ON-OFF THRESHOLD MODELS OF SOCIAL CONTAGION

A Thesis Presented

by

Kameron Decker Harris

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Mathematics

October, 2012

Accepted by the Faculty of the Graduate College, The University of Vermont, in partial fulfillment of the requirements for the degree of Master of Science, specializing in Mathematics.

Thesis Examination Committee:

_____ Advisor
Peter Sheridan Dodds, Ph.D.

_____
Christopher M. Danforth, Ph.D.

_____ Chairperson
Joshua Bongard, Ph.D.

_____ Dean, Graduate College
Domenico Grasso, Ph.D.

August 15, 2012

Abstract

We study binary state contagion dynamics on a social network where nodes act in response to the average state of their neighborhood. We model the competing tendencies of imitation and nonconformity by incorporating an off-threshold into standard threshold models of behavior. In this way, we attempt to capture important aspects of fashions and general societal trends. Allowing varying amounts of stochasticity in both the network and node responses, we find different outcomes in the random and deterministic versions of the model. In the limit of a large, dense network, however, we show that these dynamics coincide. The dynamical behavior of the system ranges from steady state to chaotic depending on network connectivity and update synchronicity. We construct a mean field theory for general random networks. In the undirected case, the mean field theory predicts that the dynamics on the network are a smoothed version of the average node response dynamics. We compare our theory to extensive simulations on Poisson random graphs with node responses that average to the chaotic tent map.

Table of Contents

List of Tables

List of Figures

1 Introduction

Almost universally, people enjoy observing, speculating, and arguing about their fellow humans' behavior, including what they are wearing, the music they listen to, and however else they express themselves. Fashions, trends, and fads intrigue us, whether past, present, or future. Contemporary examples of prevalent trends include skinny jeans, fixed-wheel bicycles, and products made by Apple Inc. In the 1960's, shag carpets and floral wallpaper would take their place. Trends are also prevalent in language: "hipster" was popular in the 1940's and evolved into "hippie", but it has been reappropriated today with different connotations; "groovy" was once used seriously, and now comes out tongue-in-cheek.

The dynamics of these cultural phenomena are fascinating and complex. They reflect numerous factors such as the political climate, social norms, technology, marketing, and history. These phenomena are also influenced by essentially random events. Usually, we can explain the emergence of a fad after the fact, but it is extremely difficult to predict *a priori* whether some behavior will become popular. Furthermore, those behaviors which are adopted by a large fraction of the population can lose their excitement and die out forever or later recur unpredictably.

In this work we describe a mathematical model of the rise and fall of trends. In particular, we model a social contagion process where people are influenced by the behavior of their friends. The agents in the model act according to simple competing tendencies of imitation and non-conformity. One can argue that these two ingredients are essential to all trends; indeed, Simmel, in his classic essay "Fashion" (1957), believed that these are the essential forces behind the creation and destruction of fashions.

Our model is not meant to be quantitative, except perhaps in carefully designed experiments, but it captures the features with which we are familiar: some trends take off and some do not, and some trends are stable while others vary wildly through time. Our model is closely related to the seminal work of Schelling (1971) and Granovetter (1978). Being a mathematical model, it is also connected to theories of percolation (Stauffer and Aharony, 1994), disease spreading (Newman, 2003), and magnetism (Newman, 2003; Aldana et al., 2003).

We focus on the derivation and analysis of dynamical master equations that describe the expected evolution of the system state. Through these equations we are able to predict the behavior of the social contagion process. In some cases, this can be done by hand, but most of the time we resort

to numerical methods for their iteration or solution.

This thesis is structured as follows. In Section 2, we introduce the reader to important background material relevant to the on-off threshold model. In Section 3, we define the model and its deterministic and stochastic variants. In Section 4, we provide an analysis of the model when the underlying network is fixed. Section 5 develops a mean field theory of the model in the most general kind of random graphs. In Section 6, we consider the model on Poisson random graphs with a specific kind of response function. Our analysis is then applied to this specific case, and we compare the results of simulations and theory. Finally, Section 7 presents conclusions and directions for further research.

2 Background

2.1 Basics of graphs and networks

When modeling any dynamical system of many interacting particles or agents, we are often forced to start with a simplified description of their interactions. In a solid, for example, atoms are situated on some sort of lattice and assumed to interact only with their nearest neighbors. However, in many cases the interactions we aim to model do not have the periodic structure of a lattice. Graphs, which are just a set of points connected by lines, are a more general mathematical structure. We denote a graph $\mathscr{G}$ as an ordered pair $\mathscr{G} = (V, E)$ where V is the set of vertices (also called nodes) and $E \subseteq V \times V$ the set of edges (also called links). Here, E is a set of ordered pairs, and we denote an edge from vertices i to j as $ij \in E$. The terms *graph* and *network* will be used interchangeably, but they have acquired different connotations. Hackett (2011, §1.2.1) puts this rather nicely:

> To model a complex system as a graph is to filter out the functional details of each of its components, and the idiosyncrasies of their interactions with each other, and to focus instead on the underlying structure (topology) as an inert mathematical construct. Although this technique is central also to network theory, the word network, in contrast, usually carries with it connotations of the context in which the overarching system exists, particularly when that system displays any sort of nonlinear dynamics. For example, when investigating the spread of infectious disease on a human sexual contact network it makes sense to consider the relevant sociological parameters as well as the abstract topology, and it is in such settings that the interdisciplinary aspect that distinguishes network theory comes to the fore.

In models of human behavior, interactions can be considered to occur on a social network. Each person is connected to those people they interact with. Interactions that constitute a connection

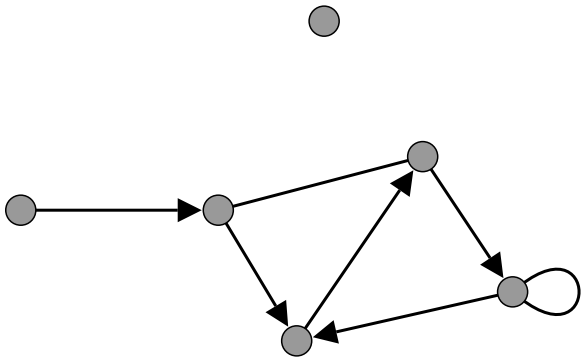

Figure 1: A graph consisting of a connected component of 5 nodes and one isolated node. The majority of the edges are directed. There is one undirected edge, shown without arrows. There is one self-loop which should be interpreted as directed.

between two people, A and B, may be defined in many ways. For example:

Case 1. A contacts B. This could mean:

(a) A sends B an email.

(b) A attends the same concert as B.

Case 2. A is B's superior in a hierarchy.

Case 3. A and B both belong to the same group.

These examples illustrate different types of edges that can be included in a graph. In Case 1a, sending an email message is a one-way communication, so it forms a *directed* edge. We represent directed edges with arrows in a drawing of the corresponding graph. In Case 1b, however, attending a concert is a symmetric relation and there is no reason to give the edge direction. Such edges are called *undirected* or bidirectional. See Figure 1 for an example of a small graph containing multiple types of edges. A simple graph, by definition, contains no directed edges or self loops (edges connecting a vertex to itself). It will be useful to represent the connectivity of the graph with the *adjacency matrix*, $A = (A_{ij})$, where $A_{ij} = 1$ if and only if $ji \in E$ (this "backwards" definition is not standard, but it is useful for the linear algebra to come). The unfamiliar reader is directed to West (2001) for a thorough introduction to graph theory.

When we wish to analyze the structure of a given network, one of the first things to examine is its degree distribution. The *degree* k_i of a vertex $i \in V$ is the number of edges incident to it, regardless of direction. The sequence $k_1, k_2, \ldots, k_N$ is called the degree sequence. In simple graphs,

this is just the size of a vertex's neighborhood. In more complicated graphs we usually speak instead of a vertex i's in-degree $k_i^{(\mathrm{i})}$, out-degree $k_i^{(\mathrm{o})}$, and undirected degree $k_i^{(\mathrm{u})}$, all defined in the obvious way (Dodds et al., 2011). The *degree distribution* p_k is a probability distribution which tells us how edges are distributed among nodes. The average degree $k_{\mathrm{avg}} = \sum_{i=0}^{\infty} kp_k$ characterizes the overall density of edges. A degree distribution which is sharply peaked about its mean indicates a relatively homogeneous network where vertices tend to have the same number of incident edges. In contrast, a skewed distribution could result in some nodes having very high degree while the majority have low degree. This is common in real networks, and the tail of the degree distribution often follows an approximate power law $p_k \sim k^{-\alpha}$ for some exponent α (Newman, 2003).

2.2 Random graphs

Random graphs are a family of models used to represent networks from the real world, although their suitability as such is questionable for reasons that will be elaborated below. Nevertheless, they are well-suited to analysis and as null models.

The simplest random graph is the binomial model introduced by Erdős and Rényi (Bollobás, 2001). Define a probability space $\mathcal{G}(N, p)$ of graphs on N vertices where any of the $\binom{N}{2}$ possible edges are chosen independently with probability p. The expected average degree in the network will be $k_{\mathrm{avg}} = p(N-1)$. We will often make statements about this network in the limit of large system size $N \to \infty$. (In physics, this is referred to as the thermodynamic limit. Many expressions are simplified in performing this approximation, and derived results can be quite accurate given a large enough, albeit finite, N.) The degree distribution is a binomial distribution $p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$ which is well-approximated by the Poisson distribution $p_k \approx (k_{\mathrm{avg}})^k \exp(-k_{\mathrm{avg}})/k!$ with parameter $k_{\mathrm{avg}} = Np$ when N is large and k_{avg} fixed. For this reason, $\mathcal{G}(N, p)$ is also called the *Poisson random graph* model. As noted previously, many networks observed in the real world have heavy-tailed distributions, so the Poisson model is not suitable for those types of networks. Furthermore, real networks often contain a high density of triangles or clustering — friendship tends to be transitive. Poisson random graphs, on the other hand, have zero clustering in the thermodynamic limit.

