

# An analytical approach to cascades on random networks

James P. Gleeson<sup>a</sup> and Diarmuid J. Cahalane<sup>b</sup>

<sup>a</sup> Applied Mathematics, University College Cork, Cork, Ireland;

<sup>b</sup> MACSI, University College Cork, Cork, Ireland.

## ABSTRACT

The expected steady-state fraction of active nodes in Watts' model of threshold dynamics on random networks is determined analytically. The analysis applies to random graphs with arbitrary degree distributions, and includes the effect of finite seed fractions. The seed fraction is shown to have a strong impact upon the existence of global cascades and Watts' cascade condition is extended to include these effects.

**Keywords:** Complex networks, cascades, threshold dynamics.

## 1. INTRODUCTION

Dynamical models on networks have attracted a great deal of recent research interest, see the reviews<sup>1-3</sup> and references therein. Of particular interest are the effects of the topological structure of the network upon the time evolution of dynamical quantities. Under certain circumstances a change of state at a small number of network nodes may cause a cascade of changes over the whole network — such cascades may occur as, for example, overload failures,<sup>4-9</sup> avalanches in sandpile models,<sup>10,11</sup> evolution of species,<sup>12</sup> or the spread of rumours and fads.<sup>13-15</sup> In this paper we analyze and generalize the model of cascades on networks introduced by Watts.<sup>13</sup> Watts' model defines the dynamics of binary-valued nodes in a network. The network consists of  $N$  nodes (vertices of a graph) and the connections between nodes (the edges of the graph) determined by one of a variety of common topologies (e.g. Erdős-Rényi random graph, small-world network, etc.). One characteristic of network structures which is of interest is the probability  $p_k$  that a randomly-selected node has degree  $k$ , i.e. that it has  $k$  edges connecting it to neighbouring nodes. The dependence of  $p_k$  upon  $k$  varies with different network topologies, but the mean connectivity  $z$  defined by

$$z = \sum_{k=0}^{\infty} k p_k \quad (1)$$

gives the average of the degree distribution and will appear frequently in our work.

The dynamics modelled by Watts is based on the class of problems known as *binary decisions with externalities*.<sup>16</sup> Each node in the network has a *state*  $v_i \in \{0, 1\}$ . Node  $i$  is called *active* if  $v_i = 1$  and *inactive* if  $v_i = 0$ , and the state of any given node may change over time. Each node also has a threshold  $r_i$  chosen from a distribution  $P(r)$ ; note the threshold of a node does not change over time. The Watts model begins by setting all but a (randomly-chosen) fraction  $\rho_0$  of the  $N$  nodes in a given network to the inactive state  $v_i = 0$ , with the remaining nodes initialized to active, with  $v_i = 1$ . Updating is asynchronous: at each discrete timestep a randomly-chosen node is updated by comparing the fraction of its neighbours which are currently active to its threshold  $r_i$ . If the active fraction of neighbours (the “neighbourhood average”) exceeds the node's threshold, then the node becomes active, otherwise its state is not changed, i.e.

$$v_i \text{ is updated to } \begin{cases} 1 & \text{if } \frac{1}{k_i} \sum_{j \in \mathcal{N}(i)} v_j > r_i, \\ \text{unchanged} & \text{if } \frac{1}{k_i} \sum_{j \in \mathcal{N}(i)} v_j \leq r_i. \end{cases} \quad (2)$$

Here the set of neighbours of node  $i$  is denoted by  $\mathcal{N}(i)$ , and  $k_i$  is the degree of node  $i$ . Neighbourless nodes, i.e. those with degree 0 are also left unchanged. Note that alternative update rules are also possible, in particular we will examine in later work the effect of altering (2) by setting  $v_i$  to 0 if  $\frac{1}{k_i} \sum_{j \in \mathcal{N}(i)} v_j \leq r_i$ . However, Watts'

---

Send correspondence to JPG: E-mail: j.gleeson@ucc.ie, Telephone: +353 21 490 3410

original model is simpler to analyze, because under its update rule a node which becomes active cannot later become inactive. As a result, the overall fraction of active nodes

$$\rho = \frac{1}{N} \sum_{i=1}^N v_i \quad (3)$$

is non-decreasing in time for Watts' model.

Watts' model thus begins with a fraction  $\rho_0$  of active nodes, and the asynchronous updating may then lead to further nodes being activated. Watts introduced this model to study the dynamics of cascades on networks, and in particular to examine the influence of network topology upon cascade propagation. His main result is the *cascade condition* (equation (5) of Ref. 13):

$$\sum_{k=1}^{\infty} k(k-1)p_k F\left(\frac{1}{k}\right) > z, \quad (4)$$

which must hold if a small initial fraction  $\rho_0$  of active nodes is to generate a global cascade in which almost all nodes in the network are activated. In (4),  $z$  and  $p_k$  are the mean connectivity and degree distribution already introduced in equation (1), while  $F(R)$  is the probability that a node's threshold  $r_i$  is less than  $R$ :

$$F(R) = \int_{-\infty}^R P(r) dr. \quad (5)$$

Watts derived (4) using percolation arguments and provided numerical evidence (using  $\rho_0 = 10^{-4}$ , see Fig. 1 of Ref. 13) that the cascade condition is quite accurate. With networks of size  $N = 10^4$ , some small discrepancies between theory and simulation were noted in Ref. 13 — these were attributed to the finite size of the system.

