Diffusion and Cascading Behavior in Random Networks

Marc Lelarge*

Abstract

The spread of new ideas, behaviors or technologies has been extensively studied using epidemic models. Here we consider a model of diffusion where the individuals' behavior is the result of a strategic choice. We study a simple coordination game with binary choice and give a condition for a new action to become widespread in a random network. We also analyze the possible equilibria of this game and identify conditions for the coexistence of both strategies in large connected sets. Finally we look at how can firms use social networks to promote their goals with limited information.

Our results differ strongly from the one derived with epidemic models. In particular, we show that connectivity plays an ambiguous role: while it allows the diffusion to spread, when the network is highly connected, the diffusion is also limited by high-degree nodes which are very stable. In the case of a sparse random network of interacting agents, we compute the contagion threshold for a general diffusion model and show the existence of (continuous and discontinuous) phase transitions. We also compute the minimal size of a seed of new adopters in order to trigger a global cascade if these new adopters can only be sampled without any information on the graph. We show that this minimal size has a non-trivial behavior as a function of the connectivity. Our analysis extends methods developed in the random graphs literature based on the properties of empirical distributions of independent random variables, and leads to simple proofs.

*INRIA - ENS, Marc.Lelarge@ens.fr

1 Introduction

There is a vast literature on epidemics on complex networks (see [34] for a review). Most of the epidemic models consider a transmission mechanism which is independent of the local condition faced by the agents concerned. However, if there is a factor of coordination or persuasion involved, relative considerations tend to be important in understanding whether some new belief or behavior is adopted [36]. To fix ideas, it may be useful to think of the diffusion process as modeling the adoption of a new technology. In this case, when an agent is confronted with the possibility of adoption, her decision depends on the persuasion effort exerted on her by each of her neighbors in the social network. More formally, those neighborhood effects can be captured as follows: the probability that an agent adopts the new technology when r out of her d neighbors have already adopted can be modeled by a threshold function: the probability is $v \in (0,1]$ if $r/d \geq \theta$ and it is zero otherwise. For $\theta = 1/2$ and $v = 1$, this would correspond to a local majority rule. More generally, some simple models for the diffusion of a new behavior have been proposed in term of a basic underlying model of individual decision-making: as individuals make decision based on the choices of their neighbors, a particular pattern of behavior can begin to spread across the links of the network [36], [9].

To illustrate our purpose, consider the basic game-theoretic diffusion model proposed by Morris [33]. Consider a graph G in which the nodes are the individuals in the population and there is an edge (i,j) if i and j can interact with each other. Each node has a choice between two possible behaviors labeled A and B. On each edge (i,j), there is an incentive for i and j to have their behaviors match, which is modeled as the following coordination game parameterized by a real number $q \in (0,1)$: if i and j choose A (resp. B), they each receive a payoff of q (resp. $(1-q)$); if they choose opposite strategies, then they receive a payoff of 0. Then the total payoff of a player is the sum of the payoffs with each of her neighbors. If the degree of node i is d_i and N_i^B is her number of neighbors playing B, then the payoff to i from choosing A is $q(d_i - N_i^B)$ while the payoff from choosing B is $(1-q)N_i^B$. Hence i should adopt B if $N_i^B > qd_i$ and A if $N_i^B \leq qd_i$. A number of qualitative insights can be derived from this simple diffusion model. Clearly, a network where all nodes play A is an equilibrium of the game as is the state where all nodes play B. Consider a network where all nodes initially play A. If a small number of nodes are forced to adopt strategy B (constituting the seed) and we apply best-response updates to other nodes in the network, then these nodes will be repeatedly applying the following rule: switch to B if enough of your neighbors have already adopted B. In this paper, we assume that updates are made simultaneously for each players. There can be a cascading sequence of nodes switching to B such that a network-wide equilibrium is reached in the limit. Most of the results on this model are restricted to deterministic (possibly infinite) graphs [33], [24]. In this work, we analyze the diffusion in the large population limit when the underlying graph is a random network $G(n, \mathbf{d})$ with n vertices and where $\mathbf{d} = (d_i)_1^n$ is a given degree sequence (see Section 3.1 for a detailed definition). Although random graphs are not considered to be highly realistic models of most real-world networks, they are often used as first approximation and are a natural first choice for a sparse interaction network in the absence of any known geometry of the problem [15]. There is now a large literature on the study of complex networks across different disciplines such as physics [34], mathematics [21], sociology [37], networking [27] or economics [36]. There is also a large literature on local interaction and adoption externalities [16], [31], [13], [38]. Similarly to these works, we study a game where players' payoffs depend on the actions taken by their neighbors in the network but not on the specific identities of these

neighbors. Our most general framework allows to deal with threshold games of complements, i.e. player has an increasing incentive to take an action as more neighbors take the action.

The main contribution of our paper is a model which allows to study rigorously semi-anonymous threshold games of complements with local interactions on a complex network. While most of the literature following [10], [5] is concerned with stochastic versions of best response dynamics on fixed networks, we focused here on properties of deterministic best response dynamics on random networks with given vertex degrees. Our analysis yields several insights into how the diffusion propagates and as a consequence into how to design optimal firm strategies. The diffusion of information has been an active research area recently [14], [29]. Empirical analysis of the topological patterns of cascades in the context of a large product recommendation networks is studied in [28] and [30]. Our results extend the previous analysis of global cascades made in [37] using a threshold model. They differ greatly from the study of standard epidemics models used for the analysis of the spread of viruses [3] where an entity begin as 'susceptible' and may become 'infected' and infectious through contacts with her neighbors with a given probability. Already in the simple model of Morris [33] presented above, we show that connectivity (i.e. the average number of neighbors of a typical agent) plays an ambiguous role: while it allows the diffusion to spread, when the network is highly connected, the diffusion is also limited by high-degree nodes which are very stable. These nodes require a lot of their neighbors to switch to B in order to play B themselves. In the case of a sparse random network of interacting agents, we compute the contagion threshold in Section 2.1 and show the existence of (continuous and discontinuous) phase transitions. In Section 2.2, we also compute the minimal size of a seed of new adopters in order to trigger a global cascade if these adopters can only be sampled without any information on the graph. We show that this minimal size has a non-trivial behavior as a function of the connectivity. Our results allow also to explain why social networks can display a great stability in the presence of continual small shocks that are as large as the shocks that ultimately generate a global cascade. Cascades can therefore be regarded as a specific manifestation of the robust yet fragile nature of many complex systems [7]: a system may appear stable for long periods of time and withstand many external shocks (robustness), then suddenly and apparently inexplicably exhibit a large scale cascade (fragility). In Section 2.3, we analyze the possible equilibria of the game for low values of q. In particular, we give conditions under which an equilibrium with coexistence of large (i.e. containing a positive fraction of the total population) connected sets of players A and B is possible. In Section 2.4, we give a heuristic argument allowing to recover the technical results which gives some intuition behind our formulas.

Our paper is divided into two parts. The first part is contained in the next section. We apply our technical results to the particular case of the model of Morris [33] presented above. For Erdős-Rényi random graphs, we describe our main findings and provide heuristics and intuitions for them. These results are direct consequences of our main Theorems stated and proved in the second part of the paper. This second part starts with Section 3, where we present in details the most general model of diffusion that we are able to analyze. We also state our main technical results: Theorem 9 and Theorem 11. Their proofs can be found in Sections 4 and 5 respectively.

Probability asymptotics: in this paper, we consider sequences of (random) graphs and asymptotics as the number of vertices n tends to infinity. For notational simplicity we will usually not show the dependency on n explicitly. All unspecified limits and other asymptotics statement are for $n \to \infty$. For example, w.h.p. (with high probability) means with probability tending to 1 as

$n \to \infty$ and $\to^p$ means convergence in probability as $n \to \infty$. Similarly, we use o_p , Ω_p and O_p in a standard way. For example, if $X^{(n)} = X$ is a parameter of the random graph, $X = o_p(n)$ means that $\mathbb{P}(X > \epsilon n) \to 0$ as $n \to \infty$ for every $\epsilon > 0$, equivalently $X/n \to^p 0$, or for every $\epsilon > 0$, $|X| < \epsilon n$ w.h.p.

2 Analysis of a simple model of cascades

2.1 Contagion threshold for random networks

An interesting perspective is to understand how different network structures are more or less hospitable to cascades. Going back to previous model, we see that the lower q is, the easiest the diffusion spreads. In [33], the contagion threshold of a connected infinite network (called the cascade capacity in [9]) is defined as the maximum threshold q_c at which a finite set of initial adopters can cause a complete cascade, i.e. the resulting cascade of adoptions of B eventually causes every node to switch from A to B. There are two possible models to consider depending whether the initial adopters changing from A to B apply or not best-response update. It is shown in [33] that the same contagion threshold arises in both models. In this section, we restrict ourselves to the model where the initial adopters are forced to play B forever. In this case, the diffusion is monotone and the number of nodes playing B is non-decreasing. We say that this case corresponds to the permanent adoption model: a player playing B will never play A again. We will discuss the other model in Section 2.3.

We now compute the contagion threshold for a sequence of random networks. Since a random network is finite and not necessarily connected, we first need to adapt the definition of contagion threshold to our context. For a graph $G = (V, E)$ and a parameter q, we consider the largest connected component of the induced subgraph in which we keep only vertices of degree strictly less than q^{-1}. We call the vertices in this component pivotal players: if only one pivotal player switches from A to B then the whole set of pivotal players will eventually switch to B in the permanent adoption model. For a player $v \in V$, we denote by $C(v, q)$ the final number of players B in the permanent adoption model with parameter q, when the initial state consists of only v playing B, all other players playing A. Informally, we say that $C(v, q)$ is the size of the cascade induced by player v.

Proposition 1. *Consider the random graph $G(n, \mathbf{d})$ satisfying Conditions 8 and 10 below with asymptotic degree distribution $\boldsymbol{p} = (p_r)_{r=0}^{\infty}$, and define q_c by:*

$$q_c(\boldsymbol{p}) = q_c \quad = \quad \sup \left\{ q : \sum_{2 \le r < q^{-1}} r(r-1)p_r > \sum_{1 \le r} r p_r \right\}. \tag{1}$$

Let $P^{(n)}$ be the set of pivotal players in $G(n, \mathbf{d})$.

(i) For $q < q_c$, there are constants $0 < \gamma(q, \boldsymbol{p}) \le s(q, \boldsymbol{p})$ such that w.h.p. $\lim_n \frac{|P^{(n)}|}{n} = \gamma(q, \boldsymbol{p})$ and for any $v \in P^{(n)}$, $\liminf_n \frac{C(v,q)}{n} \ge s(q, \boldsymbol{p})$.

(ii) For $q > q_c$, for an uniformly chosen player v, we have $C(v, q) = o_p(n)$.

Note that we can rewrite (1) as follows: let D be a random variable with distribution $\boldsymbol{p}$, i.e. $\mathbb{P}(D = r) = p_r$, then

$$q_c = \sup \left\{ q : \mathbb{E}\left[D(D-1)\mathbb{1}(D < q^{-1}) \right] > \mathbb{E}[D] \right\}.$$

We can restate Proposition 1 as follows: let $C^{(n)}$ be the size of a cascade induced by switching a random player from A to B. Proposition 1 implies that for $q > q_c$, we have $\mathbb{P}(C^{(n)} > \epsilon n) \to 0$ as $n \to \infty$ for every $\epsilon > 0$ whereas for $q < q_c$, there exists constants $s(q, \boldsymbol{p}) \geq \gamma(q, \boldsymbol{p}) > 0$ depending only on q and the parameters of the graph such that $\liminf_n \mathbb{P}(C^{(n)} \geq (s(q, \boldsymbol{p}) - \epsilon)n) \geq \gamma(q, \boldsymbol{p})$ for every $\epsilon > 0$. Informally, we will say that global cascades (i.e. reaching a positive fraction of the population) occur when $q < q_c$ and do not occur when $q > q_c$. We call q_c defined by (1) the contagion threshold for the sequence of random networks with degree distribution (p_r). This result is in accordance with the heuristic result of [37] (see in particular the cascade condition Eq. 5 in [37]). Proposition 1 follows from Theorem 11 below which also gives estimates for the parameters $s(q, \boldsymbol{p})$ and $\gamma(q, \boldsymbol{p})$ when $q < q_c$. Under additional technical conditions, we have $\lim_n \frac{C(v,q)}{n} = s(q, \boldsymbol{p})$ in case (i).

Note that for $q > 1/3$, we have, $\sum_{r<q^{-1}} r(r-1)p_r \leq 2p_2 \leq \sum_r rp_r$, as soon as $p_0+p_1+p_2 < 1$. Hence we have the following elementary corollary:

Corollary 2. *For any random graph $G(n, \mathbf{d})$ satisfying Conditions 8 and 10 below and with at least a positive fraction of nodes of degree greater than 3, we have $q_c \leq 1/3$.*

Recall that for general graphs, it is shown in [33] that $q_c \leq 1/2$. Note also that our bound in Corollary 2 is tight: for a r-regular network chosen uniformly at random, Proposition 1 implies that $q_c = r^{-1}$ corresponding to the contagion threshold of a r-regular tree (see [33]). For $r \geq 3$ and $q < r^{-1}$, an initial adopter will cause a complete cascade (i.e. reaching all players) in a random r-regular graph w.h.p. In particular, in this case, the only possible equilibria of the game are all players playing A or all players playing B. We will come back to the analysis of equilibria of the game in Section 2.3.

We now consider some examples. First the case of Erdős-Rényi random graphs $G(n, \lambda/n)$ where each of the $\binom{n}{2}$ edges is present with probability λ/n for a fixed parameter λ. In this case, we can apply our results with $p_r = e^{-\lambda}\frac{\lambda^r}{r!}$ for all $r \geq 0$. As shown in Figure 1 for the case of Erdős-Rényi random graphs $G(n, \lambda/n)$, q_c is well below $1/3$, indeed we have $q_c \leq 1/4$ for any value of λ. As shown in Figure 2, we see that q_c is a non-decreasing function of the average degree in the graph λ for $\lambda \leq 2$. Clearly on Figure 1, q_c is a non-increasing function of λ, for $\lambda \geq 4$.

The second curve in Figure 1 corresponds to the contagion threshold for a scale-free random network whose degree distribution $p_r = \frac{r^{-\gamma}}{\zeta(\gamma)}$ (with $\zeta(\gamma) = \sum r^{-\gamma}$) is parameterized by the decay parameter $\gamma > 1$. We see that in this case we have $q_c \leq 1/9$. In other words, in an Erdős-Rényi random graph, in order to have a global cascade, the parameter q must be such that any node with no more than four neighbors must be able to adopt B even if it has a single adopting neighbor. In the case of the scale free random network considered, the parameter q must be much lower and any node with no more than nine neighbors must be able to adopt B with a single adopting neighbor. This simply reflects the intuitive idea that for widespread diffusion to occur there must be a sufficient high frequency of nodes that are certain to propagate the adoption.