A more flexible generalization is the configuration model. In this model, either the degree sequence (Molloy and Reed, 1995, 1998) or the expected degree sequence (Chung and Lu, 2002) is given in advance (also see Newman, 2003); the models produce similar networks with subtle differences. The configuration model of can be thought of as a random wiring process as follows.

First, draw a degree sequence $k_1, \ldots, k_N$ independently according to the desired degree distribution p_k. We assign k_i *stubs* (half-edges) to each vertex $i \in V$. Next, choose a pair of stubs at random, connect them, remove both from the queue of stubs, and continue until all stubs are connected. Finally, pairs of edges are chosen and their endpoints shuffled in the manner of Milo et al. (2002) to ensure uniform sampling from all graphs with the given degree sequence. The resulting graph will have the imposed degree sequence, so long as the sum of the degrees is even. There may be self-loops or repeated edges, neither of which are allowable for simple graphs. However, we expect these to occur increasingly rarely for large N, and they can be removed without affecting the resulting graph much. Degree-degree correlations can be introduced by shuffling edges as described in Melnik et al. (2011) and Payne et al. (2011). The configuration model also lacks triangles and higher-order cliques in the thermodynamic limit, although some work has been done to create random graph models with clustering (see Hackett, 2011, §§4-5).

2.3 Dynamical processes on networks

Networks provide a structure on top of which all kinds of dynamical processes may take place. In many cases, the structure of the network itself heavily influences the dynamics. For instance, ferromagnetic materials can be modeled as atoms in a lattice network, where the bulk magnetization is the result of interactions between individual atoms' magnetic moments; strikingly different results occur in different dimensional lattices. Another example is a food web, a network of species connected by trophic interactions (who eats whom). The populations of the species in the ecosystem can be described by dynamical equations that reflect the structure of the network, and this can affect ecosystem stability. These are just two examples where dynamical processes on networks are a reasonable way to model system behavior. Since networks are ubiquitous structures, these models appear across all disciplines (Vespignani, 2012).

2.3.1 Random Boolean networks

Consider Boolean or binary state dynamics, where each node can be either "on" or "off" (in various contexts this can mean active/inactive, infected/susceptible, or spin up/spin down). The state of node i is encoded by a variable $x_i \in \{0, 1\}$, and the system state is $\mathbf{x} = (x_i)$. At each time step, nodes receive input from their neighbors in the (undirected) network. They then compute a function of that input, i.e., $f_i(x_{j_1}, x_{j_2}, \ldots, x_{j_{k_i}})$ where $j_1, \ldots, j_{k_i}$ are the neighbors of node i. This determines their

state in the next time step. Because the state space $\{0,1\}^N$ is finite, all trajectories are eventually periodic. The detailed structure of the cycles depends on the specific details of the network and the Boolean functions.

The system described above, for general $f_i : \{0,1\}^{k_i} \to \{0,1\}$, is known as a Boolean network. Boolean networks were first studied by Kauffman (1969) as a model for dynamical behavior within cells. See the review by Aldana et al. (2003). Most researchers have considered the dynamics on random K-regular graphs with a parameter $\tilde{p}$ that determines the bias between 0s and 1s in the output of the update functions f_i, which are otherwise randomly chosen Boolean functions.

As mentioned, deterministic Boolean network models must be eventually periodic. However, the behavior of the transient and the structure of the basins of attraction are different for different parameters K and $\tilde{p}$. In particular, there is a critical value of the connectivity $K_c(\tilde{p})$ that separates the transient dynamics into two phases (Aldana et al., 2003):

1. Frozen, $K < K_c$: The distance between nearby trajectories $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ decays exponentially with time.

2. Critical, $K = K_c$: The temporal evolution of distance between trajectories is determined by fluctuations.

3. Chaotic, $K > K_c$: The distance between nearby trajectories grows exponentially with time.

2.3.2 Social models

When modeling social systems with a Boolean network, the nodes represent people and their states encode whether or not they participate in a behavior, possess a certain belief, etc. This could be rioting or not rioting (Granovetter, 1978), buying a particular style of tie (Granovetter and Soong, 1986), liking a particular band or style of music, or believing in some unintuitive or controversial idea, e.g. climate change. The state can represent any behavior with only two mutually exclusive possibilities.

The function f_i that determines how node i changes state is called its *response function* in sociological contexts. Schelling (1971, 1973) and Granovetter (1978) pioneered the use of threshold response functions in models of collective social behavior (although they were not the first; see the citations in their papers). This was based on the intuition that, for a person to adopt some new behavior, the fraction of the population exhibiting that behavior might need to exceed some

critical value, the person's threshold. (Mathematically, a threshold response function $f(\phi; \phi_{\text{on}})$ with threshold ϕ_{on} returns 0 if $\phi < \phi_{\text{on}}$ and 1 if $\phi \geq \phi_{\text{on}}$[1].) These models were generalized to the case where the dynamics take place on a network by Watts (2002). If we initialize a social network with some fraction of active nodes, some of their neighbors' thresholds may be exceeded and the activity can spread (depending on the distribution of thresholds and network structure).

Standard threshold models have simple dynamical behavior: in a word, "spreading." If a single activation, on average, leads to more than one subsequent activation, then the spreading will be successful. The activity will increase in a sigmoid fashion until some final fraction of the network is activated. These kinds of spreading are often studied using branching processes. See the book by Harris (1963) for an overview of branching processes and the widely-used generating function formalism. Branching processes have been used to model, among other things, extinction of families, species, and genes, neutron cascades (as happens during nuclear chain reactions), and high energy particle showers caused by cosmic rays. The generating function formalism developed in part by Newman (as in Newman, 2003; Watts, 2002) to analyze spreading on networks is a straightforward application of the classical theory of branching processes.

We note that threshold random Boolean network models have been studied for the purpose of modeling neural networks (Aldana et al., 2003). Those models take place on signed, weighted graphs, which differ from the networks considered here, and the problems considered are different. It would be interesting to explore the connections between our on-off threshold model (Section 3) and other threshold Boolean network models.

2.4 Some notation

The Bachmann-Landau asymptotic notations are used throughout this thesis. When used carefully, asymptotic notation greatly improves the readability of analytic statements and proofs. It is also widely used in probability. For an overview of the notation's history and usage, see Knuth (1976) and the references therein. The notations we have used here are [citing from Knuth (1976)]:

- $O(f(n))$ is the set of all $g(n)$ such that there exist positive constants C and n_0 with $|g(n)| \leq Cf(n)$ for all $n \geq n_0$.

[1] The edge case could be defined differently, but this will not influence the dynamics except in carefully constructed "pathological" scenarios.

- $\Omega(f(n))$ is the set of all $g(n)$ such that there exist positive constants C and n_0 with $g(n) \geq Cf(n)$ for all $n \geq n_0$.
- $o(f(n))$ is the set of all $g(n)$ such that $g(n)/f(n) \to 0$ as $n \to \infty$.
- $g(n) \sim f(n)$ if $g(n)/f(n) \to 1$ as $n \to \infty$.

Formally, each of the above define sets of functions, but we often use statements such as "f is $O(g)$" (read as "f is big-oh of g") or "$f = O(g)$" to mean "$f \in O(g)$."

When expressing probabilities we will use notation of the form $P(x)$, $P(x|y)$, etc. Here, $P(x)$ is the probability that the random variable associated with x equals the specific value x. We leave out the random variables to avoid introducing unnecessary clutter.

Symbols in boldface represent vector quantities or vector-valued functions, e.g., $\mathbf{x} = (x_i)$. Subscripts have been left out in places for clarity.

3 The on-off threshold model

Here we study a simple extension of the classical threshold models (such as Schelling, 1971, 1973; Granovetter, 1978; Watts, 2002; Dodds and Watts, 2004, among others): the response function also includes an off-threshold. See Figure 2 for an example on-off threshold response function. This is exactly the model of Granovetter and Soong (1986), but on a network. We motivate this choice with the following (also see Granovetter and Soong, 1986). (1) Imitation: the on state becomes favored as the fraction of active neighbors surpasses the on-threshold (bandwagon effect). (2) Non-conformity: the on state is eventually less favorable with the fraction of active neighbors past the off-threshold (reverse bandwagon, snob effect). (3) Simplicity: in the absence of any raw data of "actual" response functions, which are surely highly context-dependent and variable, we choose arguably the simplest deterministic functions which capture imitation and non-conformity.

Let $\mathscr{G} = (V, E)$ be a graph with $N = |V|$. Assign each vertex $i \in V$ an on-threshold $\phi_{\text{on},i}$ and an off-threshold $\phi_{\text{off},i}$ with $0 \leq \phi_{\text{on},i} \leq \phi_{\text{off},i} \leq 1$. Then that node's response function $f_i(\phi_i; \phi_{\text{on},i}, \phi_{\text{off},i})$ is 1 if $\phi_{\text{on},i} \leq \phi_i \leq \phi_{\text{off},i}$ and 0 otherwise. Let $\mathbf{x}(0) \in \{0,1\}^N$ be the initial states of all nodes. At time step t, each node i computes the fraction $\phi_i(t)$ of their neighbors in $\mathscr{G}$ who are active and takes the state $x_i(t+1) = f_i\left(\phi_i(t); \phi_{\text{on},i}, \phi_{\text{off},i}\right)$ at the next time step. The above defines a deterministic dynamical system for a fixed graph and fixed thresholds.

We now make some quick remarks about the on-off threshold model. First, our model is a particular kind of Boolean network (Section 2.3.1). Note that each node reacts only to the fraction of its neighbors who are active, rather than the absolute number, and the input varies from 0 to 1 in steps of $1/k_i$, where k_i is node i's degree. Note that if $\phi_{\text{on},i} = 0$ the node activates spontaneously, and if $\phi_{\text{off},i} = 1$ we have the usual kind of threshold response function (without an off-threshold).

