MULTI-PARAMETER MODELS OF INNOVATION DIFFUSION ON COMPLEX NETWORKS*

N. J. McCullen†, A. M. Rucklidge†, C. S. E. Bale‡, T. J. Foxon§, and W. F. Gale‡

Abstract. A model, applicable to a range of innovation diffusion applications with a strong peer to peer component, is developed and studied, along with methods for its investigation and analysis. A particular application is to individual households deciding whether to install an energy efficiency measure in their home. The model represents these individuals as nodes on a network, each with a variable representing their current state of adoption of the innovation. The motivation to adopt is composed of three terms, representing personal preference, an average of each individual's network neighbours' states and a system average, which is a measure of the current social trend. The adoption state of a node changes if a weighted linear combination of these factors exceeds some threshold. Numerical simulations have been carried out, computing the average uptake after a sufficient number of time-steps over many realisations at a range of model parameter values, on various network topologies, including random (Erdős–Rényi), small world (Watts–Strogatz) and (Newman's) highly clustered, community-based networks. An analytical and probabilistic approach has been developed to account for the observed behaviour, which explains the results of the numerical calculations.

1. Introduction. Social phenomena, such as the spread of a technological or behavioural innovation through communities, can be modelled as dynamical processes on networks [1–5,10]. Our model, introduced in §2, builds on previous threshold diffusion models (e.g. [4, 13, 17]) by incorporating sociologically realistic factors, yet remains simple enough for mathematical insights to be developed.

An example of a particular application of this model is to the adoption of innovations related to energy behaviours and technologies by individual households. These innovations are often not directly visible to an adopter's peers but communication of the benefits of adoption may occur through interaction between individuals. The decision to adopt is therefore based on multiple factors, taking into account not only individual preferences, but also whether or not an individual's social circle has adopted the innovation. As such, the spread of the innovation will be influenced by the network of social contacts between individuals, including both social peers and wider social trends. The real topologies of interpersonal networks are not known exactly and are constantly changing, further adding to the uncertainty in models of social phenomena. Given these challenges, useful models and methods are required that can assess the likely outcomes of particular scenarios, such as the effect of an intervention to persuade population to adopt an innovation, without being excessively sensitive to particular choices of parameter values of the model or the precise structure of the network. Here, we present numerical investigations of the model, studying its behaviour statistically, as well as presenting mathematical analysis that gives a deeper understanding of the observed uptake of the innovation over the network using probabilistic arguments.

The numerical methods used in §3 take the ensemble averages of simulation outcomes

---

†Department of Applied Mathematics, University of Leeds.

‡Energy Research Institute, University of Leeds.

§School of Earth and Environment, University of Leeds.

over many realisations, each with different conditions, such as the initial seed of adopters or the precise details of the network structure. This is carried out over a range of parameter values and on different types of network topology in order to study both the stability of the model with respect to minor variations and the sensitivity to structural details of the models used. These methods can, therefore, be applied to various dynamical models on a variety of networks to assess likely differences in outcomes for alternative scenarios in a statistical sense, rather than relying on individual predictions.

We take the applied dynamical systems approach of looking for lines in parameter space where the behaviour of the dynamical process changes. Mathematical analysis is presented in §4 explaining observations for the simulation results in certain simplified cases. Specifically, assuming that all individuals are homogeneous in their model parameters, we can use probabilistic arguments for the conditions of the neighbourhoods of the individuals required to trigger uptake, and calculate the likelihood of success given a choice of parameters on certain types of networks.

We have found that the likelihood of success, defined as adoption by the majority of individuals, depends strongly on both the choice of model parameters and the topology of the network. Two particularly important factors are found to be the *node degree*, i.e., the number of connections of the nodes with their neighbours, and the network's *transitivity*, which is the degree of clustering related to the correlation between the neighbourhoods of connected individuals. This makes it possible to assess what changes to the model would result in an increased probability of success, and how this can be interpreted in terms of policy decisions intent on maximising adoption in the real world.

2. Modelling social behaviour as dynamical systems on networks. A system of N individuals (or groups such as households) can be represented by N *nodes* on a *network*. Possible interactions linking individuals are represented by the *edges* on the network. The network topology can be represented in the *adjacency matrix A*, an N^2 matrix with entries:

$$A_{ij} = \begin{cases} 1 & \text{if node } j \text{ influences } i, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Information or influence is passed along these edges, which could also be given weights to show the relative probability or strength of the interaction. In this work all influences are taken to be of equal weight and symmetric in both directions. States are assigned to the nodes, describing the properties of the individuals, and deterministic or probabilistic equations or rules can be used to describe the evolution of these states over time.

2.1. A multi-parameter model for innovation uptake. The following model describes the purchase or adoption of a single innovation (not considering competition between similar items) but could apply equally well to the spread of any property over a network, such as a rumour or behaviour. The model used in this work assigns each node i with a binary variable representing their current state, $x_i = 1, 0$, indicating whether or not the individual has adopted the innovation. The nodes in state $x = 1$ are referred to as *activated* nodes. The initial state is a fixed proportion of activated nodes, randomly distributed across the network. The decision to invest in the innovation is determined by the perceived usefulness, or *utility*, to the individual (which can include subjective judgements). When this utility outweighs the

barriers to adoption (including financial costs), the individual i adopts the innovation, so x_i changes from 0 to 1. As this is a one-way process, unlike voter models [3], the state at the next time-step is determined by:

$$x_i(t+1) = \begin{cases} 1 & \text{if } x_i(t) = 1, \\ 1 & \text{if } x_i(t) = 0 \text{ and } u_i(t) > \theta_i, \\ 0 & \text{otherwise}, \end{cases} \qquad (2.2)$$

where θ_i is the threshold and t is time. The total utility u_i is a combination of both personal and social benefit [5]. The personal benefit p_i is a measure of the perceived intrinsic benefit to the individual of acquiring the innovation. The social benefit can be split into both the direct influence from an individual's peer group and the (possibly indirect) influence from society in general [16], which could be affected by the desire to fit in with the social norm on different levels. Thus our utility model has three factors that can be given relative weightings α_i, β_i and γ_i, with $\alpha_i + \beta_i + \gamma_i = 1$. The total utility is, therefore, given by:

$$u_i(t) = \alpha_i p_i + \beta_i s_i(t) + \gamma_i m(t), \qquad (2.3)$$

where s_i is the average of x within a node's *network neighbourhood*:

$$s_i(t) = \frac{1}{k_i} \sum_j^N A_{ij} x_j(t), \qquad (2.4)$$

and the *degree* of node i is $k_i = \sum_{j=1}^N A_{ij}$. The mainstream social norm m is the average of x over the entire population:

$$m(t) = \frac{1}{N} \sum_i^N x_i(t). \qquad (2.5)$$

This can be made equivalent to previous threshold diffusion models such as [17] by setting $(\alpha, \beta, \gamma) = (0, 1, 0)$. We define $m_0 = m(t = 0)$ to be the initial fraction of adopters. For the rest of the discussion we make the simplifying assumption that all nodes take the same values of α, β, γ, p and θ, so that only the value of $s_i(t)$ differs between the nodes, which corresponds to the network effects.