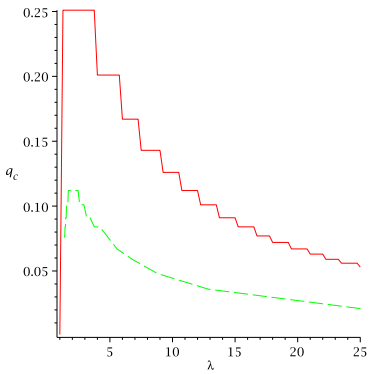

Figure 1: $q_c(\lambda)$ for Erdős-Rényi random graphs and for power law graphs (dashed curve) as a function of λ the average degree.

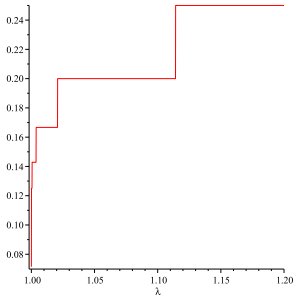

Figure 2: $q_c(\lambda)$ for Erdős-Rényi random graphs as a function of λ in the range $\lambda \in [1; 1.2]$.

We also observe that in both cases, for q sufficiently low, there are two critical values for the parameter λ, $1 < \lambda_i(q) < \lambda_s(q)$ such that a global cascade for a fixed q is only possible for $\lambda \in (\lambda_i(q); \lambda_s(q))$. The heuristic reason for these two thresholds is that a cascade can be prematurely stopped at high-degree nodes. For Erdős-Rényi random graphs, when $1 \leq \lambda < \lambda_i(q)$, there exists a "giant component", i.e. a connected component containing a positive fraction of the nodes. The high-degree nodes are quite infrequent so that the diffusion should spread easily. However, for λ close to one, the diffusion does not branch much and progresses along a very thin tree, "almost a line", so that its progression is stopped as soon as it encounters a high-degree node. Due to the variability of the Poisson distribution, this happens before the diffusion becomes too big for $\lambda < \lambda_i(q)$. Nevertheless the condition $\lambda > \lambda_i(q)$ is not sufficient for a global cascade. Global diffusion also requires that the network not be too highly connected. This is reflected by the existence of the second threshold $\lambda_s(q)$ where a further transition occurs, now in the opposite direction. For $\lambda > \lambda_s(q)$, the diffusion will not reach a positive fraction of the population. The intuition here is clear: the frequency of high-degree nodes is so large that diffusion cannot avoid them and typically stops there since it is unlikely that a high enough fraction of their many neighbors eventually adopts. Following [37], we say that these nodes are locally stable.

The proof of our Theorem 11 makes this heuristic rigorous for a more general model of diffusion and gives also more insights on the nature of the phase transitions. We describe it now. The lower curve in Figure 3 represents the number of pivotal players in an Erdős-Rényi

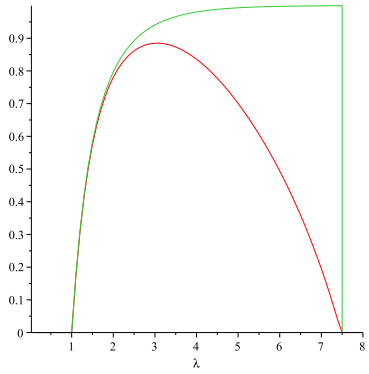

Figure 3: Size $s(q, \lambda)$ of the cascade (in percent of the total population) for Erdős-Rényi random graphs as a function of λ the average degree for a fixed $q = 0.15$. The lower curve gives the asymptotic fraction of pivotal players $\gamma(q, \lambda)$.

random graphs as a function of λ the average connectivity for $q^{-1} = 6.666...$: hence we keep only the largest connected component of an Erdős-Rényi random graph where we removed all vertices of degree greater than 6. By the same heuristic argument as above, we expect two phase transitions for the size of the set of pivotal players. In the proof of Theorem 11, we show that it is indeed the case: the phase transitions occur at the same values $\lambda_i(q)$ and $\lambda_s(q)$ as can be seen on Figure 3 where the normalized size $\gamma(q, \lambda)$ of the set of pivotal players is positive only for $\lambda \in (\lambda_i(q), \lambda_s(q))$. Hence a cascade is possible if and only if there is a 'giant' component of pivotal players. Note also that both phase transitions for the pivotal players are continuous, in the sense that the function $\lambda \mapsto \gamma(q, \lambda)$ is continuous. This is not the case for the second phase transition for the size of the cascade: the function $\lambda \mapsto s(q, \lambda)$ is continuous in $\lambda_i(q)$ but not in $\lambda_s(q)$ as depicted on Figure 3. This has important consequences: around $\lambda_i(q)$ the propagation of cascades is limited by the connectivity of the network as in standard epidemic models. But around $\lambda_s(q)$, the propagation of cascades is not limited by the connectivity but by the high-degree nodes which are locally stable.

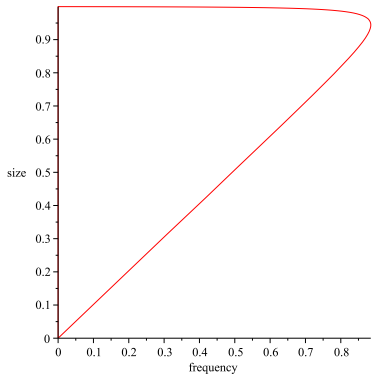

Figure 4: Size $s(q, \lambda)$ of the cascade (in percent of the total population) for Erdős-Rényi random graphs as a function of the frequency of the cascade.

To better understand this second phase transition, consider a dynamical model where at each step a player is chosen at random and switch to B. Once the corresponding cascade is triggered, we come back to the initial state where every node play A before going to the next step. Then for λ less than but very close to $\lambda_s(q)$, most cascades die out before spreading very far. However, a set of pivotal players still exists, so very rarely a cascade will be triggered by such a player in which case the high connectivity of the network ensures that it will be extremely large. As λ approaches $\lambda_s(q)$ from below, global cascades become larger but increasingly rare until they disappear implying a discontinuous phase transition as shown in Figure 4: for low values of λ, the cascades are infrequent and small. As λ increases, their frequencies and sizes also increase until a point where the cascade reaches almost all vertices of the giant component of the graph. Then as λ increases, their sizes remain almost constant but their frequencies are decreasing.

2.2 Advertising with word of mouth communication

We consider now scenarios where $\lambda \notin [\lambda_i(q), \lambda_s(q)]$ and the initial set of adopters grows linearly with the total population n. More precisely, consider now a firm advertising to a group of consumers, who share product information among themselves: potential buyers are not aware of the existence of the product and the firm undertakes costly informative advertising. The firm chooses the fraction of individuals who receive advertisements. Individuals are located in a social network modeled by a random network $G(n, \mathbf{d})$ with given vertex degrees as in previous section. However contrary to most work on viral marketing [35], [23], we assume that the advertiser has limited knowledge about the network: the firm only knows the proportions of individuals having different degrees in the social network. One possibility for the firm is to sample individuals randomly and to decide the costly advertising for this individual based on her degree (i.e. her number of neighbors). The action of the firm is then encoded in a vector $\boldsymbol{\alpha} = (\alpha_d)$, where α_d represents the fraction of individuals with degree d which are directly targeted by the firm. These individuals will constitute the seed and we call them the early adopters. Note that the case $\alpha_d = \alpha$ for all d corresponds to a case where the firm samples individuals uniformly. This might be one possibility if it is unable to observe their degrees. In order to optimize its strategy, the firm needs to compute the expected payoff of its marketing strategy as a function of $\boldsymbol{\alpha}$. Our results allows to estimate this function in terms of $\boldsymbol{\alpha}$ and the degree distribution in the social network.

We assume that a buyer might buy either if she receives advertisement from the firm or if she receives information via word of mouth communication [11]. More precisely, following [12], we consider the following general model for the diffusion of information: a buyer obtains information as soon as one of her neighbors buys the product but she decides to buy the product when $\mathrm{Bi}(k, \pi) > K(s)$ where s is her number of neighbors and k the number of neighbors having the product, $\mathrm{Bi}(k, \pi)$ is a Binomial random variable with parameters k and $\pi \in [0, 1]$ and $K(s)$ is a general random variable. In words, π is the probability that a particular neighbor does influence the possible buyer. This possible buyer does actually buy when the number of influential neighbors having bought the product exceeds a threshold $K(s)$. Thus, the thresholds $K(s)$ represent the different propensity of nodes to buy the new product when their neighbors do. The fact that these are possibly randomly selected is intended to model our lack of knowledge of their values and a possibly heterogeneous population. Note that for $K(s) = 0$ and $\pi \in [0, 1]$, our model of diffusion is exactly a contact process with probability of contagion between neighbors

equals to π. This model is also called the SI (susceptible-infected) model in mathematical epidemiology [3]. Also for $\pi = 1$ and $K(s) = qs$, we recover the model of Morris [33] described previously.

We now give the asymptotic for the final number of buyers for the case $\pi \in [0,1]$ and $K(s) = qs$, with $q \leq 1$ (the general case with a random threshold is given in Theorem 9). We first need to introduce some notations. For integers $s \geq 0$ and $0 \leq r \leq s$ let b_{sr} denote the binomial probabilities $b_{sr}(p) := \mathbb{P}(\mathrm{Bi}(s,p) = r) = \binom{s}{r} p^r (1-p)^{s-r}$. Given a distribution $\boldsymbol{p} = (p_s)_{s \in \mathbb{N}}$, we define the functions:

$$\begin{aligned} h(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &:= \sum_s (1-\alpha_s) p_s \sum_{r \geq s - \lfloor sq \rfloor} r b_{sr}(1 - \pi + \pi z), \\ g(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &:= \lambda(\boldsymbol{p}) z (1 - \pi + \pi z) - h(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi), \\ h_1(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &:= \sum_s (1-\alpha_s) p_s \sum_{r \geq s - \lfloor sq \rfloor} b_{sr}(1 - \pi + \pi z), \end{aligned}$$

where $\lambda(\boldsymbol{p}) = \sum_s s p_s$. We define

$$\hat{z}(\boldsymbol{\alpha}, \boldsymbol{p}, \pi) := \max \{ z \in [0,1] : \ g(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) = 0 \}.$$

We refer to Section 2.4 for an intuition behind the definitions of the functions h, h_1, g and $\hat{z}$ in terms of a branching process approximation of the local structure of the graph.

Proposition 3. *Consider the random graph $G(n, \boldsymbol{d})$ for a sequence $(d_i)_1^n$ satisfying Condition 8 below. If the strategy of the firm is given by $\boldsymbol{\alpha}$, then the final number of buyers is given by $(1 - h_1(\hat{z}, \boldsymbol{\alpha}, \boldsymbol{p}, \pi))n + o_p(n)$ provided $\hat{z}(\boldsymbol{\alpha}, \boldsymbol{p}, \pi) = 0$, or $\hat{z}(\boldsymbol{\alpha}, \boldsymbol{p}, \pi) \in (0,1]$, and further $g(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) < 0$ for any z in some interval $(\hat{z} - \epsilon, \hat{z})$.*

To illustrate this result, we consider the simple case of Erdős-Rényi random graphs $G(n, \lambda/n)$ with $\pi = 1$, $\alpha_d = \alpha$ for all d, and $q > q_c(\lambda)$ where $q_c(\lambda)$ is the contagion threshold for this network (defined in previous section). Figure 5 gives a plot of the corresponding function $z \mapsto g(z, \alpha)$ for three different values of α (the parameter λ is fixed). We see that $\alpha \mapsto \hat{z}(\alpha)$ is discontinuous at a value $\alpha_c(\lambda) \approx 6\%$.

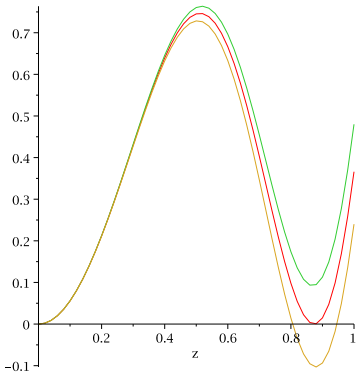

Figure 5: Function $z \mapsto g(z)$ for three different values of $\alpha \approx 4; 6; 8\%$ ($\lambda = 6$ and $q = 0.3$).

For $\alpha < \alpha_c(\lambda)$, the value of $\hat{z}(\alpha)$ is close to one whereas for $\alpha > \alpha_c(\lambda)$, it is close to zero. As a result for $\alpha < \alpha_c(\lambda)$, the size of the diffusion is rather small whereas it becomes very large

(reaching more than 99% of the population) for $\alpha > \alpha_c(\lambda)$. This is reflected in Figure 6, where the asymptotic for the final fraction of buyers is plotted as a function of α.

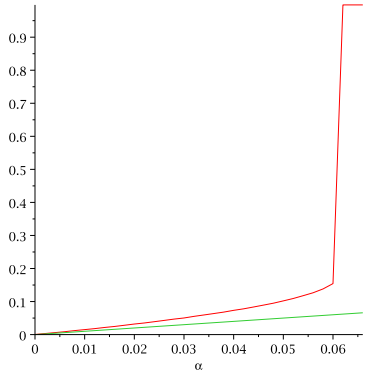

Figure 6: Function $\alpha \mapsto 1 - h_1(\hat{z}, \alpha)$ (the green function is $z \mapsto z$, the size of the initial seed).

There is a phase transition in α and clearly the advertising firm has a high interest in reaching the "critical mass" of $\alpha_c(\lambda)n$ early adopters. This phase transition is reminiscent of the one described in previous section. Recall that we are now in a setting where $q > q_c(\lambda)$ so that a global cascade triggered by a single (or a small number of) individual(s) uniformly sampled is not possible. Hence, the firm will have to advertise to a positive fraction of individuals constituting the seed for the diffusion. The intuition behind Figure 6 is the following: if the seed is too small, then each early adopter starts a small cascade which is stopped by high-degree nodes. When the seed reaches the critical mass, then the cascades "coalesce" and are able to overcome barriers constituted by high-degree nodes so that a large fraction of the nodes in the "giant component" of the graph adopt. Then increasing the size of the seed has very little effect since most of the time, the new early adopters (chosen at random by the firm) are already reached by the global diffusion. Our Theorem 9 makes this heuristic rigorous for the general model of diffusion described above (with random thresholds).

