

On Watts' Cascade Model with Random Link Weights

T. R. Hurd¹, James P. Gleeson²

¹ *Department of Mathematics, McMaster University, Canada*

² *MACSI, Department of Mathematics & Statistics,
University of Limerick, Ireland*

November 24, 2012

Abstract

We study an extension of D. Watts 2002 model of information cascades in social networks where edge weights are taken to be random, an innovation motivated by recent applications of cascade analysis to systemic risk in financial networks. The main result is a probabilistic analysis that characterizes the cascade in an infinite network as the fixed point of a vector-valued mapping, explicit in terms of convolution integrals that can be efficiently evaluated numerically using the fast Fourier transform algorithm. A second result gives an approximate probabilistic analysis of cascades on “real world networks”, finite networks based on a fixed deterministic graph. Extensive cross testing with Monte Carlo estimates shows that this approximate analysis performs surprisingly well, and provides a flexible “microscope” that can be used to investigate properties of information cascades in real world networks over a wide range of model parameters.

Key words: Contagion, random graph, stochastic network, information cascade, cascade condition, percolation, systemic risk, banking network, domino effect.

AMS Subject Classification: 05C80, 05C82, 90B15, 91B74, 91G50

1 Introduction

The main concern of the present paper is to study a specific generalization of the well-known random cascade model introduced by Duncan Watts [2002] that represents the spread of an “infectious idea”, such as a fad, a rumour, a song, a political opinion, or a new technology, on a random social network. In Watts' model, individuals and their friendship relations are modelled as nodes and undirected links of a “random network”, i.e. a generalization of the random graph model of Erdős and Rényi [1959]. Each node is assumed to have a deterministic or random threshold Φ , such that a

node will “adopt” the new idea as soon the number of its friends already adopting this idea exceeds the threshold Φ . Watts showed under some strong assumptions that the new idea will “percolate” or “cascade” through the infinite network. More precisely, this means that the set of adopting nodes triggered by any infinitesimal fraction of seed nodes will grow to a positive fraction of the network with positive probability if and only if a certain analytical condition on the network parameters holds.

In the present paper, we no longer assume that the strengths of social interactions between friends (i.e. linked nodes) are equal and deterministic, but instead take edge weights to be random variables whose distributions may be dependent on the connectedness of the nodes. This means that if ℓ denotes a social link between two friends who have friendship degrees k, k' (that is, these friends each have in total k and k' friends respectively), we take the strength of their friendship to be w_ℓ , a random variable drawn from a distribution $G_{kk'}$ that depends on k, k' . This new type of randomness reflects a genuine uncertainty we have in our observations of real social networks.

The problem of random edge weights appears to be little studied in the vast literature that has built up around the Watts Cascade Model. However, the theory has moved substantially in other directions and promising new areas of application have been found in Watts and Dodds [2007], Dodds and Payne [2009], and Dodds et al. [2012]. These developments are reviewed in Newman [2010]. The Watts percolation theory itself has been extended to far more general random graph ensembles, for example in Boguñá and Serrano [2005].

Our motivation to study the Watts model in this way comes from a relatively new thread of economic research into “systemic risk” in financial networks. This term refers to financial system crises that involve domino-type cascades of defaulting or insolvent banks, triggered by an initial shock. Eisenberg and Noe [2001] provided a highly stylized network picture of a system of banks with balance sheets that distinguish external and interbank lending and borrowing. In this network, each directed edge between banks has a weight that denotes the size of an interbank loan. An initial shock to the balance sheets of the system may cause the insolvency of some banks. Insolvent bank assets are liquidated at less than their book value, causing balance sheet shocks to their creditor banks. Such shocks may lead to secondary insolvencies: understanding the resultant “insolvency cascade” is the focus of systemic risk research. Nier et al. [2007], in an influential paper, provided a random specification of this type of system, and through Monte Carlo simulations, discovered a number of non-intuitive, non-monotonic relationships between model parameters and the resultant cascade size probability distribution. Gai and Kapadia [2010] then introduced an elegant systemic risk model based on a directed graph version of Watts’ setup, and by adapting Watts percolation results, were able to determine cascade sizes by analytical means. This paper was extended in several directions in Amini et al. [2012] and Hurd and Gleeson [2011], who proved more general versions of the percolation theorems.

This new analytical theory of systemic risk hits a major roadblock whenever a

realistic specification of interbank lending is made: until now, such models have not been able to handle edge weights that are random or depend deterministically on the size of banks. It is this roadblock in understanding systemic risk that we address here. To keep everything as simple as possible, we turn away from financial systems modelled by directed networks, and focus instead on information cascades in undirected networks that is the standard interpretation of Watts model.

Our version of the Watts undirected cascade model allows for a lot of modelling flexibility, since the underlying network (the “skeleton graph”), the adoption thresholds, the edge weights and the initial set of “seed nodes” are all allowed to be stochastic. One might take the view that the theory is now far from parsimonious. Making use of this additional flexibility in practice comes at a cost: statistical estimation or “calibration” of the myriad parameters to any specific real world network may turn out to be prohibitively difficult. It is this concern that lead us to the second main contribution of our paper: we are able to specialize the framework to any deterministic finite skeleton graph, thereby reducing complexity, whenever one accurately observes the nodes and links (as one does for example in Facebook networks, Traud et al. [2011]). The resultant picture is then of a known “real world network”, but with randomness in its node thresholds, edge weights, and initial shocks that directly reflects our remaining uncertainties. As we will see in this paper, the Watts type cascade analysis extends naturally to this more realistic setting.

Modelling real world networks directly in this way contrasts with prior work on real world networks that makes use of the infinite network analysis, Melnik et al. [2011]. Typically, these papers start with network data thought of as an empirical probability distribution and as a first step, match this empirical distribution to an estimated distribution, taken from a parametric family of infinite random graph models. The cascade analysis is then applied to this estimated infinite graph model. Finally, the resultant cascade probabilities are taken as indicators of the cascade risk inherent in the original real world network. The conclusions of such an indirect approach will likely be weak, since the structure of a given real observed network is often very unlike any of the standard families of infinite random graph models.

By comparison, our probabilistic cascade analysis works directly on the empirical finite network. Whenever the analytic hypotheses are verified, or shown to have insignificant impact, we obtain measures of cascade risk that directly reflect the actual topology and structure of the real world network. We can then hope that our predicted cascade measures will accurately reflect reality. In this paper we will see whether this hope is justified by comparing our analytical formulas to Monte Carlo estimates under a range of modelling specifications.

During this work certain subtleties concerning the dependence amongst random variables lead us to reevaluate the logical and probabilistic arguments for cascading that are typically invoked in random graph theory. In the literature, for example Newman [2010], such arguments have rested heavily on the use of generating functions and underplayed the role of conditional probability. The logical framework is certainly elegant and efficient. However, we have found it preferable to use a more direct and

transparent probabilistic style of argument initiated in Hurd and Gleeson [2011]. It would be a demanding exercise to translate our probabilistic analysis into the more familiar generating function language!

Our analytical results only hold under an important and restrictive assumption that generalizes a property of the random configuration graph model of Bollobás [1980] that was the setting of Watts’ original paper. Called the locally tree-like independence assumption (LTIA) in Hurd and Gleeson [2011], this is the formal equivalent to the so-called “mean-field” assumptions often made in the literature. It holds in the infinite graph limit in the class of configuration models, but fails to hold in many other important classes of random graph models. The second technical innovation of our paper, the WOR condition, makes sense only under the LTIA assumption. It provides the key to finding the somewhat subtle unravelling of dependent random variables needed to prove Theorem 1.

The flexibility of our analytical framework opens up a lot of questions for study. Space permits us to address only a few: Do we observe qualitatively new network effects such as phase transitions, due to the randomness of edge weights? Do results from the infinite graph setting accurately reflect what is observed in finite networks? When does the LTIA assumption lead to accurate results for real world networks? We hope readers will find the methods developed here compelling enough to program for themselves, and invite them to use such methods as a computer “microscope” to investigate further detailed properties of random networks.

2 The Extended Watts Cascade Model

The famous paper Watts [2002] introduces a model of a social network (network of “friendships”): the problem is to study if and how a newly introduced technology “percolates” through the network. Such information cascades can be viewed as a model for the spread of ideas, rumours, a social fad, or any number of analogous phenomena. This paper has important antecedents in the literature, notably Granovetter [1978] and Schelling [1973] (see also Easley and Kleinberg [2010] and references therein). The setup of the original model starts with a “skeleton graph” on N nodes labelled by $\mathcal{N} = \{1, 2, \dots, N\}$ (where N may be infinite), that is, an undirected random graph $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$. Each undirected link of the graph represents the existence of a friendship connection between two individuals or nodes. For the next layer of model structure, each node $v \in \mathcal{N}$ is assigned a random threshold $\Phi_v \in \mathbb{R}$. At the top level, a random subset $\mathcal{A}_0 \in \mathcal{N}$ of “early adopter” nodes is selected that represent those individuals that adopt the new technology before being influenced by friends. For each realization of the above set of random variables, the cascade describing the adoption of technology by nodes proceeds through a sequence of deterministic steps. At each step, a node v “adopts” if at least Φ_v of its neighbours were in “adopted” state in the previous step. The primary aim of the model is to characterize the probability of nodes eventually adopting the technology, and in particular to determine the “phases” in which the new technology will be taken up by a large fraction of the

network.