2.2. Modelling social networks. For these investigations we connected the nodes using several network models. *Erdős–Rényi* random graphs [7], where edges are connected with a probability p_e between pairs of nodes, were investigated for their analytical simplicity, due to the lack of correlations between neighbourhoods of nodes, allowing probabilistic arguments to be used [9]. The *Watts–Strogatz* scheme [18], where n-nearest-neighbour lattices have edges removed and reassigned according to a probability p_r, allowing interpolation between regular (possibly highly clustered) topologies and random networks, were also investigated, as they are frequently studied and provide a simple way to vary the clustering (*transitivity*) of the system.

Community based models, such as Newman's highly clustered networks [14], can better represent real-world social interactions in a number of ways. Here, links between nodes are assigned via their association with G *groups* each, which are randomly chosen from a set of W groups in total. Association with two or more groups per node is necessary for the network to be fully connected and total transmission to be possible. Nodes are linked to L other nodes within each group, so the total number of groups W and resulting number of members of each group $M = GN/W$ provides a means of varying the clustering while keeping other network parameters fixed. These networks also have locally higher clustering between members of a group than out to the rest of the network. For these models it would also be possible to include some geographical information to the nodes and groups in the network (c.f. Hamill and Gilbert's *Social Circles* model [12]), as discussed further in §5.

Although widely used, we do not consider *scale-free* networks, such as the *Barabási–Albert* (preferential attachment) model, here. This is because, as well as blurring the numerical results and complicating the analytical treatment by introducing a distribution of the node degrees, most individuals are only likely to communicate with a small selection of their contacts about adoption of innovation.

3. Numerical methods. Given a particular network, initial set of adopters and parameter values $(\alpha, \beta, \gamma, p, \theta)$, we ask whether or not the innovation *succeeds*, that is, do most individuals adopt the innovation. Aside from the challenges inherent in modelling the complex factors affecting behaviour, as well as in quantifying models from social data, dynamical processes on complex networks themselves often exhibit dependence on the starting conditions, such as which nodes are chosen as the initial seed. To illustrate this, the simulation results for a particular network and a single choice of parameters are given in Figure 3.1 (a), showing the uptake over 36 time steps for 100 random realisations of the initial seed. In each case, 5% of the nodes are chosen as *seed nodes* at random and are set to 1, with the rest initialised at 0. It can clearly be seen that while many runs lead to successful adoption, defined as most or all nodes becoming active, there is a clear subset of about 44% of initial conditions that stagnates and does not take over an appreciable proportion of the network.

Given this dependence, the ensemble average uptake over all 100 realisations of the choice of seed-nodes is computed for each chosen set of parameter values. The detailed structure of the network is randomised at each set of parameter values. In order to look at the influence of the parameters on the expected uptake, this calculation is repeated over a range of parameter values and plotted in the β, γ parameter space, with $0 \leqslant \beta \leqslant 1$, $0 \leqslant \gamma \leqslant 1$ and $\alpha = 1 - \beta - \gamma$. An example plot is shown in Figure 3.1 (b). Two lines can be seen dividing regions of the parameter space where adoption is total (black), where uptake is about 65% (blue) and where it stagnates at levels up to only slightly above the initial seeding level of 5% (yellow).

4. Results and analysis. The procedure described in the previous section was carried out for different network topologies in order to investigate the general behaviour of the model. The results all show distinct regions of either near total success or stagnation, with a few exceptions where there is a region of intermediate probability of success. In this section we present the results of these simulations along with an analysis into the origin of the dividing lines between these regions in parameter space. This is presented first in terms of Erdős–Rényi random (ER) networks, and later for other network models.
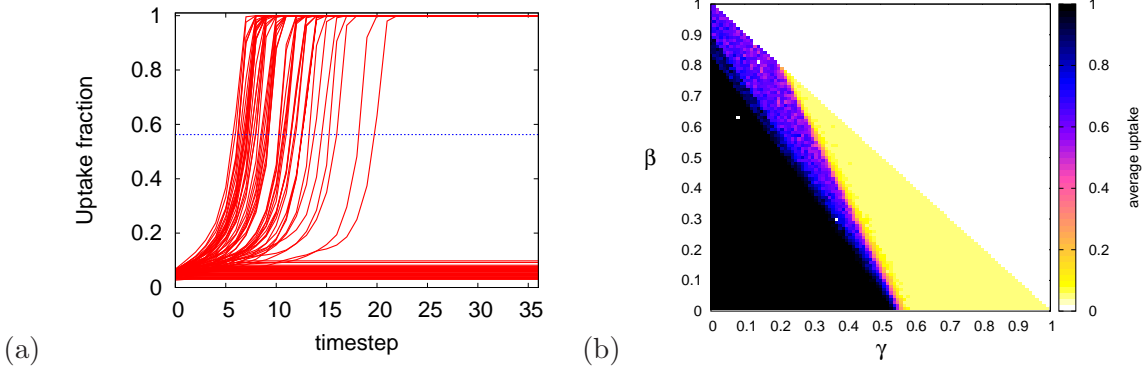
(a)   (b)

Figure 3.1. *Simulations of the adoption model on a community based network, with $N = 500$ nodes linked to others via their shared association with groups. Each node is randomly assigned to $G = 2$ groups, from a total of $W = 100$, and is linked with $L = 5$ other nodes per group. (a) Runs starting from 100 realisations of a 5% level of initial adoption ($m_0 = 0.05$), with model parameters $\alpha = 0.05, \beta = 0.8, \gamma = 0.15, p = 0.5, \theta = 0.25$. The mean uptake over the 100 realisations at time-step 36 is 56%. (b) The mean uptake at values in the β, γ parameter space, with $\alpha = 1 - \beta - \gamma$. All nodes take the same parameter values as each other.*