We now give some more insights by exploring different scenarios for different values of q. Consider first a case where $q < 1/4$, then thanks to the results of previous section, we know that there exists $1 < \lambda_i(q) < \lambda_s(q)$ such that for any $\lambda \in (\lambda_i(q), \lambda_s(q))$, a global cascade is possible with positive probability if only one random player switches to B. In particular, if a fraction α of individuals uniformly sampled are playing A, then for any $\alpha > 0$, such a global cascade will occur with high probability. For $\lambda > \lambda_s(q)$, we know that a single player cannot trigger a global cascade. More precisely, if players are chosen at random without any information on the underlying network, any set of initial adopters with size $o(n)$ cannot trigger a global cascade, as shown by Theorem 11. However the final size of the set of players playing B is a discontinuous function of α the size of the initial seed. If $\alpha_c(\lambda)$ is defined as the point at which this function is discontinuous, we have: for $\alpha < \alpha_c(\lambda)$, the final set of players B will be only slightly larger than α but if $\alpha > \alpha_c(\lambda)$ the final set of players B will be very large.

Hence we will say that there is a global cascade when $\alpha > \alpha_c(\lambda)$ and that there is no global cascade when $\alpha < \alpha_c(\lambda)$. As shown in Figure 7, for $q < 1/4$, we have $\alpha_c(\lambda) = 0$ for $\lambda \in [\lambda_i(q), \lambda_s(q)]$ and $\alpha_c(\lambda) > 0$ for $\lambda > \lambda_s(q)$. In Figure 7, we also see that our definition of global cascade when $\lambda > \lambda_s(q)$ is consistent with our previous definition since the function

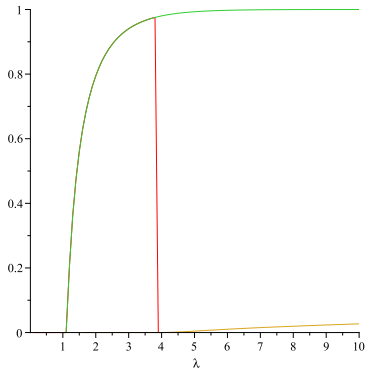

Figure 7: Case $q < 1/4$. The green (upper) curve is the size of the final set of players B for $\alpha = \alpha_c(\lambda)$ given by $\lambda \mapsto 1 - h_1(\hat{z}, \alpha_c(\lambda))$ as a function of λ. The brown (lower) curve corresponds to the minimal size of the seed for a global cascade $\lambda \mapsto \alpha_c(\lambda)$. The red curve corresponds to the size of a global cascade when $\alpha = 0$ similar as in Figure 3.

$\lambda \mapsto 1 - h_1(\hat{z}, \alpha_c(\lambda))$ giving the size of the final set of players B when the seed has normalized size $\alpha_c(\lambda)$ is continuous in λ and agrees with the previous curve for the size of the final set of players B when $\lambda < \lambda_s(q)$ and with only one early adopter.

We now consider the case where $q > 1/4$. In this case, thanks to the results of previous section, we know that for any value of λ, a single initial player B (sampled uniformly) cannot trigger a global cascade. But our definition of $\alpha_c(\lambda)$ still makes sense and we now have $\alpha_c(\lambda) > 0$ for all λ. Figure 8 gives a plot of the function $\lambda \mapsto \alpha_c(\lambda)$. We see that again there are two regimes associated with the low/high connectivity of the graph. For low values of λ, the function $\lambda \mapsto \alpha_c(\lambda)$ is non-increasing in λ. This situation corresponds to the intuition that is correct for standard epidemic models according to which an increase in the connectivity makes the diffusion easier and hence the size of the critical initial seed will decrease.

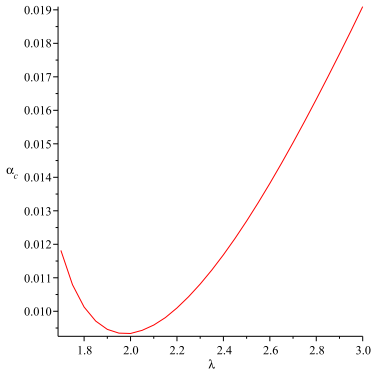

Figure 8: Case $q > 1/4$. Minimal size of the seed for a global cascade $\lambda \mapsto \alpha_c(\lambda)$ as a function of λ.

Figure 9 shows the size $1 - h_1(\hat{z}, \alpha)$ of the final set of players B as a function of the size of the initial seed α for small values of λ. We see that for the smallest values of λ, there is no

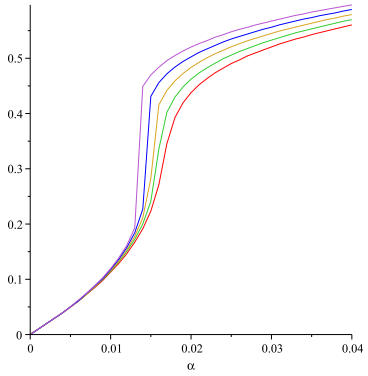

Figure 9: Case $q > 1/4$, low connectivity regime. Size of the final set of players B, $\alpha \mapsto 1 - h_1(\hat{z}, \alpha)$ as a function of the size of the initial seed α for different values of $\lambda = 1.60, 1.61, 1.62, 1.63, 1.64$.

discontinuity for the function $\alpha \mapsto 1-h_1(\hat{z}, \alpha)$. In this case $\alpha_c(\lambda)$ is not defined. We also see that there is a natural monotonicity in λ: as the connectivity increases, the diffusion also increases. However Figure 8 shows that for $\lambda \geq 2$, the function $\lambda \mapsto \alpha_c(\lambda)$ becomes non-decreasing in λ. Hence even if the connectivity increases, in order to trigger a global cascade the size of the initial seed has to increase too. The intuition for this phenomenon is now clear: in order to trigger a global cascade, the seed has to overcome the local stability of high-degree nodes which now dominates the effect of the increase in connectivity.

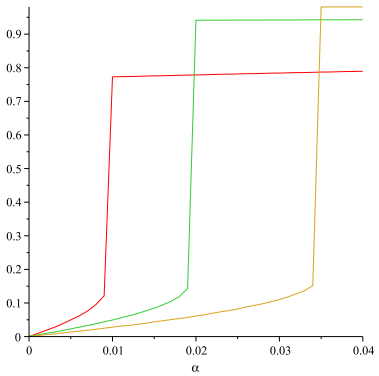

Figure 10: Case $q > 1/4$, high connectivity regime. Size of the final set of players B, $\alpha \mapsto 1 - h_1(\hat{z}, \alpha)$ as a function of the size of the initial seed α for different values of $\lambda = 2, 3, 4$ (red, green, brown respectively).

Figure 10 shows the size $1 - h_1(\hat{z}, \alpha)$ of the final set of players B as a function of the size of the initial seed α for various value of λ. Here we see that connectivity hurts the start of the diffusion: for small value of α, an increase in λ results in a lower size for the final set of players B! However when α reaches the critical value $\alpha_c(\lambda)$, then a global cascade occurs and its size is an increasing function of λ. In other words, high connectivity inhibits the global cascade but once it occurs, it facilitates its spread.

2.3 Equilibria of the game and coexistence

We considered so far the permanent adoption model: the only possible transitions are from playing A to B. There is another possible model in which the initial adopters playing B also apply best-response update. We call this model the non-monotonic model. In this model, if the dynamic converges, the final state will be an equilibrium of the game. An equilibrium of the game is a fixed point of the best response dynamics. For example, the states in which all players play A or all players play B are trivial equilibria of the game. Note that the permanent adoption model does not necessarily yield to an equilibrium of the game as the initial seed does not apply best response dynamics. To illustrate the differences between the two models consider the graph of Figure 11 for a value of $q = 1/3$: if the circled player switches to B, the whole network will eventually switch to B in the permanent adoption model, whereas the dynamic for the non-monotonic model will oscillate between the state where only the circled player plays B and the state where only his two neighbors of degree two play B.

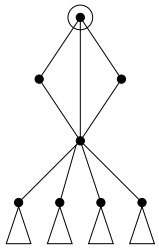

Figure 11: The circled player induces a global cascade in the permanent adoption model but not in the non-monotonic model for $q = 1/3$ (each triangle represents a long line of players).

Clearly if a player induces a global cascade for the non-monotonic model, it will also induce a global cascade in the permanent adoption model. As illustrated by the case of Figure 11, the converse is not true in general. Hence, a priori, a pivotal player as defined in previous section might not induce a global cascade in the non-monotonic model. In [33], it is shown that if one can find a finite set of initial adopters causing a complete cascade for the permanent adoption model, it is also possible to find another, possibly larger but still finite, set of initial adopters leading to a complete cascade for the non-monotonic model. Hence the contagion threshold as defined in [33] is the same for both models. In our case, we see that if we switch from A to B two pivotal players who are neighbors, then the whole set of pivotal players will eventually switch to B in the non-monotonic model. In our case also, we say that the contagion threshold is the same in both models. Moreover, both models will have exactly the same dynamics if started with the set of pivotal players playing B and all other players playing A. In particular, it shows that the dynamic converges and reaches an equilibrium of the game. Hence we have the following corollary:

Corollary 4. *Consider the random graph $G(n, \mathbf{d})$ satisfying Conditions 8 and 10 below with asymptotic degree distribution $\boldsymbol{p} = (p_r)_{r=0}^{\infty}$. For $q < q_c$, there exists w.h.p. an equilibrium of the game in which the number of players B is more than $s(q, \boldsymbol{p})n$ (defined in Proposition 1) and it can be obtained from the trivial equilibrium with all players playing A by switching only two neighboring pivotal players. We call this equilibrium the pivotal equilibrium.*

Hence for $q < q_c$, the trivial equilibrium all A, is rather 'weak' since two pivotal players can induce a global cascade and there are $\Omega_p(n)$ such players so that switching two neighbors at random will after a finite number of trials (i.e. not increasing with n) induce such a global cascade. Figure 12 shows the average number of trials required. It goes to infinity at both extreme $\lambda_i(q)$ and $\lambda_s(q)$. Moreover, we see that for most values of λ inside this interval, the average number of trials is less than 2. If $q > q_c$, then by definition if there are pivotal players, their number must be $o_p(n)$. Indeed, in the case of r-regular graphs, there are no pivotal players for $q > q_c = r^{-1}$. Hence, it is either impossible or very difficult (by sampling) to find a set of players with cardinality bounded in n leading to a global cascade since in all cases, their number is $o_p(n)$.

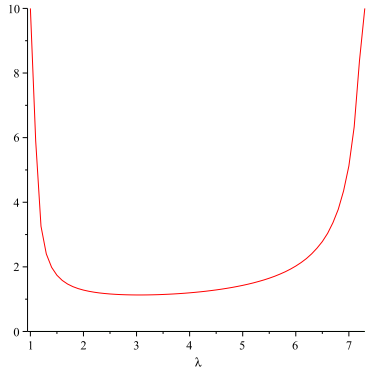

Figure 12: Average number of trials required to trigger a global cascade for Erdős-Rényi random graphs $G(n, \lambda/n)$ as a function of λ.

In the case $q < q_c$, for values of λ close to $\lambda_s(q)$, the global cascade reaches almost all nodes of the giant component. However for λ close to $\lambda_i(q)$, this is not the case as shown by the following proposition:

Proposition 5. *In an Erdős-Rényi random graph $G(n, \lambda/n)$, for $q < q_c$, there exists $\lambda_c(q) \in [\lambda_i(q), \lambda_s(q)]$ such that:*

- *for $\lambda \in (\lambda_i(q), \lambda_c(q))$, in the pivotal equilibrium, there is coexistence of a giant component of players A and a giant component of players B.*
- *for $\lambda > \lambda_c(q)$, in the pivotal equilibrium, there is no giant component of players A, although there might be a positive fraction of players A.*

Here a giant component is a subset of vertices containing a positive fraction (as n tends to infinity) of the total size of the graph such that the corresponding induced graph is connected. Figure 13 illustrates this proposition in the case of Erdős-Rényi random graphs $G(n, \lambda/n)$. The difference between the upper (red) and lower (green) curve is exactly the fractions of players A in the pivotal equilibrium while the brown curve represent the size (in percentage of the total population) of the largest connected component of players A in the pivotal equilibrium. This curve reaches zero exactly at $\lambda_c(q)$. Hence for $\lambda > \lambda_c(q)$, we see that there is still a positive fraction of the population playing A, but the set of players A is divided in small (i.e. of size $o(n)$) connected components, like 'islands' of players A in a 'sea' of players B. Note also that

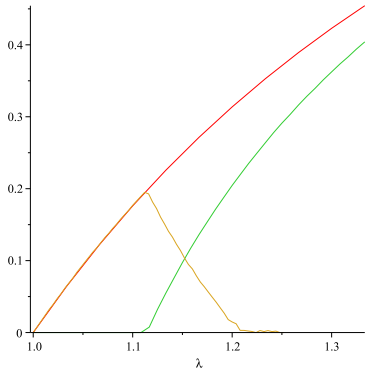

Figure 13: Coexistence: the upper (red) curve is the fraction of players in the giant component as a function of the average degree λ. The lower (green) curve is the fraction of players B in the pivotal equilibrium as a function of λ (for $q = 0.2$). The last (brown) curve is the size of the giant component of players A in the pivotal equilibrium as a function of λ.

for Erdős-Rényi random graphs, the value of $\lambda_c(q)$ is close to $\lambda_i(q)$. Proposition 5 follows from the following proposition whose proof is given in Section 6.

We first introduce some notations: for D a random variable, we define for $0 \leq x \leq 1$, D_x the thinning of D obtained by taking D points and then randomly and independently keeping each of them with probability x, i.e. given $D = d$, $D_x \sim \mathrm{Bi}(d, x)$.

We define $g(z, \boldsymbol{p}) = \sum_s p_s \sum_{r \geq s - \lfloor sq \rfloor} r b_{sr}(z)$ (corresponding to the function g defined in previous section for **alpha** $= 0$ and $\pi = 1$), $\lambda(\boldsymbol{p}) = \sum_s s p_s$ and

$$\xi = \sup \left\{ z < 1,\ \lambda(\boldsymbol{p}) z^2 = g(z, \boldsymbol{p}) \right\}.$$

Proposition 6. *Consider the random graph $G(n, \mathbf{d})$ satisfying Conditions 8 and 10 below with asymptotic degree distribution $\boldsymbol{p} = (p_r)_{r=0}^{\infty}$. Let D be a random variable with distribution $\boldsymbol{p}$. If $q < q_c$, then $\xi < 1$. Assume moreover that there exists $\epsilon > 0$ such that $g(z, \boldsymbol{p}) < 0$ for $z \in (\xi - \epsilon, \xi)$. There is coexistence of a giant component of players A and a giant component of players B in the pivotal equilibrium if*

$$\mathbb{E}\left[D_\xi (D_\xi - 2) \mathbb{1}(qD \geq D - D_\xi) \right] > 0.$$

2.4 Local Mean Field approximation

In this subsection, we say that a player A (or a non-buyer) is inactive and a player B (or a buyer) is active. We describe an approximation to the local structure of the graph and present a heuristic argument which leads quickly to a prediction for the asymptotic probability of being active. This heuristic allows to recover the formulas for the functions h, g, h_1, but results fron this section are not needed for the rigorous proofs given in Sections 4 and 5. This branching process approximation is standard in the random graphs literature [8] and is called Local Mean Field (LMF) in [27, 26] for contact epidemics. In these particular cases, the LMF approximation was turned into a rigorous argument. For our model of diffusion with threshold, this seems unlikely to be straightforward and our proof will not use the LMF approximation.