A crucial difference between our model and many related threshold models is that, in those models, an activated node can never reenter the susceptible state. Gleeson and Cahalane (2007) call this the permanently active property and elaborate on its importance to their analysis. Such models must eventually reach a steady state. When the dynamics are deterministic, this will be a fixed point, and in the presence of stochasticity the steady state is characterized by some fixed fraction of active nodes subject to fluctuations. The introduction of the off-threshold builds in a mechanism for node deactivation. Because nodes can now recurrently transition between on and off states, the deterministic dynamics can exhibit a chaotic transient (see Section 2.3.1), and the long time behavior can be periodic with potentially high period. With stochasticity, the dynamics can be truly chaotic and never repeat.

In the rest of this Section, we will describe how this model is different from the random Boolean networks in the literature. This is mainly due to the on-off threshold response functions we consider, but also the type of random graph on which the dynamics take place, varying amounts of stochasticity which we introduce in the networks and response functions, and the possibility of asynchronous updates.

3.1 The networks considered

The mean field analysis in Section 5 is applicable to any network which can be characterized by its degree distribution. As mentioned before, the vast majority of the theory of random Boolean networks assumes a regular random graph. Fortunately, such theories are easily generalized to other types of graphs with independent edges, such as Poisson and configuration model random graphs. Some specific results are given for the Poisson random graph $\mathcal{G}(N, k_{\text{avg}}/N)$, and these are the networks considered in Section 6.

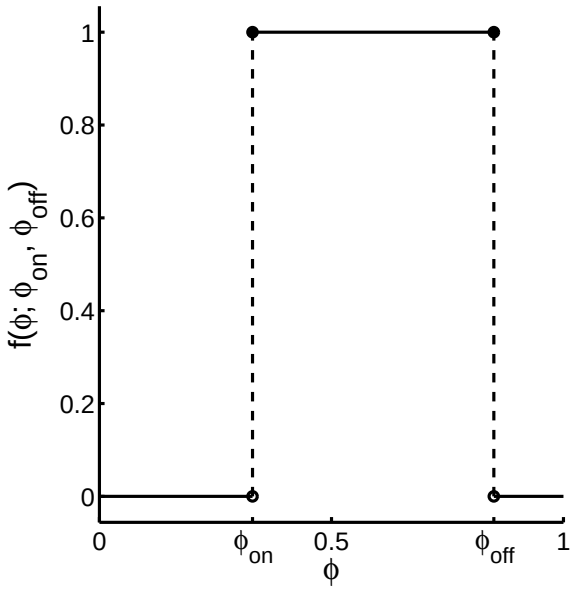

Figure 2: An example on-off threshold response function. Here, $\phi_{\text{on}} = 0.33$ and $\phi_{\text{off}} = 0.85$. The node will activate if $\phi_{\text{on}} \leq \phi \leq \phi_{\text{off}}$, where ϕ is the fraction of its neighbors who are active. Otherwise it turns off.

3.2 Stochastic variants

The specific graph and node thresholds determine exactly which behaviors are possible. These are chosen from some distribution of graphs, such as $\mathcal{G}(N, k_{\text{avg}}/N)$, and some distribution of thresholds, given by the joint density $P(\phi_{\text{on}}, \phi_{\text{off}})$. The specific graph and thresholds define a realization of the model (see Aldana et al., 2003). When these are fixed for all time, we have, in principle, full knowledge of the possible model dynamics. Given an initial condition $\mathbf{x}(0)$, the dynamics $\mathbf{x}(t)$ are deterministic and known for all $t \geq 0$.

With the introduction of noise, the system is no longer eventually periodic. Fluctuations at the node level allow a greater exploration of state space, and the behavior is comparable to that of the general class of discrete-time maps. Roughly speaking, the mean field theory we develop in Section 5 becomes more accurate as we introduce more stochasticity.

We introduce randomness in two parts of the model: the network and/or the response functions. Allowing for the network and responses to be either fixed for all time or resampled each time step and taking all possible combinations yields four different designs (see Table 1).

	Rewiring network	**F**ixed network
Probabilistic response	P-R	P-F
Deterministic response	D-R	D-F

Table 1: The four different ways the model can be realized. These are the combinations of fixed or rewired networks and probabilistic or deterministic response functions. In the thermodynamic limit of the fully stochastic version (P-R), where the graph and response functions change every time step, the mean field theory is exact (see Sec. 5).

3.2.1 Rewired graphs

First, the network itself can change every time step. This is the rewiring (R), as opposed to fixed (F), network case. For example, we could draw a new graph from $\mathcal{G}(N, k_{\text{avg}}/N)$ every time step. This amounts to rewiring the links while keeping the degree distribution fixed, and it is alternately known as a mean field, "annealed", or random mixing variant of the fixed or "quenched" model (Aldana et al., 2003).

3.2.2 Probabilistic responses

Second, the response functions can change every time step. This is the probabilistic (P), as opposed to the deterministic (D), response function case. Again, we will need a well-defined distribution $P(\phi_{\text{on}}, \phi_{\text{off}})$ for the thresholds. This amounts to having a single response function, the expected response function

$$f(\phi) = \int d\phi_{\text{on}} \int d\phi_{\text{off}}\ P(\phi_{\text{on}}, \phi_{\text{off}}) f(\phi; \phi_{\text{on}}, \phi_{\text{off}}). \tag{3.1}$$

We call $f : [0,1] \to [0,1]$ the *probabilistic response function*. Its interpretation is the following. For an updating node with a fraction ϕ of active neighbors at the current time step, then, at the next time step, the node assumes the state 1 with probability $f(\phi)$ and the state 0 with probability $1 - f(\phi)$.

3.2.3 The concept of "temperature" in the system

In this thesis, the network and response functions are either fixed for all time or resampled every time step. One could tune smoothly between the two extremes by introducing rates at which these reconfigurations occur. These rates are inversely related to quantities that behave like temperature (one for the network and another for the response functions). For the fixed case, the temperature is zero, and there are no fluctuations, while in the stochastic case, the temperature is very large or

infinite, and fluctuations occur every time step.

3.3 Synchronicity of the update

Finally, we introduce a parameter α for the probability that a given node updates. When $\alpha = 1$, all nodes update every time step, and the update rule is said to be synchronous. When $\alpha = 1/N$, only one node is expected to update with each time step, and the update rule is said to be effectively asynchronous. This is equivalent to a randomly ordered sequential update. For intermediate values, α is the expected fraction of nodes which update each time step.

4 Fixed networks

Take the case where the response functions and graph are fixed (D-F), but the update may be synchronous or asynchronous. Let $x_i(t)$ be the probability that node i is in state 1 at time t, and let $f_i(\phi) = f_i(\phi; \phi_{\text{on},i}, \phi_{\text{off},i})$. The dynamics follow the master equation

$$x_i(t+1) = \alpha f_i \left(\frac{\sum_{j=0}^{N} A_{ij} x_j(t)}{\sum_{j=0}^{N} A_{ij}} \right) + (1-\alpha) x_i(t), \tag{4.1}$$

which can be written in matrix-vector notation as

$$\mathbf{x}(t+1) = \alpha \mathbf{f}\left(T\mathbf{x}(t)\right) + (1-\alpha)\mathbf{x}(t). \tag{4.2}$$

Here $T = D^{-1}A$ is sometimes called the transition probability matrix (in the context of a random walker), D is the diagonal degree matrix, and $\mathbf{f} = (f_i)$[2]. Note that if $\alpha = 1$ we recover the fully deterministic response function dynamics, and $x_i(t) = 0$ or 1 for all t.

4.1 Asynchronous limit

Here, we show that when $\alpha \approx 1/N$, time is effectively continuous and the dynamics can be described by an ordinary differential equation. This is similar to the analysis of Gleeson (2008). Consider

[2] Eqns. (4.1) and (4.2) are not entirely correct when there are isolated nodes. In that case, $k_i = \sum_j A_{ij} = 0$ for certain i, thus the denominator in (4.1) is zero and D^{-1} undefined. If the initial network contains isolated nodes, we set all entries in the corresponding rows of T to zero.

Eqn. 4.2. Subtracting $\mathbf{x}(t)$ from both sides and setting $\Delta\mathbf{x}(t) = \mathbf{x}(t+1) - \mathbf{x}(t)$ and $\Delta t = 1$ yields

$$\frac{\Delta\mathbf{x}(t)}{\Delta t} = \alpha\left(\mathbf{f}(T\mathbf{x}(t)) - \mathbf{x}(t)\right). \tag{4.3}$$

Since α is assumed small, the right hand side is small, and thus $\Delta\mathbf{x}(t)$ is also small. Making the continuum approximation $d\mathbf{x}(t)/dt \approx \Delta\mathbf{x}(t)/\Delta t$ yields the differential equation

$$\frac{d\mathbf{x}}{dt} = \alpha\left(\mathbf{f}(T\mathbf{x}) - \mathbf{x}\right). \tag{4.4}$$

The parameter α sets the time scale for the system. From their form, similar asynchronous, continuous time limits apply to the dynamical equations in the densely connected case, Eqn. (4.5), and in the mean field theory, Eqns. (5.3) and (5.4).