4.1. Number of neighbours needed to induce uptake. In order to understand which parameter values are likely to lead to a higher level of adoption, it is important to consider the properties of the nodes' local network environment. Each node experiences the same p and m in Equation (2.3), so the only difference between the nodes is s_i, the fraction of its neighbours that are active. For any set of parameter values and current adoption level $m(t)$, there is a critical fraction $s^*(t)$, defined by:

$$s^*(t) = \frac{\theta - \alpha p - \gamma m(t)}{\beta}, \tag{4.1}$$

such that if $s_i \geqslant s^*$, the node will adopt the innovation in the next time-step. The required number of active contacts Y_i^* (an integer) is given by:

$$Y_i^*(t) = \lceil k_i s_i^*(t) \rceil, \tag{4.2}$$

where $\lceil \bullet \rceil$ denotes the smallest integer $\geqslant \bullet$. If the actual number of active contacts

$$Y_i = \sum_j A_{ij} x_j \tag{4.3}$$

is at least Y_i^*, then node i will take up the innovation in the next time-step. To simplify matters, we consider first networks where the nodes all have the same (or similar) degree $\bar{k}$, defined by:

$$\bar{k} = \frac{1}{N} \sum_{ij} A_{ij}. \tag{4.4}$$

In this case we can define the critical number of active contacts to be

$$Y^* = \lceil \bar{k} s^*(t) \rceil = \left\lceil \bar{k} \left(\frac{\theta - \alpha p - \gamma m(t)}{\beta} \right) \right\rceil, \tag{4.5}$$

and the success or otherwise of the innovation uptake across the network can be understood in terms of Y^*. If $Y^* = 0$, then all nodes immediately adopt the innovation. If $Y^* = 1$, then any node connected to one of the initial seed nodes adopts the innovation, and the innovation uptake will be successful in a fully connected network. If $Y^* = 2$, then a node must be connected to two active nodes in order to adopt. We show below that for reasonable values of m_0, and for a wide class of networks, this leads to success. The case $Y^* = 3$ will also be discussed below.

Using $\alpha = 1 - \beta - \gamma$, Equation (4.5) can be inverted to give:

$$\beta = \bar{k} \left(\frac{\gamma(m(t) - p) + p - \theta}{\bar{k}p - Y^*} \right). \tag{4.6}$$

This gives, for different $m(t)$, and for different integers Y^*, a sequence of lines in the (β, γ) parameter plane. These lines (with $m(t) = m_0$) are overlaid on numerical results in Figure 4.1 for random (ER) networks of different size ($N = 500$, $N = 2000$) and different average degree ($\bar{k} = 6$, $\bar{k} = 15$). In Figure 4.1(a) ($N = 500$, $\bar{k} = 6$), uptake is successful for most values of (β, γ) to the left of the $Y^* = 2$ line. This boundary becomes sharper for $N = 2000$ and agrees more closely with the theoretical $Y^* = 2$ line in Figure 4.1(b). For $\bar{k} = 15$, $N = 500$ (Fig. 4.1(c)), it is the $Y^* = 3$ line that separates success from failure.

4.2. Probability of induced uptake in random networks. We begin by calculating the probability that the condition $Y_i \geqslant Y^*$ is satisfied. For random networks with independent random assignment of initial conditions, where all nodes have statistically homogeneous local network environment and the same node degree k, this is:

$$P(Y \geq Y^*) = \sum_{n=Y^*}^{k} \binom{k}{n} m^n (1 - m)^{(k-n)}. \tag{4.7}$$

This is just the probability of at least Y^* nodes being adopters, from k nodes in the neighbourhood, given that each node has a probability m of being an adopter. Clearly $P(Y \geq Y^*)$ is zero if $Y^* > k$.

For the specific examples in Figure 4.1 the outcomes in the different sectors of the parameter space can be understood using these ideas. If we define Z to be the number of nodes in the whole network whose neighbourhood influences them to adopt, then the probability that at least one further node will be influenced somewhere in the network is $P(Z \geq 1) = 1 - P(Z = 0)$, where the probability that no neighbourhoods are influential is $P(Z = 0) = (P(Y < Y^*))^N$, and $P(Y < Y^*) = 1 - P(Y \geqslant Y^*)$ which, using (4.7) gives the expression:

$$P(Z \geq 1) = 1 - \left(1 - \sum_{n=Y^*}^{k} \binom{k}{n} m^n (1 - m)^{(k-n)} \right)^N, \tag{4.8}$$

which, for small m approximates to:

$$P(Z \geq 1) \approx 1 - \left(1 - \binom{k}{Y^*} m^{Y^*} \right)^N. \tag{4.9}$$
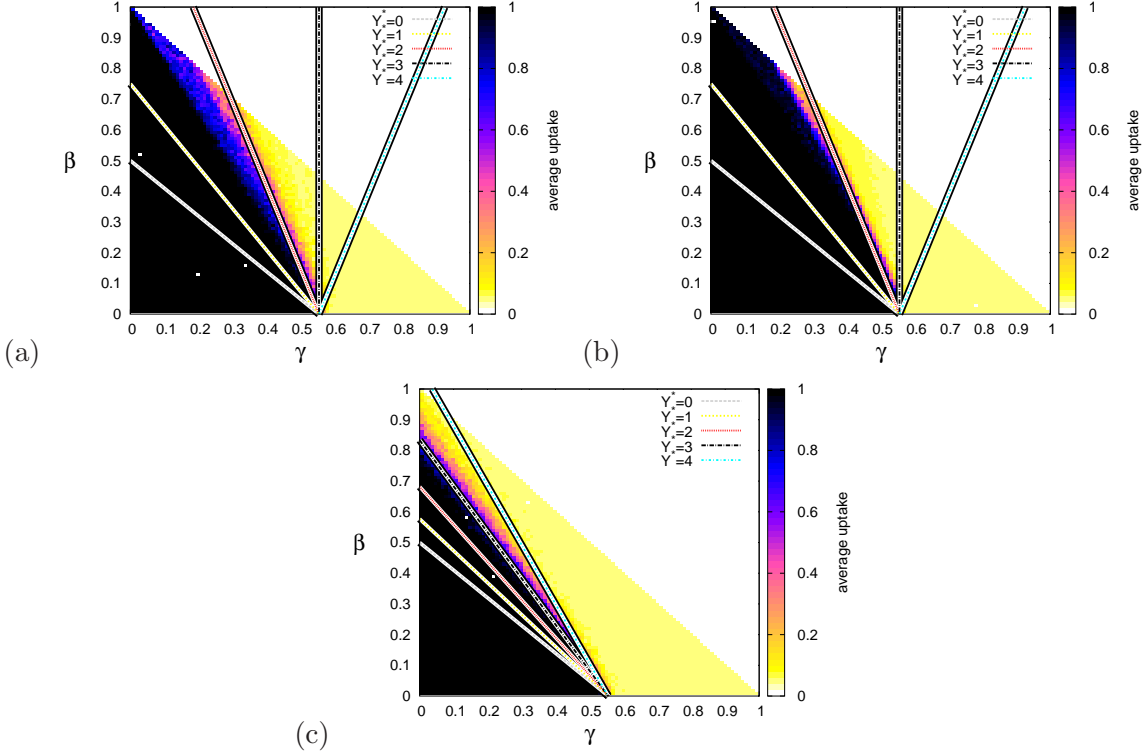
(a)

