

Diffusion of innovations on community based Small Worlds: the role of correlation between social spheres

Emily TANIMURA

Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon-la Sorbonne-CNRS,

106-112, Bld de l'hôpital, 75013, Paris

e-mail: emily.tanimura@univ-paris1.fr

Abstract: Which types of networks favor the diffusion of innovations in the sense that an innovation that has higher intrinsic benefits than an established choice will be able to replace the latter in a large population when it is initially used by a small part of the population? For deterministic and regular networks there are characterizations, based on a coordination game model of the diffusion of innovations. Here we analyze this question for a class of irregular random networks, Small world networks, which are of interest as more realistic models of social networks. We consider a parametric community based model which covers as particular cases several well known models such as Watts' Small world. We show that there are different types of Small World graphs some which favor diffusion others that do not. Our study suggests that the kinds of ties that exist between different communities of an individual play an important role. We interpret Watts' Small World as one with high correlation between social spheres of individuals and favorable to diffusion. In other Small Worlds where the communities of individuals are uncorrelated diffusion succeeds only for very large payoff benefits in favor of the innovation.

Keywords and phrases: diffusion of innovations, Small World networks, contagion threshold, community structure, social networks.

1. Motivations

A number of classical studies, such as the one by Coleman et al [8] of physicians' willingness to prescribe new drugs, or by Ryan and Gross [36] of the introduction of agricultural innovations in rural communities have shown that new technologies, norms or other innovations are often spread through social networks and are gradually adopted under the influence of friends of acquaintances who have already converted to them. This means that not only the inherent benefits of some behavior but also the social network through which the agents can influence each other becomes relevant for determining

whether a new norm will spread or a new technology will take hold. A natural question to ask and which has been the object of studies by Morris [31] or Young [46] among others is to determine which are the types of networks that allow new technologies or behaviors with greater inherent benefits to establish themselves in a population where some other pre-existing technology, norm etc is prevalent. Whether we are interested in the spread of new norms, the introduction of new products or innovations, from a formal point of view, the problem boils down to the study of a diffusion process on some social network. Moreover, given the nature of the question: whether spread can occur from a small clique of initial adopters to a larger groups or population, it is natural to restrict our attention to networks of large size.

In reality, it is plausible that influence between network structure and decisions to adopt some form of behavior goes in two directions. Network neighbors may influence each other to make similar choices, but conversely, agents might decide to reinforce or reduce their interactions depending on their affinity, that is the similarity of their decisions, leading to a complex interplay between network evolution and adoption. This approach has been taken for example by Steyer and Zimmermann [37]. However, the more basic question of how a fixed network structure influences this type of coordination process is already interesting. This has been the most common approach in the literature and the one we will follow here. To address the question of whether choices with intrinsic benefits could spread in a population, Blume([1],or Ellison ,[13])considered a coordination game in which players were assumed to play a myopic best response or trembling hand best response to their neighbors previous choices. These early studies showed contrasting outcomes on different types of networks: low dimensional lattices allow the payoff dominant action to spread easily while regular trees of high degree make it difficult for spread to occur unless the benefits of the innovation are unrealistically high. Taking a somewhat different approach, Morris [31] asks whether there are criteria that can be applied to any graph and which measure how easy it is for diffusion or "contagion" to occur on it. For this purpose, he introduces a notion "of contagion threshold" which measures how easily an innovation spreads in a graph. The contagion threshold is related to the lowest payoff benefit that allows the innovation to spread. Indeed, if the innovation is by far superior to the established norm, it should be able to spread in any network. If the innovation is only marginally better than the previous norm, then its spread relies more heavily on social factors. This is thus reflected in the contagion threshold. Morris shows that the contagion threshold can be determined based on the rate of neighborhood

growth under some assumptions about the homogeneity of the network. Similarly, Young establishes criteria for contagion but in the case where agents use trembling hand best response and not pure best response in the coordination game. Both studies are concerned with analysis for large sized networks. Morris uses a setting with an infinity of agents interacting in a regular graph, Young obtains results for N goes to infinity.

However, the criteria developed in [31] or Young rely on the fact that the networks are highly regular. Moreover, they apply only to deterministic graphs. This excludes many interesting models of large sized networks. In particular, a number of recent random graph models which have been developed to capture more accurately certain empirical properties of social networks.

A number of empirical studies and experiments (see for example Granovetter [19] or Freeman and Thompson [16]) have allowed to determine that many networks, in particular social networks are characterized by what is referred to as Small World properties: high clustering and short distance between agents . This makes it interesting to explore the behavior of social dynamics such as the diffusion of innovations, on these types of networks whose properties are closer to those of real networks than for example lattices,

A first model that captured the small world properties was proposed and studied by Watts [43] . Subsequently, a number of other models for capturing these properties have appeared in the literature (a survey can be found in Bonato [3]). However, many of these are aimed at applications in computer science and not suitable for describing social networks. Moreover, none have been as influential as that of Watts. In fact to some extent a small world graph has almost come to be identified with this particular model.

As a consequence, most studies of games or other types of interactions on Small World graphs are, in reality, studies on a particular model, that of Watts. However, going back to Watts original definition, it should be remembered that small world graphs are in fact a class of graphs sharing a certain number of asymptotic scaling properties. For a mathematical definition of these properties, see Cont and Tanimura [9]. Therefore, in order to infer something about behavior on real networks, one must show that results are generic to this class of graphs.

As far as the diffusion of innovations is concerned, the results in this paper are rather a counter example: they show that the behavior of this type

of diffusion, even described in a simplified manner, cannot be characterized based on the Small World properties alone. We identify different types of small world networks which exhibit different behaviors with regards to spread. However, these comparisons allow us to identify different types of small world networks in which diffusion is more or less facilitated and relate these differences to a quantity that has a meaningful sociological interpretation.

The class of Small world models on which we will study the diffusion of innovations, will be random graph models which are all based on an underlying community structures or social group structure. Community structures are present in many real network, rather obviously in social ones. The random graph modeling we propose does not supply a strategic foundation for the models but it is possible that the resulting networks could be obtained as equilibria of some network formation game. Here our objective is to study the impact of these structures on the diffusion of innovations and not the origin of the structures themselves.

Compared to many other random graphs models, the ones we consider here are specifically aimed at social networks. Structures and parameters are easy to interpret from the point of view of social science. Another appealing aspect of the model is that it englobes as particular cases several well known models that have been studied previously. Like Watts' Small World, the model we present here is regulated by the variation of a parameter that has a precise empirical meaning: it expresses the correlation between the different social circles of the individual which can be either overlapping or disjoint, depending on the type of sociability. For example, sociologists (see Degenne and Forsé [10]) note that in rural environments, the different social spheres of an individual tend to overlap, whereas urbanites often participate in several disjoint spheres of sociability.

As a function of the parameter that regulates correlation, we can obtain, for example, a lattice like structure, a Small World model with properties close to that of Watts' one or to the "island model" that appeared as the equilibrium of a strategic game in Jackson and Rogers [21] or a model with independent communities previously studied in Cont and Tanimura [9]. In particular, we can now interpret Watts' Small World as a network in which the correlation between social spheres of individuals is high but not total.

In the parametric community model, we study a notion of asymptotic contagion threshold, similar to that of Morris but for random graphs. We note that in inhomogenous networks, such as the ones we consider, it is

possible that spread would not occur from a typical, i.e. randomly chosen connected set of initial adopters but that some set of initial users from which spread would occur exists and could be found if one has perfect knowledge of the network. Whether such a group of agents exists could be of interest for example in marketing. For this reason, we also consider this possibility when bounding the contagion thresholds, although it complicates proofs. We obtain asymptotic analytical results for the extreme values of the correlation parameter. Numerical illustrations for finite N are given in simulations. We also consider the case of trembling hand best response, following Young. Within this class of community based models, our results suggest that the ability of the innovation to spread at a reasonable level of payoff benefits is closely linked to the level of correlation between the different communities of an individual. This suggests that in order to predict the success of an innovation, it is relevant to measure this quantity rather than the clustering which is a necessary but not sufficient condition for spread in this type of process.

2. Organization of the material

We begin, in section 3, by presenting the general community based model and motivations for it. We also presents the special cases of interest. In section 4, we recall the coordination game presented by Morris, the version of this model with trembling hand best response, following Young [46] and some of the main existing results. In section 5.1, we discuss how to define the contagion threshold of a random graph in an analogous way to that of Morris in the deterministic setting with infinite node space. We establish results about asymptotic contagion thresholds in the uncorrelated community model and in Watts' Small world. In section 5.2, we provide numerical examples based on simulations in a pre asymptotic setting. The purpose of this is to confront the asymptotic with behavior in large finite sized systems. In 5.3, we analyze the diffusion of innovation with trembling hand best response by adapting the criteria of Young. Section 6 concludes.

3. The networks: community based small worlds random graphs

As the name suggests, the community based model we will now present is based on the idea that individuals in a social network belong to one or several communities. These communities could be of a professional nature,

families, students in the same course or individuals who practice some activity together. This corresponds to a vision of networks in which the building blocks are communities rather than individuals which we think is reasonable for the general social network of an individual, that is the totality of his friends and acquaintances. For example, Wellman et al. [45], in a study of social networks in East York, found that many links are not autonomous but are embedded in larger clustered groups such as the group of co-workers or families. They estimated 81 % of links to be of this type. The importance of social circles as building blocks in networks have been emphasized by a number of sociologists, see for example the review in Degenne et Forsé [10], chapter 2. We also find a vast literature that deals with algorithms that identify community structures in network data (minimum-cut, hierarchical clustering, see Johnson et al [24], or more recent methods such as that of Girvan et al. [17]). However, many of these assume that each individual belongs to a single community which is to be identified. In our model, the assumption that an individual may belong to several communities is important. Moreover, it seems to us that the literature that identifies communities in empirical data has not found a counterpart in the modeling literature. At least random graph models have not attached much attention to community structures but have rather focused on clustering. It should be pointed out that while networks with communities are necessarily clustered, the presence of clustering does not necessarily mean that there is a clear community structure. For example, lattices are clustered but it is hard to define the community of an individual in an unequivocal manner.

In the following section we present our model. Let us point out that the construction as such does not make claims to any realism. It is the resulting graph that should be a reasonable representation of a social network.

3.1. Defining the community model

Consider M complete graphs $\mathcal{G}_1, \dots, \mathcal{G}_M$ with disjoint node sets, which we will call *clusters*. These clusters can be seen as 'communities of origin' of the nodes. We shall take them to be of equal size δ for simplicity, though this assumption can be relaxed. Starting from the (disconnected) graph $\bigcup_{m=1}^M \mathcal{G}_m$ with $N = M\delta$ nodes, we add random links to it in the following way.

We associate to every node $i = 1..N$ a cluster X_i drawn among $\mathcal{G}_1, \dots, \mathcal{G}_M$. X_i can be viewed as the 'secondary community' of the node i . In the final construction, all nodes who share a primary or secondary community will be linked. We want to determine the laws of the $(X_i)_{i=1}^N$ so as to allow for correlation between the first and secondary cluster of an agent in such a

way that agents who are in the same community of origin may have a higher probability of sharing their secondary community as well. We do this in the following way: For every $m = 1, \dots, M$, we order the nodes in \mathcal{G}_m in some arbitrary order, j_1, \dots, j_δ . For $1 \leq k \leq \delta$, we define $n_k^m = \sum_{l=1}^{l=k-1} \mathbf{1}_{X_{j_l} = \mathcal{G}_m}$.

