OPINION FORMATION MODELS IN STATIC AND DYNAMIC SOCIAL NETWORKS

By

Pramesh Singh

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: PHYSICS

Examining Committee:

Gyorgy Korniss, Dissertation Adviser

Boleslaw Szymanski, Dissertation Adviser

Chjan Lim, Member

Saroj Nayak, Member

Rensselaer Polytechnic Institute
Troy, New York

July 2014
(For Graduation August 2014)

© Copyright 2014
by
Pramesh Singh
All Rights Reserved

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

I would like to express my thanks to my advisers Prof. Gyorgy Korniss and Prof. Boleslaw Szymanski for their guidance, help, and support. I thank my committee members Prof. Chjan Lim and Prof. Saroj Nayak for useful discussions and suggestions. I would also like to thank Dr. Sameet Sreenivasan for his suggestions and comments in the preparation of this thesis. I also express my gratitude to my family members, friends, and colleagues for their constant encouragement.

This work was supported in part by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053 and by the Office of Naval Research Grant No. N00014-09-1-0607. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies either expressed or implied of the Army Research Laboratory or the U.S. Government.

ABSTRACT

We study models of opinion formation on static as well as dynamic networks where interaction among individuals is governed by widely accepted social theories. In particular, three models of competing opinions based on distinct interaction mechanisms are studied. A common feature in all of these models is the existence of a *tipping point* in terms of a model parameter beyond which a rapid consensus is reached.

In the first model that we study on a static network, a node adopts a particular state (opinion) if a threshold fraction of its neighbors are already in that state. We introduce a few initiator nodes which are in state '1' in a population where every node is in state '0'. Thus, opinion '1' spreads through the population until no further influence is possible. Size of the spread is greatly affected by how these initiator nodes are selected. We find that there exists a critical fraction of initiators p_c that is needed to trigger global cascades for a given threshold ϕ. We also study heuristic strategies for selecting a set of initiator nodes in order to maximize the cascade size. The structural properties of networks also play an important role in the spreading process. We study how the dynamics is affected by changing the clustering in a network. It turns out that local clustering is helpful in spreading.

Next, we studied a model where the network is dynamic and interactions are *homophilic*. We find that homophily-driven rewiring impedes the reaching of consensus and in the absence of committed nodes (nodes that are not influenceable on their opinion), consensus time T_c diverges exponentially with network size N. As we introduce a fraction of committed nodes, beyond a critical value, the scaling of T_c becomes logarithmic in N. We also find that slight change in the interaction rule can produce strikingly different scaling behaviors of T_c. However, introducing committed agents in the system drastically improves the scaling of the consensus time regardless of the interaction rules considered.

Finally, a three-state (*leftist, rightist, centrist*) model that couples the dynamics of social balance with an external deradicalizing field is studied. The mean-field

analysis shows that for a weak external field, the system exhibits a metastable fixed point and a saddle point in addition to a stable fixed point. However, if the strength of the external field is sufficiently large (larger than a critical value), there is only one (stable) fixed point which corresponds to an all-centrist consensus state (absorbing state). In the weak-field regime, the convergence time to the absorbing state is evaluated using the quasi-stationary(QS) distribution and is found to be in good agreement with the results obtained by numerical simulations.

CHAPTER 1
Introduction

Social systems, despite lying outside the domain of traditional physics, have been of great interest to statistical physicists. The framework offered by statistical physics and network science allows us to investigate the dynamics of social systems using simplified yet insightful models. For example, in the case of binary opinions, one can envision the Ising model of spin systems as a model for opinion evolution in social systems. An individual's opinion, is thus represented an individual spin state and the ordered state of opinions (*consensus*) can be modeled as ferromagnetic ordering in the Ising model [1]. Systems composed of a large number of socially interacting constituents, exhibit macroscopic properties which can be understood in terms of local microscopic interactions, similar to systems studied in traditional statistical mechanics subject to physical interactions. However, the driving forces of social dynamics are quite different in nature from the forces governing the dynamics of interacting particles in physical systems. Social dynamics is usually driven by forces such as *influence* [2], *homophily* [3], *reciprocity* [4], *structural balance* [5] etc. Specifically, study of evolution of opinions (*opinion dynamics*) is also governed by the same set of driving forces. Typically, attainment of an ordered state ('consensus' in the context of opinions) and question of how the system evolves from an initial disordered (multiple competing opinions) are of particular interest in such interacting systems.

One of the simplest models of opinion dynamics is the *voter model* [6, 7]. According to this model, every node in the system can be in one of the two possible states and at each time step a random node is selected which copies the state of one of its random neighbors. The only absorbing states (consensus) of the system under this model are when all nodes adopt one state or the other. Once every node adopts the same state, the dynamics freezes. For a finite system, consensus is always reached in this model. On regular topologies, the consensus is reached by coarsening if the dimensionality of the system $d \leq 2$, while for $d > 2$ consensus is reached only

by fluctuation [2].

The voter model has been studied with a number of modification and on various network topologies [2] as well as on adaptive networks [8]. In the case of regular networks, the order in which a node and its neighbors is selected, does not make any difference. However in networks with heterogeneous degree distributions, the order in which they are selected becomes important due to the fact that high degree nodes are more likely to be the neighbors of the initially chosen node. In this work, we will discuss the effect of selection order for a slightly different model in Chapter 3.

A Naming Game (NG) model studied by Baronchelli *et al.* [9] describes the agreement dynamics in a system where each node possesses a list words associated to objects. The model in a somewhat simplified and crude fashion, describes the formation of languages. Upon interaction, a user (chosen as speaker) speaks a word for a particular object from its list of words (vocabulary) for that object to the listener. If the speaker's list happens to be empty at that time then it invents a new word for the object. The listener adds to its vocabulary the spoken word if it does not have the word. If the listener already has the same word in its list of words for that object then they both (speaker and the listener) keep that word for the object while deleting all other words for that object. The authors, starting from a completely empty vocabulary for each node, show the existence of a disorder/order transition where a consensus state is reached eventually.

A simpler case of the above model where number of possible words are restricted to two (A or B), has been applied to study systems of two competing opinions (*binary agreement model*) [10,11]. Under this simplification, each node can either be in one of the two opposite opinion states (A or B) or in an intermediate state (AB). Thus, for this model, a node changes its from A to B (or B to A) by passing through an intermediate state AB. Hence, as opposed to the voter model, in the binary agreement model individuals have some inertia to change their state. Introduction of this intermediate state affects the dynamics significantly. For the complete graph (where every node is connected with every other node), the time to reach consensus scales logarithmically with system size [12] whereas it scales linearly

with system size in the voter model dynamics [2].

The focus of this work is to describe how opinions evolve in a social network (a network of individuals connected by links representing interpersonal relations) where microscopic interactions among nodes are prescribed by a set of dynamical rules consistent with widely accepted fundamental social theories. In particular, a few interaction mechanisms (thought to drive the social dynamics) are considered while modeling the opinion dynamics. Typically each node in the network is assigned an 'opinion state' (which is usually discrete but can also be continuous as in [13]) and starting from an arbitrary opinion distribution the system's evolution is studied. Although the actual dynamical process may vary from one model to another, which we will discuss in the following chapters, the models studied in this thesis show the existence of a *tipping point* – a critical value of a model variable – such that the system displays strikingly different behavior above and below the *tipping point*.

The first model that we discuss (Chapter 2) is the *threshold model* [14–17] in which a node turns active if a sufficient number (greater than the adoption threshold of the node) of its neighbors are already active. It is one of the simplest models that captures the dynamics of adoption of behavior and to some extent it is similar to Susceptible-Infected-Susceptible or contact-process like models [18–27]. The model is inherently asymmetric as when two nodes interact only the inactive node becomes active (adopts the new behavior) thus resulting in a spread like dynamics. Although studied mainly in the context of adoption of new behavior or information flow [16,28], the model can well be applied to the dynamics of opinion change [29]. For a single initiator (i.e. initially active node), or vanishingly small set of initiators, previous studies [16, 29, 30] found that if the adoption threshold of the nodes (which is same for all nodes) large enough, spread is only minuscule. In this study we find that large spread (*global cascades*) can be achieved if the number of initiators is greater than a critical fraction which we identify as the tipping point for multiple initiators. By introducing a fraction of initiators larger than the critical fraction global cascades can be triggered in the system.

An important application oriented goal of the study of spreading processes is to attain a maximum spread size (or minimum depending upon the problem). In

this context not only the fraction but also the particular choice of these initiators plays an important role in the spreading process. In Ref. [31, 32] authors studied greedy and scalable maximization algorithms for selecting initiators. However in these studies the initiator fraction is small [$O(1)$] and does not explore the regime where a critical initiator fraction is required to trigger cascades. In contrast, our approach to selection strategy for initiators is heuristic. Three strategies for selecting initiator nodes are compared and we find that selecting initiators by their degree (highest first) works best.

Structural properties of the network significantly affect the dynamics occurring on it [11, 30, 33, 34]. The role of network topology, clustering and community structure in particular, for the threshold dynamics for a small initiator fraction is studied in [30, 33]. They found that local-clustering (high density of triads) in the network facilitates spreading. In Ref. [35], the threshold model dynamics has recently been studied on multiplex networks (multiple interdependent networks [36–38]). The authors discover that due to the coupling among different networks, it is possible to facilitate spreading and trigger cascades in the network layers which may otherwise (in the absence of multiplexity) not be susceptible to global cascades. In this thesis, we study the threshold dynamics on an empirical high-school network [39] and a few randomized versions of this network. The results imply (consistent with earlier findings [30, 33]) that clustering facilitates spreading even in the case of multiple [$O(N)$] initiators.

The next opinion formation model described in this thesis (Chapter 3) is driven by homophily [3, 40]. The underlying idea of homophily is that individuals tend to form social relations with those who are similar to them. Axelrod [41] studied a model of cultural evolution in which, an individual's state is represented by a set of cultural features or attributes and each of these attributes can take a value from a set of available traits for that attribute. According to this model, when two nodes interact, with probability proportional to their similarity, one of them adopts a cultural trait from the other and thus increasing their mutual similarity. Interactions among the nodes lead to formation of multiple cultural regions. The number of stable cultural regions strongly depends on initial conditions, for example the number of

attributes and the number of possible traits per attribute. This model addresses the question of why cultural differences persist despite the fact that interaction among individuals tends to make them more similar.

The model studied in Chapter 3 is similar in spirit to the Axelrod model. However, in our model the interaction network is considered to be dynamic i.e. the links evolve in time. Thus, in this model the network topology coevolves with opinion updates. In addition to the models described in Refs. [42–47], previously studied models of coevolution of structure with states of the nodes include a voter model on adaptive network with a tunable rewiring probability by Nardini *et al.* [8]. In this model, competing effects of influence and rewiring result in an exponentially long convergence time. In [48, 49], authors study the dynamics a voter in a two-party system and observe the existence of metastable states in addition to consensus state. An Axelrod-like [41] model with rewiring was studied by Vazquez *et al.* [50] (A similar model was also studied in Ref. [51]). This is very closely related to our model which also happens to be a variant of the Axelrod model. The main feature that distinguishes our model from the rest is the introduction of *committed nodes*. Committed nodes (or agents) are rigid in their opinion and can not be influenced to change it [52–62]. Commitment is an important aspect of *minority influence* where a consistent minority population influences the majority to adopt a new idea or behavior [63]. Real world examples of minority influence in history include the American Civil Right Movement and the Suffragette Movement. The effect of committed nodes have been studied extensively [1, 10, 64–66] on static networks. These studies find the existence of a critical committed fraction (tipping point) such that the system is characterized by metastability when the committed fraction is below a critical value.

Our model studies the effect of homophily, coevolution, and the role of committed nodes on opinion evolution simultaneously. In our model, just like in the Axelrod model, the state of a node is described by a set (or a vector) of attributes (which can take values from a set of traits) and the opinion of an individual is also plays the role of one of its attributes. This allows us to compute the similarity between two nodes. Nodes can influence other nodes if they are compatible (satisfy a similarity

threshold) and in turn become more similar or rewire randomly if they happen to be incompatible. Committed nodes can influence and get influenced by other nodes on any attribute except their opinion on which they stay uninfluenceable.

The time taken to reach consensus in this model diverges exponentially with system size due to the competition between influence and rewiring. A similar divergent behavior is also observed in [8]. However, this divergence in the consensus time disappears and the scaling of the consensus time becomes logarithmic with system size when a fraction of committed nodes above a critical value are introduced [67]. As found in [8], the selection order of influencer and adopter nodes in dynamical rules of the model is important in this case and with slight change in the rules a different scaling behavior of consensus time is observed. But introduction of committed nodes drastically improves the consensus time regardless of the selection order considered.

Another key driving force of social dynamics is structural balance (or social balance) [5]. The basic conditions of the theory of social balance are: (i) a friend of my friend and an enemy of my enemy is my friend, and (ii) a friend of my enemy and an enemy of my friend is my enemy. In order to be labeled 'balanced' a triad of nodes must satisfy these conditions [68]. In other words, social balance describes the tendency of individuals to like what their friends like (and enemies dislike) and dislike what their friends dislike (and enemies like) [69]. In a network where links can be classified as friendly or unfriendly, a triangle containing an odd number of negative links (one or three) is, by definition, unbalanced. According to the balance theory, these unbalanced triangles have a tendency to attain a balanced configuration. The balancing dynamics in a social network drives it to a balanced state by updating the states of the unbalanced triangles. The theory of social balance is particularly useful in study of international relations. Antal *et al.* in [70] illustrated how the evolution of the European alliance in the late 19th century and early 20th century was motivated by structural balance. A study [71] validates the idea of social balance by monitoring the links connecting the nodes in a network of players in a multiplayer online game. Imbalance in social networks can also be thought of in terms of (social) energy where unbalanced triads occupy higher energy states than the balanced ones.

The role of the balancing dynamics, thus, is to drive the system to a minimum energy state (analogous to minimization of energy). Certain (balanced) configurations of triads are energetically favored. Marvel *et al.* studied the energy landscape of social balance in the context of fully-connected networks containing negative as well as positive links and demonstrated the existence of local minima (*jammed states*) of varying levels in the landscape.

These studies focus on the states of the links (positive or negative) and treat them as independent of the states of the nodes. In the model that we study (Chapter 4), the states of the links are coupled to the states of the nodes i.e. the nature of a link depends on the opinion-states of the nodes at its ends. Specifically, we consider a three-state (*leftist*, *rightist*, *centrist*) model in which a link between two extreme opinions of opposite kind (between leftist and rightist) is negative. Three-state opinion models have been analyzed in the context of influence [10, 72, 73]. However, the main goal of our model is to study the effect of an external moderating field, which persuades the nodes holding radical viewpoint (extreme opinion) to adopt a centrist view, with some predefined probability (equivalently the strength of the field), on the concurrent balancing dynamics in the network in which nodes update their states to balance the triads that they are part of.

In contrast to the models considered in Chapters 2 and 3, this model takes into account triadic interactions. The system, under this dynamics shows a consensus state and a metastable fixed point separated by a saddle point [10]. Using the mean-field analysis, it is shown that the metastability in the system is removed if the strength of the field is greater than a critical value. In this strong field regime, the system only possesses only one fixed point–the consensus state where every node adopts the centrist opinion.

CHAPTER 2
Threshold-limited spreading with multiple initiators

Adoption of a new behavior by a node is strongly affected by its neighborhood. It has long been known through empirical studies that in a population of socially interacting individuals where each individual node holds an opinion from a binary set, a small fraction of *initiators* holding opinion opposite to the one held by the majority can trigger large cascades and eventually result in a dominant majority holding the initiators' opinion. Some recent studies have investigated such phenomena in the context of the adoption of scientific, cultural, and commercial products [74, 75]. A study (Ref. [15]) was performed on groups of participants where each participant had to choose between two alternatives to predict which alternative the majority of the participating population chooses. Each individual in this population was assigned a fixed number (four) of connections (neighbors). After first round of making a choice, the participants were allowed to discuss which of the alternatives their neighbors chose through exchange of emails and given a chance to revise their initial opinion. The results this empirical study showed that individuals (who initially hold one of the two possible opinions) are very likely to change their opinion to the opposite opinion if the number of their interacting neighbors holding the opposite opinions is greater than a certain number (two out of four neighbors).

One of the simplest models that captures adoption dynamics, irrespective of context, is the threshold model [14, 16, 28, 30]. According to the threshold model, an individual changes its opinion only if a critical fraction of its neighbors have already adopted the new opinion. This required fraction of new adoptees in the neighborhood is designated the *adoption threshold* [14, 15]. Here, we denote the adoption threshold by ϕ. Since its introduction [14], the threshold model has been studied extensively on complex networks to analyze the conditions under which a vanishingly small fraction (of the total system size) of initiators is capable of triggering a

Portions of this chapter previously appeared as: P. Singh, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Threshold-limited spreading in social networks with multiple initiators," Sci. Rep. **3,** 2330 (2013) [29].

cascade of opinion change [16, 17, 30]. In particular, these studies considered initial conditions with a single "active" node [16] or an active connected clique (a single node and all of its neighbors) [30] as initiators. In this scenario, the condition for global cascades in connected sparse random networks is $\phi < 1/\langle k \rangle$ [16,17,30], where $\langle k \rangle$ is the average degree of the network. However, with a few exceptions [31,32,76], little attention has been paid to the question of how the size and the selection of this initiator fraction affects the spreading of an opinion in the network, in particular, in the regime where a single active node or a small clique is insufficient to trigger global cascades. An optimal selection of multiple initiators which maximizes the spread remains an open question. However, in this study, three heuristic strategies to select a set of initiators are investigated on Erdős-Rényi (ER) random networks [77] and an empirical high-school friendship network [39]. Previous studies of the threshold model [16, 17] found that in the case of single initiator, when the average degree is too low or too high, large cascades are not triggered. However, within an intermediate range of $\langle k \rangle$, large cascades are realized. This range is referred to as the *cascade window* [28]. A similar cascade window is observed in case of multiple initiators. The cascade window was found to be significantly wider when the initiators were selected by their degree (highest first). It was also found that the total time taken for the cascade to terminate is shortest for this selection strategy.

Furthermore, we discuss, using empirical and stylized models of social networks, how the size of the initiator fraction affects the cascade size in the final state. For a fixed ϕ, effect of systematically varying initiator p on cascade size is studied. For $p > p_c$ (tipping point), cascades are triggered for any given ϕ even in the regime ($\phi > 1/\langle k \rangle$) where cascades cannot be triggered by a single-node of a small initiator set. The dependence of p_c on ϕ turns out to be a smooth curve separating the two phases, one in which cascades are observed and the other where cascades cannot be triggered. An exact analytical formulation of p_c for any given ϕ can be difficult. However, in Ref. [76], assuming locally tree-like structures, the authors developed an asymptotic approach to approximate the size of the cascades. This method is expected to work better for random graphs with small average degree (with negligible presence of triads) and to gradually break down for graphs with higher $\langle k \rangle$. In

our study we make use of this approximation method to estimate the tipping point p_c.

Structural properties besides $\langle k \rangle$, such as community structure and presence of triads within the network, also play an important role in the spreading process [78] and can affect the dynamics significantly in social [11, 30, 33] as well as infrastructure networks [34]. To understand the effect of clustering, the threshold dynamics is studied on an empirical network (high clustering) and its randomized versions (minimal clustering). The results show that local clustering in the network, similar to the case of a single node (or single-clique) initiator [30, 33], facilitates opinion spread to a larger extent than a homogeneous random network.

2.1 The model

In the threshold model, every node in the network can be in one of the two possible states, 0 (inactive) or 1 (active), that can be also be thought of as signifying distinct binary opinions on an issue. The typical initial condition for studying threshold model dynamics is one where all nodes except - the initiators - are in state 0. Then, the dynamics proceeds as follows. At each time step, a node is selected at random. If the node is inactive, it becomes active if at least a threshold fraction ϕ of its neighboring nodes are active i.e. in state 1. The dynamics of the threshold model is asymmetric in the sense that the active state is assumed to be permanent i.e. once a node becomes active it remains active indefinitely. The evolution of the system continues until a (steady) state is reached where no further activations can occur. The individual adoption threshold ϕ, in general, can be different for every node but for simplicity, only the case where every node has the same threshold is considered. The size of the cascade at any point during its evolution or after it has terminated, is quantified by the fraction of active nodes in the network. In the following sections we discuss the simulation of this dynamics for various network topologies.

2.2 Selection strategies

The activation of an initially inactive (state 0) node depends on the fraction of its active neighbors and the adoption threshold ϕ (which we consider to be the same for all nodes). The node turns active if it satisfies the *threshold condition* i. e. the fraction of its active neighborhood exceeds ϕ. As a result of this *threshold condition* a node's degree plays an important role in determining how easily it can be influenced. The threshold condition is more easily satisfied for a low-degree node than a high-degree node, since the former requires fewer active nodes to be present than the latter, given a fixed adoption threshold ϕ for all nodes. As discussed before, for a small fraction of initiators there is a critical value of $\phi_c = 1/\langle k \rangle$ such that global cascades are possible only if $\phi < \phi_c$. For a fixed number of initiators, high degree nodes are less likely to get influenced because it is more difficult for them to meet the threshold condition (high degree nodes require more active neighbors to become active). A high $\langle k \rangle$ is therefore not a desirable condition for cascades and for a sufficiently high $\langle k \rangle$, it becomes impossible to trigger cascades. On the other hand, for low $\langle k \rangle$, the network consists of disconnected clusters of sizes less than $O(N)$, and cascades remain confined to one or few of these clusters. As a result, global cascades only become possible in an intermediate range of $\langle k \rangle$ - the *cascade window*. In general, cascade window sizes depend on both, the threshold ϕ, and the initiator fraction p.

The precise choice of initiators also plays an important role in the size of the cascade and consequently the cascade window itself. A strategic selection of initiators can dramatically increase the average size of the spread, which is denoted by S. Here, we compare three heuristic strategies for selecting a set of initiators constituting a fraction p of the total network size: (*i*) random selection, (*ii*) selecting nodes in the descending order of their degrees, and (*iii*) selection in the descending order of k-shell index [79]. In (*ii*) and (*iii*), the choice of initiators may not be unique. If there are many sets of initiators that can be selected for the same degree (or k-shell), one of these sets is selected at random.