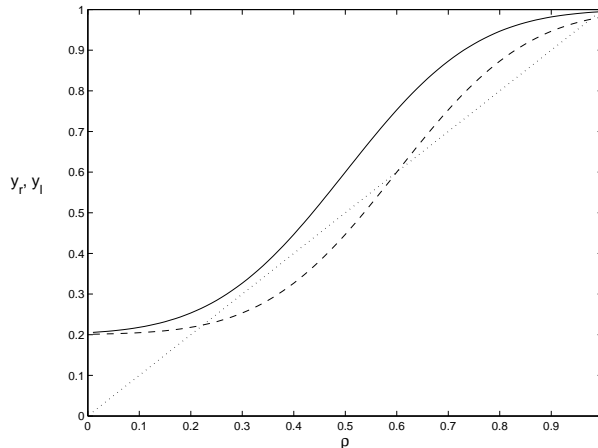
In this paper we extend Watts' result in several ways. First, we provide an analytical approach to calculating the expected long-time limit  $\rho_{\infty}$  of the fraction of active nodes:

$$\rho_{\infty} = \lim_{t \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N v_i. \quad (6)$$

The fraction  $\rho_{\infty}$  is used to quantify the size of global cascades (when they occur). Secondly, our analysis allows us to extend Watts' cascade condition (4) to include the effects of initially-activated fractions  $\rho_0$  which are not infinitesimally small. This is important for applications where the eventual steady-state popularity of a new product may be strongly influenced by the fraction  $\rho_0$  of early adopters.<sup>15</sup> The extended cascade condition can give dramatically different results to (4) for  $\rho_0$  values as low as 0.1%. Finally, our approach is valid for any distribution  $P(r)$  of thresholds, and is not restricted (as in Ref. 13) to distributions which have support only in the interval  $r \in [0, 1]$ . The remainder of this paper is organized as follows. In section 2 we examine the useful limiting case of all-to-all connected networks (the mean field case). Section 3 introduces our theoretical approach for random graphs of arbitrary degree distribution, and predictions are compared with numerical simulations in section 4.

## 2. ALL-TO-ALL CONNECTIVITY

To motivate our approach and introduce some of the techniques used in the remainder of the paper, we begin by considering the relatively simple case of all-to-all connected networks. In these networks (also known as "complete graphs"<sup>17</sup> or "mean-field models"<sup>18</sup>) every node is connected to every other node (including itself). Thus all nodes have the same neighbourhood average, and this is equal to the fraction  $\rho = 1/N \sum_{i=1}^N v_i$  of active nodes in the network. In all-to-all networks, the update rule (2) is simplified to: set  $v_i$  to 1 if  $\rho > r_i$ , and leave  $v_i$  unchanged otherwise. To simplify the analysis, we replace the asynchronous updating used by Watts with a synchronous update. This means that at each timestep all  $N$  nodes are simultaneously updated according to (2). In the Appendix we prove that synchronous and asynchronous updating both lead to the same steady-state results, at least under update rule (2).



**Figure 1.** The curves  $y_\ell = \rho$  (dotted),  $y_r$  for threshold mean  $r^* = 0.5$  (solid), and  $y_r$  for  $r^* = 0.6$  (dashed) giving the solution of the fixed point equation (8). Standard deviation  $\sigma = 0.2$  and initial activated fraction  $\rho_0 = 0.2$  for both cases.

Consider the update from timestep  $n$  (with active fraction  $\rho_n$  of nodes) to timestep  $n+1$  (with active fraction  $\rho_{n+1}$ ) under a synchronous application of rule (2). A fraction  $\rho_0$  of the nodes were initially activated at timestep 0 and these are therefore still active at timestep  $n+1$ . Of the remaining fraction  $(1 - \rho_0)$  of the nodes, only those with thresholds less than the neighbourhood average are activated. Since all neighbourhood averages equal the global fraction  $\rho_n$  in the all-to-all connected network, the probability that a threshold is less than  $\rho_n$  is simply  $F(\rho_n)$ . Thus the total fraction of nodes active at timestep  $n+1$  is given by

$$\rho_{n+1} = \rho_0 + (1 - \rho_0)F(\rho_n). \quad (7)$$

Beginning with the initial condition  $\rho_0$ , equation (7) is a nonlinear map governing the growth of the active fraction to its final limiting value  $\rho_\infty$ . From the point of view of cascades, it is interesting to ask whether (7) permits the growth of a small initial fraction to a global cascade (i.e.  $\rho_\infty$  of order one).

Analysis of equation (7) is aided by plotting the curves  $y_r = \rho_0 + (1 - \rho_0)F(\rho)$  and  $y_\ell = \rho$  (corresponding to the right and left hand sides of equation (7)) on the same axes. The lowest intersection point of these curves occurs at the limiting value of the active fraction  $\rho_\infty$  since this is a fixed point of the map (7):

$$\rho_\infty = \rho_0 + (1 - \rho_0)F(\rho_\infty). \quad (8)$$

In Figure 1 we show the curves  $y_r$  and  $y_\ell$  for two different threshold distributions  $P(r)$ , both with  $\rho_0 = 0.2$ . The solid line is  $y_r$  in the case where  $P(r)$  is a Gaussian distribution with mean  $r^* = 0.5$  and standard deviation  $\sigma = 0.2$ . The value of  $\rho_\infty$  in this case is  $\approx 0.995$ . The dashed line shows  $y_r$  for a Gaussian  $P(r)$  with a different mean,  $r^* = 0.6$ , but the same standard deviation. In this case the curve  $y_r$  intersects the line  $y_\ell = \rho$  at three points, instead of the single intersection seen for the solid curve. Consequently, the value of  $\rho_\infty$  in this case is  $\approx 0.224$ , considerably lower than in the other case, even though the parameter  $r^*$  has changed only by a relatively small amount. It is clear that the dependence of  $\rho_\infty$  upon  $r^*$  undergoes a transition at some critical value  $r_t^*$ , with  $\rho_\infty$  jumping discontinuously from a high value for  $r^* < r_t^*$  to a low value for  $r^* > r_t^*$ . Moreover, the existence of such a transition depends upon the standard deviation  $\sigma$  of the threshold distribution. If  $\sigma$  is large, the function  $y_r$  has only one intersection with  $y_\ell$  for all values of  $r^*$ , while for lower  $\sigma$  values there may be multiple intersection points. Therefore, there exists a critical standard deviation  $\sigma_c$ , with discontinuous transitions of the type described above existing only for  $\sigma < \sigma_c$ .