Now, we draw X_{j_l} so that $P(X_{j_l} = \mathcal{G}_m) = \frac{e^{\alpha n_k^m}}{\sum_{i=1}^{i=M} e^{\alpha n_k^i}}$. This means that the

secondary community of X_{j_l} is drawn uniformly at random and for all other nodes in \mathcal{G}_m , the secondary community of is drawn preferentially among those to whom other nodes in \mathcal{G}_m are linked. When $\alpha = 0$, we have the independent case. When α goes to infinity, the nodes in the same initial cluster will almost surely be in the same secondary cluster. As a function of the parameter α , we obtain as particular cases:

- **Independent case** ($\alpha = 0$) This is the model with independence between the two communities of an individual. This is the model that was studied in Cont and Tanimura [9]. The $(X_i)_{i=1}^{i=N}$ are drawn independently and uniformly at random. The resulting graph is obtained as follows: Define for $m = 1..M$

$$A_m = [\mathcal{G}_m] \cup \{i = 1..N, X_i = \mathcal{G}_m\} \quad (1)$$

Then

$$\Gamma^N = \bigcup_{m=1}^M A_m \times A_m. \quad (2)$$

- **Islands model**

The islands model is obtained as a minor variation of the previous model. As before, the $(X_i)_{i=1}^{i=N}$ are drawn independently and uniformly at random. However, this time we will link a node to only one (or more generally less than δ nodes) node in his secondary community: For every $m = 1, ..M$ we consider the nodes $\{i | X_i = \mathcal{G}_m\}$ that will have a link to some node in \mathcal{G}_m . We distribute the links among nodes in \mathcal{G}_m by drawing for each node in $\{i | X_i = \mathcal{G}_m\}$, a node Y_i chosen uniformly among those chosen the fewest times previously, and create the link (i, Y_i) . Thus we have

$$\Gamma^N = \bigcup_{m=1}^M \mathcal{G}_m \times \mathcal{G}_m \bigcup_{i=1}^{i=N} (i, Y_i) \quad (3)$$

This generates a graph based on communities or "islands" connected by single links which resembles the equilibrium graph obtained in a

strategic network formation game proposed by Jackson and Rogers [21]

- **The perfectly correlated case** ($\alpha \rightarrow \infty$) In section 3.2, we show that there is a community based model with perfect correlation between secondary communities of nodes within the same initial community which has a structure similar to a one-dimensional ring lattice
- **high correlation (Large fixed α) and relation to Watts' model** In section 3.2, we will show that the community model with strong but not perfect correlation between communities has properties similar, in a sense that we will define, to Watts' Small World. Recall that Watts' model is defined by superposing on a set of N nodes, a one dimensional ring lattice L and an Erdős–Renyi random graph E . In the ring lattice, each node is connected to the $2k$ neighbors within distance k in the lattice. In the Erdős–Renyi graph each link exists with probability $\frac{p}{N}$. The resulting graph Γ_N has links defined by $(i, j) \in \Gamma_N$ if and only if $(i, j) \in L$ or $(i, j) \in E$.

3.2. Comparison of Watts' model and the strong correlation case

We will consider the general model described above but with a minor modification where secondary communities are drawn uniformly at random in $(\mathcal{G}_m)_{m=1}^{m=M}$ but without replacement. In this case we obtain a graph that is almost identical to Watts' Small World as we see from proposition (1). If secondary communities are drawn uniformly at random with replacement, we obtain a connected component with a "locally lattice like structure". This case is treated in the appendix. First, let us consider the case $\alpha = 0$. In this case, all agents in the same initial community or cluster are also in the same secondary cluster. Suppose that we draw $(X_i)_{m=1}^{m=M}$, uniformly among $(\mathcal{G}_m)_{m=1}^{m=M}$ but without replacement, so that each cluster is the secondary cluster of exactly one other cluster. It is then easy to see that we have a ring structure. Since all nodes in \mathcal{G}_m are linked to all the nodes in cluster X_m , it is easy to see that we have the following proposition:

Proposition 1. *Let the size of the initial clusters be δ . Then the graph Γ_N defined by the procedure above has a link set $\{(i, j) | (i, j) \in \Gamma_N\}$ that includes the link set of a one dimensional ring lattice where agents are linked to all neighbors at distance at most δ and is included in the link set of a one dimensional ring lattice where agents are linked to all neighbors at distance at most 2δ .*

Consequently, fixing a large α , the probability that some agent has a

secondary cluster different from the other agents in his initial community is small but remains positive. This secondary cluster being drawn uniformly at random, corresponds to the uniformly random links or "shortcuts" in Watts' model. In this sense, it is natural to interpret a one dimensional lattice as a community model with perfect correlation between the secondary communities of agents who share some community, and Watts' model as a network in which correlation is high but not complete.

4. Diffusion of innovations modeled as a coordination game in a population

Consider a coordination game with the payoff matrix below (we could consider somewhat more general payoffs as [31],[46] but this changes nothing in the analysis) :

$$\begin{vmatrix} u(1,1),u(1,1) & 0,0 \\ 0,0 & u(0,0),u(0,0) \end{vmatrix}$$

where $u(1, 1) > u(0, 0) > 0$. If an agent plays this coordination game with all of his network neighbors and averages his payoffs, then the utility he derives from his choice a_i depends on the intrinsic value of this choice and on his neighbors' actions:

$$U^i(a_i, \cdot) = \frac{1}{\text{Card}(V(i))} \sum_{j \in V(i)} u(a_i, a_j) \quad (4)$$

Following Morris [31], we assume that agents perform myopic utility optimization, choosing at time t the best response to their neighbors' choices at $t - 1$. There is a critical smallest value q such that if at least the fraction q of an agents neighbors chose 1 in the previous period, 1 will be the agents best response.

$$U_t^i(1) \geq U_t^i(0) \Leftrightarrow \frac{1}{\text{Card}(V(i))} \sum_{j \in V(i)} a_{t-1}^j \geq \frac{u(0,0)}{u(0,0) + u(1,1)} = q \quad (5)$$

In a graph G , where the initial set of nodes in state 1 is I^0 , the set of agents who will choose 1 at time k is defined recursively: $(I^q)^k(G) = \{j \in X \mid \frac{\text{Card}(V(j) \cap (\bigcup_{l < k} I^l))}{\text{card}(V(j))} \geq q\}$. These sets depend on the graph G through the neighborhoods $V(j)$.

Whenever $u(1, 1) > u(0, 0)$, it would be optimal if the population coordinated on 1, but this is not a guaranteed outcome if few agents chose 1

initially. Morris considers the case where the node space X is infinite. The contagion threshold of a graph, which is a measure of how easy it is for an action with better payoffs to spread in the graph, is defined as

$$\eta(G) = \max_q \{q \mid \bigcup_{k \geq 0} (I^q)^k(G) \text{ is cofinite, for some finite } I^0\} \quad (6)$$

(a set is cofinite if its complement is a finite set). Thus the contagion threshold of a graph is the largest value of q in the coordination for which the payoff dominant action is able to spread from a finite set to the whole (infinite) system. The contagion threshold is always inferior to $\frac{1}{2}$. The closer it is to this upper bound, the easier it is for the payoff dominant action to spread in the graph. We note that if η is low, the action 1 can only spread if $u(1, 1) \gg u(0, 0)$. If $\eta = \frac{1}{2}$, action 1 will spread whenever $u(1, 1) > u(0, 0)$. (We note that in a coordination game, it is reasonable that the payoff difference between the actions should not be too great so that q is fairly high. Thus, although the description is formally the same as an infection dynamic, in the latter, it is sufficient to know one person infected by a disease to catch it, which corresponds to a low q . The spread of norms on the other hand requires that an individual has a sizeable fraction of his neighbors who adopted the norm in order for him to adopt it, unless one of the choices has payoffs greatly superior to those of the other one. If a graph has a low contagion threshold, the payoff dominant action would not be able to spread for most reasonable payoffs. For example, a contagion threshold of $1/10$ would require the payoff dominant action to have a payoff 9 times greater than the other action.)

4.1. Coordination game with trembling hand best response

As before, agents use a myopic best response to their neighbors previous choices but the response is smoothed so that the best response is chosen with high probability but the other action has a small probability of being chosen. In terms of choice probability,

$$P(1|h) = \frac{e^{\beta(hu(1,1)-(1-h)u(0,0))}}{1 + e^{\beta(hu(1,1)-(1-h)u(0,0))}} \quad (7)$$

This dynamic is an ergodic, homogeneous markov chain with a potential function. This guarantees that the state where all agents choose the payoff dominant action will be reached since this state uniquely maximizes the potential function. However, the time before reaching the invariant distribution

may depend on the size of the graph N . This dynamic has been studied on different types of graphs (Blume, Ellison). Young has established a criteria, close knittedness that guarantees that spread is rapid in the sense that the time it takes is bounded independently of the size of the graph, which is important for large graphs. In the stochastic diffusion, closeknittedness of a graph is comparable to to the contagion threshold in the deterministic version, in the sense that in an s close knit graph, every diffusion in which $\frac{u(0,0)}{u(0,0)+u(1,1)} \leq s$ will with high probability reach a state where most agents play 1 rapidly (ie in a time independent of N).

Definition 1. *A graph is rs -close knit if every node belongs to a set S of size at most r such that for each $S' \subset S$, $\frac{\text{Card}(\{(i,j) \in \Gamma | i \in S, j \in S'\})}{\text{Card}(\{(i,k) \in \Gamma | i \in S'\})} \geq s$.*

It is proved in [46] that

Theorem 1. *(adapted from Young) If the payoff parameters in the diffusion are such that $\frac{u(0,0)}{u(0,0)+u(1,1)} < s$, then for every $\epsilon > 0$, there is a β_0 such that for every $\beta > \beta_0$, there is a waiting time $T(r, k, \beta)$ that does not depend on N such that we have for every rs close knitted graph*

$$\mathbf{P}\left(\sum_{i=1}^{i=N} a_i^t \geq N(1 - \epsilon)\right) > 1 - \epsilon. \quad (8)$$

whenever $t > T$.

5. asymptotic bounds of contagion thresholds in different small world models

Here we will consider deterministic diffusion as described in section 1.2 If we specify a set of nodes that are initially in state 1, I^0 , then we define, for the threshold q , the nodes that enter state 1 at time t recursively as:

$$(I^q)^t = \left\{i \mid \frac{\text{Card}(V(i) \cap (\bigcup_{s < t} (I^q)^s))}{\text{Card}(V(i))} \geq q\right\} - \bigcup_{s < t} (I^q)^s \quad (9)$$

We note that $(I^q)^t = (I^q)^t(G_N)$ through the dependance on G_N through $V(i)$ but to simplify notation, we will often simply write $(I^q)^t$, when the context is clear.

The definition of contagion threshold (6) applies to deterministic graphs when the node set is infinite. We want to give an analogous definition that applies to sequences of random graphs indexed by increasing size. This definition should capture the idea that a small set of initially infected nodes

can spread to the whole graph asymptotically when N is large. We propose two possible increasingly strong definitions

Definition 2. *Let I_0 be any set fixed prior to the realization of Γ^N such that $\text{Card}(I_0) = o(1)$. We say that the asymptotic contagion threshold, with respect to a typical I_0 is inferior to η if for all $q > \eta$ and for every fixed $\epsilon > 0$*

$$\limsup_{N \rightarrow \infty} \mathbf{P}\left(\frac{\text{Card}(\bigcup_{t \geq 0} (I^q)^t(\Gamma^N))}{N} \geq \epsilon\right) = 0 \quad (10)$$

We can also require spread to be bounded by a constant, a stronger requirement.