4.2 Dense network limit for Poisson random graphs

The following result is particular to Poisson random graphs, but similar results are possible for other random graphs with dense limits. The normalized Laplacian matrix is defined as $\mathcal{L} \equiv I - D^{-1/2}AD^{-1/2}$, where I is the identity (West, 2001). So $T = D^{-1/2}(I - \mathcal{L})D^{1/2}$. By Oliveira (2010), when k_{avg} is $\Omega(\log N)$ there exists a typical Laplacian matrix $\mathcal{L}^{\text{typ}} = I_N - \mathbf{1}_N\mathbf{1}_N^\dagger/N$ [we let $\mathbf{1}_N$ denote the length-N vector of ones and $(\cdot)^\dagger$ the matrix transpose] such that the actual $\mathcal{L} \approx \mathcal{L}^{\text{typ}}$ in the induced 2-norm (spectral norm) with high probability. In this limit, if we assume uniform degrees $k_i = k_{\text{avg}}$ for all $i \in V$, then $T \approx T^{\text{typ}} = \mathbf{1}_N\mathbf{1}_N^\dagger/N$. So T effectively averages the node states: $T\mathbf{x}(t) \approx T^{\text{typ}}\mathbf{x}(t) = \sum_{i=1}^N x_i(t)/N \equiv \phi(t)$. Without a subscript, $\phi(t)$ denotes the active fraction of the network at time t. We make the above approximation in Eqn. 4.2 and average that equation over all nodes, finding

$$\phi(t+1) = \alpha f(\phi(t)) + (1-\alpha)\phi(t) \equiv \Phi(\phi(t); \alpha, f), \tag{4.5}$$

where we have assumed that N is large and the average of nodes' individual response functions $\sum_{i=1}^N f_i/N$ converges in a suitable sense to the stochastic response function f, Eqn. (3.1). This amounts to assuming a law of large numbers for the response functions, i.e., that the sample average converges to the expected function. Note that α tunes between the probabilistic response function $\Phi(\phi; 1) = f(\phi)$ and the 45° line $\Phi(\phi; 0) = \phi$. Also, the fixed points of Φ are fixed points of f, but

their stability will depend on α.

When the network is dense, it ceases to affect the dynamics, since each node sees a large number of other nodes. Thus the network is effectively the complete graph. In this way we recover the map models of Granovetter and Soong (1986).

5 Mean field theory

In physics, making a "mean field" calculation refers to replacing the complicated interactions among many particles by a single interaction with some effective external field. There are analogous techniques for understanding networks dynamics. Instead of considering the $|E|$ interactions among the N nodes, network mean field theories derive self-consistent expressions for the overall behavior of the network, after averaging over large sets of nodes. These have been fruitful in the study of random Boolean networks (Derrida and Pomeau, 1986) and can be surprisingly effective when networks are non-random (Melnik et al., 2011).

We derive a mean field theory, in the thermodynamic limit, for the dynamics of the on-off threshold model by blocking nodes according to their degree class. This is equivalent to nodes retaining their degree but rewiring edges every time step. The model is then part of the well-known class of random mixing models with non-uniform contact rates. Probabilistic (P-R) and deterministic (D-R) response functions result in equivalent behavior for these random mixing models. The important state variables end up being the active density of stubs. In an undirected network without degree-degree correlations, the state is described by a single variable $\rho(t)$. In the presence of correlations we must introduce more variables, i.e., $\rho_k(t), \rho_{k'}(t), \ldots$, to deal with the relevant degree classes.

5.1 Undirected networks

To derive the mean field equations in the simplest case — undirected, uncorrelated random graphs — consider a degree k node at time t. The probability that the node is in the 1 state at time $t+1$ given a density ρ of active stubs is

$$F_k(\rho; f) = \sum_{j=0}^{k} \binom{k}{j} \rho^j (1-\rho)^{k-j} f(j/k), \tag{5.1}$$

where each term in the sum counts the contributions from having 0, 1, ..., k active neighbors. Now, the probability of choosing a random stub which ends at a degree k node is $q_k = kp_k/k_{\text{avg}}$ in an uncorrelated random network (Newman, 2003). This is sometimes called the edge-degree distribution. So if all of the nodes update synchronously, the active density of stubs at $t+1$ will be

$$g(\rho; p_k, f) = \sum_{k=1}^{\infty} q_k F_k(\rho; f) = \sum_{k=1}^{\infty} \frac{kp_k}{k_{\text{avg}}} F_k(\rho; f). \tag{5.2}$$

Finally, if each node only updates with probability α, we have the following map for the density of active stubs:

$$\rho(t+1) = \alpha\, g\left(\rho(t); p_k, f\right) + (1-\alpha)\rho(t) \equiv G(\rho(t); p_k, f, \alpha). \tag{5.3}$$

By a similar argument, the active density of nodes is given by

$$\phi(t+1) = \alpha\, h(\rho(t); p_k, f) + (1-\alpha)\phi(t) \equiv H(\rho(t), \phi(t); p_k, f, \alpha), \tag{5.4}$$

where

$$h(\rho; p_k, f) = \sum_{k=0}^{\infty} p_k F_k(\rho; f). \tag{5.5}$$

Note that the edge-oriented state variable ρ contains all of the dynamically important information, rather than the vertex-oriented variable ϕ.

5.2 Analysis of the map equation

The function $F_k(\rho; f)$ is known in polynomial approximation theory as the kth Bernstein polynomial (in the variable ρ) of f (Phillips, 2003; Peña, 1999). These are approximating polynomials which have applications in computer graphics due to their "shape-preserving properties." The Bernstein operator $\mathbb{B}_k$ takes $f \mapsto F_k$. This is a linear, positive operator which preserves convexity for all k and exactly interpolates the endpoints $f(0)$ and $f(1)$. Immediate consequences include that each F_k is a smooth function and the kth derivatives $F_k^{(k)}(x) \to f^{(k)}(x)$ where $f^{(k)}(x)$ exists. For concave f (such as the tent or logistic maps), we have concave F_k for all k and $F_k \nearrow f$ uniformly. This convergence is typically slow. Importantly, $F_k \nearrow f$ implies that $g(\rho; p_k, f) \leq f$ for any degree distribution p_k.

In some cases, the dynamics of the undirected mean field theory given by $\rho(t+1) = G(\rho(t))$, Eqn. (5.3), are effectively those of the map Φ, from the dense limit Eqn. (4.5). We see that g,

Eqn. (5.2), can be seen as the expectation of a sequence of random functions F_k under the edge-degree distribution q_k (indeed, this is how it was derived). From the convergence of the F_k's, we expect that $g(\rho; p_k, f) \approx f(\rho)$ if the average degree k_{avg} is "large enough" and the edge-degree distribution has a "sharp enough" peak about k_{avg} (we will clarify this soon). Then as $k_{\text{avg}} \to \infty$, the mean field coincides with the dense network limit we found for Poisson random graphs, Eqn. 4.5. Some thought leads to a sufficient condition for this kind of convergence: the standard deviation $\sigma(k_{\text{avg}})$ of the degree distribution must be $o(k_{\text{avg}})$. In Appendix A we prove this as Lemma 1.

In general, if the original degree distribution p_k is characterized by having mean k_{avg}, variance σ^2, and skewness γ_1, then the edge-degree distribution q_k will have mean $k_{\text{avg}} + \sigma^2/k_{\text{avg}}$ and variance $\sigma^2[1 + \gamma_1\sigma/k_{\text{avg}} - (\sigma/k_{\text{avg}})^2]$. Considering the behavior as $k_{\text{avg}} \to \infty$, we can conclude that requiring $\sigma \in o(k_{\text{avg}})$ and $\gamma_1 \in o(1)$ are sufficient conditions on p_k to apply Lemma 1. Poisson degree distributions ($\sigma = \sqrt{k_{\text{avg}}}$ and $\gamma_1 = k_{\text{avg}}^{-1/2}$) fit these criteria.

5.3 Generalized random networks

In the most general kind of random networks, edges can be undirected or directed, and we then denote node degree by a vector $\mathbf{k} = (k^{(\text{u})}, k^{(\text{i})}, k^{(\text{o})})^\dagger$. The degree distribution is written as $p_{\mathbf{k}} \equiv P(\mathbf{k})$. There may also be correlations between node degrees. Correlations of this type are encoded by the conditional probabilities

$$
\begin{aligned}
p^{(\text{u})}_{\mathbf{k},\mathbf{k}'} &\equiv P(\mathbf{k}, \text{undirected}|\mathbf{k}') \\
p^{(\text{i})}_{\mathbf{k},\mathbf{k}'} &\equiv P(\mathbf{k}, \text{incoming}|\mathbf{k}') \\
p^{(\text{o})}_{\mathbf{k},\mathbf{k}'} &\equiv P(\mathbf{k}, \text{outgoing}|\mathbf{k}'),
\end{aligned}
$$

the probability that an edge starting at a degree $\mathbf{k}'$ node ends at a degree $\mathbf{k}$ node and is, respectively, undirected, incoming, or outgoing relative to the destination degree $\mathbf{k}$ node. We introduced this convention in a series of papers (Payne et al., 2011; Dodds et al., 2011). These conditional probabilities can also be defined in terms of the joint distributions of node types connected by undirected and directed edges. The mean field equations for the on-off threshold model are closely related to the equations for the time evolution of a contagion process (Payne et al., 2011, Eqns. (13–15)). We omit a detailed derivation, since it is similar to that in Section 5.1 (see also Gleeson and Cahalane, 2007; Payne et al., 2011). The result is a coupled system of equations for the density of active stubs which

now may depend on node type ($\mathbf{k}$) and edge type (undirected or directed):