(b)

(c)

Figure 4.1. *Comparison of simulations of the adoption model on random Erdős–Rényi networks. As in all cases $\theta = 0.25$ and $p = 0.5$. The mean node degree $\bar{k}$ and network size N values are: (a) $\bar{k} = 6$, $N = 500$; (b) $\bar{k} = 6$, $N = 2000$; (c) $\bar{k} = 15$, $N = 500$. The results of the analysis in §4 are also plotted for $Y^* = 0, 1, 2, 3, 4$. The lines of Y^* in the (β, γ) plane show the upper limits of (β, γ) for each value of the critical number of neighbours Y^*. For example, two active neighbours are required to trigger adoption to the left of the $Y^* = 2$ line but this changes to three on the right side of the line.*

For $k = 6$, $N = 500$ and $m_0 = 0.05$ this gives $P(Z \geq 1) \approx 1$ for both $Y^* = 1, 2$, but for $Y^* = 3$, $P(Z \geq 1) \approx 0.7$. For the $Y^* = 2$ case, the probability of a site having an influential neighbourhood is $P(Y \geq Y^*) \approx 0.0375$ and the expected number of nodes satisfying this condition is this probability multiplied by the number of remaining (inactive) nodes $N(1 - m)P(Y \geq Y^*) \approx 18$, doubling adoption at each time-step. However, in the $Y^* = 3$ sector, the probability of any chosen node having the required number of neighbours is $P(Y \geq Y^*) \approx 0.0025$, resulting in only one expected new active node with a network size $N = 500$, easily leading to stagnation. The $Y^* = 2$ line becomes a sharper dividing line for larger networks (Fig. 4.1 (b)), as finite size fluctuation effects become less important. The value of Y^* which is the dividing line changes for $\bar{k} = 15$ (Fig. 4.1 (c)), where $N(1 - m)P(Y \geq Y^* = 3) \approx 27$, again of the order of the initial seeding level, and so adoption is successful for $Y^* = 3$, but fails when $Y^* = 4$.

4.3. Propagation on random networks. In Equation (4.5), Y^* is a decreasing function of time, since $m(t)$ can only increase. As argued above, if $Y^* = 0$ or 1, uptake will be successful. In this section, we estimate the likelihood that Y^* will decrease from a larger initial value to

1, and how long this will take.

The expected number of nodes that will become adopters in this time-step is the number of nodes that have not adopted the technology, $N(1-m)$, multiplied by this probability. The overall fraction of active nodes will therefore increase by:

$$\Delta m = (1-m)P(Y \geq Y^*), \tag{4.10}$$

giving the updated value of $m(t+1)$ as a function of the current value $f(m(t))$:

$$m(t+1) = m(t) + (1-m(t))P(Y \geq Y^*) \equiv f(m(t)). \tag{4.11}$$

Y^* is also a function of m but, being an integer, can only take discrete values, so $f(m)$ is discontinuous when Y^* changes.

This gives a macroscopic, system level, version of the microscopic (individual) dynamics given by Equation (2.2). The functions in (4.11) and (4.10) are shown in Figure 4.2 for some example parameter values. In Figure 4.2 (a) and (c) we have $Y^* > k = 6$ for $m < 0.1$, so
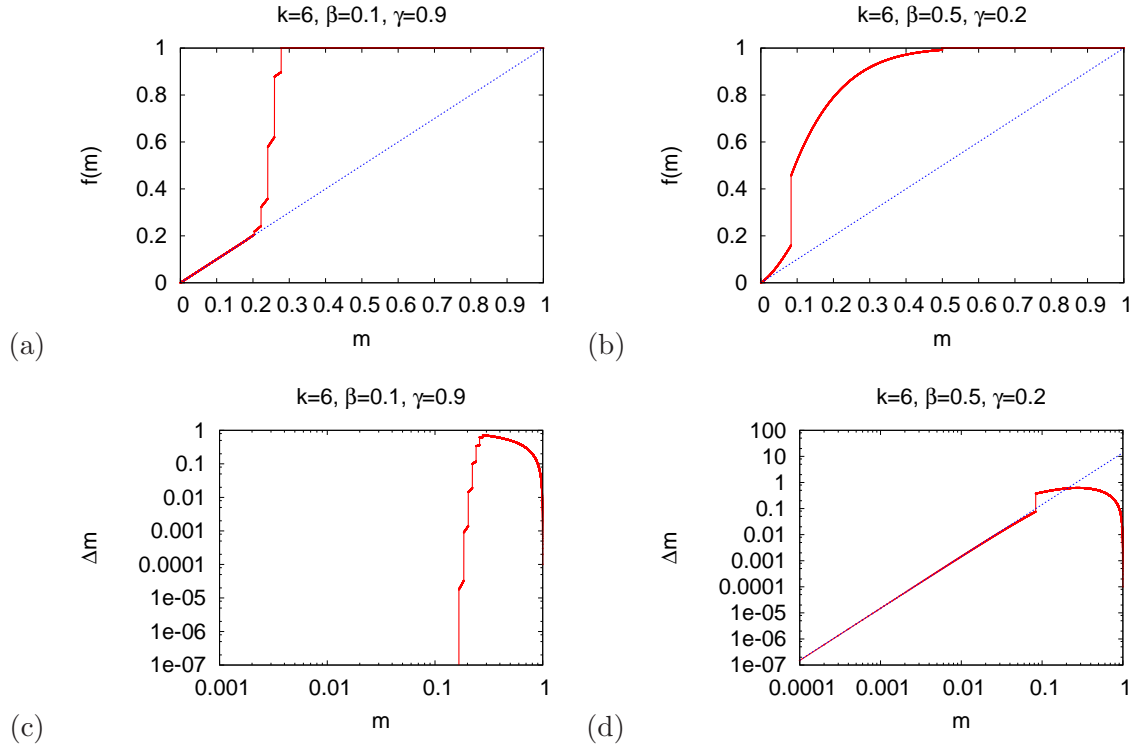