Definition 3. *Let I_0 be any set fixed prior to the realization of Γ^N such that $\text{Card}(I_0) = o(1)$. We say that the asymptotic contagion threshold, with respect to a typical I_0 is inferior to η in the strong sense, if there is a constant k , which may depend on $\text{card}(I_0)$ but not on N , such that for all $q > \eta$*

$$\limsup_{N \rightarrow \infty} \mathbf{P}\left(\text{Card}\left(\bigcup_{t \geq 0} (I^q)^t(\Gamma^N)\right) > k\right) = 0 \quad (11)$$

Moreover, we can ask that spread does not occur, no matter how we choose the initial set of adopters. The definition below can also be given with respect to spread that is limited in the strong sense ((11))

Definition 4. *We say that the asymptotic contagion threshold, with respect to any I_0 is inferior to η if for all $q > \eta$, for every $s = o(1)$, and for every fixed $\epsilon > 0$,*

$$\limsup_{N \rightarrow \infty} \mathbf{P}\left(\bigcup_{\{I^0 | \text{Card}(I^0) \leq s\}} \frac{\text{Card}(\bigcup_{t \geq 0} I^t(\Gamma^N))}{N} \geq \epsilon\right) = 0 \quad (12)$$

The stronger definition, (12) guarantees that diffusion dies out over the contagion threshold for *any* choice of the set of initial users in the graph, even one chosen intentionally to maximize total spread. We provide upper bounds of the contagion threshold for the three types of small world random graphs presented in the previous section: the perturbed lattice (Watts), the independent community model and an island model. The main results of this section are essentially that the asymptotic contagion thresholds are very low in all the small world models we consider, at least when the connectivity is not too low, when we apply the weak definition (10). The same is not true for the strong definition. In Watts model we have:

Theorem 2. *In Watts' Small World, the contagion threshold with respect to a typical I_0 , with respect to the weak limitation of spread (10) is no greater than $1/2k$.*

However, if we require that spread is limited in a stronger sense, we see that the asymptotic contagion threshold is much higher:

Theorem 3. *In Watts' Small World, the contagion threshold with respect to a typical I_0 , with respect to the strong limitation of spread (11) equals $1/2$.*

In the independent community model and island model, we will show that contagion fails except at low thresholds in the strong sense (11) and for any choice of the set of initial adopters:

Theorem 4. *For any $q > 1/2\delta$ contagion fails in the independent community model with cluster size δ and island model with island size 2δ , in the sense of definition (12). Moreover, for any $k = O(1)$, there is a bounded function $f(k)$ such that*

$$\mathbf{P}(\limsup_{N \rightarrow \infty} [\bigcup_{\{I^0 | \text{Card}(I^0) \leq k\}} \text{Card}(\bigcup_{t \geq 0} (I^q)^t(\Gamma^N)) > f(k)]) = 0 \quad (13)$$

Proof theorem 2 and 3

In the proof of theorem 3 and 4, we will use the notation $O(1)$ for a quantity that is bounded by a constant independently of N . This facilitates notation, since the exact values of these quantities do not intervene in the proof. Consider an arbitrary set I_0 such that $\text{Card}(I_0) = o(1)$. We have $I_0 = \bigcup_{l < L} I_l$ where each I_l is an interval in the lattice. Let $f(N)$ be any function verifying $\lim_{N \rightarrow \infty} f(N) = \infty$. In what follows, we may suppose that $\lim_{N \rightarrow \infty} f(N)/\ln N = 0$ because whenever contagion fails for such an $f(N)$, necessarily it fails for any $f(N)$ orders of magnitude greater.

Let I^0 be an arbitrary fixed interval in the lattice. Fix $q > 1/2k$. If $2k$ is the degree of a node in the lattice, define $K =: \min[d|\frac{k}{2k+d} < q]$. In the lattice small world, spread occurs locally from the ends of the interval. Call l_{min} and l_{max} the endpoints of the interval I_l . We want to give an upper bound of the probability that all the nodes in $[l_{min} - f(N), l_{max} + f(N)]$ become infected through local spread. We then address spread through shortcuts. A sufficient condition for the local spread to stop is that the endpoint of the interval meets an interval of k consecutive nodes that all have a high fraction of non local neighbors, ie whose degree in the Erdos-Renyi graph is at least d . Whenever at most $O(f(N))$ nodes are in state 1, for any node

$i \in [l_{min} - f(N), l_{max} + f(N)]$, the probability of having at least K links in the erdos renyi graph to nodes that are not in state 1 is given by

$$\mathbf{P}(deg_{ER}(l) \geq K) \geq \binom{N - o(f(N))}{K} \left(\frac{p}{N}\right)^K \left(1 - \frac{p}{N}\right)^{N - o(f(N)) - K} =: q > 0 \quad (14)$$

Consider the nodes in $[l_{max}, l_{max} + f(N)] - I^0$. $Card([l_{max}, l_{max} + f(N)] - I^0) = O(f(N))$ and we can divide these nodes into $O(f(N))$ disjoint intervals each containing k nodes. For each such interval, the probability that every node in it has degree at least K is at least q^k and the probability that none of the intervals in $[l_{max}, l_{max} + f(N)] - I^0$ has this property is inferior to $(1 - q^k)^{O(f(N))}$. We can apply the same argument to $[l_{min}, l_{min} - f(N)]$. Thus the probability that the local spread continues at one or both ends of the interval is inferior to $(1 - q^k)^{O(f(N))}$. This argument applies to each of the intervals I_l . There are a finite number L of intervals and thus the probability that one of them is able to spread is also at most $(1 - q^k)^{O(f(N))}$

We also need to consider non local spread. In what follows, we use the notation $S(f(N)) = \bigcup_l [l_{min} - f(N), l_{max} + f(N)]$ For $i = 1, ..N$, define $C_i = Card((i, j) \in E | j \in S(f(N)))$. We note that if local spread is contained to $S(f(N))$, a sufficient (although not necessary) condition for spread to be contained to $S(f(N))$ is that (a):For every $i \notin S(f(N))$ $C_i^t \leq 2$ and (b) for every $i \in S(f(N))$, $C_i^t = 0$.

For the event (a) we have

$$\begin{aligned} \mathbf{P}\left(\bigcap_{i \notin S(f(N))} C_i^t < 2\right) &= \prod_{i \notin S(f(N))} \mathbf{P}(C_i \leq 2) \\ &= \prod_{i \notin S(f(N))} \left(1 - \sum_{k \geq 2} \binom{O(f(N))}{k} \left(\frac{p}{N}\right)^k \left(1 - \frac{p}{N}\right)^{o(f(N))}\right) \\ &\geq \prod_{i \notin S(f(N))} \left(1 - o(f(N)) * \max_{k \geq 2} \left(\frac{o(f(N))}{N}\right)^k\right) \\ &\geq \left(1 - o\left(\frac{f(N)^3}{N^2}\right)\right)^N = 1 - O\left(\frac{f(N)^3}{N}\right) \end{aligned} \quad (15)$$

As for (b)

$$\begin{aligned} \mathbf{P}\left(\bigcap_{i \neq j \in (S(f(N)))^2} (i, j) \notin E\right) &\geq \\ \prod_{i \neq j \in (S(f(N)))^2} \mathbf{P}((i, j) \notin E) &\geq \left(1 - \frac{1}{N}\right)^{O(f(N)^2)} = O\left(\frac{f(N)^2}{N}\right) \end{aligned} \quad (16)$$

Together with the previous, we conclude that

$$\limsup_{N \rightarrow \infty} \mathbf{P}\left(\frac{\text{card}(\bigcup_t I^t)}{N} \leq \frac{\text{card}(S(f(N)))}{N}\right) = 1 \quad (17)$$

The result in theorem 3 is obvious. For any constant k , the probability that each node in an interval of k nodes has degree 0 in the Erdos-Renyi graph is $(1 - \frac{p}{N})^{Nk}$ and $\lim_{N \rightarrow \infty} (1 - \frac{p}{N})^{Nk} > 0$.

Proof independent community model

In what follows, we discuss the steps involved in proving theorem ?? and present a number of lemmas which intervene. We have chosen to defer the proofs of these lemmas to the end of the section. At time 0, the contagion process starts from a set of initially infected nodes I^0 . We will consider only initially infected sets that consist of a number of completely infected clusters since any set that is not of this type is included in such a set. Thus, we may assume that initially infected sets are of the form $S(k)$ for some k , where

$$S(k) = \{(\mathcal{G}_1, \dots, \mathcal{G}_M)^k \mid \mathcal{G}_{m_1} \neq \mathcal{G}_{m_2} \dots \neq \mathcal{G}_{m_k}\}. \quad (18)$$

We want to show that for every k , there is a constant $K(k)$ that does not depend on N such that

$$\mathbf{P}(\limsup_{N \rightarrow \infty} [\bigcup_{\{I^0 \mid I^0 \in S(k)\}} \text{Card}(\bigcup_{t \geq 0} I^t(\Gamma^N)) > K(k)]) = 0. \quad (19)$$

We have

$$\begin{aligned} \mathbf{P}\left(\bigcup_{I^0 \in S(k)} \text{Card}(\bigcup_{t \geq 0} I^t(\Gamma^N)) > K(k)\right) &\leq \sum_{I^0 \in S(k)} \mathbf{P}(\text{Card}(\bigcup_{t \geq 0} I^t(\Gamma^N)) > K(k)) = \\ &\text{Card}(S(k)) \mathbf{P}(\text{Card}(\bigcup_{t \geq 0} I^t(\Gamma^N)) > K(k)) \quad . \end{aligned}$$

The last equality is due to the fact that the law of the graph Γ^N is invariant to permutation of the clusters. The probability of spread ex-ante is the same from any set in $S(k)$. Thus, we fix an arbitrary $I^0 \in S(k)$ independently of Γ^N . Since $\text{Card}(S(k)) < \binom{M}{k}$, it is sufficient to show that we can find a

$K(k)$ such that

$$\mathbf{P}(\text{Card}(\bigcup_{t \geq 0} I^t) > K(k)) \leq \frac{1}{N^{1+\alpha}} \binom{M}{k} \quad (20)$$

We index the contagion process by $t = 0, \dots$. To go from step $t - 1$ to step t in the contagion process, i.e. to determine the nodes that become infected at t , we need to have knowledge of all the links that involve nodes in I^{t-1} . Thus, between times $t - 1$ and t in the contagion process, we draw variables that determine links involving the nodes in I^{t-1} . We denote the variables drawn in the graph process by $(Y_s)_{s \geq 0}$. We will describe later what these variables are. We denote by \mathcal{F}_s the sigma algebra generated by the variables $(Y_l)_{l \leq s}$. Variables in the graph process are always drawn between two steps in the contagion process. We denote by $Y_{s_{t-1}}$ the first variable drawn between $t - 1$ and t and Y_{s_t} the last one.

The independent community model is uniquely defined by the realization of the variables $(X_i)_{i=1}^{i=N}$ which follow a uniform and independent law on $\mathcal{G}_1, \dots, \mathcal{G}_M$. In order to study the contagion process, we will not draw the variables X_i directly. Instead, we use the following alternative way of drawing the variables $(X_i)_{i=1}^{i=N}$

Remark 1. *We recall that in the independent community model for each node $i \in 1, 2, \dots, N$ we draw a variable X_i following a uniform law on $\mathcal{G}_1, \dots, \mathcal{G}_M$. It is possible to draw a graph Γ^N with the same law in the following way: For every $i = 1, \dots, N$, let $R_i^0 = \emptyset$ be the set of "rejected" nodes at time 0 in the graph process. At $s > 0$, consider any set J_s of clusters such that $J_s \cap R_i^{s-1} = \emptyset$ and denote $\text{Card}(J_s) = n_s$. We will now determine whether $X_i \in J_s$, according to the probability:*

$$\mathbf{P}(X_i \notin J_s) = \frac{M - n_s - \text{Card}(R_i^{s-1})}{M - \text{Card}(R_i^{s-1})}. \quad (21)$$

If $X_i \notin J_s$, we put $R_i^s = R_i^{s-1} \cup J_s$ and say that J_s is rejected by i . Otherwise, we draw X_i uniformly among the clusters in J_s . The law of X_i determined this way is clearly uniform on $\mathcal{G}_1, \dots, \mathcal{G}_M$.

We will use this alternative way of drawing the successors $(X_i)_{i=1}^{i=N}$ in order to monitor the contagion dynamic. Between two steps $t - 1$ and t in the contagion dynamics, we determine for every node $i = 1, \dots, N$ whether it

”rejects” the set of clusters containing previously infected nodes. If it does, in the following steps, we know that its law then follows a uniform law on the complement of the sets that have been previously rejected. If it does not, we determine its successor.

In the independent community model, each node belongs to a unique initial cluster \mathcal{G}_m and nodes are linked only if they are in the same A_m for some m . We will see that this makes it convenient to monitor infected nodes by monitoring initial clusters. An initial cluster \mathcal{G}_m will be assigned to a set depending on how many infected nodes the associated set A_m contains. Let us begin by defining these sets and how the clusters move between them.

