indicates that the system is attracted to a metastable state. Stochastically, the system will also randomly fluctuate within this interval and always has a non-zero probability of achieving consensus. So, even though the system is attracted to the metastable distribution with high probability, it will eventually achieve the consensus state in finite time. The consensus time for the metastable case, however, is exponential with N. In the special case when $\gamma_2 = \beta_2 = 0$ should be interpreted as $\omega = \pm\infty$ and correspond to a biased voter model [21]. In this case, the consensus time may not be exponential because consensus of both A and B are stable.

For $\omega < 0$, the origin is stable, and ρ_2 is unstable. In this case, the solution to equation 3.46 is the logistic function. For $\bar{\rho}$ small, the equation resembles a linear model. This suggests that replacing the nonlinear system with a linear urn model near consensus can capture its behavior. That is, since the drift portion of the dynamics is small, the consensus time when $\omega < 0$ is $E[\tau] \sim \ln N/(\beta_1 - \alpha_1)$. So, we will only focus on the two remaining cases: the phase transition itself ($\omega = 0$) and the metastable case ($\omega > 0$).

When $\omega = 0$, the two roots are coincident, which indicates a transcritical bifurcation [43]. Fig. 3.3 shows the bifurcation diagram of these nonlinear 2-urn models.

Although the consensus time is proportional to $\ln N$ when $\omega < 0$, we will show that when $\omega = 0$, the consensus time increases to $\sqrt{N}$ for $\gamma_2 + \beta_2 \neq 0$.

3.3.1.1 Phase Transition: A Model of Consensus Formation

Socially, we consider a model that may correspond to the decay of a popular trend, A, that dominates the system. Those who do are not involved with the trending fad are considered in urn B. Those who do not accept A can be convinced to adopt it with probability α_1 when exposed to it. Similarly, people with A can be dissuaded from it with probability β_1 when exposed to those without A. Also, an individual may be turned away from the fad if they are exposed to it for long enough. That is, if an individual with A is exposed to many individuals with A, then the novelty of A fades and the person no longer adopts it, which is a sociologically

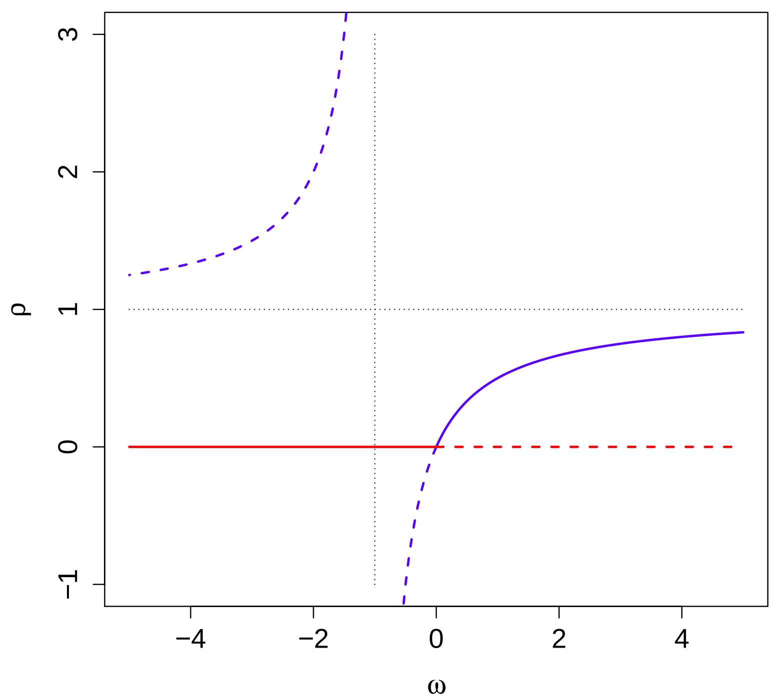

Figure 3.3: Bifurcation diagram showing the stability of the stationary points of equation 3.46 for nonlinear consensus models. The solid lines indicate that the point is stable dashed line indicate that it is unstable. The red lines correspond to the root at the origin, and the blue curve corresponds to the second root, ρ_2. For $\omega > 0$, the system attracts to a metastable state and for $\omega < 0$, the systems resemble linear models.

observed tendency [44]. If A is perceived as a fad, then people will be less likely to adopt it [45]. We model this by assuming that if A speaks with another A, then the fad is rejected with probability β_2 and the listener becomes B. It is also reasonable to assume the both individuals could reject the fad simultaneously, which would occur with probability γ_2, but for simplicity, we set $\gamma_2 = 0$.

We wish to find the amount of time until nobody accepts A, and everyone is in state B. To simplify the analysis, we will consider the case in which $\alpha_1 = \beta_1 = \beta_2 = 1$, although any configuration of parameters could be considered, provided $\omega = 0$. Let $T(\rho)$ be the time to consensus. Using first step analysis, we find that T must satisfy

$$T(\rho) = p(\rho)T(\rho + \Delta\rho) + r(\rho)T(\rho) + q(\rho)T(\rho - \Delta\rho) + 1, \tag{3.49}$$

where $\Delta\rho = 1/N$ and

$$p(\rho) = \rho(1 - \rho) \tag{3.50}$$

$$q(\rho) = \rho(1 - \rho) + \rho^2 \tag{3.51}$$

$$r(\rho) = 1 - p(\rho) - q(\rho). \tag{3.52}$$

$$\tag{3.53}$$

We also have the boundary condition $T(0) = 0$. Expanding by Taylor's theorem to second order gives

$$-1 \sim v(\rho)\frac{dT}{d\rho} + \frac{1}{2N}D(\rho)\frac{d^2T}{d\rho^2} \tag{3.54}$$

where

$$v(\rho) = -\rho^2 \tag{3.55}$$

$$D(\rho) = 2\rho(1 - \rho) + \rho^2. \tag{3.56}$$

The system is dominated by the drift term, v, when the system is not near consensus. Therefore, ρ approaches consensus in $O(1)$ time. After this time has passed, we assume that these two terms in equation 3.54 balance. Let $\rho = \delta\xi$ for some $\delta \ll 1$ that balances the drift and diffusion terms. The derivatives are then transformed by

$$\frac{d}{d\rho} = \frac{1}{\delta}\frac{d}{d\xi}. \tag{3.57}$$

So, after dropping higher order terms, equation 3.54 is given by

$$-1 \sim -\delta\xi^2\frac{dT}{d\xi} + \frac{\xi}{N\delta}\frac{d^2T}{d\xi^2}. \tag{3.58}$$

For these terms to balance, we choose $\delta = 1/\sqrt{N}$. Now, we wish to solve

$$\xi \frac{d^2T}{d\xi^2} - \xi^2 \frac{dT}{d\xi} = -\sqrt{N}. \tag{3.59}$$

We are given one boundary condition, $T(0) = 0$, yet we wish to solve a second order equation. Since the time for the system to approach consensus is $O(1)$ and that the consensus time is monotonic, the derivative of consensus time with respect to ρ is at most $O(1)$ when $\rho = O(1)$. So, the derivative of T with respect to ξ tends to zero as $\xi \to \infty$. This is stated mathematically by

$$\lim_{\xi \to \infty} \frac{dT}{d\xi} = 0, \tag{3.60}$$

which supplies the second boundary condition. Now, the solution to equation 3.59 is

$$T(\xi) = \sqrt{N} \int_0^{\xi} \int_u^{\infty} \frac{1}{s} e^{-\frac{1}{2}(s^2 - u^2)} ds du. \tag{3.61}$$

We are particularly interested in the case when A initially dominates the system. Furthermore, since the integral converges exponentially, we can take $\xi \to \infty$ without significantly changing the value. So, to find $T(\infty)$, we take

$$s = \sqrt{2}r \sec\theta \tag{3.62}$$

$$u = \sqrt{2}r \tan\theta. \tag{3.63}$$

Making this change of variables, the integral can be simplified to

$$T \sim \sqrt{\frac{\pi^3 N}{8}}. \tag{3.64}$$

This shows that when $\omega = 0$, the consensus time increases from the order of $\ln N$ to $\sqrt{N}$.

3.3.2 Metastable Consensus Time

The other interesting case is when $\omega > 0$, in which there is a stable fixed point of the drift equation in the region $0 < \rho < 1$. The interesting feature of models such as these is that the system is attracted to the root in the interior, ρ_2, and is repelled from the root at $\rho_1 = 0$. However, for any N, the probability that $n_A = 0$ tends to 1 as $m \to \infty$. In other words, the system almost certainly reaches consensus in finite time. Consensus of B is stable with probability 1, yet ρ_2 is stable only with probability close to 1. Therefore, ρ_2 is the mean of a metastable distribution that exists apart from consensus. In terms of the Markov chain model, consensus corresponds to the eigenvalue 1 of the transition matrix and the metastable distribution is an eigenvector with eigenvalue that is exponentially close to 1.

We have two major goals of this section. The first is to find the metastable distribution for a particular metastable urn process and the second is to find the corresponding consensus time. As in the above case, we restrict our study to a particular parameter configuration to serve as a canonical example of the solution. Once the balls have been selected from the urns, the model that we consider is given by the following rules:

1. If the balls came from different urns, place both in A.

2. If the balls came from the same urn, place both in B.

It is evident from these rules that consensus of B is an absorbing state. If all balls are in urn B, then the balls are always drawn from the same urn and replaced in B with probability 1. This model has the parameter configuration $\alpha_1 = 2$, $\gamma_2 = 1$, and all others equal to 0. By equation 3.47, we have that $\omega = 1$ and therefore, $\rho_2 = 1/2$. Since $\omega > 0$, the bifurcation diagram in Fig. 3.3 indicates that $\rho_2 = 1/2$ is stable and constitutes a metastable state.

Let λ be the eigenvalue that corresponds to the metastable distribution, ϕ_i. Since λ is exponentially close to 1, $1 - \lambda$ is a transcendentally small term. This observation offers evidence of the conjecture stated in section 3.1.1 that no other change of variables could ascertain the exponentially small spectral gap. Since ϕ_i are components of an eigenvector, we can choose how scale it. Conditioned on the information that the system is not in consensus, then $\phi_i = Pr\{n_A = i\}$. So, for ϕ_i to constitute a probability distribution, we need the sum of the components to be 1. So, we let c_i be the components of the unscaled eigenvector corresponding to λ

and scale it so that

$$\phi_i = \frac{c_i}{\sum_{i=1}^{N} c_i}. \tag{3.65}$$

With this, the recursion relation for c_i from the eigenvalue problem for c_i after dropping the $1-\lambda$ terms is given by

$$0 = p_{i-1}^{(1)} c_{i-1} - (p_i^{(1)} + q_i^{(2)}) c_i + q_{i+2}^{(2)} c_{i+2}. \tag{3.66}$$

We also demand that $c_i = 0$ when $i < 0$ and $i > N$. Also, we set $c_N = 1$ to begin the recursion. We know that $c_N \neq 0$ because if it were, it would imply that the system can never achieve this state. So, we take equation 3.66 and solve for c_{i-1} to find

$$c_{i-1} = \frac{(p_i^{(1)} + q_i^{(2)}) c_i - q_{i+2}^{(2)} c_{i+2}}{p_{i-1}^{(1)}}. \tag{3.67}$$

Given $c_N = 1$ and $c_i = 0$ for $i > N$, we can use equation 3.67 to find c_i for $1 \leq i \leq N-1$. After rescaling to obtain ϕ_i, we obtain the distribution shown in figure 3.4. It should be noted that the distribution asymptotically approaches a Gaussian function to leading order.

We now calculate the consensus time for the metastable case. The strategy is to use equation 3.34, which gives the consensus time in terms of the eigenvalues and eigenvectors. When using equation 3.34, we take $d_1 = 1$ and all other $d_k = 0$ for $[\mathbf{v_1}]_i = \phi_i$ and λ_1 being the eigenvalue for ϕ_i. That is, equation 3.34 reduces to

$$E[\tau] = \frac{2\phi_2}{N^2(N-1)(1-\lambda)^2}. \tag{3.68}$$

Since we found the dominant eigenvector in the system, all we need is the corresponding eigenvalue. We find it by considering the row-sum of the transition matrix $\mathbf{T}$. Since all of the rows sum to 1, we multiply both sides of $\mathbf{T}\mathbf{v} = \lambda\mathbf{v}$ by $\mathbf{1}^T$ on the left to obtain $\mathbf{1}^T\mathbf{v} = \lambda\mathbf{1}^T\mathbf{v}$. Unless $\lambda = 1$, we must have that the sum of the components of $\mathbf{v}$ is zero. This is important because $\sum_{i=1}^{N} \phi_i = 1$ in order to be a probability distribution conditioned that the system has not reached consensus. This implies that $\phi_0 = -1$ in order for the sum of the components from $i = 0 \ldots N$

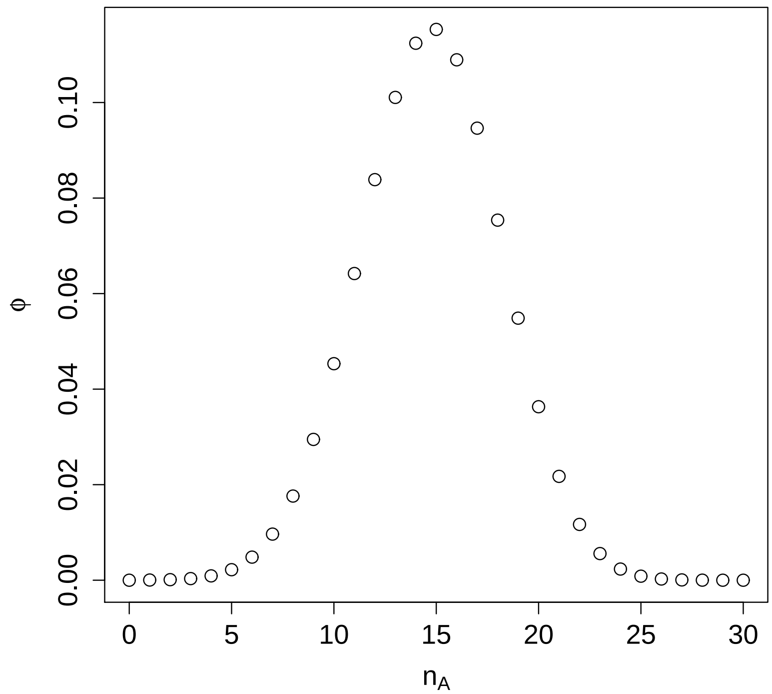

Figure 3.4: Solution for the metastable distribution generated by equation 3.67 for $N = 30$. The distribution is asymptotically a Gaussian function because the diffusivity is smooth. Near $\rho_2 = 1/2$, which is the mean, the diffusivity can be approximated by a constant, which yields a Gaussian.

to be zero. If we take $i = 0$ in the eigenvalue problem, we find that

$$\lambda\phi_0 = \phi_0 + q_2^{(2)}\phi_2. \tag{3.69}$$

For $\phi_0 = -1$, we have that $1-\lambda = q_2^{(2)}\phi_2$. Substituting this into equation 3.68 gives

$$E[\tau] = \frac{N-1}{2\phi_2} \tag{3.70}$$

where ϕ_2 is easily found by 3.67 and 3.65. Figure 3.5 compares simulation data against the results in equation 3.70.

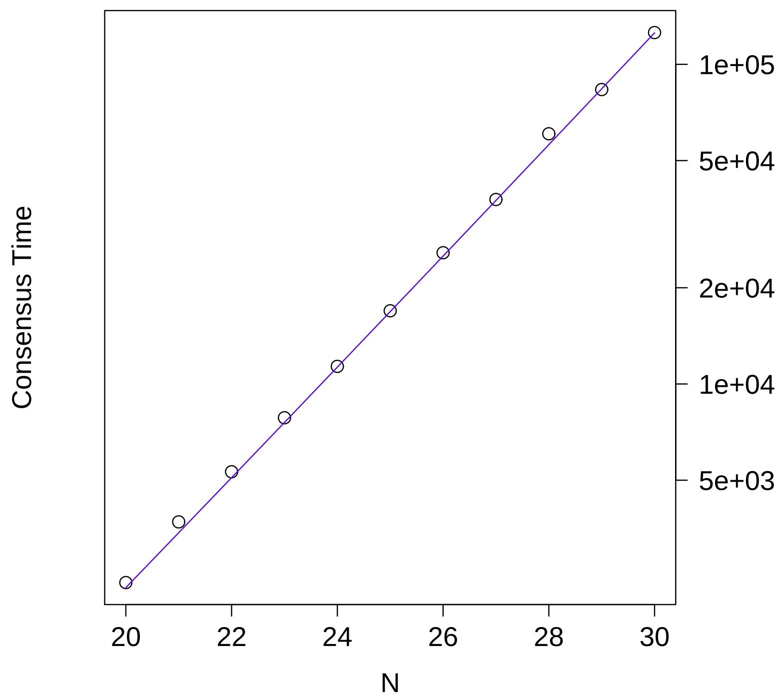

Figure 3.5: Simulation data ($\circ$) of the consensus time averaged over 500 runs is plotted against the exact solution shown in blue, given by equation 3.70. The horizontal axis is on a linear scale and the vertical axis is on a logarithmic scale. The apparent linear relationship indicates the exponential nature of the consensus time.

CHAPTER 4
Solution of Urn Models on Incomplete Networks

In the previous chapters, we considered social models on the complete graph, in which every node is connected to every other node. This corresponds to the mean field analysis, which suggests that network effects neglected. Although the mean field network case is valid in many cases [12], the effect of an incomplete network structure isn't always negligible [21–23, 42, 46–49]. We examine here some of these cases in which the network topology has a significant role in the system. In particular, we consider bipartite models and scale free models.

We offer a natural extension of the 2-urn models on networks by using a "speaker-listener" interpretation. The first ball that is chosen in the 2-urn process can be thought of as a randomly chosen node in a network. Now, instead of drawing a random second node anywhere in the network, choose a node that is connected to the speaker instead. This constitutes the listener, which would be the second ball in the 2-urn construction. The nodes then update their states according to the states of the listener and speaker. As with the 2-urn models, we first consider the voter model, and then more general parameterized cases. These networks are of particular interest because we reduce the dynamics of the model to a multi-urn model, where there is a pair of urns for each degree k in the network.

