Percolation on fitness dependent networks with heterogeneous resilience

K. Hoppe, G. J. Rodgers
*Department of Mathematical Sciences, Brunel University,
Uxbridge, Middlesex UB8 3PH, United Kingdom*

(Dated: August 27, 2014)

The ability to understand the impact of adversarial processes on networks is crucial to various disciplines. The objects of study in this article are fitness driven networks. Fitness dependent networks are fully described by a probability distribution of fitness and an attachment kernel. Every node in the network is endowed with a fitness value and the attachment kernel translates the fitness of two nodes into the probability that these two nodes share an edge. This concept is also known as mutual attractiveness. In the present article, fitness does not only serve as a measure of attractiveness, but also as a measure of a node's robustness against failure. The probability that a node fails increases with the number of failures in its direct neighborhood and decreases with higher fitness. Both static and dynamic network models are considered. Analytical results for the percolation threshold and the occupied fraction are derived. One of the results is that the distinction between the dynamic and the static model has a profound impact on the way failures spread over the network. Additionally, we find that the introduction of mutual attractiveness stabilizes the network compared to a pure random attachment.

I. INTRODUCTION

The investigation of contagious processes is a vital part of many fields in academia, whether it is the transmission of sexual diseases, studied in social science [1], the spread of financial distress in economics [2, 3] or the spread of viruses in epidemiology [4]. The most simple transmission models assume perfect mixing of the underlying population. In other words, contacts between banks, sexual partners or pedestrians are purely random. This simplification is made in order to be able to formulate a system of coupled nonlinear differential equations that can be solved with ease [5]. However, this assumption is an oversimplification of the underlying problem. Contact patterns are largely heterogeneous [6]. This heterogeneity has been endogenized by considering underlying networks that describe different contact motifs [7–10].

Beyond the introduction of a topology of contacts, further aspects have been introduced, such as timing and local transmission probabilities [5], in order to account for real-life phenomena. The spreading behavior in multi-layered network architectures and networks of networks has also been investigated [11–14]. Other features, such as awareness and vaccinations were considered too [15, 16]. In other studies, more complex models that attempt to explain default cascades, occurring for instance in financial markets, have been considered [3]. In [3] a model was set up that takes into account different kinds of shocks, as well as correlations, not only degree-degree but also degree-robustness, and a model specific market illiquidity parameter.

The present article considers a flat network, in the sense that it consists of one layer. The underlying topology is not chosen to be arbitrary, it is derived from a hidden variable model [17–21]. In contrast to node degree based models, such as the Barabási-Albert model [22], or models, that combine local attractiveness together with node degree as an attractor for new edges [23], fitness/hidden-variable models are purely driven by static node intrinsic fitness. The two functions that determine the topology of the network and all related quantities are the attachment kernel $f(x, y)$, that describes the probability that a node with fitness x originates an edge towards a node with fitness y, and the probability density $\rho(x)$, that describes the distribution of fitness in the system. The literature on pure fitness models is split into two parts, static and dynamic models, which are both considered in this paper. Static networks have been identified as the correct model in various applications such as the interbank lending market and the world trade network [24–26]. Their dynamic counterpart, introduced in [20], relies on assumptions that are met in several real-world examples. The attractiveness of fitness driven models in general is the amount of information that a local agent is assumed to have. While degree dependent growth models assume that new nodes have information about the connectivity of the entire network, it is sufficient for fitness dependent models to assume that nodes have information about the ranking of some derived, topological independent quantity. Consider for instance the network of investments. The balance sheet structure of an enterprise is more likely to be accessible to an investor than the absolute number of other investors invested in a particular company.

In this paper, we couple for the first time attractiveness and resilience. Fitness determines the propensity to acquire more edges in the network, but also measures robustness against failures in the direct neighborhood of a node. This association of attractiveness and robustness is not just interesting as an exercise for a theoretician, it is of great interest to organizations that can influence the underlying free parameters, such as the distribution of fitness, as well as the average node-degree in the system. Consider for example the interbank lending market. Assuming that the selection process is fixed, in the sense that $f(x, y)$ is given and cannot be influenced, the policy maker or regulator can create incentives to stimulate interaction between existing agents in the form of new edges that increase the average node degree, or they can reshape the fitness distribution $\rho(x)$ by the introduction of a tax. It is shown later in the paper that these two quantities largely influence the percolation threshold as well as the fraction of the network that is occupied by the giant vulnerable component.

Attractiveness and robustness cannot always be assumed to correspond one-to-one, such as in the case of the sexual contact network that has been investigated in [1]. However there are many scenarios in which this correspondence appears realistic, consider again the interbank-lending market. Here, attractiveness is derived directly from the likelihood to file for bankruptcy. Another possible model is the network of buyer-supplier relationships in the technological sector. Think of the arrival of a new technology as an epidemic process. In this case large suppliers are attractive due to their size, but also control the adoption of new technologies. One example is the Flash-technology, that has never been adopted by Apple on its mobile devices.

The remainder of this article is organized as follows. In Section II the two models under investigation are presented. In Section III the core results on percolation for the given models are derived. In Section IV the main results are presented. Section V closes the article with concluding remarks.

II. MODELS

In this paper we study the exposure of contact networks, that are not necessarily heterogeneous, to random failures. We consider two different classes of network formation models. These are static fitness models [17] and their dynamic counterpart [20]. The distinctive characteristics of these two networks is the set of nodes. The dynamic model is characterized by a constant addition of nodes, while the static model comprises a fixed number of nodes. The timescale of the network formation is assumed to be longer than the typical timescale of the spread of a failure, such that the formation of the network and its occupation with a large failed component can be considered separately. Fitness models are characterized by an attachment kernel $f(x, y)$ and a fitness distribution $\rho(x)$. The attachment kernel $f(x, y)$ describes the probability that a node with fitness x originates a new edge toward a node with fitness y. The fitness of each node is static over the lifetime of the network and is drawn from the probability density $\rho(x)$. Fitness is assumed to be distributed over the unit interval $[0, 1]$. If fitness is distributed over another interval in a particular application it can always be normalized to the interval $[0, 1]$.