At time $t = 0$, $s = 0$,

$$I_0^0 = \{i \in \mathcal{G} \in S(k)\}, \mathcal{D}_0^0 = D_0^0 = S(k) \quad (22)$$

$$C_0^0 = C_0^0 = \emptyset, \mathcal{B}_0^0 = B_0^0 = \mathcal{G}_1, \dots, \mathcal{G}_M - S(k). \quad (23)$$

The upper index is for the contagion process and the lower one for the graph process. All these sets except I_s^t are sets of clusters. However, we may by abuse of notation, write for example $i \in D^t$ meaning $i \in \mathcal{G} \in D^t$. It may be useful to bear in mind in what follows that we want to define these sets in such a way that at time t we will have $\mathcal{G}_m \in B^t$ if A_m contains no infected node, $\mathcal{G}_m \in C^t$ if A_m contains one infected node and $\mathcal{G}_m \in D^t$ if A_m contains at least two infected nodes. These properties have, of course, to be verified which is done in the lemma 5.2 after we define the process that describes how clusters move between these sets.

Graph process

For $i = 1, \dots, N$, we denote by R_i^s i 's set of rejected clusters at time s in the graph process, with $R_i^0 = \emptyset$. To go from step $t - 1$ to t in the contagion process, we realize the following variables $(Y_s)_{s=s_{t-1}}^{s=s_t}$.

- We determine (if they are previously unknown) the successors of the most recently infected nodes; for $j \in I_{s-1}^{t-1}$, we draw X_j uniformly in the set $\mathcal{G}_1, \dots, \mathcal{G}_M - R_j^{s-1}$.
- For every node i whose successor is not determined at s , we draw whether $X_i \in \mathcal{D}_{s-1}^{t-1}$ or not:

$$\mathbf{P}(X_i \notin \mathcal{D}_{s-1}^{t-1}) = \frac{M - \text{Card}(\mathcal{D}_{s-1}^{t-1}) - \text{Card}(R_i^{s-1})}{M - \text{Card}(R_i^{s-1})}. \quad (24)$$

If $X_i \notin \mathcal{D}_{s-1}^{t-1}$, we put $R_i^s = R_i^{s-1} \cup \mathcal{D}_{s-1}^{t-1}$. Otherwise we draw X_i uniformly in \mathcal{D}_{s-1}^{t-1} .

- When all the aforementioned variables have been drawn, we create a temporary set

$$T_s^t = [\mathcal{G} \in B_s^{t-1} | \exists! j \in \mathcal{G}, X_j \in \mathcal{D}_s^{t-1}] \sqcup [\mathcal{G} \in B_s^{t-1} | \exists! i \in I^{t-1}, X_i = \mathcal{G}]. \quad (25)$$

This set contains clusters that will end up either in C^t or D^t but we need more information to determine which one of them.

- Thus, for every node $i \in C_s^{t-1} \cup T_s^t$, whose successor is not determined yet, we draw whether $X_i \in (C_s^{t-1} \cup T_s^t)$

$$\mathbf{P}(X_i \notin (C_s^{t-1} \cup T_s^t)) = \frac{M - \text{Card}(C_s^{t-1} \cup T_s^t) - \text{Card}(R_i^s)}{M - \text{Card}(R_i^s)}. \quad (26)$$

If $X_i \notin (C_s^{t-1} \cup T_s^t)$, we put $R_i^{s+1} = R_i^s \cup (C_s^{t-1} \cup T_s^t)$. Otherwise, we draw X_i uniformly in $C_s^{t-1} \cup T_s^t$.

Updating of the sets

The sets we have defined will essentially be updated with the contagion process and not the graph process. The only exception is that at after any draw of a variable in the graph process, if there are very high degree nodes and clusters, we "remove" these from the process. Indeed, we have the following lemma which shows that such nodes cannot be infected before T .

Lemma 5.1. *Suppose that $q > 1/2\delta$, and that $\text{Card}(I^0) \leq k\delta$. Then there is a constant $c(k)$ that depends only on k such that for every $t \leq T$ and for any realization of the random variables drawn prior to t , i.e. for any \mathcal{F}_{s_t} . We have*

$$\text{Card}(I^t) \leq c(k), \text{ and } \max\{\text{deg}_i | i \in I^t\} \leq c(k). \quad (27)$$

The proof can be found at the end of the section. We note that the statement in lemma 5.1 would not be verified for all types of graphs and diffusions. For example, it would not hold for dynamics on scale free graphs where it is sufficient to have one infected neighbor to become infected.

The degrees of the nodes will not be known until all the variables $(X_i)_{i=1}^{i=N}$ have been drawn, but over time some of these successors are determined and

if there is an $\bar{s} \geq 0$ such that

$$\mathbf{P}(\text{Card}(A_m) \geq c(k) | \mathcal{F}_{\bar{s}}) = 1 \quad (28)$$

$$\text{then for all } s \geq \bar{s}, \mathcal{G}_m \in B_s \quad (29)$$

Moreover, if

$$\mathbf{P}(\text{deg}_i \geq c(k) | \mathcal{F}_{\bar{s}}) = 1 \quad (30)$$

$$\text{then for all } s \geq \bar{s}, i \notin I_s \quad (31)$$

This updating according to the graph process "re-assigns" nodes and clusters when their degree exceeds $c(k)$ since they cannot according to the lemma 5.1 be infected before T . With this exception, the sets of clusters that we have defined evolve with the contagion process and not with the graph process.

Updating with the contagion process At t , the sets B^t, C^t, D^t and I^t are thus defined recursively. As we have seen previously, these sets depend on the index in the graph process only through the fact that high degree nodes may have been re-assigned. In order to simplify notations, in what follows, we will omit it.

- For $\mathcal{G}_m \in B^{t-1}$: If $\mathcal{D}^{t-1} \in R_i^{s_t}$ for every $i \in \mathcal{G}_m$, and $X_j \neq \mathcal{G}_m$ for every $j \in I^{t-1}$, then $\mathcal{G}_m \in \mathcal{B}^t$. If $\mathcal{G}_m \in T^t$, and if $C^{t-1} \cup T^t \in R_i^{s_t}$ for every $i \in \mathcal{G}_m$ and if there is no $j \in C^{t-1} \cup T^t$ such that $X_j = \mathcal{G}_m$, then $\mathcal{G}_m \in C^t$. In all other cases, $\mathcal{G}_m \in \mathcal{D}^t$.
- For $\mathcal{G}_m \in C^{t-1}$: if $C^{t-1} \cup T^t \cup \mathcal{D}^{t-1} \in R_i^{s_t}$ for every $i \in \mathcal{G}_m$ and if there is no $j \in C^{t-1} \cup T^t \cup \mathcal{D}^{t-1}$ such that $X_j = \mathcal{G}_m$, then $\mathcal{G}_m \in C^t$. In all other cases, $\mathcal{G}_m \in \mathcal{D}^t$.
- For $\mathcal{G}_m \in D^{t-1}$, we have $\mathcal{G}_m \in D^t$

It is also useful to define the clusters which enter a give set for the first time at t :

$$\mathcal{B}^t = B^t - B^{t-1}, \mathcal{C}^t = C^t - C^{t-1}, \mathcal{D}^t = D^t - D^{t-1} \quad (32)$$

As for the recursive definition of the nodes infected at time t , we will modify the actual dynamics slightly in such a way that the number of infected nodes may be strictly greater but never fewer than in the actual diffusion: We define

$$I^t = \{i \in D^{t-1}\} \cup \{i | \text{Card}(V(i) \cap I^{t-1}) \geq 2\} \cup \{i | X_i \in D^{t-1}\} \quad (33)$$

Thus any node with links to at least two neighbors infected at $t-1$ will be infected at t . We note that in the actual dynamics, an agent is infected

at t if a proportion greater than $1/2\delta$ of his neighbors were infected at $t - 1$. Each cluster contains δ nodes and each node belongs to two clusters. Thus, in the original dynamic, having two infected neighbors is necessary but not sufficient to become infected. Moreover, we also immediately infect some nodes that will actually be infected at a later time with probability one. These modification have for purpose to simplify the monitoring of the states. The following lemma gives some properties of the sets B^t , C^t and D^t .

Lemma 5.2. • *Property 1: If $\mathcal{G}_m \in B^t$, then $A_m \cap I^t = \emptyset$.*

- *Property 2: For $t \geq 1$, If $\mathcal{G}_m \in C^t$, then $\text{Card}(A_m \cap I^t) \leq 1$. More precisely, we may have either $\text{Card}([\mathcal{G}_m] \cap I^t) = 1$ or $\text{Card}((A_m - [\mathcal{G}_m]) \cap I^t) = 1$ but not both. In the first case, \mathcal{G}_m contains exactly one infected node at t , in the second case it contains none.*
- *Property 3: $I^t \subset [i | i \in \mathcal{G}_m \in C^t \cap D^t]$*

Proof deferred to the end of the chapter.

By the above lemma, all infected nodes at t belong to initial clusters in C^t and D^t . Thus, if C^t and D^t are both empty for some t , we also have $I^t - I^{t-1} \neq \emptyset$, so that no new nodes are infected at (or after) t . The following lemma shows that at a given $t < T$ the number of clusters in C^t and D^t are bounded independently of N for any realization of the graph process will be useful. The lemma is similar to lemma 5.3 but applies to the modified dynamics.

Lemma 5.3. *Suppose that that $\text{Card}(I^0) \leq k$ and consider the modified contagion dynamics defined by (33). We have a constant c that does not depend on N such that for any realization of the variables drawn before $s < s_t$ and $t \leq T$, we have*

$$\text{card}(I^t) \leq \text{card}(i | i \in C^t \cup D^t) < c \quad (34)$$

We also have

$$\text{Card}(R_i^s) \leq \text{Card}(C^t \cup D^t) \leq c. \quad (35)$$

Now we would like to give a lower bound on the probability that there is a $t \leq T$ such that $C^t = D^t = \emptyset$. The dynamics are such that new nodes are infected if they have links to previously infected nodes, ie if their successors

are in C^{t-1} or D^{t-1} . The lemma 5.3 gave a bound of the number of clusters that are in these sets. Since the successors of the nodes follow a uniform law, the probability of having successors in a given set depends only on its size. A bound on this probability is given in the following lemma which will be used repeatedly.

Lemma 5.4. *Let H and J be two sets of clusters verifying $\text{Card}(H) < c$ and $\text{Card}(J) < c$ where c is the constant in the lemma 5.3. Let $s \leq s_t$. If there was no node in H with a successor in J at time $s - 1$, we have $\mathbf{P}(J \subset \bigcap_{l \in H} R_l^s) \geq \prod_{l \in H} \frac{M - \text{Card}(J) - \text{Card}(R_l^{s-1})}{M - \text{Card}(R_l^{s-1})} \geq 1 - O(\frac{1}{M})$*

The proof is quite immediate. Now, to be able to conclude, finally, we use the two lemmas stated below whose proofs can be found at the end of the section.