4.1 Voter Model on Complete Bipartite Graphs

Here, we will consider the complete bipartite graph. In this case, nodes in the network are separated into two groups. Every node in one group is connected to every node in the opposite group [50]. The complete bipartite graph can also be defined as the complement of two complete graphs. Let N_1 be the number of nodes in the first group and N_2 be the total number of nodes in the second group. Also, let $n_A^{(1)}(m)$ and $n_A^{(2)}(m)$ be the number of nodes with opinion A in groups 1 and 2

Portions of this chapter previously appeared as: W. Pickering and C. Lim, "Solution of the voter model by spectral analysis," *Phys. Rev. E*, vol. 91, no. 1, p. 012812, Jan. 2015.

respectively. Letting

$$a_{ij}^{(m)} = Pr\{n_A^{(1)}(m) = i,\ n_A^{(2)}(m) = j\} \tag{4.1}$$

and

$$Q^{(m)}(x, y, u, v) = \sum_{i,j} a_{ij}^{(m)} x^i, y^{N_1 - i} u^j v^{N_2 - j}, \tag{4.2}$$

we can use the procedure laid out in section chapter 2 to find the single step propagator for the model:

$$\left[\frac{u(x-y)}{NN_2} - \frac{y(u-v)}{NN_1}\right] Q_{yu}^{(m)} + \left[\frac{-v(x-y)}{NN_2} + \frac{x(u-v)}{NN_1}\right] Q_{xv}^{(m)} = \Delta_{+m} Q^{(m)}. \tag{4.3}$$

Unlike for the complete graph, the closed form expression for the solution to the spectral problem for this equation will not be found with these methods. However, we can apply some assumptions to the system that can reduce propagator to the region of the (i, j) grid that is diffusion dominant. Then, we wish to find the approximate size of the spectral gap, since this governs the expected time to consensus. Since we need to restrict the region to estimate the spectrum, it will not allow us to find the eigenvectors exactly. As such, the more detailed solutions that depend on the behavior of the eigenvectors, such as local times, will not be found.

If we take the single step propagator to continuous time, it is known that the system approaches equilibrium when $n_A^{(1)}/N_1 - n_A^{(2)}/N_2 \sim 0$ [21]. Furthermore, the time to reach this equilibrium state is negligible compared to the time to reach consensus. Along this line, diffusion governs the motion of the macro-state of the system instead of drift. We inject this assumption into the generating function method by noting that

$$Q_{yu}^{(m)} = \sum_{i,j} c_{ij} j(N_1 - i) x^i y^{N_1-i-1} u^{j-1} v^{N_2-j} \tag{4.4}$$

$$= \sum_{i,j} c_{ij} N_1 N_2 \frac{j}{N_2} \left(1 - \frac{i}{N_1}\right) x^i y^{N_1-i-1} u^{j-1} v^{N_2-j} \tag{4.5}$$

$$\approx \sum_{i,j} c_{ij} N_1 N_2 \frac{i}{N_1} \left(1 - \frac{j}{N_2}\right) x^{i-1} y^{N_1-i} u^j v^{N_2-j-1} \tag{4.6}$$

$$= \sum_{i,j} c_{ij} i(N_2 - j) x^{i-1} y^{N_1-i} u^j v^{N_2-j-1} \tag{4.7}$$

$$= Q_{xv}^{(m)} \tag{4.8}$$

We substitute $Q_{yu}^{(m)}$ in for $Q_{xv}^{(m)}$ to obtain the reduced propagator given by

$$\frac{(u-v)(x-y)}{N_1 N_2} Q_{yu}^{(m)} \approx \Delta_{+m} Q^{(m)}. \tag{4.9}$$

Letting $G(x, y, u, v) = \sum_{ij} c_{ij} x^i y^{N_1-i} u^j v^{N_2-j}$, the spectral problem is given by

$$(u-v)(x-y) G_{yu} \approx N_1 N_2 (\lambda - 1) G. \tag{4.10}$$

Let $s = u - v$, $r = x - y$, and $H(r, y, s, v) = G(x, y, u, v)$. As with the complete graph, this is a linear transformation, so we expect solutions of the same form: $H(r, y, s, v) = \sum_{ij} b_{ij} r^i y^{N_1-i} s^j v^{N_2-j}$. The differential equation becomes

$$rs(H_{sy} - H_{rs}) \approx N_1 N_2 (\lambda - 1) H. \tag{4.11}$$

The corresponding difference equation for the coefficients, b_{ij}, is

$$j(N_1 - i + 1) b_{i-1,j} - ij b_{ij} \approx N_1 N_2 (\lambda - 1) b_{ij}. \tag{4.12}$$

With similar arguments for the complete graph, we require a singularity in this difference equation so that the solution is not trivial. Also, we assume that both sides of 4.12 agree to within an $O(1)$ factor since this is an approximation of the propagator. So, we write that $O(ij) = N_1 N_2 (\lambda - 1)$. Taking i, j to be small yields the approximate size of the spectral gap:

$$\lambda_{ij} = 1 - O\left(\frac{1}{N_1 N_2}\right). \tag{4.13}$$

Note that in procedure, the same spectrum would be recovered if one were to use G_{xv} instead of G_{yu} in equation 4.10.

We can find the moments of consensus times by using equation 2.47 from chapter 2. Taking $p = 1$, this suggests that the expected time to consensus is $O(N_1 N_2/N)$. This is consistent with continuous time analysis, which shows that the expected number iterations to reach consensus is

$$E[\tau] = \frac{4N_1 N_2}{N}\left[(1-\omega)\ln\frac{1}{1-\omega} + \omega\ln\frac{1}{\omega}\right], \tag{4.14}$$

where ω is the degree weighted mean of microstates [21]. As before, this bound is valid for all initial probability distributions. However, without detailed information about the eigenvectors, we cannot extract more detailed information about the propagator than the bound on moments of consensus time.

4.2 Voter Model on Uncorrelated Degree Sequence Networks

Here we will look into the spectrum when the Voter model is imposed on networks with arbitrary, fixed degree sequences. The expected time to reach consensus on these networks is known to be $O(N\mu_1^2/\mu_2)$, where μ_1 and μ_2 are the first and second moments of the degree sequence respectively [21, 22]. We shall verify this result by investigating the spectrum of the transition matrix, which we have shown to give control over the expected time to consensus. Furthermore, we can utilize this solution to find the moments of consensus time.

To begin, we establish the following notation. Let $\mathbf{n} \in \mathbb{Z}_2^N$ be a vector consisting of the microstates of each node in the network. Components of $\mathbf{n}$ take value 1 when the node has state A and takes value 0 when it has state B. Let $\mathbf{A}$ be the adjacency matrix of the network, and $\mathbf{k} \in \mathbb{R}^N$ contain the degrees of each node. Also, let component k of $\mathbf{n}_A(m)$ be the number of nodes with degree k that have opinion A. Let N_k be the total number of nodes with degree k. We also take $\{\mathbf{e}_i\}_{i=1}^N$ to be the standard basis vectors for $\mathbb{R}^N$.

Let $a_{\boldsymbol{\alpha}}^{(m)} = Pr\{\mathbf{n}_A(m) = \boldsymbol{\alpha}\}$. Then, the single step propagator is found to be

$$\Delta_{+m} a_{\boldsymbol{\alpha}}^{(m)} = \sum_k [p_{k\boldsymbol{\alpha}-\boldsymbol{e}_k} a_{\boldsymbol{\alpha}-\boldsymbol{e}_k}^{(m)} - (p_{k\boldsymbol{\alpha}} + q_{k\boldsymbol{\alpha}}) a_{\boldsymbol{\alpha}}^{(m)} + q_{k\boldsymbol{\alpha}+\boldsymbol{e}_k} a_{\boldsymbol{\alpha}+\boldsymbol{e}_k}^{(m)}]. \tag{4.15}$$

Here, the transition probabilities are

$$p_{k\boldsymbol{\alpha}} = \frac{1}{Nk} \sum_{i \in K} \mathbf{A}\mathbf{e}_i \cdot \mathbf{n}(1 - n_i) \tag{4.16}$$

$$q_{k\boldsymbol{\alpha}} = \frac{1}{Nk} \sum_{i \in K} \mathbf{A}\mathbf{e}_i \cdot (\mathbf{1} - \mathbf{n}) n_i, \tag{4.17}$$

where K is the set of all nodes with degree k. As with the complete bipartite graph, let us assume that the system is diffusion driven. The system will quickly approach this state in $O(N)$ steps and move towards consensus on a much slower time scale [21]. Furthermore, we replace $\mathbf{A}$ with the mean over all adjacency matrices to give an expected time to consensus over random networks. This suggests that we take

$$A_{ij} \to \frac{k_i k_j}{N\mu_1}. \tag{4.18}$$

With these assumptions, we require that for each k,

$$\frac{\alpha_k}{N_k} \approx \frac{\mathbf{k} \cdot \mathbf{n}}{N\mu_1} \tag{4.19}$$

$$q_{k\boldsymbol{\alpha}} \sim p_{k\boldsymbol{\alpha}} \sim \frac{\alpha_k (N_k - \alpha_k)}{N N_k}. \tag{4.20}$$

The approximation sign in equation 4.19 is in the sense that both sides agree to within an $O(1)$ factor. Using these simplifications, we now write the reduced propagator as

$$\Delta a_{\boldsymbol{\alpha}}^{(m)} \sim \sum_k \Delta_k^2 (p_{k\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}}). \tag{4.21}$$

Here, Δ_k^2 is the second centered difference operator over α_k. Let $\mathbf{N}$ be a vector with components N_k. We can write the reduced propagator in generating function form by letting

$$Q^{(m)}(\mathbf{x},\mathbf{y}) = \sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}}^{(m)} \mathbf{x}^{\boldsymbol{\alpha}} \mathbf{y}^{\mathbf{N}-\boldsymbol{\alpha}}. \tag{4.22}$$

Here, the vector powers of $\mathbf{x}, \mathbf{y}$ is the multi-index notation, defined by $\mathbf{x}^{\boldsymbol{\alpha}} = \prod_i x_i^{\alpha_i}$ [51]. We can write the reduced propagator in terms of this function and obtain the following equation for the spectral problem:

$$(\lambda - 1)G \sim \sum_k \frac{(x_k - y_k)^2}{NN_k} \frac{\partial^2 G}{\partial x_k \partial y_k}. \tag{4.23}$$

Here, G takes the same form as $Q^{(m)}$, and the coefficients are approximate values of the eigenvectors as we had done before. To solve this, let $u_k = x_k - y_k$ and $G(\mathbf{x},\mathbf{y}) = H(\mathbf{u},\mathbf{y})$ to obtain

$$(\lambda - 1)H \sim \sum_k \frac{u_k^2}{NN_k} \left(\frac{\partial^2 H}{\partial u_k \partial y_k} - \frac{\partial^2 H}{\partial u_k^2} \right). \tag{4.24}$$

This translates to the finite difference equation for the coefficients of H:

$$(\lambda - 1)b_{\boldsymbol{\alpha}} \sim \sum_k \left[\frac{(\alpha_k - 1)(N_k - \alpha_k + 1)b_{\boldsymbol{\alpha}-k}}{NN_k} - \frac{\alpha_k(\alpha_k - 1)b_{\boldsymbol{\alpha}}}{NN_k} \right]. \tag{4.25}$$

By requiring that there is a singularity in the difference equation, we find a general form of the spectrum as

$$\lambda \sim 1 - \sum_k \frac{\alpha_k(\alpha_k - 1)}{NN_k} \tag{4.26}$$

$$\approx 1 - \sum_k \left(\frac{\alpha_k}{N_k} \right)^2 \frac{N_k}{N}. \tag{4.27}$$

To proceed, we need to apply the restriction on α_k/N_k given in equation 4.19. In particular, we are looking for small values of α_k, which will maximize λ such that $\lambda < 1$. Notice that when α_k is small and non-zero, we find from equation 4.19 that

$$\frac{\alpha_k}{N_k} = O\left(\frac{k}{N\mu_1}\right) \tag{4.28}$$

Use this in equation 4.27 to find that

$$\lambda \approx 1 - O\left(\frac{1}{N^2\mu_1^2}\sum_k k^2 \frac{N_k}{N}\right) \tag{4.29}$$

$$1 - \lambda = O\left(\frac{\mu_2}{N^2\mu_1^2}\right). \tag{4.30}$$

This gives an asymptotic approximation of the spectral gap. The calculation for finding the moments of consensus time given in chapter 2 can be applied here to attain the result given in equation 2.47. For the expected time to consensus, this result states that $E[\tau] = O\left(\frac{\mu_1^2}{\mu_2}N\right)$, which is consistent with known results regarding consensus times on uncorrelated heterogeneous networks [21, 22].

For each network that we consider, we test the analytical result for the moments of consensus times against Monte Carlo simulation data. To this end, we conduct two simulations. In the first simulation, we take $N = 100$, and run the model 100 times for each network we have considered. For the bipartite graph, we fix $N_1 = 4N_2$ in all simulations. For the uncorrelated network, we use the Erdős Rényi model with the probability of linking two nodes set to be $5/N$ [6]. We let T_p be the numerical result for the pth moment of consensus time for the complete and bipartite simulations. For the Erdős Rényi networks, we normalize the time to consensus by μ_1^2/μ_2 prior to computing T_p. We do this since equation 2.47 predicts that there is a linear relationship between $\ln(T_p/p!)$ and p in each network. In figure 4.1, we see that the simulation provides a linear agreement between these quantities.

The second simulation that we provide tests the asymptotic relationship between an arbitrary moment of consensus time with N. For each network, we simulate the voter model for $N = 10, 20, ..., 100$. For each of these values of N, we run the model 100 times and numerically compute the $5th$ moment in each case. This is to test the dependence of the moments of consensus time with N. In equation 2.47, we

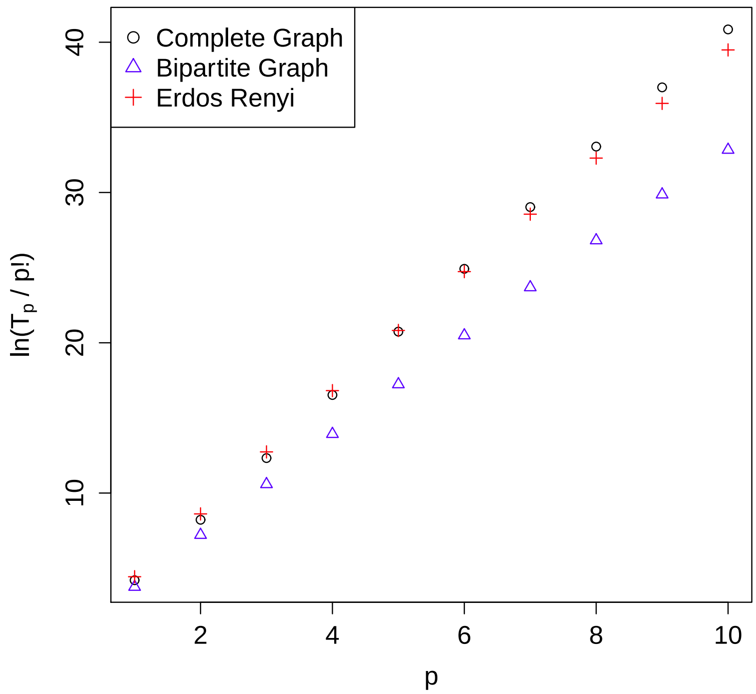

Figure 4.1: Monte Carlo simulation of the voter model on the complete graph, complete bipartite graph, and uncorrelated networks. We fix $N = 100$ and let p range from $1 \ldots 10$. We expect each network to exhibit linear behavior in the moment p.

expect to see that there is a linear relationship between $\ln(T_p/p!)$ and $\ln N$. Furthermore, since we take $p = 5$ for each N, we expect that the slope should be about 10 in each case. Figure 4.2 shows that the trend is in fact linear and that the prediction about the slope is accurate.

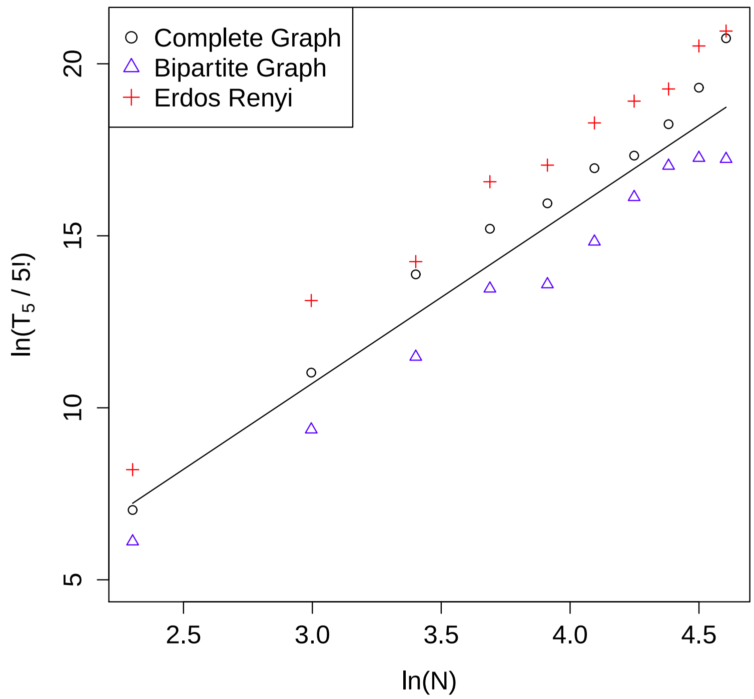

Figure 4.2: Monte Carlo simulation of the voter model on the complete graph, complete bipartite graph, and Erdős Rényi networks. In this simulation, we fix $p = 5$ and let N range from $10, 20, ..., 100$. In addition, a reference line of slope p is plotted, which is the prediction given by the theory.

4.3 Extension of the 2-Urn Models on Networks

In the above analysis, we considered social models exclusively on the complete graph. It has been observed that the ordering dynamics of the Naming Game – a similar social model – do not significantly change when comparing real world networks against complete graphs [12]. This gives great credibility to using complete graph analysis on sparse real world networks. However, the use of a fully connected graph may not be an accurate approximation of the dynamics on certain networks. For the voter model, which is a case of the above general social framework, the system orders differently based on network structure [15, 21, 22, 42].

The structure of the 2-urn models yields a natural extension to incomplete

Table 4.1: Parameters of the Network Model

Input				
AA	AB	BA	BB	Output
$-$	α_{12}	α_{11}	γ_1	AA
β_{22}	$-$	r_0	α_{21}	AB
β_{21}	r_1	$-$	α_{22}	BA
γ_2	β_{11}	β_{12}	$-$	BB

graphs. Each node has one of two states: A or B. For the general network model, we choose a node randomly and the choose a random neighbor. The nodes then update their spins with probabilities that depend on the order that the nodes were selected and their spins. We now have twelve parameters that characterize each model, each of which are transition probabilities. Table 4.1 shows the probability of accepting a particular update as a function of the spins and the order. So, if the first node has state B and the second node has state A, then the probability that they both become A is α_{11}. For parameters that have two subscripts (e.g. α_{11}), the second digit refers to which node is updated. The first digit is for consistency with the complete graph notation.