The LMF model is characterized by its connectivity distribution $\boldsymbol{p} = (p_r)$. We now construct a tree-indexed process. Let T be a Galton-Watson branching process [2] with a root which has offspring distribution p_r and all other nodes have offspring distribution p_r^* given by $p_{r-1}^* = \frac{rp_r}{\lambda}$ for all $r \geq 1$. Recall that p_{r-1}^* corresponds to the probability that an edge points to a node with degree r, see [8]. The tree T describes the local structure of the graph G (as n tends to infinity): the exploration of the successive neighborhoods of a given vertex is approximated by the branching process T as long as the exploration is local (typically restricted to a finite neighborhoods independent of n).

We denote by $\emptyset$ the root of the tree and for a node i, we denote by $\mathrm{gen}(i) \in \mathbb{N}$ the generation of i, i.e. the length of the minimal path from $\emptyset$ to i. Also we denote $i \to j$ if i belongs to the children of j, i.e. $\mathrm{gen}(i) = \mathrm{gen}(j) + 1$ and j is on the minimal path from $\emptyset$ to i. For an edge (i, j) with $i \to j$, we denote by $T_{i \to j}$ the sub-tree of T with root i obtained by the deletion of edge (i, j) from T.

We now consider the diffusion model described in Section 2.2 (with $K(s) = qs$ and probability for a neighbor to be inluential is π) on the tree T, where the initial set of active nodes is given by a vector $\sigma = (\sigma_i)$. In our model the σ_i's are independent Bernoulli random variables with parameter α_{d_i} where d_i is the degree of node i in the tree. Thanks to the tree structure, it is possible to compute the probability of being active recursively as follows: for any node $i \neq \emptyset$, let $Y_i = 1$ if node i is active on the sub-graph $T_{i \to j} \cup \{(i, j)\}$ with initial set of active nodes given by the restriction of σ to individuals in $T_{i \to j}$ and with individual j held fix in the inactive state. Then for any node $i \neq \emptyset$ of degree d_i, i becomes active if the number of active influential children exceeds qd_i. Hence, we get

$$Y_i = 1 - (1 - \sigma_i)\mathbb{1}\left(\sum_{\ell \to i} B_{\ell i} Y_\ell \leq q d_i\right), \tag{2}$$

where the $B_{\ell i}$'s are independent Bernoulli random variables with parameter π. Then the state of the root is given by

$$X_\emptyset = 1 - (1 - \sigma_\emptyset)\mathbb{1}\left(\sum_{i \to \emptyset} B_{i\emptyset} Y_i \leq q d_\emptyset\right). \tag{3}$$

In order to compute the distribution of $X_\emptyset$, we first solve the Recursive Distributional Equation (RDE) associated to the Y_i's: thanks to the tree structure, the random variables Y_ℓ in (2) are i.i.d. and have the same distribution as Y_i. Hence their distribution solve the RDE given by

$$Y \overset{d}{=} 1 - (1 - \sigma(D^* + 1))\mathbb{1}\left(\sum_{i=1}^{D^*} B_i Y_i \leq q(D^* + 1)\right), \tag{4}$$

where for a given d, the random variable $\sigma(d)$ is Bernoulli with parameter α_d, B_i's are independent Bernoulli with parameter π, D^* has distribution p_r^*, Y and the Y_i are i.i.d. copies (with unknown distribution). To solve the RDE (4), we need to compute only the mean of the Bernoulli random variable Y. Hence taking expectation in (4) directly gives a fixed-point equation for this mean and the following lemma follows (its proof is defered to Section 7.1):

Lemma 7. *Let $x = \mathbb{P}[Y = 1]$, where the distribution of Y solves the RDE (4), then we have $\lambda(1 - x)(1 - x\pi) = h(1 - x; \boldsymbol{\alpha}, \boldsymbol{p}, \pi)$. Moreover $X_\emptyset$ in (3) is a Bernoulli random variable with parameter $1 - h_1(1 - x; \boldsymbol{\alpha}, \boldsymbol{p}, \pi)$.*

By the change of variable $z = 1-x$, we see that Lemma 7 is consistent with Proposition 3 and allow to recover the functions h, g, h_1. Clearly the crucial point allowing previous computation is the fact that in recursion (2) the Y_i can be computed "bottom-up" so that the Y_i's of a given generation (from the root) are independent. The Y_i's in (3) encode the information that i is activated by a node in the subtree of T "below" i (and not by the root). If one considers a node in the original graph and runs a directed contagion model on a local neighborhood of this node where only 'directed' contagion toward this node are allowed, then the state of the graph seen from this node is well approximated by the Y_i's.

3 General model and main results

We first present the model for the underlying graph, then the model for the diffusion process and finally our main results for the spread of the diffusion.

3.1 Graphs: the configuration model

Let $n \in \mathbb{N}$ and let $\boldsymbol{d} = (d_i^{(n)})_1^n = (d_i)_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. For notational simplicity we will usually not show the dependency on n explicitly. We define a random multigraph with given degree sequence $\boldsymbol{d}$, denoted by $G^*(n, \boldsymbol{d})$ by the configuration model [6]: take a set of d_i half-edges for each vertex i and combine the half-edges into pairs by a uniformly random matching of the set of all half-edges. Conditioned on the multigraph $G^*(n, \boldsymbol{d})$ being a simple graph, we obtain a uniformly distributed random graph with the given degree sequence, which we denote by $G(n, \boldsymbol{d})$ [18].

We will let $n \to \infty$ and assume that we are given $\boldsymbol{d} = (d_i)_1^n$ satisfying the following regularity conditions, cf Molloy and Reed [32]:

Condition 8. *For each n, $\boldsymbol{d} = (d_i^{(n)})_1^n$ is a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even and, for some probability distribution $\boldsymbol{p} = (p_r)_{r=0}^\infty$ independent of n,*

(i) $|\{i : d_i = r\}|/n \to p_r$ *for every* $r \geq 0$ *as* $n \to \infty$*;*

(ii) $\lambda := \sum_r r p_r \in (0, \infty)$*;*

(iii) $\sum_i d_i^2 = O(n)$*.*

The results of this work can be applied to some other random graphs models too by conditioning on the vertex degrees (see [19, 20]). For example, for the Erdős-Rényi graph $G(n, \lambda/n)$ with $\lambda \in (0, \infty)$, the assumptions hold with the probability distribution $\boldsymbol{p} = (p_r)$ being a Poisson random variable with mean λ.

3.2 Diffusion: percolated threshold model

In this section, we describe the diffusion for any given finite graph G with vertex set $[1, n]$. We still denote by d_i the degree of node i. From now on, a vertex i is either active or inactive. In our model, the initial set of active nodes S (the seed) will remain active during the whole process

of the diffusion. We will consider **single activation** or **degree based activation:** randomly activate vertex i with probability α_{d_i}, where d_i is the degree of the vertex. In other words, each node i draws independently of each other a Bernoulli random variable σ_i with parameter α_{d_i} and is considered as initially active if $\sigma_i = 1$ and not initially active otherwise. In the case of degree based activation, we denote $\boldsymbol{\alpha} = (\alpha_d)$ the parameters of this activation. In particular, if $\alpha_d = \alpha$ for all d, then a fraction α chosen uniformly at random among the population is activated before the diffusion takes place.

Symmetric threshold model: We first present the symmetric threshold model which generalizes the bootstrap percolation [4]: a node becomes active when a certain threshold fraction of neighbors are already active. We allow the threshold fraction to be a random variable with distribution depending on the degree of the node and such that thresholds are independent among nodes. Formally, we define for each node i a sequence of random variables in $\mathbb{N}$ denoted by $(K_i(d))_{d=1}^{\infty}$. The threshold associated to node i is $k_i = K_i(d_i)$ where d_i is the degree of node i. We assume that for any two vertices i and j, the sequences $(K_i(d))$ and $K_j(d)$ are independent and have the same law as a sequence denoted by $(K(d))$. For $\ell \leq s$, we denote $t_{s\ell} = \mathbb{P}(K(s) = \ell)$ the probability distribution of the threshold for a node of degree s. For example in the model of Morris [33] described in the introduction, we take $K(d) = \lfloor qd \rfloor$ so that $t_{s\ell} = \mathbb{1}(\lfloor qs \rfloor = \ell)$. We will use the notation $(\boldsymbol{d}, \boldsymbol{k}) = (d_i, k_i)_1^n$ and $\boldsymbol{t} = (t_{s\ell})$ denotes the distribution of thresholds.

Now the progressive dynamic of the diffusion operates as follows: some set of nodes S starts out being active; all other nodes are inactive. Time operates in discrete steps $t = 1, 2, 3, \ldots$. At a given time t, any inactive node i becomes active if its number of active neighbors is at least $K_i(d_i) + 1$. This in turn may cause other nodes to become active.

It is easy to see that the final set of active nodes (after n time steps if the network is of size n) only depends on the initial set S (and not on the order of the activations) and can be obtained as follows: set $X_i = \mathbb{1}(i \in S)$ for all i. Then as long as there exists i such that $\sum_{j \sim i} X_j > K_i(d_i)$, set $X_i = 1$, where $j \sim i$ means that i and j share an edge in G. When this algorithm finishes, the final state of node i is represented by X_i: $X_i = 1$ if node i is active and $X_i = 0$ otherwise.

Note that we allow for the possibility $d_i = K_i(d_i)$ in which case, node i is never activated unless it belongs to the initial set S. Note also that the condition $K_i(d_i) \geq 0$ is actually not restrictive. If we wish to consider a model where $K_i(d_i) < 0$ is possible, we just need to modify the initial seed S so as to put node i in S if $K_i(d_i) < 0$. Hence this model is equivalent to ours if we increase α_{d_i} accordingly.

Percolated threshold model: this model depends on a parameter $\pi \in [0, 1]$ and a distribution of random thresholds $(K(d))_{d \in \mathbb{N}}$ given by $\boldsymbol{t} = (t_{s\ell})_{\ell \leq s}$ as described above. Given any graph G and initial set S, we now proceed in two phases.

- bond percolation: randomly delete each edge with probability $1 - \pi$ independently of all other edges. We denote the resulting random graph by G_π;
- apply the symmetric threshold model with thresholds $K(d)$: set $X_i = \mathbb{1}(i \in S)$ and then as long as there is i such that $\sum_{j \sim_\pi i} X_j > K_i(d_i)$, set $X_i = 1$, where $j \sim_\pi i$ means that i and j share an edge in G_π and d_i is the degree of node i in the original graph G.

Clearly if $\pi = 1$, this is exactly the symmetric threshold model. If in addition $K(d) = k$, then this model is known as bootstrap percolation [4]. On the other hand if $\pi \in (0, 1)$ and $K(d) = 0$

for any d, then this is the contact process with probability of activation π on each edge. Note that the percolated threshold model is not equivalent to the symmetric threshold model on the (bond) percolated graph since threshold depends on the degree in the original graph.

3.3 Diffusion with degree based activation

Recall that for integers $s \geq 0$ and $0 \leq r \leq s$, b_{sr} denotes the binomial probabilities $b_{sr}(p) := \mathbb{P}(\mathrm{Bi}(s,p) = r) = \binom{s}{r} p^r (1-p)^{s-r}$.

For a graph G, let $v(G)$ and $e(G)$ denote the numbers of vertices and edges in G respectively. The subgraph of $G(n, \boldsymbol{d}, \boldsymbol{k})$ induced by the activated (resp. inactive) nodes at the end of the diffusion is denoted by $H = H(n, \boldsymbol{d}, \boldsymbol{k}, \boldsymbol{\alpha}, \pi)$ (resp. $I = I(n, \boldsymbol{d}, \boldsymbol{k}, \boldsymbol{\alpha}, \pi)$). For $r \leq s$, we denote by $v_{sr}(I)$ the number of vertices in I with degree s in G and r in I, i.e. the number of vertices with degree s in G which are not activated and with r neighbors which are not activated either. We denote by $v_s(H)$ the number of activated vertices of degree s in G (and with possibly lower degree in H).

Given a distribution of degrees $\boldsymbol{p} = (p_s)$ and of thresholds $\boldsymbol{t} = (t_{s\ell})_{\ell \leq s}$, we define the functions:

$$\begin{aligned}
h(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &:= \sum_s (1-\alpha_s) p_s \sum_{\ell \leq s} t_{s\ell} \sum_{r \geq s-\ell} r b_{sr}(1 - \pi + \pi z), \\
g(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &:= \lambda(\boldsymbol{p}) z (1 - \pi + \pi z) - h(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi), \\
h_1(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &:= \sum_s (1-\alpha_s) p_s \sum_{\ell \leq s} t_{s\ell} \sum_{r \geq s-\ell} b_{sr}(1 - \pi + \pi z),
\end{aligned}$$

where $\lambda(\boldsymbol{p}) = \sum_s s p_s$. We define

$$\hat{z} = \hat{z}(\boldsymbol{\alpha}, \boldsymbol{p}, \pi) := \max \{ z \in [0,1] : \, g(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) = 0 \}. \tag{5}$$

We refer to Section 7.2 for a justification of the use of the max in (5).

Theorem 9. *Consider the random graph $G(n, \boldsymbol{d}, \boldsymbol{k})$ for a sequence $\boldsymbol{d} = (d_i)_1^n$ satisfying Condition 8 and let $\boldsymbol{t}$ be a distribution of thresholds as defined above. For the percolated threshold diffusion on the graph $G(n, \boldsymbol{d}, \boldsymbol{k})$, we have: if $\hat{z} = 0$, or if $\hat{z} \in (0,1]$, and further $g(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) < 0$ for any z in some interval $(\hat{z} - \epsilon, \hat{z})$, then*

$$\begin{aligned}
v(H)/n &\xrightarrow{p} 1 - h_1(\hat{z}, \boldsymbol{\alpha}, \boldsymbol{p}, \pi), \\
v_{sr}(I)/n &\xrightarrow{p} \sum_{i + \ell \geq s - r} (1-\alpha_s) p_s t_{s\ell} b_{sr}(\hat{z}) b_{s-r,i}(1-\pi), \\
v_s(H)/n &\xrightarrow{p} p_s - \sum_{k \geq s-\ell} (1-\alpha_s) p_s t_{s\ell} b_{sk}(1 - \pi + \pi\hat{z}), \\
e(I)/n &\xrightarrow{p} \left(\mathbb{1}(\pi \neq 1) \frac{\hat{z}}{2(1 - \pi + \pi\hat{z})} + \mathbb{1}(\pi = 1) \frac{1}{2} \right) h(\hat{z}; \boldsymbol{\alpha}, \boldsymbol{p}, \pi).
\end{aligned}$$

If we condition the induced graph $I^* = I^*(n, \boldsymbol{d}, \boldsymbol{k}, \boldsymbol{\alpha}, \pi)$ of inactive nodes in $G^*(n, \boldsymbol{d}, \boldsymbol{k})$ on its degree sequence $\boldsymbol{d}^{I^*}$ and let n^{I^*} be the number of its vertices, then I^* has the distribution of $G^*(n^{I^*}, \boldsymbol{d}^{I^*})$.