The critical value  $\sigma_c$  may be determined by seeking a solution of the fixed point equation (8) which is a double root. This requires that the fixed point  $\rho_\infty$  satisfies both (8) and the slope-matching condition

$$1 = (1 - \rho_0)P(\rho_\infty), \quad (9)$$

where we have used the fact that  $\frac{dF(\rho)}{d\rho} = P(\rho)$ . Equation (9) gives a simple criterion for existence of solutions: it requires that the standard deviation  $\sigma$  be small enough so that  $P(r) = 1/(1 - \rho_0)$  for some value  $r$  in the interval  $[0, 1]$ . Noting that the peak value of a Gaussian distribution occurs at its mean, and equals  $1/\sqrt{2\pi}\sigma$ , we obtain the criterion

$$\sigma \leq \sigma_c \equiv \frac{1 - \rho_0}{\sqrt{2\pi}} \quad (10)$$

for discontinuous transitions to exist. The discontinuity in the  $r^*$  dependence of  $\rho_\infty$  which first appears at  $\sigma = \sigma_c$  arises at a value  $r_c^*$  of the mean  $r^*$ , and a value of  $\rho_\infty$  corresponding to the mean of the distribution. This implies a double root of equation (8) at  $\rho_\infty = r_c^*$ , with  $F(\rho_\infty) = 1/2$ , yielding

$$r_c^* = \frac{1}{2}(1 + \rho_0). \quad (11)$$

Thus when  $\sigma < \sigma_c$ , the discontinuity in the  $r^*$ -dependence of  $\rho_\infty$  occurs at some value  $r_t^*$ , with  $r_t^*$  less than or equal to the  $r_c^*$  value given by (11). We note that (10) is a generalization of the  $\rho_0 = 0$  result used by several authors on related problems,<sup>18</sup> while the  $\rho_0 = 0$  version of (11) is analogous to the result found for the random-field Ising model on Bethe lattices by Dhar et al.<sup>19</sup>

### 3. RANDOM GRAPHS

The results of the previous section apply when the network is all-to-all connected, so that every node has  $N$  neighbours. In the  $N \rightarrow \infty$  limit, the mean connectivity  $z = N$  therefore becomes infinitely large. In the remainder of the paper we will examine the case where  $z$  is finite, while taking the limit of infinitely large networks  $N \rightarrow \infty$ . A good example of such a network is the well-known Erdős-Rényi (or Poisson) random graph with degree distribution

$$p_k = e^{-z} \frac{z^k}{k!}. \quad (12)$$

We perform numerical simulations of the dynamics on networks with this degree distribution, however our analytical results can also be applied to random graphs with arbitrary degree distributions.

Our analytical approach is based on methods introduced by Dhar et al. to study the zero-temperature random-field Ising model (RFIM) on Bethe lattices.<sup>19</sup> The RFIM is a spin-based model of magnetism, and its zero-temperature limit has been extensively studied as a model of hysteresis and Barkhausen noise.<sup>18</sup> A Bethe lattice of coordination number  $z$  (for integer  $z$ ) is an infinite tree where every node has exactly  $z$  neighbours. Dhar et al derive analytical results valid on Bethe lattices, but their numerical simulations show that the theory also applies very accurately to random graphs where every node has exactly  $z$  neighbours, provided that short-distance loops are rare. To analyze Watts' model we extend the approach of Ref. 19 in two ways. First, we consider tree-like random graphs with arbitrary degree distributions, rather than the Bethe lattices of Ref. 19. Secondly, we account for the difference between Watts' update rule (2) and standard RFIM dynamics — in the latter, a node is updated to state 0 if its local field (the equivalent of the neighbourhood average  $1/k_i \sum_{j \in \mathcal{N}(i)} v_j$ ) does not exceed its threshold  $r_i$ , but in Watts' model such nodes are left unchanged. This difference between the update rules is crucial to our derivation of the  $\rho_0$ -dependence of the activated fraction  $\rho_\infty$  (and therefore the cascade condition) below.

We begin by replacing the given random graph (with degree distribution  $p_k$ ) by a tree structure. The top level of the tree is a single node with degree  $k$ , and this is connected to its  $k$  neighbours at the next lower level of the tree. Each of these nodes is in turn connected to  $k_i - 1$  neighbours at the next lower level, where  $k_i$  is the degree of node  $i$ . The degree distribution of the nodes in the tree is given by

$$\tilde{p}_k = \frac{k}{z} p_k, \quad (13)$$

which is the distribution for the number of nearest neighbours in a connected graph.<sup>1,20</sup> To find the final density  $\rho_\infty$  of active nodes, we label the levels of the tree from  $n = 0$  at the bottom, with the top node at an infinitely high level ( $n \rightarrow \infty$ ). Define  $q_n$  as the conditional probability that a node on level  $n$  is active, conditioned on

its parent (on level  $n + 1$ ) being inactive. Consider updating a node on level  $n + 1$ , assuming that the nodes on all lower levels have already been updated. With probability  $\tilde{p}_k$  the chosen node has  $k$  neighbours: one of these is its parent (on level  $n + 2$ ), and the remaining  $k - 1$  are its children (on level  $n$ ). Since a fraction  $\rho_0$  of nodes were initially set to be active, there is a probability  $\rho_0$  that we have chosen one of these nodes. In this case the state of the node remains unchanged. On the other hand, with probability  $(1 - \rho_0)$  the node in question is inactive. In this case we must consider its neighbours. Each of the  $k - 1$  children are active with probability  $q_n$ . We also assume its parent is inactive. Thus the node has  $m$  active neighbours (and therefore  $k - 1 - m$  inactive neighbours) with probability  $\binom{k-1}{m} q_n^m (1 - q_n)^{k-1-m}$ . The probability that its threshold  $r_i$  is less than its neighbourhood average  $m/k$  is given by  $F(m/k)$ , and combining the independent probabilities yields the equation for  $q_{n+1}$ :

$$q_{n+1} = \sum_{k=1}^{\infty} \frac{k}{z} p_k \left[ \rho_0 + (1 - \rho_0) \sum_{m=0}^{k-1} \binom{k-1}{m} q_n^m (1 - q_n)^{k-1-m} F\left(\frac{m}{k}\right) \right] \text{ for } n = 0, 1, 2, \dots \quad (14)$$

with  $q_0 = \rho_0$ . This can be written in a form similar to the all-to-all connected case (7):

$$\begin{aligned} q_{n+1} &= \rho_0 + (1 - \rho_0)G(q_n) \text{ for } n = 0, 1, 2, \dots, \\ q_0 &= \rho_0, \end{aligned} \quad (15)$$

with the nonlinear function  $G$  defined by

$$G(q) = \sum_{k=1}^{\infty} \frac{k}{z} p_k \sum_{m=0}^{k-1} \binom{k-1}{m} q^m (1 - q)^{k-1-m} F\left(\frac{m}{k}\right). \quad (16)$$