These models have a very wide range of applications. For instance, the voter model has $\alpha_{11} = \beta_{11} = 1$ and all others zero, while the invasion process has $\alpha_{12} = \beta_{12} = 1$, and all others equal to zero [21]. Generalized versions of the Moran model that include mutations have parameters $\beta_{21} = \mu_1$, $\alpha_{21} = \mu_2$, $\beta_{11} = 1 - \mu_2$, and $\alpha_{11} = 1 - \mu_1$ [52]. We can also cast movement of individuals, or particles, by interpreting state A as occupied and B as unoccupied. Of particular interest in the literature is the coalescing random walk, in which particles move from site to site and cluster together when moving into an occupied site [30, 42, 53]. This is accomplished in our framework by $r_1 = 1, \beta_{21} = 1$ and all other parameters set to zero. The annihilating random walk is similar to the coalescing random walk, but is characterized by the fact that the two particles are removed if one moves into an occupied site [53, 54]. This is given by $r_1 = 1, \gamma_2 = 1$ and all other parameters set to zero. Furthermore, we can consider directed networks which determines the flow of information throughout the network. These represent a fraction of the possible applications of the general network model.

Clearly not all parameter configuration on all networks can be solved as swiftly and concisely as the above 2-urn models. Even a generalize treatment of each case is beyond the scope of this manuscript. However, we explore some of the previous examples imposed on a popular class of networks that we call *uncorrelated degree heterogeneous networks*. These are random networks that are generated with a given, fixed degree sequence. The network is then chosen uniformly from the set of all graphs with that degree sequence [21, 55–58].

4.3.1 Consensus Model: Listener First

Here we explore a model of consensus formation on the above random networks. We interpret the situation in the following way. During a single update, the first node that is chosen is called the "listener". This individual chooses one of its neighbors to be the second node, called the "speaker". Then, we assume that only the listener is allowed to update its state. We also assume that consensus of the B state is an absorbing macrostate. That is, if every node is in state B, every node will remain in state B with probability 1.

Now we express these ideas as a parameter configuration in the above framework. Since only the first node – the listener – is allowed to change its state in an update, we set $\alpha_{12} = \gamma_1 = \beta_{22} = r_0 = r_1 = \alpha_{22} = \gamma_2 = \beta_{12} = 0$. By assumption, consensus of B is an absorbing state, so we additionally take $\alpha_{21} = 0$. This leaves us with parameters $\alpha_{11}, \beta_{21}, \beta_{11}$. We then drop the second digit in the subscripts since it is implied that only the first node is allowed to change. This corresponds to a 3-parameter system: $\alpha_1, \beta_2, \beta_1$.

We wish to express the transition probabilities of the system in terms of known quantities. First, we must establish the notation used in the following discussion. Let $\mathbf{A}$ be the adjacency matrix of the network and let N_k be the number of nodes with degree. Also let k_i be the degree of node i and let $\mathbf{k}$ be a vector with components k_i. Let $\eta_i = 1$ if node i has state A and $\eta_i = 0$ if node i has state B. The vector $\boldsymbol{\eta}$ takes components η_i. Similarly, let n_k be the number of nodes of degree k of state A and $\mathbf{n}$ takes components n_k. Also, let $\rho_k = n_k/N_k$ and $\rho = \sum_k n_k/N$. Let D_k be the set of all nodes with degree k. Let μ_p be the pth moment of the degree distribution of the network, which is given by

$$\mu_p = \sum_k \frac{N_k}{N} k. \tag{4.31}$$

Also, we define the quantities z_p to expected value of the microstate of the system, whose probability measure is proportional to the degree to the pth power. This is written explicitly by

$$z_p = \frac{\mathbf{k}^p \cdot \boldsymbol{\eta}}{N\mu_p}. \tag{4.32}$$

Here, $\mathbf{k}^p$ is the component-wise power of $\mathbf{k}$. That is, the ith component of $\mathbf{k}^p$ is k_i^p. Finally, we can express the transition probabilities of the network model:

$$Pr\{\Delta n_k = 1\} = \sum_{i \in D_k} \sum_j \frac{\alpha_1 A_{ij}}{N k_i} (1 - \eta_i)\eta_j \tag{4.33}$$

$$Pr\{\Delta n_k = -1\} = \sum_{i \in D_k} \sum_j \frac{A_{ij}}{N k_i} [\beta_1 \eta_i (1 - \eta_j) + \beta_2 \eta_i \eta_j]. \tag{4.34}$$

Now we make a mean-field assumption and replace A_{ij} with mean adjacency matrix of the uncorrelated degree heterogeneous networks [21]. Since the edges of the network are uncorrelated, the probability that nodes i and j have an edge is proportional to their respective degrees. The expected value of the components of the adjacency matrix, therefore, is

$$E[A_{ij}] = \frac{k_i k_j}{N\mu_1}. \tag{4.35}$$

Upon substitution of $A_{ij} \to E[A_{ij}]$ and simplification, we express the transition probabilities succinctly as

$$p(\rho_k, z_1) = \alpha_1 \frac{N_k}{N} (1 - \rho_k) z_1 \tag{4.36}$$

$$q(\rho_k, z_1) = \beta_1 \frac{N_k}{N} \rho_k (1 - z_1) + \beta_2 \frac{N_k}{N} \rho_k z_1. \tag{4.37}$$

We use the transition probabilities to generate an ODE for the time evolution of the expected value of ρ_k. Taking $\dot{\bar{\rho}}_k \sim E[\Delta\rho_k/(1/N)]$, we acquire the following ODE system:

$$\dot{\bar{\rho}}_k = [\alpha_1(1-\bar{\rho}_k) - \beta_2\bar{\rho}_k]z_1 - \beta_1\bar{\rho}_k(1-z_1) \tag{4.38}$$

Note that this system is coupled together only by the variable z_1. We will first show that $\rho_k \to z_1$ as the system evolves. This is found by expressing the equation for $\dot{\bar{\rho}}_k - \dot{z}_1$. To do so, we need $\dot{z}_1$. Multiplying equation 4.38 by $\frac{N_k}{N\mu_1}$ and summing over k gives

$$\dot{z}_1 = [\alpha_1(1-z_1) - \beta_2 z_1]z_1 - \beta_1 z_1(1-z_1) \tag{4.39}$$

Note that equation 4.39 can be solved, with solution

$$z_1(t) = \frac{\omega(1+\omega)^{-1}}{1+\left(\frac{\omega}{(1+\omega)z_1(0)} - 1\right)e^{-(\alpha_1-\beta_1)t}}. \tag{4.40}$$

where ω is given in Eq 3.47. Note that when $\omega > 0$, $z_1 \to 0$ as $t \to \infty$. When $\omega < 0$, we have $z_1 \to \omega(1+\omega)^{-1}$ instead. This solution holds only if $\omega \neq 0$. The point $\omega = 0$ is a bifurcation point, in which the solution is

$$z_1(t) = \left[\beta_2 t + z_1(0)^{-1}\right]^{-1}. \tag{4.41}$$

Once we show that each $\rho_k \to z_1$, we will have explicit solutions of each ρ_k. Now, $\dot{\bar{\rho}}_k - \dot{z}_1$ can be expressed as

$$\dot{\bar{\rho}}_k - \dot{z}_1 = -(\bar{\rho}_k - z_1)[(\alpha_1+\beta_2)z_1 + \beta_1(1-z_1)] \tag{4.42}$$

Note that the expression in the bracket of Eq 4.42 is always positive and bounded away from 0. Therefore, equation 4.42 shows an exponential convergence of $\bar{\rho}_k$ to z_1 for each k. This principle is depicted in figure 4.3.

Now that we understand how the system orders, we now calculate the consensus time, $T(\boldsymbol{\rho})$, where $\boldsymbol{\rho}$ takes components ρ_k. The procedure is to relate the problem to a complete graph and then use previously established techniques. To simplify this, we take $\alpha_1 = \beta_1 = \beta_2 = 1$. We first establish the backwards equation

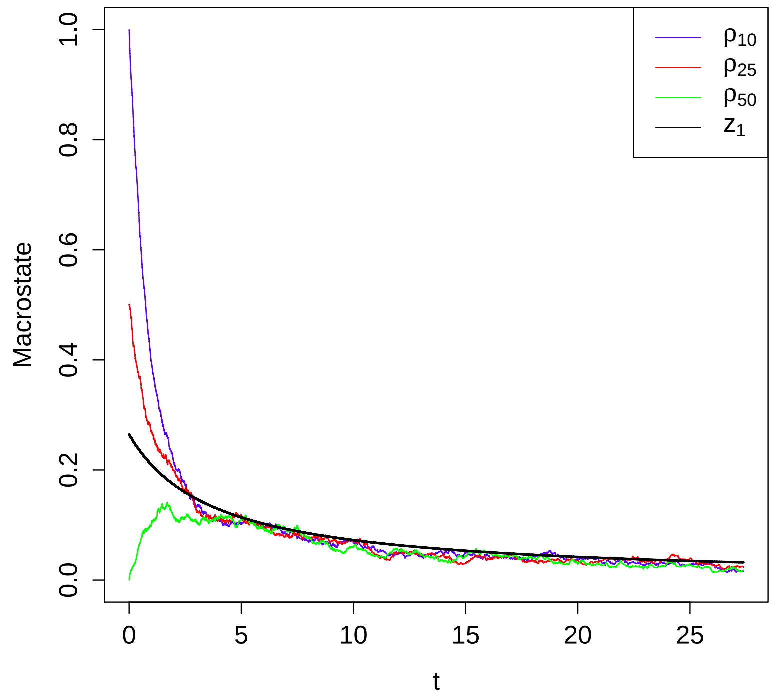

Figure 4.3: The listener only consensus model is simulated for a network with $N = 3000$ with $N_{10} = N_{25} = N_{50} = 1000$. The model was simulated with $\alpha_1 = \beta_1 = \beta_2 = 1$. The initial condition is $\rho_{10} = 1$, $\rho_{25} = 0.5$, and $\rho_{50} = 0$. Also plotted is the exact solution for z_1 given by equation 4.41. This shows that each ρ_k converges to z_1 as the system evolves.

for T:

$$-\frac{1}{N} = \sum_k \Big\{ q(\rho_k, z_1)T(\rho_k - 1/N_k) + [1 - q(\rho_k, z_1) - p(\rho_k, z_1)]T(\rho_k) + p(\rho_k, z_1)T(\rho_k + 1/N_k) \Big\} \quad (4.43)$$

We expand T to two terms as express the forward equation as

$$-1 = \sum_k \left(v_k \frac{\partial T}{\partial \rho_k} + \frac{D_k}{2N_k} \frac{\partial^2 T}{\partial \rho_k^2} \right) \quad (4.44)$$

where

$$v_k = (1-\rho_k)z_1 - \rho_k(1-z_1) - \rho_k z_1 \tag{4.45}$$

$$D_k = (1-\rho_k)z_1 + \rho_k(1-z_1) + \rho_k z_1 \tag{4.46}$$

Now we use the fact that each ρ_k converges to z_1 exponentially for each k. That is, we take each $\rho_k = z_1$ and substitute into the backward equation for T. The change of variables affects the derivatives, giving

$$\frac{\partial}{\partial \rho_k} \to \frac{kN_k}{N\mu_1}\frac{\partial}{\partial z_1}. \tag{4.47}$$

Now the backward equation is

$$-1 = \sum_k \left[v(z_1)\frac{kN_k}{N\mu_1}\frac{dT}{dz_1} + \frac{k^2 N_k}{2N^2\mu_1^2}D(z_1)\frac{d^2T}{dz_1^2}\right]. \tag{4.48}$$

Here,

$$v(z_1) = -z_1^2 \tag{4.49}$$

$$D(z_1) = 2z_1(1-z_1) + z_1^2 \tag{4.50}$$

This simplifies to

$$-1 = v(z_1)\frac{dT}{dz_1} + \frac{\mu_2}{2N\mu_1^2}D(z_1)\frac{d^2T}{dz_1^2}. \tag{4.51}$$

Note that equation 4.51 has exactly the same form as equation 3.54, which we have solved exactly. The primary difference is the appearance of μ_2/μ_1^2. To solve this, we follow the same procedure as the complete graph model and simply replace N with $N\frac{\mu_1^2}{\mu_2}$. This gives that the consensus time is

$$T \sim \sqrt{\frac{\pi^3}{8}\frac{N\mu_1^2}{\mu_2}}. \tag{4.52}$$

There are a few things to note about this equation. Firstly, the form of the solution can be separated into the complete graph solution in equation 3.64 multiplied by the topological parameter $\sqrt{\mu_1^2/\mu_2}$. This may indicate that similar network models can be decomposed into a separable solution. That is, the consensus time for some models might be decomposed into the complete graph solution multiplied by an appropriate topological parameter. Secondly, the fact that $\mu_1^2/\mu_2 \leq 1$ indicates that sparse network topology only decreases the time to consensus. This implies that the complete graph solution is an upper bound for the consensus time for these models. Figure 4.4 depicts this solution for the complete graph and two scale free networks.

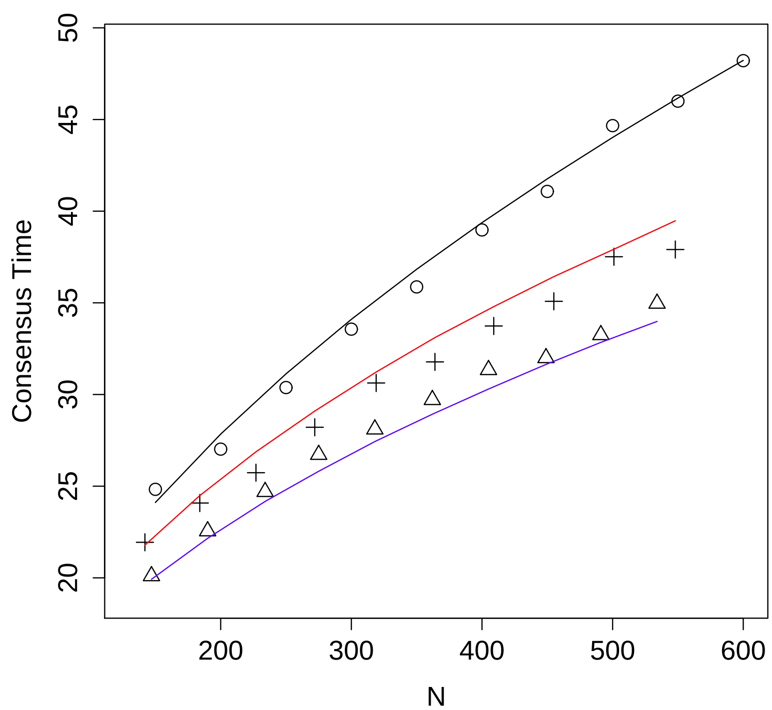

Figure 4.4: Solutions for the consensus time given by equation 3.64 compared with simulation data. The lines are the theoretical prediction and the points are simulation data averaged over 1000 runs. Three networks are considered: the complete graph (black, $\circ$), and two scale free networks with $\nu = 1$ (blue, $\triangle$) and $\nu = 2$ (red, $+$).

4.3.2 Consensus Model: Speaker First

We now consider a different model from the above listener first model. Instead of assigning the first node to be the listener, we designate the first node to be the speaker. The speaker then chooses a neighbor (the listener) to convey its message to. We assume that only the listener updates their state based on this interaction. We also assume that consensus is an absorbing state, as in the listener first case.

These assumptions produce a model with parameters $\alpha_{12}, \beta_{12}, \beta_{22}$. All other parameters equal zero, and for convenience, we drop the second subscript. We briefly outline the analysis, as it follows the same paradigm as the listener first model. The transition probabilities for the model are given by

$$Pr\{\Delta n_k = 1\} = \sum_i \sum_{j \in D_k} \frac{\alpha_1 A_{ij}}{N k_i} \eta_i (1 - \eta_j) \tag{4.53}$$

$$Pr\{\Delta n_k = -1\} = \sum_i \sum_{j \in D_k} \frac{A_{ij}}{N k_i} [\beta_1 (1 - \eta_i) \eta_j + \beta_2 \eta_i \eta_j]. \tag{4.54}$$

We substitute $A_{ij} \to E[A_{ij}]$ and simplify the transition probabilities to obtain

$$p_k(\rho_k, \rho) = \alpha_1 \frac{k}{\mu_1} \frac{N_k}{N} \rho (1 - \rho_k) \tag{4.55}$$

$$q_k(\rho_k, \rho) = \frac{\beta_1 k}{\mu_1} \frac{N_k}{N} (1 - \rho) \rho_k + \frac{\beta_2 k}{\mu_1} \frac{N_k}{N} \rho \rho_k \tag{4.56}$$

We now generate the mean field ODE system for $\bar{\rho}_k$, and obtain

$$\dot{\bar{\rho}}_k = \frac{k}{\mu_1} [\alpha_1 \rho (1 - \rho_k) - \beta_1 (1 - \rho) \rho_k - \beta_2 \rho \rho_k] \tag{4.57}$$

To simplify the analysis, we take $\alpha_1 = \beta_1 = \beta_2 = 1$. Note that the rate of change is proportional to k. We now multiply equation 4.57 by $\frac{N_k}{k N \mu_{-1}}$ and sum over k. Since $\alpha_1 = \beta_1 = 1$, this yields

$$\dot{z}_{-1} = -\frac{1}{\mu_1 \mu_{-1}} \rho^2 \tag{4.58}$$

Since z_{-1} is monotonically decreasing, the system will globally converge to $\rho_k = 0$. This indicates that the convergence is very slow if the system is near consensus. This is due to the fact that the parameters are chosen at a phase transition, similar to the complete graph. The slow convergence is also indicative of a stable center manifold. Linearizing the system for $\dot{\bar{\rho}}_k$ around $\rho_k = 0$ gives

$$\dot{\bar{\rho}}_k = \frac{k}{\mu_1}(\rho - \bar{\rho}_k) \tag{4.59}$$

This system has a Jacobian matrix given by

$$\begin{aligned} \frac{\partial \dot{\bar{\rho}}_k}{\partial \bar{\rho}_l} &= \frac{kN_l}{N\mu_1} \qquad\qquad l \neq k \\ \frac{\partial \dot{\bar{\rho}}_k}{\partial \bar{\rho}_k} &= \frac{k}{\mu_1}\left(\frac{N_k}{N} - 1\right). \end{aligned} \tag{4.60}$$

It is a simple exercise to show that **1** is an eigenvector of the above Jacobian, whose eigenvalue is 0. In addition, by the Gershgorin circle theorem [59], all other eigenvalues have negative real part. This indicates that the system tends to the center manifold near consensus, and the system then converges slowly to consensus. Furthermore, the eigenvector **1** indicates that each ρ_k tends to the same value. Making this substitution into equation 4.58 gives

$$\dot{z}_{-1} = -\frac{1}{\mu_1\mu_{-1}}z_{-1}^2, \tag{4.61}$$

which has solution,

$$z_{-1} = \left(C + \frac{t}{\mu_1\mu_{-1}}\right)^{-1} \tag{4.62}$$

for some constant C. Now, we determine the consensus time by writing the backwards equation for T. The equation takes the form given in equation 4.44, however v_k and D_k are given by

$$v_k = \frac{k}{\mu_1}[\rho(1-\rho_k) - (1-\rho)\rho_k - \rho\rho_k] \tag{4.63}$$

$$D_k = \frac{k}{\mu_1}[\rho(1-\rho_k) + (1-\rho)\rho_k + \rho\rho_k]. \tag{4.64}$$

We now apply the idea that each ρ_k tends to a common value given by the slow variable z_{-1}. We make the change of variables $\rho_k = z_{-1}$. The partial derivatives are now transformed by

$$\frac{\partial}{\partial \rho_k} \to \frac{k^{-1} N_k}{N \mu_{-1}} \tag{4.65}$$

Substituting these into equation 4.44 yields the ODE for the consensus time given by

$$-\mu_1 \mu_{-1} = v(z_{-1}) \frac{dT}{dz_{-1}} + \frac{1}{2N} D(z_{-1}) \frac{d^2 T}{dz_{-1}^2}, \tag{4.66}$$

where

$$v(z_{-1}) = -z_{-1}^2 \tag{4.67}$$

$$D(z_{-1}) = 2z_{-1}(1 - z_{-1}) + z_{-1}^2. \tag{4.68}$$

The reduced equation for T is remarkably similar to equation 3.54. Thus, we have translated the corresponding problem to a 2-urn case with a single topological parameter $\mu_1 \mu_{-1}$. Following the same steps as the complete graph case, we have shown that the consensus time for the speaker first model is

$$T \sim \mu_1 \mu_{-1} \sqrt{\frac{\pi^3 N}{8}} \tag{4.69}$$

It is simple to show that for every degree sequence, $\mu_1 \mu_{-1} \geq 1$. This implies that this sparse graph structure will only slow down the rate of convergence, which is in contrast to the listener first case which showed a decrease in consensus time. Since the listener first model converges more quickly, it appears that imposing one's state upon others is less effective in achieving consensus than listening to the messages of others. Also, the solution once again exhibits a "separable" form, in which a single topological parameter is multiplied by the complete graph solution.