There are many possible generalizations of the Watts model. In this paper we focus on the particular generalization where the edges ℓ are given random weights w_ℓ that describe the degree of friendship (or influence) between friends. The random nature of these weights reflects the uncertainty we have in knowing the strength of friendships in a network. One is now interested in questions such as:

- Is there a major effect on the cascade if w_ℓ has fat tails (such as a Pareto distribution)?
- Does the percolation picture of cascades survive in this generalization?
- Are the cascade results computable in closed form, or must they be replaced by numerical techniques?
- Do random edge weights lead to interesting new phase structures as model parameters are varied?

The full specification of the extended model consists of the following ingredients.

1. The skeleton graph \mathcal{E} is a random Bollobás configuration graph with a specified degree distribution $P_k = \mathbb{P}[\deg(v) = k]$, $k \in \mathcal{K}$. For simplicity, we assume that \mathcal{K} is a finite set of non-negative integers. We also let $z = \sum_k kP_k$ denote the mean degree.
2. The random variables $\Phi_v \in \mathbb{R}^+$ have cumulative distribution functions (CDFs) $F_k(x) = \mathbb{P}[\Phi_v \leq x | k_v = k]$ that depend only on the degree k_v . We suppose that the set of initially adopting nodes \mathcal{A}_0 is defined by the condition $\Phi_v = 0$, which occurs with probability $A_k^{(0)} = F_k(0)$ if $k_v = k$.
3. The random weights $w_{vv'} \in \mathbb{R}^+$ for edges $(v, v') \in \mathcal{E}$ have CDFs $G_{kk'}(x)$ which depend on the degrees $k_v = k, k_{v'} = k'$.
4. The condition for a node v to have “adopted” at the n th step of the cascade, written $v \in \mathcal{A}_n$, is

$$\sum_{v' \in \mathcal{N}_v \cap \mathcal{A}_{n-1}} w_{v'v} \geq \Phi_v \tag{1}$$

for any $n > 0$.

Finally, we assume a strong form of conditional independence: conditioned on the skeleton graph \mathcal{E} , the family of random variables $\{\Phi_v, w_{vv'}\}$ for all v, v' form a mutually independent collection.

Note that the original Watts model arises by taking \mathcal{E} to be a configuration random graph with general degree distribution, $w_\ell = 1$ deterministically and $F_k(x) = A_k^{(0)} + (1 - A_k^{(0)})F(x/k)$ for a single function $F(x)$.

Our aim is now to track the evolution of the cascade through the increasing sequence of sets of nodes:

$$\mathcal{A}_0 \rightarrow \mathcal{A}_1 \rightarrow \dots \rightarrow \mathcal{A}_n \rightarrow \dots$$

and to determine the probabilities that a node of given degree eventually adopts:

$$A_k^{(\infty)} = \lim_{n \rightarrow \infty} \mathbb{P}[v \in \mathcal{A}_n | k_v = k] .$$

The mathematical analysis that follows is based on a single assumption we will make throughout this paper.

The “locally tree-like” independence assumption (LTIA): Consider an extended Watts model on \mathcal{N} defined by a collection of random variables $(\mathcal{E}, \{\Phi\}, \{w\})$. Let $\mathcal{N}_1, \mathcal{N}_2 \subset \mathcal{N}$ be any two subsets that share exactly one node $\mathcal{N}_1 \cap \mathcal{N}_2 = \{v\}$ and let X_1, X_2 be any pair of random variables where for each $i = 1, 2$, X_i is determined by the information on \mathcal{N}_i . Then, conditioned on information located at the node v , that is conditioned on Φ_v or k_v , X_1 and X_2 are independent.¹

LTIA is a concise formalization of conditions that appear extensively in the literature under the name of “mean field” approximations. In our setting, two distinct model properties are captured by LTIA:

1. The random graph \mathcal{E} should be tree-like in the sense that the density of triangles and higher order cycles is zero. This is true if \mathcal{E} is an actual finite tree (if N is finite), or if it is drawn from an infinite Bollobás configuration model.
2. The collection of random variables $\Phi_v, w_{vv'}$ form a mutually independent collection, conditioned on the random skeleton graph \mathcal{E} .

When the LTIA condition is exactly true, the analytical results we present here are exact. In practice, in a model where LTIA is not true, our analytical results are an approximation to the true behaviour in the model.

Certain mathematical considerations arise in the following analysis. First, one needs to know that the degree distribution of v conditioned on v being the end-node of an edge is not P_k but

$$Q_k = \mathbb{P}[k_v = k | v \in \mathcal{N}_{v'} \text{ for some } v'] = \frac{kP_k}{z}, k \in \mathcal{K}$$

where $z = \sum_k kP_k$ is the mean degree.

The second consideration is to have a compact representation for the probability that $X \geq Y$ for two independent non-negative random variables. Let $\langle \cdot, \cdot \rangle$ denote the standard Hermitian inner product on functions on $(-\infty, \infty)$:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} \bar{f}(x)g(x)dx .$$

¹The following is a concise mathematical expression of this independence assumption. To any subset of nodes $\mathcal{N}' \subset \mathcal{N}$ we associate the sigma-algebra \mathcal{G}' generated by the thresholds Φ_v and degrees k_v of nodes $v \in \mathcal{N}'$ and edges (v, v') in $\mathcal{N}' \times \mathcal{N}'$. If we let $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_v$ denote the corresponding sigma-algebras, then \mathcal{G}_1 and \mathcal{G}_2 are independent sigma-algebras conditioned on \mathcal{G}_v : that is, $(\mathcal{G}_1 \perp\!\!\!\perp \mathcal{G}_2) | \mathcal{G}_v$.

Assuming that X has a density $f_X(x)$ and the CDF F_Y of Y is continuous, and letting these functions have Fourier transforms $\mathcal{F}(f_X), \mathcal{F}(F_Y)$, then by the Parseval Identity

$$\mathbb{P}[X \geq Y] = \int_0^\infty f_X(x)F_Y(x)dx = \langle F_Y, f_X \rangle = \frac{1}{2\pi} \langle \mathcal{F}(F_Y), \mathcal{F}(f_X) \rangle. \quad (2)$$

The last of these formulas will be useful in computing such probabilities when X is itself a sum of independent random variables $X = X_1 + X_2$ because whereas $f_X(x)$ is a convolution

$$f_X(x) = \int_0^x f_{X_1}(y)f_{X_2}(x-y)dy ,$$

its Fourier transform is a simple product, $\mathcal{F}(f_X) = \mathcal{F}(f_{X_1})\mathcal{F}(f_{X_2})$.

The main result of this section is a theorem that determines the double sequence of probabilities

$$A_k^{(n)} = \mathbb{P}[v \in \mathcal{A}_n | k_v = k]$$

beginning with the initial vector $A_k^{(0)}, k \in \mathcal{K}$. The theorem is inductive on the steps $n = 0, 1, 2, \dots$ of the cascade. To motivate the result, we consider the defining property (1) of the set \mathcal{A}_n and note this implies

$$A_k^{(n)} = \mathbb{P}[\Phi_v \leq \sum_{v' \in \mathcal{N}_v \cap \mathcal{A}_{n-1}} w_{v'v} | k_v = k] . \quad (3)$$

The subtle point is that the random variables $\Phi_v, \mathcal{A}_{n-1}$ and $w_{vv'}$ occurring in this expression have a complicated dependence, and the probability cannot be evaluated directly using (2). The key to unravelling this dependence is a notion that takes care of the causal relations in the order of adopting nodes. We say that a node property \mathcal{P} holds for a node v “without regarding” a neighbouring node $v' \in \mathcal{N}_v$, and write “ \mathcal{P} holds for v WOR v' ”, if this property of v holds in the subgraph of \mathcal{E} obtained by deleting the edge to v' and the entire connected subgraph rooted on this edge. It is important to note that the concept “ \mathcal{P} holds for v WOR v' ” makes sense only under the LTIA assumption: otherwise with positive probability the deleted subgraph will include v itself via a cycle in the network.