Figure 4.2. *Plots of Eqn. (4.11), for $k = 6$, (a) $\beta = 0.1, \gamma = 0.9$ and (b) $\beta = 0.5, \gamma = 0.2$, and Eqn. (4.10) at the same values in (c) and (d) (on a log scale). In (a) and (c) $Y^* > k$ for small m so there is no further uptake, and in (b) and (d) $Y^* = 2$ for small m so uptake is disproportionately slow for small levels of seeding. The diagonal is drawn for comparison in (a) and (b), while in (d) the fit to the initial slope is given.*

$\Delta m = 0$ in this range, and starting at an initial seed level of $m_0 = 0.05$ will lead to stagnation. If $m_0 \geqslant 0.2$, Δm is quite large and success will come quickly. In contrast, in Figure 4.2 (b) and

(d), $Y^* = 2$ for small m, and uptake can only happen if the network size N is large enough to allow further nodes to be influenced.

For small values of the initial adoption $m \ll 1$ the following approximation can be made:

$$P(Y \geq Y^*) \approx \Delta m \approx \binom{k}{Y^*} m^{Y^*}, \tag{4.12}$$

with $Y^* \approx \left\lceil \bar{k} \left(\frac{\theta - \alpha p}{\beta} \right) \right\rceil$ for small enough m. If $Y^* > 1$ then the super-linear power of m results in very small returns for small initial seeding, and it takes a time on the order of $m_0^{1-Y^*}$ for a successful outcome, for small m_0.

4.4. Structured networks. The approach above works for random networks, where independence of neighbourhoods can be assumed, and for randomly distributed initial seeds. For networks with high clustering, where the clustering coefficient or *transitivity*, defined as

$$c = \left(\frac{\#(\text{length 2 paths})}{3 \times \#(\text{triangles})} \right), \tag{4.13}$$

is close to 1, this is not the case. This difference is demonstrated in the numerical simulations shown in Figure 4.3. For these results the Watts–Strogatz scheme was used to rewire a highly clustered one dimensional lattice, where nodes are initially laid out in a ring and linked to their six nearest neighbours. Due to the large number of triangles, this network has a high transitivity $c = 0.6$, indicating a 60% probability that any two neighbours of a node are also neighbours of each other. For comparison, a completely random (ER) network of this size and edge density would have $c = 0.01$.

The probability of rewiring is p_r, and for small values of p_r where there is still significant clustering, the results show successful uptake in around 50% of cases in the $Y^* = 3$ sector, with either successful or stagnated adoption for Y^* values below or above this, respectively. This behaviour for $Y^* = 3$ would not be expected in random networks at these values, where the random probability of any particular neighbourhood triggering an adoption $P(Y \geq Y^*) \approx 0.002$, resulting in only a single expected further uptake per realisation. This would result in very slow uptake or stagnation on finite networks, which is indeed what happens for large p_r, where c is small and the network is again more random in structure (Fig. 4.3 (d)).

The enhanced success of uptake in these cases is due to the clustered structure favouring continuing propagation from any influential neighbourhood. In clustered networks, any neighbour j of a newly activated node i is likely to also be a neighbour of the nodes that influenced i to adopt, and so j has an enhanced chance of itself becoming an adopter. If a suitable neighbourhood is present under these circumstances, one uptake will then trigger another in an adjacent position due to this overlap of neighbourhoods (see Fig. 4.4). If clustering is sufficient that there is overlap between neighbourhoods for a sizeable portion of the whole network, then it can take only a single critical neighbourhood to induce a successful uptake for the system. The 50% chance of success seen in Figure 4.3 (a) in the $Y^* = 3$ sector then represents the chance of randomly seeding a group of nodes in such a way as to trigger this process. The probability given by (4.8), of at least one node in a random system having a critical neighbourhood, yields $P(Z \geq 1) \approx 0.67$ for $k = 6$, $Y^* = 3$ and $N = 500$. Equation
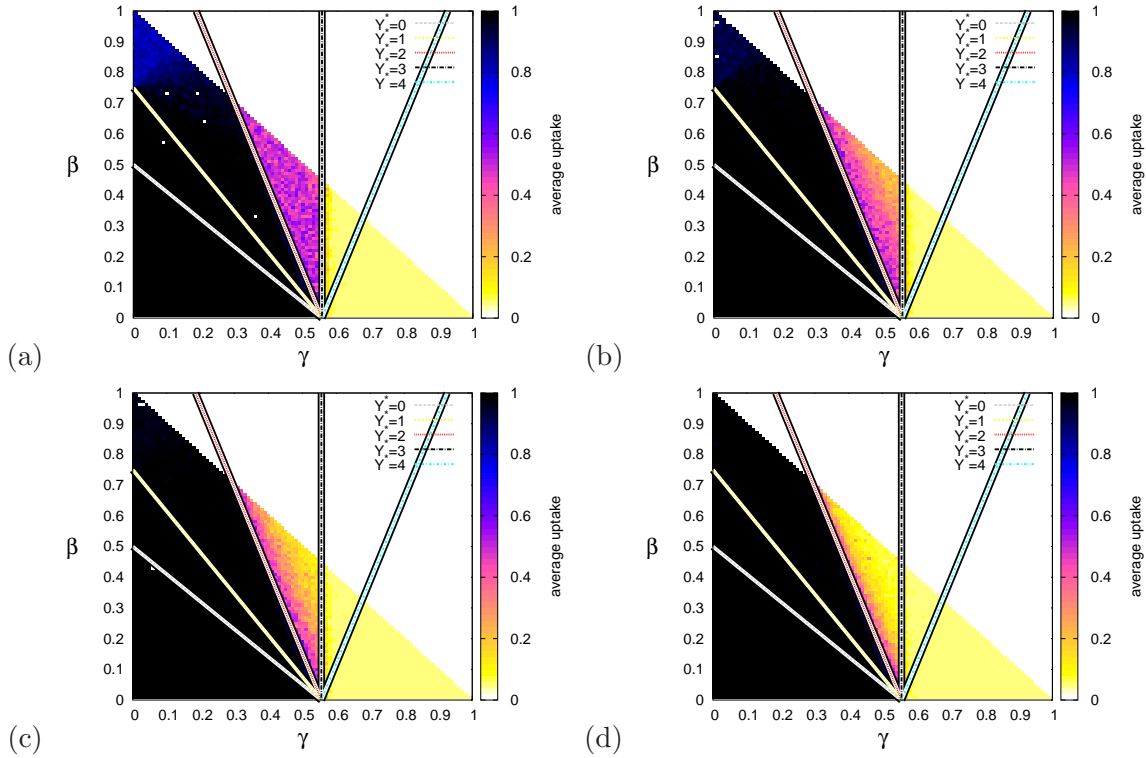
Figure 4.3. *Results on Watts–Strogatz networks with nodes initially connected to their six nearest neighbours on a one dimensional lattice and different rewiring probabilities p_r, resulting in different transitivity c: (a) $p_r = 0$, $c = 0.6$; (b) $p_r = 0.01$, $c = 0.57$; (c) $p_r = 0.02$, $c = 0.53$; (d) $p_r = 0.05$, $c = 0.45$. Model parameters are $p = 0.5$ and $\theta = 0.25$, with $N = 500$. Edges were swapped in pairs to preserve the node degree $k = 6$.*