CHAPTER 5
Exact Solution of the Multi-State Voter Model on the Complete Graph

In the binary case of the voter model studied in chapter 2, each node in a network is endowed with one of two states. In a single update, a node is chosen randomly and adopts the state of a randomly chosen neighbor. Although the model typically specifies that the states are binary, we study the case in which there are initially M states [41, 60–62]. This is only one avenue of extended study of the voter model. Other variants that have been considered include nonlinear update probabilities [63–67] and the introduction of committed minorities (zealots) [68,69]. The majority vote model is similar to the voter model, with the key difference being that the opinion of every neighbor is considered during an update [70–72].

We assume that M can take any value from 2 to N, where N is the number of nodes in the network. As the system evolves, it inevitably eliminates a state completely. That is, there will almost surely be $M-1$ distinct states in the network at some finite future time. This process is repeated until consensus is reached and one opinion dominates the network.

We solve this model by casting it as an equivalent urn model. Unlike the binary voter model, there are M urns in the system, but the interaction rules remain the same. Using innovative methods of analysis, we will confirm that when $M = N$, the expected time to consensus is barely larger than when $M = 2$ [41]. Furthermore, it will be shown that $O(N)$ opinion states will be eliminated in $O(1)$ time. These results indicate that the opinions of individuals quickly condense together into a few dominating groups and that the system reaches consensus at a much slower rate. We will also solve for the variance of the consensus time and show that it is $\frac{1}{3}(\pi^2 - 9)(N-1)^2$.

Portions of this chapter previously appeared as: W. Pickering and C. Lim, "Solution of the multistate voter model and application to strong neutrals in the naming game," *Phys. Rev. E*, vol. 93, no. 3, p. 032318, Mar. 2016.

5.1 The Multi-State Voter Model

We assume throughout that there are N nodes and that every node is connected to all other nodes. We apply the model on the complete graph primarily for analytical tractability. Each node in the network is initially endowed with one of M possible opinion states denoted by $A_1, A_2, \ldots, A_M$. We define the components of $\mathbf{n}(m)$ to be total number of nodes with opinion A_j. The voter model prescribes the random walk for the macrostate vector $\mathbf{n}$. That is, we can write for time step m,

$$\begin{pmatrix} n_1(m+1) \\ \vdots \\ n_M(m+1) \end{pmatrix} = \begin{pmatrix} n_1(m) \\ \vdots \\ n_M(m) \end{pmatrix} + \begin{pmatrix} \Delta n_1(m) \\ \vdots \\ \Delta n_M(m) \end{pmatrix}. \tag{5.1}$$

Here Δn contains the random nature of the walk at time step m. In a single update, $\Delta n_i = 1$ and $\Delta n_j = -1$ for some i and j, which implies that $\Delta \mathbf{n} = \mathbf{e}_i - \mathbf{e}_j$ for standard basis vectors $\mathbf{e}_k$. The probability that $\Delta \mathbf{n}$ takes this value is prescribed by the rules of the voter model and is given by

$$\Pr\{\Delta \mathbf{n} = \mathbf{e}_i - \mathbf{e}_j | \mathbf{n}(m) = \boldsymbol{\alpha}\} = \frac{\alpha_i \alpha_j}{N(N-1)}. \tag{5.2}$$

This accounts for all lazy steps in the system as well since there is a nonzero probability that $\Delta \mathbf{n} = \mathbf{0}$.

We define the macro-state probability distribution by $a_{\boldsymbol{\alpha}}^{(m)} = \Pr\{\mathbf{n}(m) = \boldsymbol{\alpha}\}$. We now define a generating function for the probability distribution of macro-states at time step m as

$$Q^{(m)}(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|=N} a_{\boldsymbol{\alpha}}^{(m)} \boldsymbol{x}^{\boldsymbol{\alpha}}. \tag{5.3}$$

The vector power is interpreted in the sense of the multi-index notation of Schwartz [51], which will be used extensively. This generating function allows us to very easily find a succinct expression for the Markov transition matrix for a single step of the multi voter model. The form of the generating function allows us to determine the shift and differentiation properties of $Q^{(m)}$. These properties are given by $\frac{\alpha_i \alpha_j}{N(N-1)} a_{\alpha}^{(m)} \longrightarrow \frac{x_i x_j}{N(N-1)} Q_{x_i x_j}^{(m)}$ and $a_{\boldsymbol{\alpha}-\boldsymbol{e}_i+\boldsymbol{e}_j}^{(m)} \longrightarrow \frac{x_i}{x_j} Q^{(m)}$ [15, 25–27]. Using these properties, we can rewrite the spectral problem as an equivalent partial differential equation for the generating function of the macro-state probability distribution as

$$Q^{(m+1)} - Q^{(m)} = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \frac{(x_i - x_j)^2}{N(N-1)} \frac{\partial^2 Q^{(m)}}{\partial x_i \partial x_j}. \tag{5.4}$$

This constitutes a transition matrix that we wish to diagonalize. Given the diagonalization of the transition matrix, we can find all future macrostate probability distributions explicitly, which yield several exact solutions. To accomplish this, we proceed to solve for all of its eigenvalues and eigenvectors.

5.2 Spectral Solution

We can solve the partial differential equation given in equation 5.4 exactly for all eigenvalues and eigenfunctions. To do this, we write the eigenvalue problem in generating function form. For eigenvalue λ with eigenvector $\mathbf{v}$ with components $c_{\boldsymbol{\alpha}}$. Let

$$G(\mathbf{x}) = \sum_{|\boldsymbol{\alpha}|=N} c_{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}} \tag{5.5}$$

be the generating function for the eigenvector $\mathbf{v}$. Furthermore, we require that each component in the vector $\boldsymbol{\alpha}$ is non-negative. This is because the index $\boldsymbol{\alpha}$ directly represents the number of individuals with each opinion type. We can rewrite the eigenvalue problem for equation 5.4 as

$$N(N-1)(\lambda - 1)G = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} (x_i - x_j)^2 \frac{\partial^2 G}{\partial x_i \partial x_j}. \tag{5.6}$$

We solve for both λ and G by utilizing a linear change of variables $\mathbf{x} \to \mathbf{u}$ and $G(\mathbf{x}) = H(\mathbf{u})$. Since the change of variables is linear, we expect H to have the same form as G. So we define $b_{\boldsymbol{\alpha}}$ so that

$$H(\mathbf{u}) = \sum_{|\boldsymbol{\alpha}|=N} b_{\boldsymbol{\alpha}} \mathbf{u}^{\boldsymbol{\alpha}}. \tag{5.7}$$

The change of variables is chosen so that the resulting difference equation for $b_{\boldsymbol{\alpha}}$ is explicit. Having an explicit equation for $b_{\boldsymbol{\alpha}}$ will allow us to to find all eigenvalues and eigenvectors exactly. The change of variables that accomplishes this is given to be

$$u_1 = x_1 - x_M \tag{5.8}$$

$$\vdots \tag{5.9}$$

$$u_{M-1} = x_{M-1} - x_M \tag{5.10}$$

$$u_M = x_M. \tag{5.11}$$

With this change of variables we can write equation 5.6 as

$$N(N-1)(\lambda-1)H = \sum_{i=1}^{M-1}\left[u_i^2\left(H_{u_iu_M} - \sum_{j=1}^{M-1} H_{u_iu_j}\right) + \sum_{j=i+1}^{M-1}(u_i-u_j)^2 H_{u_iu_j}\right]. \tag{5.12}$$

To simplify this equation, we use the following identity:

$$\sum_{i=1}^{M-1}\sum_{j=i+1}^{M-1}(u_i^2+u_j^2)H_{u_iu_j} = \sum_{i=1}^{M-1}\left(\sum_{j=1}^{M-1} u_i^2 H_{u_iu_j} - u_i^2 H_{u_iu_i}\right). \tag{5.13}$$

The proof of this identity is given at the end of this chapter. Applying the identity and canceling like terms reduces equation 5.13 to

$$N(N-1)(\lambda-1)H = \sum_{i=1}^{M-1}\left[-u_i^2 H_{u_iu_i} - \sum_{j=i+1}^{M-1} 2u_iu_j H_{u_iu_j} + u_i^2 H_{u_iu_M}\right]. \tag{5.14}$$

We now rewrite this as a difference equation for the coefficients of H. By equation 5.7 we obtain

$$N(N-1)(\lambda-1)b_{\boldsymbol{\alpha}} = \sum_{i=1}^{M-1}\left[-\alpha_i(\alpha_i-1)b_{\boldsymbol{\alpha}} - \sum_{j=i+1}^{M-1} 2\alpha_i\alpha_j b_{\boldsymbol{\alpha}} + (\alpha_i-1)(\alpha_M+1)b_{\boldsymbol{\alpha}-\boldsymbol{e_i}+\boldsymbol{e_M}}\right]. \tag{5.15}$$

Let $w(\boldsymbol{\alpha}) = \sum_{i=1}^{M-1}\alpha_i$. We use this to reduce equation 5.15 to an explicit form given by

$$b_{\boldsymbol{\alpha}} = \frac{(\alpha_M + 1)\sum_{i=1}^{M-1}(\alpha_i - 1)b_{\boldsymbol{\alpha}-\boldsymbol{e_i}+\boldsymbol{e_M}}}{N(N-1)(\lambda - 1) + w(\boldsymbol{\alpha})[w(\boldsymbol{\alpha}) - 1]}. \tag{5.16}$$

Recall that each $\alpha_i \geq 0$.

Observe that if equation 5.16 is nonsingular for every $\boldsymbol{\alpha}$, then every $b_{\boldsymbol{\alpha}} = 0$. Since this corresponds to the trivial solution to the eigenvalue problem, we discard these solutions. Requiring a singularity in equation 5.16 implies that the eigenvalues for the transition matrix of the M state voter model are

$$\lambda_{w(\boldsymbol{\alpha})} = 1 - \frac{w(\boldsymbol{\alpha})[w(\boldsymbol{\alpha}) - 1]}{N(N-1)}. \tag{5.17}$$

Since many values of $\boldsymbol{\alpha}$ will yield the same value for $w(\boldsymbol{\alpha})$, it is clear that there will be many repeated eigenvalues. Given that $w(\boldsymbol{\alpha})$ ranges from 0 to N, the set of eigenvalues is the same for all values of M. However, the multiplicities of each eigenvalue will vary with M.

The components of the eigenvectors can be found by transforming back from $H(\boldsymbol{u})$ to $G(\boldsymbol{x})$. This will yield a relationship between $c_{\boldsymbol{\alpha}}$ and b_α. Using generating function techniques, this relationship is given by

$$c_{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\beta}|=N} b_{\boldsymbol{\beta}}(-1)^{\alpha_M - \beta_M} \prod_{i=1}^{M-1} \binom{\beta_i}{\alpha_i}. \tag{5.18}$$

Here $\boldsymbol{\beta}$ is a multi-index that has M non-negative components, similar to $\boldsymbol{\alpha}$. The mathematical derivation of equation 5.18 is given at the end of this chapter.

5.3 Applications

With the solution to the spectral problem available, we can exactly calculate several quantities and estimate others. Below we define and calculate the expected collapse times, the moments of consensus time, the expected local times, and the expected number of states over time. We also consider the special case when $M = N$, which is when all nodes begin with a distinct opinion. The connection between the spectral problem and these exact solutions is as follows. With all eigenvalues and eigenvectors, we can explicitly diagonalize the Markov transition matrix for the macrostates of the system. The probability of achieving each macrostate governs each of the following quantities, so having an exact m-step propagator allows us to exactly calculate these solutions.

5.3.1 Moments of Collapse Times

The collapse times τ_k are the amount of scaled time m/N until a state is eliminated from the system. That is, if there are k states in the system, then the collapse time is the time until only $k-1$ states are present. Once eliminated, a state can never be reintroduced in the system. This process is repeated until each individual adopts a single consensus state.

Here we estimate the collapse times for each k. We do this by estimating the probability that k have survived by time m. We call this the survival probability and denote it by $S_k(m)$. Given that the system is martingale and that the solution of the spectral problem is known, the survival probability for M states can be bounded as follows:

$$S_k(m) = O(\lambda_k^m). \tag{5.19}$$

One can verify that the dominant eigenvalues are λ_k by considering the eigenvector $c_{\boldsymbol{\alpha}} = 1$ and assuming that the index $\boldsymbol{\alpha}$ corresponds to macrostates where k states have survived. This eigenvector provides a uniform upper bound for the survival probability the system.

To find the moments of the time to collapse, we use this to estimate the probability that the system collapses at time m. This is equal to the difference between the survival probabilities from time step $m-1$ to m. That is, by equation 5.19, the probability of collapse at time m is $O[\lambda_k^{m-1}(1-\lambda_k)]$. Therefore, we can write the pth moments of collapse time as

$$E[\tau_k^p] = (1-\lambda_k)\sum_{m=1}^{\infty} O\left(\lambda_k^m \left(\frac{m}{N}\right)^p\right) \tag{5.20}$$

$$= O\left\{p!\left[\frac{N-1}{k(k-1)}\right]^p\right\}. \tag{5.21}$$

The estimate varies for various initial conditions, but is asymptotically correct for all N, k, and p. We will make use of this result when determining the asymptotic behavior of the moments of consensus time and the expected number of states over time.

5.3.2 Moments of Consensus Time

The consensus time τ is the amount of scaled time until every individual adopts a single opinion. All consensus states are absorbing, so once this state has been achieved, all dynamics in the system halt. We can use the solution to the spectral problem to find all moments of the consensus time. Furthermore, we will use estimates to find the asymptotic behavior for large N.

To find the consensus time, we define $l^{(m)}$ to be the probability that the system reaches consensus at time m. This is equal to the probability that the system has only one individual that has a different state than all of the others and then adopts the majority opinion. Therefore, we write

$$l^{(m)} = \frac{1}{N} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} a^{(m)}_{\boldsymbol{e_i}+(N-1)\boldsymbol{e_j}}. \tag{5.22}$$

Given the solution of the spectral problem, we can represent the macrostate probability as

$$a^{(m)}_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}} d_{\boldsymbol{\beta}} \lambda^m_{\boldsymbol{\beta}} [\mathbf{v}_{\boldsymbol{\beta}}]_{\boldsymbol{\alpha}}, \tag{5.23}$$

where $d_{\boldsymbol{\alpha}}$ is the initial distribution expressed in the eigenbasis. Let

$$s_{\boldsymbol{\beta}} = \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} d_{\boldsymbol{\beta}} [\mathbf{v}_{\boldsymbol{\beta}}]_{\boldsymbol{e_i}+(N-1)\boldsymbol{e_j}}. \tag{5.24}$$

With this, the moments of the consensus time are given by

$$E[\tau^p] = \sum_{m=1}^{\infty} l^{(m)} m^p \tag{5.25}$$

$$= \frac{1}{N} \sum_{m=1}^{\infty} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} \sum_{\boldsymbol{\beta}} d_{\boldsymbol{\beta}} \lambda_{\boldsymbol{\beta}}^{m} [\mathbf{v}_{\boldsymbol{\beta}}]_{\boldsymbol{e_i}+(N-1)\boldsymbol{e_j}} m^p \tag{5.26}$$

$$\sim \frac{1}{N} \sum_{\boldsymbol{\beta}} \frac{p! s_{\boldsymbol{\beta}}}{(1-\lambda_{\boldsymbol{\beta}})^{p+1}}. \tag{5.27}$$

This is an exact solution for the moments of consensus time. Note that the quantity $s_{\boldsymbol{\beta}}$ depends on the initial distribution through $d_{\boldsymbol{\alpha}}$. The eigenvectors, $\boldsymbol{v}_{\boldsymbol{\beta}}$ can be determined componentwise by equation 5.18.

The formula given in equation 5.27 is exact for all N, p, and initial conditions. We now extract asymptotic information about the moments of consensus time. In particular, observe that the consensus time is the sum of all collapse times. Therefore,

$$\tau^p = \left(\sum_{k=2}^{M} \tau_k \right)^p \tag{5.28}$$

$$= \sum_{|\boldsymbol{\gamma}|=p} \binom{p}{\boldsymbol{\gamma}} \prod_{k=2}^{M} \tau_k^{\gamma_k}. \tag{5.29}$$

Here $\boldsymbol{\gamma}$ is a vector with components $\gamma_2, \ldots, \gamma_M$. The multi-index notation is used to denote the multinomial coefficients as well. Also, the collapse times are independent random variables. So, when taking the expected value of τ^p, we obtain

$$E[\tau^p] = \sum_{|\boldsymbol{\gamma}|=p} \binom{p}{\boldsymbol{\gamma}} \prod_{k=2}^{M} E[\tau_k^{\gamma_k}]. \tag{5.30}$$

Using the estimate in equation 5.21, this becomes

$$E[\tau^p] = \sum_{|\boldsymbol{\gamma}|=p} \binom{p}{\boldsymbol{\gamma}} \prod_{k=2}^{M} O\left\{\frac{\gamma_k!}{[N(1-\lambda_k)]^{\gamma_k}}\right\} \tag{5.31}$$

$$= p!N^{-p} \sum_{|\boldsymbol{\gamma}|=p} \prod_{k=2}^{M} O\left[\frac{1}{(1-\lambda_k)^{\gamma_k}}\right] \tag{5.32}$$

$$= O\left(p!N^p \sum_{|\boldsymbol{\gamma}|=p} \prod_{k=2}^{M} \left[\frac{1}{k(k-1)}\right]^{\gamma_k}\right) \tag{5.33}$$

$$= O\left(p!N^p 2^{-p} \sum_{|\boldsymbol{\gamma}|=p} \prod_{k=3}^{M} \left[\frac{2}{k(k-1)}\right]^{\gamma_k}\right). \tag{5.34}$$

We take the big O outside of the product because as the system evolves, the macrostate probability distribution tends to the uniform distribution, which corresponds to the dominant eigenvalue in the system. This means that the estimate given in equation 5.21 without the big O is the exact solution. Furthermore, the moments are bounded by the dynamics when $M = N$, which is examined in Sec. 5.3.5. The initial condition also may provide further dependence on M, however this dependence is bounded, which does not affect the validity of the result.