Lemma 5.5. *For every $t < T$, and for every history \mathcal{F}_{s_t} , $\mathbf{P}(I^t - I^{t-1} \neq \emptyset | \mathcal{D}^{t-1} = \emptyset, \mathcal{F}_{s_t}) \leq O(1/M)$*

Lemma 5.6. *For every $t < T$, and for every history \mathcal{F}_{s_t} , $\mathbf{P}(\mathcal{D}^t \neq \emptyset | \mathcal{F}_{s_t}) < O(1/M)$.*

We will now use the lemmas 5.5 and 5.6 to show that

$$\mathbf{P}\left(\bigcap_{j \leq T} I^j - I^{j-1} \neq \emptyset\right) \leq O(1/M^T) \quad (36)$$

First, we show that with probability greater than $1 - O(1/M^{T/2})$, there is a sequence of times $(t_k)_{k=1}^{k=T/2}$, such that $t_1 < t_2 < \dots < t_{T/2} \leq T$ and such that $\mathcal{D}^{t_k} = \emptyset$. Indeed, if no such sequence existed, there would be an increasing sequence of times $(t_l)_{l=1}^{l=T/2}$ such that $\mathcal{D}^{t_l} \neq \emptyset$. For any such sequence, we have by Lemma 5.6 that

$$\mathbf{P}\left(\bigcap_{l=1}^{l=T/2} \mathcal{D}^{t_l} \neq \emptyset\right) = \prod_{l=1}^{l=T/2} \mathbf{P}(\mathcal{D}^{t_l} \neq \emptyset | \bigcap_{k < l} \mathcal{D}^{t_k} \neq \emptyset) \leq (O(1/M))^{T/2} \quad (37)$$

The number of possible sequences is also bounded independently of M . Thus

$$\mathbf{P}(\exists (t_l)_{l=1}^{T/2} \text{ such that } \bigcap_{l=1}^{l=T/2} \mathcal{D}^{t_l} \neq \emptyset) \leq (O(1/M))^{T/2}. \quad (38)$$

By lemma 5.5, we have

$$\begin{aligned}
& \mathbf{P}\left(\bigcap_{k=1}^{k=T/2} I^{t_k+1} - I^{t_k} \neq \emptyset \mid \bigcap_{k=1}^{k=T/2} \mathcal{D}^{t_k} = \emptyset\right) = \\
& \prod_{k=1}^{k=T/2} \mathbf{P}(I^{t_k+1} - I^{t_k} \neq \emptyset \mid \mathcal{D}^{t_k} = \emptyset, \bigcap_{l \leq k} I^{t_l} - I^{t_l-1} \neq \emptyset) \\
& = \prod_{k=1}^{k=T/2} \mathbf{P}(I^{t_k+1} - I^{t_k} \neq \emptyset \mid \mathcal{D}^{t_k} = \emptyset, \mathcal{F}_{t_k}) \leq O(1/M^{T/2})
\end{aligned}$$

Together with the previous,

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{j \leq T} (I^j - I^{j-1} \neq \emptyset)\right) & \leq \mathbf{P}\left(\bigcap_{k=1}^{k=T/2} I^{t_k+1} - I^{t_k} \neq \emptyset \mid \bigcap_{k=1}^{k=T/2} \mathcal{D}^{t_k} = \emptyset\right) + \\
& \mathbf{P}(\exists (s_l)_{l=1}^{l=T/2} \text{ such that } \bigcap_{l=1}^{l=T/2} \mathcal{D}^{s_l} \neq \emptyset) \leq O(1/M^{T/2}) \quad (39)
\end{aligned}$$

We choose $T = 2k + 4$ to obtain

$$\mathbf{P}(I^T - I^{T-1} \neq \emptyset) \leq O\left(\frac{1}{M^{k+2}}\right) \quad (40)$$

Finally, we have:

$$\begin{aligned}
& \mathbf{P}\left(\bigcup_{I^0 \in \mathcal{S}(k)} \text{Card}\left(\bigcup_{t \geq 0} I^t(\Gamma^N)\right) > c\right) \leq \binom{M}{k} \mathbf{P}\left(\text{Card}\left(\bigcup_{t \geq 0} I^t(\Gamma^N)\right) > c\right) \\
& \leq \binom{M}{k} \mathbf{P}\left(\text{Card}\left(\bigcup_{t \geq 0} I^t(\Gamma^N)\right) > c \mid I^T - I^{T-1} \neq \emptyset\right) \mathbf{P}(I^T - I^{T-1} \neq \emptyset) \leq \\
& \qquad \qquad \qquad \binom{M}{k} O\left(\frac{1}{M^{k+2}}\right) \leq O\left(\frac{1}{M^2}\right)
\end{aligned}$$

We note that by lemma 5.3, necessarily,

$$\mathbf{P}\left(\text{Card}\left(\bigcup_{t \geq 0} I^t(\Gamma^N)\right) > c \mid I^T - I^{T-1} = \emptyset\right) = 0. \quad (41)$$

It follows almost immediately from the proof we have given that the same property is verified for the island model with cluster size 2δ since the links in this model are a subset of the links in the independent community model with the same realization of $(X_i)_{i=1}^{i=N}$.

5.1. Simulation set up and quantities of interest

In the previous section, we established some analytical bounds of the asymptotic contagion thresholds of the different Small World models. In this section, we will consider simulations of the diffusion dynamics in order to understand some aspects of its behavior that are not answered by the asymptotic analysis. On one hand, there is the question of the exact asymptotic contagion threshold of which we provided only an upper bound. However, this has little consequence from a practical point of view. The upper bounds provided in section 2 are already so low that diffusion would fail for most reasonable payoffs and even if the actual threshold is smaller, this would have little effect on outcomes. A question of greater interest concerns the behavior in large finite sized systems compared to the asymptotic. Indeed, we know from the previous analysis that for some large size N the fraction of infected nodes will be small, but we do not know for what values of N we will start to see behavior that resembles the asymptotic. In practice, it is possible that the behavior in graphs even of large size is different from asymptotic behavior. If this is the case, it is of interest to understand at what size the asymptotic is valid, and also to explain behavior pre-asymptotically.

In order to compare the behavior of the diffusion dynamic in a finite system of size N with the results in the previous section, we need to define a notion of contagion threshold for a graph of size N . The definitions (10),(12) of the contagion threshold are asymptotic and do not apply to finite sized graphs. In a graph of fixed size, we will define the contagion threshold as $\max_{\{q \in \frac{1}{100}\{1, \dots, 50\}\}} [q | \text{Card}(\bigcup_t (I^q)^t) = N]$, thus as the largest q (with a precision of 0,01) that allows all nodes to reach state 1 in the long run.

The behavior of the simulations confirms that it is reasonable to define the contagion threshold in this way in a finite sized system. Indeed, in all the parameter ranges and for all the realizations in our simulations we see sharp transitions. Either only a small fraction of the nodes are infected or all of them are. If the transition had been very gradual, it would be difficult to define an equivalent of the asymptotic notion.

In our simulations, we compute the average of the observed contagion threshold over several realizations of the graph. We note that since the contagion dynamic is deterministic, we only need to average over graph realizations. For graphs of size $N < 50000$ we have averaged over five realizations. For graphs larger than $N = 50000$ we have used only two realizations. Averaging over relatively few realization is justified by the fact that we have observed very minor fluctuations of the observed contagion threshold between realizations for all parameter values and values of N we have considered.

Often no fluctuations were observed over the five realizations.

independent community model

The independent community model depends on only one parameter, the size of the clusters. We will consider the observed threshold in a fixed size graph as a function of cluster size and also how the threshold varies with graph size N for fixed cluster size. The figures we discuss below can be found on pages 21 – 23

- Figure 1 represents the average observed contagion threshold as a function of cluster size in the independent community model of size $N = 10000$. We have also represented the asymptotic upper bound. For this size, the contagion threshold is already below our asymptotic upper bound.
- Figure 4 shows the average observed contagion threshold of the independent community model with $\delta = 4$ and of watts models of different connectivity when graph size ranges between 1000 and 100000. The contagion threshold of the independent community model is significantly lower than that of the watts model with similar degree ($2k = 14$) and in fact significantly lower than the watts model in all the connectivity ranges that are shown here.

We see from the simulations, that in the independent community model, the observed contagion threshold is small already in graphs of moderate size. As can be expected from the asymptotic bound, the threshold decreases with cluster size. The simulations show that the upper bound in section 2 is indeed not exact. The actual contagion threshold is lower. This is expected. Indeed, our proof essentially bounded the probability that new nodes were linked to more than one of the previously infected nodes. Using the fact that nodes in the independent community model have a minimal degree of 2δ , this bounds the probability that diffusion succeeds for thresholds over $\frac{1}{2\delta}$. However, 2δ is the minimal degree. The average degree is higher and for this reason, the actual contagion threshold is likely to be below our bound. This is confirmed by simulations. Generally, the simulations of the diffusion dynamic on the independent community model do not reveal anything surprising. There is fair agreement between the simulations and the behavior that can be expected from the asymptotic bound. Even in graphs of moderately large size ($N = 1000$) the behavior is not radically different from the asymptotic.

Watts' Small World

The watts model involves two parameters, the degree of the local links in the lattice $2k$ and the average number of random (erdos renyi) links. We will explore the effect of both by fixing the local degree $2k$ and varying the

fraction of random links, and inversely varying the local degree when the fraction of random links is fixed. We will also analyze how the threshold changes with the size of the graph in these different cases. The figures can be found on pages 21 – 23.

- Graphic 2 represents the average observed contagion threshold as a function of the lattice degree. The size of the graph is $N = 10000$ and the fraction of random links has been fixed at 10 per cent. We have also represented the asymptotic upper bound. As we see, the local degree $2k$ has a strong influence on the contagion threshold which is significantly higher for larger values of $2k$.
- Graphic 3: Here we have fixed the local degree at $2k = 14$ and we let the fraction of random links vary between 1 and 25 per cent. The size of the graph is $N = 10000$.
- Graphic 4: The fraction of random links is 10 per cent. We consider the contagion threshold in the watts model with $2k = 4, 14, 20$ respectively, for values of N ranging from 1000 to 100000. For $2k = 4$, we are close to the asymptotic bound when $N = 100000$. For $2k = 14$ and 20 we remain far from the asymptotic bound even for $N = 100000$.
- Figure 5: This graph shows the fraction of infected nodes as a function of q . As we see the transition is sharp.
- The last three graphs (6, 7, 8) show the contagion threshold as a function of size in the watts models with $2k = 4, 14, 20$ respectively. In each graph, we also show the mean field model threshold and the asymptotic bound. We see that in the case $2k = 4$ we are close to the asymptotic bound for large N . In the other cases the threshold remains closer to the mean field threshold and is far from the asymptotic bound even for $N = 100000$.

The simulation of the diffusion dynamic on the Watts model shows behavior that is very different from the asymptotic in graphs of size up to $N = 100000$, except in the case where the local lattice degree is very low. The observed contagion threshold is much higher than the one we should see asymptotically. In the graphs where we have simulated the diffusion, we see higher contagion threshold for higher local connectivity (lattice degree), while the asymptotic bound predicts the opposite. We will attempt to explain these observations heuristically.

Mean field case We begin by analyzing the mean field case, ie the case where each node has a degree in the erdos renyi graph that is equal to the expected degree. Suppose that p is the average degree of the nodes in the erdos renyi link or equivalently the number of shortcuts/non local degree.

In this case a necessary and sufficient condition for the diffusion to spread is that $q = \frac{u(0,0)}{u(0,0)+u(1,1)} \leq \frac{k}{2k+p}$. Indeed, the nodes that are immediately to the left (or right) of the interval of nodes in state 1, will have exactly proportion $\frac{k}{2k+p}$ of their neighbors in state 1. If these nodes do not switch to 1, none of the consecutive nodes will either. Thus, if all nodes had a degree equal to expected degree, then the maximal q for which we can expect the diffusion to spread would be $\frac{k}{2k+p}$. The mean field case seems relevant for the behavior in the simulations. Indeed, spread does not seem to occur for $q > \frac{k}{2k+p}$ even in relatively small graphs. Moreover, when the local degree $2k$ is large, the observed threshold remains closer to the mean field case value than to the asymptotic bound even for large N .

The role of the local (lattice) degree

We see from figure 2 and 4 that the local degree in the graph ($2k$ the number of lattice neighbors) has a significant impact on the threshold for all values of N . In the simulations that are shown in figure 2 and 4, we have, instead of fixing p , the average degree in the erdos renyi graph, fixed the proportion of random links. Thus graphs with a higher number of local neighbors also has a higher number of random links. This is necessary in order to compare the graphs with respect to the impact of local degree. Suppose that a fraction r of random links is fixed so that $\frac{p}{2k+p} = r \Leftrightarrow p = \frac{r2k}{1-r}$. Thus the maximal value of q for which a node with typical degree p will be infected verifies $q = \frac{k}{2k + \frac{r2k}{1-r}} = \frac{1-r}{2}$. Thus a node of typical degree will be infected at the same threshold independently of k whenever the fraction of random links is fixed. When the proportion of random links is fixed, all graphs would have the same contagion threshold independently of k if there were no fluctuations in degree. However, the way the threshold reacts to deviations from the mean depends on k : put $c =: \frac{2}{1-r}$ and define $d' = d - p$, the deviation from the mean. Then a node with degree d that is next to the interval of nodes in state 1 will switch to 1 only if q verifies $f(k, d') = \frac{k}{kc+d'} \leq q$. Thus for any fixed value of the deviation from the mean d' , the function $f(k, d')$ is increasing in k , so that the maximal value of q for which a high degree nodes will adopt 1 increases with $2k$.