We will show that (3) is expressible in terms of intermediate WOR probabilities defined by

$$C_k^{(n)} = \mathbb{P}[v \in \mathcal{A}_n \text{ WOR } v' | k_v = k, v' \in \mathcal{N}_v] \quad (4)$$

where v, v' are random nodes selected from \mathcal{N} . At the outset,

$$C_k^{(0)} = \mathbb{P}[\Phi_v = 0 | k_v = k, v' \in \mathcal{N}_v] = A_k^{(0)} . \quad (5)$$

We begin by noting that for $n > 0$, $v \in \mathcal{A}_n$ is true if and only if

$$\Phi_v \leq \sum_{v'' \in \mathcal{N}_v} \tilde{w}_{v'',v}^{(n-1)}$$

where

$$\tilde{w}_{v'',v}^{(n-1)} = w_{vv''} \mathbf{1}(v'' \in \mathcal{A}_{n-1} \text{ WOR } v) .$$

Next we observe that under the condition $k_v = k$, LTIA implies that the k random variables $\tilde{w}_{v'',v}^{(n-1)}$ for $v'' \in \mathcal{N}_v$ are independent. They are also identically distributed with CDF given by

$$\begin{aligned} \mathbb{P}[\tilde{w}_{v'',v}^{(n-1)} \leq x | k_v = k, v'' \in \mathcal{N}_v] &= \sum_{k''} \mathbb{P}[\tilde{w}_{v'',v}^{(n-1)} \leq x | k_v = k, k_{v''} = k'', v'' \in \mathcal{N}_v] Q_{k''} \\ &= \sum_{k''} \left[(1 - C_{k''}^{(n-1)}) + C_{k''}^{(n-1)} G_{kk''}(x) \right] Q_{k''} . \end{aligned}$$

In the above step, we have used LTIA again: conditioned on $k_v = k, k_{v''} = k''$, the random variables $w_{vv''}$ and $\mathbf{1}(v'' \in \mathcal{A}_{n-1} \text{ WOR } v)$ are independent. Hence it follows that

$$\begin{aligned} A_k^{(n)} &= \mathbb{P} \left[\sum_{v'' \in \mathcal{N}_v} w_{vv''} \mathbf{1}(v'' \in \mathcal{A}_{n-1} \text{ WOR } v) \geq \Phi_v | k_v = k \right] \\ &= \frac{1}{2\pi} \langle \mathcal{F}(F_k), (\mathcal{F}(g_k^{(n-1)}))^k \rangle \end{aligned} \quad (6)$$

where $\mathcal{F}(g_k^{(n-1)})$ is the Fourier transform of

$$g_k^{(n-1)}(x) = \sum_{k'} \left[(1 - C_{k'}^{(n-1)}) \delta_0(x) + C_{k'}^{(n-1)} g_{kk'}(x) \right] Q_{k'} \quad (7)$$

and δ_0 denotes the Dirac delta function centred at 0.

We now suppose the vector $\{C_k^{(n-1)}, k \in \mathcal{K}\}$ is known for some $n > 0$, and seek a formula for the vector $C_k^{(n)}$. The logic is the same as in the previous paragraph, with one difference. Given $k_v = k, v' \in \mathcal{N}_v$, the sum over v'' is over the set $\mathcal{N}_v \setminus v'$ instead of the set \mathcal{N}_v . It follows that

$$C_k^{(n)} = \frac{1}{2\pi} \langle \mathcal{F}(F_k), (\mathcal{F}(g_k^{(n-1)}))^{k-1} \rangle \quad (8)$$

with $g_k^{(n-1)}$ given again by (7).

The above arguments prove our first main result:

Theorem 1. *Consider an extended Watts model on \mathcal{N} defined by a collection of random variables $(\mathcal{E}, \{\Phi\}, \{w\})$ that satisfies LTIA. Then:*

1. *The probabilities $C_k^{(n)} = \mathbb{P}[v \in \mathcal{A}_n \text{ WOR } v' | k_v = k, v' \in \mathcal{N}_v]$ satisfy the recursive formula*

$$C_k^{(n)} = \frac{1}{2\pi} \langle \mathcal{F}(F_k), (\mathcal{F}(g_k^{(n-1)}))^{k-1} \rangle \quad (9)$$

with $F_k(x) = \mathbb{P}[\Phi_v \leq x | k_v = k]$, $g_k^{(n-1)}(x)$ given by (7). The initial values are $C_k^{(0)} = A_k^{(0)} = \mathbb{P}[\Phi_v = 0 | k_v = k]$.

2. *The probabilities $A_k^{(n)} = \mathbb{P}[v \in \mathcal{A}_n | k_v = k]$ are given in terms of $C_k^{(n-1)}$ by*

$$A_k^{(n)} = \frac{1}{2\pi} \langle \mathcal{F}(F_k), (\mathcal{F}(g_k^{(n-1)}))^k \rangle. \quad (10)$$

3. The probability that a node of degree k eventually adopts, $A_k^{(\infty)}$, is given by (10) in terms of a vector $C^{(\infty)} = \{C_k^{(\infty)}\}_{k \in \mathcal{K}}$ that satisfies the \mathbb{R}^K valued fixed point equation defined by the recursion given in part (1). The existence (but not uniqueness) of such a fixed point is guaranteed by the Tarski-Knaster Fixed Point Theorem.

Remark 2. It is interesting that nearest-neighbour correlations in this model are determined in terms of the C and A probabilities. In fact, under LTIA, the joint probability of eventual non-adoption of a nearest neighbour node pair v, v' with $k_v = k, k_{v'} = k'$ is simply:

$$\begin{aligned} \mathbb{P}[v, v' \notin \mathcal{A}_\infty | k_v = k, k_{v'} = k'] &= \mathbb{P}[v \notin \mathcal{A}_\infty \text{ WOR } v', v' \notin \mathcal{A}_\infty \text{ WOR } v | k_v = k, k_{v'} = k'] \\ &= [1 - C_{k'}^{(\infty)}][1 - C_k^{(\infty)}]. \end{aligned}$$

A result of this type is found in Gleeson and Melnik [2009].

It is well-known in network science, see e.g. Newman [2010][Chapter 13.11], that the frequency of global cascades in infinite random graphs is related to the fractional size of the so-called in-component associated to the giant vulnerable cluster. In the present context, a vulnerable cluster has the meaning of a connected subgraph of the network consisting of “vulnerable” directed edges, where a vulnerable directed edge is an edge whose weight is sufficient to exceed the adoption threshold of its forward node. We define:

- $\mathcal{E}_V \subset \mathcal{E}$, the set of vulnerable directed edges;
- \mathcal{E}_s , the giant strongly connected set of vulnerable edges (the “giant vulnerable cluster”);
- \mathcal{E}_i , the “in-component” of the giant vulnerable cluster, i.e. the set of vulnerable edges that are connected to \mathcal{E}_s by a directed path of vulnerable edges;
- $1 - b_{k'} := \mathbb{P}[(v', v) \in \mathcal{E}_i | k_{v'} = k']$, a conditional probability of an edge being in \mathcal{E}_i ;
- $a_{k'k} = \mathbb{P}[\Phi_v \leq w_{v'v} | k_{v'} = k', k_v = k]$, the conditional probability of a node being vulnerable.

Now note that $(v', v) \in \mathcal{E}_i^c$ (i.e. the complement of \mathcal{E}_i) means either $\Phi_v > w_{v'v}$ or $\Phi_v \leq w_{v'v}$ and all the $k_v - 1$ “downstream” directed edges (v, v'') are in the set \mathcal{E}_i^c . Thus, invoking the LTIA, one determines that

$$\begin{aligned} b_{k'} &= \sum_k \mathbb{P}[k_v = k | k_{v'} = k', v \in \mathcal{N}_{v'}^-] \mathbb{P}[(v', v) \in \mathcal{E}_i | k_{v'} = k', k_v = k] \\ &= \sum_k \frac{k P_k}{z} \left[(1 - a_{k'k}) + a_{k'k} b_k^{k-1} \right] := h_{k'}(\mathbf{b}) \end{aligned}$$

In other words, the vector $\mathbf{b} = \{b_k\}_{k \in \mathcal{K}}$ satisfies the fixed point equation $\mathbf{b} = \mathbf{h}(\mathbf{b})$ where

$$h_{k'}(\mathbf{b}) = \sum_k \frac{k P_k}{z} \left[(1 - a_{k'k}) + a_{k'k} b_k^{k-1} \right], \quad k' \in \mathcal{K}. \quad (11)$$

The cascade condition is the condition on a random network that determines whether or not it has a giant vulnerable cluster, or equivalently, whether a randomly selected initial adopting node has a positive probability of triggering a cascade that reaches a positive fraction of the infinite network. As shown in a similar setting in Dodds et al. [2012], Amini et al. [2012] and Hurd and Gleeson [2011], this boils down to a condition on the spectral radius $\|D\|$ of the derivative matrix $D = (D_{k'k})_{k',k \in \mathcal{K}}$. Here $D_{k'k} = \partial h_{k'}/\partial b_k$ is evaluated at the trivial fixed point $\mathbf{b} = [1, 1, \dots, 1]$, and is given by an explicit formula:

$$D_{kk'} = \frac{a_{k'k}k(k-1)P_k}{z} . \quad (12)$$

Recall that the spectral radius of a square matrix is the magnitude of its largest eigenvalue.