Let

$$\eta(M,p) = \sum_{|\boldsymbol{\gamma}|=p} \prod_{k=3}^{M} \left[\frac{2}{k(k-1)}\right]^{\gamma_k} \tag{5.35}$$

with

$$\boldsymbol{\gamma} = (\gamma_2, \gamma_3, \ldots, \gamma_M). \tag{5.36}$$

We can therefore write

$$E[\tau^p] = O[p!N^p 2^{-p} \eta(M,p)]. \tag{5.37}$$

We will now provide some of the fundamental properties of $\eta(M,p)$. Intuitively, the meaning of $\eta(M,p)$ is the correction made to the estimate by changing M. First, we show that $\eta(M,p)$ is bounded above:

$$\eta(M,p) \leq \sum_{0\leq\gamma\leq p} \prod_{k=3}^{M} \left[\frac{2}{k(k-1)}\right]^{\gamma_k} \tag{5.38}$$

$$= \prod_{k=3}^{M} \frac{1-\left[\frac{2}{k(k-1)}\right]^{p+1}}{1-\left[\frac{2}{k(k-1)}\right]} \tag{5.39}$$

$$\leq \prod_{k=3}^{M} \frac{k(k-1)}{k(k-1)-2} \tag{5.40}$$

$$= 3\left(\frac{M-1}{M+1}\right) \tag{5.41}$$

This result shows that the moments of consensus time can be estimated uniformly in M by $O(p!N^p 2^{-p})$ (See Fig. 1). The dependence on M affects an $O(1)$ factor of the moments of consensus time that one may not wish to casually ignore. For instance, when $M = 2$, we have $\eta(2,p) = 1$, whereas $\eta(M,p) \leq 3$ for large M. This suggests that the uniform estimate can be up to three times as high as the exact solution as M changes. Furthermore, for fixed M and as $p \to \infty$, the estimate given by equation 5.41 is exact. Therefore, we have that

$$\eta(M,\infty) = 3\left(\frac{M-1}{M+1}\right). \tag{5.42}$$

For the first and second moments, evaluations of $\eta(M,p)$ are also given:

$$\eta(M,1) = 2\left(1-\frac{1}{M}\right) \tag{5.43}$$

$$\eta(M,2) = \frac{2}{3}(\pi^2-9) + 2\left(1-\frac{1}{M}\right)^2 - \frac{2}{3M^3} + O\left(M^{-4}\right). \tag{5.44}$$

Equation 5.43 shows that the expected consensus time is always $O(N)$ regardless of the number of initial states M. These particular cases will be used explicitly in Sec. 5.3.5 when studying the exact solutions of the moments of the consensus.

For small M, $\eta(M,p)$ can be easily calculated exactly. For $M = 2, 3, 4$ we have

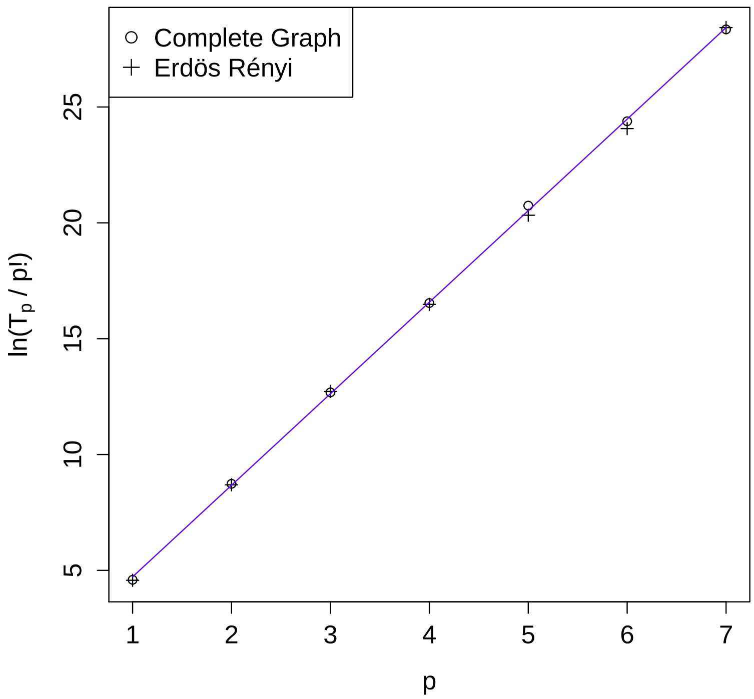

Figure 5.1: Simulation of the voter model for $N = 100$ and $M = 50$ on complete and Erdős-Rényi networks $G(N, 1/3)$ plotted with p. For each p, the simulation is averaged over $1\,000$ runs. Since η is bounded, equation 5.37 predicts a linear relationship with p with a slope of $\ln(100) - \ln(2) \approx 3.912$. The best fit line for the data is given, which has slope 3.951. Furthermore, the data for these sparse networks accurately follow the complete graph paradigm.

$$\eta(2, p) = 1, \tag{5.45}$$

$$\eta(3, p) = \frac{3}{2} - \frac{3^{-p}}{2}, \tag{5.46}$$

$$\eta(4, p) = \frac{9}{5} - 3^{-p} + \frac{6^{-p}}{5}. \tag{5.47}$$

Note that as $p \to \infty$, these solutions for $\eta(M, p)$ exactly match the upper bound given in equation 5.41. The dependence on p for each $\eta(M, p)$ always takes the form

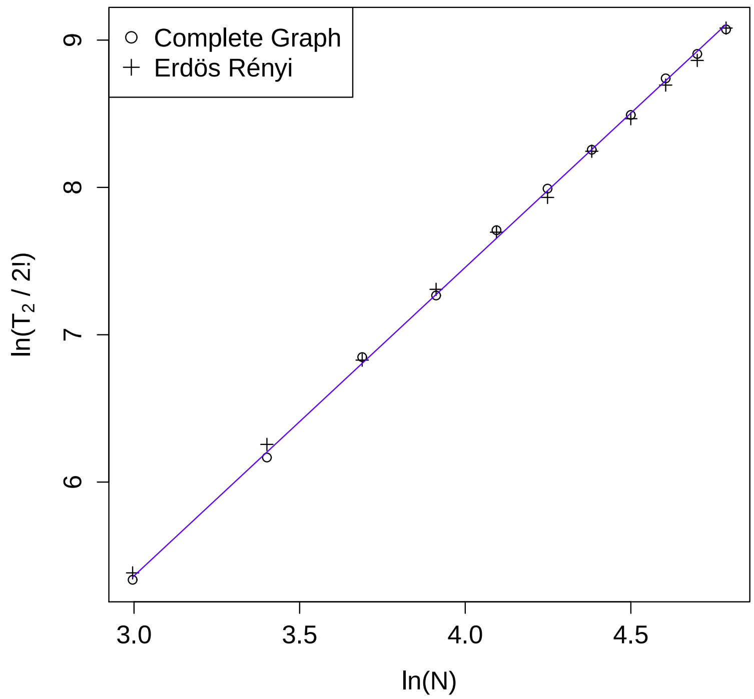

Figure 5.2: Simulation of the voter model on complete and Erdős-Rényi networks $G(N, 1/3)$ plotted with $\ln N$. Data are averaged over $2\,000$ runs with $p = 2$ and $M = 20$. For the complete graph, equation 5.37 predicts a linear relationship between the second moment and $\ln N$ with a slope of 2. The best fit line for the data on the complete graph is given, which has a slope of 2.094. Data for these sparse networks are also accurately predicted by complete graph results.

of an exponential attraction to the upper bound in equation 5.41.

5.3.3 Expected Local Times

The local time is defined as the amount of scaled time m/N spent at each macrostate $\boldsymbol{n}$ prior to consensus. If $M_{\boldsymbol{\alpha}}(m)$ is the number of times state $\boldsymbol{n} = \boldsymbol{\alpha}$ has been visited by time m, then one can construct a random walk model for each $M_{\boldsymbol{\alpha}}$. That is, we write

$$M_{\boldsymbol{\alpha}}(m+1) = M_{\boldsymbol{\alpha}}(m) + \Delta M_{\boldsymbol{\alpha}}(m). \tag{5.48}$$

The expected local time, therefore, is $E[M_{\boldsymbol{\alpha}}(\infty)]$ with this notation. Taking the expected value of equation 5.48 and summing from $m = 0$ to $m = \infty$, we get

$$E[M_{\boldsymbol{\alpha}}(\infty)] = E[M_{\boldsymbol{\alpha}}(0)] + \sum_{m=0}^{\infty} E[\Delta M_{\alpha}(m)]. \tag{5.49}$$

Now $M_{\boldsymbol{\alpha}}(0) = 1$ if $\boldsymbol{n}(0) = \boldsymbol{\alpha}$ and 0 otherwise. Therefore, $E[M_{\boldsymbol{\alpha}}(0)] = a_{\boldsymbol{\alpha}}^{(0)}$, which is given by the initial condition. Similarly, $\Delta M_{\alpha}(m) = 1$ if $\boldsymbol{n}(m+1) = \boldsymbol{\alpha}$ and 0 otherwise. The probability that $\Delta M_{\alpha}(m) = 1$ is $a_{\alpha}^{(m+1)}$. So, the local time for state $\boldsymbol{n} = \boldsymbol{\alpha}$ is (see Fig. 2)

$$E[M_{\boldsymbol{\alpha}}(\infty)] = a_{\boldsymbol{\alpha}}^{(0)} + \sum_{m=0}^{\infty} a_{\boldsymbol{\alpha}}^{(m+1)} \tag{5.50}$$

$$= \sum_{m=0}^{\infty} a_{\boldsymbol{\alpha}}^{(m)}. \tag{5.51}$$

We use the diagonalization given in equation 5.23 to compute this. We ignore the terms that have eigenvalue 1, however, because these correspond to consensus states. We only consider nonabsorbing states when considering local time. Let $E[\boldsymbol{M}]$ take components $E[M_{\boldsymbol{\alpha}}(\infty)]$. Therefore, the local time reduces to

$$E[\boldsymbol{M}] = \sum_{\substack{\boldsymbol{\beta} \\ \lambda_{\boldsymbol{\beta}} \neq 1}} \frac{d_{\boldsymbol{\beta}} \boldsymbol{v}_{\boldsymbol{\beta}}}{1 - \lambda_{\boldsymbol{\beta}}}. \tag{5.52}$$

The components of $\boldsymbol{M}$ that correspond to consensus states are meaningless, as it is understood that when the system enters a consensus state, the dynamics halt entirely. The other components are exactly equal to the expected local time for their respective macrostates (see Fig. 3).

5.3.4 Expected States Over Time

Given the solution to the spectral problem and the expected collapse times, finding the expected number of states over time $s(t)$ is straightforward. To do this, we sum the collapse times τ_k from $k = s + 1$ to $k = M$. Using equation 5.21, we

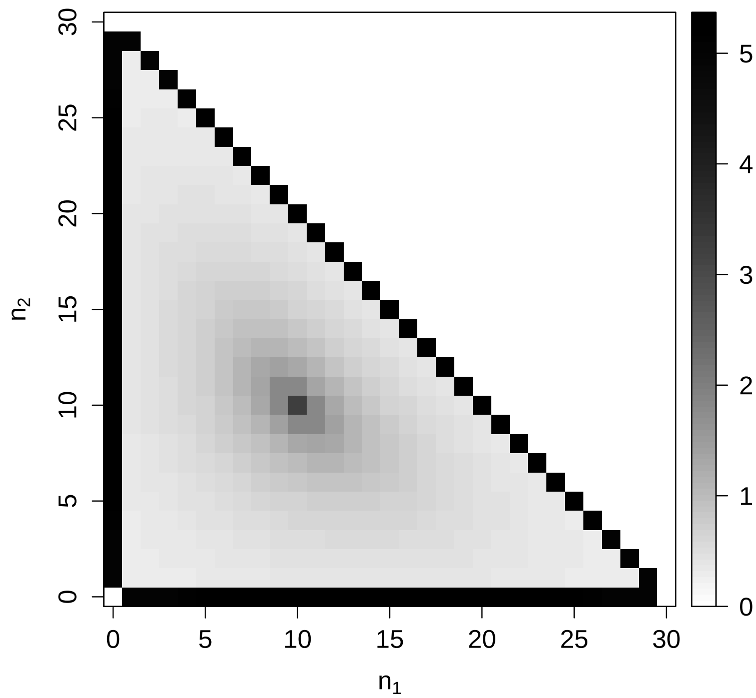

Figure 5.3: Example of expected local times for $N = 30$, and $M = 3$. The initial condition is $n_1(0) = n_2(0) = 10$. Most of the time is spent on the boundary where one of the states had been eliminated. Each macrostate on the boundary has nearly equal local time.

show that the time for s states to exist in the system is

$$t = \sum_{k=s+1}^{M} O\left(\frac{N}{k(k-1)}\right) \tag{5.53}$$

$$= O\left[N\left(\frac{1}{s} - \frac{1}{M}\right)\right] \tag{5.54}$$

Here t is interpreted as the scaled time m/N. Solving for S shows that the expected number of states as a function of time is

$$s(t) = \left(\frac{1}{M} + \frac{ct}{N}\right)^{-1} \tag{5.55}$$

for a constant rate c. This result is in agreement with the literature regarding the multistate voter model [41].

5.3.5 Ordering Dynamics for Uniform Distributions and $M = N$

While the above solutions hold for all M, N, p, and initial condition $a_{\boldsymbol{\alpha}}^{(0)}$, the ordering dynamics of the model reduce significantly in the special case where the initial condition is uniformly distributed. This is because the uniform distribution is also an eigenvector for all M. The eigenvalue that corresponds to this eigenvector is λ_k when there are k distinct opinions in the system. Therefore, the diagonalization reduces considerably, which allows us to find simplified expressions for the above quantities.

A special case of a uniformly distributed initial condition is when $M = N$. This is when each individual adopts a unique personal opinion state prior to global discussion. In this case, there is only one possible initial condition and therefore it is uniformly distributed. Also, notice that in the next iteration of the model, one state will have been eliminated with probability 1. In this time step, two individuals will have the same state while the others possess distinct states. The probability distribution at this time step is uniform (constant). That is, each state is equally likely to have the two individuals than any other state during the first time step.

We can compute τ_k exactly for each k for the uniform case. Because the probability distribution of the macrostates is an eigenvector, the estimates we calculated above are exact. In particular, the survival probability is given to be $S_k(m) = \lambda_k^m$. Thus, making this substitution into the derivation given in Sec. 5.3.1, we find that the expected time to collapse from k states to $k-1$ states is given to be exactly

$$E[\tau_k] = \frac{N-1}{k(k-1)}. \tag{5.56}$$

We now use this to find the exact number of states over time. Recall from Sec. 5.3.4 that the time to achieve s states is the sum of collapse times from $k = s+1$ to $k = M$. Therefore, we obtain

$$t = \sum_{k=s+1}^{M} \frac{N-1}{k(k-1)} \tag{5.57}$$

$$= (N-1)\left(\frac{1}{s} - \frac{1}{M}\right) \tag{5.58}$$

Therefore, the expected number of states is given to be

$$s(t) = \left(\frac{t}{N-1} + \frac{1}{M}\right)^{-1}. \tag{5.59}$$

When $M = O(N)$, this shows that $O(N)$ states will be eliminated in $O(1)$ time. For example, take $t = 1$ and $M = N$ and observe that $s(1) \sim N/2$. This shows that the system retains only half of its initial number of states at $t = 1$, which corresponds to a sweep of nodes in the network. For any $t = O(1)$, we find that only a fraction of the initial number of states remain, so $O(N)$ states were eliminated in this time. This shows that these systems quickly converge to $O(1)$ states relative to the consensus time (see Fig. 4).

Taking $s = 1$, the resulting value of t is the expected time to reach consensus. Doing so shows that the expected time to consensus is

$$E[\tau] = \frac{(N-1)(M-1)}{M}. \tag{5.60}$$

For $M = N$, this result also shows that the expected number of interactions between individuals until consensus is reached is exactly $(N-1)^2$ and that the consensus time is close to N. When $M = 2$, the expected consensus time is at most $N \ln 2$ [21], which is not much less than the $M = N$ case.

We also expand the methods in Sec. 5.3.2 to find all moments of consensus time. By combining the observation in equation 5.30 with equation 5.56, we can find that

$$E[\tau^p] = p!(N-1)^p 2^{-p} \eta(M, p). \tag{5.61}$$

By using equation 5.43, note that the expected time to consensus given in equation 5.60 agrees with this result.

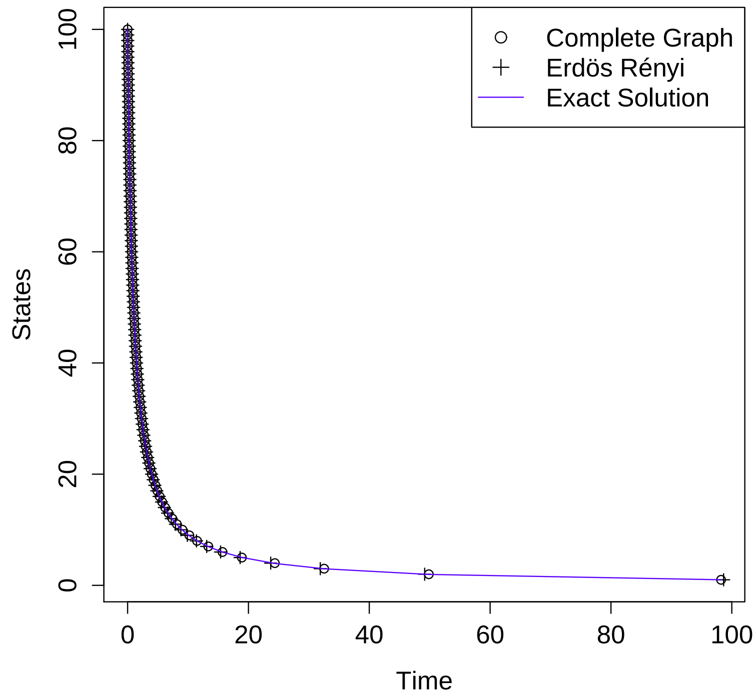

Figure 5.4: Data for the number of states in the system over time plotted for the voter model on the complete graph and Erdős-Rényi networks $G(N, 1/3)$. In each case, we take $N = 100$ and $M = 100$ and average the results over $1\,000$ runs. The exact solution given by equation 5.59 is also plotted. This numerically suggests that the complete graph paradigm accurately describes the dynamics on these sparse networks.