$$\begin{aligned}
\rho^{(\mathrm{u})}_{\mathbf{k}}(t+1) =& \alpha \sum_{\mathbf{k}'} p^{(\mathrm{u})}_{\mathbf{k},\mathbf{k}'} \sum_{j_u=0}^{k^{(\mathrm{u})\prime}} \sum_{j_i=0}^{k^{(\mathrm{i})\prime}} \binom{k^{(\mathrm{u})\prime}}{j_u} \binom{k^{(\mathrm{i})\prime}}{j_i} \\
& \times \left[\rho^{(\mathrm{u})}_{\mathbf{k}'}(t)\right]^{j_u} \left[1-\rho^{(\mathrm{u})}_{\mathbf{k}'}(t)\right]^{(k^{(\mathrm{u})\prime}-j_u)} \\
& \times \left[\rho^{(\mathrm{i})}_{\mathbf{k}'}(t)\right]^{j_i} \left[1-\rho^{(\mathrm{i})}_{\mathbf{k}'}(t)\right]^{(k^{(\mathrm{i})\prime}-j_i)} \; f\left(\frac{j_u+j_i}{k^{(\mathrm{u})\prime}+k^{(\mathrm{i})\prime}}\right) \\
& + (1-\alpha)\rho^{(\mathrm{u})}_{\mathbf{k}}(t) \qquad (5.6) \\
\rho^{(\mathrm{i})}_{\mathbf{k}}(t+1) =& \alpha \sum_{\mathbf{k}'} p^{(\mathrm{i})}_{\mathbf{k},\mathbf{k}'} \sum_{j_u=0}^{k^{(\mathrm{u})\prime}} \sum_{j_i=0}^{k^{(\mathrm{i})\prime}} \binom{k^{(\mathrm{u})\prime}}{j_u} \binom{k^{(\mathrm{i})\prime}}{j_i} \\
& \times \left[\rho^{(\mathrm{u})}_{\mathbf{k}'}(t)\right]^{j_u} \left[1-\rho^{(\mathrm{u})}_{\mathbf{k}'}(t)\right]^{(k^{(\mathrm{u})\prime}-j_u)} \\
& \times \left[\rho^{(\mathrm{i})}_{\mathbf{k}'}(t)\right]^{j_i} \left[1-\rho^{(\mathrm{i})}_{\mathbf{k}'}(t)\right]^{(k^{(\mathrm{i})\prime}-j_i)} \; f\left(\frac{j_u+j_i}{k^{(\mathrm{u})\prime}+k^{(\mathrm{i})\prime}}\right) \\
& + (1-\alpha)\rho^{(\mathrm{i})}_{\mathbf{k}}(t)\,. \qquad (5.7)
\end{aligned}$$

The active fraction of nodes at a given time is given by:

$$\begin{aligned}
\phi(t+1) =& \alpha \sum_{\mathbf{k}} p_{\mathbf{k}} \sum_{j_u=0}^{k^{(\mathrm{u})}} \sum_{j_i=0}^{k^{(\mathrm{i})}} \binom{k^{(\mathrm{u})}}{j_u} \binom{k^{(\mathrm{i})}}{j_i} \\
& \times \left[\rho^{(\mathrm{u})}_{\mathbf{k}}(t)\right]^{j_u} \left[1-\rho^{(\mathrm{u})}_{\mathbf{k}}(t)\right]^{(k^{(\mathrm{u})}-j_u)} \\
& \times \left[\rho^{(\mathrm{i})}_{\mathbf{k}}(t)\right]^{j_i} \left[1-\rho^{(\mathrm{i})}_{\mathbf{k}}(t)\right]^{(k^{(\mathrm{i})}-j_i)} \; f\left(\frac{j_u+j_i}{k^{(\mathrm{u})}+k^{(\mathrm{i})}}\right) \\
& + (1-\alpha)\phi(t)\,. \qquad (5.8)
\end{aligned}$$

6 Poisson random graphs with tent map average response function

The results so far have been entirely general, in the sense that the underlying network and thresholds are arbitrary. Now we apply the general theory to the case of Poisson random graphs with a simple distribution of thresholds.

The networks we consider are Poisson random graphs from $\mathcal{G}(N, k_{\mathrm{avg}}/N)$. The thresholds ϕ_{on} and ϕ_{off} are now distributed uniformly on $[0, 1/2)$ and $[1/2, 1)$, respectively. This distribution results

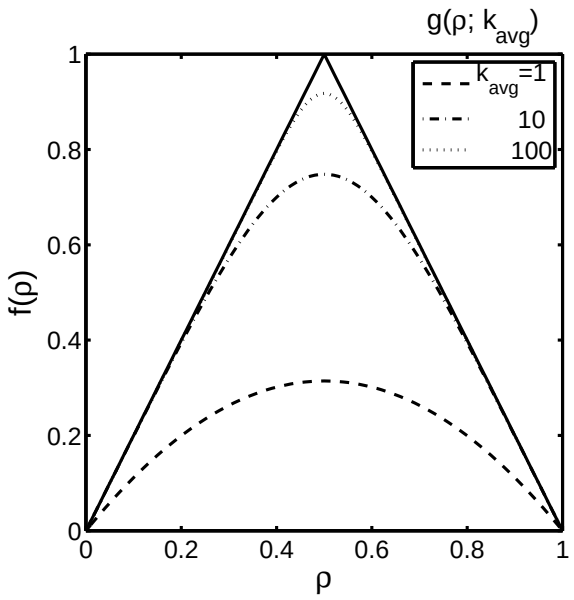

Figure 3: The tent map probabilistic response function $f(\rho)$, Eqn. (6.1). Note that we use the argument ρ for comparison with the edge maps $g(\rho; k_{\text{avg}}) = g(\rho; p_k, f)$, Eqn. (5.2), shown for $k_{\text{avg}} = 1, 10, 100$. These p_k are Poisson distributions with mean k_{avg}. As k_{avg} increases, $g(\rho; k_{\text{avg}})$ increases to $f(\rho)$.

in the probabilistic response function (see Figure 3):

$$f(\phi) = \begin{cases} 2\phi & \text{if} \quad 0 \leq \phi < 1/2, \\ 2 - 2\phi & \text{if} \quad 1/2 \leq \phi \leq 1. \end{cases} \tag{6.1}$$

The tent map is a well-known chaotic map of the unit interval (Alligood et al., 1996). We thus expect the on-off threshold model with this probabilistic response function to exhibit similarly interesting behavior.

6.1 Analysis of the dense limit

When the network is in the dense limit (Section 4.2), the dynamics follow $\phi(t+1) = \Phi(\phi(t); \alpha)$, where

$$\Phi(\phi; \alpha) = \alpha f(\phi) + (1 - \alpha)\phi = \begin{cases} (1+\alpha)\phi & \text{if} \quad 0 \leq \phi < 1/2, \\ (1-3\alpha)\phi + 2\alpha & \text{if} \quad 1/2 \leq \phi \leq 1. \end{cases} \tag{6.2}$$

Solving for the fixed points of $\Phi(\phi; \alpha)$, we find one at $\phi = 0$ and another at $\phi = 2/3$. When $\alpha < 2/3$, the nonzero fixed point is attracting for all initial conditions except $\phi = 0$. When $\alpha = 2/3$, $[1/2, 5/6]$

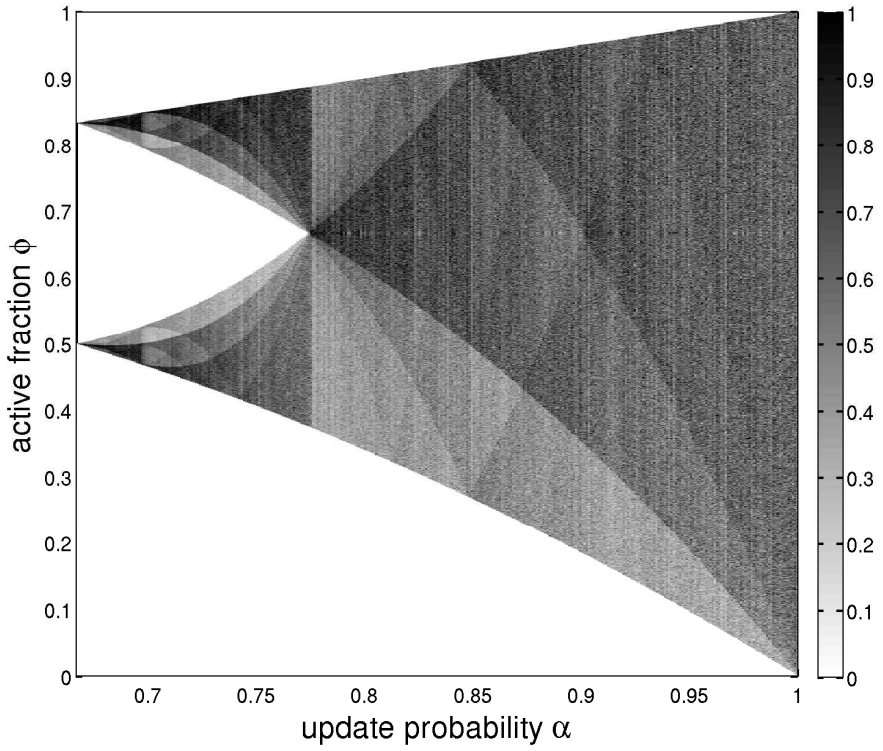

Figure 4: Bifurcation diagram for the dense map $\Phi(\phi;\alpha)$, Eqn. (6.2). This was generated by iterating the map at 1000 α values between 0 and 1. The iteration was carried out with 3 random initial conditions for 10000 time steps each, discarding the first 1000. The ϕ-axis contains 1000 bins and the invariant density, shown by the grayscale value, is normalized by the maximum for each α. With $\alpha < 2/3$ (not shown), all trajectories go to the fixed point at $\phi = 2/3$.

is an interval of period 2 centers. Any orbit will eventually land on one of these period 2 orbits. When $\alpha > 2/3$, this interval of period 2 orbits ceases to exist, and more complicated behavior ensues. Figure 4 shows the bifurcation diagram for $\Phi(\phi;\alpha)$. From the bifurcation diagram, the orbit appears to cover dense subsets of the unit interval when $\alpha > 2/3$. The bifurcation diagram appears like that of the tent map (not shown; see Alligood et al., 1996; Dodds et al., 2012), except that the branches to the right of the first bifurcation point are separated here by the interval of period 2 orbits.

6.1.1 The effect of conformists, an aside

Suppose some fraction c of the population is made up of individuals without any off-threshold (alternatively, each of their off-thresholds $\phi_{\text{off}} = 1$). These individuals are conformist or "purely pro-social" in the sense that they are perfectly happy being part of the majority. For simplicity,

assume $\alpha = 1$. The map $\Phi(\phi; c) = 2\phi$ for $0 \leq \phi < 1/2$ and $2 - 2(1-c)\phi$ for $1/2 \leq \phi \leq 1$. If $c > 1/2$, then the equilibrium at $2/3$ is stable. Pure conformists, then, can have a stabilizing effect on the process. We expect a similar effect when the network is not dense.