The network that we consider is directed and every agent is in one of two states, bankrupt or solvent / dead or alive / technology-adapter or -refuser, / infected or susceptible etc. Although the model can be understood in terms of many applications, it is in general a binary rule as it was introduced in [27]. The two states will be referred to as solvent and bankrupt hereafter. Every agent i is initially solvent and changes its state to bankrupt if a critical fraction of the agents in the agent's neighborhood have gone bankrupt. This critical fraction is given by its fitness x_i. If at least a fraction x_i of a node's neighbors has changed its state to bankrupt, node i also changes its state from solvent to bankrupt. The networks under investigation here are both directed. In the context of the interbank lending market, an edge from a bank i to a bank j represents a money flow from i to j. If too many banks that have borrowed money from node i fail, bank i will also fail. The critical fraction is hence coupled to the out-degree of a node, since exposure flows against edge-directionality. Consider now a randomly induced initial bankruptcy somewhere in the network. This bankruptcy can only propagate to a neighbor that has a fitness value such that $x_i \leq 1/k_i^{\text{out}}$. Nodes that fulfill this condition are referred to as vulnerable. Once the initially solvent network is perturbed with the state change of a single bankruptcy, the spread of bankruptcies over the network develops asynchronously in accordance the simple threshold rule $x_i \leq 1/k_i^{\text{out}}$.

The vulnerability condition $x \leq 1/k$ implies that a bank distributes its lent money uniformly over all its obligors, which might be different in the real world but we will abstract from this possibility for the sake of simplicity. If the vulnerability condition is fulfilled, a single failing bank in the neighborhood of a node i can cause i to fail as well.

A. Static model

The static fitness model describes a network that comprises a fixed number of nodes N and edges M. Since failures/adaptations can only spread against the direction of edges, the quantity of interest here is the out-degree of a node. Analytical results on the degree distribution for this model have been obtained earlier, see for instance Refs. [17, 19]. In order to consider percolation in this network, it is necessary to obtain the fitness-conditional out-degree distribution. This quantity, also referred to as the propagator has been found to be Poissonian in Ref. [18]. While the propagator is derived constructively in Ref. [18], a derivation from first principles is described below.

Although the network is static i.e. it contains a fixed number of nodes, the process in which edges are added can be understood as a dynamic procedure. Edges are deployed one by one, not all at the same time. An edge is added to a pair of nodes $i \to j$ with probability $f(x_i, x_j)/\sum_{k,l} f(x_k, x_l)$. The networks under consideration are sparse, in the sense that $M \ll N^2$, thus the possibility of adding an edge to a pair of nodes that is already connected is negligible in the thermodynamic limit. The probability that a node with fitness x increases its out-degree by one during an edge-addition step is defined as

$$\lambda(x,N)=\sum_{l=1}^{N}\frac{f(x,x_l)}{\sum_{i=1}^{N}\sum_{j=1}^{N}f(x_i,x_j)}. \tag{1}$$

That is the properly normalized probability that an edge is deployed originating from a node with fitness x. For large N, fitness can be considered as a continuous variable, which leads to

$$\lambda(x,N)=\frac{1}{N}\int_0^1\frac{f(x,y)\rho(y)dy}{\int_0^\infty\int_0^\infty f(\xi,\eta)\rho(\xi)\rho(\eta)d\xi d\eta}. \tag{2}$$

Since $1/N$ is a factor of $\lambda(x,N)$, we can define $\lambda(x)=N\lambda(x,N)$. Using the postulate that the edge addition procedure can be understood as a sequential process, the fitness-conditional degree distribution can be found using a rate equation approach. The probability that a node with fitness x has out-degree k in a network with N nodes and M edges will be denoted with $p^{(\mathrm{s})}_{M,N}(k|x)$ and evolves as

$$\begin{aligned}p^{(\mathrm{s})}_{M+1,N}(k|x)&=p^{(\mathrm{s})}_{M,N}(k|x)\left(1-\lambda(x,N)\right)\\&\quad+p^{(\mathrm{s})}_{M,N}(k-1|x)\,\lambda(x,N).\end{aligned} \tag{3}$$

The first part of Eq. (3) corresponds to the an edge update that occurs at a node with fitness unequal to x, the second part corresponds to an edge update that increases the out-degree of a node with fitness x by one. The solution of Eq. (3) is given by

$$p^{(\mathrm{s})}_{M,N}(k|x)=\frac{e^{-M/N\,\lambda(x)}(\frac{M}{N}\lambda(x))^k}{\Gamma(k+1)}. \tag{4}$$

The details of the derivation can be found in Appendix A. This result is sufficient to characterize vulnerable nodes in the network. The probability that a randomly chosen node has degree k and is vulnerable, i.e. whose fitness value is less than the reciprocal of its out degree, is defined as $Q^{(\mathrm{s})}(k)=\mathbb{P}[k_i^{\mathrm{out}}=k\cap x_i\leq 1/k]$ and can be expressed in terms of $p^{(\mathrm{s})}_{M,N}(k|x)$:

$$\begin{aligned}Q^{(\mathrm{s})}(k)&=\int_0^{1/k}p^{(\mathrm{s})}_{M,N}(k|x)\rho(x)dx &(5)\\&=\frac{1}{\Gamma(k+1)}\int_0^{1/k}e^{-M/N\,\lambda(x)}\left(\frac{M}{N}\lambda(x)\right)^k\rho(x)dx. &(6)\end{aligned}$$

This expression is in excellent agreement with results from numerical simulations of the network assembly process. Fig. 1 compares the prediction of Eq. (6) with numerical simulations for various cases of $\rho(x)$,$f(x,y)$ and M/N. Another observation from Fig. 1 is that the qualitative differences between $Q^{(\mathrm{s})}(k)$ for exponential- and Pareto distributed fitness are minimal. The qualitative differences between random attachment, i.e. $f(x,y)=1$ and mutual attractiveness with $f(x,y)=xy$ become more evident for large M/N. The average out-degree changes the behavior of $Q^{(\mathrm{s})}(k)$ significantly. The distribution becomes generally broader for higher M/N. One more interesting aspect of the network is the effect of diversification. Fig. 2(a) shows the conditional probability that a randomly chosen node is vulnerable, given that it has degree k. That is

$$\mathbb{P}[x\leq 1/k\,|\,k_i^{\mathrm{out}}=k]=\frac{Q(k)}{P(k)} \tag{7}$$

with $P(k)=\int_0^1 p_{M,N}(k|x)\rho(x)dx$. Fig. 2(a) shows that for the case of mutual attractiveness, the transition between finding a vulnerable node with certainty and not finding any vulnerable node for a given degree is occurring at higher degrees for larger $\langle k\rangle$. This implies that adding new edges destabilizes the system. This adverse diversification effect is at first glance counterintuitive. It implies that the higher the average degree is in the system, the larger is the probability that a node with a given degree is vulnerable. In other words, given a node with degree $k'>1$, it is more likely that this node is vulnerable in a system with $\langle k\rangle=k'$, than in a system with $\langle k\rangle=1$. This is because in a system with average degree $\langle k\rangle=1$, a node has to have a greater than average fitness to attract k' edges, and therefore is more unlikely to fail. In a system with $\langle k\rangle=k'$ however, a node only needs to have average fitness to attract k' many edges. Due to its lower fitness, it is therefore more prone to failures. The case of random attachment, i.e. $f(x,y)=1$ shows an entirely different behavior. The quotient $Q^{(\mathrm{s})}(k)/P^{(\mathrm{s})}(k)$ is almost the same over the whole range of k. This arises from the form of $\rho(x)$, for $f(x,y)=1$, the quotient is given by

$$\frac{Q(k)}{P(k)}=\int_0^{1/k}\rho(x)dx. \tag{8}$$

The case of random attachment is different because fitness does not have an influence on the degree of a node. The probability for a randomly chosen node to be vulnerable simply decays therefore with the cumulative distribution of fitness as shown in Eq. (8). Fig. 2(b) shows that adding additional edges to the network has a positive effect overall, as expected. The fraction of vulnerable nodes decays with increasing $\langle k\rangle$.