The proof of this theorem is given in Section 4. Note that Proposition 3 follows easily.

3.4 Diffusion with a single activation

In this section, we look at the diffusion with one (or a small number $o(n)$ of) initial active node(s) in $G(n, \boldsymbol{d}, \boldsymbol{k})$. For simplicity, we make the following additional assumption:

Condition 10. *(iv)* $\sum_i d_i^3 = O(n)$.

For $u \in [1, n]$, let $\mathcal{C}(u)$ (resp. $\mathcal{C}(1, \dots, k)$) be the subgraph induced by the final active nodes with initial active node u (resp. initial active nodes $1, \dots, k$). We also define $\mathcal{I}(u)$ as the subgraph induced by the inactive nodes with initial active node u. The set of vertices of $\mathcal{C}(u)$ and $\mathcal{I}(u)$ is a partition of the vertices of the original graph. To ease notation, we denote $h(z) = h(z; 0, \boldsymbol{p}, \pi)$, $g(z) = g(z; 0, \boldsymbol{p}, \pi)$ and $h_1(z) = h_1(z; 0, \boldsymbol{p}, \pi)$. We define

$$\xi := \sup \{z \in [0, 1) : \ g(z) = 0\}. \tag{6}$$

We also define

$$\begin{aligned} \bar{g}(z) &= (1 - \pi + \pi z) \left(\lambda z - \sum_s s p_s (1 - t_{s0}) - \sum_s s p_s t_{s0} (1 - \pi + \pi z)^{s-1} \right), \\ \bar{h}_1(z) &= \sum_s p_s t_{s0} (1 - \pi + \pi z)^s + \sum_s p_s (1 - t_{s0}), \end{aligned}$$

and $\overline{\xi} = \sup \{z \in [0, 1) : \ \bar{g}(z) = 0\}$.

We call the following condition the cascade condition:

$$\pi \sum_r r(r-1) r_r t_{r0} > \sum_r r p_r, \tag{7}$$

which can be rewritten as $\pi \mathbb{E}[D(D-1)\mathbb{1}(K(D) = 0)] > \mathbb{E}[D]$ where D is a random variable with distribution $\boldsymbol{p}$.

We denote by $P = P(n, \boldsymbol{d}, \boldsymbol{k})$ the largest connected component of the graph $G(n, \boldsymbol{d}, \boldsymbol{k})$ on which we apply a bond percolation with parameter π (i.e. we remove each edge independently with probability $1 - \pi$) and then apply a site percolation by removing all vertices with $k_i \geq 1$. The vertices of the connected graph P are called pivotal vertices: for any $u \in P$, we have $P \subset \mathcal{C}(u)$.

Theorem 11. *Consider the random graph $G(n, \boldsymbol{d}, \boldsymbol{k})$ for a sequence $\boldsymbol{d} = (d_i)_1^n$ satisfying Conditions 8 and 10 and let $\boldsymbol{t}$ be a distribution of thresholds as defined above.*

(i) If the cascade condition (7) is satisfied, then

$$\lim_n \frac{v(P)}{n} = 1 - \bar{h}_1(\overline{\xi}) > 0.$$

Moreover, for any $u \in P$, we have w.h.p.

$$\liminf_n \frac{v(\mathcal{C}(u))}{n} = \liminf \frac{v(\cap_{u \in P} \mathcal{C}(u))}{n} \geq 1 - h_1(\xi) > 0,$$

where ξ is defined by (6). Moreover if $\xi = 0$ or ξ is such that there exists $\epsilon > 0$ with $g(z) < 0$ for $z \in (\xi - \epsilon, \xi)$, then we have for any $u \in P$:

$$\begin{aligned} v(\mathcal{C}(u))/n &\xrightarrow{p} 1 - h_1(\xi), \\ v_{sr}(\mathcal{I}(u))/n &\xrightarrow{p} \sum_{i+\ell \geq s-r} p_s t_{s\ell} b_{sr}(\xi) b_{s-r,i}(1-\pi), \\ v_s(\mathcal{C}(u))/n &\xrightarrow{p} p_s - \sum_{k \geq s-\ell} p_s t_{s\ell} b_{sk}(1 - \pi + \pi\xi), \\ e(\mathcal{I}(u))/n &\xrightarrow{p} \left(\mathbb{1}(\pi \neq 1) \frac{\xi}{2(1-\pi+\pi\xi)} + \mathbb{1}(\pi = 1)\frac{1}{2} \right) h(\xi). \end{aligned}$$

If we condition the induced graph $\mathcal{I}^*(u)$ of inactive nodes in $G^*(n, \boldsymbol{d}, \boldsymbol{k})$ on its degree sequence $\boldsymbol{d}^{\mathcal{I}^*(u)}$ and let $n^{\mathcal{I}^*(u)}$ be the number of its vertices, then $\mathcal{I}^*(u)$ has the distribution of $G^*(n^{\mathcal{I}^*(u)}, \boldsymbol{d}^{\mathcal{I}^*(u)})$.

(ii) If $\pi \sum_r r(r-1) r_r t_{r0} < \sum_r r p_r$, then for any $k = o(n)$, $v(\mathcal{C}(1, \ldots, k)) = o_p(n)$.

A proof of this theorem is given in Section 5. We end this section with some remarks: if $K(d)$ does not depend on d, then the cascade condition becomes with D a random variable with distribution $\boldsymbol{p}$:

$$\pi \mathbb{P}(K = 0) > \frac{\mathbb{E}[D]}{\mathbb{E}[D(D-1)]}.$$

In particular, if $K = 0$, then we find the well-known condition for the existence of a 'giant component'. This corresponds to existing results in the literature see in particular Theorem 3.9 in [17] which extend the standard result of Molloy and Reed [32]. More generally, in the case $\pi \in [0, 1]$ and $K = 0$ (corresponding to the contact process), a simple computation shows that

$$\begin{aligned} h(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &= (1-\alpha)(1 - \pi + \pi z) \phi_D'(1 - \pi + \pi z) \\ h_1(z; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) &= (1-\alpha) \phi_D(1 - \pi + \pi z), \end{aligned}$$

where $\phi_D(x) = \mathbb{E}[x^D]$ is the generating function of the asymptotic degree distribution. Applying Theorems 9 and 11 allow to obtain results for the contact process. Similarly, the bootstrap percolation has been studied in random regular graphs [4] and random graphs with given vertex degrees [1]. The bootstrap percolation corresponds to the particular case of the percolated threshold model with $\pi = 1$ and $K(d) = \theta \geq 0$ and our Theorems 9 and 11 allow to recover results for the size of the diffusion. Finally, the case where $K(d) = qd$ and $\pi = 1$ implies directly Proposition 1.

4 Proof of Theorem 9

4.1 Sketch of the proof

It is well-known that it is often simpler to study the random multigraph $G^*(n, \boldsymbol{d})$ with given vertex sequence $\boldsymbol{d} = (d_i)_1^n$ defined in Section 3.1. We consider asymptotics as the number of vertices tends to infinity and thus assume throughout the paper that we are given, for each n, a

sequence $\boldsymbol{d} = (d_i^{(n)})_1^n$ with $\sum_i d_i^{(n)}$ even. We may obtain $G(n, \boldsymbol{d})$ by conditioning the multigraph $G^*(n, \boldsymbol{d})$ on being a simple graph. By [18], we know that the condition $\sum_i d_i^2 = O(n)$ implies $\liminf \mathbb{P}(G^*(n, \boldsymbol{d}) \text{ is simple}) > 0$. In this case, many results transfer immediately from $G^*(n, \boldsymbol{d})$ to $G(n, \boldsymbol{d})$, for example, every result of the type $\mathbb{P}(\mathcal{E}_n) \to 0$ for some events $\mathcal{E}_n$, and thus every result saying that some parameter converges in probability to some non-random value. This includes every results in the present paper. Henceforth, we will in this paper study the random multigraph $G^*(n, \boldsymbol{d})$ and in a last step (left to the reader) transfer the results to $G(n, \boldsymbol{d})$ by conditioning. More precisely, we define the graph $G^*(n, \boldsymbol{d}, \boldsymbol{k})$, i.e. we add a label to each vertex corresponding to its threshold and say that it is the configuration model associated to the degree-threshold sequence $(\boldsymbol{d}, \boldsymbol{k})$ and asymptotic degree and threshold distributions $\boldsymbol{p} = (p_s)$ and $\boldsymbol{t} = (t_{s\ell})_{s\ell}$.

We run the dynamic of the diffusion of Section 3.2 on a general graph $G^*(n, \boldsymbol{d}, \boldsymbol{k})$ in order to compute the final size of the diffusion in a similar way as in [19]. The main point here consists in coupling the construction of the graph with the dynamic of the diffusion. This is done in Section 4.3. The proof of Theorem 9 follows then easily. In order to prove Theorem 11, we use the same idea of coupling (in a similar spirit as in [20] for the analysis of the giant component) but we have to deal with an additional difficulty due to the following lack of symmetry: if $\mathcal{C}(u)$ is the final set of the diffusion with only u as initial active node, then for any $v \in \mathcal{C}(u)$, we do not have in general $\mathcal{C}(u) = \mathcal{C}(v)$. We take care of this difficulty in Section 5. In the next section, we present a preliminary lemma that will be used in the proofs.

4.2 A Lemma for death processes

A pure death process with rate 1 is a process that starts with some number of balls whose lifetime are i.i.d. rate 1 exponentials. Now consider n bins with independent rate 1 death processes. To each bin, we attach a couple (d_i, k_i) where d_i is the number of balls at time 0 in the bin and $k_i = K_i(d_i)$ is the threshold corresponding to the bin. We now modify the death process as follows: all balls are initially white. For any living ball, when it dies, with probability $1 - \pi$ and independently of everything else, we color it green instead of removing it from the bin. Let $W_j^{(n)}(t)$ and $G_j^{(n)}(t)$ denote the number of white and green balls respectively in bin j at time t, where $j = 1, \dots, n$ and $t \geq 0$.

Let $U_{sri,\ell}^{(n)}(t)$ be the number of bins that have s balls at time 0 and r white balls, i green balls at time t and threshold ℓ, i.e. $U_{sri,\ell}^{(n)}(t) = |\{j \in [1, n],\, W_j^{(n)}(t) = r,\, G_j^{(n)}(t) = i,\, d_j = s,\, k_j = \ell\}|$. In what follows we suppress the superscripts to lighten the notation. The following lemma is an extension of Lemma 4.4 in [19]:

Lemma 12. *Consider the n independent processes defined above and assume that the sequence $(d_i)_1^n$ satisfies Condition 8 where $(p_r)_{r=0}^\infty$ can be a defective probability distribution: $\sum_r p_r \leq 1$. We assume that the distribution of the thresholds is $\boldsymbol{t} = (t_{s\ell})$. Then, with the above notation, as $n \to \infty$,*

$$\sup_{t \geq 0} \sum_{s,\ell} \sum_{r=0}^{s} \sum_{i=0}^{s-r} r \left| U_{sri,\ell}(t)/n - p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi) \right| \xrightarrow{p} 0. \tag{8}$$

Proof. Let $n_{s\ell} = |\{i : d_i = s, k_i = \ell\}|$. In particular $\sum_{r,i} U_{sri,\ell}^{(n)}(0) = U_{ss0,\ell}^{(n)}(0) = n_{s\ell}$. First

fix integers s, ℓ and j with $0 \le j \le s$. Consider the $n_{s\ell}$ bins that start with s balls and with threshold ℓ. For $k = 1, \dots, n_{s\ell}$, let T_k be the time of death or recoloring of the jth ball in the kth such bin. Then $|\{k : T_k \le t\}| = \sum_{r=0}^{s-j} \sum_{i=0}^{s-r} U_{sri,\ell}(t)$. Moreover, for the kth bin, we have $\mathbb{P}(T_k \le t) = \sum_{r=0}^{s-j} b_{sr}(e^{-t})$. Multiplying by $n_{s\ell}/n$ and using Glivenko-Cantelli theorem (see e.g. Proposition 4.24 in [22]), we have

$$\sup_{t \ge 0} \left| \frac{1}{n} \sum_{r=0}^{s-j} \sum_{i=0}^{s-r} U_{sri,\ell}(t) - \frac{n_{s\ell}}{n} \sum_{r=0}^{s-j} b_{sr}(e^{-t}) \right| \xrightarrow{p} 0.$$

Since $n_{s\ell}/n \to^p p_s t_{s\ell}$ (by Condition 8(i) and the law of large numbers), we see that

$$\sup_{t \ge 0} \left| \frac{1}{n} \sum_{i=0}^{s-r} U_{sri,\ell}(t) - p_s t_{s\ell} b_{sr}(e^{-t}) \right| \xrightarrow{p} 0$$

The law of $G_k(t)$ given $d_k = s$ and $W_k(t) = r \le s$ is a Binomial distribution with parameters $s - r$ and $1 - \pi$. Hence by the law of large numbers, we have

$$\sup_{t \ge 0} \left| \frac{U_{sri,\ell}(t)}{n} - p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi) \right| \xrightarrow{p} 0.$$

Hence each term in (8) tends to 0 in probability. Hence the same holds for any finite partial sum. Let $\epsilon > 0$ and S be such that $\sum_{s=S}^{\infty} s p_s < \epsilon$. By Condition 8(iii), we have $\sum_{s,\ell} s n_{s\ell}/n = \sum_s d_s/n \to \lambda = \sum_s s p_s$. Hence, also $\sum_{s \ge S} \sum_\ell s n_{s\ell}/n \to \sum_{s=S}^{\infty} s p_s < \epsilon$. So that for sufficiently large n, we get $\sum_{s \ge S} \sum_\ell s n_{s\ell}/n < \epsilon$ and

$$\begin{aligned}
&\sup_{t \ge 0} \sum_{s \ge S} \sum_{\ell} \sum_{r=0}^{s} \sum_{i=0}^{s-r} r \left| U_{sri,\ell}(t)/n - p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi) \right| \\
\le &\sup_{t \ge 0} \sum_{s \ge S} \sum_{\ell} \sum_{r=0}^{s} \sum_{i=0}^{s-r} r \left(U_{sri,\ell}(t)/n + p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi) \right) \\
&\le \sum_{s \ge S} \sum_{\ell} s(n_{s\ell}/n + p_s t_{s\ell}) < 2\epsilon.
\end{aligned}$$

Hence (8) holds. □

4.3 Proof of the diffusion spread

Our proof of Theorem 9 is an adaptation of the coupling argument in [19]. We start by analyzing the symmetric threshold model. We can view the algorithm of Section 3.2 as follows: start with the graph G and remove vertices from S. As a result, if vertex i has not been removed, its degree has been lowered. We denote by d_i^A the degree of i in the evolving graph. Then recursively remove vertices i such that $d_i^A < d_i - K_i(d_i)$. All removed vertices at the end of this procedure are active and all vertices left are inactive. It is easily seen that we obtain the same result by removing edges where one endpoint satisfy $d_i^A < d_i - K_i(d_i)$, until no such edge remains, and finally removing all isolated vertices, which correspond to inactive nodes.