Cascade Condition: If $\|D\| > 1$, then with positive probability, a single random seed will generate a cascade that reaches a positive fraction of the infinite network. If $\|D\| < 1$, then with probability 1, the cascade generated by a single random seed will be a zero fraction of the infinite network.

For a global cascade to arise from a random single seed, it is sufficient for the seed v to trigger at least one edge $(v, v') \in \mathcal{E}_i$. Therefore, by conditioning on the degree of v one finds that the frequency of global cascades is bounded below:

$$f \geq \sum_{k>0} [1 - b_k^k] P_k .$$

3 Discrete Probability Distributions

The structure of (9) and similar equations is problematic from the point of view of numerical approximations. Numerical evaluation of the implied integrals leads to truncation errors and discretization errors, both of which will be difficult to handle in our setting. In this section, we analyze the case where the random variables $\{\Phi_v, w_\ell\}$ all take values in a specific finite discrete set $\mathcal{M} = \{0, 1, \dots, (M-1)\}$ with a large value M . In such a situation, the convolutions in (6) can be performed exactly and efficiently by use of the discrete Fast Fourier Transform (FFT).

Let X, Y be two independent random variables with probability mass functions (PMF) p_X, p_Y taking values on the non-negative integers $\{0, 1, 2, \dots\}$. Then the random variable $X + Y$ also takes values on this set and has the probability mass function (PMF) $p_{X+Y} = p_X * p_Y$ where the convolution of two functions f, g is defined to be

$$(f * g)(n) = \sum_{m=0}^n f(m)g(n-m) . \quad (13)$$

Note that p_{X+Y} will not necessarily have support on the finite set \mathcal{M} if p_X, p_Y have support on \mathcal{M} . This discrepancy leads to the difficulty called ‘‘aliasing’’ in the next section.

3.1 The Discrete Fourier Transform

In this section we consider space \mathbb{C}^M of \mathbb{C} -valued functions on $\mathcal{M} = \{0, 1, \dots, M-1\}$. The discrete Fourier transform, or fast Fourier transform (FFT), is the linear mapping $\mathcal{F} : a = [a_0, \dots, a_{M-1}] \in \mathbb{C}^M \rightarrow \hat{a} = \mathcal{F}(a) \in \mathbb{C}^M$ defined by

$$\hat{a}_k = \sum_{l \in \mathcal{M}} \zeta_{kl} a_l, k \in \mathcal{M}.$$

where the coefficient matrix $Z = (\zeta_{kl})$ has entries $\zeta_{kl} = e^{-2\pi i k l / M}$.

It is easy to prove that its inverse, the ‘‘inverse FFT’’ (IFFT), is given by the map $a \rightarrow \tilde{a} = \mathcal{G}(a)$ where

$$\tilde{a}_k = \frac{1}{M} \sum_{l \in \mathcal{M}} \bar{\zeta}_{kl} a_l, k \in \mathcal{M}.$$

If we let \bar{a} denote the complex conjugate of a , we can define the Hermitian inner product between

$$\langle a, b \rangle := \sum_{m \in \mathcal{M}} \bar{a}_m b_m.$$

We also define the convolution product of two vectors:

$$(a * b)(n) = \sum_{m \in \mathcal{M}} a(m) b(n - m \text{ modulo } M), \quad n \in \mathcal{M}.$$

Now we note the following easy-to-prove identities which hold for all $a, b \in \mathbb{C}^M$:

1. Inverse mappings:

$$a = \mathcal{G}(\mathcal{F}(a)) = \mathcal{F}(\mathcal{G}(a));$$

2. Conjugation:

$$\overline{\mathcal{G}(a)} = \frac{1}{M} \mathcal{F}(\bar{a});$$

3. Parseval Identity:

$$\langle a, b \rangle = M \langle \tilde{a}, \tilde{b} \rangle = \frac{1}{M} \langle \hat{a}, \hat{b} \rangle;$$

4. Convolution Identities:

$$\tilde{a} \cdot * \tilde{b} = \widetilde{(a * b)}; \quad \hat{a} \cdot * \hat{b} = \widehat{(a * b)}$$

where $\cdot *$ denotes the component-wise product.

We note that if the support of the sum of two \mathcal{M} -valued random variables X, Y is itself in \mathcal{M} , that is, if $\text{supp}(X + Y) = \{n | \exists m \in \text{supp}(X) \text{ s.t. } n - m \in \text{supp}(Y)\} \in \mathcal{M}$, then p_{X+Y} , the PMF of $X + Y$, is given by $p_X * p_Y$. On the other hand, if there are $m \in \text{supp}(X), n \in \text{supp}(Y)$ such that $m + n \geq M$, then $p_{X+Y} - p_X * p_Y$ is not zero: such a difference is called an ‘‘aliasing error’’. In what follows we assume that all aliasing errors can be made zero by choosing M sufficiently large but finite.

We now consider the probability

$$P = \mathbb{P}[Y \leq \sum_{i=1}^k X_i].$$

for independent random variables Y, X_1, X_2, \dots, X_k where X_i are distributed with PMF $g_i(m)$ and Y has PMF $f(m)$. We suppose that f has support on \mathcal{M} while each g_i has support on $\{0, 1, \dots, \lfloor M - 1/k \rfloor\}$. We also define the CDF $F(n) = \sum_{m=0}^n f(m) = \mathbb{P}(Y \leq n)$ for $n \in \mathcal{M}$. Then using the above identities, we are lead to an efficient means to compute P for large k involving matrix operations and the FFT:

$$\begin{aligned} &= \mathbb{P}[Y \leq \sum_{i=1}^k X_i] = \sum_{m \in \mathcal{M}} (g_1 * \dots * g_k)(m) \sum_{n=0}^m f(n) = \sum_{m \in \mathcal{M}} (g_1 * \dots * g_k)(m) F(m) \\ &= \langle F, g_1 * \dots * g_k \rangle = \frac{1}{M} \langle \hat{F}, \widehat{(g_1 * \dots * g_k)} \rangle = \frac{1}{M} (\hat{F})' * (\hat{g}_1 * \dots * \hat{g}_k). \end{aligned} \quad (14)$$

Here in the final expression, A' denotes the conjugate transpose of a matrix A , $*$ denotes matrix multiplication between a size $[1, M]$ matrix and a size $[M, 1]$ matrix, and $*$ denotes component-wise multiplication of vectors and matrices.

Remark 3. *There is no aliasing problem and the identity (14) is true if $g_1 * \dots * g_k$ has support on \mathcal{M} . We formalise this requirement as the Aliasing Assumption.*

Remark 4. *The overall computation time for any numerical implementation of the cascade formulas in Theorem 1 will be dominated by computing convolutions such as those in (10). In particular, one can check that negligible total time will be taken in computing FFTs (each of which take of the order of $M \ln M$ additions and multiplications). One can compare the efficiency of our recommended FFT approach to the direct approach by considering a single evaluation of the convolution: to compute $f * g$ using (13) for M -vectors f, g requires M^2 additions and multiplications, whereas to compute $\hat{f} * \hat{g}$ requires only M multiplications. All else being equal, we can expect a speedup by a factor of $O(M)$ when the FFT method is used instead of the direct method. Since in our implementations we often have $M \gtrsim 2^{10}$, this is a huge improvement. Of course, this factor is too optimistic: it will be worsened considerably when one takes care of the aliasing problem.*

4 Fixed Random Graphs

We now consider how the formalism of Section 2 can be applied in an approximate fashion in the important special class of problems when the skeleton graph is actually known (i.e. deterministic) and finite while the thresholds and edge weights are random. This will allow us to address the question of systemic risk in tractable models of real observed financial networks.

Let $A = (A_{vv'}), v, v' \in \mathcal{N}$ be the symmetric adjacency matrix of the fixed undirected graph \mathcal{E} . We number the nodes in \mathcal{N} by $v = 1, 2, \dots, N$ and the links in \mathcal{E} by $\ell = 1, 2, \dots, L$ where $L = \sum_{1 \leq v < v' \leq N} A_{vv'}$. The discrete threshold and edge weight random variables Φ_v, w_ℓ are assumed to have support on the set $\mathcal{M} = \{0, 1, \dots, M - 1\}$. The probability that node v is an early adopter is $A_v^{(0)} = \mathbb{P}[\Phi_v = 0] = \mathbb{P}[v \in \mathcal{A}_0]$. The CDF of Φ_v is an M -vector $F_v(m) = \mathbb{P}[\Phi_v \leq m]$ for $m \in \mathcal{M}$, and we let $\hat{F}_v = \mathcal{F}(F_v)$ denote its discrete Fourier transform (FFT). Similarly, $\hat{g}_\ell(m), m \in \mathcal{M}$ denotes the FFT of the M -vector $g_\ell(m) = \mathbb{P}[w_\ell = m]$, the probability mass function of w_ℓ . To ensure LTIA, it is necessary but not sufficient that $\{\Phi_v, w_\ell\}$ is a collection of mutually independent random variables. The LTIA is exact only when \mathcal{E} is a tree graph or if the edge weights and thresholds are deterministic.