To find the second moment of consensus time, we apply equation 5.44 to equation 5.61 to show that

$$E[\tau^2] = (N-1)^2 \left[\frac{1}{3}(\pi^2 - 9) + \left(1 - \frac{1}{M}\right)^2 - \frac{1}{3M^3} + O\left(\frac{1}{M^4}\right) \right]. \tag{5.62}$$

Figure 5.5 features this result. We can also use this to find the variance of the consensus time. We combine equation 5.62 with the $p = 1$ case to find that

$$\mathrm{Var}(\tau) = (N-1)^2 \left[\frac{1}{3}(\pi^2 - 9) - \frac{1}{3M^3} + O\left(\frac{1}{M^4}\right) \right]. \tag{5.63}$$

This shows that the variance of the consensus time for uniform distributions does not change much with M. Furthermore, taking $M = N$, the first term in the expansion makes for a good estimate, with higher-order terms being $O(N^{-1})$. So, for $M = N$, we have $\mathrm{Var}(\tau) \sim \frac{1}{3}(\pi^2 - 9)(N-1)^2$.

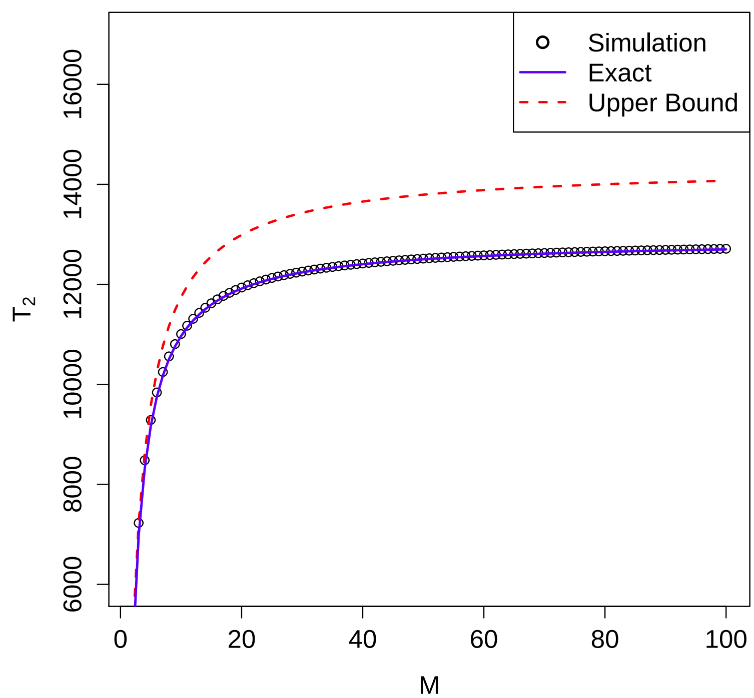

Figure 5.5: Simulation data for the second moment of consensus time T_2 with $N = 100$ over $30\,000$ runs. The exact solution given in equation 5.62 is given as the solid curve. In addition, the upper bound found by applying equation 5.39 to equation 5.61 is plotted as the dashed line. The upper bound overestimates the data by a factor of $1.107\,007$ at most.

5.4 Proof of Identity 5.13

Here we prove the identity given in equation 5.13 that was utilized to solve for all eigenvalues and eigenvectors of the multistate voter model. To begin, consider

$$\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} u_i^2 H_{u_iu_j} = \sum_{i=1}^{M-1}\left(\sum_{j=1}^{i} u_i^2 H_{u_iu_j} + \sum_{j=i+1}^{M-1} u_i^2 H_{u_iu_j}\right). \tag{5.64}$$

For the first double sum on the right-hand side, we interchange the sums. Because the sums are dependent, we obtain

$$\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} u_i^2 H_{u_iu_j} = \sum_{j=1}^{M-1}\sum_{i=j}^{M-1} u_i^2 H_{u_iu_j} + \sum_{i=1}^{M-1}\sum_{j=i+1}^{M-1} u_i^2 H_{u_iu_j}. \tag{5.65}$$

We relabel $i \leftrightarrow j$ in the first sum on the right hand side and separate the $i = j$ term to obtain

$$\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} u_i^2 H_{u_iu_j} = \sum_{i=1}^{M-1}\left(u_i^2 H_{u_iu_i} + \sum_{j=i+1}^{M-1} u_j^2 H_{u_iu_j}\right) + \sum_{i=1}^{M-1}\sum_{j=i+1}^{M-1} u_i^2 H_{u_iu_j}. \tag{5.66}$$

Rearranging terms in this equation shows that

$$\sum_{j=i+1}^{M-1} (u_i^2 + u_j^2) H_{u_iu_j} = \sum_{i=1}^{M-1}\left(\sum_{j=1}^{M-1} u_i^2 H_{u_iu_j} - u_i^2 H_{u_iu_i}\right). \tag{5.67}$$

This concludes the proof of the identity of equation 5.13.

5.5 Calculation of Eigenvector Components

In this section we utilize generating function techniques to relate $b_{\boldsymbol{\alpha}}$ to $c_{\boldsymbol{\alpha}}$. One strategy is to substitute $\boldsymbol{u} \to \boldsymbol{x}$ into the definition of $H(\boldsymbol{u})$ and combine all terms together. By definition, this must equal $G(\boldsymbol{x})$ and so the resulting coefficients must be $c_{\boldsymbol{\alpha}}$. For large M, this becomes cumbersome, so we propose a more general means of finding the relationship for any M using differentiation properties of the generating functions. In particular, for multi-index derivative operator $D^{\boldsymbol{\alpha}}$, where

$|\boldsymbol{\alpha}| = N$, we have

$$D^{\boldsymbol{\alpha}} G(x) = \sum_{|\boldsymbol{\beta}|=N} c_{\boldsymbol{\beta}} (\boldsymbol{\beta})_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}-\boldsymbol{\alpha}}. \tag{5.68}$$

Here $(\boldsymbol{\beta})_{\boldsymbol{\alpha}}$ is the multi-index Pochhammer symbol, which is defined by $(\beta_1)_{\alpha_1}, \ldots, (\beta_M)_{\alpha_M}$. Because we defined $|\boldsymbol{\alpha}| = N$, the only term in the sum that is nonzero is when $\boldsymbol{\beta} = \boldsymbol{\alpha}$. Therefore, $D^{\boldsymbol{\alpha}} G = \boldsymbol{\alpha}! c_{\boldsymbol{\alpha}}$, where $\boldsymbol{\alpha}! = \alpha_1!, \ldots, \alpha_M!$. With this observation, we use the definition of H to obtain

$$\begin{aligned} G(\boldsymbol{x}) &= H(\boldsymbol{u}(\boldsymbol{x})) && (5.69) \\ &= \sum_{|\boldsymbol{\beta}|=N} b_{\boldsymbol{\beta}} \left[\prod_{i=1}^{M-1} (x_i - x_M)^{\beta_i} \right] x_M^{\beta_M} && (5.70) \\ &= \sum_{|\boldsymbol{\beta}|=N} b_{\boldsymbol{\beta}} \left[\prod_{i=1}^{M-1} \sum_{\gamma_i=0}^{\beta_i} \binom{\beta_i}{\gamma_i} (-1)^{\beta_i - \gamma_i} x_i^{\gamma_i} x_M^{\beta_i - \gamma_i} \right] x_M^{\beta_M}. && (5.71) \end{aligned}$$

Simplifying the expression on the right-hand side gives

$$G(\boldsymbol{x}) = \sum_{|\boldsymbol{\beta}|=N} b_{\boldsymbol{\beta}} \sum_{0 \le \boldsymbol{\gamma} \le \boldsymbol{\beta}} \left[\prod_{i=1}^{M-1} \binom{\beta_i}{\gamma_i} x_i^{\gamma_i} \right] \times (-1)^{N-\beta_M-|\boldsymbol{\gamma}|} x_M^{N-|\boldsymbol{\gamma}|}. \tag{5.72}$$

Now we take $D^{\boldsymbol{\alpha}}$ of this equation for $|\boldsymbol{\alpha}| = N$. We found that on the left-hand side, we get $\boldsymbol{\alpha}! c_{\boldsymbol{\alpha}}$. Therefore, we obtain

$$\boldsymbol{\alpha}! c_{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\beta}|=N} b_{\boldsymbol{\beta}} \left[\prod_{i=1}^{M-1} \binom{\beta_i}{\gamma_i} \alpha_i! \right] (-1)^{\alpha_M - \beta_M} \alpha_M!. \tag{5.73}$$

Therefore, we find that

$$c_{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\beta}|=N} b_{\boldsymbol{\beta}} (-1)^{\alpha_M - \beta_M} \prod_{i=1}^{M-1} \binom{\beta_i}{\gamma_i}, \tag{5.74}$$

which is the desired relationship stated in equation 5.18.

CHAPTER 6
Solution of the K-Word Naming Game: The Effect of Opinion Diversity

Similar to the urn models on networks and the multi-state voter model, the K-word naming game on the complete graph reduces to an urn model with more than 2 urns. In the model, there are K words, or opinions, that are present. However, there are $2^K - 1$ states that an individual can adopt, since the model assumes that an individual can adopt multiple opinions simultaneously [7–13]. In this chapter, we will discuss the effect of large K on social properties of the system, such as the consensus time and stability [10, 12, 73–75].

The naming game with committed agents can account for several historical precedents in which the majority opinion was overtaken by a committed minority (e.g., the suffragette movement in the early 20th century, and the adoption of the American civil-rights in 1960's [10]). Such processes are known in sociology under the term minority influence [76]. When the committed minority fraction of the population is small, their opinion will still be suppressed by an existing majority opinion [77]. Yet, when this fraction exceeds a modest tipping point value [10,78,79], the minority opinion will spread.

To gain these insights into the dynamics of social systems, we solve the critical problem of complexity for the naming game. For K opinions, the system of ODEs that describe relative population sizes has $2^K - 1$ equations, which is numerically and analytically difficult to study [12]. Furthermore, if the number of opinions also becomes infinite with N, then these ODE methods fail. By applying more robust methods of analysis, we solve the problem of exponential complexity, and by doing so, demonstrate the potential of solving other highly complex systems by these means.

In the model, there are K words (opinions), which we call $A_1, A_2, \ldots, A_K$.

Portions of this chapter previously appeared as: W. Pickering, B. K. Szymanski, and C. Lim, "Analysis of the high dimensional naming game with committed minorities," *Phys. Rev. E*, vol. 93, no. 5, p. 052311, May 2016.

There are N individuals, each with a word list, which is a set of words. The individuals update their word lists as they change their opinions in response to messages from others. We also assume that any individual may speak to any other individual. This means that the social network considered here is a complete graph, which is a common assumption and corresponds to mean-field network analysis [10, 12, 13, 80, 81], although other networks have also been considered [82, 83]. It has been observed that dynamics of the naming game on real world networks are qualitatively similar to complete graph results [12].

Time is discretized so that one interaction of individuals takes place within a time step. In one step, an individual is chosen uniformly at random to be the speaker and another is chosen uniformly at random as the listener. Let W_s and W_l be the word lists of the speaker and the listener respectively. The speaker chooses a random word, $A_s \in W_s$, to transmit to the listener. If none of the two is committed, they update their word lists according to the following rules:

1. If $A_s \notin W_l$, then $W_l \to W_l \cup \{A_s\}$.
2. If $A_s \in W_l$, then $W_s, W_l \to \{A_s\}, \{A_s\}$ [8].

In addition, we also may include committed agents (aka zealots) in the system. They never update their word lists and only adopt a single word. We consider two cases when these committed minorities are present. We first consider the case when there are n' zealots of one word. Then, we consider the case when there are n' zealots for each word. We show below that there are similar rates of convergence for both cases. The critical fraction of committed agents is the value of n'/N that yields a phase transition in the system. When this fraction of zealots is below this critical value, the opinion of the committed minorities will be suppressed by the majority. When the committed fraction is above the critical value, the minority opinion overcomes the majority. We are also interested in the time until all individuals have the same opinion, which we define as the consensus time.

The measurement of diversity of opinion that we first consider is the entropy of the system. The Shannon entropy in particular measures the uncertainty of a random variable, such as a message [40]. If the message has high Shannon entropy, then a listener has a significant probability of hearing a diverse range of opinions. This also means that there greater competition among the opinions for dominance in the system. There is more dissent, disorder, and disagreement in high entropy systems. Low entropy systems have more consistency in the messages that are transmitted, so there is more agreement and less diversity. These systems are predictable, ordered, and united. Let the probability of speaking A_s be P_s. The Shannon entropy of the system is

$$\mathcal{H} = -\sum_{s=1}^{K} P_s \ln P_s. \tag{6.1}$$

We take the natural logarithm in Eq. 6.1 for convenience. We aim to demonstrate the following principle:

1. The consensus time is expected to increase as $\mathcal{H}$ increases.
2. The critical fraction of committed agents is expected to decrease as $\mathcal{H}$ decreases.

In the case of the voter model with two opinions, the consensus time on the complete graph is exactly equal to $\mathcal{H}$ scaled by N [21], which reinforces the above principle. If there is greater dissent in a population, then it is also easier for a minority of zealots to dominate the system. This reinforces the rule *if divided then conquered* since it is easier to dominate the system in the presence of internal conflict. These claims are demonstrated in Fig. 6.1.

The rate of convergence is central to the solutions that we obtain. Below, we show that for each case we consider, the dominant eigenvalues can be estimated by

$$\lambda = 1 - O\left(\frac{K(\theta + n')}{N^2}\right), \tag{6.2}$$

where θ reflects the state of the system. Even though the naming game is nonlinear, it is shown that the propagator can be subdivided into a sum of many linear operators. This is shown at the end of the chapter.

We start with the simple case when the system does not have committed minorities. Also, we assume that each word has near equal representation in the initial condition. Now, θ can be as large as $O(N)$ since the system saturates itself

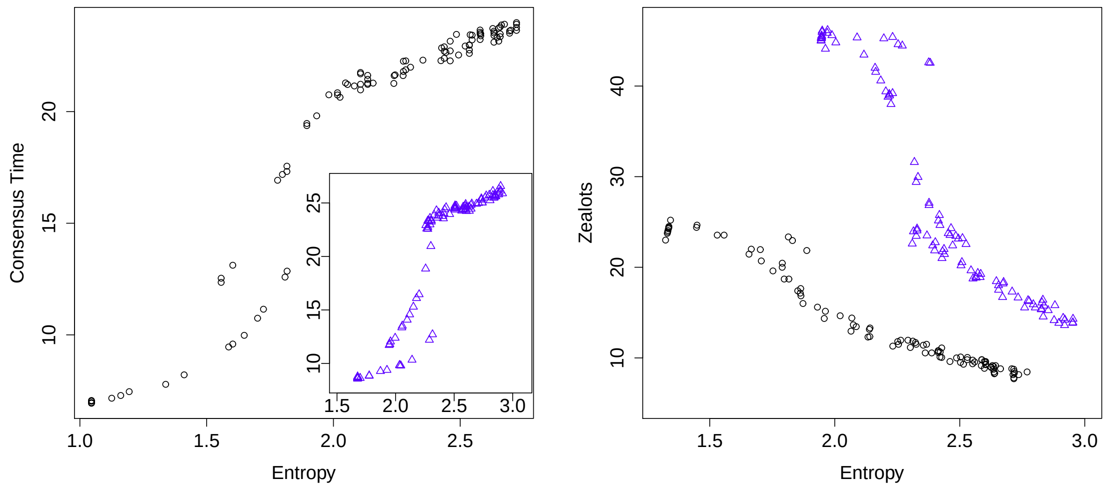

Figure 6.1: Plots of the consensus time (top) and the critical number of zealots (bottom) against the Shannon entropy of various initial conditions. A committed minority of word A_1 is introduced only in the bottom figure. Data are shown for $N = 200$, $K = 20$ ($\circ$) and $N = 400$, $K = 40$ ($\triangle$).

with $O(N)$ individuals with words lists of length 2 or more. This gives $1 - \lambda_K = O(K/N)$ as the rate of convergence.

We define the collapse time as the amount of scaled time until a word is eliminated from the system. Scaled time is the number of discrete time steps divided by the number of nodes in the network. That is, the scaled time t is defined as $t = m/N$. We wish to find the amount of time until a word is expected to be eliminated from the system, then sum these to calculate the consensus time.

If the system is not near consensus (outer region), then the entire probability distribution cannot be estimated by the dominant eigenvalue. Instead, we take the survival probability, which is the probability that there are k words in the system at time t, and set it to $1/N$. When this is the case, it is expected that less than one individual will have one of the K words. Given that $k \leq K$ words are present at scaled time t, the survival probability is λ_k^{tN}. So, setting $\lambda_k^{tN} = 1/N$ and solving for

t gives the collapse time, $\tau_k^{(outer)}$, which is

$$\tau_k^{(outer)} = O\left(\frac{\ln N}{k}\right). \tag{6.3}$$

When the system is near consensus (inner region), the system is diffusion-like. We use the infinite series to calculate the expected value. That is, the collapse time near consensus, $\tau_k^{(inner)}$, is given by

$$\tau_k^{(inner)} = \sum_{m=0}^{\infty} s_m \frac{m}{N} \tag{6.4}$$

where s_m is the probability of collapse. The probability of collapse is the change in the survival probabilities: $s_m \sim \lambda_k^{m-1} - \lambda_k^m$ [15, 24]. Making this substitution gives

$$\tau_k^{(inner)} = O\left(\frac{1}{k}\right). \tag{6.5}$$

As the system approaches consensus, the $\ln N$ factor in the collapse time tends to $O(1)$ as the system transitions between regions.

We estimate the number of words in the system over time as well as the time to consensus by summing the collapse times for their respective regions from $k = S$ to $k = K$. Solving for $S(t)$, which is the number of words in the system over time, yields

$$S_{outer}(t) \leq K \exp\left(-\frac{\alpha t}{\ln N}\right) \tag{6.6}$$

$$S_{inner}(t) \leq K \exp(-\alpha t). \tag{6.7}$$

The inner region converges on a faster time scale than the outer region. The convergence will, however, accelerate as the system approaches consensus. These results are shown numerically in Fig. 6.2.