Other quantities of interest are the conditional and the unconditional expectation of the out-degree of a randomly chosen node $\mathbb{E}[k|x]$ and $\langle k\rangle$ respectively. These are obtained as

$$\mathbb{E}[k|x]=\sum_{k\geq 0}kp^{(\mathrm{s})}_{M,N}(k|x)=\frac{M}{N}\lambda(x)\quad\text{and} \tag{9}$$

$$\langle k\rangle=\int_0^1\mathbb{E}[k|x]\rho(x)dx=\frac{M}{N}\int_0^1\lambda(x)\rho(x)dx. \tag{10}$$

Since the average out-degree of a node is determined by the number of edges M divided by the number of nodes

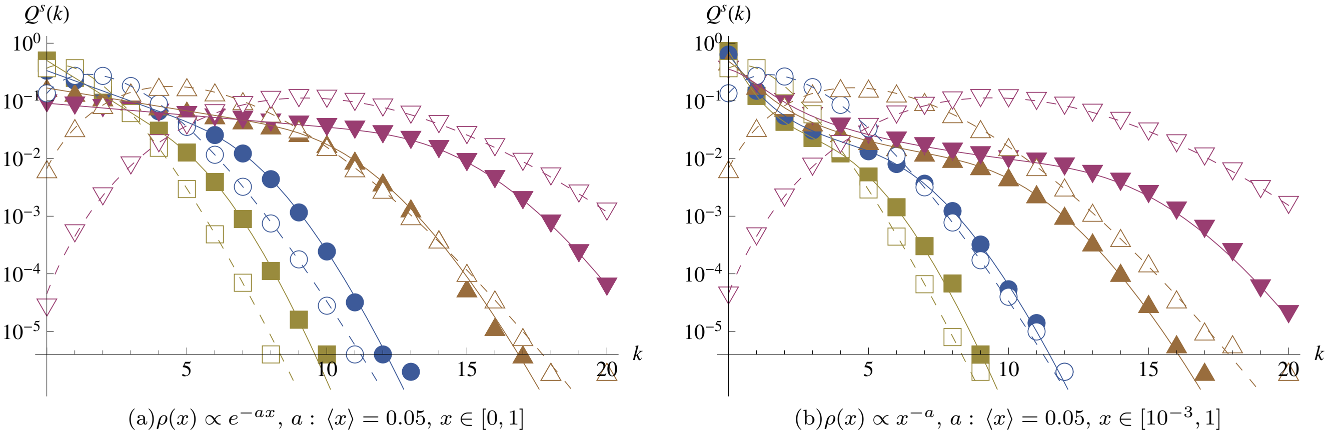


FIG. 1. (Color online) The plots show $Q^{(s)}(k)$ for different fitness distributions: (a) shows different configurations with fitness distributed exponentially on the unit-interval, with average fitness $\langle x\rangle = 0.05$. (b) shows different configurations with Pareto distributed fitness on the interval $[10^{-3}, 1]$, whereby the zero is excluded to avoid the singularity. Random selection, i.e $f(x, y) = 1$ is marked will hollow symbols, mutual attractiveness, i.e. $f(x, y) = xy$ with solid markers. The different plot markers, distinguish different average degrees: $\langle k\rangle = 1$ squares (yellow), $\langle k\rangle = 2$ circles (blue), $\langle k\rangle = 5$ triangles (brown), $\langle k\rangle = 10$ downward triangles (purple). The lines indicate analytical results from Eq. (6), whereby solid lines represent the case of $f(x, y) = xy$ and dashed lines $f(x, y) = 1$. Numerical results are obtained in a network with $N = 10^4$ nodes, averaged over 50 network realizations.

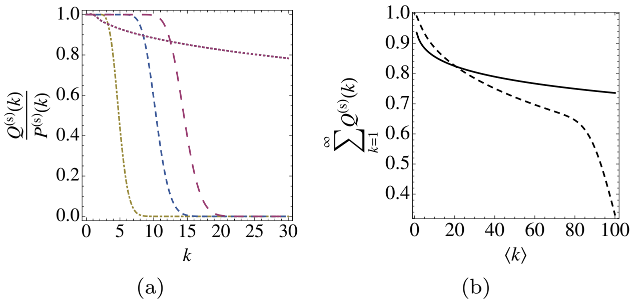


FIG. 2. (Color online) Plot (a) shows degree-conditional probabilities of vulnerability for different average degrees $M/N = 1$: dot-dashed (mustard), $M/N = 5$: fine-dashed (blue), $M/N = 10$ coarse-dashed (purple). Examples for $f(x, y) = xy$ are indicated with different dashing patterns, random attachment $f(x, y) = 1$ as dotted lines. The dotted lines are all overlapping, indicating that for random attachment, the conditional probability to find a vulnerable node is independent of M/N. This is analytically confirmed in Eq. (8). Plot (b) illustrates the relationship between the fraction of vulnerable nodes in the network and the average degree $\langle k\rangle$. The configuration in this example is power-law fitness, i.e. $\rho(x) = x^{-a}$, $x \in [10^{-3}, 1]\, a : \langle x\rangle = 0.05$. The solid line shows the result for $f(x, y) = xy$, the dashed line represents random attachment, i.e. $f(x, y) = 1$.

that are in the network N, Eq. (10) implies a normalization condition for $\rho(x)$ and $\lambda(x)$, such that

$$\int_0^1 \lambda(x)\rho(x)dx = 1. \tag{11}$$

B. Dynamic model

The dynamic model, as opposed to the static model, is characterized by sequential additions of nodes and edges. Thus, the discrete time, Markov chain approach that is used in the previous subsection is employed here again. The dynamic model that is investigated here is similar to the model in Ref. [20]. While an undirected network is considered in Ref. [20], the network in the present paper is directed. The evolution is as follows. At every timestep, one of two things can occur:

(i) With probability q, a new node is created and endowed with a fitness value x, drawn from a probability density ρ. A node inside the network is then chosen randomly to form a new edge toward this new node. The origin of this edge depends mutually on the new node's fitness x and the originating node's fitness y. The probability for such an edge is $f(y, x)/\sum_{l=1}^{N(t)} f(x_l, x)$.