Equation (15) is a nonlinear mapping for  $q_n$ , with fixed point  $q_\infty$  defined by the lowest intersection of the curves  $y = q$  and  $y = \rho_0 + (1 - \rho_0)G(q)$ . The probability that the single node at the top of the tree is active is given by adding the probabilities for two independent cases: either it is already active as part of the originally activated fraction (with probability  $\rho_0$ ), or it was initially inactive (with probability  $1 - \rho_0$ ). In the latter case, it will become active if sufficiently many of its  $k$  neighbours (all on the next lower level) are active. Noting that the top node has degree  $k$  with probability  $p_k$ , we conclude that the total probability of it being active is given by

$$\rho_\infty = \rho_0 + (1 - \rho_0) \sum_{k=1}^{\infty} p_k \sum_{m=0}^k \binom{k}{m} q_\infty^m (1 - q_\infty)^{k-m} F\left(\frac{m}{k}\right). \quad (17)$$

Equations (15) and (17) determine the expected value of the active fraction of nodes in the network as  $t \rightarrow \infty$ . Although the theory is defined in terms of level-by-level updating on a tree, we show that these results also apply to the random-graph Watts model, provided that: (i) the network structure is locally tree-like, and (ii) the state of each node is altered at most once. We note that condition (i) is true in topologies such as the Erdős-Rényi random graph, while (ii) is guaranteed by the update rule (2).

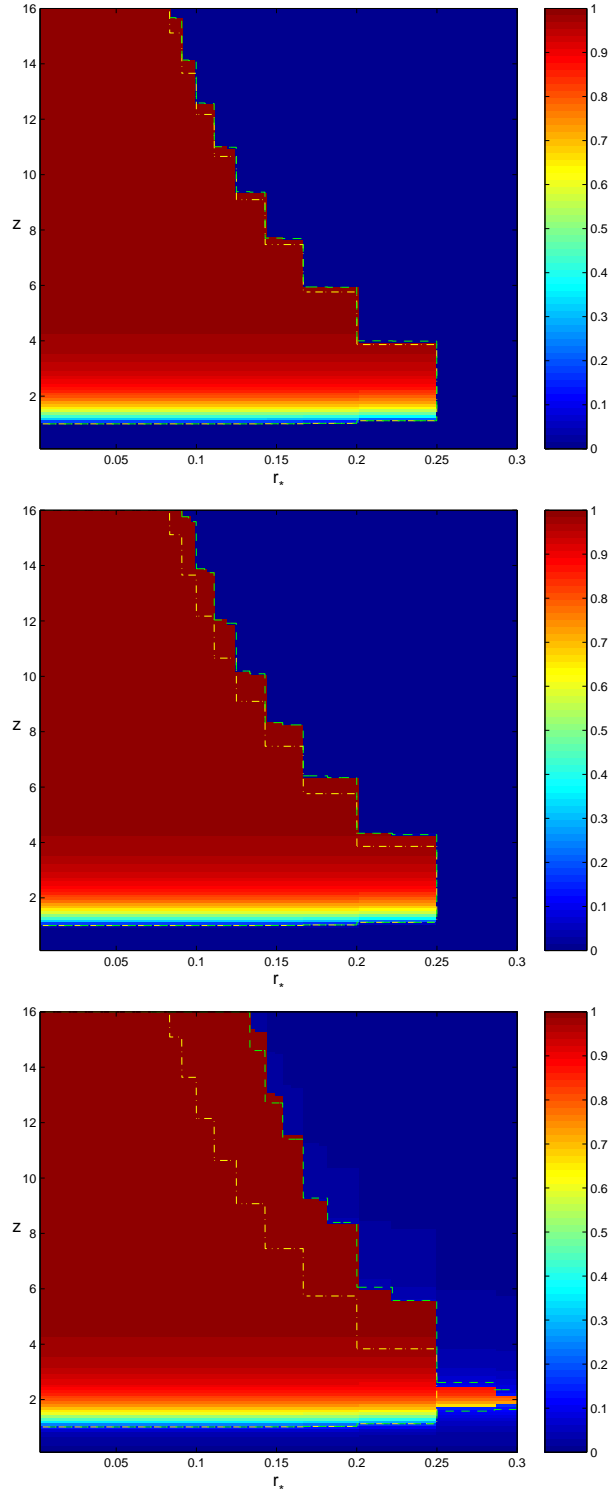
## 4. RESULTS

### 4.1. Uniform thresholds

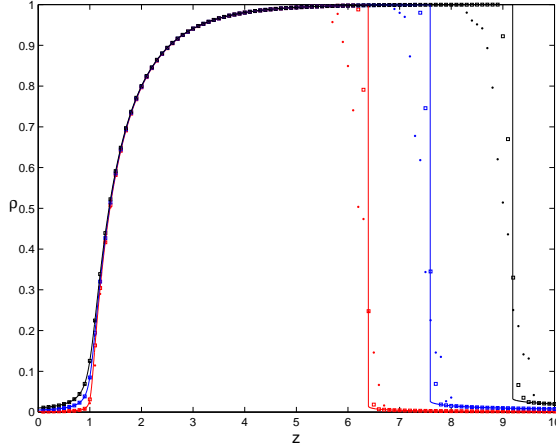
We first present some results of using the threshold distribution

$$P(r) = \delta(r - r^*), \quad (18)$$

where the Dirac delta function implies that every node has the same threshold  $r_i = r^*$ . We focus on the Erdős-Rényi random graph topology, with degree distribution given by the Poisson distribution (12). In Figure 2 we plot the values of  $\rho_\infty$  resulting from (15) and (17), as a function of the threshold  $r^*$  and mean network connectivity  $z$ . Note the smooth transition from  $\rho_\infty = 0$  to  $\rho_\infty \approx 1$  near  $z = 1$ , and the abrupt transition back to  $\rho_\infty \approx 0$