The local degree in the graph also affects the contagion threshold through a different mechanism. If the local degree is very low, a single node with atypically high non local degree is sufficient to stop the local spread. If the local degree is high, on the other hand, this is not sufficient. Indeed, suppose that node j is the first node to the left (or right) of an interval of nodes in state 1 and that $\frac{k}{2k+deg_j} \leq q$. If the nodes to the right of j , $j+1, \dots, j+m$

are such that $\frac{k-1}{2k+deg_{j+m}} \geq q$, the local spread can still continue. If k is large, local spread will stop only when there are several nodes with atypically high degree close to each other.

It is the small variance of the degree distribution in the Erdos renyi graph that explains why we do not observe the asymptotic contagion threshold even in fairly large systems ($N = 100000$), at least for higher local degrees. This degree distribution follows a binomial law and the probability that a node has a degree that is significantly higher than the expected one is small. If the size of the graph is sufficiently large, such nodes will be present, but even a graph of size $N = 100000$ may still be too small for this to occur.

If the local degree is very small, behavior resembles the asymptotic one for $N \approx 100000$. If the local degree is higher, on the other hand, behavior remains close to that in the mean field analysis of the system even for large N , and the contagion threshold is significantly higher than it is asymptotically. Essentially, the asymptotic is based on events (high degree nodes) that are rare but will always occur in a large enough system. In moderately large systems, on the other hand, behavior seems to be more accurately described by an analysis based on the typical behavior of the degrees.

Comparison of pre-asymptotic contagion thresholds in Watts' Small World and the independent community model

While the analysis in the previous section shows that asymptotically, the contagion thresholds in all small world models are very low, the behavior of the diffusion dynamic in graphs of fairly large fixed size ($N = 100000$) differs significantly between the models. In the independent community model, the contagion threshold is close to its asymptotic value even when the size of the graph is not very large ($N \approx 1000$). In the Watts model, the contagion threshold is much higher than its asymptotic value even when the size of the graph is fairly large ($N = 100000$). Thus, the independent community model has a much lower contagion threshold than the watts model with the same connectivity in all the size ranges considered. ($N \in [1000, 100000]$). This is true for all reasonable choices of the fraction of non local links in the watts model (between 1 and 25 per cent). This difference pervades even as the size of the graph grows and the difference in the observed contagion threshold remains significant at $N = 100000$. In graphs of this size, the differences between contagion thresholds are significant. In the independent community model, they are so low that the payoff dominant action could only spread if it had payoffs at least ten times greater than the other action, meaning that in all practical applications spread would not occur. In the Watts model on the other hand, an action presenting reasonable payoff benefits would be

able to spread in a moderately large graph. Thus despite theoretical large limit contagion thresholds that are comparable, these models exhibit very different behavior with respect to the contagion threshold even when the graphs are large.

5.2. Stochastic diffusion on small worlds

We have seen that a step in the proof of the main theorem in Young, shows that if payoffs verify $\frac{u(0,0)}{u(0,0)+u(1,1)} = q$, then for every $\epsilon > 0$, there is a $\beta(r, k)$ and there is a time $T(r, k, \beta)$ independent of N such that for every node i that belongs to an r, s close knitted set with $s \geq q$,

$$\mathbf{P}(a_i^t = 1) \geq 1 - \epsilon \quad (42)$$

whenever $t > T$. This criteria is useful in random graphs where we will bound the probability that a large fraction of the nodes belong to such sets. The analysis in [46] does not indicate how to find a close knitted set containing i . However, we should look for a set that is not too large, since T is increasing with the size r . In practice, it is natural to consider the local communities of the nodes if there are such, these would be, in the models we have considered, lattice intervals in the watts model, clusters and islands respectively in the cluster and islands models. In what follows, we will assume that β has been fixed so that the invariant distribution is concentrated on 1.

It follows from Young that

Proposition 2. *In the island model, for every $\epsilon > 0$, for every fixed $l \geq 1$, if $\frac{u(0,0)}{u(0,0)+u(1,1)} \leq \frac{1}{2} - \frac{l+2}{2(\delta+1+l)}$ then there is a time $T(k, r, \delta)$ independent of N , such that*

$$E\left[\sum_{i=1}^{i=N} a_i^t\right] \geq N(1 - \epsilon)(1 - 1/l) \quad (43)$$

whenever $t > T$

Proof We bound the average close knittedness of the clusters: We take $S = [\mathcal{G}]$. $\text{Card}([\mathcal{G}]) = \delta$. Consider $S' \subset S$. We have $\text{Card}((i, j) \in \Gamma | i \in S, j \in S') = \text{Card}(S')(\delta - \text{Card}(S')) + (\text{Card}(S') - 1)\text{Card}(S')/2$. Every node in $[\mathcal{G}_m]$ has a degree bounded by $\delta + \frac{1}{\delta}\text{Card}(i | X_i = \mathcal{G}_m) + 1$ since it is linked to $\delta - 1$ nodes in $[\mathcal{G}_m]$ and at most one other node, and the nodes with in links to \mathcal{G}_m are distributed among the nodes in \mathcal{G}_m . Thus

$$\begin{aligned} \min_{S' \subset S} \frac{\text{Card}(\{(i, j) \in \Gamma | i \in S, j \in S'\})}{\text{Card}(\{(i, k) \in \Gamma | i \in S'\})} &\geq \\ \frac{\text{Card}(S')(\delta - \text{Card}(S') + (\text{Card}(S') - 1)/2)}{\text{Card}(S')(\delta + 1 + \frac{\text{Card}(i|X_i=\mathcal{G})}{\delta})} &\end{aligned} \quad (44)$$

This quantity is minimized for $S' = S$, when it equals $\frac{(\delta-1)/2}{\delta+1+\frac{\text{Card}(i|X_i=\mathcal{G})}{\delta}}$. If we put $r = \frac{\text{Card}(i|X_i=\mathcal{G})}{\delta}$, we can write $\frac{(\delta-1)/2}{\delta+1+\frac{\text{Card}(i|X_i=\mathcal{G})}{\delta}} = \frac{1}{2} - \frac{1}{2}(\frac{2+r}{\delta+1+r})$. Since $\sum_m \text{Card}(i|X_i = \mathcal{G}_m) = N$, we can use this to conclude that at least $(1 - \frac{1}{l})M$ of the commentates are at least $\frac{1}{2} - \frac{1}{2}(\frac{2+l}{\delta+1+l})$ close-knitted. This bound is useful when δ , the number of members in the group is large enough, so that the bound comes close to $\frac{1}{2}$.

Proposition 3. *In the watts model, for every $\epsilon > 0$, whenever $\frac{u(0,0)}{u(0,0)+u(1,1)} \leq \frac{1}{2}$, we can choose $p > 0$ such that if the probability of each random link is p/N then there is a time $T(k, r, \epsilon)$ independent of N , such that*

$$E\left[\sum_{i=1}^{i=N} a_i^t\right] \geq N(1 - \epsilon)^2(1 - (p + \epsilon)(2k + 1)) \quad (45)$$

Proof

Let the probability of a link in the Erdos reyi graph be P/N . Then by standard results, for any $\epsilon > 0$, and $\epsilon > 0$, we can find N_0 such that for $N \geq N_0$

$$\mathbf{P}[\text{Card}(\{(i, j) | (i, j) \in E\}) \leq \frac{N-1}{2}(p + \epsilon)] \geq 1 - \epsilon \quad (46)$$

equivalently, (46) bounds the probability that more than $N(p + \epsilon)$ nodes

are endpoints of links in the erdos renyi graph. Consider the N intervals $([i - k, i + k])_{i=1}^{i=N}$. Whenever $\text{Card}(\{(i, j) | (i, j) \in E\}) \leq \frac{N-1}{2}(p + \epsilon)$, no more than $N(p + \delta)2k + 1$ of these intervals contain a node which is endpoint of a link in the erdos renyi graph. The remaining $N[1 - (p + \delta)(2k + 1)]$ nodes belong to $1/2$ close knit sets of size $2k + 1$. Applying the result (42) to all nodes in close knitted sets, we can conclude that for $N > N_0$, there is a T

such that

$$E\left[\sum_{i=1}^{i=N} a_i^t\right] \geq E\left[\sum_{i=1}^{i=N} a_i^t \text{Card}(\{(i, j) | (i, j) \in E\})\right] \leq \frac{N-1}{2}(p+\epsilon) * \\ \mathbf{P}[\text{Card}(\{(i, j) | (i, j) \in E\}) \leq \frac{N-1}{2}(p+\epsilon)] \geq N(1-\epsilon)^2(1-(p+\epsilon)(2k+1)) \quad (47)$$

In the independent community model, it is more complicated to bound the close knittedness of the sets A_m . We will limit ourselves to heuristics and simulations based results. Consider a set A_m and an arbitrary subset $S' \subset A_m$. We have

$$\frac{\text{Card}(\{(i, j) \in \Gamma | i \in S, j \in S'\})}{\text{Card}(\{(i, k) \in \Gamma | i \in S'\})} = \\ \frac{\text{Card}(S')(\text{card}(A_m) - \text{Card}(S')) + (\text{Card}(S') - 1)\text{Card}(S')/2}{\text{Card}(S')(\text{card}(A_m) - \frac{\text{Card}(S')+1}{2}) + \sum_{s \in S'} \text{Card}(A_n) \mathbf{1}_{s \in A_n}} \quad (48)$$

Since every node i belongs to two sets in $(A_m)_{m=1}^m=M$, we can consider for i the most close knit of its two communities: we have

$$\min_{\{S' \subset S\}} \max_{\{A_m | i \in A_m\}} \frac{\text{Card}(S')(\text{card}(A_m) - \frac{\text{Card}(S')+1}{2})}{\text{Card}(S')(\text{card}(A_m) - \frac{\text{Card}(S')+1}{2}) + \sum_{s \in S'} \text{Card}(A_n) \mathbf{1}_{s \in A_n}} \quad (49)$$

Since we know the distribution of the sizes of $(A_m)_{m \leq M}$, we could bound this expression more explicitly. We note that if there had been no variation in cluster size so that $\text{Card}(A_m) = 2\delta$ for every $m = 1, \dots, M$, then the closeknittedness s verifies $s = \frac{(2\delta-1)2\delta/2}{(2\delta-1)2\delta/2+(2\delta)^2} = \frac{2\delta-1}{6\delta-1}$, which is close to $\frac{1}{3}$ for large δ . This should be a good approximation since the binomial distribution of cluster sizes has low variance. Computations based on simulations estimate the average closeknittedness to be around 0,35 with little variation.

6. Conclusion

This analysis of the diffusion of innovations on different types of Small World networks draws a rather complex picture of the phenomenon which indicates that outcomes can be very sensitive to small modifications of the dynamics itself (best response or trembling hand best response) as well as to finer

differences between networks that share many of their statistical properties. Moreover, from a technical point of view, the asymptotic behavior with respect to the size of the graph is not always well approximated by the behavior in even very large finite sized networks.