(ii) With probability $1 - q$, a new edge between two existing nodes inside the network is created. The probability for a new edge $i \to j$ is $f(x_i, x_j)/\sum_{k,l}^{N(t)} f(x_k, x_l)$.

New nodes are added to the network with an incoming edge, this implies that for $q = 1$, the network is a perfect tree, with edges pointing from the root node toward the leafs. Consider again the interbank lending market. The rule above implies that a bank enters the market with a loan request. Alternatively, consider the network of cashflows in buyer-supplier relationships. In this case, the growth rule translates into a scenario in which a company enters the market as a seller.

The fitness-conditional degree distribution for the dynamic networks is similarly derived as it is presented in Ref. [20]. The main difference is the introduction of directionality. Also, while Ref. [20] compressed the distinction between the attachment kernel $f(x,y)$ and the fitness distribution $\rho(x)$, these two functions remain explicit in the calculation below, to facilitate later analysis.

The central quantity of this model is the probability that a node with fitness x originates a new edge at time t. This quantity is defined as

$$\begin{aligned}\theta_q(x,t) = q\int_0^1 \frac{f(x,y)}{\sum_{i=1}^{N(t)} f(x_i,y)}\rho(y)dy \\ + (1-q)\frac{\sum_{l=1}^{N(t)} f(x,x_l)}{\sum_{i=1}^{N(t)}\sum_{j=1}^{N(t)} f(x_i,x_j)}.\end{aligned} \tag{12}$$

The first part of Eq. (12) corresponds to a node addition step. The term inside the integral is the properly normalized probability that the node with fitness x initiates an edge toward the new node, averaged over all possible fitness of the newly added node. The second term corresponds to the case of edge addition. The fraction is the properly normalized probability that a node with fitness x receives a new outward pointing edge. The sums run over all nodes that are present in the network at time t. We consider the case of a sparse network, such that the number of edges is at all times $O(N(t))$. Thus, the problem of edge duplications can be neglected in the leading order approximation. Using that the distribution of fitness inside the network $(x_i)_{i=1}^{N(t)}$ is in the limit of large t simply described by $\rho(x)$ and that for $t\to\infty$ the expected value of $N(t)$ is given by qt, Eq. (12) can be written in terms of continuous fitness as

$$\begin{aligned}\theta_q(x,t) = \frac{1}{t}\Bigg[\int_0^1 \frac{f(x,y)\rho(y)}{\int_0^1 f(\xi,y)\rho(\xi)d\xi}dy \\ + \frac{(1-q)}{q}\frac{\int_0^1 f(x,y)\rho(y)dy}{\int_0^1\int_0^1 f(\xi,\eta)\rho(\xi)\rho(\eta)d\xi d\eta}\Bigg].\end{aligned} \tag{13}$$

Since $1/t$ is a factor of $\theta_q(x,t)$ in the $t\to\infty$ limit, we can use $\theta_q(x) = t\theta_q(x,t)$ in the following. The evolution of the out-degree distribution of a node with fitness x that joined at τ, $p_t^{(\mathrm{d})}(k|x,\tau)$ obeys

$$\begin{aligned}p_t^{(\mathrm{d})}(k|x,\tau) = p_t^{(\mathrm{d})}(k|x,\tau)\left(1-\theta_q(x,t)\right) \\ + p_t^{(\mathrm{d})}(k-1|x,\tau)\,\theta_q(x,t).\end{aligned} \tag{14}$$

The first term in Eq. (14) corresponds to the situation in which an edge is added to a node that has a fitness different from x. The second term corresponds to an update around a node with fitness x. The solution of Eq. (14) is given by

$$p^{(\mathrm{d})}(k|x) = \left(\frac{\theta_q(x)}{1+\theta_q(x)}\right)^k \frac{1}{1+\theta_q(x)}. \tag{15}$$

The details of the derivation can be found in Appendix B. The probability that a randomly chosen node has degree k and is vulnerable is calculated in the same way as in the previous subsections and is given by

$$\begin{aligned}Q^{(\mathrm{d})}(k) &= \int_0^{1/k} p^{(\mathrm{d})}(k|x)\rho(x)dx & (16)\\ &= \int_0^{1/k}\left(\frac{\theta_q(x)}{1+\theta_q(x)}\right)^k \frac{\rho(x)dx}{1+\theta_q(x)}. & (17)\end{aligned}$$

The conditional expectation of the out-degree is given by

$$\mathbb{E}[k|x] = \sum_{k\geq 0} kp^{(\mathrm{d})}(k|x) = \theta_q(x) \tag{18}$$

and the average degree is

$$\langle k\rangle = \int_0^1 \mathbb{E}[k|x]\rho(x)dx = \frac{1}{q}. \tag{19}$$

As in the previous section, this can be verified by considering that the average out-degree must equal the number of edges divided by the number of nodes in the network. This fraction is for $t\to\infty$ given by $(qt+(1-q)t)/qt = 1/q$. Fig. 3 shows the comparison of numerical simulations to the analytical prediction of Eq. (17). The agreement between them is excellent. Compared to Fig. 1, the probability $Q^{(\mathrm{d})}(k)$ is broader than $Q^{(\mathrm{s})}(k)$. Thus vulnerable nodes with high degrees are more likely to exist in the dynamic model than in the static model. Another interesting graphical indicator is the degree-conditional probability that a randomly chosen node is vulnerable. This quantity corrects for implications of the degree distribution. Fig. 4(a) displays this probability, given by $Q^{(\mathrm{d})}(k)/P^{(\mathrm{d})}(k)$. Compared to Fig. 2(a), the decay of $Q^{(\mathrm{d})}(k)/P^{(\mathrm{d})}(k)$ occurs at smaller values of k for all different $\langle k\rangle$ configurations. Also, the decay is slower than in the static case and largely influenced by different values of $\langle k\rangle$. Qualitatively, Figs. 4(a) and 2(a) are similar, they both display the phenomenon of adverse diversification, that is discussed above. However, also in the dynamic case is the overall fraction of vulnerable nodes decaying in increasing $\langle k\rangle$, see Fig. 4(b).

III. PERCOLATION

After the main topological property of interest, the degree distribution of vulnerable nodes has been derived in the previous section, this section reviews the methodology that is used to calculate the main properties of the vulnerable component. The approach we take has been used elsewhere [5, 27–29] and is reviewed here for completeness.