The simulation results are shown in Fig. 2.1(a) for a fixed fraction of initiators $p = 0.01$ on an ER graph with $N = 1000$ and $\phi = 0.18$. We first look at the average

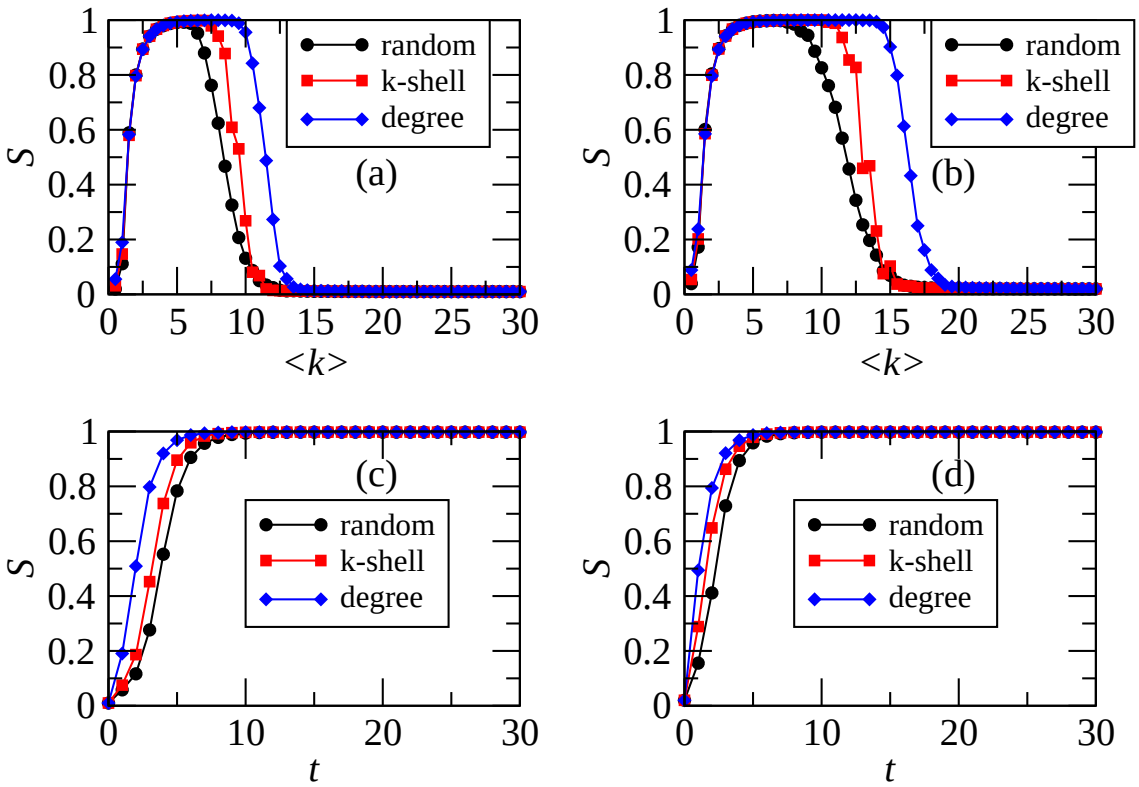

Figure 2.1: Cascade size S as a function of the average degree on ER networks of $N = 1000$ nodes with threshold $\phi = 0.18$ for different selection strategies of multiple initiators for (a) $p = 0.01$; for (b) $p = 0.02$. Time evolution of the average cascade size S on ER networks of $N = 1000$ nodes with average degree $\langle k \rangle = 6.0$ and threshold $\phi = 0.18$ for different selection strategies of multiple initiators for (c) $p = 0.01$; for (d) $p = 0.02$.

spread size as a function of average degree $\langle k \rangle$ on an ER random graph as shown in Figure 2.1(a). When $\langle k \rangle$ is small, all three strategies perform equally well because the network consists only of small clusters without a giant component and hence spread is localized to those clusters. As soon as $\langle k \rangle$ becomes large enough for a giant component to arise, the spread covers a large portion of the network. Further increasing $\langle k \rangle$ makes it harder for the nodes to satisfy the threshold condition and S decreases again.

To understand the differences in the performance of these heuristics, we first note that there are two distinct aspects determining the efficacy of a node as an initiator. First, it must be capable of influencing a large number of nodes, i.e. it

should have a large degree. Second, it must be connected to nodes which have an easily satisfiable threshold condition i.e. the degrees of its neighbors must be sufficiently low. Additionally, and related to the first point, it also makes sense to choose the highest-degree nodes as initiators, since they are the hardest to influence. In light of these arguments, the highest-degree selection strategy appears to be a natural choice for generating large cascades. It would appear that high k-shell nodes are a comparably good choice, since high k-shell nodes also possess a high degree. However, by construction, nodes in the highest k-shells are a special subset of the high-degree nodes that are predominantly connected to other nodes of high-degree. In other words, nodes selected in descending order of their k-shell index have fewer easily influenceable neighbors than nodes selected purely on the basis of degree. This qualitatively explains why the k-shell method does not perform as well as the high-degree selection. Finally, the random selection works the poorest since it largely selects low-degree nodes which trigger a small number of cascades many of which frequently terminate when they encounter a high-degree node.

An increase in the initiator fraction p makes the cascade window wider by allowing cascades to occur for even higher $\langle k \rangle$ values as shown in Fig. 2.1(b) where p is increased to 0.02. The selection strategies follow the same ranking in this case as well.

Results obtained from simulations indicate that highest degree method also works better (followed by the k-shell method) in terms of the speed of the cascade. The results for $p = 0.01$ and $p = 0.02$ are shown in Figs. 2.1(c) and (d), respectively.

2.3 Tipping point for multiple initiators

As discussed in the previous section, for a small ($O(1)$-size) seed of initiators, cascades can only occur if ϕ is smaller than a critical value ($\phi < 1/\langle k \rangle$ for sparse random graphs [16,17,30]). However, this does not hold if we introduce a sufficiently large fraction of initiators in the system. It is clear from the results of the previous section that the range of $\langle k \rangle$ up to which cascades can be triggered, can be increased by increasing p. This also implies that the critical threshold can be increased by increasing the initiator fraction p thereby allowing cascades in the previously for-

bidden regime. We look at the quantity S (average fraction of nodes in state 1) as a function of p. (We will refer to S as cascade size for short.) Gradually increasing p shows that in the beginning when $p \ll 1$, (global) cascades are not observed. When p reaches a critical value p_c, a discontinuous transition occurs and large cascades are seen immediately as shown in Fig. 2.2(a). The need for a minimum critical fraction of committed nodes for consensus has been observed in different models of influence [10, 66, 67].

Since starting with a finite p itself accounts for a large number of nodes in state 1, the relevant quantity to look at is the number of nodes that were initially in state 0 and eventually adopted state 1 (i.e., excluding the initiators). Thus, we define

$$\tilde{S} = \frac{S - p}{1 - p} \,, \tag{2.1}$$

which measures the fraction of non-initiator nodes that participate in the cascade. Transitions in $\tilde{S}$ are shown in Fig. 2.2(b) for different ϕ values and several network sizes. It can clearly be seen that the transition only depends upon ϕ and is independent of system size N. This transition (the emergence of the tipping point) is quite generic in the threshold model, and can be observed in networks with different sizes and average degrees, as well as for different selection methods for initiators (Fig. 2.3).

The critical point p_c in each case is calculated by numerically computing the derivative of $\tilde{S}$ with respect to p and finding its maximum. Having calculated p_c allows us to explicitly look at the relationship between p_c and ϕ as shown in Fig. 2.4(a) for different average degrees $\langle k \rangle$. As $\langle k \rangle$ increases, all curves appear to converge to the limiting case of the fully-connected network (complete graph) for which $p_c = \phi$. Therefore, for a given threshold ϕ the minimum number of initiators needed to trigger large cascades can be estimated.

A previously developed asymptotic method [76] is employed to estimate $p_c(\phi)$ semi-analytically (see Appendix A.1). This method uses a tree-approximation for the network structure and calculates the cascade size by assuming a progressive, directed activation of nodes from the surface of the tree to the root. Consequently, the prescribed approximation method works well only for low $\langle k \rangle$ and low p. For

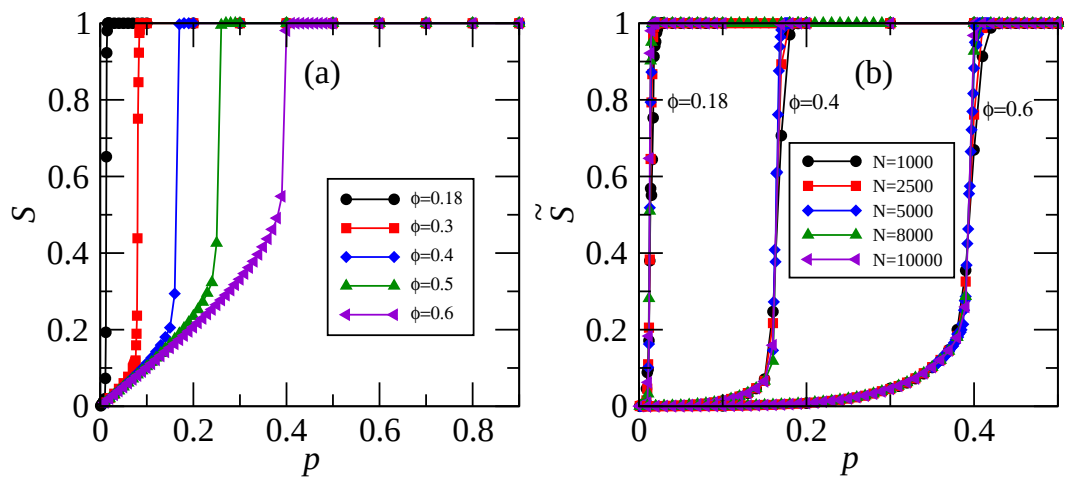

Figure 2.2: Cascade size and scaled cascade size as a function of initiators on ER networks with $\langle k \rangle = 10.0$. (a) Cascade size S as a function of initiators p for ER networks with $N = 10000$ for different values of ϕ. (b) Scaled cascade size $\tilde{S}$ [Eq. (2.1)] vs. p for ER networks with different network sizes N and ϕ values.

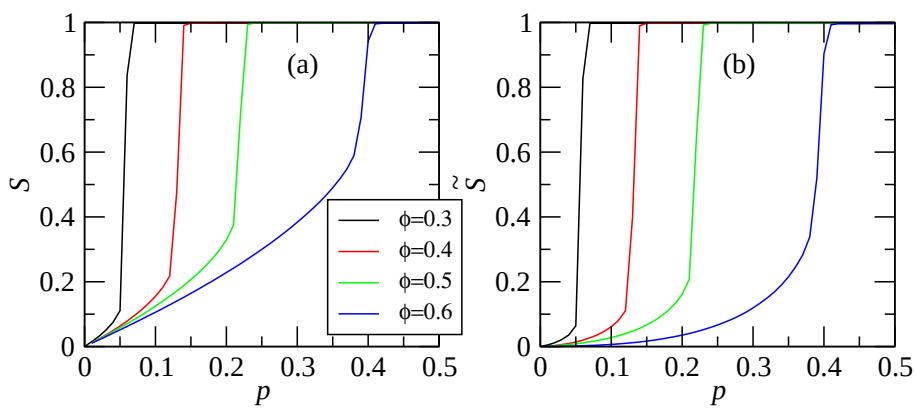

Figure 2.3: (a) Cascade size S as a function of p for ER networks with $N = 5000$ and $\langle k \rangle = 6.0$. (b) Scaled cascade size $\tilde{S}$ for the same set of parameters

large $\langle k \rangle$, the tree-approximation breaks down, while for large p, deviations from the assumed progressive and directed activation of levels, become significant. The comparison of the analytically predicted p_c using this method to values obtained from simulations clearly show regions of approximation validity and breakdown [Fig. 2.4(a)].

For a fixed $\langle k \rangle = 10$ and $N = 5000$, how the selection of initiators (selecting

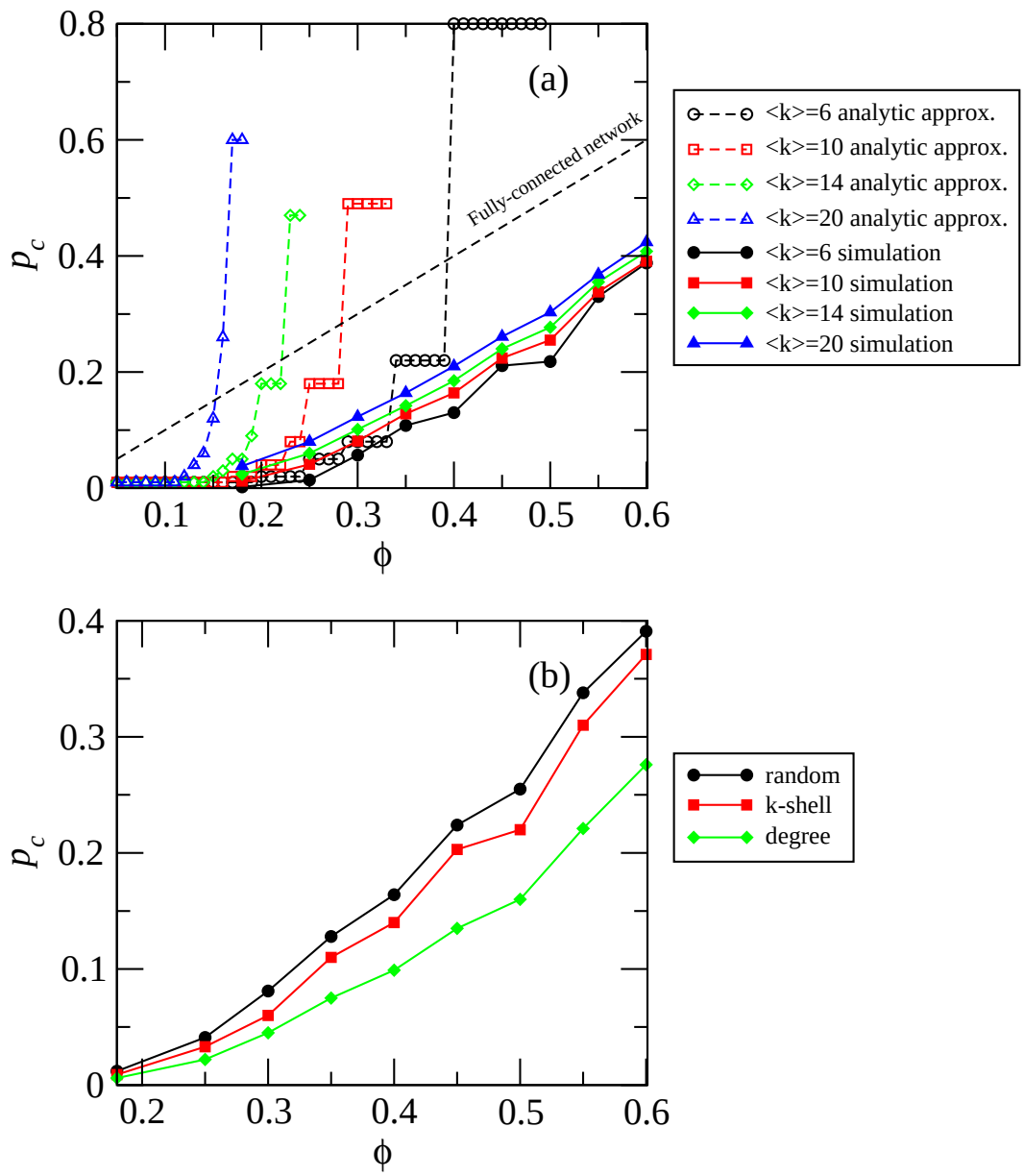

Figure 2.4: (a) Critical fraction of initiators (obtained by simulation and analytic approximation) for global cascades p_c as a function of the local threshold-value ϕ for ER networks of size $N = 5000$ with various values of the average degree. The dashed line corresponds to the exact limiting case on large complete graphs (fully-connected networks), $p_c \simeq \phi$. (b) Critical fraction of initiators for three different selection strategies for ER networks of $\langle k \rangle = 10$ and $N = 5000$.

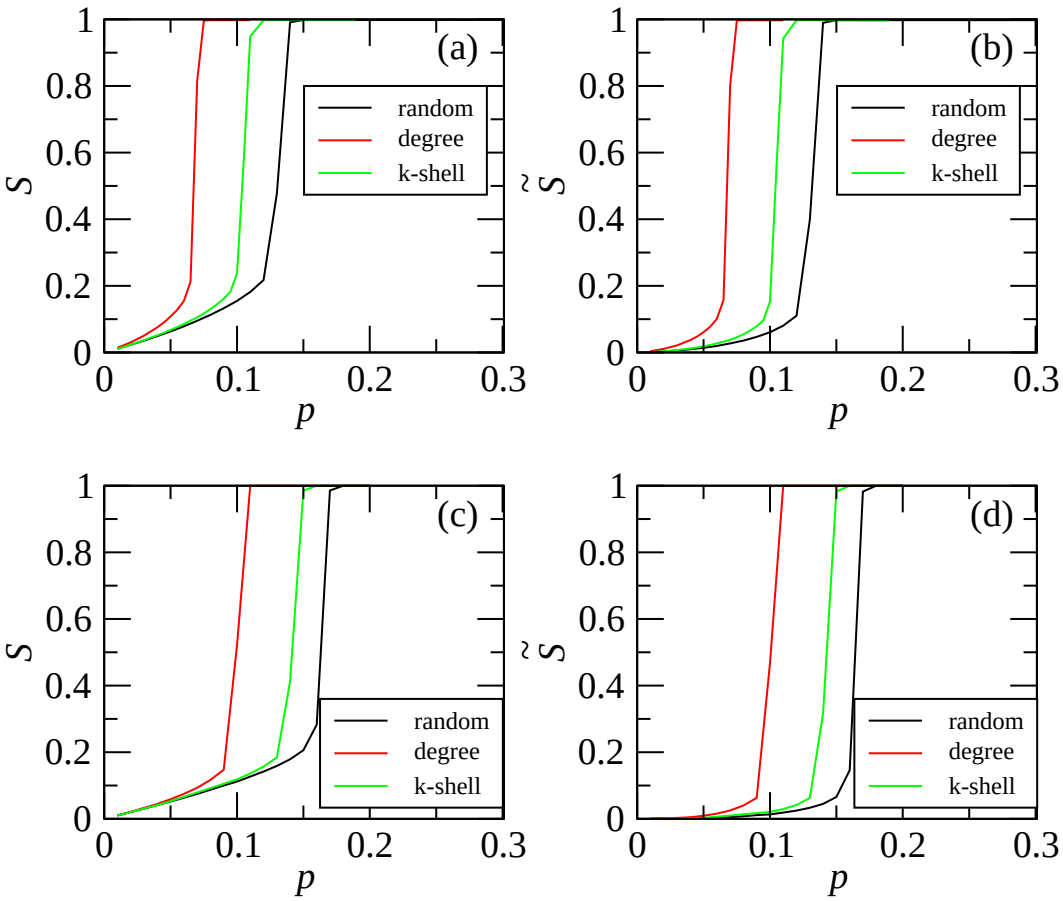

Figure 2.5: Three strategies of selecting the initiator nodes (a) S as a function of p (b) Scaled $\tilde{S}$ as a function of p, for ER networks with $\langle k \rangle = 6.0$. (c) S as a function of p (d) Scaled $\tilde{S}$ as a function of p, for ER networks with $\langle k \rangle = 10.0$. All results are for $\phi = 0.4$ and network size $N = 5000$.

initiators by their degrees, k-shell scores, and randomly) affect the critical fraction p_c, was also studied by simulation. Simulation results in Fig. 2.4(b) show that selection of initiators by their degree works better than the other two methods across the range of threshold ϕ. For a particular choice of $\phi = 0.4$, these transition are shown in Fig. 2.5. It can clearly be seen that p_c is the smallest (transition happens for a smaller p) for the highest degree method. This is consistent with our finding that selection by degree works best.

2.4 Impact of network structure and clustering

Structural changes in the network have a significant impact on the dynamics of the the threshold model. Centola *et al.* [30] studied the threshold dynamics with a small single clique of initiators and demonstrated that in this case where nodes need simultaneous exposure to multiple active sources in order to become active, short distance links are indeed helpful in triggering cascades. Hence they found that clustering facilitates spreading process occurring on the network. In Ref. [80], an experimental study was conducted on a population embedded in an online network with various network topologies. In this experiment, each individual has the same number of pre-assigned neighbors and can receive recommendations from their neighbors to sign up for a health behavior. The study was aimed to study the effect of clustering on the adoption behaviors of the individuals. The study found that in a network with higher clustering the spread of behavior was greater in extent compared to random networks.

In this section we study how the dynamics of the threshold model is affected by structural changes in the network. We study the dynamics on an empirical high-school friendship network, using one particular network from the Add Health data set [39] (also employed in [11]) and a few degree-sequence preserving randomized versions of it. For simplicity, only the giant component of the high-school network ($N = 921$ and $\langle k \rangle \approx 5.96$) is considered and refer to it as the high-school network hereafter. The network contains two large communities which are roughly equal in size. We generate two distinct ensembles of networks from this high-school network by employing the following randomization methods:

1. The link swap method (henceforth referred to as *x-swap*) in which two links are selected at random and then one end point of a link is swapped with the end point of the other link. An x-swap step is disallowed if it results in fragmentation of the network. This swapping is done repeatedly so that the network is randomized to an extent that any community structure, local clustering, or degree-degree correlation is eliminated [81–83].

2. The exact sampling method by Del Genio et al. (DKTB) [84], a connected network is constructed from the degree sequence of the original network. The

Table 2.1: Average clustering coefficient of the high-school network before and after randomization for a single run. For the x-swapped network, C was measured after $t = 6000$ time steps (when it stabilizes).