The aim of this section is to use LTIA to derive an approximate formula for the marginal likelihoods $A_v^{(\infty)} := \mathbb{P}[v \text{ eventually adopts}]$. In the general case, we expect that the approximation that takes the dependence between different branches of the random graph to be negligible should improve as N gets large, all else being equal. We will find that the dependence of the accuracy on other parameters is not easy to understand, but that overall assuming LTIA leads to surprisingly good performance.

The analysis will determine the required probabilities $A_v^{(n)} = \mathbb{P}[v \in \mathcal{A}_n]$ via the intermediate probabilities

$$C_{v,v'}^{(n)} = \mathbb{P}[v \in \mathcal{A}_n \text{ WOR } v']$$

that will be computed for each of the $2L$ directed links $\ell = (v, v')$ and $n = 1, 2, \dots$. At the outset, $C_{v,v'}^{(0)} = A_v^{(0)}$.

The analysis for a finite real graph exactly parallels the argument presented in Section 2. First we note that $v \in \mathcal{A}_n$ for $n > 0$ means that

$$\sum_{v'' \in \mathcal{N}_v} \tilde{w}_{v'',v} \geq \Phi_v$$

where

$$\tilde{w}_{v'',v} := w_{vv''} \mathbf{1}(v'' \in \mathcal{A}_{n-1} \text{ WOR } v).$$

By LTIA, we assume that $w_{vv''}$ and $\mathbf{1}(v'' \in \mathcal{A}_{n-1} \text{ WOR } v)$ are independent for all $v'' \in \mathcal{N}_v$. Therefore, the sum of the \tilde{w} over v'' is a sum of independent random variables, each with moment generating function (MGF)²

$$\hat{g}_{v'',v}^{(n-1)}(m) = (1 - C_{v'',v}^{(n-1)}) + C_{v'',v}^{(n-1)} \hat{g}_{v'',v}(m), m \in \mathcal{M}. \quad (15)$$

From these facts, and (14), and making use of the Aliasing Assumption (Remark 3), we deduce the formula for $A_v^{(n)}$:

$$A_v^{(n)} = \mathbb{P}\left[\sum_{v'' \in \mathcal{N}_v} \tilde{w}_{v'',v} \geq \Phi_v\right] = \left\langle \hat{F}_v, \prod_{v'' \in \mathcal{N}_v} \hat{g}_{v'',v}^{(n-1)} \right\rangle. \quad (16)$$

²Note: we have used the fact that the MGF of the random variable 0 is the M -vector $[1, 1, \dots, 1]$.

When it comes to computing $C_{v,v'}^{(n)}$ in terms of the set $\{C^{(n-1)}\}$, we may repeat this argument. The only change is in the sum over v'' : it now runs over $\mathcal{N}_v \setminus \{v'\}$ not \mathcal{N}_v . It follows that

$$C_{v,v'}^{(n)} = \left\langle \hat{F}_v, \prod_{v'' \in \mathcal{N}_v \setminus \{v'\}} \hat{g}_{v'',v}^{(n-1)} \right\rangle. \quad (17)$$

The above argument rests on two assumptions, and leads directly to the main result concerning real world graphs:

Theorem 5. *Consider an extended Watts model on \mathcal{N} on a fixed finite graph \mathcal{E} with N nodes and L edges. Let the thresholds and edge weights be a collection of independent random variables $\{\Phi_v, w_\ell\}$ supported on $\mathcal{M} = \{0, 1, \dots, M-1\}$. Assuming LTIA and the Aliasing Assumption (Remark 3), the following formulas hold:*

1. *For each $n > 0$ and directed edge v, v' , the probabilities $C_{v,v'}^{(n)} = \mathbb{P}[v \in \mathcal{A}_n \text{ WOR } v']$ satisfy the recursive formula*

$$C_{v,v'}^{(n)} = \left\langle \hat{F}_v, \prod_{v'' \in \mathcal{N}_v \setminus \{v'\}} \hat{g}_{v'',v}^{(n-1)} \right\rangle$$

with $F_v(m) = \mathbb{P}[\Phi_v \leq m]$ and $\hat{g}_{v'',v}^{(n-1)}(x)$ given by (15). The initial values are $C_{v,v'}^{(0)} = A_v^{(0)} = \mathbb{P}[\Phi_v = 0]$.

2. *The probabilities $A_v^{(n)} = \mathbb{P}[v \in \mathcal{A}_n]$ are given in terms of the collection $C_\ell^{(n-1)}$ by (16).*
3. *The probability that a node v eventually adopts, $A_v^{(\infty)}$, is given by (16) in terms of the $2L$ -vector $C^{(\infty)} = \{C_{v,v'}^{(\infty)}\}$ that satisfies the \mathcal{M}^{2L} -valued fixed point equation defined by the recursion given in part (1). The existence (but not uniqueness) of such a fixed point is guaranteed by the Tarski-Knaster Fixed Point Theorem.*

Remark 6. Caveats: *In practice, we can arrange for the Aliasing Assumption to be approximately true by choosing M large enough that the probabilities $\mathbb{P}[\sum_{v'' \in \mathcal{N}_v} \tilde{w}_{v'',v} \geq M]$ are kept sufficiently small: aliasing errors of this order of smallness can thus be expected to appear in any implementation of our formulas. However, we cannot generally arrange for LTIA to be even approximately true. The outstanding difficulty in understanding the true nature of the adoption cascade will therefore be in understanding the biases and errors that relate the LTIA formulas to the “true values” that can only be estimated by slow Monte Carlo simulations of the real world network.*

The percolation argument of Section 3 has a straightforward analogue for real world graphs that satisfy the LTIA approximately. As before we define:

- vulnerable directed edges $\ell = (v', v)$ by the condition $\Phi_v \leq W_{v',v}$, \mathcal{E}_V to be the set of such edges, and $a_\ell = \mathbb{P}[\ell \in \mathcal{E}_V]$;

- \mathcal{E}_s , a strongly connected cluster of vulnerable edges;
- \mathcal{E}_i , the “in-component” of \mathcal{E}_s ;
- $1 - b_\ell = \mathbb{P}[\ell \in \mathcal{E}_i]$.

Treating LTIA now as an approximation, one can then show that

$$b_{v'v} \sim a_{v'v} \prod_{v'' \in \mathcal{N}_v \setminus v'} b_{vv''} + (1 - a_{v'v}).$$

That is, the sequence $\mathbf{b} = \{b_{v'v}\}$ is an approximate solution of a fixed point equation $\mathbf{b} = \mathbf{h}(\mathbf{b})$. Direct differentiation at the trivial fixed point $\mathbf{b} = [1, 1, \dots, 1]$ yields the result

$$D_{\ell, \ell'} = \begin{cases} a_\ell & \text{if } \ell' \in \mathcal{E}_\ell \setminus \ell \\ 0 & \text{else} \end{cases}$$

When the spectral radius of D is greater than 1 we can expect to see large cascades from the initial random seed with frequency approximately

$$f \geq \sum_{v \in \mathcal{N}} \left[1 - \prod_{v' \in \mathcal{N}_v} b_{vv'} \right] A_v^{(0)}.$$

5 Numerical Investigations

In this section, we report results of some numerical experiments we have made testing the above analysis. We divided our experiments into those that addressed the infinite random graphs of Section 2, and those that addressed real world graphs as described in Section 4. In our numerical implementations, all random variables are taken to be either deterministic or discretized lognormals. The lattice spacing δx and lattice size M were always chosen carefully to reduce any bias between the Monte Carlo and analytic implementations.

5.1 Experiment 1: Infinite Graphs

For simplicity, we considered a Poisson random graph model as used in the experiments of Watts [2002], with Poisson degree distributions parametrized by $z \in [0, 10]$ and threshold CDFs

$$F_k(x) = A_k^{(0)} + (1 - A_k^{(0)})\mathbf{1}(x \geq \phi_* k_v)$$

with $\phi_* = 0.18$ and $A_k^{(0)} = 10^{-4}$. In Experiment 1A, the only change was to allow the weight of an edge to depend deterministically on the degrees of its nodes, i.e. its “type”, while keeping the average edge weight to be 1. We treated the original specification $w_\ell = 1$ as the benchmark case, and two additional cases where the weight $\bar{w} = \bar{w}(k, k', \alpha)$ of an edge of type kk' is proportional to $(kk')^{-\alpha}$ for some

α . The first alternative with $\alpha = -0.5$ corresponds to a social network in which the strength of edges increases as the connectivities k, k' increase. The second alternative with $\alpha = 0.5$ corresponds to a social network in which the edge strength of nodes decreases as the connectivities k, k' increases. The case $\alpha = 0$ recovers the original benchmark case.