First, define the generating function of the probability that a given node has degree k and is vulnerable $Q^{(i)}(k)$, as

$$G_0^{(i)}(s) = \sum_{k\geq 0} Q^{(i)}(k)s^k, \quad i\in\{\mathrm{s,d}\}. \tag{20}$$

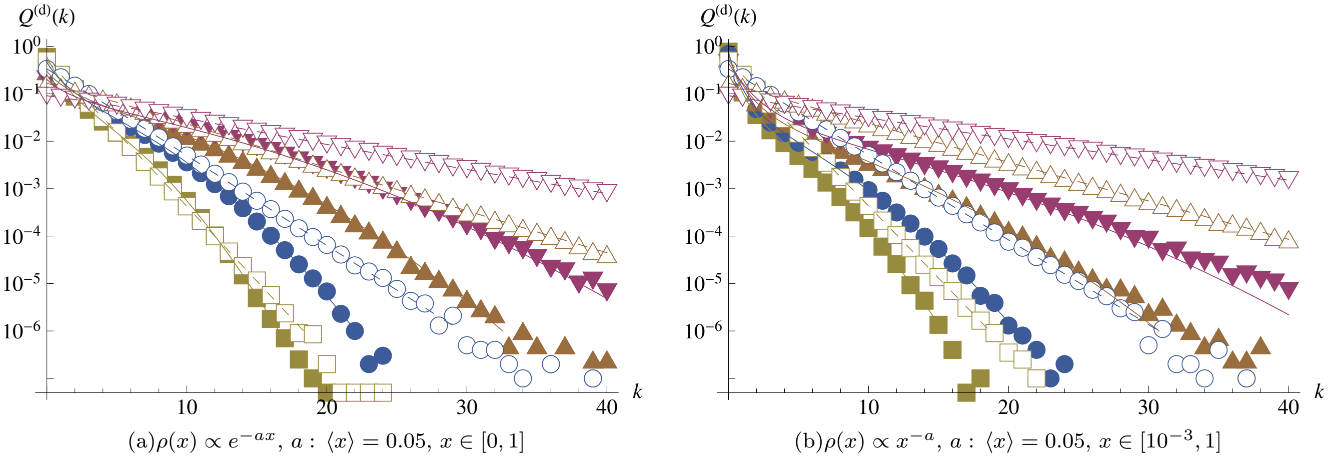


FIG. 3. (Color online) The plots show $Q^{(\mathrm{d})}(k)$ for different fitness distributions: (a) shows different configurations with fitness distributed exponentially on the unit-interval, with average fitness $\langle x\rangle = 0.05$. (b) shows different configurations with Pareto distributed fitness on the interval $[10^{-3}, 1]$, whereby the zero is excluded to avoid the singularity. Random selection, i.e $f(x, y) = 1$ is marked will hollow symbols, mutual attractiveness, i.e. $f(x, y) = xy$ with solid markers. The different plot markers, distinguish different average degrees: $\langle k\rangle = 1$ squares (yellow), $\langle k\rangle = 2$ circles (blue), $\langle k\rangle = 5$ triangles (brown), $\langle k\rangle = 10$ downward triangles (purple). The lines indicate analytical results from Eq. (17), whereby solid lines represent the case of $f(x, y) = xy$ and dashed lines $f(x, y) = 1$. Numerical results are obtained in a network with $N = 10^6$ nodes, averaged over 20 independent network realizations.

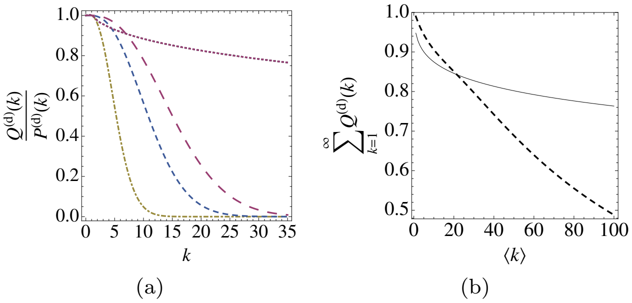


FIG. 4. (Color online) Part (a) shows degree-conditional probabilities of vulnerability for different average degrees $\langle k\rangle = 1$ dot-dashed line (mustard), $\langle k\rangle = 5$ fine dashed (blue), $\langle k\rangle = 10$ coarse dashed (purple). The dotted line is an overlap of all $\langle k\rangle$ configurations with random attachment, i.e. $f(x, y) = 1$. All other lines correspond to $f(x, y) = xy$. The underlying fitness distribution is a power law with $\langle x\rangle = 0.05$. The overlapping lines indicate that for random attachment, the conditional probability to find a vulnerable node is independent of the average degree $\langle k\rangle$. Figure (b) illustrates the relationship between the fraction of vulnerable nodes in the network and the average degree. The underlying configuration is again a power-law on $[10^{-3}, 1]$ with $\langle x\rangle = 0.05$. Solid (dashed) line corresponds to $f(x, y) = xy$ $(f(x, y) = 1)$.

The superscript notation that indicates the specific model is suppressed in the following. Furthermore, the excess degree distribution [29] describes the probability that a randomly chosen neighbor of a randomly picked node is vulnerable and has $k + 1$ neighbors in total (k many neighbors, without the randomly picked node). To be more precise, for this case of a directed graph, it is the probability to find a vulnerable node with out-degree $k + 1$, following a randomly chosen edge against its direction. This probability is proportional to the number of edges that are emitted from vulnerable nodes with degree $k + 1$, correctly normalized with the average degree in the network:

$$R(k) = \frac{(k + 1)Q(k + 1)}{\sum_{k\geq 0} kP(k)}. \tag{21}$$

The normalization is correct, since the node is a random choice taken from the set of all nodes, and not just from the set of vulnerable ones. Also, the generating function for $R(k)$ is defined as

$$G_1(s) = \sum_{j\geq 0} \frac{(j + 1)Q(j + 1)}{\langle k\rangle} = \frac{G_0'(s)}{\langle k\rangle}. \tag{22}$$

Denote the number of nodes that can be reached following only connected nodes along the direction of their edges with t and the distribution of t with $\phi(t)$. The corresponding generating function is defined as

$$H_1(x) = \sum_{t\geq 1} x^t \phi(t). \tag{23}$$

Additionally, $H_0(x)$ is the generating function for the probability that a randomly chosen node belongs to a component of size t. $H_0(x)$ and $H_1(x)$ can be calculated directly, and are given by

$$H_0(x) = 1 - G_0(1) + xG_0(H_1(x)) \tag{24}$$
$$H_1(x) = 1 - G_1(1) + xG_1(H_1(x)). \tag{25}$$

The details of the derivation can be found in Appendix C. The form of these two equations is standard. Equivalent results can also be found in Refs. [27, 30] for example.