6.2 Mean field

6.2.1 Analysis

In this specific example, we can write the degree-dependent map $F_k(\rho; f)$ in terms of incomplete regularized beta functions (NIST, 2012). Since f is understood to be the tent map, we will write $F_k(\rho; f) = F_k(\rho)$. First, use the piecewise form of Eqn. (6.1) to write

$$\begin{aligned} F_k(\rho) &= \sum_{j=0}^{M} \binom{k}{j} \rho^j (1-\rho)^{k-j} \left(\frac{2j}{k}\right) + \sum_{j=M+1}^{k} \binom{k}{j} \rho^j (1-\rho)^{k-j} \left(2 - \frac{2j}{k}\right) \\ &= 2 - 2\rho - 2\sum_{j=0}^{M} \binom{k}{j} \rho^j (1-\rho)^{k-j} + \left(\frac{4}{k}\right) \sum_{j=0}^{M} \binom{k}{j} \rho^j (1-\rho)^{k-j} j. \end{aligned}$$

We have let $M = \lfloor k/2 \rfloor$ for clarity ($\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling functions) and used the fact that the binomial distribution $\binom{k}{j}\rho^j(1-\rho)^{k-j}$ sums to one and has mean $k\rho$. For $n \leq M$, we have the identity

$$\sum_{j=0}^{M} (j)_n \binom{k}{j} \rho^j (1-\rho)^{k-j} = \rho^n (k)_n I_{1-\rho}(k-M, M-n+1) \tag{6.3}$$

where $I_x(a, b)$ is the regularized incomplete beta function and $(k)_n = k(k-1)\cdots(k-(n-1))$ is the falling factorial (Winkler et al., 1972; NIST, 2012). This is an expression for the partial (up to M) nth factorial moment of the binomial distribution with parameters k and ρ. Note that when $n = 0$ we recover the well-known expression for the binomial cumulative distribution function. So,

$$F_k(\rho) = 2\rho - 4\rho I_\rho(M, k-M) + 2I_\rho(M+1, k-M) \tag{6.4a}$$

$$= (2 - 2\rho) - (2I_{1-\rho}(k-M, M+1) - 4\rho I_{1-\rho}(k-M, M)). \tag{6.4b}$$

When $0 \leq \rho \leq 1/2$, the form of Eqn. (6.4a) can be used to show directly that $F_k(\rho)$ is bounded above by the tent map $f(\rho)$, which we already knew from the properties of the Bernstein polynomials. A similar approach works for the region $1/2 \leq \rho \leq 1$ using Eqn. (6.4b). We find a weak bound for

the rate at which $F_k(\rho)$ converges to $f(\rho)$. For $\rho < 1/2$, using Eqn. (6.4a),

$$\begin{aligned} f(\rho) - F_k(\rho) &= 4\rho I_\rho(M, k-M) - 2I_\rho(M+1, k-M) \\ &\leq 2\left(I_\rho(M, k-M) - I_\rho(M+1, k-M)\right) \\ &\leq 2\frac{\rho^M(1-\rho)^{k-M}}{MB(M, k-M)} \\ &\leq \frac{2(1/2)^k}{MB(M, k-M)} \\ &\leq \frac{4\Gamma(k)}{k2^k[\Gamma(k/2)]^2} \qquad \text{(for even } k\text{)} \\ &\leq 2\pi^{-1/2}\frac{1}{k}\cdot\frac{\Gamma(k/2+1/2)}{\Gamma(k/2)} = O(k^{-1/2}), \end{aligned}$$

where we have used identities from NIST (2012). We find the same $O(k^{-1/2})$ behavior using Eqn. (6.4b) in the region $\rho > 1/2$.

Finally, note that the active edge fraction $\rho(t) \approx \phi(t)$, the active node fraction. This is because Poisson random graphs are highly regular, with $q_k = kp_k/k_{\text{avg}} = p_{k-1} \approx p_k$. Thus the mean field dynamics for active edge density are effectively the same as for active node density.

6.2.2 Numerical algorithm

The map $g(\rho; p_k, f)$ is parametrized here by the network parameter k_{avg}, since p_k is fixed as a Poisson distribution with mean k_{avg} and f is the tent map, and we write it as simply $g(\rho; k_{\text{avg}})$. To evaluate $g(\rho; k_{\text{avg}})$, we compute $F_k(\rho)$ using Eqn. (6.4) and constrain the sum in Eqn. (5.2) to values of k with $\lfloor k_{\text{avg}} - 3\sqrt{k_{\text{avg}}} \rfloor \leq k \leq \lceil k_{\text{avg}} + 3\sqrt{k_{\text{avg}}} \rceil$. This computes contributions to within three standard deviations of the average degree in the graph, requiring only $O(\sqrt{k_{\text{avg}}})$ evaluations of Eqn. (6.4). The representation in Eqn. (6.4) allows for quick numerical evaluation of $F_k(\rho)$ for any k, which we performed in MATLAB using the built-in routines for the incomplete beta function.

In Figure 3, we show $g(\rho; k_{\text{avg}})$ for $k_{\text{avg}} = 1, 10$, and 100. We confirm the conclusions of Section 5.2: $g(\rho; k_{\text{avg}})$ is bounded above by $f(\rho)$, and $g(\rho; k_{\text{avg}}) \nearrow f(\rho)$ as $k_{\text{avg}} \to \infty$. Convergence is slowest at $\rho = 1/2$, and the kink that the tent map has there has been smoothed out by the effect of the Bernstein operator.

6.3 Simulations

We performed direct simulations of the on-off threshold model for the D-F, P-F, and P-R designs, in the abbreviations of Table 1. Unless otherwise noted, $N = 10^4$. For all of the bifurcation diagrams, the first 3000 time steps were considered transient and discarded, and the invariant density of ρ was calculated from the following 1000 points. For plotting purposes, the invariant density was normalized by its maximum at those parameters. For example, in Figure 4 we plot $P(\phi|\alpha)/\max_\phi P(\phi|\alpha)$ rather than the raw density $P(\phi|\alpha)$.

To compare the mean field theory to those simulations, we numerically iterated the edge map $\rho(t+1) = G(\rho(t); k_{\text{avg}}, \alpha)$ for different values of α and k_{avg}. We then created bifurcation diagrams of the possible behavior in the mean field as was done for the simulations.

6.4 Results

We show results for the deterministic model on a small network in Figure 5. Here, $N = 100$ and $k_{\text{avg}} = 17$. Starting from an initial active node at $t = 0$, the active population grows monotonically over the next 6 time steps. From $t = 6$ to $t \approx 80$, the transient time, the active population fluctuates in a similar manner to the stochastic case. After $t \approx 80$, the state collapses into a period 4 orbit. We call the overall period of the system its "macroperiod." Individual nodes may exhibit different "microperiods." Note that the macroperiod is the lowest common multiple of the individual nodes' microperiods. In Figure 5, we observe microperiods 1, 2, and 4 in the timeseries of individual node activity. A majority of the nodes end up frozen in the on or off state, with approximately 20% of the nodes exhibiting cyclical behavior after collapse. The focus of this thesis has been the analysis of the on-off threshold model, and the D-F case has not been as amenable to analysis as the stochastic cases. A deeper examination through simulation of the deterministic case will appear in Dodds et al. (2012).

We explore the mean field dynamics by examining the limiting behavior of the active edge fraction ρ under the map $G(\rho; k_{\text{avg}}, \alpha)$. We simulated the map dynamics for a mesh of points in the (k_{avg}, α) plane. We plotted the 3-dimensional (3-d; N-d denotes N-dimensional) bifurcation structure of the mean field theory in Figure 6. We also made 2-d bifurcation plots for fixed k_{avg} and α slices through this volume, shown in Figures 7 and 8. In all cases, the invariant density of ρ is normalized by its maximum for that (k_{avg}, α) pair and indicated by the grayscale value.

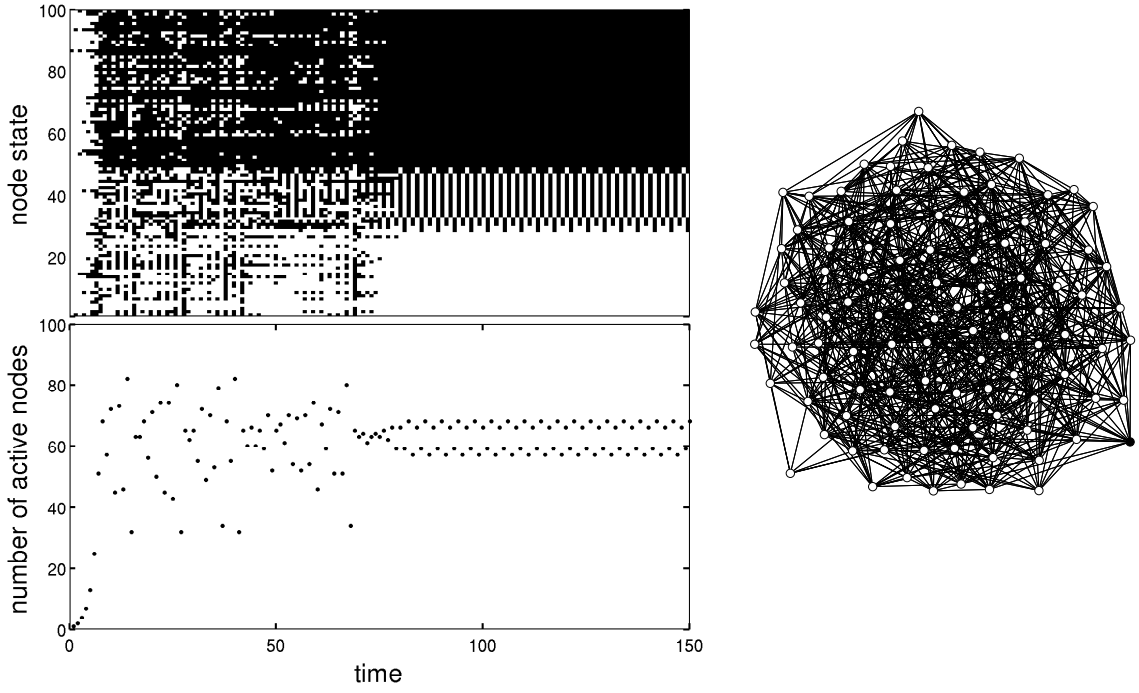

Figure 5: Deterministic (D-F) dynamics on a small graph. Here, $N = 100$ and $k_{\text{avg}} = 17$. On the left, we plot the state evolution over time. The upper plot shows individual node states (black = active) sorted by their eventual level of activity, and the lower plot shows the total number of active nodes. We see that the contagion takes off, followed by a transient period of unstable behavior until around time step 80, when the system enters a macroperiod 4 orbit. Note that individual nodes exhibit different microperiods (see Sec. 6.4). On the right, we show the network itself with the initial seed node in black in the lower right.