Figure 1 shows the results. First we note that for the default case $\alpha = 0$, the curve replicates the results shown in Figure 2b of Watts [2002]. The case with $\alpha < 0$ does not differ dramatically from the constant weight case. Both show a lower cascade threshold very near to $z = 1$.

In contrast, in the case with $\alpha > 0$, where highly connected nodes have relatively less influence on their friends, we see clearly that the cascade window is shifted to higher values of z . The upward shift in the lower cascade threshold away from $z = 1$ is somewhat surprising, and begs an intuitive explanation. The expected transition near $z = 1$ for $\alpha \leq 0$ coincides almost exactly with the percolation point at $z = 1$ for Erdős-Renyi (ER) random graphs because, as one can check, for these values of α and $z \sim 1$ the vulnerable edge condition for \mathcal{S}_V is satisfied with high probability for all nodes. By contrast, when $\alpha = 0.5$ and $z \sim 1$, there is a significant positive probability for non-vulnerable edges of low degree, and consequently a significant upward shift in the transition point.

In Experiment 1B, we explored the effect of additional randomness, by taking the random variables w_ℓ to be lognormally distributed, with means $\bar{w} = \bar{w}(k, k', \alpha)$ matching Experiment 1A and standard deviation $0.5 \times \bar{w}$, whenever the type of ℓ is (k, k') . Surprisingly, the effect of randomness is counter to our expectation that the critical cascade region would shift to lower z : as Figure 2 shows, the opposite behaviour occurs.

In both Experiments 1A and 1B we also compared the $N = \infty$ analytical results to large scale Monte Carlo estimates computed for $N = 10^5$, Figures 1 and 2 show the expected excellent agreement.

In Experiment 1C, we observed the dependence of the average cascade size $\bar{A}^{(\infty)}$ on z and ϕ when the edge weights were lognormal with mean and standard deviation $(1, 0.5)$. We can see in Figure 3 that a jump discontinuity in $\bar{A}^{(\infty)}$ arises as a function of ϕ , as z increases above 3.5. In the next section we will observe that this discontinuity is strongly associated with discrepancies between the LTIA analytics and the true values for finite real world networks.

5.2 Experiment 2: Finite Erdős-Renyi Graphs

The aim of Experiment 2 was to investigate how the analytical formulas of Theorem 5, proved under the LTIA assumption, compare to the results of Monte Carlo simulations, on specific real world networks. In Experiment 2A we considered twelve Erdős-Renyi skeleton graphs, of three sizes $N = 10, 100, 1000$ and four different values of the connectivity parameter $z = 1.5, 3.5, 5.5, 7.5$. In each case, the default seed was a random set of $N/10$ nodes, and the weights were log-normally distributed with

mean and standard deviation $(1, 0.5)$. The threshold distribution for a node of degree k was log-normal with mean and standard deviation $(\phi k, \phi k/2)$ for variable ϕ .

In each of the three subplots of Figure 4, the four curves show the analytical values of mean cascade size as functions of the threshold parameter ϕ , for the four different values of z . The three subplots are for the three different values of N . On each figure, the circles and squares show the corresponding Monte Carlo values computed with 10000 realizations.

The Monte Carlo values shown on these figures give a very clear demonstration of how the $N \rightarrow \infty$ limit as shown in Figure 3 is approached through finite Erdős-Renyi graphs. One sees little discrepancy when $N = 1000$, while for smaller N the discontinuities in ϕ are smoothed out over a wider range of ϕ . When one compares analytic results derived from Theorem 5 to the Monte Carlo values, one finds discrepancies over a wide range of ϕ when $N = 10$. However, for $N = 1000$, the range over which a significant discrepancy occurs has narrowed considerably, and is confined to the range where the cascade size varies rapidly with ϕ . Further numerical experimentation shows that this type of behaviour is quite typical.

Experiment 2B tested whether the LTIA analytic formula can accurately determine adoption probabilities on a node-by-node basis. The four panels of Figure 5 show results for a single Erdős-Renyi skeleton graph with $N = 100$ and $z = 3.5$ and four different threshold parameter values $\phi = 0.3, 0.4, 0.5, 0.6$. The panels plot the node-by-node $A_v^{(\infty)}$ probabilities as computed by both analytic (red crosses) and Monte Carlo methods (blue pluses), in increasing order of their analytic values. Taken together, they back up our general conclusion that the LTIA analytics show reasonable accuracy on a node-by-node basis, and that accuracy improves as one moves away from medium cascade regions. The green circles represent the node degree as a fraction of the maximum node degree: one notices that there is very little indication that node degree is significantly correlated with adoption probability.

Remark 7. *We also ran consistency checks to verify that the Monte Carlo estimated cascades were unbiased estimators of the LTIA analytic cascades for a wide range of finite tree graphs. We created graphs similar to Figure 5 that in all cases showed the almost perfect agreement one expects for tree graphs. These results are not shown here.*

6 Concluding Remarks

The main result of this paper, the Cascade Theorem 1 for models with random edge weights, is not a straightforward extension of original Watts Model. Unlike earlier versions of the theory, the inductive derivation of Theorem 1 requires a subtle decoupling of the dependence of edge and threshold random variables. The key to this decoupling is the introduction of both the “locally tree-like assumption” (LTIA) and the “without regarding” (WOR) condition. We believe that this particular argument is new in the literature. For simplicity we have confined the discussion to information

cascades: it is clear that the ideas extend to the more complex models of cascades in financial systems.

We have shown both in infinite networks and in finite real world networks that the analytics of Theorems 1 and 5 give correct answers where appropriate, and are a meaningful approximation where the LTIA assumption fails to hold. We have used these results to look at the effects random edge weights have on information cascades, and have found evidence of interesting transitions and “tipping points” deserving of further study. These theorems give us the possibility to explore the totality of models of this general type with efficient computer algorithms. Although our primary motivations come from the models of Nier et al. [2007] and Gai and Kapadia [2010] for systemic risk in financial networks, it is possible that our results are applicable to a wide range of network phenomena.

7 Acknowledgements

This work was funded by awards from Science Foundation Ireland (11/PI/1026 and MACSI 06/MI/005), from the FET-Proactive project PLEXMATH (grant no. 317614) funded by the European Commission, and from the Natural Sciences and Engineering Research Council of Canada. We acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities. We are indebted to our summer research intern, Lionel Cassier of École Polytechnique France, who developed much of the preliminary numerical code that underlies the graphics in this work.

References

- H. Amini, R. Cont, and A. Minca. Stress Testing the Resilience of Financial Networks. *Int. J. Theor. App. Fin.*, 15:1–20, 2012.
- M. Boguñá and M. A. Serrano. Generalized percolation in random directed networks. *Phys. Rev. E*, 72:016106, 2005.
- B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Comb.*, 1:311, 1980.
- P. S. Dodds and J. L. Payne. Analysis of a threshold model of social contagion on degree-correlated networks. *Phys. Rev. E*, 79:066115, 2009.
- P. S. Dodds, K. D. Harris, and C. M. Danforth. Limited imitation contagion on random networks: Chaos, universality, and unpredictability. arXiv: 1208.0255, 2012.

- P. S. Dodds, K. D. Harris, and J. L. Payne. Direct, physically motivated derivation of the contagion condition for spreading processes on generalized random networks. *Phys. Rev. E*, 83:056122, 2011.
- D. Easley and J. Kleinberg. *Networks, Crowds and Markets*. Oxford University Press, Oxford, 2010.
- L. Eisenberg and T. H. Noe. Systemic risk in financial systems. *Management Sci.*, 47(2):236–249, 2001.
- P. Erdős and A. Rényi. On random graphs. *I. Publ. Math. Debrecen*, 6:290–297, 1959.
- P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61, 1960.
- P. Gai and S. Kapadia. Contagion in financial networks. *Proc. Roy. Soc. A*, 466(2120):2401–2423, 2010.
- J. P. Gleeson and S. Melnik. Analytical results for bond percolation and k -core sizes on clustered networks. *Phys. Rev. E*, 80:046121–046133, 2009.
- M. Granovetter. Threshold Models of Collective Behavior. *Am. J. Soc.*, 83:1420-1443, 1978.
- T. R. Hurd and J. P. Gleeson. A framework for analyzing contagion in banking networks. arXiv:1110.4312 [q-fin.GN], October 2011.
- S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson. The unreasonable effectiveness of tree-based theory for networks with clustering. *Phys. Rev. E*, 83:036112–036123, 2011.
- M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.
- E. Nier, J. Yang, T. Yorulmazer, and A. Alentorn. Network Models and Financial Stability. *J. Econ. Dyn. Control*, 31:2033–2060, 2007.
- T. C. Schelling. Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices With Externalities. *J. Conflict Resolution*, 17:381-428, 1973.
- A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter. Comparing community structure to characteristics in online collegiate social networks. *SIAM Rev.*, 53:526–543, 2011.
- D. J. Watts and P. S. Dodds. Influentials, networks, and public opinion formation. *J. Consumer Res.*, 34:441–458, 2007.
- D. J. Watts. A simple model of global cascades on random networks. *PNAS*, 99(9):5766–5771, 2002.