Network	C (before randomization)	C (after randomization)
X-swapped	0.125441	0.007648
Exact sampling	0.125441	0.008533

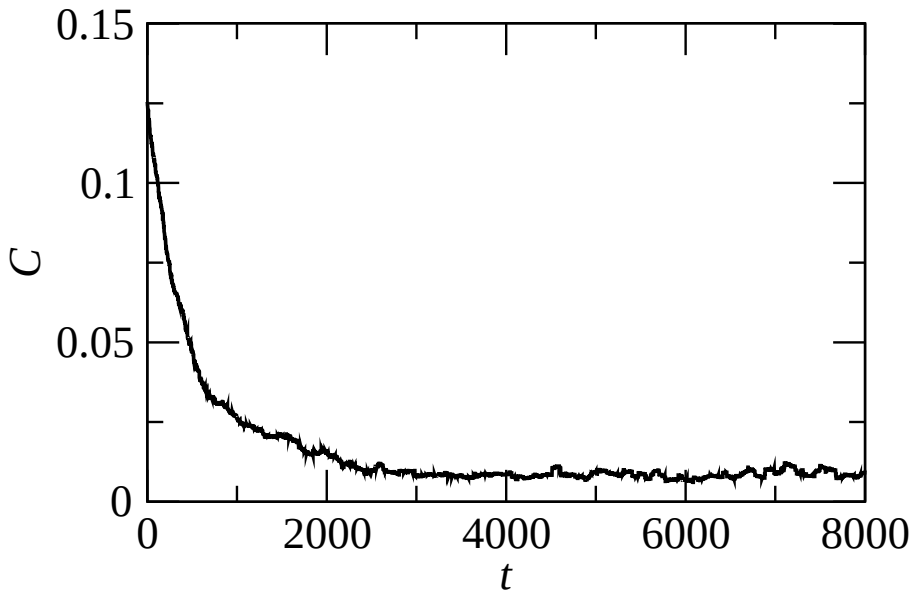

Figure 2.6: Time evolution of the average clustering coefficient of the high-school network during X-swap randomization.

algorithm takes as input the exact degree sequence of the network and joins the link stubs from different nodes until every stub has been paired with another stub [84, 85].

Both methods of randomization leave the original degree sequence unchanged. However during this process, local clustering present in the network is destroyed (as shown in Table 2.1). Fig. 2.6 shows, starting from the original high-school network, how the average clustering coefficient C (see Appendix A.2) changes at each time step during X-swap randomization.

Results for x-swapped and exact sampling [84] are very similar and therefore only one of them (the former) is shown in detail. We look at the size of spread S as a function of time for $p = 0.01$ in the original high-school network Fig. 2.7(a) and

the x-swapped high-school network Fig. 2.7(b), while Fig. 2.7(c) shows the direct comparison between the corresponding ensemble-averaged time series. Analogous plots for a larger initiator fraction $p = 0.02$ are shown in Figs. 2.7(d–f). For the empirical high-school network, some runs reveal the existence of community structure in the network where spread is faster in one community compared to other. More specifically, in some of these runs, the cascade first sweeps one of the communities (while the other one resists) before it becomes global. This can be seen by the step-like evolution in the corresponding time series in Fig. 2.7(a) [randomized networks do not exhibit this behavior, see Figs. 2.7(b)]. The same phenomena can also be observed in the configuration snapshots in Fig. 2.8(a), while their randomized counterparts do not show this behavior [Fig. 2.8(b,c)].

In general, the results show that triggered cascades are larger and more likely for a network with high local clustering than for a randomized network with the same degree sequence [Fig. 2.7], although the impact of clustering is diminishing for larger values of p. Note that the clustering coefficient of the original high-school (HS) graph is $C_{HS} \approx 0.125$; for its randomized versions obtained by x-swaps (XS) and exact-degree sequence (DKTB [84]) construction are $C_{XS} \approx C_{DKTB} \approx 0.008$.

The average cascade size S [Fig. 2.9(a) and (b)] and the probability of global cascades P_c [Fig. 2.9(c) and (d)] as a function of threshold ϕ also indicate that strong clustering (present in empirical networks) facilitates threshold-limited spreading. (We define a global cascade as a cascade that covers at least 60% percent of the network size N.) Hence, this important feature of threshold-limited spreading [30,33] is preserved for the case of multiple initiators studied here.

The temporal evolution of the average cascade size in the original HS network, its two randomized versions, and an ER network of the same size and with the same average degree (although the ER network has a different degree sequence and distribution) is shown in Fig. 2.10. The two methods of randomization (x-swap and exact sampling) roughly give the same cascade size S. In case of randomized networks, for some realizations spread reaches the full network [Fig. 2.8(c)] and for some realizations spread is minuscule [Fig. 2.8(b)] and therefore $S < 1$.

Finally, analogous to Fig. 2.2, we show the emergence of global cascades (at

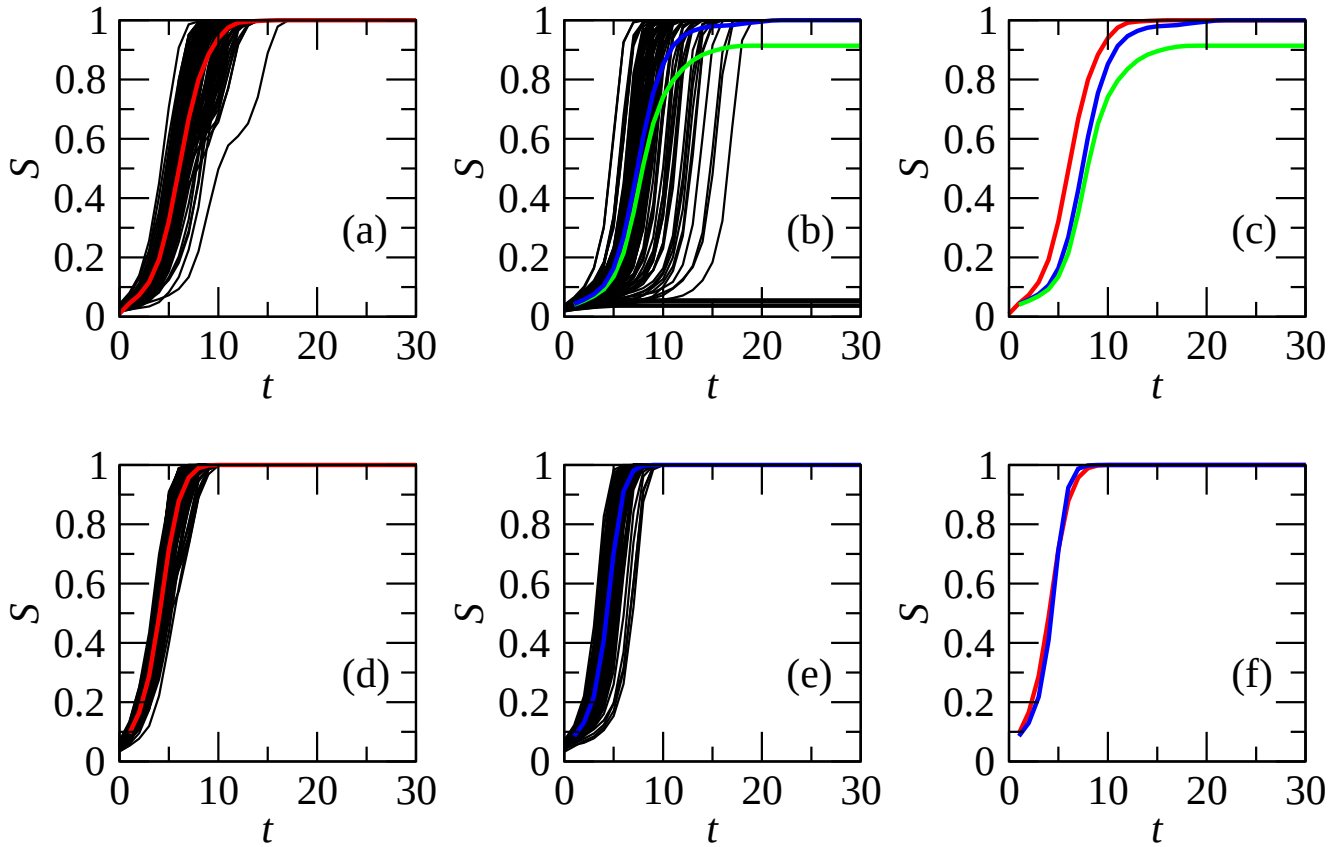

Figure 2.7: Time evolution of the size of the cascades S on the high-school (HS) network and its randomized version by X-swaps with identical degree sequence, with $N = 921$, $\langle k \rangle = 5.96$, and $\phi = 0.18$ for two different values of fraction of initiators. (a) HS friendship network and (b) its x-swapped randomized version. (c) Direct comparison of the ensemble-averaged time series for the original HS network (red solid curve) and for its x-swapped randomized version (green solid curve); blue solid curves represent conditional average over runs for which the spread reaches the entire network. Thin black curves in (a) and (b) are individual time series. The fraction of initiators for (a–c) is $p = 0.01$. (d) HS friendship network and (e) its x-swapped randomized version. (f) Direct comparison of the ensemble-averaged time series for the original HS network (red solid curve) and for its x-swapped randomized version; blue solid curves represent conditional average over runs for which the spread reaches the entire network. Thin black curves in (d) and (e) are individual time series. The fraction of initiators for (d–f) is $p = 0.02$.

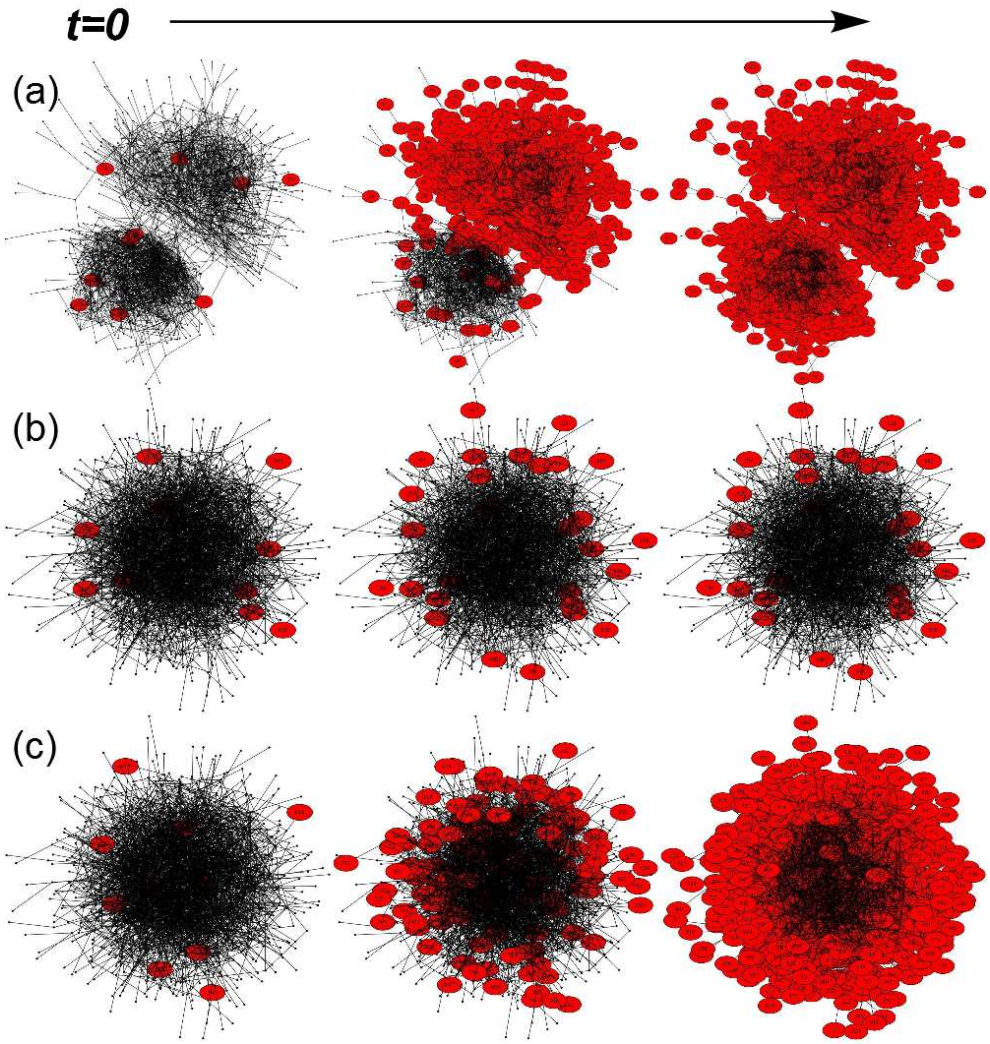

Figure 2.8: Visualizations of spreading in the threshold model (typical individual runs) for various networks at different times during evolution (arrow on top indicates the direction of time evolution). $N = 921$, $\langle k \rangle = 5.96$, $p = 0.01$ and $\phi = 0.18$. Nodes in state 1 (active nodes) are colored red. (a) Original high-school network; (b) Randomized network (by X-Swap) when eventual spread is local; and (c) The same randomized network but for a run that reaches the whole network.

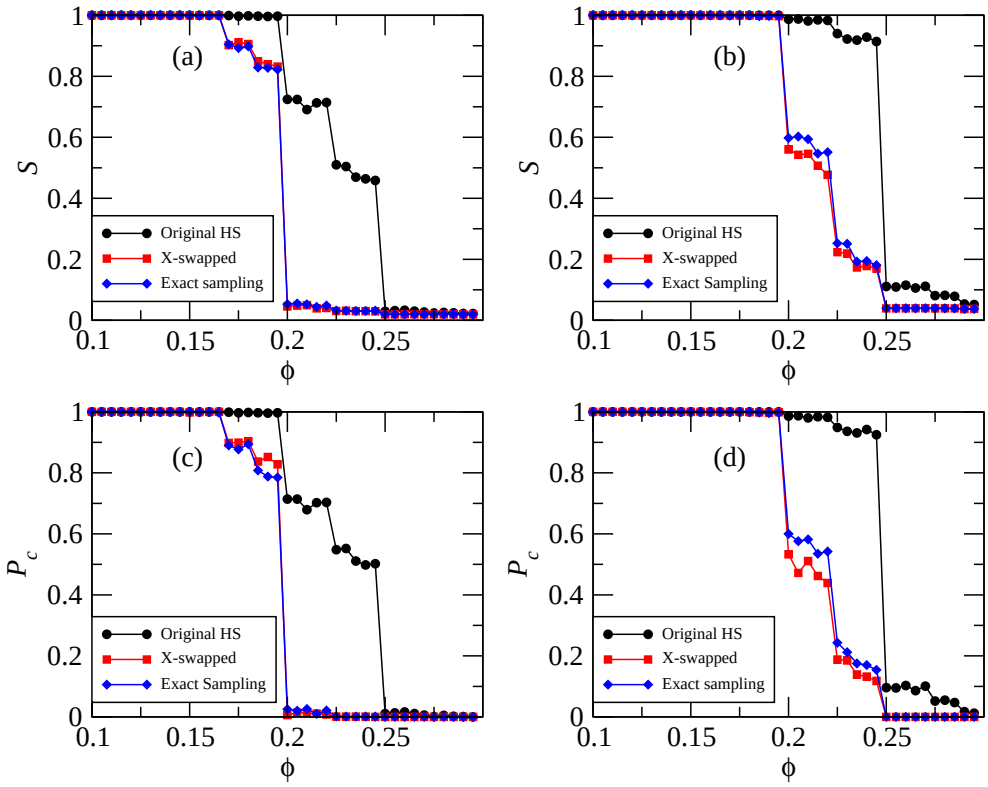

Figure 2.9: Cascade size and probability of global cascades in the high-school (HS) network and its two randomized versions with identical degree sequence (x-swapped and exact sampling [84]) with $N = 921$ and $\langle k \rangle = 5.96$. Cascade size S as a function of ϕ for (a) $p = 0.01$ and for (b) $p = 0.02$. Probability of global cascades P_c as a function of ϕ for (c) $p = 0.01$ and for (d) $p = 0.02$.

the tipping point p_c) in the high-school network, as the density of initiators is varied [Fig. 2.11].

2.5 Conclusions

This study focuses on identifying a tipping point for multiple initiators in the threshold model. For high individual thresholds ($\phi > 1/\langle k \rangle$), a small initiator fraction is incapable of triggering cascades. Global cascades of opinion can be triggered in the system, for arbitrarily high values of ϕ, if the initiator fraction exceeds a critical value p_c. The existence of such a tipping point is shown in the case of ER networks and an empirical high-school network in a similar way.

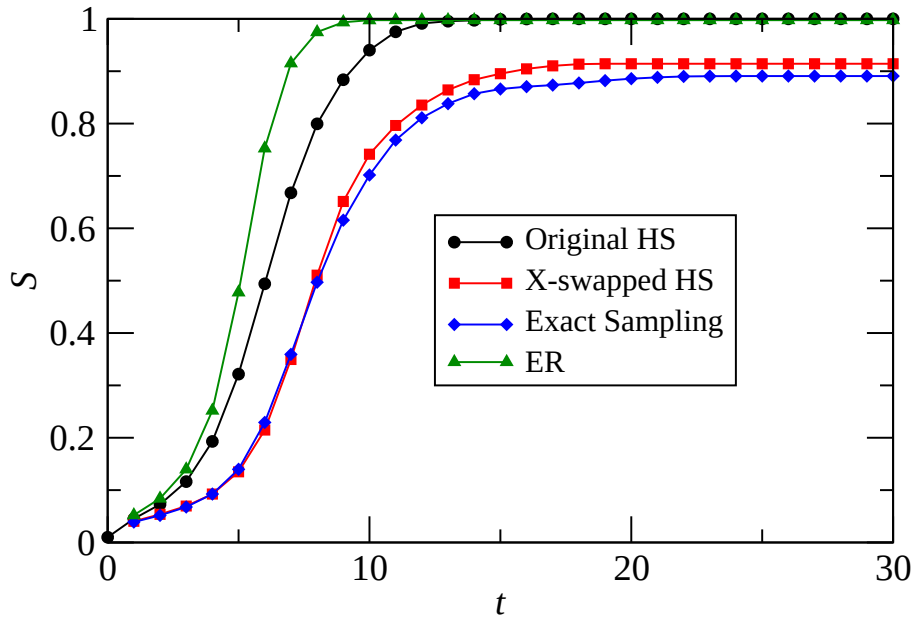

Figure 2.10: **Average cascade size as a function of time in the high-school network (HS), in its two randomized versions with identical degree sequence (x-swapped and exact sampling [84]), and in ER networks with the same average degree.** $N = 921$, $\langle k \rangle = 5.96$, $p = 0.01$ **and** $\phi = 0.18$ **in all cases.**

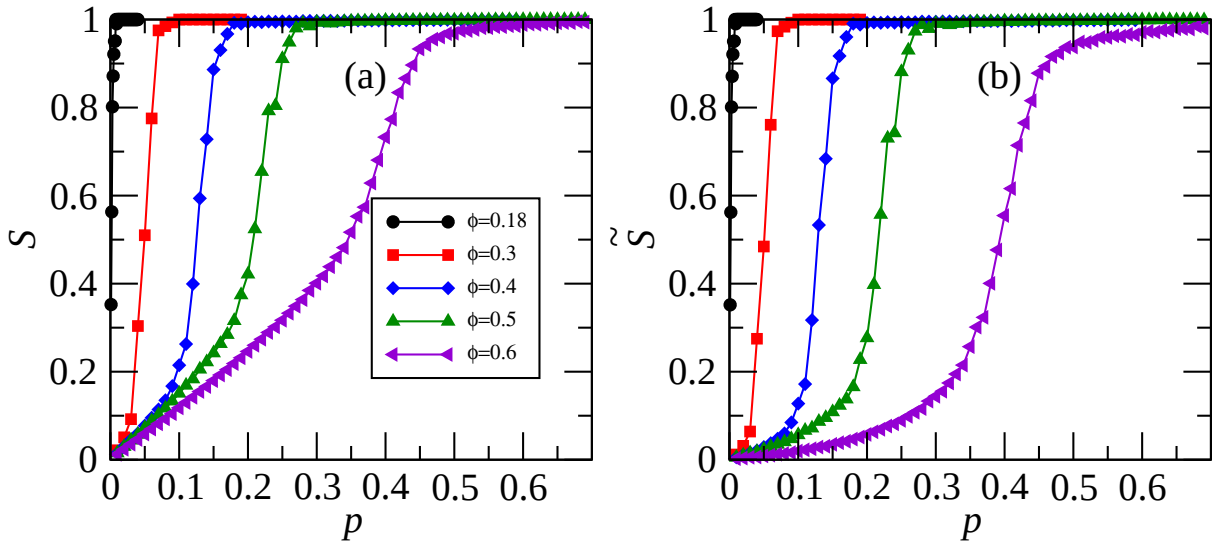

Figure 2.11: **Cascade size and scaled cascade size as a function of initiators on the high-school network** ($N = 921$, $\langle k \rangle = 5.96$)**. (a) Cascade size** S **as a function of initiators** p **for different values of** ϕ**. (b) Scaled cascade size** $\tilde{S}$ **[Eq. (2.1)] vs.** p **for different values of** ϕ**.**

Further, three different heuristic strategies to select a fraction of initiator nodes were compared on ER networks and an empirical network. The simulation results demonstrate that selecting initiators by descending order of their degree results in fastest as well as largest cascades.

Effects of network structure, in particular community structure and local clustering, were studied by simulating the threshold dynamics on an empirical network and a few degree-sequence preserving randomized versions of it. The cascade window was found to be narrower in the case of randomized networks. The results show that local clustering in the network helps in spreading an opinion where the adoption of the new opinion by a node depends on its local threshold.

CHAPTER 3
Consensus in coevolving homophilous networks

In the previous chapter, a model of influence (the threshold model) in social networks was studied where the structure of the social network during the spreading dynamics was considered to be static. But, more generally, social networks are not necessarily static and may be evolving in time. Considering social networks as static greatly simplifies the analysis and is a valid assumption in the case where the network structure either does not change at all or the rate at which structural changes happen is much slower than the actual dynamical process occurring on the network and does not play a significant role. In the model discussed in this chapter, the network is considered to be structurally evolving. The evolution of network structure in this model is driven by homophily. The idea of homophily was introduced by Lazarsfeld and Merton [3, 40], according to which individuals tend to form social connections with those who are similar to them. Complementarily, the persistence of ties is also thought to depend strongly on the similarity of the individuals they connect [86–88].