Regard each edge as consisting of two half-edges, each half-edge having one endpoint. We introduce types of vertices. We set the type of vertices in the seed S to B. Say that a vertex (not

in S) is of type A if $d_i^A \geq d_i - K_i(d_i)$ and of type B otherwise. In particular at the beginning of the process, all vertices not in S are of type A since $d_i^A = d_i$ and all vertices in S are by definition of type B. As the algorithm evolves, d_i^A decreases so that some type A vertices become of type B during the execution of the algorithm. Similarly, say that a half-edge is of type A or B when its endpoint is. As long as there is any half-edge of type B, choose one such half-edge uniformly at random and remove the edge it belongs to. This may change the other endpoint from A to B (by decreasing d^A) and thus create new half-edges of type B. When there are no half-edges of type B left, we stop. Then the final set of active nodes is the set of vertices of type B (which are all isolated).

As in [19], we regard vertices as bins and half-edges as balls. At each step, we remove first one random ball from the set of balls in B-bins and then a random ball without restriction. We stop when there are no non-empty B-bins. We thus alternately remove a random B-ball and a random ball. We may just as well say that we first remove a random B-ball. We then remove balls in pairs, first a random ball and then a random B-ball, and stop with the random ball, leaving no B-ball to remove. We change the description a little by introducing colors. Initially all balls are white, and we begin again by removing one random B-ball. Subsequently, in each deletion step we first remove a random white ball and then recolor a random white B-ball red; this is repeated until no more white B-balls remain.

We now run this deletion process in continuous time such that, if there are j white balls remaining, then we wait an exponential time with mean $1/j$ until the next pair of deletions. In other words, we make deletions at rate j. This means that each white ball is deleted with rate 1 and that, when we delete a white ball, we also color a random white B-ball red. Let $A(t)$ and $B(t)$ denote the numbers of white A-balls and white B-balls at time t, respectively, and $A_1(t)$ denotes the number of A-bins at time t. Since red balls are ignored, we may make a final change of rules, and say that all balls are removed at rate 1 and that, when a white ball is removed, a random white B-ball is colored red; we stop when we should recolor a white B-ball but there is no such ball.

Let τ be the stopping time of this process. First consider the white balls only. There are no white B-balls left at τ, so $B(\tau)$ has reached zero. However, let us consider the last deletion and recoloring step as completed by redefining $B(\tau) = -1$; we then see that τ is characterized by $B(\tau) = -1$ and $B(t) \geq 0$ for $0 \leq t \leq \tau$. Moreover, the A-balls left at τ (which are all white) are exactly the half-edges in the induced subgraph I of inactive nodes. Hence, the number of edges in this subgraph is $\frac{1}{2}A(\tau)$, while the number of nodes not activated is $A_1(\tau)$.

Moreover, if we consider only the total number $A(t)+B(t)$ of white balls in the bins, ignoring the types, the process is as follows: each ball dies at rate 1 and upon its death another ball is also sacrificed. The process $A(t) + B(t)$ is a death process with rate 2 (up to time τ). Consequently, by Lemma 4.3 of [19] (or Lemma 12 above), we have

$$\sup_{t\leq\tau} \left| A(t) + B(t) - n\lambda e^{-2t} \right| = o_p(n), \tag{9}$$

since Condition 8 (iii) implies $\sum_r r|\{i:\ d_i = r\}|/n \to \lambda$.

Now if we consider the final version of the process restricted to A-bins, it corresponds exactly to the death process studied in Section 4.2 above with $\pi = 1$. We need only to compute the initial condition for this process. For a degree based activation, each vertex of degree s is activated (i.e. the corresponding bin becomes a B-bin) with probability α_s. Hence by the law of large numbers,

the number of A-bins with initially s balls and threshold ℓ is $n_{s\ell} = (1-\alpha_s)p_s t_{s\ell} n + o_p(n)$. With the notation of Lemma 12, we have

$$A(t) = \sum_{s\geq 1,\, r\geq s-\ell} rU_{sr0,\ell}(t), \quad \text{and,} \quad A_1(t) = \sum_{s\geq 1,\, r\geq s-\ell} U_{sr0,\ell}(t),$$

with the defective probability distribution $(1-\alpha_s)p_s$. Hence by Lemma 12 we get (recall that $\pi = 1$ here):

$$\sup_{t\leq\tau}\left|\frac{A(t)}{n} - \sum_{s\geq 1, r\geq s-\ell} r(1-\alpha_s)p_s t_{s\ell} b_{sr}(e^{-t})\right| \xrightarrow{p} 0.$$

It is then easy to finish the proof as in [19] for this model. In particular, it ends the proof of Theorem 9 for the case $\pi = 1$.

We now consider the percolated threshold model with $\pi < 1$. We modify the process as follows: for any white A-ball when it dies, with probability $1-\pi$, we color it green instead of removing it. A bin is of type A if $r + i \geq s - \ell$, where r is the number of withe balls in the bin, i the number of green balls (which did not transmit the diffusion) and s and ℓ are the initial degree and threshold. Let $A(t)$ be the number of white A-balls. By Lemma 12, we now have:

$$\sup_{t\leq\tau}\left|\frac{A(t)}{n} - \sum_{s, r+i\geq s-\ell} r(1-\alpha_s)p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi)\right| \xrightarrow{p} 0,$$

$$\sup_{t\leq\tau}\left|\frac{A_1(t)}{n} - \sum_{s, r+i\geq s-\ell} (1-\alpha_s)p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi)\right| \xrightarrow{p} 0.$$

In particular, we have thanks to Lemma 14 (in Section 7.1) for $t \leq \tau$,

$$A(t)/n = \frac{e^{-t}}{1-\pi+\pi e^{-t}} h(e^{-t}; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) + o_p(n), \quad A_1(t)/n = h_1(e^{-t}; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) + o_p(n). \tag{10}$$

By looking at white balls (without taking types in consideration), we see that Equation (9) is still valid. Hence, we have

$$\sup_{t\leq\tau}\left|\frac{B(t)}{n} - \frac{e^{-t}}{1-\pi+\pi e^{-t}} g(e^{-t}; \boldsymbol{\alpha}, \boldsymbol{p}, \pi)\right| \xrightarrow{p} 0. \tag{11}$$

Assume now that $t_1 > 0$ is a constant independent of n with $t_1 < -\ln \hat{z}$ so that $\hat{z} < 1$ and $g(1; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) = g(1) > 0$. Hence, we have $g(z) > 0$ for $z \in (\hat{z}, 1]$ and thus $g(e^{-t}) > 0$ for $t \leq t_1$. We can find some $c > 0$ such that $g(e^{-t}) \geq c$ for $t \leq t_1$. But $B(\tau) = -1$, so if $\tau \leq t_1$ then $\frac{e^{-\tau}}{1-\pi+\pi e^{-\tau}} g(e^{-\tau}) - B(\tau)/n > c\frac{e^{-t_1}}{1-\pi+\pi e^{-t_1}}$ and from (11), we have $\mathbb{P}(\tau \leq t_1) \to 0$. In case $\hat{z} = 0$, we may take any finite t_1 and find $\tau \to^p \infty$ and (10) with $t \to \infty$, yields that

$$\lim_{t\to\infty} A(t) = o_p(n), \quad \lim_{t\to\infty} A_1(t) = nh_1(0; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) + o_p(n).$$

In case $\hat{z} > 0$, by the hypothesis we can find $t_2 \in (-\ln \hat{z}, -\ln(\hat{z} - \epsilon))$ such that $g(e^{-t_2}) = -c < 0$. If $\tau > t_2$, then $B(t_2) \geq 0$ and thus $B(t_2)/n - \frac{e^{-t_2}}{1-\pi+\pi e^{-t_2}} g(e^{-t_2}) \geq c\frac{e^{-t_2}}{1-\pi+\pi e^{-t_2}}$. Hence by (11),

we have $\mathbb{P}(\tau \geq t_2) \to 0$. Since we can choose t_1 and t_2 arbitrarily close to $-\ln \hat{z}$, we have $\tau \to^p -\ln \hat{z}$ and (10) with $t = \tau$ yields that

$$A(\tau) = \frac{n\hat{z}}{1-\pi+\pi\hat{z}} h(\hat{z}; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) + o_p(n), \quad A_1(\tau) = n h_1(\hat{z}; \boldsymbol{\alpha}, \boldsymbol{p}, \pi) + o_p(n).$$

Note that $v(H) = n - A_1(\tau)$ and $e(I) = \frac{1}{2}A(\tau)$. Hence we proved Theorem 9 for $v(H)/n$ and $e(I)/n$. The results for $v_{sr}(I)/n$ and $v_s(H)/n$ follows from the same argument, once we note that

$$v_{sr}(I) = \sum_{r+i \geq s-\ell} U_{sri,\ell}(\tau) \text{ and, } v_s(H) = |\{j : d_j = s\}| - \sum_{r+i \geq s-\ell} U_{sri,\ell}(\tau).$$

Finally, the statement concerning the distribution of the induced subgraph I follows from the fact that this subgraph has not been explored when previous algorithm stops.

5 Proof of Theorem 11

We start this section with some simple calculations. We define for $z \in [0,1]$,

$$\begin{aligned}
a(z) &= \sum_s p_s \sum_{\ell \leq s} t_{s\ell} \sum_{r \geq s-\ell} r b_{sr}(z), \\
h(z) &= a(1-\pi+\pi z), \\
h_1(z) &= \sum_{s, r \geq s-\ell} p_s t_{s\ell} b_{sr}(1-\pi+\pi z), \\
h_2(z) &= \frac{z}{1-\pi+\pi z} h(z) \text{ if } \pi < 1 \text{ and } h_2(z) = h(z) \text{ otherwise}, \\
g(z) &= \lambda z(1-\pi+\pi z) - h(z), \\
f(z) &= \frac{z}{1-\pi+\pi z} g(z) \text{ if } \pi < 1 \text{ and } f(z) = g(z) \text{ otherwise}.
\end{aligned}$$

For $s \geq 1$ and $\ell \geq 1$, we have

$$\begin{aligned}
\frac{d}{dz} \sum_{r \geq s-\ell} r b_{sr}(z) &= s^2 z^{s-1} + \sum_{s-1 \geq r \geq s-\ell} r(r-sz) \binom{s}{r} z^{r-1}(1-z)^{s-r-1} \\
&= s^2 z^{s-1} - \sum_{s-1 \geq r \geq s-\ell} r(sz-r) \frac{s}{z(s-r)} b_{s-1r}(z),
\end{aligned}$$

so that for $z \in [0,1]$, we have $\left| \frac{d}{dz} \sum_{r \geq s-\ell} r b_{sr}(z) \right| \leq s^2 + s^3$. Hence by Condition 10, a is differentiable on $[0,1]$ and we have

$$a'(z) = \sum_{s,\ell} p_s t_{s\ell} \left(s^2 z^{s-1} - \sum_{s-1 \geq r \geq s-\ell} r(sz-r) \binom{s}{r} z^{r-1}(1-z)^{s-1-r} \right).$$

In particular, we have

$$\begin{aligned}
a'(1) &= \sum_{s,\ell} p_s t_{s\ell} \left(s^2 - \mathbf{1}(\ell \geq 1) s(s-1) \right) \\
&= \sum_s p_s t_{s0} s(s-1) + \lambda,
\end{aligned}$$

so that we have

$$\begin{aligned} g'(1) &= \lambda(1+\pi) - \pi a'(1) \\ &= \mathbb{E}[D] - \pi \mathbb{E}\left[D(D-1)\mathbb{1}(K(D)=0)\right]. \end{aligned}$$

Note also that $f(1) = g(1) = 0$ and $f'(1) = g'(1)$.

Consider now the case (ii) where $\pi\mathbb{E}\left[D(D-1)\mathbb{1}(K(D)=0)\right] < \mathbb{E}[D]$, so that $g'(1) = f'(1) > 0$. The proof for an upper bound on $n^{-1}v(\mathcal{C}(1,\ldots,k))$ follows easily from Theorem 9. Take a parameter $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathbb{N}}$ with $\alpha_d = \alpha > 0$ for all d. Clearly the final set of active nodes $H(\alpha)$ will be greater than for any seed with size $o_p(n)$. Now when α goes to zero, the fact that $f'(1) < 0$ ensures that $\hat{z}(\alpha) \to 1$ in Theorem 9 so that $\lim_{\alpha\to 0}\lim_{n\to\infty} v(H(\alpha))/n = 0$. Hence point (ii) in Theorem 11 follows.

We now concentrate on the case where the cascade condition holds. In particular we have $g'(1) < 0$ so that ξ defined in (6) by $\xi = \sup\{z \in [0,1), g(z) = 0\}$ is strictly less than one and we have

$$f(z) > 0,\ \forall z \in (\xi, 1). \tag{12}$$

Also as soon as there exists $\epsilon > 0$ such that $g(z) < 0$ for $z \in (\xi - \epsilon, \xi)$, we can use the same argument as above. Since, we have in this case $\hat{z}(\alpha) \to \xi$ as $\alpha \to 0$, it gives an upper bound that matches with the statement (i) of Theorem 11. In order to prove a lower bound, we follow the general approach of [20]. We modify the algorithm defined in Section 4.3 as follows: the initial set S now contains only one vertex. When there is no half-edge (or ball) of type B, we say that we make an exception and we select a vertex (or a bin) of type A uniformly at random among all vertices of type A. We declare its white half-edges of type B and remove its other half-edges, i.e. remove the green balls contained in the corresponding bin if there are any. Exceptions are done instantaneously.

For any set of nodes $v_1,\ldots,v_k$, let $\mathcal{C}(v_1,\ldots,v_k)$ be the subgraph induced by the final active nodes with initial active nodes $v_1,\ldots,v_k$. If $S = \{v_1\}$, then clearly when the algorithm has exhausted the half-edges of type B, it removed the subgraph $\mathcal{C}(v_1)$ from the graph and all edges with one endpoint in $\mathcal{C}(v_1)$. Then an exception is made by selecting a vertex say v_2 in $G\backslash\mathcal{C}(v_1)$. Similarly when the algorithm exhausted the half-edges of type B, it removed the subgraph $\mathcal{C}(v_1,v_2)$ and all edges with one endpoint in this set of vertices. More generally, if k exceptions are made consisting of selecting nodes $v_1,\ldots v_k$, then before the $k+1$-th exception is made (or at termination of the algorithm if there is no more exception made), the algorithm removed the subgraph $\mathcal{C}(v_1,\ldots,v_k)$ and all edges with one endpoint in this set of vertices.