**Figure 2.** Final density  $\rho_\infty$  of activated nodes in a Poisson random graph of mean connectivity  $z$ , and uniform threshold value  $r^*$ . The initial density of activated nodes is (a)  $\rho_0 = 10^{-4}$ , (b)  $\rho_0 = 10^{-3}$ , (c)  $\rho_0 = 10^{-2}$ . The yellow dot-dashed line is the first-order (Watts) cascade condition (4); the green dashed line is the extended cascade condition (25).



**Figure 3.** Final fraction  $\rho_\infty$  of activated nodes in a Poisson random graph with uniform threshold  $r^* = 0.18$ , as a function of the mean network connectivity  $z$ . The initially activated fraction is  $\rho_0 = 10^{-3}$  (red),  $\rho_0 = 5 \times 10^{-3}$  (blue), and  $\rho_0 = 10^{-2}$  (black). Lines are the theoretical predictions using (15) and (17); symbols are from numerical simulations (over 100 network realizations) with  $N = 10^4$  (points) and  $N = 10^5$  (squares) nodes.

at larger  $z$  values. Parts (a), (b), and (c) of Figure 2 show the effects of the initially-active fraction  $\rho_0$  upon the final results, with  $\rho_0$  taking values  $10^{-4}$ ,  $10^{-3}$ , and  $10^{-2}$ , respectively. Note that Watts' cascade boundary (4) is plotted as a yellow dot-dashed line in these figures — while (4) is accurate in the limit of infinitesimally small  $\rho_0$ , it clearly requires modification for the finite (although still very small) values used here.

Another view of the dependence upon  $\rho_0$  is provided in Figure 3, where  $\rho_\infty$  is plotted as a function of  $z$ , while the threshold value is kept at  $r^* = 0.18$ . Again we note the difference between small- $\rho_0$  results and Watts' prediction of the cascade boundary (at  $z \approx 5.76$  in this case). Also note the agreement of the theoretical predictions of (15) and (17) with the numerical simulations on Poisson random graphs. Each simulation result is the average over 100 network realizations, and matches theory extremely well except near the discontinuous transition at high  $z$ . Increasing the number of nodes in the simulated network from  $N = 10^4$  (points) to  $N = 10^5$  (squares) leads to improved agreement between theory and simulation.

## 4.2. Extending the cascade condition

It is clearly desirable to understand how Watts' cascade condition relates to the analytical approach used here to derive (15) and (17). Moreover, the results of Figures 2 and 3 motivate us to extend Watts' result to cases where the initialized fraction  $\rho_0$  is small but finite.

A global cascade occurs when the activation of a small initial fraction  $\rho_0$  of the nodes leads to a much larger  $\rho_\infty$ , i.e.  $\rho_\infty \gg \rho_0$ . While  $\rho_\infty$  is ultimately determined by (17), its magnitude can be estimated by that of the conditional probability  $q_\infty$  which is the fixed point of equation (15). We will determine a cascade condition based on the magnitude of  $q_\infty$ , assuming that a cascade occurs in cases where  $q_\infty \gg \rho_0 = q_0$ .

As a first approximation we linearize  $G(q)$ , retaining only two terms of its Taylor series about  $q = 0$ . The resulting approximation to (15) gives the fixed point equation

$$q_\infty \approx \rho_0 + (1 - \rho_0) [G(0) + q_\infty G'(0)], \quad (19)$$

which is readily solved to yield

$$q_\infty \approx -\frac{(1 - \rho_0)G(0) + \rho_0}{(1 - \rho_0)G'(0) - 1}. \quad (20)$$

We now argue as follows: if the denominator of (20) is negative, then the resulting (approximation to)  $q_\infty$  is positive and of order  $\rho_0$ . In this case the conditional probability  $q_\infty$  is of the same order of  $\rho_0$ , i.e.  $q_\infty \ll 1$ , and so a global cascade has not occurred. We deduce the *first-order cascade condition*:

$$(1 - \rho_0)G'(0) - 1 > 0 \quad (21)$$

which must hold for cascades to occur. Noting that (16) gives  $G(0) = F(0)$  and

$$G'(0) = \sum_{k=1}^{\infty} \frac{k}{z} p_k (k-1) \left[ F\left(\frac{1}{k}\right) - F(0) \right], \quad (22)$$

we see that Watts' cascade condition (4) is a special case of (21) when  $\rho_0 = 0$  and  $F(0) = 0$ .

The first-order cascade condition (21) is plotted as a yellow dot-dashed line in Figure 2 (see also Figure 4). It is clear that this condition does not accurately represent the effects of non-zero  $\rho_0$  or  $F(0)$ . The reason for this is that the function  $G$  is not well-approximated by a straight line near the critical parameters for cascades.