As the individuals influence one another to adopt new behaviors, opinions, ideas or other traits, the similarity between any pair of nodes keeps evolving. Thus, one can envision the structure of a social network as being in a constant state of flux: links between dissimilar individuals decay with time while new ties between similar individuals form at some rate. This continuous death and birth of links is presumably balanced in such a way that on average, at any given time, the mean number of connections ascribed to any individual is roughly constant, or at least, bounded from above [89].

The present model is a variant of the Axelrod model [41] and similar to [50] where nodes are assigned a set of attributes to quantify similarity and a homophily-driven rewiring rule is imposed. However, the focus of this study is to understand

Portions of this chapter previously appeared as: P. Singh, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Accelerating consensus on coevolving networks: The effect of committed individuals," Phys. Rev. E **85,** 046104 (2012) [67].

the effect of *committed agents* - individuals who are selectively immune to influence on a given issue, and who hold the same opinion on that issue - on the scaling of consensus time with the system size N in dynamical networks. The effect of committed individuals, all holding the same opinion, has been previously studied on structurally static networks [1, 10, 65, 66].

The key finding in these studies was the existence of a critical committed fraction below which the deterministic evolution equations admit a mixed steady state in addition to the always present consensus steady state, with a saddle point separating the two. As a consequence, on finite networks [10, 65], the time to attain consensus scales exponentially with system size when the committed fraction is below the critical value. This is consistent with the known scaling of transition times between deterministically stable states in stochastic bistable systems [90]. In stark contrast, above the critical value of the committed fraction, the consensus process is essentially deterministic with consensus times logarithmically dependent on network size, and where the only steady-state solution to the deterministic equations is the consensus state. Further analysis has also revealed [10, 64] that this "critical" point or threshold is a spinodal point associated with an underlying first-order transition [91] in the phase diagram of the model.

3.1 The model

In this model, individuals are represented by nodes on a network and every node is assigned a set of F independent attributes that constitute the node's *state*. Each attribute can take one of q distinct traits, represented by integers in $[0, q-1]$. Thus the node's state is represented by an F component vector. Initially, each attribute of each node is assigned one of the q values randomly. The structure of the network connecting the individuals is initialized to an Erdős-Rényi (ER) random network with a given average degree . Next, we define the rules governing the evolution of individual states as well as the structure of the network connecting them. At each time step: a node i is selected at random, and one of its neighbors j is selected at random. We then compute the *similarity* between nodes i and j, where similarity is defined as the number of attributes for which i and j possess the

same trait. Then,

1. If the similarity is found to be equal to or above a given threshold ϕ, node j adopts the trait possessed by node i for a randomly chosen attribute from among those for which they currently do not share the same trait. We refer to this as the *influence step*.

2. Otherwise, the link between i and j is severed and node i randomly selects a node k in the network from among those to which it is currently not connected, and forms a link to it. We refer to this as the *rewiring step*.

Our model clearly is a variant of the Axelrod model with the following distinctions: in the Axelrod model, the influence step occurs with probability proportional to the similarity between the nodes whereas in our model influence occurs only when similarity exceeds a hard threshold. Secondly, given that the chosen nodes are not similar on all attributes, if influence does not occur, then a rewiring step necessarily does. This hard threshold also distinguishes our model from that of [50], although as shown below the qualitative behavior of both models is similar. It is worth pointing out that the rewiring step is designed such that the total number of links (or average degree) in the network is conserved. Also, the update rule as defined above always assumes that the node chosen first is the *influencer* and the node chosen second is the *adopter*. While most of the results described in this study are restricted to this order in choosing the influencer and adopter, we discuss alternate orders of selection (motivated by the results in [8]) in Section 3.3.

We first examine the effect of the number of traits per attribute q in determining the steady state structure of the network, and confirm that our results are similar to those found in [50, 92]. We fix $F = 5$ and set the similarity threshold to be $\phi = 3$. In particular, three phases differing in the steady state network structure are observed, as q is varied, as shown in Fig. 4.2. In the first phase, observed for low values of q, the system evolves to a state where the network structure is static and global consensus is achieved i.e. for each attribute, each node possesses the same trait as all other nodes. As the number of traits is increased, at a critical value $q = q_c$, the system undergoes a phase transition to phase 2 where the steady state of

the network consists of disconnected fragments with each fragment coming to consensus locally. While the first two phases differ in the eventual network structure, they are similar in that the system eventually reaches a *frozen* state in which neither the node traits nor the network structure are evolving. However, further increasing q beyond phase 2, eventually reveals a third phase where the initial dissimilarity among the nodes is so large that the system does not end up in a frozen state and rewiring continues indefinitely (Fig. 4.2). In this phase, at any given time, there exists a giant component; however, in the asymptotic limit of network size, the system will never reach consensus, so long as the average degree is not too small [50]. This three-phase behavior is expected to be seen for different choices of F and ϕ. For fixed F, as ϕ is increased the transitions occur at smaller values of q. For fixed ϕ, as F is increased the transition points move to higher values of q (Fig. 3.2).

Interestingly, in a related model studied by Garcia-Lazaro *et al.* [93], for which the probability of rewiring is governed by a *dissimilarity* threshold, the behavior of q_c as a function of ϕ is observed to be nonmonotonic, in contrast to the trend shown in Fig 3.2. In particular, for intermediate to low dissimilarity thresholds (analogous to going from intermediate to large ϕ in our model), they find that q_c progressively increases. The origin of this difference is in the inherently probabilistic update rules prescribed in Ref. [93], as opposed to the deterministic ones in our model. In other words, the low probabilities of rewiring a link between two very similar nodes and of influence occurring between two highly dissimilar neighbors, both of which are present in the model in Ref. [93], unlike in ours, are responsible for the nonmonotonic behavior of q_c that they observe.

Figure 3.3 shows the convergence time T_c for the choice of $F = 5$, $q = 2$, as a function of ϕ. The nonmonotonicity in convergence times seen here is a consequence of the fragmentation transition that occurs as ϕ is increased, keeping F and q fixed. Indeed, as Fig. 3.2 shows, as ϕ is increased, the fragmentation transition occurs progressively earlier in q. As ϕ is increased from 0 to 3, the rate at which an influence step occurs decreases, and since the rewiring rate is not high enough to fragment the network, the system converges when there is global consensus. This takes progressively longer from $\phi = 0$ to $\phi = 3$, as the influence rate is decreasing,

and the maximal value of convergence time is obtained for $\phi = 3$. Beyond $\phi = 3$, the initial dissimilarities (with regard to ϕ) are sufficiently high to cause the network to fragment, and nodes in each component are similar enough that consensus within each component is reached in a relatively short time. Thus, convergence times are identical to the time to reach global consensus only when $\phi \leq 3$. We point out that a similar nonmonotonic behavior for convergence times was also observed in Ref. [44]. From here onward, in order to study how the approach to consensus can be speeded up, we continue with $F = 5$ and $\phi = 3$ and keep the number of traits per attribute q fixed at 2, so that the system is confined to phase 1, where both a giant component and global consensus are guaranteed. Furthermore, as is clear from the preceding discussion, for $F = 5$ and $q = 2$, the choice $\phi = 3$ allows us to explore the worst-case scenario from the point of view of consensus times within the phase where an unfragmented network is guaranteed. Henceforth, for notational simplicity, we use T_c to mean consensus time. Since the focus of our study in what follows is confined to phase 1 where consensus is global, this terminology does not lead to any ambiguity.

For these parameters, when a pair of neighboring nodes is selected at time $t = 0$, the probability of them meeting the similarity threshold, and hence the probability of influence, as one can see by elementary combinatorics, is exactly $\frac{1}{2}$ and equal to the probability of rewiring at time $t = 0$. Since for the chosen parameters the system is in phase 1, the steady state reached is one where the network consists of a single connected component and the states of all nodes in this component are identical i.e. nodes are similar in the traits they possess for all F attributes. Thus the steady state is one where a consensus is reached. Shown in Fig. 3.4 is the scaling of the consensus time T_c as a function of the system size. We contrast the behavior of a system where node attributes and network structure coevolve to the behavior of a system where the node attributes evolve on a purely static network. (In other words, in the latter case, if the chosen nodes in an update step do not meet the similarity threshold, nothing happens.) As shown clearly, the effect of rewiring is detrimental to consensus times. With rewiring present, T_c is exponential in N, in contrast to a linear scaling found for the static network. The divergence of T_c with

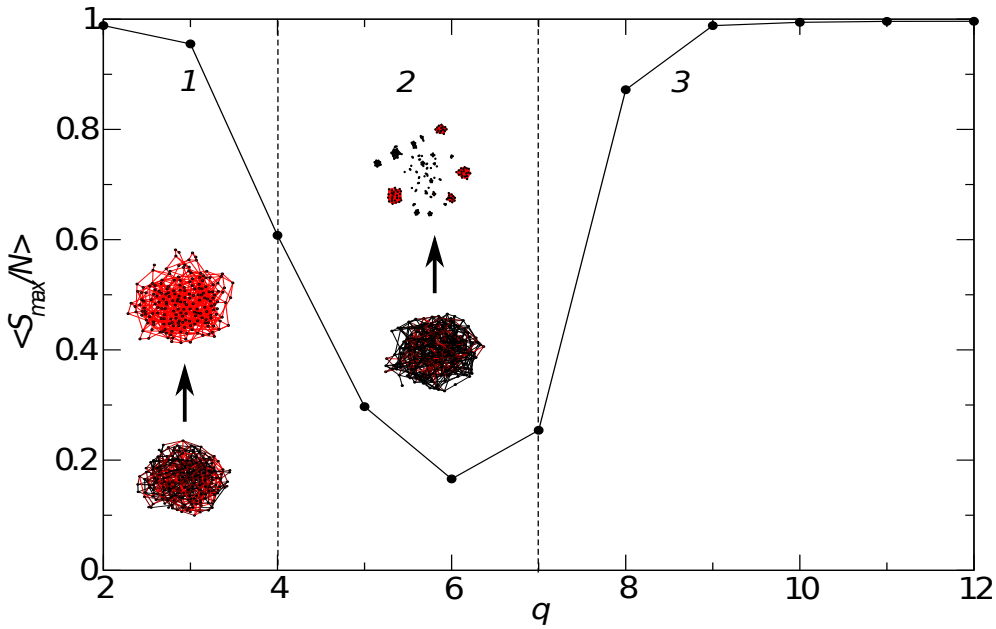

Figure 3.1: The average size of the largest connected component (cluster), $\langle S_{max} \rangle$, in the final state of the system as a function of q, starting from an ER Network with $N = 200$, $\langle k \rangle = 6.0$ and threshold $\phi = 3$. The average is taken over 100 realizations of network evolution. The plot shows three different phases characterized by distinct steady states of network structure (see text). This phase diagram is analogous to the one shown in Vazquez et al. [50] for a related model. Also shown for phases 1 and 2 are initial (lower) and steady-state (upper) network snapshots for a single realization of co-evolution. The colors on the nodes represent the traits for a given attribute that is being tracked in the visualization. Snapshots in phase 1 (2) have $q = 2$ ($q = 7$) traits per attribute, and initially begin with each node having equal probability, $1/q$, of possessing any trait per attribute. We represent trait 0 for the tracked attribute by red, and all other traits by black. Therefore, in the steady-state snapshot of phase 2, each black cluster represents that all of its constituent nodes have adopted the same trait for the tracked attribute, but this common trait is different from 0.

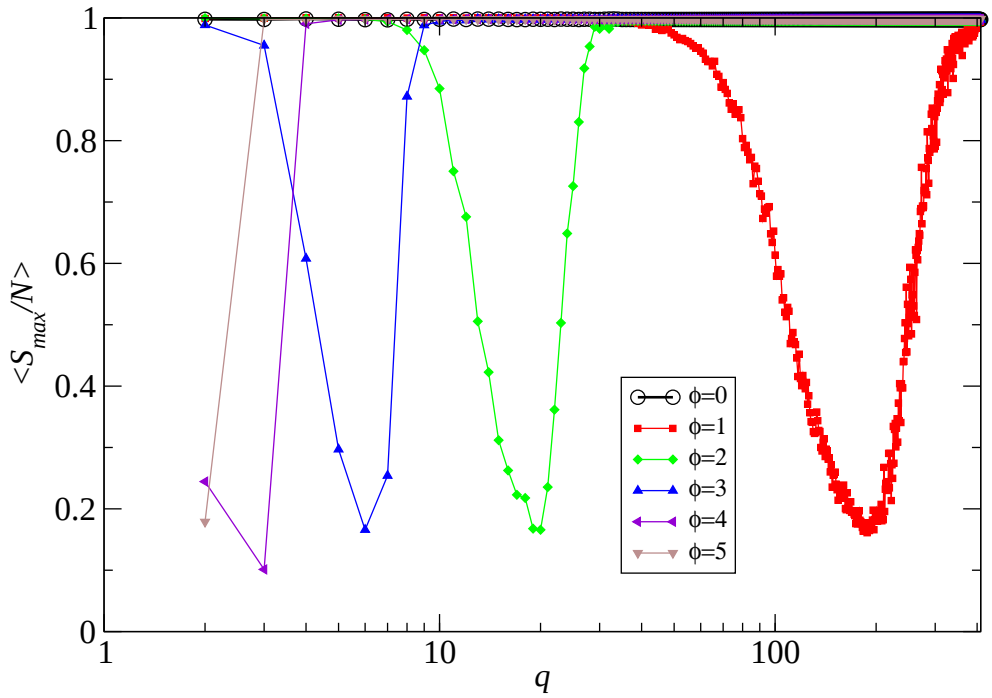

Figure 3.2: Phase diagrams of the system as in Fig. 4.2 for all possible values of ϕ (for $F = 5$). As shown, the fragmentation transition takes place at successively higher values of q as ϕ is decreased, $\rightarrow \infty$ as $\phi \rightarrow 0$. When ϕ is large (4 or 5) network is fragmented even for the smallest non-trivial value of $q(= 2)$

N when rewiring is introduced can be qualitatively explained as follows. When a random node is chosen as the new end point of a rewired link, it is most probable that the chosen node has an attribute vector that is currently the most abundant vector - the majority state - in the population. If we assume that the attribute vector corresponding to the majority state does not change too frequently, then nodes in the majority state end up garnering a large number of rewired links. However, this comes at the detriment of the majority state, since when an influence step occurs and a random neighbor is chosen as the adopter, this neighbor is more likely to be in the majority state (due to the larger number of links leading to a node in the majority state). As a result, as soon as nodes in the majority state accrue more links than the rest of the population, they also become more likely to get influenced. This effect suppresses the growth of the majority state. Thus, the negative feedback due to rewiring strongly slows the spread of any particular state in the network. A

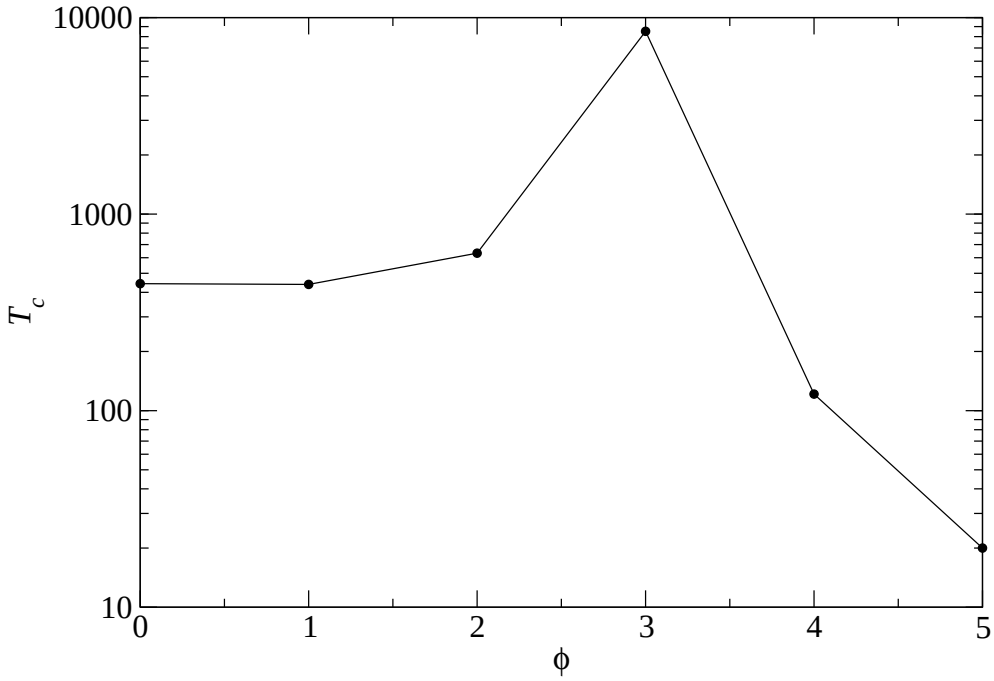

Figure 3.3: Consensus time T_c vs similarity threshold ϕ when the initial network is ER with $N = 200$ and the average degree $\langle k \rangle = 6.0$. T_c varies non-monotonically with ϕ and reaches its maximal value at $\phi = 3$.

similar negative feedback effect and exponential scaling of consensus times in the presence of rewiring was also observed in Ref. [8].

3.1.1 Committed agents

So far we have discussed how a consensus can be reached across all attributes i.e. the state vector of every node becomes identical. However, in a realistic scenario, it might be desirable to cause an entire population to adopt a given trait (opinion) for a given attribute (issue). For example, within a social network of teenage students in a high school, it is desirable that a consensus is reached where everyone views smoking as unhealthy behavior. Here, we study whether the introduction of committed agents [10, 11] can cause fast consensus on any given attribute characterizing the individuals. Without loss of generality we choose attribute 1 as the one that committed agents intend to engineer a consensus on. We refer to this attribute as the *designated* attribute (opinion). Also, we assume that committed agents rigidly adopt trait 1 for the designated attribute. Committed agents are thus considered

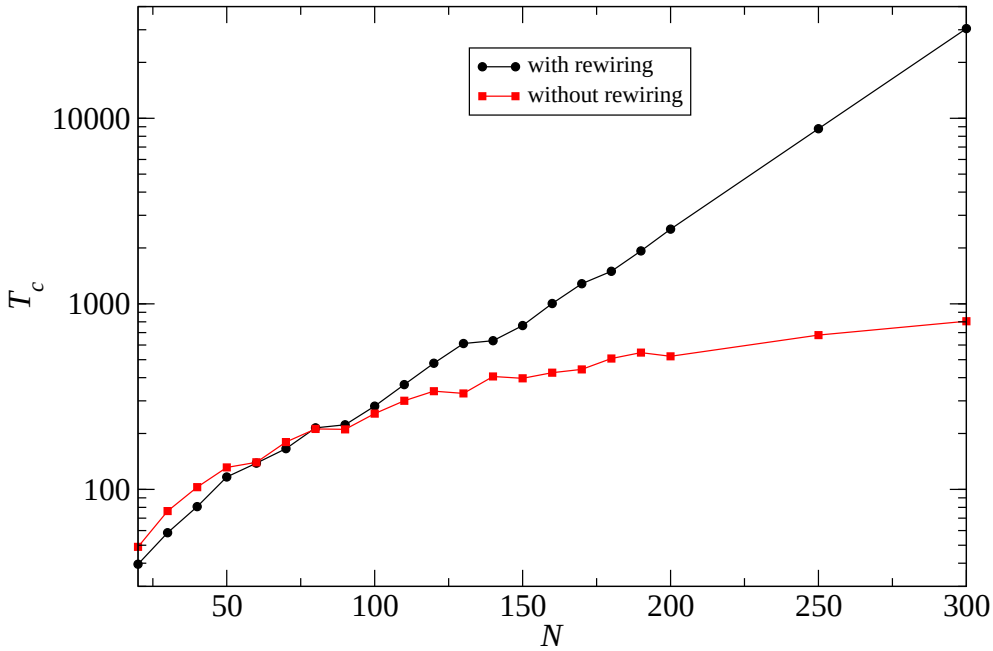

Figure 3.4: Comparison between consensus times for two different cases – when rewiring is present in the network (black circles) and when the network is static (red squares). These are results obtained from simulations, by averaging over 100 **realizations of network evolution. For each realization, the starting network is ER with** $\langle k \rangle = 6.0$**. Clearly, with rewiring** T_c **scales exponentially with** N **(** $T_c \sim \exp(\alpha N)$ **where** $\alpha \approx 0.02$**) while without rewiring the scaling is linear.**

uninfluenceable on attribute 1, but are identical in their behavior to uncommitted nodes for other attributes. In the following, we redefine consensus time T_c to mean the time taken for all nodes to adopt the trait proselytized by committed agents (i.e. 1) for the first attribute. Attributes besides the designated one for committed nodes, as well as all attributes for uncommitted nodes are initialized to 1 or 0 with equal probability.

In Fig. 3.5, we show the effect of introducing committed agents into the network. In particular, we choose randomly, a fraction p of nodes as committed agents and study how the consensus time scaling with N changes as p is varied. As expected, for $p = 0$ the exponential scaling of T_c with N is recovered. Although consensus time values decrease for any given N as p is increased, the scaling behavior remains unchanged until a critical fraction of $p = p_c \approx 0.1$ is reached. Beyond

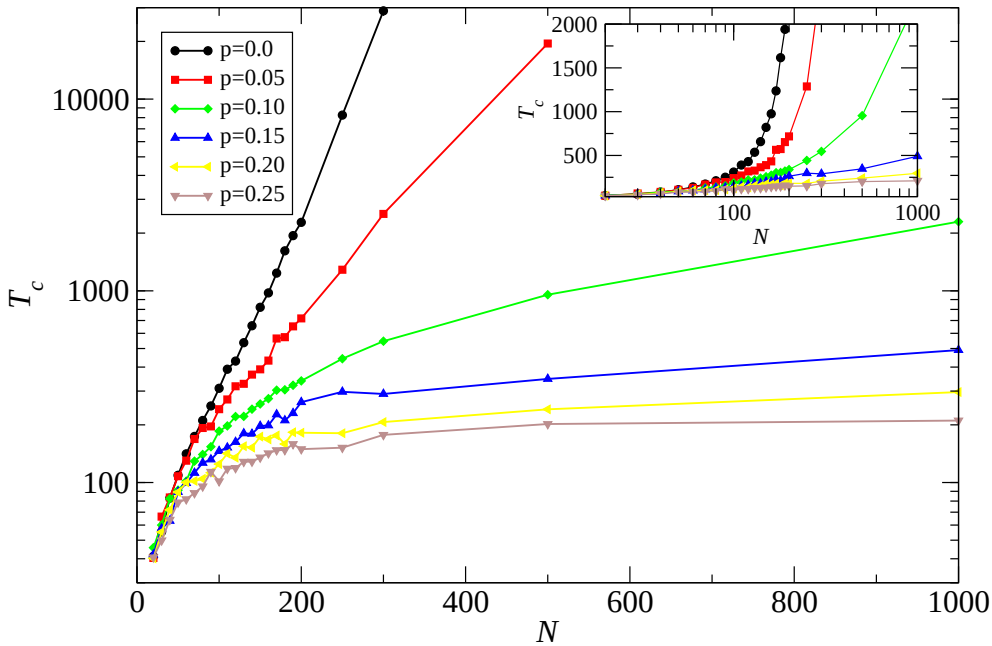

Figure 3.5: Simulation results for consensus time T_c as a function of network size N when the initial network is ER with average degree $\langle k \rangle = 6.0$ for different values of the committed fraction p. In these simulations, the influencer is chosen first, followed by the adopter. For low p, T_c diverges exponentially with N. At the critical value $p_c \approx 0.1$, the system undergoes a transition beyond which $T_c \sim \log(N)$ (see inset). Precisely at p_c, T_c appears to scale linearly with N.

this critical fraction, consensus times dramatically change their scaling behavior and become logarithmic in N. This result indicates that beyond a critical value of the committed fraction, this fraction can efficiently overcome the resistance to consensus generated by the random rewiring taking place in the network. Fig. 3.8(a) shows consensus time as a function of p, for different network sizes.