(4.8) is strictly valid only for independent neighbourhoods, and hence gives too high a value in this case, where neighbourhoods overlap, but a quantitatively accurate expression would require a far more intensive calculation of the different combinations of the initial seed for the whole system. However, the current method is accurate enough to give qualitatively correct predictions.

The propagation via overlapping neighbourhoods is a percolation phenomenon, similar to that seen in other studies [14], but percolating on *hyper-edges* connecting *hyper-nodes*, rather than via single links. Here, a hyper-node is a complete sub-graph of $Y^* + 1$ nodes, or a $(Y^* + 1)$-*clique*. Two hyper-nodes are connected if the two $(Y^* + 1)$-cliques have Y^* nodes in common. The implication is that if a hyper-node becomes active, then every hyper-node to which it is connected will also become active in the next time-step. This is shown in Figure 4.4, for propagation via 4-cliques where $Y^* = 3$, with a single uptake triggering others via their overlapping neighbourhoods.

To demonstrate the transition between cluster-enabled propagation and stagnation, the ensemble uptake within the $Y^* = 3$ sector of Figure 4.3 (after 36 time-steps) is shown in Figure 4.5, plotted as a function of both p_r and c. The regular (un-rewired) network has a clustering coefficient $c = 0.6$, which decreases smoothly with increasing rewiring probability. If the

node degree is held constant by rewiring edges in pairs (Fig. 4.5 (a)) then the average uptake after 36 time-steps fits well with the function $m(36) = \sigma \exp(\varepsilon c)$, with fitting parameters $\sigma = 2.9 \times 10^{-4}$ and $\varepsilon = 12.39$. This is not the case if edges are rewired one at a time (Fig. 4.5 (b)), where the node degree becomes distributed. In this case the uptake peaks at around 5% rewiring ($c \approx 0.5$), similar to other models showing a peak in the level of adoption at certain node degrees [1].

A further example is shown in Figure 4.6 for a community based, random clustered network [14], where the clustering can be varied by associating N individual nodes via G groups of varying sizes M. The number of communities W and number of connections L per community (here $L = 5$) determines the size and density of the communities and hence the amount of clustering. In Figure 4.6 (d) the dependence of c on W is shown alongside the mean uptake after 36 time-steps, averaged over 1000 realisations for each W. Similar results can be seen as in the previous case, with increased clustering resulting in an increased likelihood of successful uptake, despite the low number of influential neighbourhoods. However, these expected values are only valid when averaging over different network realisations as well as initial adopters. For many networks it was found that uptake was almost never achieved for most choices of initial adopters, which we believe to be related to the connectedness of cliques of size Y^*+1. However, an interesting observation is the strong functional dependence of the expected success on the
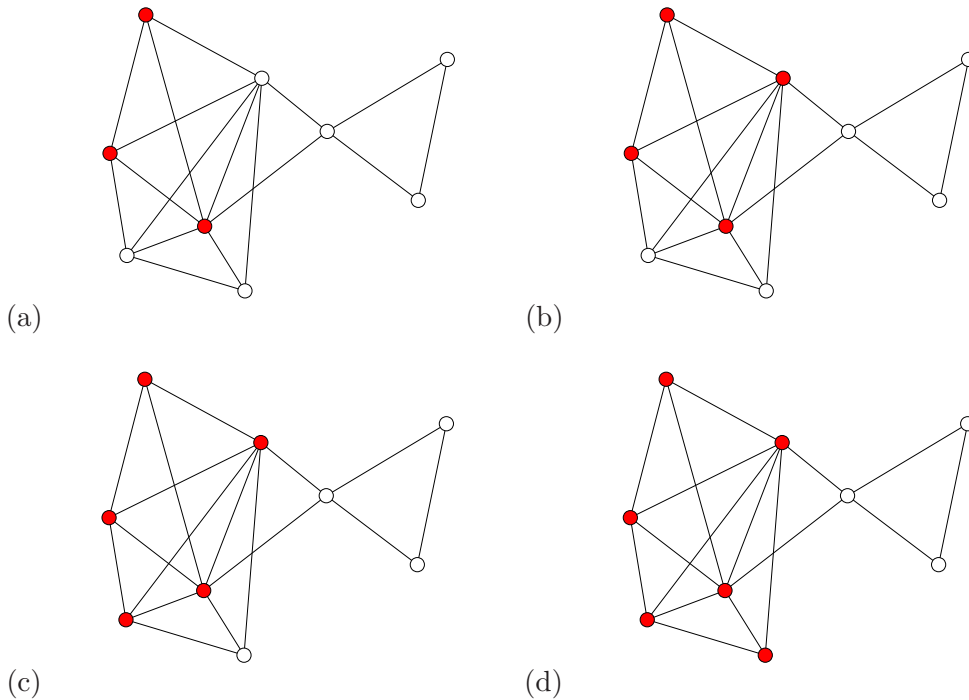