Accordingly, we extend (19) by using the Taylor series for  $G(q)$  to order  $q^2$ . This approximation results in a quadratic equation for the fixed point  $q_{\infty}$  which we represent as

$$aq_{\infty}^2 + bq_{\infty} + c = 0, \quad (23)$$

with coefficients given by

$$\begin{aligned} a &= \frac{1}{2}(1 - \rho_0)G''(0), \\ b &= (1 - \rho_0)G'(0) - 1, \\ c &= (1 - \rho_0)G(0) + \rho_0. \end{aligned} \quad (24)$$

Under the approximation of small  $q$  values, we interpret the existence of a positive root of (23) to imply  $q_{\infty} \ll 1$ , and hence the impossibility of global cascades. Note that the first-order approximation (20) given by  $q_{\infty} = -c/b$  is recoverable from (23) if  $a$  (i.e.  $G''(0)$ ) is assumed to be negligible. Since  $c \geq 0$ , the first-order cascade condition (21) then follows from demanding  $b > 0$ . However, when the nonlinear behaviour of  $G$  is important it is still possible for cascades to occur when  $b$  is negative, provided that the full solution of (23) precludes positive real roots. This extends the cascade condition to those cases where  $b \leq 0$ , but  $b^2 - 4ac < 0$ . This *extended cascade condition* may be further simplified by neglecting terms of order  $\rho_0^2$  to give:

$$-[G'(0) - 1]^2 + 2G(0)G''(0) - 2\rho_0 [G'(0) - G'(0)^2 - G''(0) + 2G(0)G''(0)] > 0 \quad (25)$$

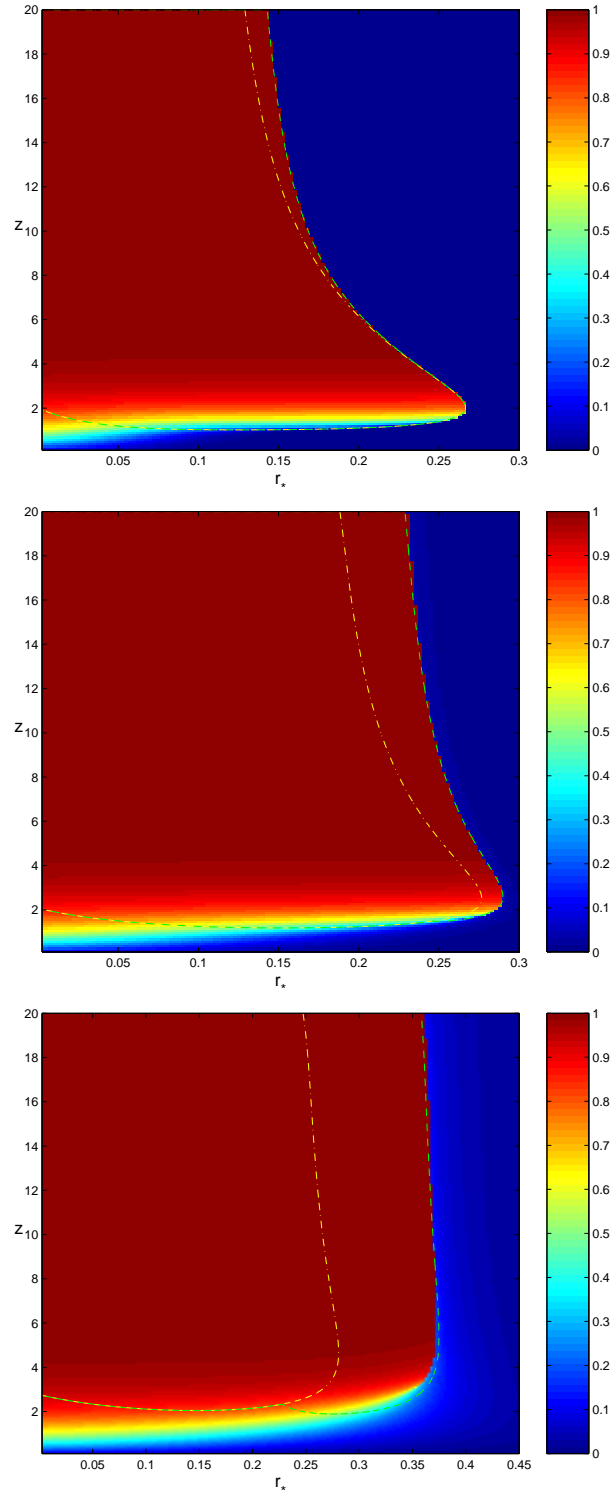
for cascades (which do not satisfy (21)) to occur. Using expressions (22) and

$$G''(0) = \sum_{k=1}^{\infty} \frac{k}{z} p_k (k-1)(k-2) \left[ F(0) - 2F\left(\frac{1}{k}\right) + F\left(\frac{2}{k}\right) \right], \quad (26)$$

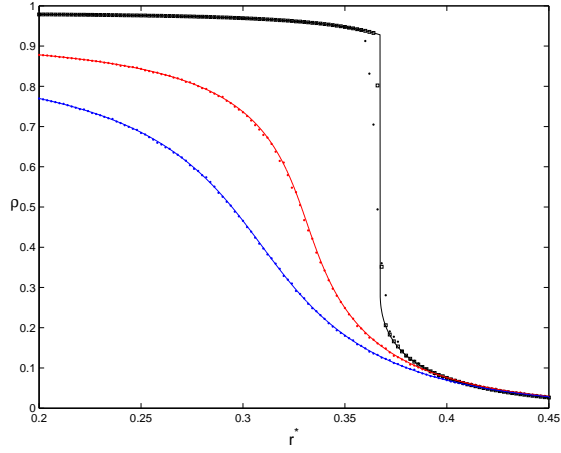
it is straightforward to plot the extended cascade condition: this is the green dashed line in Figures 2 and 4. This extended condition clearly gives much improved approximations to the actual cascade boundaries.