The average size of the vulnerable component can be obtained from $H_0'(1)$ as

$$\langle n\rangle = H_0'(1) = \frac{G_0(1) + (G_0'(1))^2}{\langle k\rangle - G_0''(1)}. \tag{26}$$

The phase transition between a finitely sized vulnerable component and a vulnerable component that spans over the entire network can be calculated from Eq. (26). The infinitely sized cluster emerges when $\langle n\rangle$ diverges, thus when $\langle k\rangle = G_0''(1)$. This condition can be written for the static model as

$$\sum_{k\geq 0} \frac{k(k-1)}{\Gamma(k+1)} \int_0^{1/k} e^{-M/N\,\lambda(x)} \left(\frac{M}{N}\lambda(x)\right)^k \rho(x)dx = \frac{M}{N}. \tag{27}$$

For the dynamic model it is given by

$$\sum_{k\geq 0} k(k-1) \int_0^{1/k} \left(\frac{\theta_q(x)}{1+\theta_q(x)}\right)^k \frac{\rho(x)dx}{1+\theta_q(x)} = \frac{1}{q}. \tag{28}$$

Eqs. (27) and (28) define the percolation threshold, in other words for any two of $\rho(x)$, M/N, $\lambda(x)$, respectively $\rho(x)$, q and $\theta_q(x)$ held fixed, these equations define the point of phase transition between an infinitely cluster and a finite percolation in terms of the third variable.

Apart from the percolation threshold, other quantities of interest can also be derived from the theory that is laid out above [27, 28, 30]. Eq. (C8) describes the generating function for the sizes of vulnerable clusters in the network. This generating function is expressed in terms of two other generating functions: $G_0(1)$, which is already calculated above, and $H_1(x)$ which is given implicitly in Eq. (C9). Remember, $H_0(x)$ is the generating function for the distribution of outbreaks outside the percolating cluster, if it exists. The fraction of nodes in the largest vulnerable component S can therefore be computed with $H_0(1)$:

$$S = 1 - H_0(1). \tag{29}$$

S can be computed using Eq. (C8) and solving Eq. (C9) numerically [30]:

$$S = G_0(1) - G_0(\xi), \tag{30}$$

whereby ξ solves

$$\xi = 1 - G_1(1) + G_1(\xi). \tag{31}$$

One trivial solution of Eq. (31) is $\xi = 1$. The function $\varphi(\xi) = G_1(\xi) - \xi + 1 - G(1)$ can maximally have one more root. This follows from the following consideration. The derivative of $\varphi(\xi)$ is $G_1'(\xi) - 1$ and the derivative of the derivative is $G_1''(\xi)$ which is strictly positive by definition of $G_1(\xi)$. Thus the derivative of $\varphi(\xi)$ must be a strictly increasing function. If a second root exists the derivative must be zero somewhere $\in (0,1)$, moreover positive at $\xi = 1$. This condition translates into

$$G_1'(1) - 1 > 0. \tag{32}$$

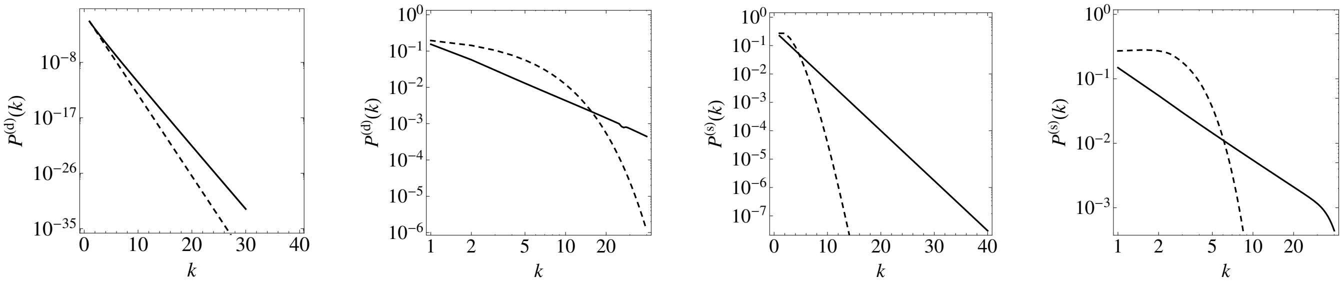
Substituting the definition of $G_1(x)$ leads to

$$G_0''(1) > \langle k\rangle\,. \tag{33}$$

As expected, this condition is equivalent to the percolation condition given in Eqs. (27) and (28). Therefore below percolation, the only solution to Eq. (31) is $\xi = 1$, which translates into $S = 0$, which makes sense since S is the size of the giant component which does not exist below percolation.