Now we will estimate the consensus time. For the naming game with two words, the consensus time is $O(\ln N)$ [8, 9, 11, 38]. However, this may increase when the number of words is large. In accordance to the entropy principle, we assume that each word is equally represented to acquire an upper bound on the consensus

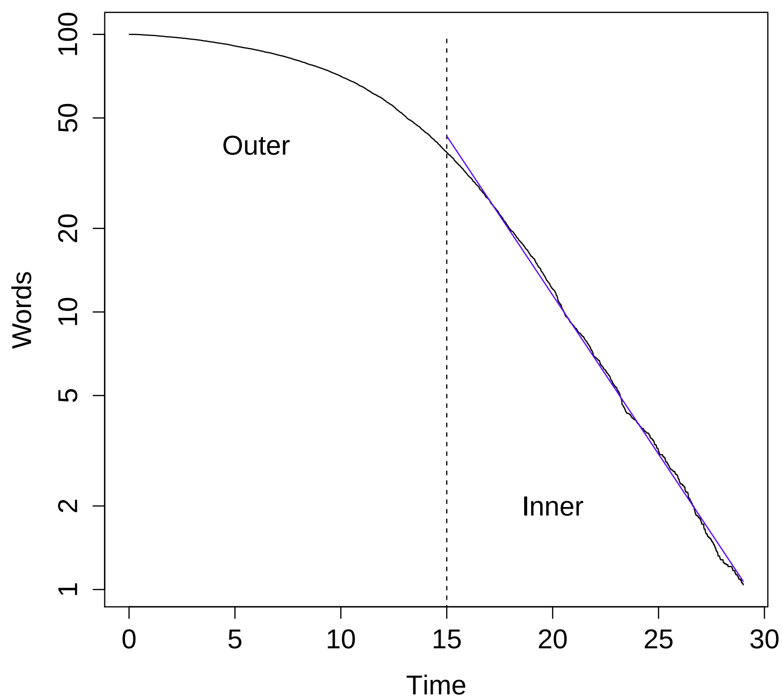

Figure 6.2: Semi-log plot of the number of words in the system as a function of scaled time. Also shown is the best fit for the inner region, which confirms exponential convergence. The naming game is averaged over 100 runs with $N = K = 100$.

time.

To find the consensus time, we estimate the time spent in the outer and inner regions. We take Eqs. 6.6 and 6.7 with $S = 1$ and solve for t. Adding these together gives the consensus time:

$$E[\tau] \sim c_1 \ln N \ln K + c_2 \ln K \tag{6.8}$$

where c_1 and c_2 are constants. This is consistent with known information regarding the case when $K = 2$, for Eq. 6.8 is $O(\ln N)$ for $K = 2$. It also accounts for cases when K takes extreme values. For $K = O(N)$, the consensus time increases to $O(\ln^2 N)$. An example of an extreme K case is given in Fig. 6.3.

Now we consider two cases with zealots in the system. The first case we

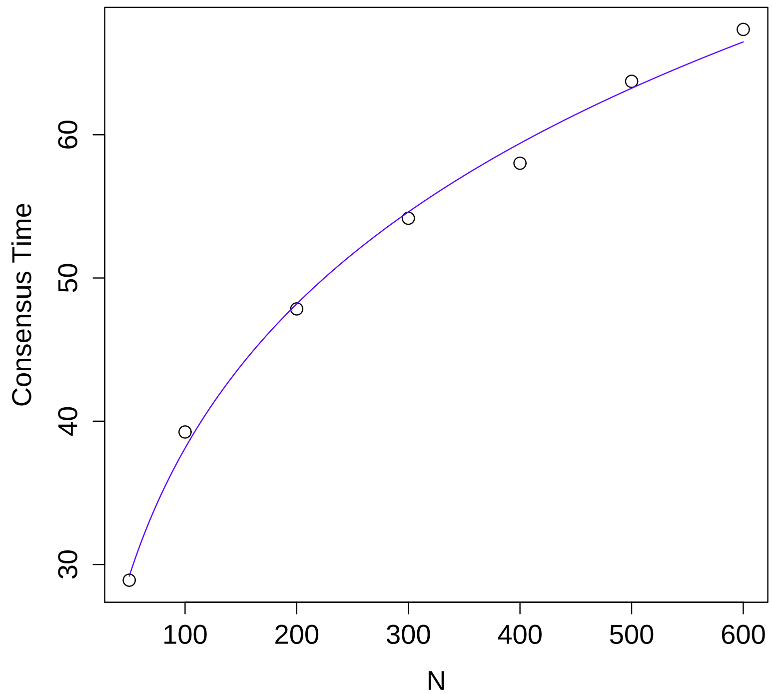

Figure 6.3: Plot of the consensus time averaged over 50 runs of the naming game for various N with $K = N$. Also plotted is the estimate given in Eq. 6.8 fitted to the data. The best fit yields $c_1 \approx 1.18$ and $c_2 \approx 2.84$.

consider is when there is only a committed minority of a single word. Without loss of generality, let us say that there are n' zealots with word A_1. It has been shown for $K = 2$ that when $n'/N \approx 10\%$ or more, there are enough zealots to quickly turn an entire population. Otherwise, the system is trapped in a metastable state, and it takes an exponential time for the population to adopt the zealots' opinion [10, 12].

We seek to extend this to cases when K is arbitrary. Particularly, we consider cases when K is large and the spectral method is required to analyze the system. This problem was briefly discussed in Waagen *et al.* [12] and their conclusion was that the same 10% critical fraction holds for all K and initial conditions to guarantee the zealots dominate the system. Their approach is to consider the worst case initial conditions and show it reduced to the $K = 2$ case. The worst case initial condition minimizes entropy, and according the the entropy principle above, this maximizes the number of zealots required. We take the analysis of Waagen *et al.* [12] a step further by assuming the opposite scenario for the initial condition: each uncommitted community is initially of equal size, which maximizes entropy.

Let C be the number of individuals initially with word A_k, where A_k is not the zealots' opinion. For the case when there are only zealots of a single type, we have $N = (K-1)C + n'$. For fixed N, this gives a dependence on C in terms of K.

To find the phase transition over n', the criterion we use is simple. This occurs when Eq. 6.2 is dominated by a different class of eigenvalues that describe a stationary distribution. This stationary distribution is the metastable state, and the system will converge to it if the rate is higher than the consensus rate. Setting these rates equal to each other gives the phase transition. The rate of convergence to the metastable state is shown below. Setting these rates equal to each other gives

$$1 - \lambda = \frac{a}{N} + \frac{bK}{N^2}, \tag{6.9}$$

where a and b are constants. We take $\theta = O(1)$ in $1-\lambda$ since the system is initially dominated by uncommitted words. Taking Eq. 6.9 and solving for n' gives

$$n'_c = \frac{aN}{K} + b. \tag{6.10}$$

Here, $a, b = O(1)$. This tells us that we expect the number of zealots required to turn a population decays as $1/K$ to a constant. We express this in terms of C by substitution. This produces a nonlinearity in n'_c, which we approximate to provide the following fit:

$$n'_c = \frac{aNC}{N + dC} + b, \tag{6.11}$$

where d is another constant. A comparison of this against simulation data is given in Fig 6.4. This result shows that as the relative sizes of the community grows larger, it takes more zealots to turn the population.

The case where each opinion has zealots follows by a similar argument. If $n' < n'_c$, then one opinion eventually will suppress all others. When $n' > n'_c$, a

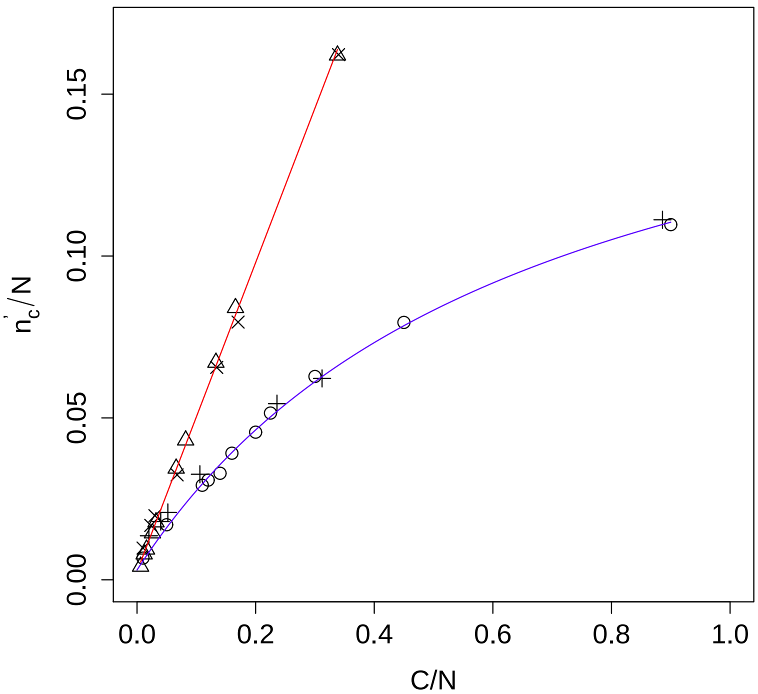

Figure 6.4: Critical fraction of committed agents plotted against C/N for $N = 1000$ $(\circ, \triangle)$ and $N = 500$ $(+, \times)$. Data for a committed minority of a single word is shown $(\circ, +)$ and the best fit of Eq. 6.11 in blue. Also shown is the case when there are committed minorities of every word $(\triangle, \times)$ with the fit of Eq. 6.12 in red.

stalemate develops and no opinion gains dominance. We still apply criterion of Eq. 6.9 along with Eq. 6.2 for the phase transition. This means that the dependence of the critical number of zealots as a function of K has the same form as Eq. 6.10. However, now we have $N = K(C + n')$. When substituting K for C, we obtain

$$n_c' = a'C + b'. \tag{6.12}$$

We use a' and b' to denote different constants from the previous case that are also both $O(1)$. Fig. 6.4 depicts this relationship in practice.

In addition to this innovative approach, these results take previous analysis further by considering the dependence on the initial condition. We demonstrate that the consensus time and the critical number of zealots have distinct correlations between the entropy of the state. This reinforces the rule *"divide and conquer"*, and also suggest that social systems with great dissent can foster many committed minority groups that may block each other from reaching a tipping point, which is high in case of the uncommitted groups sharing a few opinions only. Our results suggest that high opinion diversity among uncommitted individuals changes the dynamics. In such situations, the tipping point can be reached with the number of committed minority members being small and independent of the system size, making the system unstable and quickly transferring to the state in which uncommitted individuals adopt one of the minority opinions.

6.1 Calculation of Rate of Convergence

Our analysis of the naming game is based on the rate of convergence of the system. The rate of convergence is given by the dominant eigenvalues of the transition matrix for the probability distribution of the system. Knowing the rate of convergence, we can estimate the time until a word is eliminated (collapse time) as well as the consensus time. For the case with committed minorities, we can also use this analysis to estimate the number of zealots required until a drastic qualitative change occurs in the system. This is because, the dominant eigenvalues of the transition matrix depend on the number of zealots. When the fraction of committed minorities is high enough, these eigenvalues no longer dominate the ordering of the system. This means that different eigenvectors determine the overall shape of the probability distribution over time, and there is a significant change in qualitative behavior. Once we have the dominant eigenvalues, these solutions become easy to find.

To find the convergence rate, we first express the transition matrix componentwise. Let the $n_W(m)$ be the total number of individuals with word list W at discrete time m, and let the vector $\boldsymbol{n}$ take components n_W. Also let

$$a_{\boldsymbol{\alpha}}^{(m)} = Pr\{\boldsymbol{n}(m) = \boldsymbol{\alpha}\}. \tag{6.13}$$

We seek to express $a_{\boldsymbol{\alpha}}^{(m+1)}$ in terms of $a_{\boldsymbol{\alpha}}^{(m)}$. To do this, we must account for all possible transitions that the model allows. Although this is a complicated task for the general K word naming game, we follow a simplified model to ameliorate this issue while keeping the original qualitative properties intact. In the simplified model, only the listener updates their word list in response to a message from the speaker, as in [77]. This we call the *listener only naming game*. In every simulation, we apply the original naming game rules, which shows that there is still agreement under this modification.

Since we assume that only one individual changes their word list in a given time step, an individual with word list W may transition to having word list W' or vice versa. To account for all transitions in the stochastic matrix, we must consider all pairs of word lists (W_1, W_2) along with their respective transition probabilities. Let D be the set of all pairs of word lists. Also let $\mathcal{L}_I[\cdot]$ be the operator acting on the current macrostate that accounts for the possible transitions involving word pair $I = (W_1, W_2)$. We then write

$$a_{\boldsymbol{\alpha}}^{(m+1)} - a_{\boldsymbol{\alpha}}^{(m)} = \sum_{I \in D} \mathcal{L}_I \left[a_{\boldsymbol{\alpha}}^{(m)} \right]. \tag{6.14}$$

We estimate the rate of convergence of the model by the spectral properties of each $\mathcal{L}_I$. Summing all of them together gives the relative magnitude of $a_{\boldsymbol{\alpha}}^{(m+1)} - a_{\boldsymbol{\alpha}}^{(m)}$, which is the change in probability over a single time step. We wish to find the smallest change in probability possible that retains K words in the system. Since each $\mathcal{L}_I$ corresponds to pairs of word lists that transition to each other, we exhaust each case of pairs of word lists and find the smallest eigenvalues, many of which are zero. The meaning of each case is that we only allow the given pair of word lists, (W_1, W_2) to change their word lists in a given step.

6.1.1 Case 1: $|W_1|, |W_2| \geq 2$

These cases tend to a stationary distribution that is not the consensus state. If we only allow a pair of word lists that contain multiple words, then it is impossible to update the system in such a way that a word is eliminated. The only way for a word to be eliminated is if a listener is the only holder of it and hears and then adopts a familiar word. Since neither W_1 nor W_2 fit this criterion, we take the change to be 0 without loss of generality. Note also that this conclusion applies to the vast majority of cases for large K.

If the system does not converge to consensus, then it converges to the stationary distribution acquired from these cases. It is valuable to understand the behavior of the second largest eigenvalues in these cases, especially when considering zealots. The rate of convergence to the stationary distribution yields the criteria for the phase transition as different sets of eigenvectors starts governing the shape of the system. The stationary distribution in this case is related to the metastable distribution when the number of zealots is small. So, we seek to find the size of the rate of convergence to this stationary distribution.

The only possible means of transition in this case occurs when W_1 and W_2 differ by a single word. Otherwise it is impossible for W_1 and W_2 to transition to each other. Let $W_2 = W_1 \cup \{A_p\}$ and let S_p be the set of word lists that contain A_p. Note that there are K different choices for A_p. Let p_i be the probability of transition from W_1 to W_2 given that $n_{W_1} = i$, which is given by

$$p_i = \left(\sum_{W \in S_p} \frac{n_W}{N|W|} \right) \frac{i}{N-1}. \tag{6.15}$$

Since it is impossible to transition from W_2 to W_1 in the naming game, this constitutes a triangular transition matrix, whose spectrum is $\lambda_k = -p_k$. Let

$$\mu_p = \sum_{W \in S_p} \frac{n_W}{N|W|}, \tag{6.16}$$

which depends on the macrostate of the system and the particular word pair. The total change in probability comes from the sum of the relative changes for each word. That is, we sum Eq. 6.15 for $p = 1 \ldots K$. In doing so, we find that the sum of μ_p is at most $O(1)$ if the sum of n_W achieves its maximum value of $O(N)$. This yields a total rate of change being proportional to $1/N$ to leading order.

We are also interested in a second term in total change in probability, as it is significant for the naming game with zealots. This is attained by supposing that the sum over μ_p does not achieve its maximum value. If each n_W is only $O(1)$, then the sum of μ_p is $O(K/N^2)$. This matches the leading term for $K = O(N)$, but is smaller for $K = O(1)$. These considerations are utilized when calculating the total rate of convergence.

6.1.2 Case 2: $W_1 = \{A_k\}, |W_2| > 2, A_k \in W_2$

Here we only consider transitions in a word list that contains a single word and a word list that have 3 or more words. The size of the eigenvalues are easy to find in this case since it is only possible for W_2 to become W_1. This is because it is impossible for an individual with only a single word to adopt 3 or more words in a single step. Mathematically, this case corresponds to a triangular transition matrix, whose eigenvalues are the diagonal elements. Let p_i be the probability that an individual with word W_2 hears word A_k and thus transitions to W_1 given that there are i individuals with W_1. Since all other individuals with all other word lists are considered fixed, let

$$\mu_1 = \sum_{W \in S_k \setminus \{W_1 \cup W_2\}} \frac{n_W}{|W|}, \tag{6.17}$$

which is considered constant. Now, we express the transition probability as

$$p_i = \frac{(i + n')(N' - i)}{N(N-1)} + \frac{1}{|W_2|} \frac{N' - i}{N} \frac{N' - i - 1}{N - 1} + \mu_1 \frac{(N' - i)}{N(N-1)} \tag{6.18}$$

where $N' = n_{W_1} + n_{W_2}$, which is conserved here. Also, n' is the number of zealots corresponding to the word A_k. The eigenvalues for this case are $-p_i$, and the smallest eigenvalue that does not correspond to consensus is

$$\lambda \sim -\frac{N' + \mu_1 + n'}{N^2} \tag{6.19}$$

This can be seen by taking $i = N' - 1$. Note that μ_1 and N' captures the dependence on the state of the system on the relative change in probability.

6.1.3 Case 3: $W_1 = \{A_k\}, W_2 = \{A_k, A_l\}$

Here W_1 has only one word and W_2 has two words, one of which is A_k for some k. This is the most dynamic of the cases because W_1 can transition to W_2 and vice versa. Because of the listener only assumption, this constitutes a tridiagonal transition matrix. Let p_i and q_i be the probability n_{W_1} increases and decreases respectively, given that $n_{W_1} = i$. Let

$$\mu_2 = \sum_{W \in S_l \setminus \{W_2\}} \frac{n_W}{|W|} \tag{6.20}$$

and recall the definition of μ_1 from Eq. 6.17. The transition probabilities are then expressed as

$$\begin{aligned} p_i = &\frac{(i+n')(N'-i)}{N(N-1)} + \frac{(N'-i)(N'-i-1)}{2N(N-1)} \\ &+ \mu_1 \frac{N'-i}{N(N-1)}, \end{aligned} \tag{6.21}$$

$$q_i = \frac{i(N'-i)}{2N(N-1)} + \mu_2 \frac{i}{N(N-1)}. \tag{6.22}$$

To find the rate of convergence for this step, we wish to solve the following eigenvalue problem

$$\lambda c_i = p_{i-1}c_{i-1} + (-p_i - q_i)c_i + q_{i+1}c_{i+1} \tag{6.23}$$

In order to solve for all eigenvalues of this problem, we apply the generating function method of Ref. [15], which exactly diagonalized the voter model. We begin by expressing Eq. 6.23 in terms of a generating function $G(x, y)$, which we define as

$$G(x, y) = \sum_{i=0}^{N'} c_i x^i y^{N'-i} \tag{6.24}$$

Using shift and differentiation properties of G, we rewrite Eq. 6.23 as

$$\begin{aligned} N(N-1)\lambda G = &(x - \frac{1}{2}y)(x-y)G_{xy} + (n' + \mu_1)(x-y)G_y \\ &+ \frac{1}{2}y(x-y)G_{yy} - \mu_2(x-y)G_x \end{aligned} \tag{6.25}$$

We solve this by the change of variables $u = x - y$ and $G(x, y) = H(u, y)$. Here, we have

$$H(u, y) = \sum_{i=0}^{N'} b_i u^i y^{N'-i}. \tag{6.26}$$

Making this change gives the equivalent equation for H:

$$\begin{aligned} N(N-1)\lambda H = \left(u^2 - \frac{1}{2}uy\right) H_{uy} - u^2 H_{uu} + \frac{1}{2}uyH_{yy} \\ + (n' + \mu_1)uH_y - (n' + \mu_1 + \mu_2)uH_u. \end{aligned} \tag{6.27}$$

The above written as a difference equation for the coefficients of H gives

$$\begin{aligned} N(N-1)\lambda b_i = -\left[\frac{1}{2}i(N' - i) + i(i-1) + i(n' + \mu_1 + \mu_2)\right] b_i \\ + (N' - i + 1)\left[\frac{1}{2}N' + \frac{1}{2}i - 1 + n' + \mu_1\right] b_{i-1}. \end{aligned} \tag{6.28}$$

This constitutes a lower triangular matrix problem for b_i. If there is not a singularity in b_i for some i between 0 and N', then all $b_i = 0$, which is trivial. So, assuming that there exists a singularity at some $i = k$, we require the b_i to vanish. This yields the following result for the eigenvalues of this case:

$$\lambda_k = -\frac{k(k-1) + \frac{1}{2}k(N' - k) + (n' + \mu_1 + \mu_2)k}{N(N-1)}. \tag{6.29}$$

Note that this result depends on the number of committed agents, n'. Each b_i can be found explicitly by Eq. 6.28 by taking $b_k = 1$ and $b_i = 0$ for $i < k$. We then find the coefficient of $G(x, y)$ by calculating $H(x - y, y)$. Doing so gives

$$G(x, y) = \sum_{i=0}^{N'} \left[\sum_{j=i}^{N'} \binom{j}{i} (-1)^{j-i} b_j\right] x^i y^{N'-i}. \tag{6.30}$$

The value of c_i in terms of b_j is given in the bracket of Eq. 6.30. To find the dominant eigenvalue of this case, take $k = 1$ in Eq. 6.29, which yields

$$\lambda = -\frac{\frac{1}{2}N' + n' + \mu_1 + \mu_2}{N^2}. \tag{6.31}$$

Similar to Eq. 6.19, the change in probability depends on the state of the system.