We use the same notation as in Section 4.3. In particular, we still have:

$$\sup_{t\geq 0}\left|A(t) + B(t) - n\lambda e^{-2t}\right| = o_p(n). \tag{13}$$

We now ignore the effect of the exceptions by letting $\tilde{A}(t)$ be the number of white A balls if no exceptions were made, i.e. assuming $B(t) > 0$ for all t. If $d_{\max} = \max_i d_i$ is the maximum degree of $G^*(n,(d_i)_1^n)$, then we have $\tilde{A}(0) = A(0) \in [n - d_{\max}, n]$. By Condition 8 (iii), $d_{\max} = O(n^{1/2})$, and thus $n^{-1}d_{\max} = o_p(n)$. Hence we can apply results of previous section:

$$\sup_{t\geq 0}\left|\frac{\tilde{A}(t)}{n} - \sum_{s,r+i\geq s-\ell} r p_s t_{s\ell} b_{sr}(e^{-t}) b_{s-r,i}(1-\pi)\right| \xrightarrow{p} 0. \tag{14}$$

We now prove that:

$$0 \leq \tilde{A}(t) - A(t) < \sup_{0 \leq s \leq t} (\tilde{A}(s) - A(s) - B(s)) + d_{\max}. \tag{15}$$

The fact that $\tilde{A}(t) \geq A(t)$ is clear. Furthermore, $\tilde{A}(t) - A(t)$ increases only when exceptions are made. If an exception is made at time t, then the process B reached zero and a vertex with j white half-edges is selected so that $B(t) = j - 1 < d_{\max}$. Hence we have

$$\tilde{A}(t) - A(t) < \tilde{A}(t) - A(t) - B(t) + d_{\max}.$$

Between exceptions, if A decreases by one then so does $\tilde{A}$, hence $\tilde{A}(t) - A(t)$ does not increase. Consequently if s was the last time before t that an exception was performed, then $\tilde{A}(t) - A(t) \leq \tilde{A}(s) - A(s)$ and (15) follows.

Let $\tilde{B}(t) = A(t) + B(t) - \tilde{A}(t)$, then we have

$$\sup_{t \geq 0} \left| \frac{\tilde{B}(t)}{n} - f(e^{-t}) \right| \xrightarrow{p} 0. \tag{16}$$

Equation (15) can be written as

$$0 \leq \tilde{A}(t) - A(t) < -\inf_{s \leq t} \tilde{B}(s) + d_{\max} \tag{17}$$

We first assume that ξ given by (6) is such that $\xi > 0$ and there exists $\epsilon > 0$ such that $g(z) < 0$ and hence $f(z) < 0$ for $z \in (\xi - \epsilon, \xi)$. Let $\tau = -\ln \xi$. Then by (12), we have $f(e^{-t}) > 0$ for $0 < t < \tau$ so that $\inf_{t \leq \tau} f(e^{-t}) = f(1) = 0$ and hence by (16),

$$\inf_{t \leq \tau} n^{-1} \tilde{B}(t) \xrightarrow{p} 0. \tag{18}$$

By Condition 8 (iii), $n^{-1} d_{\max} = o_p(n)$. Consequently, (17) yields

$$\sup_{t \leq \tau} n^{-1} |\tilde{B}(t) - B(t)| = \sup_{t \leq \tau} n^{-1} |\tilde{A}(t) - A(t)| \xrightarrow{p} 0, \tag{19}$$

and thus by (16)

$$\sup_{t \leq \tau} \left| \frac{B(t)}{n} - f(e^{-t}) \right| \xrightarrow{p} 0. \tag{20}$$

By assumption, there exists $\epsilon > 0$ sufficiently small for $f(e^{-\tau-\epsilon}) < 0$. Since $f(e^{-t}) > 0$ on the interval $[\epsilon, \tau - \epsilon]$, (20) implies that w.h.p. $B(t)$ remains positive on $[\epsilon, \tau - \epsilon]$, and thus no exception is made during this interval.

On the other hand, $f(e^{-\tau-\epsilon}) < 0$ and (16) implies $n^{-1}\tilde{B}(\tau + \epsilon) = f(e^{-\tau-\epsilon}) + o_p(n)$, while $B(\tau + \epsilon) \geq 0$. Thus with $\delta = -f(e^{-\tau-\epsilon})/2 > 0$, w.h.p.

$$\tilde{A}(\tau + \epsilon) - A(\tau + \epsilon) = B(\tau + \epsilon) - \tilde{B}(\tau + \epsilon) \geq -\tilde{B}(\tau + \epsilon) > n\delta, \tag{21}$$

while (19) yields $\tilde{A}(\tau) - A(\tau) < n\delta$ w.h.p. Consequently, w.h.p. $\tilde{A}(\tau+\epsilon) - A(\tau+\epsilon) > \tilde{A}(\tau) - A(\tau)$ and an exception is performed between τ and $\tau + \epsilon$.

Let T_1 be the last time an exception was performed before $\tau/2$ and let T_2 be the next time it is performed. We have shown that for any $\epsilon > 0$, w.h.p. $0 \leq T_1 \leq \epsilon$ and $\tau - \epsilon \leq T_2 \leq \tau + \epsilon$.

Lemma 13. *Let T_1^* and T_2^* be two random times when an exception is performed, with $T_1^* \leq T_2^*$, and assume that $T_1^* \xrightarrow{p} t_1$ and $T_2^* \xrightarrow{p} t_2$, where $0 \leq t_1 \leq t_2 \leq \tau$. If $v(T_i^*)$ is the number of vertices removed by the algorithm by time T_i^*, then*

$$\frac{v(T_2^*) - v(T_1^*)}{n} \quad \xrightarrow{p} \quad h_1(e^{-t_1}) - h_1(e^{-t_2}). \tag{22}$$

In particular, if $t_1 = t_2$, then $v(T_2^*) - v(T_1^*) = o_p(n)$.

Proof. By definition, we have:

$$v(T_2^*) - v(T_1^*) = A_1(T_1^*-) - A_1(T_2^*-).$$

Since $T_2^* \xrightarrow{p} t_2 \leq \tau$ and f is continuous, $\inf_{t \leq T_2^*} f(e^{-t}) \xrightarrow{p} \inf_{t \leq t_2} f(e^{-t}) = 0$, and (16) and (17) imply, in analogy with (18) and (19), $n^{-1} \inf_{t \leq T_2^*} \tilde{B}(t) \xrightarrow{p} 0$ and

$$\sup_{t \leq T_2^*} n^{-1} |\tilde{A}(t) - A(t)| \xrightarrow{p} 0. \tag{23}$$

Let $\tilde{A}_1(t)$ be the number of bins if no exceptions were made. Then we clearly have $\tilde{A}_1(t) - A_1(t) \leq \tilde{A}(t) - A(t)$ since each time an exception is made $\tilde{A}_1(t) - A_1(t)$ increases by one while $\tilde{A}(t) - A(t)$ increases by more than one. Hence (22) follows from results in previous section. □

Let $\mathcal{C}'$ (resp. $\mathcal{C}''$) be the subgraph induced by the vertices removed by the algorithm between 0 and T_1 (resp. T_2). By Lemma 13, we have

$$v(\mathcal{C}')/n \quad \xrightarrow{p} \quad 0 \tag{24}$$

$$v(\mathcal{C}'')/n \quad \xrightarrow{p} \quad h_1(1) - h_1(\xi) = 1 - h_1(\xi). \tag{25}$$

Hence informally, the exception made at time T_1 triggers a large cascade.

Now consider the case $\xi = 0$. Note in particular that we have $h(0) = 0$ and since $h_1(z) \leq h(z)$, we also have $h_1(0) = 0$. Then with the same argument as above, we have that $B(t)$ remains positive on $[\epsilon, +\infty)$ and thus no exception is made after a last exception made at time T_1 with $T_1 \to^p 0$. Hence (24) and (25) follow with $\mathcal{C}''$ being the whole graph, $h_1(0) = 0$.

We now finish the proof of Theorem 11 in the case where $\xi = 0$ or there exists $\epsilon > 0$ such that $g(z) < 0$ for $z \in (\xi - \epsilon, \xi)$. First by [17] Theorems 3.5 and 3.9, the cascade condition implies that $v(P) = \Omega_p(n)$. The result $\lim_n \frac{v(P)}{n} = 1 - \bar{h}_1(\overline{\xi}) > 0$ could be derived from [17]. We give an alternative proof for this result at the end of this section. For now, we denote $\gamma = 1 - \bar{h}_1(\overline{\xi}) > 0$. Coming back to the diffusion process analyzed above, we clearly have $\mathcal{C}' \cap P = \emptyset$. We now prove that $P \subset \mathcal{C}''$. This is clear in the case $\xi = 0$. We now concentrate on the case $\xi > 0$. First, let T_3 be the first time after T_2 when an exception is made. Since $\tilde{A}(t) - A(t)$ increases by at most $d_{\max} = o_p(n)$ each time an exception is made, we obtain from (23):

$$\sup_{t \leq T_3} (\tilde{A}(t) - A(t)) \leq \sup_{t \leq T_2} (\tilde{A}(t) - A(t)) + d_{\max} = o_p(n).$$

Hence similarly as in (21), we have for every $\epsilon > 0$, w.h.p. $\tau + \epsilon > T_3$. Since also $T_3 > T_2 \to^p \tau$, it follows that $T_3 \to^p \tau$. If $\mathcal{C}'''$ is the subgraph removed by the algorithm between T_2 and

T_3, then Lemma 13 implies that $v(\mathcal{C}''') = o_p(n)$. Assume now that $\mathcal{C}'' \cap P = \emptyset$, then with probability at least $\gamma > 0$, the vertex chosen for the exception at T_2 belongs to P and then we have $\mathbb{P}(\mathcal{C}'''\text{has more than } \gamma n \text{ vertices}) \geq \gamma$, in contradiction with $v(\mathcal{C}''') = o_p(n)$. Hence we have $\mathcal{C}'' \cap P \neq \emptyset$ and then $P \subset \mathcal{C}''$ as claimed.

We clearly have for any $u \in P$:

$$v(\mathcal{C}(P)) = v(\mathcal{C}(u)) = v(\cap_{u \in P} \mathcal{C}(u)).$$

Hence we only need to prove that $v(\mathcal{C}(P)) \geq 1 - h_1(\xi)n + o_p(n)$. To see this, attach to each vertex i a random variable U_i uniformly distributed over $[0, 1]$. Each time an exception has to be made, pick the vertex among the remaining ones with minimal U_i so that we do not change the algorithm described at the beginning of the section. From the analysis above, we see that all exceptions made before T_1 are vertices not in P and the exception made at time T_1 belongs to P. Now consider the graph $\tilde{G}$ obtained from the original graph where $\mathcal{C}'$ has been removed but all other variables are the same as in the original graph. Since $v(\mathcal{C}') = o_p(n)$, this graph satisfies Conditions 8 and 10. Hence previous analysis applies and we have in addition that the first exception made by the algorithm belongs to P since a pivotal vertex in G is also pivotal in $\tilde{G}$. Hence the subgraph of $\tilde{G}$ removed between times $\tilde{T}_1 = 0$ and $\tilde{T}_2$ by the algorithm is exactly $\tilde{\mathcal{C}}(P)$ in $\tilde{G}$. Since $\tilde{G}$ is a subgraph of G, we have $\tilde{\mathcal{C}}(P) \subset \mathcal{C}(P)$ in the original graph. And the first claim in (i) follows from (25) applied to the graph $\tilde{G}$. The second claim in (i) follows exactly as in the proof of Theorem 9 given in previous section.

Now consider the case where $\xi > 0$ but for any $\epsilon > 0$, there exists $z \in (\xi - \epsilon, \xi)$ such that $g(z) \geq 0$ and hence $f(z) > 0$. The idea to get a lower bound is to let π vary. Since $\pi > 0$, for any $0 < \pi' < \pi$, we see that by a standard coupling argument, for any given intial seed all active nodes in the model with π' will also be active in the model with π. Hence the model with π' provides a lower bound for the number of active nodes in the model with π. Now consider the function $g(z, \pi) = \lambda z(1 - \pi(1 - z)) - a(1 - \pi(1 - z))$ as a function of π. We have

$$\frac{\partial g}{\partial \pi}(z, \pi) = (1 - z)\left[a'(1 - \pi(1 - z)) - \lambda z\right]$$

Since ξ is a local minimum of $z \mapsto g(z, \pi)$ and g is differentiable as a function of z, we have

$$\frac{\partial g}{\partial z}(\xi, \pi) = 0 \Leftrightarrow a'(1 - \pi(1 - \xi)) = \frac{\lambda}{\pi}(1 - \pi(1 - \xi)) + \lambda\xi.$$

Hence we have $\frac{\partial g}{\partial \pi}(\xi, \pi) = \frac{\lambda}{\pi}(1 - \xi)(1 - \pi(1 - \xi)) > 0$. In particular for $\pi - \epsilon < \pi' < \pi$, we have $g(\xi, \pi') < g(\xi, \pi) = 0$. Let $\xi(\pi') = \sup\{z \in [0, 1),\ g(z, \pi') = 0\}$, then we have $\xi(\pi') > \xi$ for any $\pi' < \pi$ and $g(z, \pi') < 0$ for $z \in (\xi, \xi(\pi'))$. Moreover, we have $\xi(\pi') \to \xi(\pi)$ as $\pi' \to \pi$ and previous argument is valid for the model with π' as close as desired from π showing that $1 - h_1(\xi(\pi')) \to 1 - h_1(\xi)$ is a lower bound for the fraction of active nodes in the model with π.