Although the complexity of our model makes it analytically intractable, we present a qualitative explanation of the transition gleaned from our extensive simulations. We characterize the state of the system by the order parameter $m = 2n_1 - 1$ (analogous to the magnetization in spin systems), where n_1 is the density of nodes which adopt trait 1 for the designated attribute. In the absence of committed agents, the system settles in a stable steady state at $m = 0$ [see Fig. 3.6]. As explained in the previous section and Ref. [8], the negative feedback effect of the dynamics

makes the consensus states $m = -1$ and $m = 1$ repulsive absorbing points of the system, which can be reached only via a noise-driven escape event. Thus, for a finite system, in the absence of committed agents, $m = 0$ constitutes a metastable state. When committed agents are introduced, they impart the system a bias toward the consensus state $m = 1$ favored by them. This can be seen in Figs. 3.6(b) and 3.6(c). Thus, while the absorbing states at $m = 1$ and $m = -1$ continue to be repelling, the location of the metastable state gets progressively closer to the $m = 1$ absorbing state, as p gets close to p_c. In the vicinity of p_c, however, it becomes difficult to obtain a clean estimate of the metastable state order parameter, since the growing proximity of the stable steady state to the absorbing point $m = 1$ results in increasingly quick noise-driven escapes to consensus. However, we intuitively expect that the transition occurs when the location of the metastable state coincides with $m = 1$, making this consensus state, a stable absorbing point of the system.

For a cleaner observation of the transition via simulations, we study the survival probability P_s, defined as the fraction of realizations that have not reached consensus up to some observation time T. Shown in Fig. 3.7 are individual plots for different system sizes of P_s as a function of committed agent fraction p, for different observation times T. As is clearly shown, increasing N results in the drop in survival probability becoming steeper and occurring at successively higher values of p, with p_c appearing to converge to a value around ≈ 0.1. The qualitative picture obtained from studying the order parameter trajectories (Fig. 3.6) and survival probabilities (Fig. 3.7) appears to indicate that the transition is continuous (second order).

3.2 Finite-size scaling analysis

In this section we employ finite-size scaling analysis for the consensus time [47, 94] to obtain an estimate of the critical committed fraction p_c from our simulation results. In parlance with the theory of phase transitions, we assume the following scaling ansatz for T_c:

$$T_c(p, N) = N^{\alpha} f(N^{\beta}(p - p_c)) \tag{3.1}$$

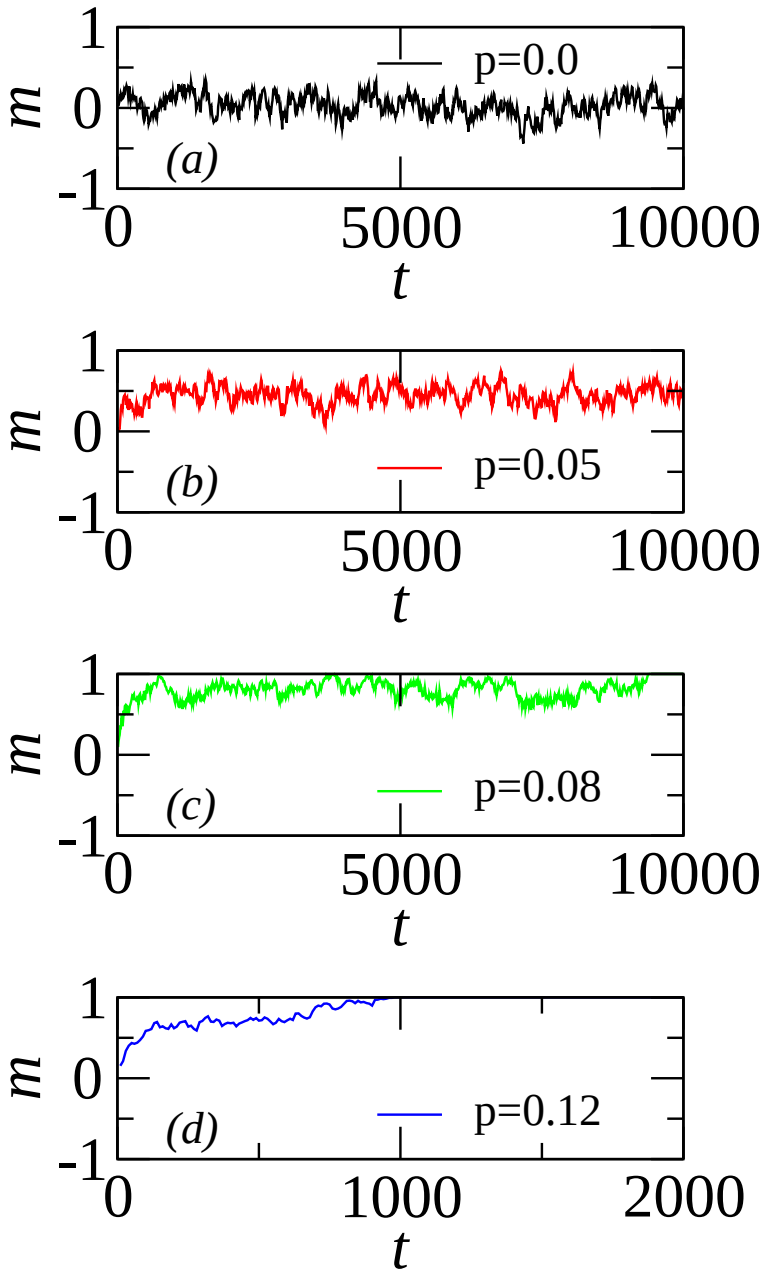

Figure 3.6: Individual trajectories of a network of size $N = 1000$ for different committed agent fractions: (a) $p = 0$, (b) $p = 0.05$, (c) $p = 0.08$, and (d) $p = 0.12$. For values of p less than $p_c \approx 0.1$ (ac), the system appears to settle to a metastable state whose order parameter depends on p, until a fluctuation results in the system's getting absorbed by the consensus state with order parameter $m = 1$. For $p > p_c$, in contrast, the order parameter steadily drifts toward the consensus state $m = 1$.

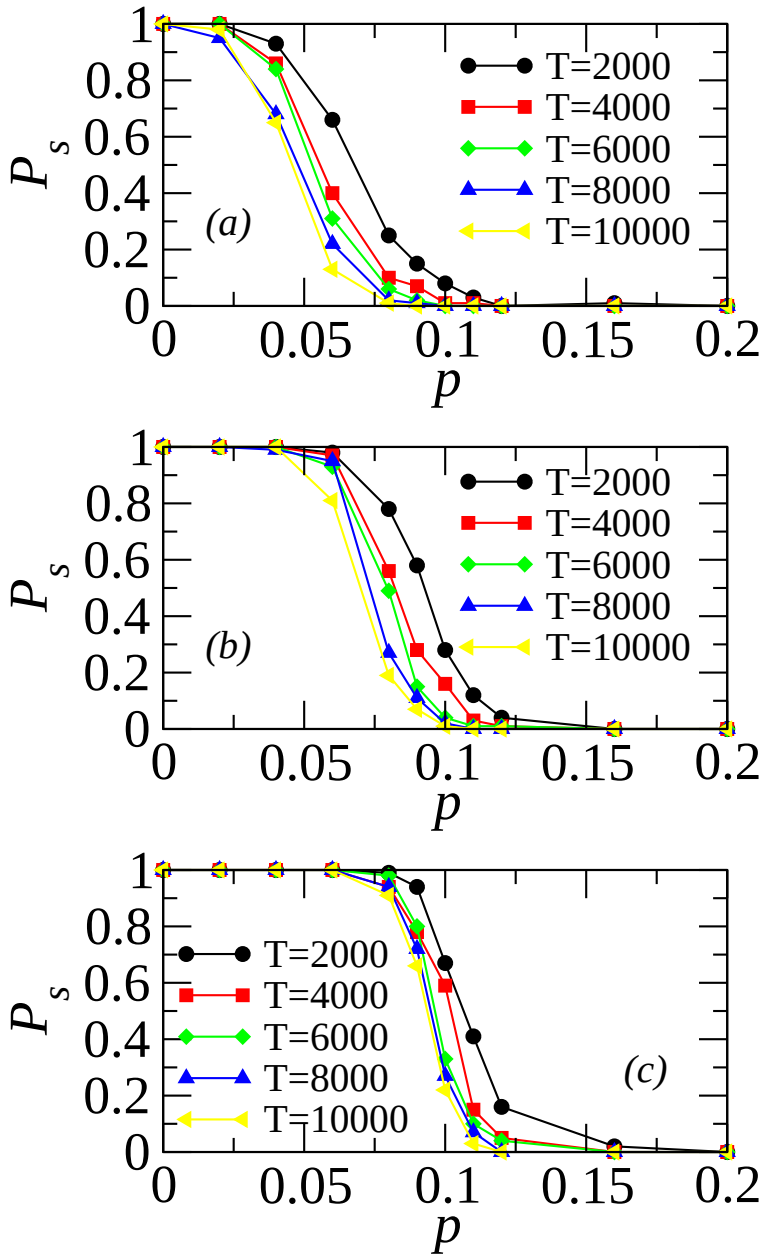

Figure 3.7: Survival probability P_s as a function of committed agent fraction p for several observation times and for system sizes (a) $N = 500$, (b) $N = 1000$, and (c) $N = 2000$. For a given N, the drop in P_s becomes steeper as T increases, while for a given T, the drop becomes steeper as N increases.

where p_c is the critical value of p and α and β are critical exponents. The function f is an unknown scaling function.

At $p = p_c$, the scaling ansatz yields $T_c N^{-\alpha} = f(0)$, which implies that if $T_c N^{-\alpha}$ is plotted as a function of p, curves for different N values must intersect at $p = p_c$. Since α itself is unknown, we progressively increase α from values close to zero until curves of $T_c N^{-\alpha}$ versus p corresponding to different N, intersect at a single point. We find that a single intersection point is obtained for $\alpha = 1$ and this point is at $p = p_c \approx 0.1$ (see Fig. 3.8)(b). To obtain the exponent β, we plot $\frac{T_c}{N^\alpha}$ as a function of $N^\beta(p - p_c)$, and vary β until a good scaling collapse is found for all curves near p_c. Following this procedure we find $\beta \approx 1$. The scaling collapse is shown in Fig. 3.8(c), indicating good agreement with Eq. (3.1).

Following the same scaling ansatz, we can also scale the survival probabilities obtained (see Fig. 3.7) for different system sizes. Assuming the survival probability to decay exponentially with time (as is expected in metastable-escape problems), $P_s \sim exp(-T/\tau)$, where τ is identical to the mean lifetime or the mean consensus time T_c. Therefore, with the assumption that $\alpha = 1$ and $\beta = 1$, we expect

$$\frac{N}{T} \ln(P_s) \sim \frac{1}{f(N(p - p_c))} \tag{3.2}$$

Fig 3.9 shows the good scaling collapse obtained for survival probabilities of different system sizes, and observation time $T = 10000$, in agreement with Eq. (3.2).

3.3 Effect of influencer-adopter selection order

In this section we study how the order of selecting the influencer and adopter nodes affects consensus times, and also the transition observed in consensus times as the committed fraction is varied. Surprisingly, such subtle changes in selection order can have a profound effect on consensus times.

3.3.1 Adopter-first selection

In adopter-first selection, a random node i is selected, followed by a randomly chosen neighbor j of i. However in contrast to the model studied so far, if the criterion for an influence step to occur is met, then i adopts the trait of j for a randomly

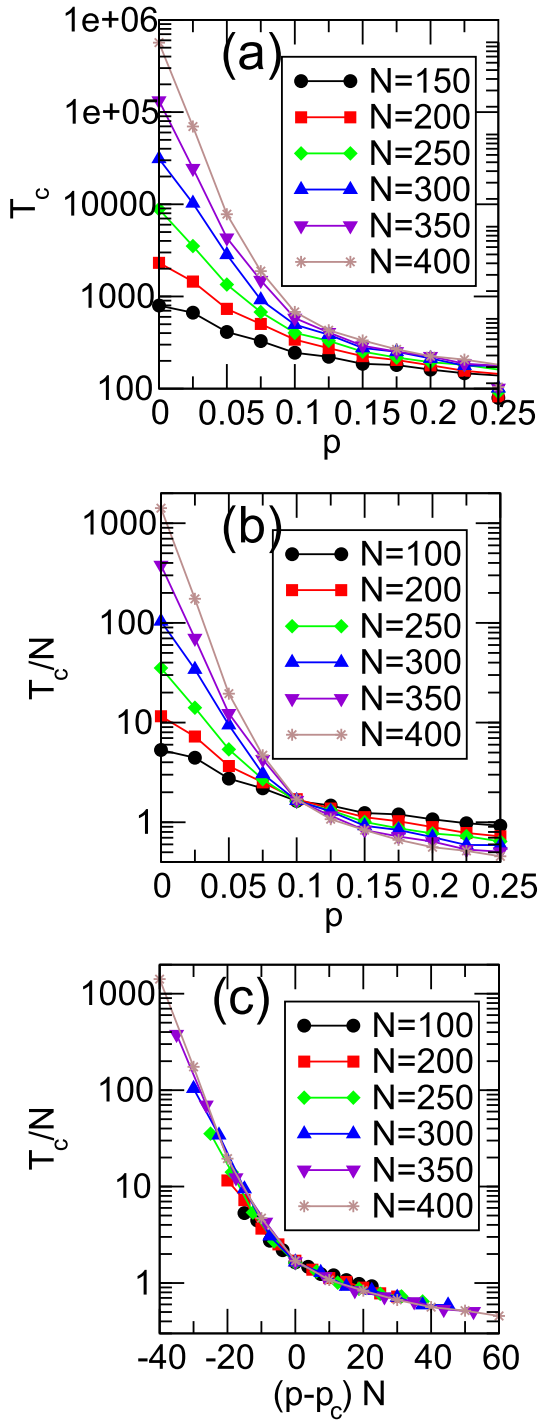

Figure 3.8: **(a)** T_c **vs** p **starting from an ER network with average degree** $\langle k \rangle = 6.0$ **for different system sizes** N**. (b)**$\frac{T_c}{N^\alpha}$ **with** $\alpha = 1.0$ **vs** p **for the data in (a). The critical point - the value of** p **at which the curves intersect - is** $p_c \approx 0.1$ **(see text). (c)** $\frac{T_c}{N^\alpha}$ $(\alpha = 1.0)$ **vs** $N^\beta(p - p_c)$ **shows good scaling collapse in the vicinity of** p_c **for** $\beta = 1.0$**.**

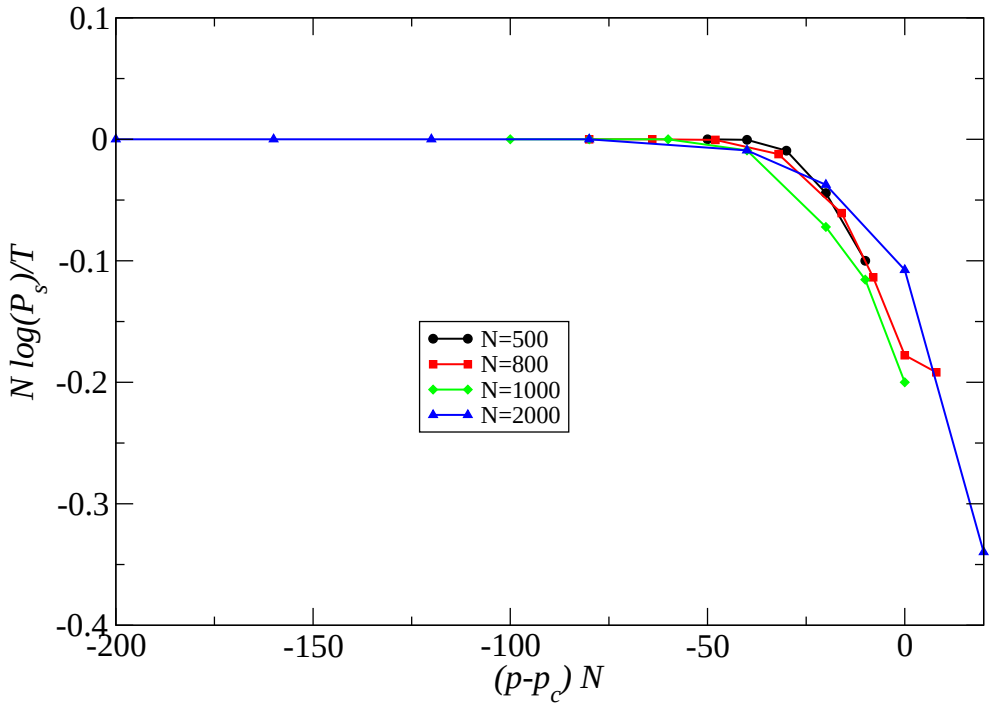

Figure 3.9: Scaling collapse for survival probabilities obtained for four system sizes ($N = 500, 800, 1000, 2000$) and observation time $T = 10000$, showing good agreement with the scaling ansatz, Eq. (3.2)

chosen attribute from among those for which they currently do not share the same trait. Thus, as the name suggests, the first chosen node is treated as the adopter and the second as the influencer. A comparison between consensus times on static networks and networks with rewiring under this updating scheme (Fig. 3.10) suggests that rewiring generates a positive feedback which aids the network in reaching consensus (also observed in [8]). This is in contrast to what was observed in the previous case of influencer-first selection.

Accordingly, for adopter-first dynamics, the previously observed exponential divergence of T_c with N is absent even in the case when the committed fraction $p = 0$. Instead, with $p = 0$, consensus time grows as a power law in N, with an exponent ≈ 0.7. When committed nodes are introduced, i.e., for $p > 0$, $T_c \sim \log N$ suggesting the existence of a transition in the scaling behavior of T_c occurring at or very close to $p = 0$ (Fig. 3.11).

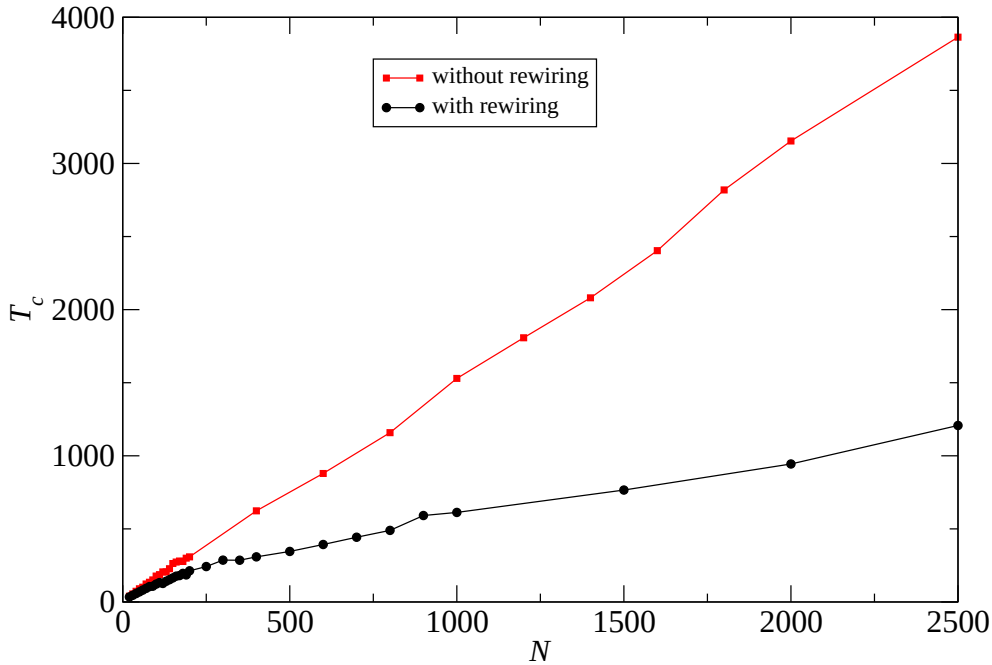

Figure 3.10: Consensus times for two different cases: when rewiring is allowed in the network (black circles) and when the network is static (red squares) when first chosen node is the adopter (and $p = 0.0$). The initial network is ER with $\langle k \rangle = 6.0$. T_c on a static network scales linearly with N but with rewiring T_c follows a power law with exponent ≈ 0.7.

3.3.2 Unbiased selection

In the selection orders considered so far, the direction of influence is always fixed beforehand. Here, we investigate the case in which, after a pair of neighboring nodes is selected, a random one among them is chosen to be the influencer and the other, the adopter (provided the criterion for an influence step is met). In other words, a pair of nodes (i, j) is selected, and with probability $\frac{1}{2}$ node i adopts one of node j's traits, and otherwise node j adopts one of node i's traits.