Figure 4.4. *An example of uptake propagation via clustering, in a network where nodes require three active neighbours to become active themselves. The sequence of uptakes is shown in (a)–(d), starting from a single node influenced by three of its neighbours, which then induces further uptake via 3-member overlap of 4-cliques.*
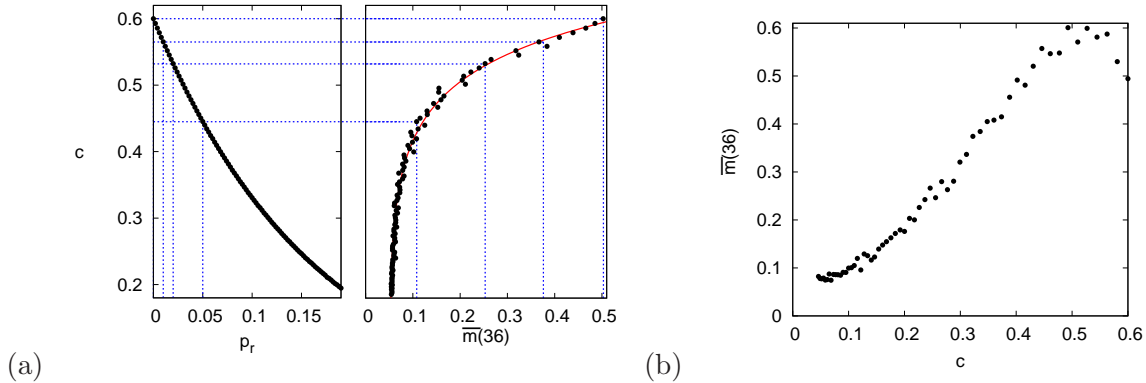
Figure 4.5. *Uptake on Watts–Strogatz rewired 6-neighbour rings, after 36 time-steps, against both rewiring probability p_r and clustering c. (a) Runs starting from 1000 random initial conditions with $m_0 = 0.05$, $\alpha = 0.1, \beta = 0.45, \gamma = 0.45, p = 0.5, \theta = 0.25$, randomising the network on each realisation. Edges are rewired in pairs to preserve the node degree $k = 6$. Dashed blue lines show the data in Fig. 4.3, with $p_r = 0, 0.01, 0.02, 0.05$ and $c = 0.6, 0.57, 0.53, 0.45$, respectively. The fitted red curve is the function $\sigma \exp(\varepsilon c)$, with $\sigma = 2.9 \times 10^{-4}$, $\varepsilon = 12.39$. (b) For comparison, the same system is used but with edges rewired individually between $0 < p_r < 0.5$, showing the effect of allowing distribution of the node degree with p_r. A peak in the uptake appears at $c \approx 0.5$.*

transitivity, which fits to about 1% accuracy to the function:

$$m(36) = 1 - \sigma \exp(-\varepsilon c), \tag{4.14}$$

with fitting parameters $\sigma = 0.86$, $\varepsilon = 11.2$. This relationship was also found for different values of (α, β, γ) within the $Y^* = 3$ sector, with different coefficients in (4.14). This suggests that it is possible to obtain an analytical expression for these systems, and should motivate further research.

The phenomenon of connectedness via more highly connected clusters than single edges is known as k-clique percolation, and results exist deriving this result for random (ER) networks [6, 15]. However, to understand the results in Figure 4.6 analytically would require results for the expected connected $(Y^* + 1)$-clique percolation component size distribution as a function of the clustering coefficient. Some progress has been made in this direction, with numerical results existing for cascades of various models, such as Watts' threshold model [17] on clustered random networks like the Newman model [11]. Additionally analytical results have also been derived for expected connected component sizes for single bond percolation on clustered networks [8]. However, obtaining the required results for k-clique percolation components remains an open and difficult problem.

The results in this section imply that, in cases where several network neighbours are required to influence adoption, greater levels of system-wide adoption can be achieved in more highly clustered networks, in agreement with earlier results on similar systems [4]. Therefore, innovations in systems with more communication between overlapping groups of individuals will be more likely to succeed than otherwise.

5. Conclusions. Numerical simulations have been carried out, computing the average uptake of an innovation by householders connected by a network of peer to peer influences.
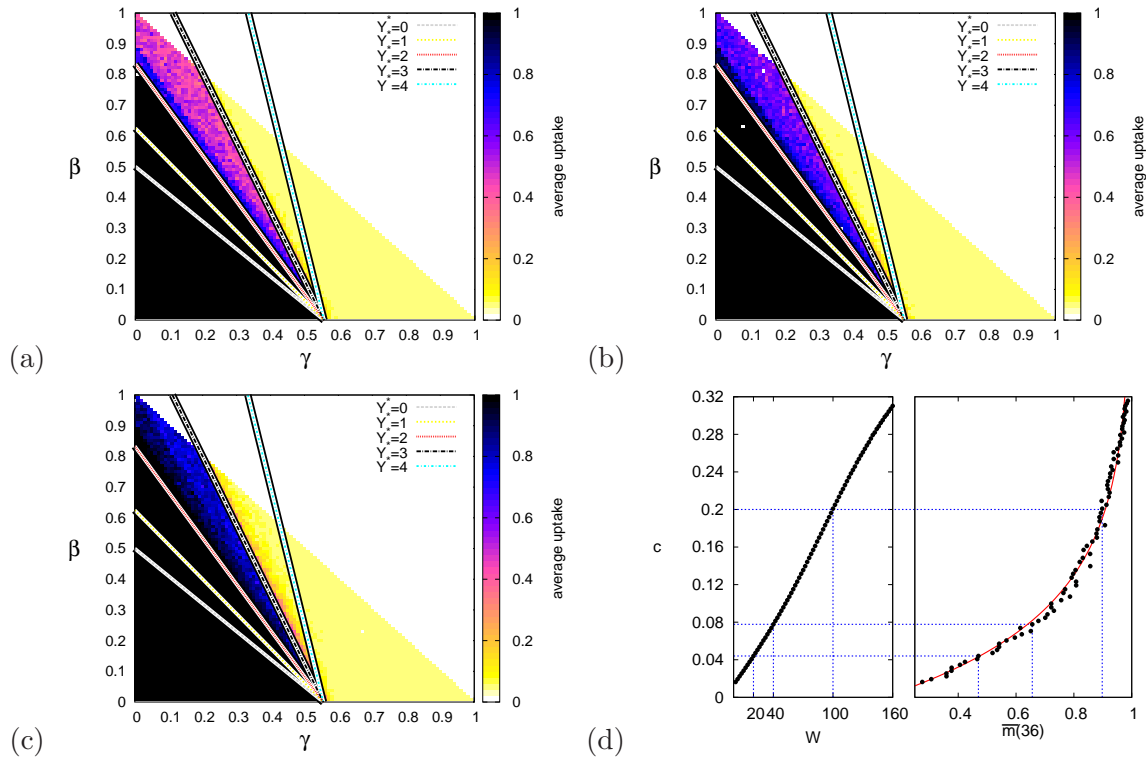
Figure 4.6. *Same as Figs. 4.3 and 4.5 but for a community-based, Newman random clustered, network model. The network size $N = 500$, so for each number of groups W there are M members, on average, resulting in a measured transitivity c: (a) $W = 20$, $M = 25$, $c = 0.044$ (b) $W = 40$, $M = 12.5$, $c = 0.078$ (c) $W = 100$, $M = 5$, $c = 0.2$ (d) Clustering as a function of W versus uptake for $\alpha = 0.15, \beta = 0.8, \gamma = 0.05$ within the $Y^* = 3$ sector. The fitted red line is the empirical function $m(36) = 1 - \sigma \exp(-\varepsilon c)$, with $\sigma = 0.86$, $\varepsilon = 11.2$.*