We finish this proof by computing the asymptotic for the size of P using our previous analysis but for a modified threshold as done in [25]. We add a bar for quantities associated to this new model. Namely, consider a modification of the original diffusion with threshold $\overline{K}_i(d_i) = (d_i + 1)\mathbb{1}(K_i(d_i) \geq 1)$. In words, a node i becomes active if one of its neighbor is active and $K_i(d_i) = 0$ in the original diffusion. Clearly the nodes that become active in this model need to have only one active neighbor in the original contagion model with parameter $K_j(d_j)$. We denote by $\overline{\mathcal{C}}(u)$ the subgraph of final active nodes with initial active node u. Note

that our algorithm is an exploration process of the components of the graph on which we apply a bond percolation with parameter π and a site percolation by removing all vertices with $k_i \geq 1$. In particular, the largest component explored is exactly P. Note that the computations made before are valid with the following functions:

$$\begin{aligned}\bar{g}(z) &= (1-\pi+\pi z)\left\{\lambda z - \sum_s s p_s (1-t_{s0}) - \sum_s s p_s t_{s0}(1-\pi+\pi z)^{s-1}\right\}\\ \bar{h}_1(z) &= \sum_s p_s(1-t_{s0}) + \sum_s p_s t_{s0}(1-\pi+\pi z)^s.\end{aligned}$$

Hence we have:

$$\overline{\xi} = \sup\{z \in [0,1],\ \bar{g}(z) = 0\}.$$

Note that if we denote $\phi(z) = \bar{g}(z)(1-\pi+\pi z)^{-1}$, then we have

$$\begin{aligned}\phi'(z) &= \lambda - \pi \sum_s s(s-1) p_s t_{s0}(1-\pi+\pi z)^{s-2},\\ \phi''(z) &= -\pi^2 \sum_s s(s-1)(s-2) p_s t_{s0}(1-\pi+\pi z)^{s-3}.\end{aligned}$$

In particular, ϕ is concave on $(0,1]$ and strictly concave unless $p_s = 0$ for $s \geq 3$. Note also that $\phi'(1) = \lambda - \pi \sum_s s(s-1) p_s t_{s0}$, so that under the cascade condition $\phi'(1) < 0$ and ϕ is strictly concave. Hence, we have $\overline{\xi} < 1$ and $\phi(x) > 0$ for $x \in (\overline{\xi}, 1)$ and if $\xi > 0$, then $\phi(x) < 0$ for $x < \overline{\xi}$. In particular, previous analysis allows to conclude that $\lim_n \frac{v(P)}{n} = 1 - \bar{h}_1(\overline{\xi}) > 0$.

6 Proof of Proposition 6

By Theorem 11, when the cascade condition holds, the set of active vertices contain the set of pivotal vertices and hence has a giant component. We denote $\mathcal{I}$ the induced subgraph of inactive vertices, in the pivotal equilibrium (i.e. in the final state when all pivotal nodes are initially active). By Theorem 11, we have for $v_r(\mathcal{I})$, the number of vertices in $\mathcal{I}$ with degree r in $\mathcal{I}$:

$$\frac{v_r(\mathcal{I})}{n} \xrightarrow{p} \sum_{r \geq s - \lfloor qs \rfloor} p_s b_{sr}(\xi),\ \text{where, } \xi = \xi(\lambda) = \max\left\{z < 1,\ \lambda z^2 = \sum_s p_s \sum_{r \geq s - \lfloor qs \rfloor} r b_{sr}(z)\right\}.$$

We denote $v_r(z) = \sum_{r \geq s - \lfloor qs \rfloor} p_s b_{sr}(z)$. Thanks to the result on the distribution of $\mathcal{I}$, there is a giant component of inactive vertices if

$$\sum_r r(r-1) v_r(\xi) > \sum_r r v_r(\xi) = \sum_s p_s \sum_{r \geq s - \lfloor qs \rfloor} r b_{sr}(\xi) = \lambda \xi^2,$$

which can be rewritten as in Proposition 6.

Now we assume that $p_r = \frac{\lambda^r}{r!} e^{-\lambda}$, $\pi = 1$ and $K(d) = \lfloor qd \rfloor$. The function $\psi(\lambda) = e^{-\lambda} \sum_{r < q^{-1}} \frac{\lambda^{r-1}}{(r-2)!}$ is increasing for $\lambda \leq \lambda^*$ and then decreasing. We assume that q is fixed suh that the cascade condition holds. Then $\psi(\lambda^*) > 1$. Then $\lambda_i(q) = \sup\{\psi(\lambda) < 1, \lambda < \lambda^*\}$ while $\lambda_s(q) = \inf\{\psi(\lambda) > 1,\ \lambda > \lambda^*\}$. We denote $\zeta(\lambda) = \max\left\{z < 1,\ \lambda z^2 = \sum_s p_s \sum_{r \geq s - \lfloor qs \rfloor} r(r-1) b_{sr}(z)\right\}$. Then, there is a giant component of inactive vertices if $\zeta(\lambda) < \xi(\lambda)$. Both functions are non-increasing in $\lambda \in (\lambda_i(q), \lambda_s(q))$ and are intersecting only once in $(\lambda_i(q), \lambda_s(q))$.

References

[1] H. Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electron. J. Combin.*, 17(1):Research Paper 25, 20, 2010.

[2] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[3] N. T. J. Bailey. *The mathematical theory of infectious diseases and its applications*. Hafner Press [Macmillan Publishing Co., Inc.] New York, second edition, 1975.

[4] J. Balogh and B. G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures Algorithms*, 30(1-2):257–286, 2007.

[5] L. E. Blume. The statistical mechanics of best-response strategy revision. *Games Econom. Behav.*, 11(2):111–145, 1995. Evolutionary game theory in biology and economics.

[6] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[7] J. Carlson and J. Doyle. Complexity and robustness. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 1):2538, 2002.

[8] R. Durrett. *Random graph dynamics*. Cambridge University Press, Cambridge, 2007.

[9] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.

[10] G. Ellison. Learning, local interaction, and coordination. *Econometrica*, 61(5):1047–1071, 1993.

[11] G. Ellison and D. Fudenberg. Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, 110(1):93–125, 1995.

[12] A. Galeotti and S. Goyal. A theory of strategic diffusion, 2008.

[13] A. Galeotti, S. Goyal, M. O. Jackson, and F. Vega-Redondo. Network games. *Rev. Econom. Stud.*, 77(1):218–244, 2010.

[14] M. Goetz, J. Leskovec, M. McGlohon, and C. Faloutsos. Modeling blog dynamics. In E. Adar, M. Hurst, T. Finin, N. S. Glance, N. Nicolov, and B. L. Tseng, editors, *ICWSM*. The AAAI Press, 2009.

[15] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA, 2008.

[16] M. O. Jackson and L. Yariv. Diffusion of behavior and equilibrium properties in network games. *The American Economic Review*, 97(2), 2007.

[17] S. Janson. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.

[18] S. Janson. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, 18(1-2):205–225, 2009.

[19] S. Janson and M. J. Luczak. A simple solution to the k-core problem. *Random Structures Algorithms*, 30(1-2):50–62, 2007.

[20] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures Algorithms*, 34(2):197–216, 2009.

[21] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[22] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[23] D. Kempe, J. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *KDD '03*.

[24] J. Kleinberg. Cascading behavior in networks: algorithmic and economic issues. In *Algorithmic game theory*, pages 613–632. Cambridge Univ. Press, Cambridge, 2007.

[25] M. Lelarge. Diffusion of innovations on random networks: Understanding the chasm. In C. H. Papadimitriou and S. Zhang, editors, *WINE*, volume 5385 of *Lecture Notes in Computer Science*, pages 178–185. Springer, 2008.

[26] M. Lelarge and J. Bolot. A local mean field analysis of security investments in networks. In J. Feigenbaum and Y. R. Yang, editors, *NetEcon*, pages 25–30. ACM, 2008.

[27] M. Lelarge and J. Bolot. Network externalities and the deployment of security features and protocols in the internet. In Z. Liu, V. Misra, and P. J. Shenoy, editors, *SIGMETRICS*, pages 37–48. ACM, 2008.

[28] J. Leskovec, L. A. Adamic, and B. A. Huberman. The dynamics of viral marketing. In J. Feigenbaum, J. C.-I. Chuang, and D. M. Pennock, editors, *ACM Conference on Electronic Commerce*, pages 228–237. ACM, 2006.

[29] J. Leskovec, L. Backstrom, and J. M. Kleinberg. Meme-tracking and the dynamics of the news cycle. In J. F. E. IV, F. Fogelman-Soulié, P. A. Flach, and M. J. Zaki, editors, *KDD*, pages 497–506. ACM, 2009.

[30] J. Leskovec, A. Singh, and J. M. Kleinberg. Patterns of influence in a recommendation network. In W. K. Ng, M. Kitsuregawa, J. Li, and K. Chang, editors, *PAKDD*, volume 3918 of *Lecture Notes in Computer Science*, pages 380–389. Springer, 2006.

[31] D. López-Pintado. Diffusion in complex social networks. *Games Econom. Behav.*, 62(2):573–590, 2008.

[32] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures Algorithms*, 6(2-3):161–179, 1995.

[33] S. Morris. Contagion. *Rev. Econom. Stud.*, 67(1):57–78, 2000.

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

[35] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 61–70, New York, NY, USA, 2002. ACM.

[36] F. Vega-Redondo. *Complex social networks*, volume 44 of *Econometric Society Monographs*. Cambridge University Press, Cambridge, 2007.

[37] D. J. Watts. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. USA*, 99(9):5766–5771 (electronic), 2002.

[38] H. Young. The spread of innovations through social learning. 2005.

7 Appendix

7.1 Technical lemma

Lemma 14. *For any $x, \pi \in [0,1]$, and $k \geq 0$ we have*

$$\frac{x}{1-\pi+x\pi}\sum_{r\geq k} r b_{s,r}(1-\pi+x\pi) = \sum_{r+i\geq k} r b_{s,r}(x) b_{s-r,i}(1-\pi).$$

Proof. This follows from the following observations for $x > 0$:

$$\begin{aligned}
\sum_{r\geq k} \frac{r}{1-\pi+x\pi} b_{s,r}(1-\pi+x\pi) &= (s-1)\sum_{r\geq k-1} b_{s-1,r}(1-\pi(1-x)) \\
&= (s-1)\sum_{r\leq s-k} b_{s-1,r}(\pi(1-x)) \\
&= (s-1)\mathbb{P}(\sum_{i=1}^{s-1} B_i Y_i \leq s-k),
\end{aligned}$$

where the B_i's and Y_i's are independent Bernoulli random variables with parameter π and $1-x$ respectively. Now we also have:

$$\begin{aligned}
\sum_{r+i\geq k} \frac{r}{x} b_{s,r}(x) b_{s-r,i}(1-\pi) &= (s-1)\sum_{r+i\geq k-1} b_{s-1,r}(x) b_{s-1-r,i}(1-\pi) \\
&= (s-1)\mathbb{P}(\sum_{i=1}^{s-1}(1-Y_i) + \sum_{i=1}^{s-1} Y_i(1-B_i) \geq k-1).
\end{aligned}$$

□

7.2 Properties of the functions h and g

We now justify the use of the max in (5). Let $\sigma(d)$ be a Bernoulli random variable with parameter α_d. Let for $0 \leq x \leq 1$, D_x be the thinning of D obtained by taking D points and then randomly and independently keeping each of them with probability x. Thus given $D = d$, $D_x \sim \mathrm{Bi}(d, x)$. With these notations, we have

$$\begin{aligned} h(z;\boldsymbol{\alpha},\boldsymbol{p},\pi) &= \mathbb{E}\left[D_{1-\pi+\pi z}(1-\sigma(D))\mathbb{1}\left(D_{1-\pi+\pi z} \geq D - K(D)\right)\right] \\ h_1(z;\boldsymbol{\alpha},\boldsymbol{p},\pi) &= \mathbb{P}\left(\sigma(D) = 0,\, D_{1-\pi+\pi z} \geq D - K(D)\right), \end{aligned}$$

so that both h and h_1 are non-decreasing in z and non-increasing in π.

Note that if $\boldsymbol{\alpha} = 0$, then $h(1;0,\boldsymbol{p},\pi) = \lambda$ so that $g(1;0,\boldsymbol{p},\pi) = 0$ and $\hat{z} = 1$. We now consider the case $\boldsymbol{\alpha} \neq 0$, so that there exists $d \geq 1$ such that $\alpha_d > 0$. In this case, we have $g(1;\boldsymbol{\alpha},\boldsymbol{p},\pi) = \lambda - \sum_s s(1-\alpha_s)p_s \geq \alpha_d p_d d > 0$. The statement then follows from the fact that the only possible jumps for $z \mapsto g(z;\boldsymbol{\alpha},\boldsymbol{p},\pi)$ are downards. More precisely, let $\hat{z} = \sup\{z \in [0,1], g(z;\boldsymbol{\alpha},\boldsymbol{p},\pi) = 0\}$. Since the function h is non-decreasing in z, its set of discontinuity points is denumerable say $\{z_i\}_{i\in\mathbb{N}}$ and h admits a left and right limit at any point z denoted by $h(z-)$ and $h(z+)$ respectively. If $\hat{z} \in \{z_i\}$, then we have $h(\hat{z}-) \leq \lambda\hat{z}(1-\pi+\pi\hat{z}) \leq h(\hat{z}+)$. In particular, we have $g(z;\boldsymbol{\alpha},\boldsymbol{p},\pi) \leq 0$ for any $z > \hat{z}$ which contradicts the fact that $g(1;\boldsymbol{\alpha},\boldsymbol{p},\pi) > 0$. Hence the functions h and g are continuous at $\hat{z}$ and the sup is attained and can be replaced by a max.

7.3 Proof of Lemma 7

Taking expectation in (4), we get with $x = \mathbb{E}[Y]$,

$$\begin{aligned} 1-x &= \sum_{s\geq 0} p_s^*(1-\alpha_{s+1})\mathbb{P}\left(\sum_{i=1}^{s} B_i Y_i \leq K(s+1)\right) \\ &= (1-\alpha_1)\frac{p_1}{\lambda} + \sum_{s\geq 1}\frac{(s+1)p_{s+1}}{\lambda}(1-\alpha_{s+1})\sum_{\ell} t_{s+1\ell}\sum_{j\leq\ell} b_{sj}(x\pi) \\ &= \frac{1}{\lambda}\left\{(1-\alpha_1)p_1 + \sum_{s\geq 1; j\geq s-\ell}(1-\alpha_{s+1})p_{s+1}t_{s+1\ell}(s+1)b_{sj}(1-x\pi)\right\}. \end{aligned}$$

Note that $(s+1)b_{sj}(1-x\pi) = \frac{j+1}{1-x\pi}b_{s+1j+1}(1-x\pi)$ for $s \geq 1$, so that

$$\begin{aligned} 1-x &= \frac{1}{\lambda}\left\{(1-\alpha_1)p_1 + \sum_{s\geq 1; j\geq s-\ell}\frac{(j+1)}{1-x\pi}(1-\alpha_{s+1})p_{s+1}t_{s+1\ell}b_{s+1j+1}(1-x\pi)\right\} \\ &= \frac{1}{\lambda(1-x\pi)}\sum_{s\geq 1; j\geq s-\ell} j(1-\alpha_s)p_s t_{s\ell} b_{sj}(1-x\pi). \end{aligned}$$

Hence, we get

$$(1-x\pi)(1-x)\lambda = \sum_{s\geq 1; j\geq s-\ell} j(1-\alpha_s)p_{s\ell}b_{sj}(1-x\pi).$$

This establishes the first part of the lemma thanks to Lemma 14. Taking expectation in (3) gives: $\mathbb{E}[X_\emptyset] = 1 - \sum_{s;j\geq s-\ell}(1-\alpha_s)p_s t_{s\ell} b_{sj}(1-x\pi)$, and the second part of the lemma follows.