### 4.3. Gaussian threshold distributions

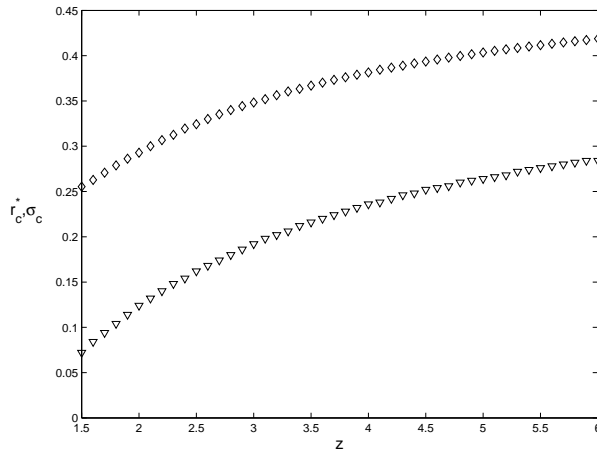
Figure 4 shows the expected activated fraction  $\rho_{\infty}$  of nodes in several cases where the threshold distribution is Gaussian, with mean  $r^*$  and standard deviation  $\sigma$ . Parts (a), (b), and (c) of the figure correspond to  $\sigma = 0.05$ ,  $\sigma = 0.1$ , and  $\sigma = 0.2$ , respectively; note the initially activated fraction  $\rho_0$  is set to zero throughout this section. The cascade conditions perform less well at low  $r^*$  and low  $z$  values in this case — this is because  $G(0)$  (and hence, from (15), the value of  $q_1$ ) is non-zero as  $r^*$  limits to zero. This renders the small- $q$  assumption of section 4.2 ineffective. However at larger values of  $r^*$  the value of  $G(0)$  is sufficiently small for the extended cascade condition to give quite accurate results. Figure 5 demonstrates the agreement of the theory (from (15) and (17)) with numerical simulations. This figure also shows that the dependence of  $\rho_{\infty}$  upon the mean threshold  $r^*$  may be discontinuous in certain circumstances. Similar arguments to those leading to equations (10) and (11) for the all-to-all case may be applied here. For sufficiently low values of the threshold standard deviation  $\sigma$  the fixed point equation resulting from (15) has three real roots. As the standard deviation  $\sigma$  is increased (while  $z$  remains unchanged), a critical point occurs where the fixed point equation has a double root. This point defines the critical values  $\sigma_c$  and  $r_c^*$  for the chosen  $z$  value. If  $\sigma$  is then further increased beyond  $\sigma_c$  the activated fraction



**Figure 4.** Final density  $\rho_\infty$  of activated nodes in a Poisson random graph of mean connectivity  $z$ , and Gaussian-distributed thresholds of mean  $r^*$ . The standard deviation of the threshold distribution is (a)  $\sigma = 0.05$ , (b)  $\sigma = 0.1$ , (c)  $\sigma = 0.2$ ; note change of  $r^*$  axis scaling for last figure, also  $\rho_0 = 0$  in all cases. The yellow dot-dashed line is the first-order cascade condition (21); the green dashed line is the extended cascade condition (25).



**Figure 5.** Final fraction  $\rho_\infty$  of activated nodes in a Poisson random graph with Gaussian-distributed thresholds of standard deviation  $\sigma = 0.2$ , as a function of the mean threshold  $r^*$ . The mean connectivity  $z$  of the network is  $z = 2$  (blue),  $z = 2.5$  (red), and  $z = 4$  (black), and  $\rho_0 = 0$  in all cases. Lines are the theoretical predictions using (15) and (17); symbols are from numerical simulations (over 100 network realizations) with  $N = 10^4$  (points) and (for  $z = 4$  only)  $N = 10^5$  (squares) nodes.



**Figure 6.** Critical values of threshold standard deviation  $\sigma_c$  (triangles) and mean  $r_c^*$  (diamonds) as a function of mean connectivity  $z$  of Poisson random networks.

$\rho_\infty$  varies smoothly with threshold mean  $r^*$ . Figure 6 shows the critical values of  $\sigma_c$  and  $r_c^*$  at various  $z$  values, as found by a numerical search for double roots of the fixed point equation. In the  $z \rightarrow \infty$  limit both quantities approach their mean-field values as given by equations (10) and (11) with  $\rho_0 = 0$ .<sup>21</sup>

These results explain the behaviour shown in Figure 5: for  $z = 2$  and  $z = 2.5$  the critical standard deviations  $\sigma_c$  are less than the value of  $\sigma = 0.2$  used in this figure. Therefore the curves corresponding to these  $z$  values show a continuous dependence of  $\rho_\infty$  upon  $r^*$ . On the other hand, Figure 6 shows that the critical value  $\sigma_c$  for  $z = 4$  is greater than  $\sigma = 0.2$ , and therefore the black curve in Fig. 5 has a discontinuity at a value of  $r^*$  less than  $r_c^* \approx 0.381$ . We note that these results generalize those found by Dhar et al<sup>19</sup> on Bethe lattices (for which the connectivity  $z$  is necessarily integer-valued).

## 5. DISCUSSION

We have presented an analytical method for determining the expected steady-state fraction of active nodes in Watts' model of threshold dynamics on a random network. Our approach includes the effect of arbitrarily large initial (seed) fractions  $\rho_0$  of activated nodes, and we have derived an extended cascade condition to explain the deviation from Watts' original condition when  $\rho_0$  is not infinitesimally small. A detailed analysis of the types of bifurcation exhibited by the analytical solution can be found in Ref. 22.

## ACKNOWLEDGMENTS

This work is supported in part by Science Foundation Ireland under Investigator Award No. 02/IN.1/IM062 and the MACSI initiative, and by the Irish Research Council for Science, Engineering and Technology under the Embark Initiative, grant reference RS/2005/16.

## APPENDIX A.

Let  $A_n^s$  be the set of active node indices at timestep  $n$ , using synchronous updating, and  $A_n^a$  be the corresponding list of active node indices when asynchronous updating is used. Define  $\mu_n^s$  to be the size of the set  $A_n^s$ , and  $\mu_n^a$  to be the size of the set  $A_n^a$ . We prove the following:

**Theorem:** On any finite network, if  $A_0^s = A_0^a$  and the updating rules ensure both  $\{\mu_n^s\}$  and  $\{\mu_n^a\}$  are non-decreasing sequences, then steady-state values  $\mu_\infty^s$  and  $\mu_\infty^a$  exist as  $n \rightarrow \infty$ , and  $\mu_\infty^s = \mu_\infty^a$ .