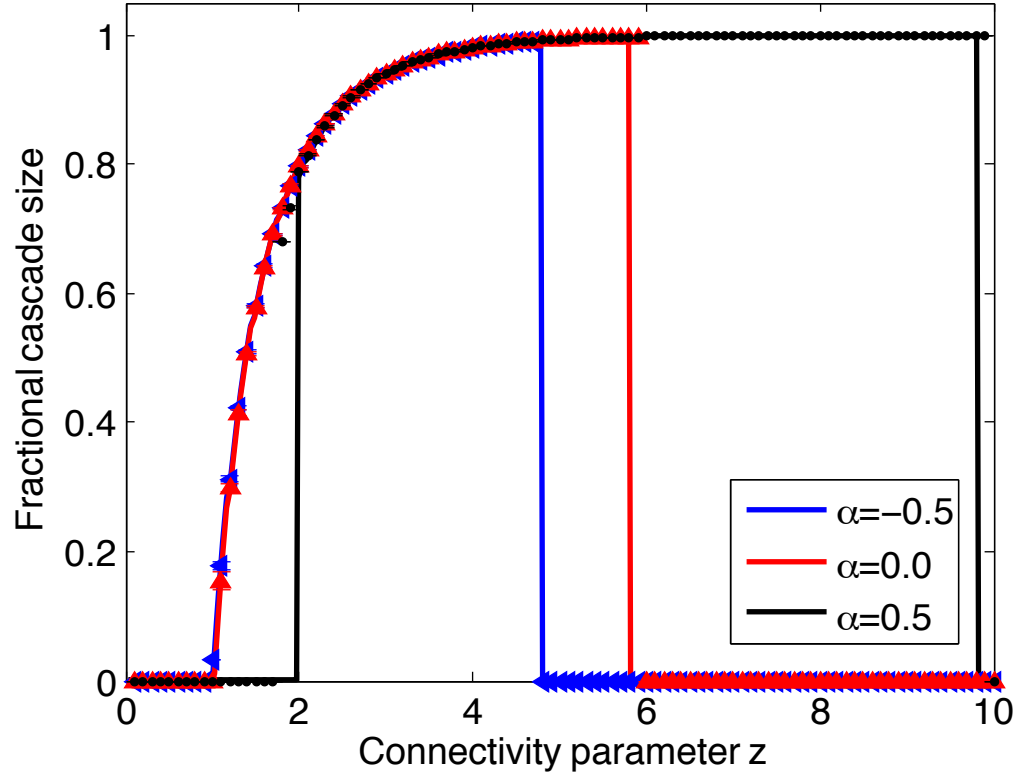


Figure 1: Experiment 1A: This figure shows how the size of global cascades on infinite networks is influenced by nonconstant but deterministic edge weights. The benchmark case with $\alpha = 0$, shown by the red triangles, is the case of constant edge weights. The cascade size remains similar in the case with $\alpha = -0.5$ (blue triangles). In both cases, the lower critical value of z is approximately 1. On the other hand, when $\alpha = 0.5$ (black pluses), the lower cascade transition is shifted up to $z \sim 1.5$, and furthermore, the upper critical value is shifted beyond 10 (in fact it occurs around $z = 14$).

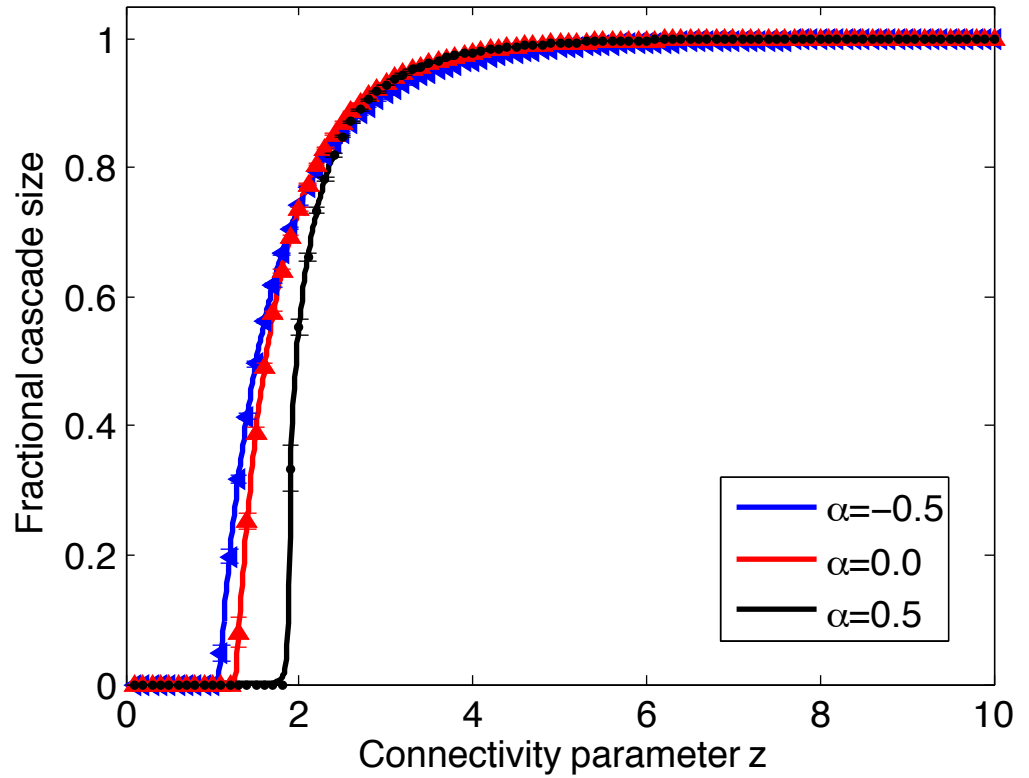


Figure 2: Experiment 1B: This figure shows how the size of global cascades on infinite networks is influenced by lognormally distributed edge weights, with means consistent with the cases of Figure 1. Compared to the deterministic cases, both the upper and lower critical values have been shifted to higher z values. All three curves, if extended beyond $z = 10$ exhibit upper critical values beyond $z = 14$.

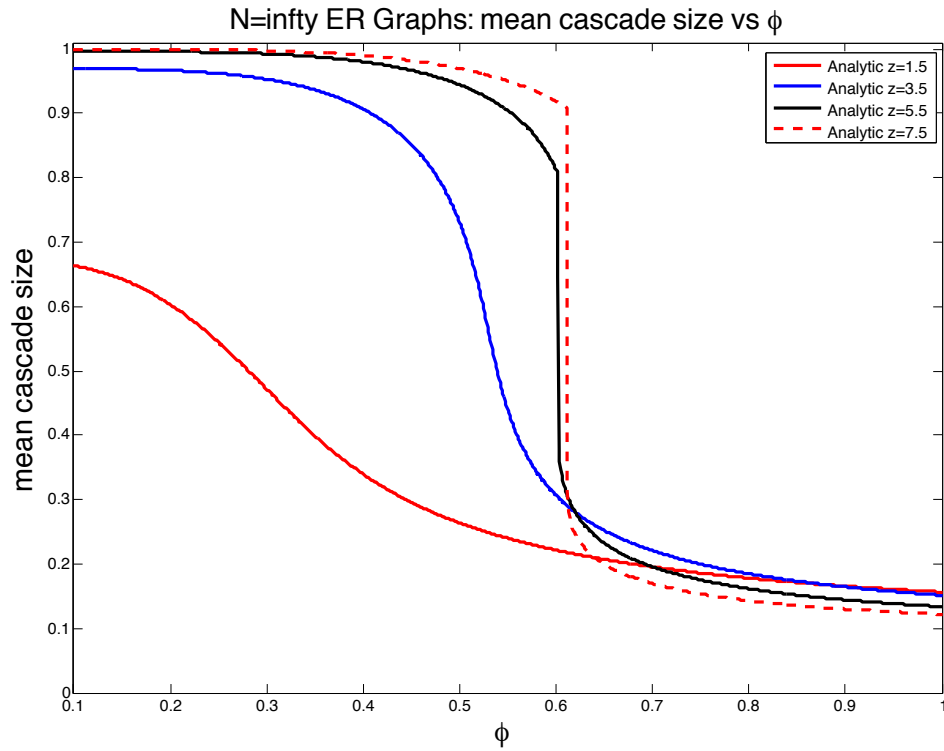


Figure 3: Experiment 1C: This figure shows how the mean cascade size depends on both the connectivity parameter z and the threshold parameter ϕ , when the weights are lognormal with mean and standard deviation $(1, 0.5)$.