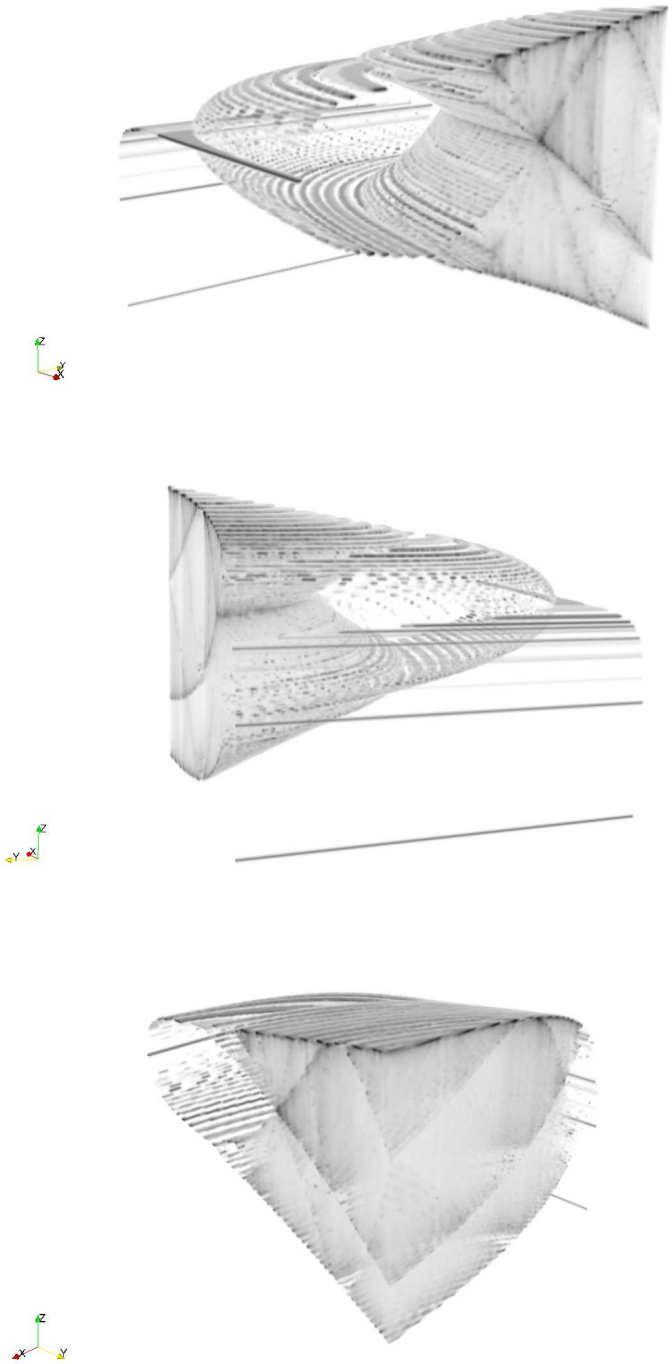

Figure 6: The 3-dimensional bifurcation diagram computed from the mean field theory. The axes X = average degree k_{avg}, Y = update probability α, and Z = active edge fraction ρ. The discontinuities of the surface are due to the limited resolution of our simulations. See Figure 7 for the parameters used. This was plotted in Paraview.

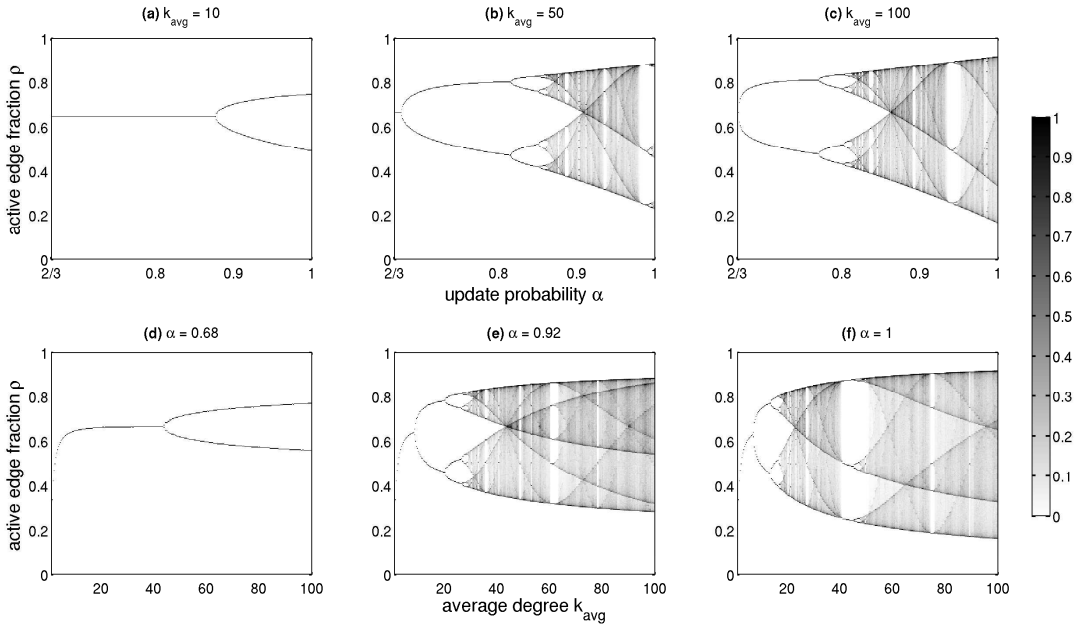

Figure 7: Mean field theory bifurcation diagram slices for various fixed values of k_{avg} and α. The top row (a–c) shows slices for fixed k_{avg}. As $k_{\text{avg}} \to \infty$, the k_{avg}-slice bifurcation diagram asymptotically approaches the bifurcation diagram for the dense map, Figure 4. Note that the first bifurcation point, near 2/3, grows steeper with increasing k_{avg}. The bottom row (d–f) shows slices for fixed α. The resolution of the simulations was $\alpha = 0.664, 0.665, \ldots, 1$, $k_{\text{avg}} = 1, 1.33, \ldots, 100$, and ρ bins were made for 1000 points between 0 and 1.

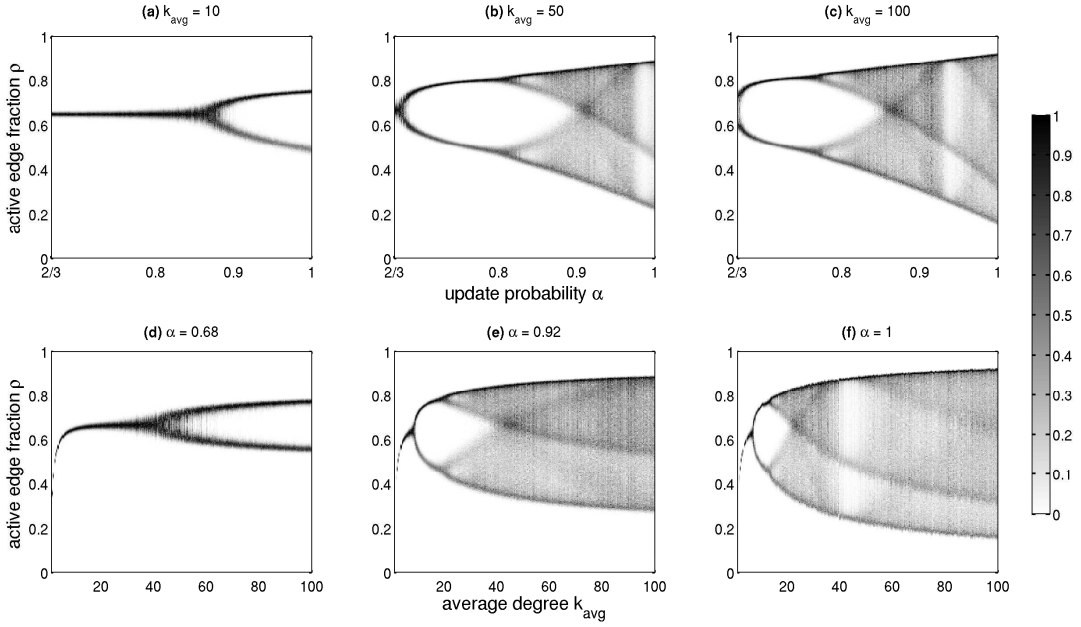

Figure 8: Bifurcation diagram from fully stochastic (P-R) simulations. The same parameters were used as in Figure 7, which has the same structure only more blurred.