6.1.4 Total Rate of Convergence

Now that we have results for each case, we put them together to obtain the convergence rate of the naming game. We will make some assumptions about the state of the system. First, we assume that initially there is symmetry in the representation of words. That is, no word initially dominates the other words in accordance to the applications given above. Second, we assume that for each word, there are individuals with only this word in their lists. The system quickly orders itself this way as long word lists are replaced by lists of length 1. This second assumption allows us to utilize Cases 2 and 3 above when determining the rate of convergence.

The rate of convergence is estimated by the smallest non-zero change given by the above cases for $\mathcal{L}_{\mathcal{I}}$. So, the rate of change of the probability distribution for a *single word*, A_k, is

$$1 - \lambda_k = O\left(\frac{\theta + n'}{N^2}\right), \tag{6.32}$$

where $\theta = N' + \mu_1 + \mu_2$, which describes the macrostate of the system. If we take this to be the total change in probability, then we have implicitly assumed that there are only two words in the system, and all others have been eliminated. So, we require that all K words are present in the system and sum the smallest change in probability given by Eq. 6.32 for each word. By symmetry, the total change in probability is K multiplied by the right hand side of 6.32. Therefore, the total rate of convergence is given by

$$1 - \lambda = O\left[\frac{K(\theta + n')}{N^2}\right]. \tag{6.33}$$

We make use of Eq. 6.33 extensively to determine the collapse time, consensus time, and the location of a phase transition over the number of zealots. We need to carefully account for the macrostate of the system when applying Eq. 6.2 due to the presence of θ. We expect the macrostate of the system to significantly affect the solution for the consensus time and phase transition.

Now we wish to find the rate of convergence to the metastable state in the presence of committed minorities. These are given by Case 1 above. The largest of these was found to be $O(1/N)$ and the next largest was $O(K/N^2)$. Since the rate of convergence is given by the sum of these cases, we find that the rate of convergence to the metastable state is

$$1 - \lambda \sim \frac{a}{N} + \frac{bK}{N^2}. \tag{6.34}$$

Here a and b are constants. When the convergence rate to the metastable state exceeds the convergence rate to consensus, the system is trapped in the metastable state. Otherwise, the system rapidly moves to consensus. This gives the criterion for the phase transition over n'.

REFERENCES

[1] P. Clifford and A. Sudbury, "A model for spatial conflict," *Biometrika*, vol. 60, no. 3, pp. 581–588, Mar. 1973.

[2] T. Liggett, *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes.* New York, NY: Springer, 1999.

[3] T. M. Liggett, *Interacting Particle Systems.* New York, NY: Springer-Verlag, 2005.

[4] P. Sen and B. K. Chakrabarti, *Sociophysics, An Introduction.* Oxford, UK: Oxford University Press, 2013.

[5] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Rev. Mod. Phys.*, vol. 81, no. 2, p. 591, May 2009.

[6] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hungar. Acad. Sci*, vol. 5, pp. 17–61, 1960.

[7] L. Steels, "A self-organizing spatial vocabulary," *Artif. Life*, vol. 2, no. 3, pp. 319–332, Apr. 1995.

[8] A. Baronchelli, M. Felici, V. Loreto, E. Caglioti, and L. Steels, "Sharp transition towards shared vocabularies in multi-agent systems," *J. Stat. Mech.: Theory E.*, vol. 2006, no. 06, p. P06014, June 2006.

[9] A. Baronchelli, V. Loreto, and L. Steels, "In-depth analysis of the naming game dynamics: the homogeneous mixing case," *Int. J. Mod. Phys. C*, vol. 19, no. 05, pp. 785–812, May 2008.

[10] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, and B. K. Szymanski, "Social consensus through the influence of committed minorities," *Phys. Rev. E*, vol. 84, no. 1, p. 011130, July 2011.

[11] W. Zhang, C. Lim, S. Sreenivasan, J. Xie, B. K. Szymanski, and G. Korniss, "Social influencing and associated random walk models: Asymptotic consensus times on the complete graph," *Chaos*, vol. 21, no. 2, p. 025115, June 2011.

[12] A. Waagen, G. Verma, K. Chan, A. Swami, and R. D'Souza, "Effect of zealotry in high-dimensional opinion dynamics models," *Phys. Rev. E*, vol. 91, no. 2, p. 022811, Feb. 2015.

[13] G. Verma, A. Swami, and K. Chan, "The impact of competing zealots on opinion dynamics," *Physica A*, vol. 395, pp. 310 – 331, Feb. 2014.

[14] W. Pickering, B. K. Szymanski, and C. Lim, "Analysis of the high-dimensional naming game with committed minorities," *Phys. Rev. E*, vol. 93, no. 5, p. 052311, May 2016.

[15] W. Pickering and C. Lim, "Solution of the voter model by spectral analysis," *Phys. Rev. E*, vol. 91, no. 1, p. 012812, Jan. 2015.

[16] P. A. P. Moran, "Random processes in genetics," *Math. Proc. Cambridge*, vol. 54, no. 1, pp. 60–71, Jan. 1958.

[17] R. A. Blythe and A. J. McKane, "Stochastic models of evolution in genetics, ecology and linguistics," *J. Stat. Mech.: Theory E.*, vol. 2007, no. 07, p. P07018, July 2007.

[18] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[19] M. D. Vose, *The Simple Genetic Algorithm, Foundations and Theory*. Cambridge, MA: MIT Press, 1999.

[20] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press, 1992.

[21] V. Sood and S. Redner, "Voter model on heterogeneous graphs," *Phys. Rev. Lett.*, vol. 94, no. 17, p. 178701, May 2005.

[22] F. Vazquez and V. M. Eguíluz, "Analytical solution of the voter model on uncorrelated networks," *New J. Phys.*, vol. 10, no. 6, p. 063011, June 2008.

[23] E. Ben-Naim, L. Frachebourg, and P. Krapivsky, "Coarsening and persistence in the voter model," *Phys. Rev. E*, vol. 53, no. 4, p. 3078, Apr. 1996.

[24] W. Pickering and C. Lim, "Solution of the multistate voter model and application to strong neutrals in the naming game," *Phys. Rev. E*, vol. 93, no. 3, p. 032318, Mar. 2016.

[25] M. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Phys. Rev. E*, vol. 64, no. 2, p. 026118, July 2001.

[26] M. Newman, "Spread of epidemic disease on networks," *Phys. Rev. E*, vol. 66, no. 1, p. 016128, July 2002.

[27] E. A. Bender and S. G. Williamson, *Foundations of Combinatorics with Applications*. Mineola, NY: Dover, 2006.

[28] I. H. Herron and M. R. Foster, *Partial Differential Equations in Fluid Dynamics*. New York, NY: Cambridge University Press, 2008.

[29] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Mineola, NY: Dover, 1964, vol. 55.

[30] C. Cooper, R. Elsasser, H. Ono, and T. Radzik, "Coalescing random walks and voting on connected graphs," *SIAM J. Discrete Math.*, vol. 27, no. 4, pp. 1748–1758, Oct. 2013.

[31] M. E. Yildiz, R. Pagliari, A. Ozdaglar, and A. Scaglione, "Voting models in random networks," in *Inform. Theory Applicat. Workshop, 2010*, Feb. 2010, pp. 1–7.

[32] M. Kac, "Random walk and the theory of brownian motion," *Amer. Math. Mon.*, vol. 54, no. 7, pp. 369–391, Sept. 1947.

[33] P. Ehrenfest and T. Ehrenfest, "Über zwei bekannte einwände gegen das boltzmannsche h-theorem," *Physik Z.*, vol. 8, pp. 311–314, 1907.

[34] E. S. Lee, "A theory of migration," *Demography*, vol. 3, no. 1, pp. 47–57, Mar. 1966.

[35] J. J. Lewer and H. Van den Berg, "A gravity model of immigration," *Econ. Lett.*, vol. 99, no. 1, pp. 164–167, Apr. 2008.

[36] J. E. Anderson, "A theoretical foundation for the gravity equation," *Amer. Econ. Rev.*, vol. 69, no. 1, pp. 106–116, Mar. 1979.

[37] A. Edelman and G. Strang, "Pascal matrices," *Amer. Math. Monthly*, vol. 111, no. 3, pp. 189–197, Mar. 2004.

[38] X. Castelló, A. Baronchelli, and V. Loreto, "Consensus and ordering in language dynamics," *Eur. Phys. J. B*, vol. 71, no. 4, pp. 557–564, Aug. 2009.

[39] W. Zhang, "Analytical approach for opinion dynamics on social networks," Ph.D. dissertation, Dept. Math. Sci., Rensselaer Polytechnic Institute, Troy, NY, 2012.

[40] C. E. Shannon, "A mathematical theory of communication," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, Jan. 2001.

[41] M. Starnini, A. Baronchelli, and R. Pastor-Satorras, "Ordering dynamics of the multi-state voter model," *J. Stat. Mech.: Theory E.*, vol. 2012, no. 10, p. P10027, Oct. 2012.

[42] J. T. Cox, "Coalescing random walks and voter model consensus times on the torus in zd," *Ann. Probab.*, vol. 17, no. 4, pp. 1333–1366, Oct. 1989.

[43] S. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering.* Boulder, CO: Westview Press, 1994, vol. 272.

[44] W. Kornblum, *Sociology in a Changing World.* Belmont, CA: Cengage Learning, 2007.

[45] J. Berger and G. Le Mens, "How adoption speed affects the abandonment of cultural tastes," *Proc. Natl. Acad. Sci.*, vol. 106, no. 20, pp. 8146–8150, Mar. 2009.

[46] D. Vilone and C. Castellano, "Solution of voter model dynamics on annealed small-world networks," *Phys. Rev. E*, vol. 69, no. 1, p. 016109, Jan. 2004.

[47] L. Frachebourg and P. Krapivsky, "Exact results for kinetics of catalytic reactions," *Phys. Rev. E*, vol. 53, no. 4, p. R3009, Apr. 1996.

[48] N. Masuda, N. Gibert, and S. Redner, "Heterogeneous voter models," *Phys. Rev. E*, vol. 82, no. 1, p. 010103, July 2010.

[49] N. Masuda, "Voter models with contrarian agents," *Phys. Rev. E*, vol. 88, no. 5, p. 052803, Nov. 2013.

[50] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications.* New York, NY: Macmillan London, 1976, vol. 290.

[51] F. John, *Partial Differential Equations.* New York, NY: Springer-Verlag, 1981.

[52] M. A. M. de Aguiar and Y. Bar-Yam, "Moran model as a dynamical process on networks and its implications for neutral speciation," *Phys. Rev. E*, vol. 84, no. 3, p. 031901, Sept. 2011.

[53] R. Arratia, "Limiting point processes for rescalings of coalescing and annihilating random walks on zd," *Ann. Probab.*, vol. 9, no. 6, pp. 909–936, Dec. 1981.

[54] J. Cardy and U. C. Täuber, "Theory of branching and annihilating random walks," *Phys. Rev. Lett.*, vol. 77, no. 23, pp. 4780–4783, Dec. 1996.

[55] T. Luczak, "Sparse random graphs with a given degree sequence," in *Proc. Symp. Random Graphs, Poznan*, 1989, pp. 165–182.

[56] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random Struct. Algor.*, vol. 6, no. 2-3, pp. 161–180, Mar. 1995.

[57] M. Molloy and B. Reed, “The size of the giant component of a random graph with a given degree sequence,” *Comb. Probab. Comp.*, vol. 7, no. 03, pp. 295–305, Sept. 1998.

[58] M. E. J. Newman, “The structure and function of complex networks,” *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Aug. 2003.

[59] S. Roman, *Advanced Linear Algebra*, 3rd ed. New York, NY: Springer-Verlag, 2008.

[60] F. Vazquez, P. L. Krapivsky, and S. Redner, “Constrained opinion dynamics: Freezing and slow evolution,” *J. Phys. A: Math. Gen.*, vol. 36, no. 3, p. L61, Jan. 2003.

[61] D. Volovik, M. Mobilia, and S. Redner, “Dynamics of strategic three-choice voting,” *Europhys. Lett.*, vol. 85, no. 4, p. 48003, Mar. 2009.

[62] X. Castelló, V. M. Eguíluz, and M. San Miguel, “Ordering dynamics with two non-excluding options: bilingualism in language competition,” *New J. Phys.*, vol. 8, no. 12, p. 308, Dec. 2006.

[63] M. De Oliveira, J. Mendes, and M. Santos, “Nonequilibrium spin models with ising universal behaviour,” *J. Phys. A: Math. Gen.*, vol. 26, no. 10, p. 2317, May 1993.

[64] J. Cox and R. Durrett, “Nonlinear voter models,” in *Random Walks, Brownian Motion, and Interacting Particle Systems*, 1991, pp. 189–201.

[65] J.-M. Drouffe and C. Godrèche, “Phase ordering and persistence in a class of stochastic processes interpolating between the ising and voter models,” *J. Phys. A: Math. Gen.*, vol. 32, no. 2, p. 249, Jan. 1999.

[66] C. Castellano, M. A. Muñoz, and R. Pastor-Satorras, “Nonlinear q-voter model,” *Phys. Rev. E*, vol. 80, no. 4, p. 041129, May 2009.

[67] F. Schweitzer and L. Behera, “Nonlinear voter models: the transition from invasion to coexistence,” *Eur. Phys. J. B*, vol. 67, no. 3, pp. 301–318, Jan. 2009.

[68] M. Mobilia, “Does a single zealot affect an infinite group of voters?” *Phys. Rev. Lett.*, vol. 91, no. 2, p. 028701, July 2003.

[69] M. Mobilia, A. Petersen, and S. Redner, “On the role of zealotry in the voter model,” *J. Stat. Mech.: Theory E.*, vol. 2007, no. 08, p. P08029, Aug. 2007.

[70] M. de Oliveira, “Isotropic majority-vote model on a square lattice,” *J. Stat. Phys.*, vol. 66, no. 1-2, pp. 273–281, July 1992.

[71] L. F. C. Pereira and F. G. B. Moreira, "Majority-vote model on random graphs," *Phys. Rev. E*, vol. 71, no. 1, p. 016123, Jan. 2005.

[72] P. R. A. Campos, V. M. de Oliveira, and F. G. B. Moreira, "Small-world effects in the majority-vote model," *Phys. Rev. E*, vol. 67, no. 2, p. 026104, Feb. 2003.

[73] D. Mistry, Q. Zhang, N. Perra, and A. Baronchelli, "Committed activists and the reshaping of status-quo social consensus," *Phys. Rev. E.*, vol. 92, no. 4, p. 042805, Oct. 2015.

[74] S. Galam, "Minority opinion spreading in random geometry," *Eur. Phys. J. B*, vol. 25, no. 4, pp. 403–406, Feb. 2002.

[75] J. Xie, J. Emenheiser, M. Kirby, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Evolution of opinions on social networks in the presence of competing committed groups," *PLoS One*, vol. 7, no. 3, p. e33215, Mar. 2012.

[76] S. Moscovici, E. Lage, and M. Naffrechoux, "Influence of a consistent minority on the responses of a majority in a color perception task," *Sociometry*, vol. 32, no. 4, pp. 365–380, Dec. 1969.

[77] S. A. Marvel, H. Hong, A. Papush, and S. H. Strogatz, "Encouraging moderation: Clues from a simple model of ideological conflict," *Phys. Rev. Lett.*, vol. 109, no. 11, p. 118702, Sept. 2012.

[78] I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, C. J. Torney, A. Hartnett, L. Conradt, S. A. Levin, and N. E. Leonard, "Uninformed individuals promote democratic consensus in animal groups," *Science*, vol. 334, no. 6062, pp. 1578–1580, Dec. 2011.

[79] M. Turalska, B. J. West, and P. Grigolini, "Role of committed minorities in times of crisis," *Sci. Rep.*, vol. 3, p. 1371, Mar. 2013.

[80] S. K. Maity and A. Mukherjee, "Emergence of dominant opinions in presence of rigid individuals," in *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*, ser. Understanding Complex Systems, A. Mehler, A. Lücking, S. Banisch, P. Blanchard, and B. Job, Eds. New York, NY: Springer Berlin Heidelberg, 2016, pp. 279–295.

[81] A. M. Thompson, B. K. Szymanski, and C. C. Lim, "Propensity and stickiness in the naming game: Tipping fractions of minorities," *Phys. Rev. E*, vol. 90, no. 4, p. 042809, Oct. 2014.

[82] Q. Lu, G. Korniss, and B. K. Szymanski, "The naming game in social networks: community formation and consensus engineering," *J. Econ. Interact. Coor.*, vol. 4, no. 2, pp. 221–235, July 2009.

[83] Q. Lu, G. Korniss, and B. K. Szymanski, "Naming games in two-dimensional and small-world-connected random geometric networks," *Phys. Rev. E*, vol. 77, no. 1, p. 016111, Jan. 2008.