Morris [31] introduced a notion of contagion threshold that applies to deterministic graphs with an infinite node set: contagion is considered to occur if only a finite number of agents who have not adopted remain. We aimed to define an analogous notion of contagion threshold but for random graphs, in the asymptotic setting with respect to the network size N . With an infinite node set, spread is either limited to a finite set or leaves only a finite set that has not adopted. In our setting, we need to distinguish orders of magnitude in relation to network size N . Therefore, we can consider that spread is limited in a weak sense if the fraction of adopters goes to zero as the size of the network grows. This does not bound the *number* of adopters, just their fraction of the total population. A stronger definition requires that the number of adopters of the innovation remains bounded by some constant independent of network size. When applying the weak definition, we found similar asymptotic contagion thresholds in all the models considered, in the sense that the innovation would not spread to a significant fraction of the population unless it was greatly superior in terms of pay-off. However, in the independent community model and islands model, spread is difficult even with respect to the stronger definition since we show that the number of adopters can be bounded by a constant independent of network size almost surely. Moreover, this result is strong in the sense that spread does not only fail to occur if the group of initial users is chosen at random, It fails even if the group of initial users is chosen strategically in order to maximize the spread. On the contrary, in Watts' Small World, except when the innovation has very small benefits, we cannot bound the number of adopters by any fixed constant, although we can bound it by a function $f(N)$ which will be orders of magnitude smaller than the network size N asymptotically.

We also see considerable differences between the various types of Small Worlds when we consider spread in large networks of fixed size. As we have seen from the pre-asymptotic simulations, in the independent community model, the behavior resembles the asymptotic even when the size of the graph is relatively small, meaning that the payoff dominant action does not spread easily unless its payoff advantage is very great. In Watts' Small World, we have seen that the innovation can easily spread to all the agents even when its advantages are modest for a network size as large as 100 000. This

suggests that the stronger definition of the asymptotic contagion threshold which would classify the independent community models and islands models as unfavorable to spread and Watts' Small World as favoring also leads to more coherence with observations of large finite sized networks.

If we consider a version of the coordination game where each agents uses trembling hand best response, as in Young [46], we see somewhat different outcomes. In this version, spread to the network as a whole always occurs so the crucial question is for which payoff benefits of the innovation the time it requires is bounded independently of network size. In some cases, adding a little randomness to the best response, can make a big difference. Applying Young's close knitedness criterion to the Island model, we find that spread can occur easily, that is rapidly, even for small payoff benefits of the innovation. This contrast with the behavior in this model when best response is deterministic. In Watts' Small World the same is true, which could be expected. Computing close knitedness in the independent community model is more complicated due to fluctuations in degree, however, based on typical degrees, we should expect spread to require higher payoff benefits of the innovation than in the other models although not as high as in the case of deterministic best response. It is in fact rather natural that spread occurs more easily on all networks in the case where each agent has some small probability of adopting the innovation even without having any neighbor who as already adopted it.

How do we explain the contrasting results on the different types of Small Worlds? Basically, what goes on is determined by the nature of the boundaries between different communities. We base this conclusion on analysis of the two extreme correlation cases: complete correlation and zero correlation between communities. In the first case, communities tend to overlap, having several common members, something that could naturally occur when individuals bring together friends and acquaintances from initially different contexts. Here these contexts or communities do not remain separate and their boundaries are permeable. In the second case, a single individual ensures the link between otherwise disjoint communities. The link is not necessarily a weak link in the sense of Watts, because the agent may be a full member of both communities, linked to all other agents in these, nevertheless, this is not sufficient for the innovation to easily cross the community barrier.

If we acknowledge that social networks tend to be based on communities, then we should attempt to understand, in real social networks, the nature of

the boundaries between the communities or the social contexts in which the individual participates, whether these are sharp or overlapping and permeable. In particular, we see that using Watts well-known Small World model to predict behavior of the diffusion of innovations or other threshold based dynamics in social sciences is suitable if there is strong correlation between the different communities of an individual but not if these are independent.

7. Appendix:proofs

Study of the case $\alpha \rightarrow \infty$, draw with replacement of secondary community

To be completed

Proof lemma 5.1

We prove the lemma by induction: $Card(I^0) = k\delta$. At $t = 1$, any node in I^t is either a successor of an $i \in I^0$ or has a successor among the clusters infected at $t = 0$. This first case concerns at most $k\delta$ nodes. Any node who has one of k initially infected clusters as successor may be infected at $t = 1$, if the total number of nodes linked to the same cluster does not exceed $2\delta^2$. Since nodes with the same successor cluster are linked to each other, their proportion of infected neighbors would be lower than $1/2\delta$ if more than $2\delta^2$ uninfected nodes had the same cluster for successor. Thus at most $2k\delta^2$ nodes may be infected this way. In total, $Card(I^1) \leq card(I^0) + k\delta + 2k\delta^2$. Moreover, at time $t = 1$, no node with degree higher than $2k\delta^2$ can be infected since it would not have a proportion greater than $\frac{1}{2\delta}$ of its neighbors in I^0 . We define $f(1, k) = card(I^0) + k\delta + 2k\delta^2$. This concludes induction step 1. For $t > 1$, by the induction hypothesis, the number of infected nodes at $t - 1$ and the maximal degree of any infected node at $t - 1$ is bounded by $f(t - 1, k)$. Thus as before, no node can be in I^t if it has degree higher than $2\delta f(t - 1, k)$. Also, at t , at most $f(t - 1, k) * f(t - 1, k)$ nodes are linked to nodes in I^{t-1} by the induction hypothesis about the degree of infected nodes. Thus $Card(I^t) \leq f(t - 1, k)^2$. We can define $f(t, k) = \max[2\delta * f(t - 1, k), (f(t - 1, k))^2 + f(t - 1, k)]$. We define $C(k) =: \max_{t \leq T} f(t, k)$. This proves the lemma.

Proof of lemma 5.2:

Property 1: we show this by induction on t : At $t = 0$ it is true. $\mathcal{G}_m \in B^t \Rightarrow \mathcal{G}_m \in B^{t-1}$. By the induction hypothesis A_m contained no infected node at time $t - 1$. The dynamics is such that $\mathcal{G}_m \in B^t$ implies that for every $i \in \mathcal{G}_m$, $D^{t-1} \in R_i^{st}$. Moreover, for every $i \in I^{t-1}$, $X_i \notin \mathcal{G}_m$. Assuming property 1 and 2 true at $t - 1$, all infected nodes at $t - 1$ are in C^t and D^t . There may be $i \in \mathcal{G}_m$ such that $X_i \in C^{t-1}$, but by property 2, the node i

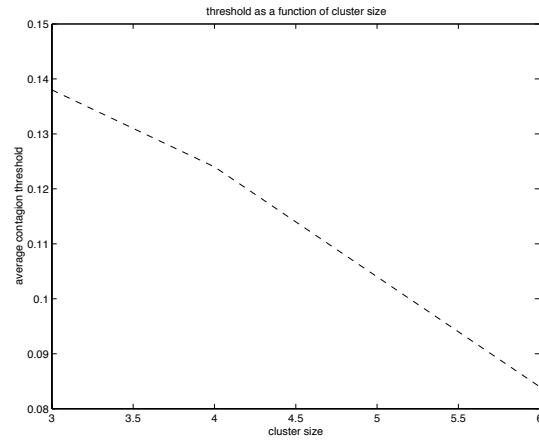


FIG 1. contagion threshold in independent community model as a function of cluster size ($N=10000$)

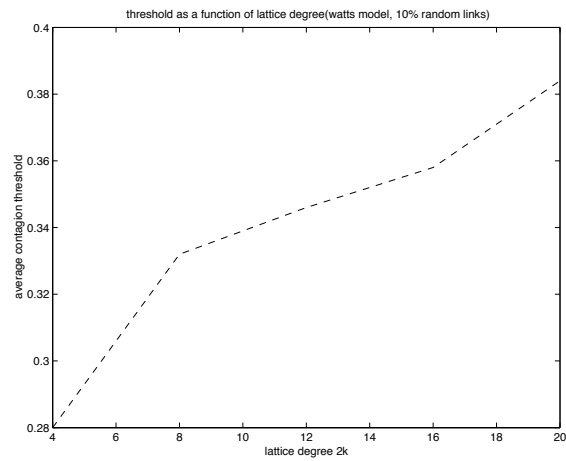


FIG 2. observed contagion threshold as a function of local degree

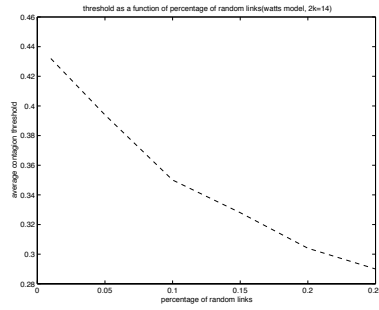


FIG 3. contagion threshold in watts model as function of fraction of shortcuts

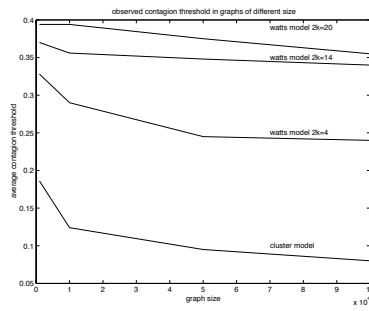


FIG 4. Contagion threshold in graphs of different size

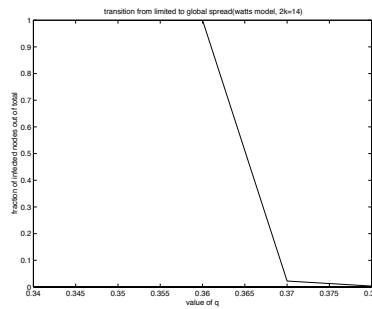


FIG 5. Transition around the contagion threshold (watts model)

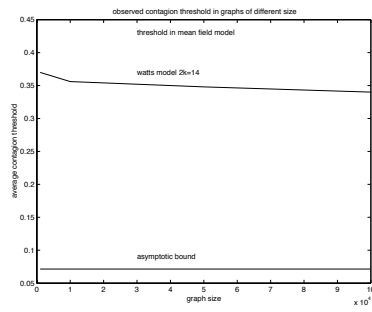


FIG 6. contagion threshold, $2k=14$, compared to mean field and asymptotic threshold

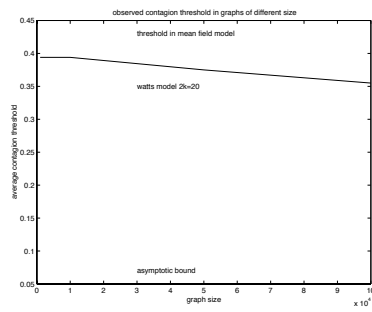


FIG 7. contagion threshold, $2k=20$, compared to mean field and asymptotic threshold

would be linked to only one infected node. Thus no node in \mathcal{G}_m is linked to more than one node in I^{t-1} and thus $A_m \cap I^{t-1} = \emptyset$

Property 2: first case: $\mathcal{G}_m \in B^{t-1} \cap C^t$. Then $A_m \cap I^{t-2} = \emptyset$. At time $t-1$, either, (a), there may be a single $j \in \mathcal{G}_m$ such that $X_j \in D^{t-1}$, or (b), there may be a node $i \in I^{t-1} - I^{t-2}$ such that $X_i = \mathcal{G}_m$, but not both. If (b) is not verified, $A_m \cap (I^{t-1} - I^{t-2}) = \emptyset$ and thus $\text{Card}(A_m - \mathcal{G}_m) \cap I^{t-1} = \text{Card}(A_m - \mathcal{G}_m) \cap I^{t-2} = 0$. Then (a) is verified. For $j \in \mathcal{G}_m$, we have $j \in I^t$ only if $X_j = D^{t-1}$. Thus $\mathcal{G}_m \cap I^t = j$ which shows the desired property at t .

Suppose that at $t-1$ (a) is not verified but (b) is. Then $(A_m - \mathcal{G}_m) \cap I^{t-1} = i$. Every node $l \in \mathcal{G}_m$ is linked to $i \in I^{t-1} \cap A_m$. Moreover, for every l , $C^{t-1} \cup T^t \cup D^{t-1} \in R_l^{st}$. Thus $V(l) \cap I^{t-1} = i$ and l cannot be infected at t . This shows that $\mathcal{G}_m \cap I^t = \emptyset$.