The mean field map dynamics exhibit period-doubling bifurcations in both parameters k_{avg} and α. Visualizing the bifurcation structure in 3-d (Figure 6) shows interlacing period-doubling cascades in the two parameter dimensions. These bifurcations are more clearly resolved when we take slices of the volume for fixed parameter values. The mean field theory (Figure 7) closely matches the P-R simulations (Figure 8). The first derivative $\partial G(\rho; k_{\text{avg}}, \alpha)/\partial\rho < \partial\Phi(\rho; \alpha)/\partial\rho$ for any finite k_{avg}, so the bifurcation point $\alpha = 2/3$ which we found for the dense map Φ is an upper bound for the first bifurcation point of G. The actual location of the first bifurcation point depends on k_{avg}, but $\alpha = 2/3$ becomes more accurate for higher k_{avg} (it is an excellent approximation in Figures 7c and 8c, where $k_{\text{avg}} = 100$). When $\alpha = 1$, the first bifurcation point occurs at $k_{\text{avg}} \approx 7$.

The bifurcation diagram slices resemble each other and evidently fall into the same universality class as the logistic map (Feigenbaum, 1978, 1979). This class contains all 1-d maps with a single, locally-quadratic maximum. Due to the properties of the Bernstein polynomials, $F_k(\rho; f)$ will universally have such a quadratic maximum for any concave, continuous f (Phillips, 2003). So this will also be true for $g(\rho; k_{\text{avg}}, f)$ with k_{avg} finite, and we see that k_{avg} partially determines the amplitude of that maximum in Figure 3. Thus k_{avg} acts as a bifurcation parameter. The parameter α tunes between $G(\rho; k_{\text{avg}}, 1) = g(\rho; k_{\text{avg}}, f)$ and $G(\rho; k_{\text{avg}}, 0) = \rho$, so it has a similar effect. Note that the tent map f and the dense limit map Φ are kinked at their maxima, so their bifurcation diagrams are qualitatively different from those of the mean field. The network, by constraining the interactions among the population, causes the mean field behavior to fall into a different universality class than the response function map.

7 Conclusions

We constructed the on-off threshold model as a simple model for social contagion resulting from limited imitation. We see that including an aversion to total conformity results in more complicated, even chaotic dynamics. This model also allows us to study the effects of differing amounts of fixedness in the social network and individual response functions, and we developed a detailed mean field theory which is exact for random mixing versions of the model. Finally, we applied the theory to a specific case, where the network is Poisson and the response functions average to the tent map.

The model exhibits rich mathematical behavior. The deterministic case, which we have barely touched on here, merits further study. In particular, we would like to characterize the distribution of

periodic sinks, how the collapse time scales with system size, and how similar the transient dynamics are to the mean field dynamics.

Furthermore, the model should be tested on realistic networks. These could include power law or small world random graphs, or real social networks gleaned from data. In a manner similar to Melnik et al. (2011), one could evaluate the accuracy of the mean field theory for real networks.

Finally, the ultimate validation of this model would emerge from a better understanding of social dynamics themselves. Characterization of people's true response functions is therefore critical (some work has gone in this direction; see Centola, 2010, 2011; Romero et al., 2011; Ugander et al., 2012). Comparison of model output to large data sets, such as observational data from social media or online experiments, is an area for further experimentation. This might lead to more complicated context- and history-dependent models. As we collect more data and refine experiments, the eventual goal of quantifiably predicting human behaviors, including fashions and trends, seems achievable.

References

Aldana, M., Coppersmith, S., and Kadanoff, L. P. (2003). Boolean dynamics with random couplings. In Kaplan, E., Marsden, J. E., and Sreenivasan, K. R., editors, *Perspectives and Problems in Nonlinear Science*, chapter 2, pages 23–90. Springer, New York. `http://arxiv.org/abs/nlin/0204062`.

Alligood, K. T., Sauer, T. D., and Yorke, J. A. (1996). *Chaos: An Introduction to Dynamical Systems*. Springer.

Bollobás, B. (2001). *Random Graphs*. Cambridge University Press, 2nd edition.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329:1194–1197.

Centola, D. (2011). An experimental study of homophily in the adoption of health behavior. *Science*, 334:1269–1272.

Chung, F. and Lu, L. (2002). The average distances in random graphs with give expected degrees. *PNAS*, 99(25):15879–15882.

Derrida, B. and Pomeau, Y. (1986). Random networks of automata: A simple annealed approximation. *Europhys. Lett.*, 1(2):45–49.

Dodds, P. S., Harris, K. D., and Danforth, C. M. (2012). Limited Imitation Contagion on Random Networks: Chaos, Universality, and Unpredictability. `http://arxiv.org/abs/1208.0255`.

Dodds, P. S., Harris, K. D., and Payne, J. L. (2011). Direct, physically motivated derivation of the contagion condition for spreading processes on generalized random networks. *Phys. Rev. E*, 83:056122.

Dodds, P. S. and Watts, D. J. (2004). Universal behavior in a generalized model of contagion. *Phys. Rev. Lett.*, 92(21):218701.

Feigenbaum, M. J. (1978). Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics*, 19(1):25–52.

Feigenbaum, M. J. (1979). The universal metric properties of nonlinear transformations. *Journal of Statistical Physics*, 21(6):669–706.

Gleeson, J. P. (2008). Cascades on correlated and modular random networks. *Phys. Rev. E*, 77:046117.

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

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443.

Granovetter, M. and Soong, R. (1986). Threshold models of interpersonal effects in consumer demand. *Journal of Economic Behavior and Organization*, 7:83–99.

Hackett, A. W. (2011). *Cascade Dynamics on Complex Networks*. PhD thesis, University of Limerick.

Harris, T. E. (1963). *The Theory of Branching Processes*. Number 119 in Die Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin.

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22:437–467.

Knuth, D. E. (1976). Big omicron and big omega and big theta. *SIGACT News*, 8(2):18–24.

Melnik, S., Hackett, A., Porter, M. A., Mucha, P. J., and Gleeson, J. P. (2011). The unreasonable effectiveness of tree-based theory for networks with clustering. *Phys. Rev. E*, 83(3):036112.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827.

Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161–179.

Molloy, M. and Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics Probability and Computing*, 7(3):295–305.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2):167–256.

NIST (2012). Digital Library of Mathematical Functions. `http://dlmf.nist.gov`.

Oliveira, R. I. (2010). Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges. `http://arxiv.org/abs/0911.0600`.

Payne, J. L., Harris, K. D., and Dodds, P. S. (2011). Exact solutions for social and biological contagion models on mixed directed and undirected, degree-correlated random networks. *Phys. Rev. E*, 84:016110.

Peña, J. M., editor (1999). *Shape Preserving Representations in Computer-Aided Geometric Design*. Nova Science Publishers, Inc.

Phillips, G. M. (2003). *Interpolation and Approximation by Polynomials*. Springer. See chapter 7.

Romero, D. M., Meeder, B., and Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. In *Proc. 20th ACM International World Wide Web Conference*.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143–186.

Schelling, T. C. (1973). Hockey helmets, concealed weapons, and daylight saving. *Journal of Conflict Resolution*, 17(3):381–428.

Simmel, G. (1957). Fashion. *American Journal of Sociology*, 62(6):541–559.

Stauffer, D. and Aharony, A. (1994). *Introduction to Percolation Theory*. Taylor and Francis, London, 2nd edition.

Ugander, J., Backstrom, L., Marlow, C., and Kleinberg, J. (2012). Structural diversity in social contagion. *PNAS*, 109(16):5962–5966.

Vespignani, A. (2012). Modelling dynamical processes in complex socio-technical systems. *Nature Physics*, 8:32–39.

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

West, D. B. (2001). *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ, 2nd edition.

Winkler, R. L., Roodman, G. M., and Britney, R. R. (1972). The determination of partial moments. *Management Science*, 19(3):290–296.

Appendix A Proof of Lemma 1

Lemma 1. *For $k \geq 1$, let f_k be continuous real-valued functions on a compact domain X with $f_k \to f$ uniformly. Let p_k be a probability mass function on $\mathbb{Z}^+$ parametrized by its mean μ and with standard deviation $\sigma(\mu)$, assumed to be $o(\mu)$. Then,*

$$\lim_{\mu\to\infty}\left(\sum_{k=0}^{\infty} p_k f_k\right) = f.$$

Proof. Suppose $0 \leq a < 1$ and let $K = \lfloor \mu - \mu^a \rfloor$. Then,

$$g = \sum_{k=0}^{\infty} p_k f_k = \sum_{k=0}^{K} p_k f_k + \sum_{k=K+1}^{\infty} p_k f_k. \tag{A.1}$$

Since $f_k \to f$ uniformly as $k \to \infty$, for any $\epsilon > 0$ we can choose μ large enough that

$$|f_k(x) - f(x)| < \epsilon \tag{A.2}$$

for all $k > K$ and all $x \in X$. Without loss of generality, assume that $|f_k| \leq 1$ for all k. Then,

$$|g - f| \leq \left(\frac{\sigma}{\mu^a}\right)^2 + \epsilon.$$

The σ/μ^a term is a consequence of the Chebyshev inequality (Bollobás, 2001) applied to the first sum in (A.1). Since σ grows sublinearly in μ, this term vanishes for some $0 \leq a < 1$ when we take the limit $\mu \to \infty$. The ϵ term comes from using (A.2) in the second sum in (A.1), and it can be made arbitrarily small. □

Appendix B Online material

To better explore the 3-d mean field bifurcation structure, we created movies of the the k_{avg} and α slices as the parameters are dialed. The videos are available at
`http://www.uvm.edu/~kharris/on-off-threshold/bifurc_movies.zip`.
Also, a VTK file with the 3-d bifurcation data, viewable in Paraview is at
`http://www.uvm.edu/~kharris/on-off-threshold/volume_normalized.vtk.zip`.

Videos of the individual-node dynamics for small networks in the D-F and P-F cases are shown for some parameters which produce interesting behavior. These are available at
`http://www.uvm.edu/~kharris/on-off-threshold/graph_movies.zip`.

The D-F, P-F, and P-R cases were implemented in Python. The code is available at
`http://www.uvm.edu/~kharris/on-off-threshold/code.zip`.