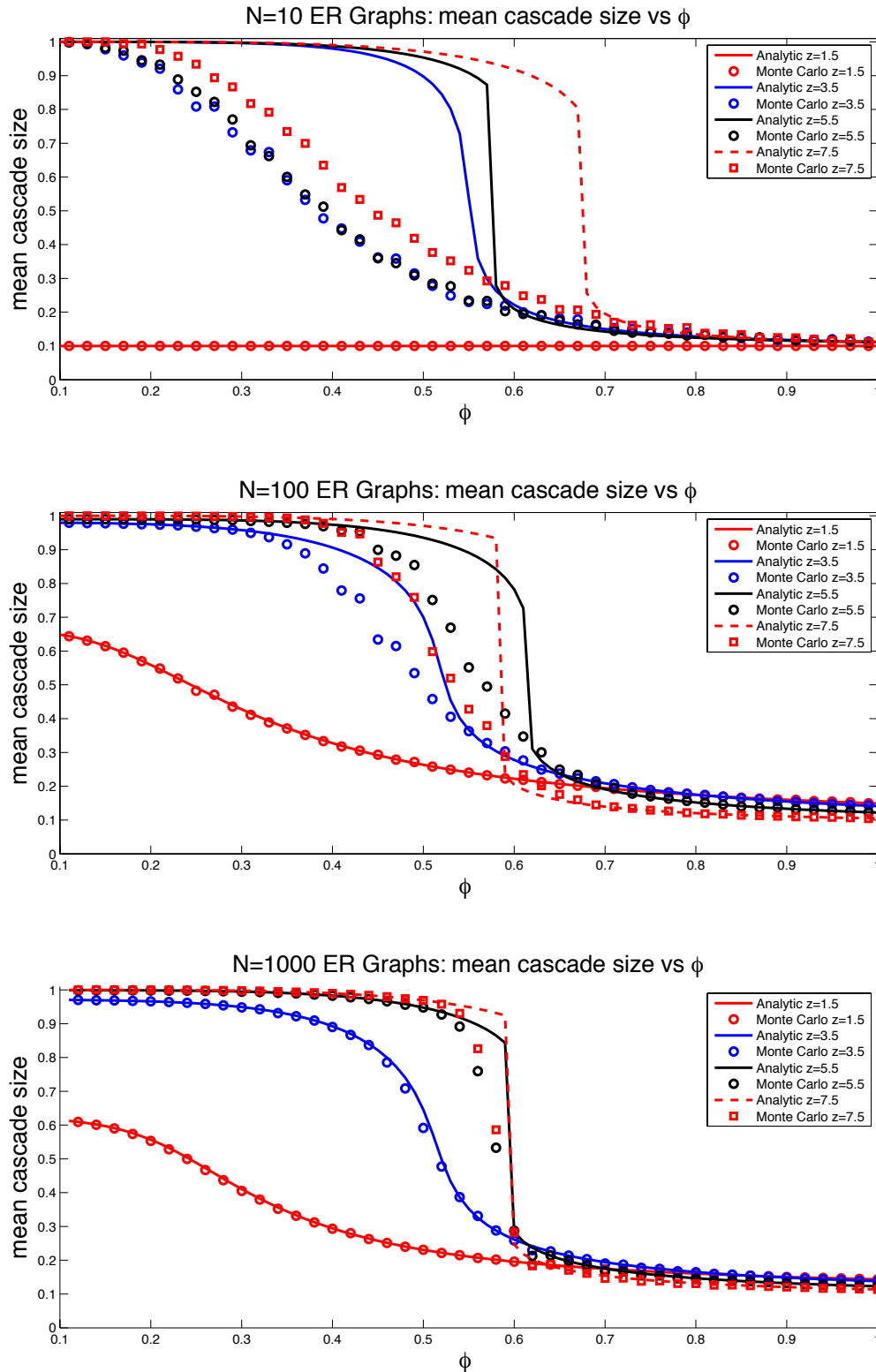


Figure 4: Experiment 2A: These three figures shows how the size of global cascades on finite Erdős-Rényi skeleton graphs is influenced by lognormally distributed edge weights, with means consistent with the cases of Figure 1. The figures from top to bottom show the results for skeleton graphs of size 10, 100, and 1000 respectively; the four curves correspond to different values of the connectivity parameter z . In each case, both LTIA and Monte Carlo values are plotted as functions of the threshold parameter ϕ .

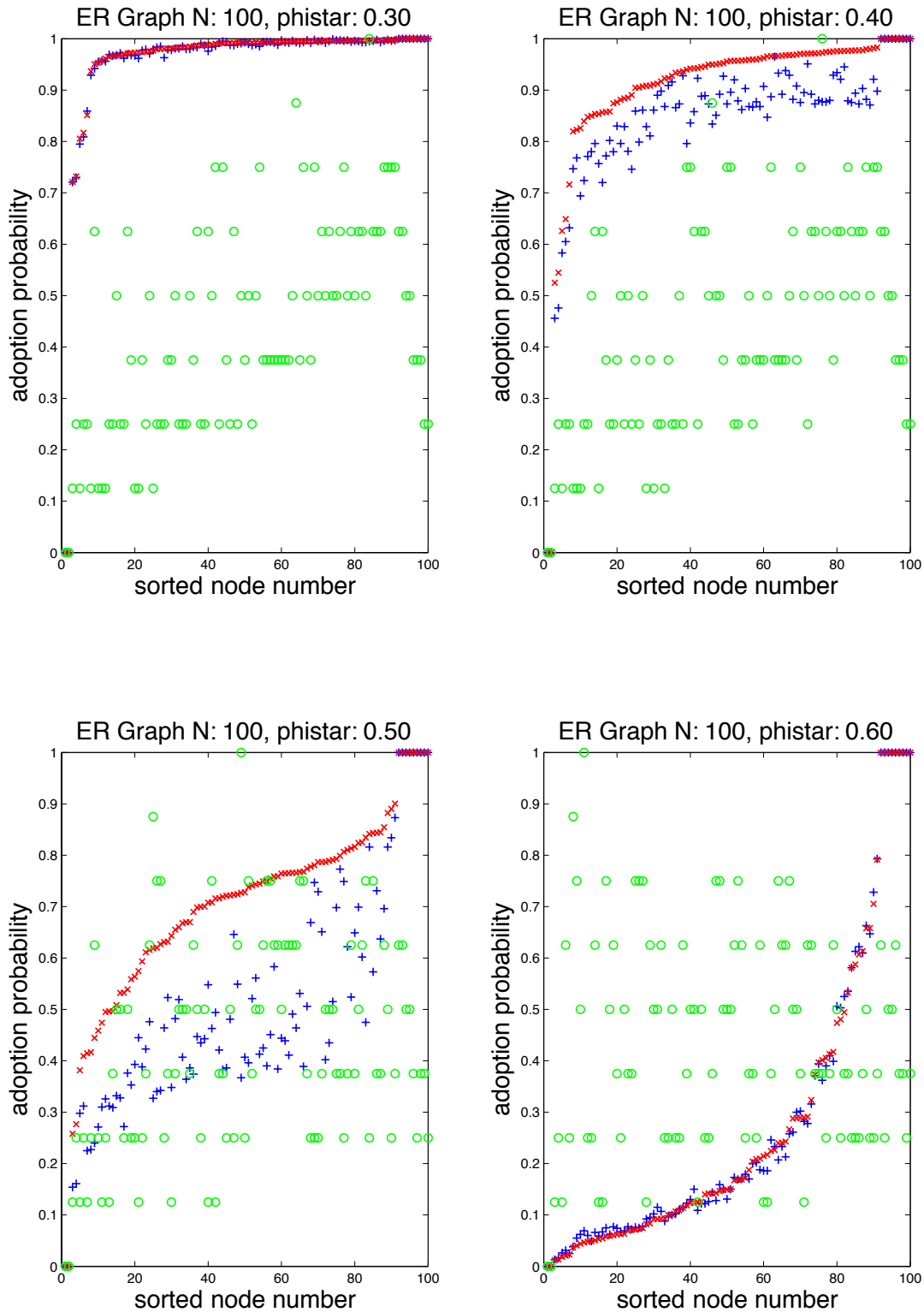


Figure 5: Experiment 2B: These four panels show results for a single Erdős-Rényi skeleton graph with $N = 100$ and $z = 3.5$ and four different threshold parameter values $\phi = 0.3, 0.4, 0.5, 0.6$. The panels plot the node-by-node $A_v^{(\infty)}$ probabilities as computed by both analytic (red crosses) and Monte Carlo methods (blue pluses), in increasing order of their analytic values. The green circles represent the node degree as a fraction of the maximum node degree.