Second case: $\mathcal{G}_m \in C^{t-1}$. Then $\text{Card}(A_m - \mathcal{G}_m) \cap I^{t-1} = \text{Card}(A_m - \mathcal{G}_m) \cap I^{t-2}$ since $\mathcal{G}_m \in R_i^{st}$ for every $i \in I^{t-1}$. Consider a $j \in \mathcal{G}_m$, such that $j \notin I^{t-1}$. At $t-1$, A_m contained at most one infected node. moreover, $R_j^{st} \supset C^{t-1} \cup D^{t-1} \cup T^t$. Thus j is linked to at most one node in I^{t-1} . Thus j is not in I^t and thus $\text{Card}(\mathcal{G}_m \cap I^t) = \text{Card}(\mathcal{G}_m \cap I^{t-1})$. We use the induction hypothesis to conclude.

Proof 5.3 We show the result by induction. Assume that at time t , there is a c_t that does not depend on N such that $\text{Card}(C^t \cup D^t) < c_t$. $c_0 = k\delta$. The dynamics is such that in order to be in C^{t+1} or D^{t+1} , a cluster must either be a successor of I^t or contain a node whose successor is in D^t . The first case concerns at most $\delta \text{Card}(I^t) < \delta \text{Card}(C^t \cup D^t) < \delta c_t$ clusters. Since clusters are re-assigned from C^t and D^t whenever $\text{Card}(A_m) \geq c(k)$ (where $c(k)$ is the constant of lemma 5.1, there are less than $c(k) * \text{Card}(D^t) < c(k) * c_t$ clusters that are in C^{t+1} or D^{t+1} because they contain a node with a successor in D^{t-1} . Consequently $\text{Card}(C^{t+1} \cup D^{t+1}) < c_t + c_t(c(k) + \delta) =: c_{t+1}$. By induction we have the desired result. We take $c = \max_{t \leq T} c_t$. end of Proof

Proof of lemma 5.5: If $\mathcal{D}^{t-1} = \emptyset$, we have $I^t \subset C^t \cup D^t$. We begin by showing that $\mathbf{P}(\mathcal{D}^t \neq \emptyset | \mathcal{D}^{t-1} = \emptyset) \leq 1 - O(1/M)$. A cluster $\mathcal{G} \in C^{t-1}$ is in C^t if $\bigcup_{s < t} C^s \cup T^t \cup D^{t-1} \in R_i^{s-t}$ for every $i \in \mathcal{G}$ and if there is no $j \in \bigcup_{s < t} C^s \cup T^t \cup D^{t-1}$ such that $X_j = \mathcal{G}$. We apply lemma 2.4 with $H = C^t$ and $J = \bigcup_{s < t} C^s \cup T^t \cup D^{t-1}$. This shows that $\mathbf{P}(C^{t-1} \subset C^t) \geq 1 - O(1/M)$.

We have $B^{t-1} \subset B^t \cup T^t$, unless there are at least two nodes in I^{t-1} that have the same cluster in B^t for successor. We have $\text{Card}(j|j \in I^{t-1}) < c(s_1)$, and $\text{Card}(B^{t-1}) > M - c(s_1)$. Thus the probability that there are two nodes in I^{t-1} with the same successor is at most $O(1/M)$. Thus $\mathbf{P}(B^{t-1} \subset B^t \cup T^t) \geq 1 - O(1/M)$. We have $T^t = [\mathcal{G} \in B^{t-1} | \exists! i \in I^{t-1}, X_i = \mathcal{G}]$. Thus $\text{Card}(T^t) < \text{card}(I^{t-1}) < c(s_1)$. For a $\mathcal{G} \in T^t$, $\mathcal{G} \in D^t$ only if there are links between T^t and $C^t \cup T^t$. We apply lemma 5.4 to $H = T^t$ and $J = C^t \cup T^t$, which shows that $\mathbf{P}(T^t \cap D^t \neq \emptyset) < O(1/M)$. Together with the previous, this allows us to conclude that $\mathbf{P}(D^t = \emptyset) \geq 1 - O(1/M)$. It remains to show that $C^t \cap I^t = \emptyset$. We note that $C^t = [\mathcal{G}_m | \exists! i \in \mathcal{G}_m, X_i \in \mathcal{D}^{t-1}] \cup [\mathcal{G}_m | \exists j \in I^{t-1}, X_j = \mathcal{G}_m]$. When $\mathcal{D}^{t-1} = \emptyset$, the first set in the disjoint union is empty. $C^t \cap I^t = \emptyset$ for elements in the second set.

This concludes the proof.

Proof of lemma 5.6: $C^{t-1} \subset C^t$ if $C^{t-1} \cup T^t \cup D^{t-1} \in R_i^{st}$ for every $i \in \mathcal{G}_m$ and if there is no $j \in T^t \cup \mathcal{D}^{t-1}$ such that $X_j = \mathcal{G}_m$. We apply the lemma with $H = C^t$ and $J = C^{t-1} \cup T^t \cup D^{t-1}$ ($J = T^t \cup \mathcal{D}^{t-1}$). It follows that $\mathcal{P}(C^{t-1} \subset C^t) > 1 - O(1/M)$.

For B^{t-1} , first we show that $\mathbf{P}(B^{t-1} \subset B^t \cup T^t) > 1 - O(1/M)$: We have seen before that the probability that two nodes in I^{t-1} have the same successor is at most $O(1/M)$. (The probability that there is a cluster in B^{t-1} that contains two nodes that do not reject D^{t-1} is also at most $O(1/M)$). Consider the set $[\mathcal{G} \in B^{t-1} | \exists! i \in I^{t-1}, X_i = \mathcal{G}]$. The probability that there is a node in this set that does not reject the set D^{t-1} is bounded by $O(1/M)$ by the lemma. Together, this shows that $\mathbf{P}(B^{t-1} \subset B^t \cup T^t) \geq 1 - O(1/M)$.

It remains to show that $\mathbf{P}(T^t \subset C^t) \geq 1 - O(1/M)$. We apply lemma 5.4 to $H = J = \bigcup_{s < t} C^s \cup T^t$. This shows that the probability that $T^t \subset C^t$ is greater than $1 - O(1/M)$.

References

- [1] Blume, L. (1995). *How Noise matters*, Mimeo, Department of Economics, Cornell University, Ithaca, NY.
- [2] Bollobas, B. (2001). *Random Graphs*, New York: Academic Press.
- [3] Bonato, A. (2005). A survey of models of the web graph, *Lecture notes in computer science*, 3405, Springer Verlag.
- [4] Boyer D. and Miramontes O. (2003). Interface motion and pinning in small-world networks, *Physical Review E*, 67, 035102(R).

- [5] Brock, W. and Durlauf, S. (2001). Interactions-based models. In Heckman, J., Leamer, E. (eds) *Handbook of Econometrics*,5, Elsevier, Amsterdam:3297–3380.
- [6] Brown, J. and Reinegen,P. (1987). Social ties and word of mouth referral behavior, *Journal of consumer dynamics*14:3:350–362.
- [7] Castellano, C., Loreto, V., Barrat, A., Cecconi F., Parisi, D. (2005). Comparison of voter and Glauber ordering dynamics on networks, *Physical review E* 71, 066107.
- [8] Coleman, J., Katz, E., Menzel,H. (1996). *Medical Innovation, A diffusion study*. New York: Bobbs Merrill.
- [9] Cont, R.
- [10] Degenne, A., Fors, M. (1994). *Les rseaux sociaux*, Paris, Colin.
- [11] Durrett, R. (2006) *Random graph dynamics*, Cambridge University Press.
- [12] Durlauf, S., Young P. (editors) (2001). *Social Dynamics*, MIT press, Cambridge, MA.
- [13] Ellison, G. (1993), Learning, Local Interaction, and Coordination, *Econometrica*, Vol.61, no 5: 1047–1071.
- [14] Erdős P. and Renyi, A. (1960). On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutato Int. Közl.* 5: 17–61.
- [15] Fisher,C.S. (1948). *To Dwell among Friends* Chicago: Chicago University Press, 2nd edn, 1982
- [16] Freeman, L., and Thompson, C. (1989). Estimating Acquaintanceship Volume, in *M. Kochen (ed), The Small World* Norwood, NJ: Ablex: 147–158.
- [17] M. Girvan and M. E. J. Newman,(1999). Community structure in social and biological networks, *Proceedings of the National Academy of Science*, USA.
- [18] Granovetter, M. (1978). Threshold models of collective behavior, *American Journal of Sociology* 83:1420-1443.
- [19] Granovetter, M. (1973). The strength of weak ties, *American Journal of Sociology*, 78: 1287-1303
- [20] Granovetter, M., (1995) *Getting a Job: a Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.
- [21] Jackson, M.,Rogers,B. (2005). The Economics of Small Worlds, *Journal of the European Economic Association*,MIT Press, Vol. 3, No. 2-3: 617–627.
- [22] Jackson, M., A Survey of Models of Network Formation: Stability and Efficiency, (2003 version, forthcoming in *Group Formaion in Economics: Networks, Clubs and Coalitions*, edited by Gabrielle Demange, Myrna

- Wooders, Cambridge University Press).
- [23] Jackson, B. (2005). Search in the Formation of Large Networks: How Random are Socially Generated Networks? *Game Theory and Information* 0503005, EconWPA.
 - [24] S. C. Johnson (1967). *Hierarchical Clustering Schemes*, Psychometrika, 2:241–254.
 - [25] Katz,E., Lazarsfeld,P.F.,*Personal Influence:the Part Played by People in the Flow of Mass Communication*,Glencoe,IL, Free Press
 - [26] Kempe,D., Kleinberg,J.,Tardos, E. (2003). Maximizing the Spread of Influence through a Social Network. *Proc. 9th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*.
 - [27] Liggett, T. (1985). *Interacting Particle Systems*, New York, Springer Verlag.
 - [28] Kandori, M, Mailath, G, Rob, R. (1993). Learning Mutation and Long Run Equilibria in Games, *Econometrica* 61: 29–56.
 - [29] Kiliworth, P., and Bernard,R. (1978). The reverse small-world experiment. *Social Networks* 1 (11/78): 159–92.
 - [30] Milgram,S (1997). The Small World Problem, *Psychology Today*,1:61–67
 - [31] Morris,S.,Contagion, (2000). *Review of Economic Studies*, vol 67,pp 57–78.
 - [32] Newman, M., Barabasi,A.,Watts,D. (2006). *The structure and Dynamics of Networks* Princeton University Press.
 - [33] Newman, A. (2002). Assortative mixing in Networks, *Physics Revue Letters*, 89, 208701.
 - [34] Rapoport,A., and Horvath, W. (1961). A Study of a Large Sociogram, *Behavioral Science* 6: 279–291.
 - [35] Rogers, E. (1983). *Diffusion of Innovations*, 3d ed. New York :Free Press.
 - [36] Ryan,B., Gross, N., (1943). The diffusion of hybrid seed corn in two Iowa communities, *Rural Sociology*,8: 15–24.
 - [37] Steyer, A., Zimmermann, J.B, (2004). Influence sociale et diffusion de l'innovation, *Mathematics and Social Sciences*,Les réseaux sociaux, number 168: 43–57.
 - [38] Tarde,G. *Les lois de l'imitation*, Reimpression Geneva: Slatkine reprints, 1979.
 - [39] Travers, J. and Milgram, S. (1969). An experimental study of the small world problem, *Sociometry*, 32, 425–443.
 - [40] Valente, T.(1996). Social Network Thresholds in the Diffusion of Informations, *Social Networks* 18:69–89.

- [41] Vega-Redondo, F., (2007). *Complex Social Networks*, Econometric Society Monographs, Cambridge University Press.
- [42] Valente, T. (1995). *Network Models of the Diffusion of Innovations* Creskill, NJ, Hampton Press.
- [43] Watts, D. J. (1999). *Small Worlds*, Princeton University Press.
- [44] Watts, D. J. and Strogatz, S.H. (1998). Collective dynamics of small world networks, *Nature*, 393:440–442.
- [45] Wellman, B., Carrington, P., Hall, A. (1998). Networks as personal communities, in B. Wellman and S.D. Berkowitz (eds), *Social Structure: a Network Approach*, Cambridge: Cambridge University Press.
- [46] Young, P. (2002). The Diffusion of Innovations in Social Networks, *Economics Working Paper Archive 437*, The Johns Hopkins University, Department of Economics.