This was carried out over many realisations, at a range of model parameter values, on various network topologies and initial conditions. It has been found that the expected level of adoption depends strongly on the topology of the network and the Y^* sector that the (β, γ) parameters are in. The node degree and the network's transitivity, related to the correlation between the neighbourhoods of connected individuals, were found to be of particular importance.

An analytical and probabilistic approach has been developed to account for the observed behaviour. The analysis explains the results of the numerical calculations well in the case of randomly connected networks, and provides useful insight in the case of networks with clustering.

The model provides a basis to simulate the behaviour seen in the real world. For example, it could form the basis for computational investigations comparing different interventions by a government to try to improve the level of uptake of innovations relating to energy efficiency. These could be represented in the model by varying the initial seeding for different roll-out strategies, or network structure for different word-of-mouth marketing campaigns. By identifying intervention strategies that have a high probability of achieving successful uptake of the technology, the model could allow government to make informed decisions of whether

to go ahead with interventions and if so, how. This will be the subject of a future paper.

Future studies can modify the community-based network models to provide more flexibility and realistic features in a number of ways. Some positional information can be ascribed to the nodes and some or all of the groups, preferentially attaching individuals to communities that are geographically closer, in a similar way to Hamill and Gilbert's *Social Circles* model [12]. In cases where none (or only one) of the groups have geographic information they again reduce to the Newman scheme. These groups could be classified into different categories, for example some could have a fixed membership number to allow variation of the clustering between groups and represent different types of group interaction (e.g., social or workplace groups). For other types there could be no geographical element, introducing long-range connections into the network, as may be the case for workplaces.

The analysis of the model made use of the assumption of a near-constant node degree, and so would not carry over simply to scale-free networks. In this work we argued that communication on innovation adoption is restricted to a select few contacts, therefore excluding wide degree distributions. An alternative would be to assume that individuals are influenced by a *number* rather than *fraction* of their network neighbours, which is independent of $\bar{k}$.

The dynamical adoption model could be made to account for the various archetypes discussed in the social science literature by assigning individuals to different groups. Individuals in different categories would be assigned different α_i, β_i and γ_i values, depending on their personality, for example with technology enthusiasts having high α values, social followers having high β, and cautious individuals having a high γ. The model could also be given a stochastic element by allowing a distribution of time-step values, and introducing a probabilistic, rather than deterministic, uptake once the utility exceeds its threshold.

The methods discussed in this paper can also be applied to multi-parameter models of different systems with similar features, and provide a framework for the investigation of many real-world systems.

REFERENCES

[1] F. ALKEMADE AND C. CASTALDI, *Strategies for the diffusion of innovations on social networks*, Computational Economics, 25 (2005), pp. 3–23.

[2] S. BIKHCHANDANI, D. HIRSHLEIFER, AND I. WELCH, *A theory of fads, fashion, custom, and cultural change as informational cascades*, Journal of Political Economy, (1992), pp. 992–1026.

[3] C. CASTELLANO, S. FORTUNATO, AND V. LORETO, *Statistical physics of social dynamics*, Reviews of Modern Physics, 81 (2009), pp. 591–646.

[4] H. CHOI, S.H. KIM, AND J. LEE, *Role of network structure and network effects in diffusion of innovations*, Industrial Marketing Management, 39 (2010), pp. 170–177.

[5] S. A. DELRE, W. JAGER, T. H. A. BIJMOLT, AND M. A. JANSSEN, *Will it spread or not? The effects of social influences and network topology on innovation diffusion*, Journal of Product Innovation Management, 27 (2010), pp. 267–282.

[6] I. DERÉNYI, G. PALLA, AND T. VICSEK, *Clique percolation in random networks*, Physical Review Letters, 94 (2005), p. 160202.

[7] P. ERDŐS AND A. RÉNYI, *On the evolution of random graphs*, Akad. Kiadó, 1960.

[8] J.P. GLEESON, *Bond percolation on a class of clustered random networks*, Physical Review E, 80 (2009), p. 036107.

[9] J.P. GLEESON, S. MELNIK, J.A. WARD, M.A. PORTER, AND P.J. MUCHA, *Accuracy of mean-field theory for dynamics on real-world networks*, Physical Review E, 85 (2012), p. 026106.

[10] M. GRANOVETTER AND R. SOONG, *Threshold models of diffusion and collective behavior*, Journal of Mathematical Sociology, 9 (1983), pp. 165–179.

[11] A. HACKETT, S. MELNIK, AND J.P. GLEESON, *Cascades on a class of clustered random networks*, Physical Review E, 83 (2011), p. 056107.

[12] L. HAMILL AND N. GILBERT, *Social circles: A simple structure for agent-based social network models*, Journal of Artificial Societies and Social Simulation, 12 (2009).

[13] E. LEE, J. LEE, AND J. LEE, *Reconsideration of the winner-take-all hypothesis: Complex networks and local bias*, Management Science, (2006), pp. 1838–1848.

[14] M.E.J. NEWMAN, *Properties of highly clustered networks*, Physical Review E, 68 (2003), p. 026121.

[15] G. PALLA, I. DERÉNYI, AND T. VICSEK, *The critical point of k-clique percolation in the erdős–rényi graph*, Journal of Statistical Physics, 128 (2007), pp. 219–227.

[16] T.W. VALENTE, *Social network thresholds in the diffusion of innovations*, Social Networks, 18 (1996), pp. 69–89.

[17] D.J. WATTS, *A simple model of global cascades on random networks*, Proceedings of the National Academy of Sciences, 99 (2002), p. 5766.

[18] D.J. WATTS AND S.H. STROGATZ, *Collective dynamics of 'small-world' networks*, Nature, 393 (1998), pp. 440–442.