**Proof:** The existence of steady states follows from the non-decreasing property of the  $\mu_n$  with  $n$ , and the fact that  $\mu_n \leq N$  in a finite network of  $N$  nodes.

Next, introduce the concept of *vulnerable nodes*: Let  $V_n^s$  be the set of vulnerable nodes (those nodes  $i$  with  $\frac{1}{k_i} \sum_{j \in \mathcal{N}(i)} v_j > r_i$ ) at timestep  $n$ , using synchronous updating. Similarly,  $V_n^a$  represents the set of vulnerable nodes at timestep  $n$  under asynchronous updating.

*Show  $\mu_\infty^a \leq \mu_\infty^s$ :* Since  $A_0^s = A_0^a$  and vulnerability of nodes follows from the activity of neighbouring nodes, we have  $V_0^s = V_0^a$ . Now consider timestep  $n = 1$ . The set  $A_1^s$  is found by taking the union of  $A_0^s$  with  $V_0^s$ . The set  $A_1^a$  contains at most one element which is not in  $A_0^a$ , this element having been chosen from  $V_0^a$ . Therefore  $A_1^a \subseteq A_1^s$  and so  $\mu_1^a \leq \mu_1^s$ . Repeating the argument leads to  $\mu_n^a \leq \mu_n^s$ , and so  $\mu_\infty^a \leq \mu_\infty^s$ .

*Show  $\mu_\infty^s \leq \mu_\infty^a$ :* As noted above, the set  $A_1^s$  is equal to  $A_0^s \cup V_0^s$ . The sets  $A_n^a$  for  $n \geq 1$  are found by appending at most one member of the set  $V_{n-1}^a$  to  $A_{n-1}^a$ , then adding newly vulnerable nodes to  $V_{n-1}^a$  gives  $V_n^a$ . Since  $V_0^s = V_0^a \subseteq V_n^a \forall n$ , the asynchronous updating will eventually lead to a timestep  $i_1$  when all elements of  $V_0^s$  have been activated (although these may only be a subset of all the active nodes in  $A_{i_1}^a$ ). Thus  $A_1^s \subseteq A_{i_1}^a$  and so  $\mu_1^s \leq \mu_{i_1}^a$ . Similar arguments lead to the sequence of time-indices  $i_2, i_3, \dots$  for which  $A_n^s \subseteq A_{i_n}^a$ , with  $\mu_n^s \leq \mu_{i_n}^a$ . It follows that  $\mu_\infty^s \leq \mu_\infty^a$ .

## REFERENCES

1. M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, **45**, pp. 167-256, 2003.
2. L. A. N. Amaral and J. M. Ottino, "Complex networks - Augmenting the framework for the study of complex systems," *European Phys. J. B*, **38**, pp. 147-162, 2004.
3. S. Boccaletti et al., "Complex networks: Structure and dynamics," *Phys. Rep.*, **424**, pp. 175-308, 2006.
4. E. J. Lee et al., "Robustness of the avalanche dynamics in data-packet transport on scale-free networks," *Phys. Rev. E*, **71**, 056108, 2005.
5. Y. C. Lai et al., "Complex networks: Dynamics and security," *Pramana - Journal of Physics*, **64**, pp. 483-502, 2005.
6. P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, **69**, 045104, 2004.
7. A. E. Motter and Y. C. Lai, "Cascade-based attacks on complex networks," *Phys. Rev. E*, **66**, 065102, 2002.
8. Y. Moreno et al., "Critical load and congestion instabilities in scale-free networks," *Europhys. Lett.*, **62**, pp. 292-298, 2003.
9. P. Holme and B. J. Kim, "Vertex overload breakdown in evolving networks," *Phys. Rev. E*, **65**, 066109, 2002.
10. K. I. Goh et al., "Cascading toppling dynamics on scale-free networks," *Physica A*, **346**, pp. 93-103, 2005.
11. K. I. Goh et al., "Sandpile on scale-free networks," *Phys. Rev. Lett.*, **91**, 148701, 2003.
12. T. Antal, S. Redner, and V. Sood, "Evolutionary dynamics on degree-heterogeneous graphs," *Phys. Rev. Lett.*, **96**, 188104, 2006.
13. D. J. Watts, "A simple mode of global cascades on random networks" *Proc. Nat. Acad. Sci.*, **99**, pp. 5766-5771, 2002.
14. P. S. Dodds and D. J. Watts, "Universal behavior in a generalized model of contagion," *Phys. Rev. Lett.*, **92**, 218701, 2004.
15. W. Q. Duan et al., "Efficient target strategies for contagion in scale-free networks," *Phys. Rev. E*, **72**, 026133, 2005.
16. T. C. Schelling, *J. Conflict Resolution*, **17**, pp. 381-428, 1973.
17. R. Dobrin, J. H. Meinke and P. M. Duxbury, "Random-field Ising model on complete graphs and trees," *J. Phys. A: Math. Gen.*, **35**, pp. L247-L254, 2002.
18. J. P. Sethna, K. Dahmen et al., "Hysteresis and hierarchies: dynamics of disorder-driven first-order phase transitions," *Phys. Rev. Lett.*, **70**, pp. 3347-3350, 1993.
19. D. Dhar, P. Shukla, and J. P. Sethna, "Zero-temperature hysteresis in the random-field Ising model on a Bethe lattice," *J. Phys. A: Math. Gen.*, **30**, pp. 5259-5267, 1997.
20. S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, 2003.
21. X. Illa, P. Shukla, and E. Vives, "Zero-temperature hysteresis in a random-field Ising model on a Bethe lattice: Approach to mean-field behavior with increasing coordination number  $z$ ," *Phys. Rev. B*, **73**, 092414, 2006.
22. J. P. Gleeson and D. J. Cahalane, "Seed size strongly affects cascades in random networks," *Phys. Rev. E.*, to appear. Preprint available from <http://euclid.ucc.ie/pages/staff/gleeson>.