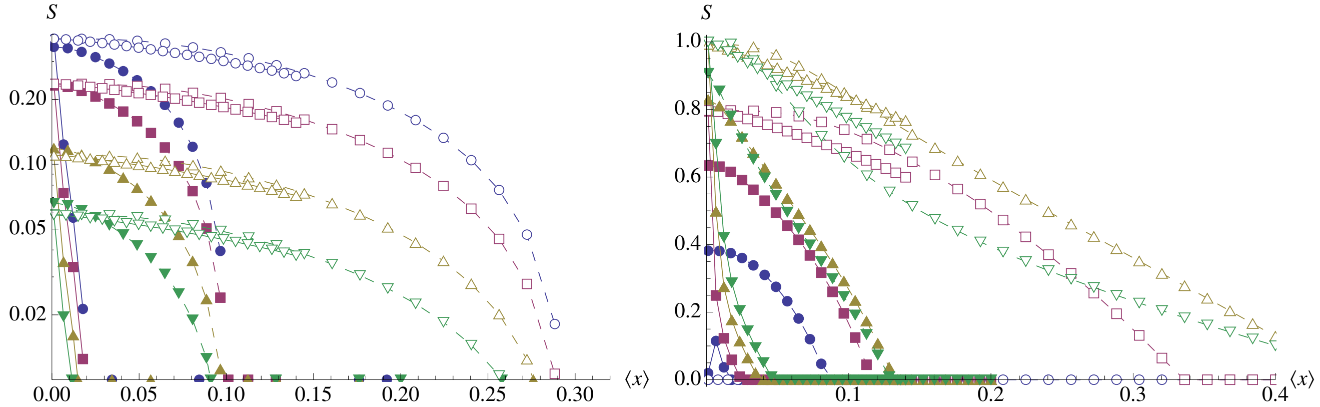
IV. RESULTS

In this section different attachment kernels and two different density distributions are considered. These are mutual and random attachment i.e. $f(x,y) = xy$ and $f(x,y) = 1$, coupled with power-law fitness $\rho(x) \propto x^{-a}$, and $\rho(x) \propto e^{-ax}$. Although correlated mutual attractiveness can possibly be investigated with the technology that is presented here, the evaluation of Eqs. (27) and (28) is very difficult numerically. Therefore, we have decided to consider random attachment compared to the most simple mutual attractiveness $f(x,y) = xy$. The two specific fitness distributions have been chosen since they allow an interesting comparison between various network motifs. Power-law probability distributions can be found in many areas of scientific interest [6, 31]. But also exponential distributions are justifiable in certain scenarios and can be found in real data [32]. Exponential distributions are considered to illustrate the effect of heavier tails on percolation in this class of networks. While the theory in the body of this article is derived for fitness distributed over the unit interval, the domain of the power-law is constrained to $[10^{-3}, 1]$ for the numerical evaluation in order to avoid the singularity at $x = 0$. The form of the resulting degree distribution of these networks follows two distinct patterns, power-law and exponential. This is illustrated in Fig. 5. It is clear from Fig. 5 that most configurations lead to an exponential or exponentially decaying degree distribution. However, as is elucidated below, the resulting spreading behavior is entirely different from case to case. Notice as well the different behavior between the dynamic and the static fitness model. While the dynamic model produces a power-law degree distribution over the whole range of k for $f(x,y) = xy$ and power-law fitness, the static model produces a power law with cutoff at large k. The fact that many configurations lead to similar degree distributions, but display an entirely different failure spreading behavior highlights the importance of the distinction between static and dynamic fitness models.



(a)$\rho(x) \propto e^{-ax}$ (b)$\rho(x) \propto x^{-a}$ (c)$\rho(x) \propto e^{-ax}$ (d)$\rho(x) \propto x^{-a}$

FIG. 5. Degree distributions for the different configurations that are under consideration. Dashed lines for random attachment-ment $f(x, y) = 1$ and solid lines for mutual attractiveness $f(x, y) = xy$ with $\langle x \rangle = 0.05$ and $\langle k \rangle = 2$ in all cases. (a) Dynamic model, exponential fitness distribution. (b) Dynamic model, power-law fitness. (c) Static model, exponential fitness, (d) Static model, power-law fitness.



(a)Dynamic model (b)Static model

FIG. 6. (Color online) Occupied fraction of the percolating vulnerable component in the dynamic (a) and static (b) network. (a) is on a log-linear scale for improved magnification of the chart area of interest. The different plot character indicate different average degrees: ○ (blue) – $\langle k \rangle = 1$, □ (magenta) – $\langle k \rangle = 2$, △ (mustard) – $\langle k \rangle = 5$, and ▽ (green) – $\langle k \rangle = 10$. The different attachment kernels are indicated with solid plot characters for $f(x, y) = xy$ and hollow characters for $f(x, y) = 1$. Finally, the two different fitness distributions are distinguished by line type. Solid line for $\rho(x) \propto x^{-a}$ and dashed line for $\rho(x) \propto e^{-ax}$.

Fig. 6 illustrates analytical results on the vulnerable fraction S depending on the average fitness $\langle x \rangle$ and average degree $\langle k \rangle$. Results for the dynamic and the static model are displayed separately. The differences between these two models are not only quantitative, but also of qualitative nature. The percolation threshold – if it exists – is (other things being equal) independent of the average node degree in the dynamic model, but differs in the static model. Additionally, the static network can become entirely occupied by the vulnerable component. This is not possible in the dynamic case. There is one artifact that does not derive from the topological properties of the network. In the case of power-law distributed fitness, coupled with random attachment, the network can never obtain a state in which the vulnerable component vanishes. This is an artifact that comes from

$$\sup_a \left\{ \int_{10^{-3}}^{1} x\rho(x; a)dx \right\} = 0.145..., \quad \rho(x; a) \propto x^{-a}. \tag{34}$$

This supremum of average fitness is not sufficient to stabilize the network. The supremum is higher for the case of exponential distributed fitness, which leads to the misleading finding that for the otherwise same configuration, the exponential fitness induced network is more stable than its power-law induced counterpart. For the case of mutual selection (solid lines), the power-law induced network (circles) is significantly more robust than the exponential network (triangles). This is evident from the onset of the large vulnerable component that occurs only at very small $\langle x \rangle$ in the case of mutual selection. A further aspect that appears in Fig. 6 is that the introduction of mutual attractiveness stabilizes the network. The percolation threshold, that marks the onset of the giant vulnerable component is notably larger in terms of $\langle x\rangle$ for random attachment. The case of the static network with $\langle k\rangle = 1$ appears to be special as percolation never occurs. This is due to the lack of overall connectivity and not caused by anything inherent to the model. A further particularity revealed in Fig. 6 is that the order of sizes of the percolating component for the same f and ρ differ for increasing $\langle k\rangle$. Take for instance the case of $f(x,y) = xy$ and $\rho(x) \propto e^{-ax}$, represented by solid lines and triangles. In the case of the dynamic model, the configuration with $\langle k\rangle = 10$ appears to be most stable, while it is the second to last stable configuration in the case of the static network. Similarly $\langle k\rangle = 1$ is most stable in the static network, and most unstable for the dynamic case. However, these comparisons are only of ordinal nature, notice that the size of the vulnerable component is significantly larger in the static network compared to the dynamic setting.

Consider additionally to the attachment kernels that were illustrated in Fig. 6, also

$$f_1(x,y) = e^{x-y} \quad \text{and} \quad f_2(x,y) = e^{y-x}. \tag{35}$$

f_1 is leading to a scenario in which exposure is flowing from fit nodes to weak ones, while f_2 induces the reverse, a network that favors weak nodes to be exposed to strong ones. The percolation thresholds differ significantly in these two scenarios. We compare configurations with exponential fitness, because the power-law distribution does not reach sufficient average fitness to stabilize the system induced by f_1 and f_2 with $q = 1$. With exponential distributed fitness, the percolation threshold for f_1 is given by $\langle x\rangle^* = 0.213...$, for f_2, it is $\langle x\rangle^* = 0.443....$ This shows that a system in which exposure flows from strong participants to weak ones is significantly more stable than a system that favors exposure flow from weak to strong. This makes sense. Consider for instance the banking sector with its flow of credit exposure. It is much more important for the system's stability that the exposed parties (creditors) are robust, that it is that the exposing side (debtor) is robust. Low fitness nodes are less likely to propagate shocks, since they are less well connected. This also means that in the case of retail banking, it is more important for the system as a whole to guarantee a strong financial position of the main participants than to employ a very rigorous customer scrutinization procedure.