First, we investigate the effect of rewiring on consensus times. As Fig. 3.12 shows, rewiring has no effect on the scaling of consensus times, and T_c scales linearly with N in both cases. This is expected since the dynamics does not favor the instantaneous growth of one state over the other in the population. However, introducing committed agents into the system (in the presence of rewiring) gives rise to a transition in the scaling behavior of T_c in this case as well: $T_c \sim N$ for

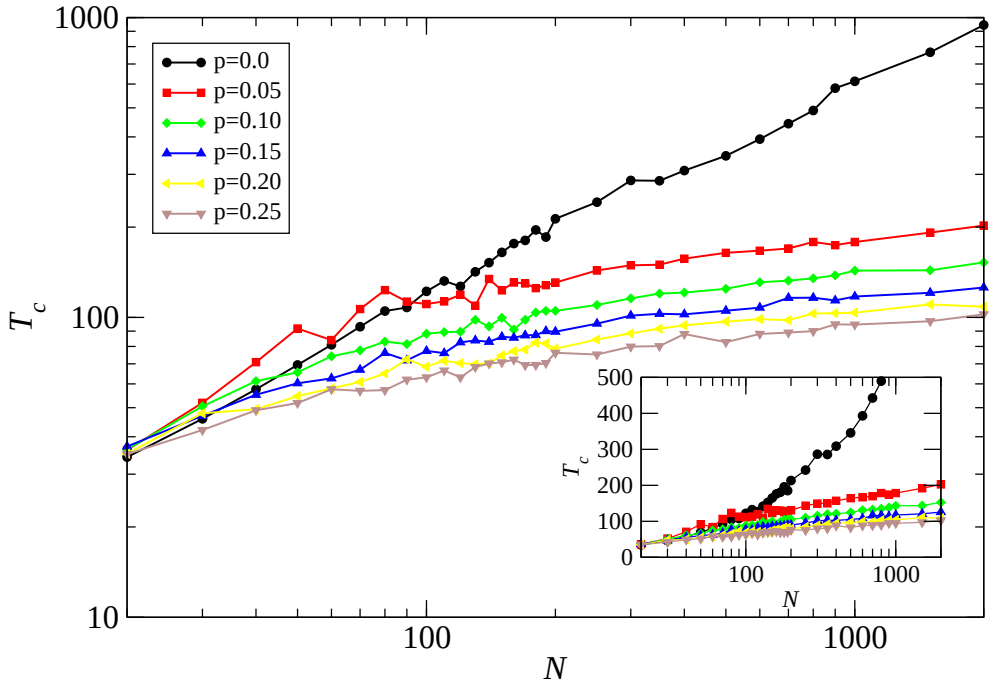

Figure 3.11: Scaling of consensus time, T_c, as a function of network size, N, for different values of the committed fraction p for adopter-first dynamics. In the absence of committed agents $T_c \sim N^{0.7}$ and is logarithmic in N for any $p > 0$ (see inset).

$p = 0$, while $T_c \sim \log N$ for $p > 0$ (Fig. 3.13).

3.4 Conclusions

In this chapter, using a variant of the Axelrod model with homophily-driven rewiring, we have studied how a small fraction of committed agents can dramatically influence the scaling of consensus times on structurally evolving networks. So far, the vast majority of models studying how a targeted change in opinion or behavior can be engineered have been confined to networks that are fixed in their topology. By considering the effect of persistent opinion holders - committed agents - on structurally evolving networks, we show that introducing a committed fraction $p > p_c$ represents a scalable method to cause widespread adoption of a given opinion on such networks. We have also considered variations to the update rule involving the selection order of influencers and adopters and shown that the transition in scaling behavior of T_c across some p_c is a consistent feature across these variations, even

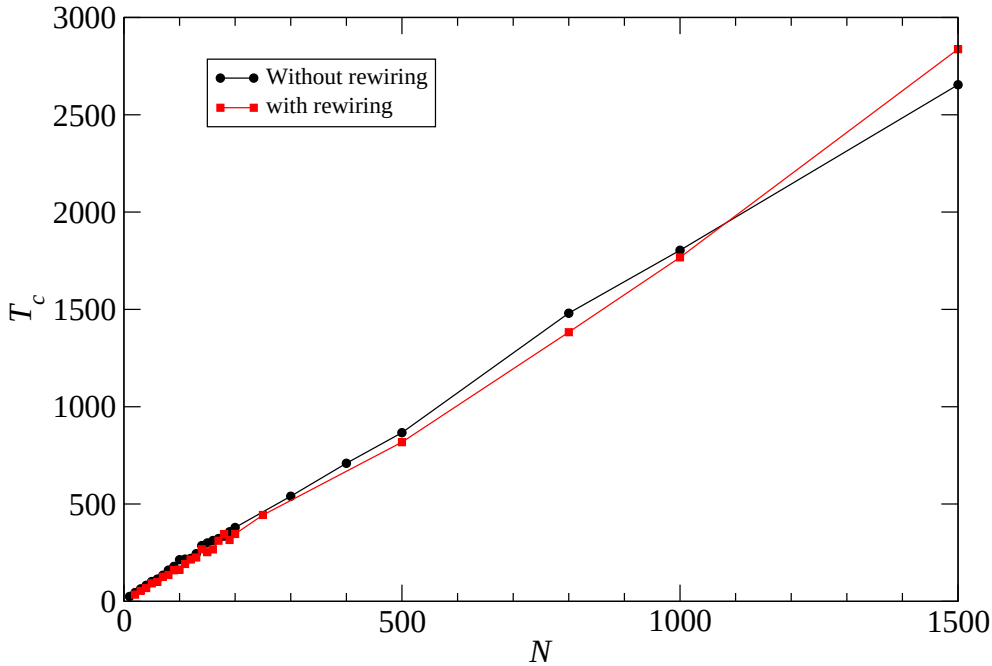

Figure 3.12: Consensus times for two different cases: when rewiring is allowed in the network (black circles) and when the network is static (red squares) for unbiased dynamics (and $p = 0.0$). The initial network is ER with $\langle k \rangle = 6.0$. T_c scales linearly with N in both cases, with nearly identical slopes.

though the precise scaling behavior is dependent on the selection order. Moreover, our results show that in the worst case scenario for consensus time - the influencer-first case - the introduction of a critical number of committed nodes can result in a dramatic reduction in consensus time.

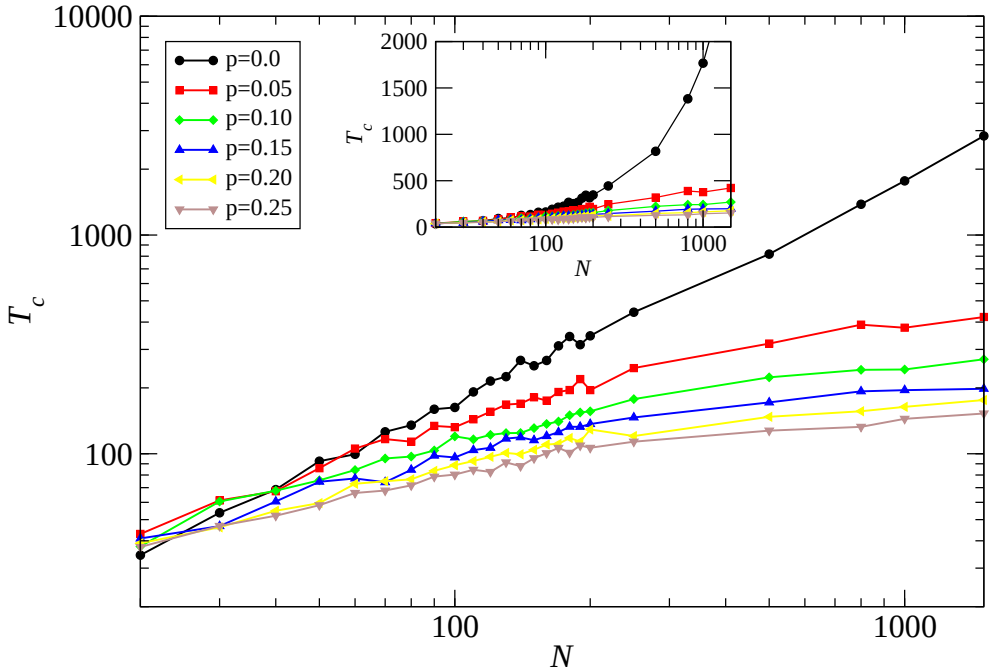

Figure 3.13: Scaling of consensus time T_c with network size N for different values of committed fraction p for unbiased dynamics. For these simulations, a microscopic update consisted of randomly picking a link, and then randomly selecting one of the endpoint nodes of the chosen link to be the adopter and the other, the influencer. For $p = 0$, T_c follows a power law with an exponent ≈ 1 (linear) and for any $p > 0$, $T_c \sim \log(N)$ (see inset).

CHAPTER 4
Structural balance and influence

So far we have restricted ourselves to the study of models with dyadic interactions. In this chapter we study a model which takes into consideration the effect of social balance and in which the interactions are triadic. In a socially interacting population, relationships among individuals (*links* in the underlying social network) can be classified as *friendly* (+) or *unfriendly* (-). Evolution of these links is governed by the theory of *structural balance*. Structural balance is one of the key driving mechanisms of social dynamics and since its introduction by Heider [5], it has been studied extensively in the context of social networks [68–70, 95]. The underlying axioms behind the theory of social balance are: (i) a friend of my friend and an enemy of my enemy is my friend, and (ii) a friend of my enemy and an enemy of my friend is my enemy. In the context of social networks, a triangle is said to be unbalanced if it contains an odd number of unfriendly links [68, 70]. According to the theory of social balance, these unbalanced triangles have a tendency to achieve balanced configurations [5, 69]. An unbalanced triad can make a transition to a balanced one by reorganizing its interpersonal links such that it satisfies the conditions of social balance. For example, a triad where two mutually antagonistic individuals have a common friend, is by definition unbalanced and can be balanced by requiring either the mutual friend to pick a side or the other two to reconcile the conflict and become friends. A structurally balanced network contains no unbalanced triangles.

Here, we study a model which couples the states of the links (friendly or unfriendly) to the opinions adopted by the nodes. Specifically, we consider three opinion states - leftist, rightist, and centrist and posit that a link between two extremists of opposite kind (namely leftist and rightist) is unfriendly while other links are friendly which gives rise to certain triadic configurations of opinions that are structurally unbalanced. Models of three opinion states have been studied in the past, however, not in the context of social balance [72, 73, 96].

Portions of this chapter to appear in: P. Singh, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Competing effects of social balance and influence," (*in preparation*).

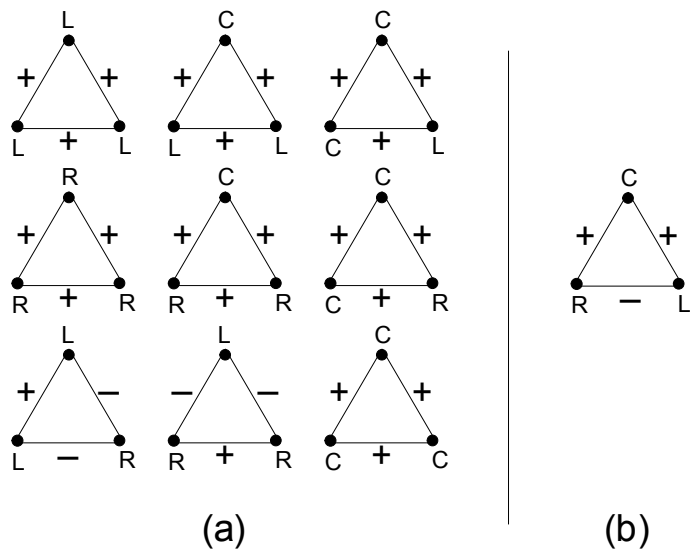

Figure 4.1: (a) Configurations of balanced triangles. (b) An unbalanced triangle.

4.1 The model

In the present model, we consider a population where each individual is in one of the three possible opinion states (*leftist*, *rightist*, and *centrist*) [72, 73]. A link that connects two *extremists* of opposite type (i.e. the link between a leftist and a rightist) is considered to be unfriendly while all other links are friendly. Thus, a triangle containing one node of each type (leftist, rightist, and centrist) is unbalanced. Fig. 4.1 shows possible balanced and unbalanced triangles (without permutations).

An unbalanced triangle can be balanced in a number of ways, each requiring a node in the triangle to update its opinion. In a model which consists of extremist and moderate opinion states of individuals, Marvel *et al.* [97] showed that moderation by external stimulus is a way to have a society adopt a moderate viewpoint. Here in our model, we consider a similar external influence field (e. g. campaigns, advertisements) which converts extremists into centrists.

The model dynamics proceeds as follows: At each time step (i) either a random node is picked with probability p and is converted to a centrist, (ii) or with the complimentary probability $(1-p)$ a random triangle is picked and if unbalanced, is balanced by either converting (with a probability α) a centrist into an extremist or (with a probability $(1-\alpha)$) an extremist into a centrist. Furthermore, since an

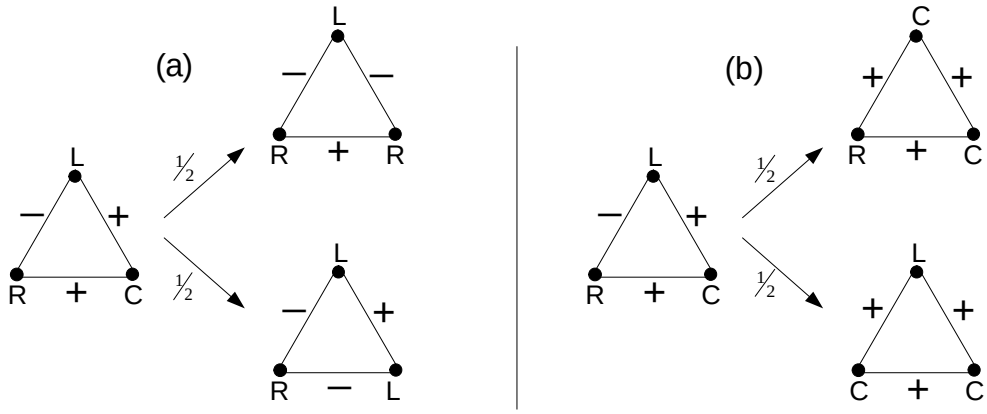

Figure 4.2: The balance step is taken with probability $(1-p)$. However, balance can be achieved in two ways: (a) with probability α a centrist is converted into an extremist and since it can either be a leftist or a rightist, one is chosen randomly with equal probability $1/2$. (b) Or with probability $(1-\alpha)$ converting an extremist (either leftist or rightist with equal probability $1/2$) to a centrist.

extremist can either be a leftist or a rightist, a choice is made with equal probability ($\frac{1}{2}$) as shown in Fig. 4.2.

4.2 Steady-state solution

At any given time the state of a system of size N can be described by two numbers - the number of leftists (X) and the number of rightists (Y) - as we can eliminate the number of centrists (Z) since $X+Y+Z=N$, and thus the evolution can be mapped onto a xy plane. A finite system will always have only one absorbing fixed point that is a consensus state where every node has adopted the centrist opinion. First we consider this dynamics on a complete graph where every node is connected to every other node. We denote the densities of leftist, rightist, and centrists by $x = X/N$, $y = Y/N$, and $z = Z/N$ respectively. The evolution of these densities (under mean-field assumption) is then governed by the following rate equations:

$$\frac{dx}{dt} = -px + 3(2\alpha-1)(1-p)xy(1-x-y) \tag{4.1}$$

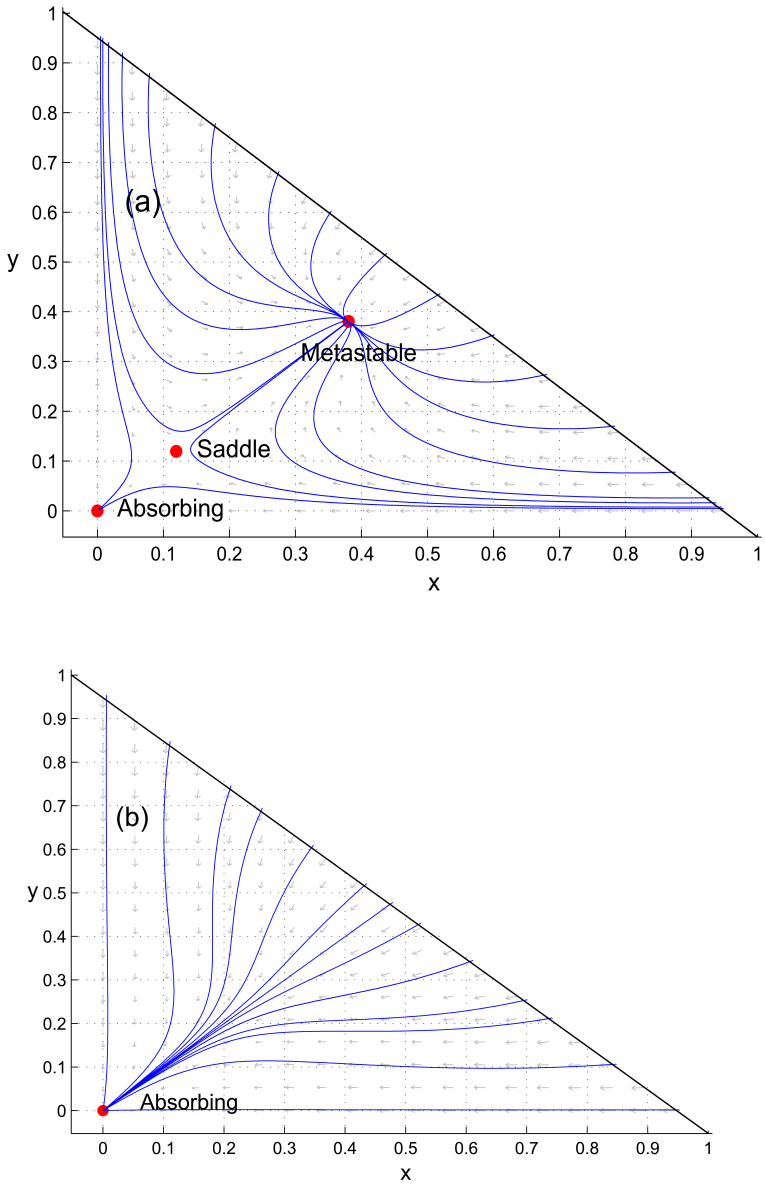

Figure 4.3: The phase diagram of the system for $\alpha = 0.75$ (a) at $p = 0.12 < p_c$, and (b) at $p = 0.20 > p_c$. For $\alpha = 0.75$, $p_c \approx 0.16$.

$$\frac{dy}{dt} = -py + 3(2\alpha - 1)(1 - p)xy(1 - x - y) \tag{4.2}$$

A trivial solution of these equations is the centrist consensus state–$x = 0$, $y = 0$ (or equivalently $z = 1$). However, the steady state solution (see Appendix B.1) of these equations, for $\alpha > \frac{1}{2}$, shows the existence of a critical point

$$p_c = \frac{3(2\alpha - 1)}{8 + 3(2\alpha - 1)} \tag{4.3}$$

such that for $p < p_c$, the system exhibits two more fixed points:

$$(x, y) = (\frac{1}{4} + \frac{1}{4}\sqrt{1 - \frac{8p}{3(2\alpha - 1)(1 - p)}}, \frac{1}{4} + \frac{1}{4}\sqrt{1 - \frac{8p}{3(2\alpha - 1)(1 - p)}}) \tag{4.4}$$

which is a metastable fixed point and

$$(x, y) = (\frac{1}{4} - \frac{1}{4}\sqrt{1 - \frac{8p}{3(2\alpha - 1)(1 - p)}}, \frac{1}{4} - \frac{1}{4}\sqrt{1 - \frac{8p}{3(2\alpha - 1)(1 - p)}}) \tag{4.5}$$

which happens to be a saddle point (unstable). Due to the competition between balancing and influencing forces, the densities fluctuate around the metastable point and the system is trapped for exponentially long times before eventually reaching the absorbing state by a large fluctuation. Here all fixed points lie on the line $y = x$ as shown in Fig. 4.3, as any asymmetry in x and y decays exponentially fast [98, 99](see Appendix B.1). The trajectories showed in Fig. 4.3 are exact only in the thermodynamic limit.

4.3 Consensus time

An all centrist (absorbing) consensus state is always reached for a finite network. Time to reach this absorbing state (consensus time T_c) can be obtained by direct simulation. This approach works well for $p > p_c$, however, for $p < p_c$ (specially when $p << p_c$ and/or $N >> 1$), T_c becomes so large that its estimation by simulation becomes difficult. We therefore use the Quasi-stationary (QS) approximation prescribed in [100] and also used in [10, 101] to estimate T_c in the region $p < p_c$.

For a fully connected network, the density of unbalanced triangles at any given time is $\frac{6\ X\ Y\ (N-X-Y)}{N(N-1)(N-2)}$. We start from the master equation that describes the time evolution of probability $P_{X,Y}$ (the probability that system has X leftists and Y rightists at time t).

$$
\begin{aligned}
\frac{1}{N}\frac{dP_{X,Y}}{dt} &= P_{X-1,Y}\frac{3\alpha(1-p)(X-1)Y(N-X+1-Y)}{N(N-1)(N-2)} \\
&+ P_{X,Y-1}\frac{3\alpha(1-p)X(Y-1)(N-X-Y+1)}{N(N-1)(N-2)} \\
&+ P_{X+1,Y}\frac{3(1-\alpha)(1-p)(X+1)Y(N-X-1-Y)}{N(N-1)(N-2)} \\
&+ P_{X+1,Y}\ p\ \frac{X+1}{N} \\
&+ P_{X,Y+1}\frac{3(1-\alpha)(1-p)X(Y+1)(N-X-Y-1)}{N(N-1)(N-2)} \\
&+ P_{X,Y+1}\ p\ \frac{Y+1}{N} \\
&- P_{X,Y}\frac{6(1-p)XY(N-X-Y)}{N(N-1)(N-2)} \\
&- P_{X,Y}\ p\ \frac{X+Y}{N} \qquad (4.6)
\end{aligned}
$$

Within the triangular region (bounded by $0 \leq X \leq (N-Y)$ and $0 \leq Y \leq (N-X)$), the allowed transitions from a state $(X,\, Y)$ are to states $(X \pm 1,\, Y)$ or $(X,\, Y \pm 1)$ with the constraint that the system stays with the bounded region. The positive and the negative terms on the right side of the master equation respectively contribute to the net flow of probability into and out of the state (X,Y). A factor of $\frac{1}{N}$ on the left-hand side appears because a microscopic step or transition from initial state to the final state takes place in a time interval $1/N$.

The QS distribution of occupation probabilities is given by $\tilde{P}_{X,Y} = P_{X,Y}(t)/P_S(t)$ where $P_S(t)$ is the survival probability. Under the QS hypothesis, the survival probability decays exponentially in time as in Eq. (4.7).

$$\frac{dP_S(t)}{dt} = -P_S(t)\tilde{Q}_0. \qquad (4.7)$$

where

$$\tilde{Q}_0 = p[\tilde{P}_{1,0} + \tilde{P}_{0,1}] \qquad (4.8)$$

which measures the flow of probability into the absorbing state $(0,0)$. We plug in $P_{X,Y}$ in terms of $\tilde{P}_{X,Y}$ into the master equation [10] to obtain the QS distribution

$$\tilde{P}_{X,Y} = \frac{\tilde{Q}_{X,Y}}{W_{X,Y} - \tilde{Q}_0} \tag{4.9}$$

where

$$W_{X,Y} = \frac{6(1-p)XY(N-X-Y)}{(N-1)(N-2)} + p(X+Y),$$

and

$$\begin{aligned}
\tilde{Q}_{X,Y} &= \tilde{P}_{X-1,Y}\frac{3\alpha(1-p)(X-1)Y(N-X+1-Y)}{(N-1)(N-2)} \\
&+ \tilde{P}_{X,Y-1}\frac{3\alpha(1-p)X(Y-1)(N-X-Y+1)}{(N-1)(N-2)} \\
&+ \tilde{P}_{X+1,Y}\frac{3(1-\alpha)(1-p)(X+1)Y(N-X-1-Y)}{(N-1)(N-2)} \\
&+ \tilde{P}_{X+1,Y}\ p\ (X+1) \\
&+ \tilde{P}_{X,Y+1}\frac{3(1-\alpha)(1-p)X(Y+1)(N-X-Y-1)}{(N-1)(N-2)} \\
&+ \tilde{P}_{X,Y+1}\ p\ (Y+1)
\end{aligned}$$

Starting from an arbitrary distribution $\tilde{P}^0_{X,Y}$, an asymptotic QS distribution $\tilde{P}_{X,Y}$ can be obtained by the iteration:

$$\tilde{P}^{i+1}_{X,Y} = a\ \tilde{P}^i_{X,Y} + (1-a)\ \frac{\tilde{Q}^i_{X,Y}}{W^i_{X,Y} - \tilde{Q}^i_0} \tag{4.10}$$

where $0 \leq a \leq 1$ is an arbitrary parameter. For the results discussed here, a satisfactory convergence was obtained in 40000 iterations for $a = 0.5$. The QS distribution for a particular system size $N = 100$ (fully-connected) and $\alpha = 0.75$ is shown in Fig 4.4. As expected from the mean-field analysis (for $\alpha = 0.75$, the metastable fixed point $(x, y) = (0.38, 0.38)$) the distribution peaks around $(X, Y) = (38, 38)$.

Once the desired QS distribution is obtained, the mean consensus time T_c is simply computed from the decay rate of the survival probability:

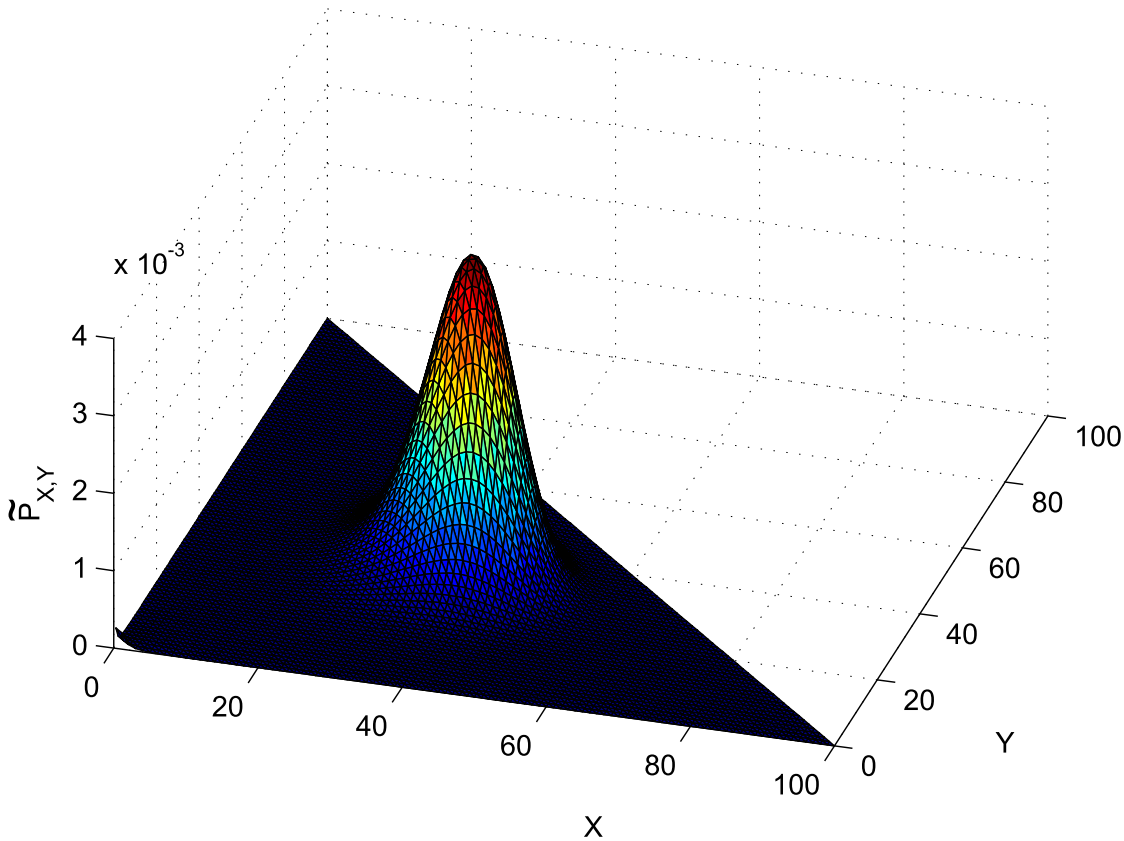

Figure 4.4: The quasi-stationary distribution for $N = 100$, $\alpha = 0.75$, $p = 0.12$. For these parameters, $p_c \approx 0.16$.

$$T_c \sim \frac{1}{p\,[\tilde{P_{1,0}} + \tilde{P}_{0,1}]} \tag{4.11}$$

We compare the T_c obtained from the QS approximation and that by direct simulation in the region $p < p_c$ for a small range of p where T_c was easily obtained (Fig 4.5 and Fig 4.6). It can be seen that there is a good agreement between the two across many system sizes and the agreement gets better as p is decreased below p_c. Fig 4.7 shows T_c as a function of p for a much longer range $p << p_c$, where it becomes almost impossible to estimate T_c through simulations.

The scaling behavior of T_c with N is also quite different in the weak ($p < p_c$) and strong field ($p > p_c$) regimes. As shown in Fig 4.8, T_c (obtained by QS approximation) scales exponentially with N for $p < p_c$, whereas for $p < p_c$, T_c

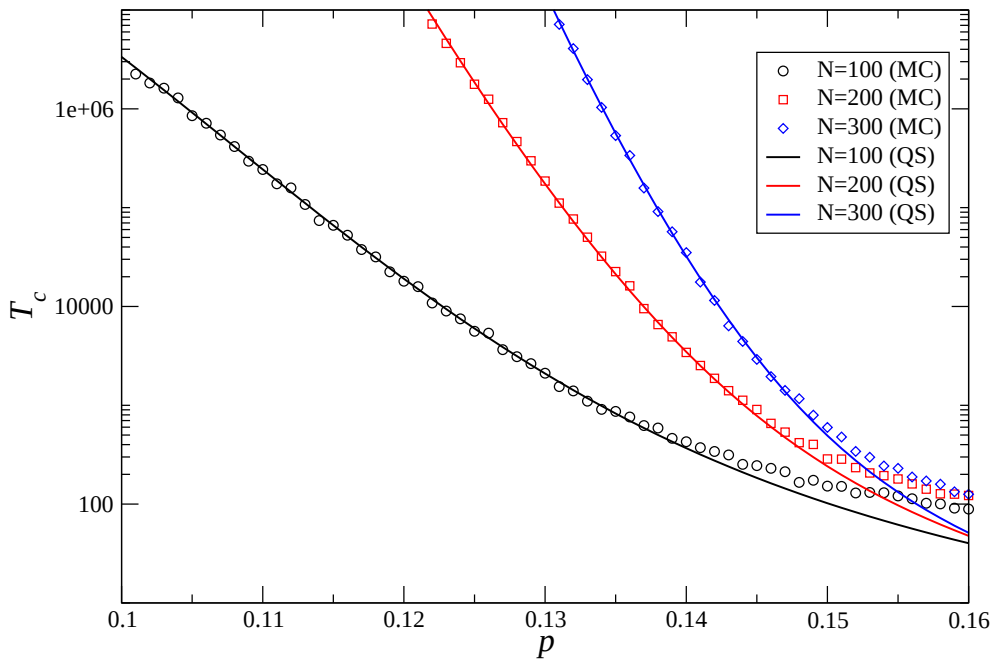

Figure 4.5: Consensus time T_c obtained from the quasi-stationary (QS) distribution and by direct monte carlo (MC) simulations for $\alpha = 0.75$

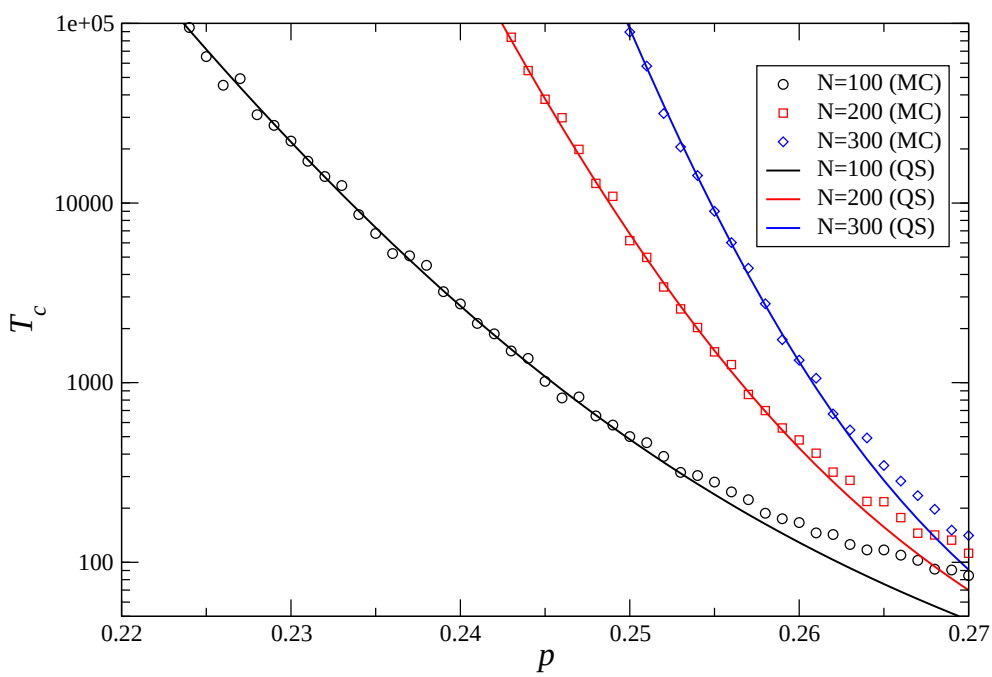

Figure 4.6: Consensus time T_c obtained from the quasi-stationary (QS) distribution and by direct monte carlo (MC) simulations for $\alpha = 1.0$.

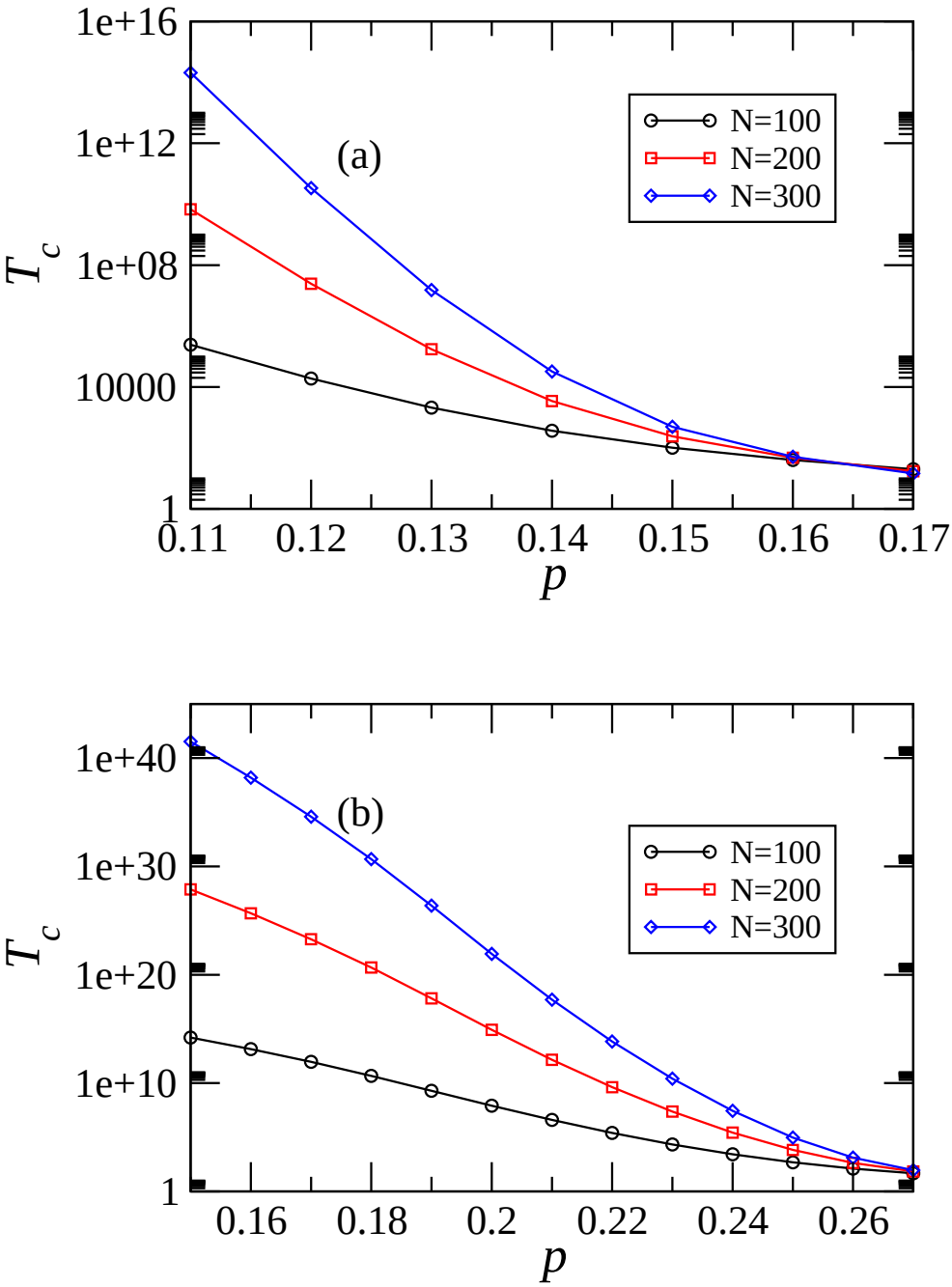

Figure 4.7: Consensus time T_c (computed by QS approximation) as a function of p in the weak field regime, (a) for $\alpha = 0.75$, and (b) for $\alpha = 1.0$.

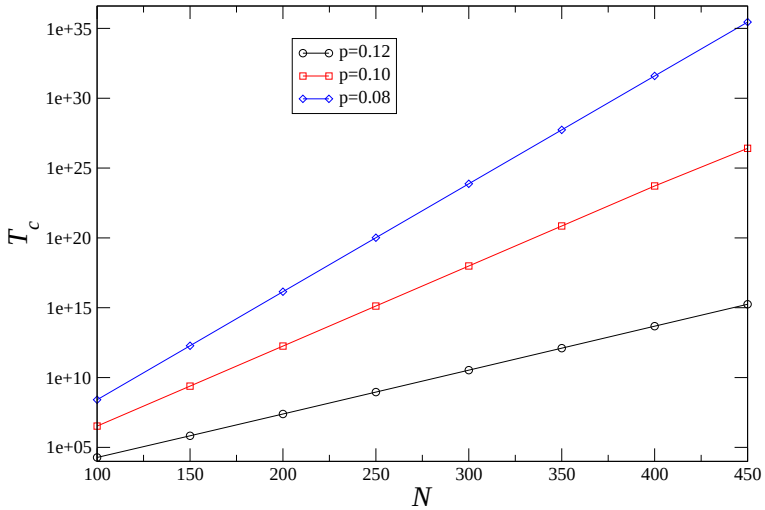

Figure 4.8: Consensus time T_c (computed by QS approximation) as a function of N for $\alpha = 0.75$

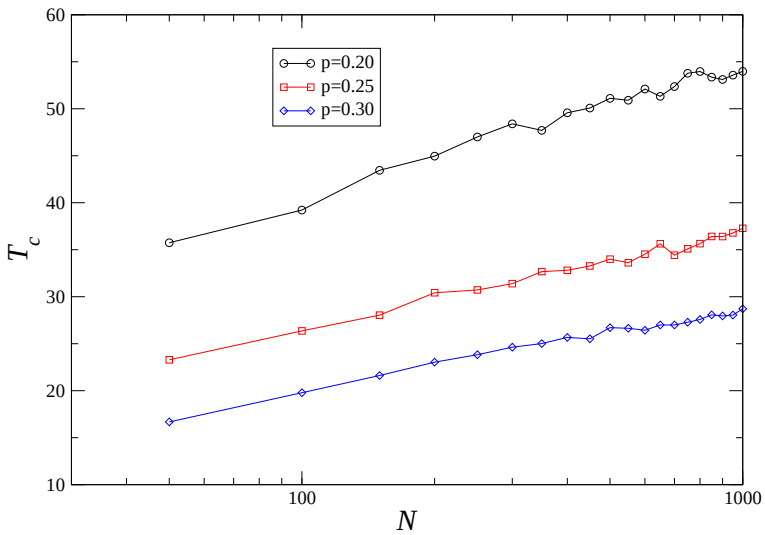

Figure 4.9: Consensus time T_c (computed by simulations) as a function of N for $\alpha = 0.75$. Initially $x = y = 0.5$.

becomes logarithmic in N (Fig 4.9).

4.4 Conclusions

In this chapter, we studied the effect of an external deradicalizing field on the dynamics of social balance in a system where the initial population comprises of three species (two radical groups in addition to a centrist population). We once again obtain a *tipping point* p_c, for the strength of the external field p such that for

$p < p_c$ system gets trapped in a metastable state while for $p > p_c$ the metastability disappears and a quick consensus is reached. To estimate the consensus time for $p < p_c$, the quasi-stationary (QS) approximation method is used and found to agree with simulation results in that regime.

CHAPTER 5
Summary

In this thesis we have investigated three distinct models of opinion formation in social networks. Each of these models considers the effect of different interaction mechanism for the constituent nodes. The evolution of the system depends on the interaction mechanism at play. Typically, a consensus (where every node or a vast majority of the nodes in the network possess the same opinion) formation process, starting from an arbitrary initial composition, is studied. The particular issues concerning opinion dynamics such as the scaling of consensus time, spread size of a particular opinion etc. are addressed. An interesting finding of this work, despite the fact that the three models are quite distinct from one another at the microscopic level, is that there exists a *tipping point* of some form that plays a crucial role in the attainment of the consensus state. Which, consequently, has an implication in realistic scenarios, where changing the model parameter characterized by a tipping point has a cost associated with it, that a minimum investment is necessary in order to reach the desired consensus state. A summary of our findings is given below.

In Chapter 2, we explored the *threshold model* in the case of multiple *initiators* where a new opinion spreads starting from initiators (source nodes) in the network and nodes adopt the new opinion according to the threshold rule as discussed. In this study we found that there exists a critical value of this initiator fraction needed in order to trigger global cascades of opinion. Given a fixed number, selecting these initiator nodes in the network to maximize the spread, remains an open problem. However, a few heuristic selection strategies of these initiators were explored and we found that selecting the nodes by the descending order of their degree works best. Through simulations on an empirical network and a few degree-sequence preserving randomized versions, effects of community structure and clustering were studied. The results indicate that community structure (clustering) is helpful in spreading the new opinion. While, for simplicity, we considered the case where every node has the same adoption threshold, it is interesting to explore the case where individuals

have a wide range of adoption threshold. Some research work aimed to understand the effect of various threshold distributions is already in progress [102].

In Chapter 3, we studied a *homophily* driven model (similar to the Axelrod model) on a structurally coevolving network. The states of the nodes in this model are described as vectors whose components correspond to independent attributes. One of these attributes is the opinion *designated attribute* of the individual node on a particular issue. *Similarity* between two nodes is quantified by the number of matching attributes. Nodes influence one another (and become more similar) if they are compatible (satisfy a similarity threshold) and incompatible nodes rewire their links. The interplay between influence and structural evolution results in a complex feedback process and exponentially long consensus times. We found that this exponential divergence in consensus time can be overcome by introducing *committed nodes*–nodes which do not change their opinion. However, the number of committed nodes should be greater than a *critical* value which defines the tipping point for this model. The scaling of consensus time with system size changes from exponential to logarithmic when the requirement of the critical number of committed nodes is fulfilled. Change in the dynamical rule (selection order of influencer and adopter) results in a different scaling of consensus time. However, even in this case, the advantage of adding committed nodes improves the scaling behavior of the consensus time in a qualitatively similar manner.

We discussed a model in Chapter 4, which studies social balance, in a three opinion-state system (system is composed of a group possessing moderate opinion and two radical groups of opposite type), in the presence of an external moderation field. The non-local moderation field in social networks can more realistically be thought of as external stimulus through media, press, advertisements, campaigns etc. The model illustrates in the mean-field limit, how system can get trapped in a metastable state characterized by the coexistence of all three species, if the external field is weaker than a critical value (tipping point). However, a consensus is attained quickly if the external field is strong enough.

It is worth noting that although the contrast shown between the consensus times above and below the tipping point is qualitatively the same, the origin of

the metastability in the system is different in Chapter 3 and 4. For the model described in Chapter 3, metastability arises due to the interplay between influence in the network and structural evolution of the network itself whereas in latter case, competing effects of external influence and the balance dynamics results in a competition among the constituent species and thus giving rise to metastability in the system.

We developed simple network models that take into account a variety of different interaction rules. However, in future, with the surge in the availability of real world network data it will be worthwhile to test and validate these models. Specifically, time-evolving network data may be very useful in improving our understanding of the evolution of opinions in social networks.

REFERENCES

[1] S. Galam, "Sociophysics: A review of Galam Models," Int. J. Mod. Phys. C **19,** 409 (2008).

[2] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," Rev. Mod. Phys. **81,** 591 (2009).

[3] P. F. Lazarsfeld and R. K. Merton, *Freedom and Control in Modern Society* (Van Nostrand, New York, 1954), Vol. 18, pp. 18–66.

[4] D. Garlaschelli and M. I. Loffredo, "Patterns of Link Reciprocity in Directed Networks," Phys. Rev. Lett. **93,** 268701 (2004).

[5] F. Heider, "Attitudes and Cognitive Organization," J. Psychol. **21,** 107 (1946).

[6] P. Clifford and A. Sudbury, "A model for spatial conflict," Biometrika **60,** 581 (1973).

[7] R. A. Holley and T. M. Liggett, "Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model," Ann. Probab. **3,** 643 (1975).

[8] C. Nardini, B. Kozma, and A. Barrat, "Whos Talking First? Consensus or Lack Thereof in Coevolving Opinion Formation Models," Phys. Rev. Lett. **100,** 158701 (2008).

[9] A. Baronchelli *et al.*, "Sharp transition towards shared vocabularies in multi-agent systems," J. Stat. Mech.: Theory Exp. **2006,** P06014 (2006).

[10] J. Xie *et al.*, "Social consensus through the influence of committed minorities," Phys. Rev. E **84,** 011130 (2011).

[11] Q. Lu, G. Korniss, and B. K. Szymanski, "The naming game in social networks: community formation and consensus engineering," J. Econ. Interact. Coord. **4,** 221 (2009).

[12] X. Castello, A. Baronchelli, and V. Loreto, "Consensus and ordering in language dynamics," Eur. Phys. J. B **71,** 557 (2009).

[13] S. Fortunato, "Universality of the threshold for complete consensus for the opinion dynamics of deffuant et al.," Int. J. Mod. Phys. C **15,** 1301 (2004).

[14] M. Granovetter, "Threshold models of collective behavior," Am. J. Sociol. **83,** 1420 (1978).

[15] B. Latané and T. L'Herrou, "Spatial clustering in the conformity game: dynamic social impact electronic groups," J. Pers. Soc. Psychol. **70,** 1218 (1996).

[16] D. J. Watts, "A simple model of global cascades on random networks," Proc. Natl. Acad. Sci. USA **99,** 5766 (2002).

[17] Q. Lu, G. Korniss, and B. K. Szymanski, "Threshold-Controlled Global Cascading in Wireless Sensor Networks," in *Proceedings of the Third International Conference on Networked Sensing Systems (INSS 2006)* (Transducer Research Foundation, San Diego, CA, 2006) pp. 164–171; http://arxiv.org/abs/cs.NI/0606054 (Accessed June 20, 2013).

[18] S. J. Handjani, "Survival of threshold Contact Processes," J. Theor. Probab. **10,** 737 (1997).

[19] P. S. Dodds and D. J. Watts, "Universal Behavior in a Generalized Model of Contagion," Phys. Rev. Lett. **92,** 218701 (2004).

[20] L. R. Fontes and R. H. Schonmann, "Threshold $\theta \geq 2$ contact processes on homogeneous trees," Probab. Theor. and Related Fields **141,** 513 (2008).

[21] T. Mountford, and R. H. Schonmann, "The survival of large dimensional threshold contact processes," Ann. Probab. **37,** 1483 (2009).

[22] T. E. Harris, "Contact interactions on a lattice," Ann. Probab. **2,** 969 (1974).

[23] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes* (Springer, New York, 1999).

[24] R. Durrett, "Some features of the spread of epidemics and information on a random graph," Proc. Natl. Acad. Sci. USA **107,** 4491 (2010).

[25] F. M. Bass, "A New Product Growth for Model Consumer Durables," Manage. Sci. **15,** 215 (1969).

[26] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic Thresholds in Real Networks," ACM Transactions on Information and System Security **10**, 13 (2008).

[27] B. A. Prakash, A. Beutel, R. Rosenfeld, and C. Faloutsos, "Winner Takes All: Competing Viruses or Ideas on fair-play Networks," in *Proceedings of the 21st International Conference on World Wide Web* (ACM New York, NY, 2012), pp. 1037–1046.

[28] D. J. Watts and P. S. Dodds, "Influentials, networks, and public opinion formation," J. Consum. Res. **34,** 441 (2007).

[29] P. Singh, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Threshold-limited spreading in social networks with multiple initiators," Sci. Rep. **3,** 2330 (2013).

[30] D. Centola, V. Eguiluz, and M. W. Macy, "Cascade dynamics of complex propagation," Physica A **374,** 449 (2007).

[31] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM New York, NY, 2003), pp. 137–146.

[32] W. Chen, Y. Yuan, and L. Zhang, "Scalable Influence Maximization in Social Networks under the Linear Threshold Model," in *Proceedings of the 2010 IEEE International Conference on Data Mining* (IEEE Computer Society, Washington, DC, 2010), pp. 88–97.

[33] Y. Ikeda, T. Hasegawa, and K. Nemoto, "Cascade dynamics on clustered network," J. Phys.: Conf. Ser. **221,** 012005 (2010).

[34] L. Huang, Y.-C. Lai, and G. Chen, "Understanding and preventing cascading breakdown in complex clustered networks," Phys. Rev. E **78,** 036116 (2008).

[35] C. D. Brummitt, K.-M. Lee, and K.-I. Goh, "Multiplexity-facilitated cascades in networks," Phys. Rev. E **85,** 045102(R) (2012).

[36] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," Nature **464,** 1025 (2010).

[37] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "Networks formed from interdependent networks," Nat. Phys. **8,** 40 (2012).

[38] C. D. Brummitt, R. M. D'Souza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," Proc. Natl. Acad. Sci. USA **109,** E680 (2012).

[39] We use the network-structure data sets from *Add Health*, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. For data files contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524, (addhealth@unc.edu, http://www.cpc.unc.edu/projects/addhealth/ (Accessed June 20, 2013).

[40] M. McPherson, L. S. Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," Ann. Rev. Sociol. **27,** 415 (2001).

[41] R. Axelrod, "The Dissemination of Culture," J. Conflict Resol. **41,** 203 (1997).

[42] D. H. Zanette and S. Gil, "Opinion spreading and agent segregation on evolving networks," Physica D **224,** 156 (2006).

[43] S. Gil and D. H. Zanette, "Coevolution of agents and networks: Opinion spreading and community disconnection," Phys. Lett. A **356,** 89 (2006).

[44] F. Vazquez, V. M. Eguíluz, and M. San Miguel, "Generic Absorbing Transition in Coevolution Dynamics," Phys. Rev. Lett. **100,** 108702 (2008).

[45] D. Kimura and Y. Hayakawa, "Coevolutionary networks with homophily and heterophily," Phys. Rev. E **78,** 016103 (2008).

[46] L.-X. Zhong, F. Ren, T. Qiu, J.-R. Xu, B.-H. Chen, and C.-F. Liu, "Effects of attachment preferences on coevolution of opinions and networks," Physica A **389,** 2557 (2010).

[47] P. Holme and M. E. J. Newman, "Nonequilibrium phase transition in the coevolution of networks and opinions," Phys. Rev. E **74,** 056108 (2006).

[48] I. J. Benczik, S. Z. Benczik, B. Schmittmann, and R. K. P. Zia, "Lack of consensus in social systems," Europhys. Lett. **82,** 48006 (2008).

[49] I. J. Benczik, S. Z. Benczik, B. Schmittmann, and R. K. P. Zia, "Opinion dynamics on an adaptive random network," Phys. Rev. E **79,** 046104 (2009).

[50] F. Vazquez, J. C. González-Avella, V. M. Eguíluz, and M. San Miguel, "Time-scale competition leading to fragmentation and recombination transitions in the coevolution of network and states," Phys. Rev. E **76,** 046120 (2007).

[51] D. Centola, J. C. González-Avella, V. M. Eguíluz, and M. San Miguel, "Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups," J. Conflict Resol. **51,** 905 (2007).

[52] W. Zhang, C. Lim, and B. K. Szymanski, "Analytic treatment of tipping points for social consensus in large random networks," Phys. Rev. E **86,** 061134 (2012).

[53] A. Halu, K. Zhao, A. Baronchelli, and G. Bianconi, "Connect and win: The role of social networks in political elections," Eurphys. Lett. **102,** 16002 (2013).

[54] M. Turalska, B. J. West, and P. Grigolini, "Role of committed minorities in times of crisis," Sci. Rep. **3,** 1371 (2013).

[55] E. Yildiz, D. Acemoglu, A. Ozdaglar, A. Saberi, and A. Scaglione, "Discrete opinion dynamics with stubborn agents," http://web.mit.edu/asuman/www/documents/voter-submit.pdf (Accessed June 20, 2013).

[56] D. Centola, R. Willer, and M. Macy, "The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms," Am. J. Sociol. **110,** 1009 (2005).

[57] M. Mobilia, "Does a Single Zealot Affect an Infinite Group of Voters?," Phys. Rev. Lett. **91,** 028701 (2003).

[58] M. Mobilia, A. Petersen, and S. Redner, "On the role of zealotry in the voter model," J. Stat. Mech.: Theory Exp. **2007,** P08029 (2007).

[59] S. Biswas and P. Sen, "Model of binary opinion dynamics: Coarsening and effect of disorder," Phys. Rev. E **80,** 027101 (2009).

[60] G. Verma, A. Swami, and K. Chan, "The Effect of Zealotry in the Naming Game Model of Opinion Dynamics," in *Proceedings of MILCOM 2012* (IEEE, 2012), pp. 1–6.

[61] Q. Li, L. A. Braunstein, S. Havlin, and H. E. Stanley, "Strategy of competition between two groups based on an inflexible contrarian opinion model," Phys. Rev. E **84,** 066101 (2011).

[62] Q. Li *et al.*, "Non-consensus Opinion Models on Complex Networks," J. Stat. Phys. **151,** 92 (2013).

[63] S. Moscovici, E. Lage, and M. Naffrechoux, "Influence of a Consistent Minority on the Responses of a Majority in a Color Perception Task," Sociometry **32,** 365 (1969).

[64] J. Xie *et al.*, "Evolution of opinions on social networks in the presence of competing committed groups," PLoS ONE **7,** e33215 (2012).

[65] 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 **21,** 025115 (2011).

[66] S. Galam and F. Jacobs, "The role of inflexible minorities in the breaking of democratic opinion dynamics," Physica A **381,** 366 (2007).

[67] P. Singh, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "Accelerating consensus on coevolving networks: The effect of committed individuals," Phys. Rev. E **85,** 046104 (2012).

[68] T. Antal, P. L. Krapivsky, and S. Redner, "Dynamics of social balance on networks," Phys. Rev. E **72,** 036121 (2005).

[69] D. Cartwright and F. Harary, "Structural balance: a generalization of Heider's theory," Psychol. Rev. **63,** 277 (1956).

[70] T. Antal, P. L. Krapivsky, and S. Redner, "Social Balance on Networks: The Dynamics of Friendship and Enmity," Physica D **224,** 130 (2006).

[71] Michael Szell, Renaud Lambiotte, and Stefan Thurner, "Multirelational organization of large-scale social networks in an online world," Proc. Natl. Acad. Sci. USA **107**, 13636 (2010).

[72] M. Mobilia, "Commitment Versus Persuasion in the Three-Party Constrained Voter Model," J. Stat. Phys. **151,** 69 (2013).

[73] F. Vazquez and S. Redner, "Ultimate fate of constrained voters," J. Phys. A: Math. Gen. **37,** 8479 (2004).

[74] S. Soderstrom, B. Uzzi, J. Fowler, and D. Diermeier, "Prerelease and Postrelease Word of Mouth and the Propensity to Adopt New Products," (unpublished).

[75] Y. Hu, S. Havlin, and H. A. Makse, "The rise and fall of social communities: Cascades of followers triggered by innovators," (unpublished).

[76] J. P. Gleeson and D. J. Cahalane, "Seed size strongly affects cascades on random networks," Phys. Rev. E **75,** 056103 (2007).

[77] P. Erdős and A. Rényi, "On the evolution of random graphs," Publ. Math. Inst. Hung. Acad. Sci. **5,** 17 (1960).

[78] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," Nature **393,** 440 (1998).

[79] M. Kitsak *et al.*, "Identification of influential spreaders in complex networks," Nat. Phys. **6,** 888 (2010).

[80] D. Centola, "The Spread of Behavior in an Online Social Network Experiment," Science **329,** 1194 (2010).

[81] S. Hanhijärvi, G. Garriga, and K. Puolamäki, "Randomization techniques for graphs," in *Proceedings of the 9th SIAM International Conference on Data Mining* (SIAM, Philadelphia, PA, 2009), pp. 780–791.

[82] A. Gionis, H. Mannila, T. Mielikäinen, and P. Tsaparas, "Assessing data mining results via swap randomization," ACM Trans. Knowl. Discov. Data **1,** 14 (2007).

[83] R. Sharan, T. Ideker, B. P. Kelley, R. Shamir, and R. M. Karp, "Identification of protein complexes by comparative analysis of yeast and bacterial protein interaction data," J. Comp. Biol. **12,** 835 (2005).

[84] C. I. Del Genio, H. Kim, Z. Toroczkai, and K. E. Bassler, "Efficient and exact sampling of simple graphs with given arbitrary degree sequence," PLoS ONE **5,** e10012 (2010).

[85] H. Kim, Z. Toroczkai, P. L. Erdős, I. Miklós, and L. A. Székely, "Degree-based graph construction," J. Phys. A: Math. Theor. **42,** 392001 (2009).

[86] I. Mergel and T. Langenberg, Harvard-Kennedy School of Government Working Paper No. PNG06-002 (2006).

[87] R. S. Burt, "Decay functions," Social Networks **22,** 1 (2000).

[88] D. Lazer, B. Rubineau, C. Chetkovich, N. Katz, and M. Neblo, "The Coevolution of Networks and Political Attitudes," Polit. Commun. **27,** 248 (2010).

[89] R. I. M. Dunbar, "Neocortex size as a constraint on group size in primates," J. Hum. Evol. **22,** 469 (1992).

[90] P. Hänggi, P. Talkner, and M. Borkovec, "Reaction-rate theory: fifty years after Kramers," Rev. Mod. Phys. **62,** 251 (1990).

[91] D. P. Landau and B. Kurt, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, UK, 2000).

[92] K. Klemm, V. M. Eguíluz, R. Toral, and M. San Miguel, "Nonequilibrium transitions in complex networks: A model of social interaction," Phys. Rev. E **67,** 026120 (2003).

[93] C. Gracia-Lázaro, F. Quijandría, L. Hernández, L. M. Floría, and Y. Moreno, "Coevolutionary network approach to cultural dynamics controlled by intolerance," Phys. Rev. E **84,** 067101 (2011).

[94] Q. Lu, G. Korniss, and B. K. Szymanski, "Naming games in two-dimensional and small-world-connected random geometric networks," Phys. Rev. E **77,** 016111 (2008).

[95] S. A. Marvel, S. H. Strogatz, and J. M. Kleinberg, "Energy Landscape of Social Balance," Phys. Rev. Lett. **103,** 198701 (2009).

[96] W. Zhang, C. Lim, G. Korniss, and B. K. Szymanski, "Opinion Dynamics and Influencing on Random Geometric Graphs," Sci. Rep. **4,** 5568 (2014).

[97] S. A. Marvel, H. Hong, A. Papush, and S. H. Strogatz, "Encouraging Moderation: Clues from a Simple Model of Ideological Conflict," Phys. Rev. Lett. **109,** 118702 (2012).

[98] G. Korniss, B. Schmittmann, and R. K. P. Zia, "Nonequilibrium phase transitions in a simple three-state lattice gas," J. Stat. Phys. **86,** 721 (1997).

[99] H. Haken, *Synergetics, An Introduction* (Springer, Berlin, 1983), 3rd ed.

[100] R. Dickman and R. Vidigal, "Quasi-stationary distributions for stochastic processes with an absorbing state," J. Phys. A: Math. Gen. **35,** 1147 (2002).

[101] M. M. de Oliveira and R. Dickman, "Quasi-stationary distributions for models of heterogeneous catalysis," Physica A **343,** 525 (2004).

[102] P. Karampourniotis, S. Sreenivasan, B.K. Szymanski, and G. Korniss, "Cascades in the Threshold Model with Multiple Initiators and Heterogeneous Threshold Values," http://meetings.aps.org/link/BAPS.2014.MAR.Y16.14 (Accessed June 19, 2014).

[103] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," Rev. Mod. Phys. **74,** 47 (2002).

APPENDIX A
Chapter 2

A.1 Analytic approximation for cascade size

Gleeson and Cahalane [76] obtained an asymptotic expression for the eventual cascade size S

$$S = p + (1-p)\sum_{k=1}^{\infty} p_k \sum_{m=0}^{k} \binom{k}{m} q_\infty^m (1-q_\infty)^{k-m} F(\frac{m}{k}), \tag{A.1}$$

where $p_k = e^{-\langle k\rangle}\langle k\rangle^k/k!$ is the degree distribution of ER networks, and q_∞ is the fixed point of the following recursion relation (for $n = 0, 1, 2, ..$ with $q_0 = p$)

$$q_{n+1} = p + (1-p)G(q_n), \tag{A.2}$$

with

$$G(q) = \sum_{k=1}^{\infty} \frac{k}{\langle k\rangle} p_k \sum_{m=0}^{k-1} \binom{k-1}{m} q^m (1-q)^{k-1-m} F(\frac{m}{k}). \tag{A.3}$$

$G(q)$ can also be written as $\sum_{l=0}^{\infty} C_l q^l$ where

$$C_l = \sum_{k=l+1}^{\infty} \sum_{n=0}^{l} \binom{k-1}{l}\binom{l}{n} (-1)^{l+n} \frac{k}{\langle k\rangle} p_k F(\frac{n}{k}). \tag{A.4}$$

For the case of uniform threshold distribution, the function F takes the form of a step function ($F(\frac{m}{k}) = 1$ if $\frac{m}{k} \geq \phi$ and $F(\frac{m}{k}) = 0$ otherwise). In the approximation where only the linear and q^2 terms are considered, the cascade conditions (respectively for the first and second order) are written as [76] (for uniform distribution of thresholds, $C_0 = 0$)

$$(1-p)C_1 > 1, \tag{A.5}$$

and

$$(C_1 - 1)^2 + 2p(C_1 - C_1^2 - 2C_2) < 0. \tag{A.6}$$

To obtain p_c, we systematically increase p from 0. The smallest value of p for which either one of the above equations (5 or 6) is satisfied, gives an estimate of p_c.

A.2 Clustering coefficient

Clustering coefficient C_i for a node in an undirected network is defined as the number of links among its neighbors divided by the total number of possible links among its neighbors [103].

$$C_i = \frac{2E_i}{k_i(k_i - 1)}, \tag{A.7}$$

where E_i is the total number of links among the neighbors of i and k_i is the degree of node i. The average clustering coefficient C for the network is the average of C_i's of all nodes in the network [78].

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{A.8}$$

APPENDIX B
Chapter 4

B.1 Steady-state solution of the rate equations

The rate equations (4.2) for the leftist and rightist densities under the mean-field assumption:

$$\begin{aligned} \frac{dx}{dt} &= -px + 3(2\alpha - 1)(1-p)xy(1-x-y) \\ \frac{dy}{dt} &= -py + 3(2\alpha - 1)(1-p)xy(1-x-y) \end{aligned} \tag{B.1}$$

By adding and subtracting these two equations, and introducing a new set of variables $u = (x + y)$ and $v = (x - y)$, one can immediately see that:

$$\frac{dv}{dt} = -pv \tag{B.2}$$

Which means that, $v \sim exp(-pt)$, or $v \to 0$ exponentially fast. Therefore we can assume that $x \approx y$. Which allows us to analyze the system in terms of single variable equation in u (and $x = y = u/2$):

$$\frac{du}{dt} = -pu + 6(2\alpha - 1)(1-p)\frac{u^2(1-u)}{4} \tag{B.3}$$

$$\tag{B.4}$$

which can be solved for stead-state $\frac{du}{dt} = 0$. A trivial solution of this equation is $u = 0$, which is the absorbing state ($x = 0$, $y = 0$). Additional roots are the solutions of the following quadratic equation:

$$u^2 - u + \frac{2p}{3(1-p)(2\alpha - 1)} = 0 \tag{B.5}$$

and are given by:

$$u = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{8p}{3(1-p)(2\alpha - 1)}} \tag{B.6}$$

These solutions make sense only when $\alpha > 1/2$, otherwise the solution will lie outside the feasible domain ($(x + y) \leq 1$). And for $\alpha > 1/2$, we obtain a critical point

$$p_c = \frac{3(2\alpha - 1)}{8 + 3(2\alpha - 1)} \tag{B.7}$$

such that the roots are real and positive for $p < p_c$. Thus, in terms of x and y the two roots (other than the absorbing state) are:

$$\begin{aligned} x = y &= \frac{1}{4} + \frac{1}{4}\sqrt{1 - \frac{8p}{3(1-p)(2\alpha - 1)}} \qquad \text{(metastable)} \\ x = y &= \frac{1}{4} - \frac{1}{4}\sqrt{1 - \frac{8p}{3(1-p)(2\alpha - 1)}} \qquad \text{(saddle)} \end{aligned} \tag{B.8}$$