V. CONCLUSIONS

We have discussed the coupling of propensity to form edges in the network with the robustness of failures using one variable called fitness. The assumption of this coupling is realistic in many scenarios because the robustness – be it financial or health – is directly related to the attractiveness of an individual in a contact process. The resulting theory can help to devise immunization strategies not through direct rewiring or protecting single nodes but rather by incentivizing individuals to form a more stable structure. It is shown in this paper that the introduction of a tax, or a stimulus to form new edges within the network can stabilize the network without violating individual preferences that are described by $f(x,y)$. Another aspect that has been illustrated in this article is the importance of the distinction between dynamic and static fitness models. The percolation behavior of these two network classes is entirely different. The percolation threshold in the dynamic model is only influenced by the attachment kernel and the distribution of fitness, while the results for the static model show that additionally the average degree has a direct impact on the onset of the percolating cluster. Furthermore, the static network can be entirely occupied with the percolating component, while this is impossible in the dynamic model. Also, the broadness of the fitness distribution has an effect on the stability of the network. A Wide variety of fitness among individuals stabilizes the network. Moreover, the attachment kernel has a profound influence on the way an epidemic spreads. Mutual attractiveness, i.e. $f(x,y) = xy$ induces a network that is significantly more resilient than a network composed by pure random attachment.

In this paper, the networks we investigated were flat so that shocks could only propagate on one layer. For future research, it would be interesting to investigate how a multiplex architecture with similar rules behaves and whether the findings are comparable to ones that were found here.

Appendix A

The recurrence relation in Eq. (3) can be solved using a generating function approach. Define

$$F^{(\mathrm{s})}_{M,N}(s|x) = \sum_{k\geq 0} p^{(\mathrm{s})}_{M,N}(k|x)s^k. \tag{A1}$$

Multiplying Eq. (3) with s^k and summing over k leads to

$$F^{(\mathrm{s})}_{M+1,N}(s|x) - F^{(\mathrm{s})}_{M,N}(s|x) = F^{(\mathrm{s})}_{M,N}(s|x)\lambda(x,N)(s-1). \tag{A2}$$

For large M, Eq. (A2) can be approximated as an ordinary differential equation in M with initial condition $F^{(\mathrm{s})}_{0,N}(s|x) = 1$. The initial condition arises from the observation that in a network without edges, the conditional degree distribution is peaked at zero: $p^{(\mathrm{s})}_{0,N}(k|x) = \delta_{k0}$, where δ_{xy} is the Kronecker Delta. Using this, and that $1/N$ is a factor of $\lambda(x,N)$, leads to

$$F_M(s|x) = e^{M/N\,\lambda(x)\,(s-1)}. \tag{A3}$$

Expanding Eq. (A3) in s around $s = 0$ leads to

$$p^{(\mathrm{s})}_{M,N}(k|x) = \frac{e^{-M/N\,\lambda(x)}(\frac{M}{N}\lambda(x))^k}{\Gamma(k+1)}. \tag{A4}$$

Appendix B

Eq. (14) can be solved using a generating function approach that has already been used in the previous appendix. Define

$$F_t^{(\mathrm{d})}(s|x,\tau) = \sum_{k\geq 0} p_t^{(\mathrm{d})}(k|x,\tau)s^k. \tag{B1}$$

Multiplying Eq. (14) with s^k and summing over k leads to

$$\begin{aligned} F_t^{(\mathrm{d})}(s|x,\tau) = F_t^{(\mathrm{d})}(s|x,\tau)(1-\theta_q(x,t)) \\ + \theta_q(x,t)s\, F_t^{(\mathrm{d})}(s|x,\tau). \end{aligned} \tag{B2}$$

Since a node enters the network with one inward pointing edge, the initial condition is given by $p_t^{(\mathrm{d})}(k|x,t) = \delta_{k0}$, so that $F_t^{(\mathrm{d})}(s|x,t) = 1$. Using this initial condition and that $1/t$ is a factor of $\theta_q(x,t)$, the solution of Eq. (B2) is given by

$$F_t^{(\mathrm{d})}(s|x,\tau) = \left(\frac{t}{\tau}\right)^{\theta_q(x)(s-1)}. \tag{B3}$$

Averaging over entry times leads to

$$\begin{aligned} F^{(\mathrm{d})}(s|x) &= \int_0^t \left(\frac{t}{\tau}\right)^{\theta_q(x)(s-1)} \frac{d\tau}{t} && \text{(B4)} \\ &= \frac{1}{1+\theta_q(x)(1-s)}. && \text{(B5)} \end{aligned}$$

The expansion of this expression in s around $s=0$ leads to the stationary conditional degree distribution

$$p^{(\mathrm{d})}(k|x) = \left(\frac{\theta_q(x)}{1+\theta_q(x)}\right)^k \frac{1}{1+\theta_q(x)}. \tag{B6}$$

Appendix C

Define the probability that a node with out-degree k belongs to a vulnerable component of size c as $\zeta(c|k)$. $\zeta(c|k)$ can be derived constructively by noting that the sum of nodes that can be reached following each of the nodes edges must sum up to $c-1$:

$$\zeta(c|k) = \sum_{t_1\geq 1}\cdots\sum_{t_k\geq 1}\delta\left(c-1, \sum_{m=1}^k t_m\right)\prod_{m=1}^k \phi(t_m) \tag{C1}$$

Where $\delta(x,y)$ is the Kronecker Delta. Now denote the probability that a randomly chosen node belongs to a vulnerable component of size c with $\pi(c)$. This is simply

$$\pi(c) = \sum_{k\geq 1} Q(k)\zeta(c|k). \tag{C2}$$

The generating function for $\pi(c)$ can be computed in the same way as for example outlined in Ref. [29]:

$$\begin{aligned} \sum_{c\geq 1}\pi(c)x^c &= \sum_{c\geq 1} x^c \sum_{k\geq 1} Q(k)\zeta(c|k) && \text{(C3)} \\ &= \sum_{k\geq 0} Q(k)\sum_{c\geq 1} x^c \sum_{t_1\geq 1}\cdots\sum_{t_k\geq 1}\delta\left(c-1,\sum_{m=1}^k t_m\right)\prod_{m=1}^k\phi(t_m) && \text{(C4)} \\ &= x\sum_{k\geq 0} Q(k)\sum_{t_1\geq 1}\cdots\sum_{t_k\geq 1} x^{\sum_{m=1}^k t_m}\prod_{m=1}^k \phi(t_m) && \text{(C5)} \\ &= x\sum_{k\geq 1} Q(k)\left(\sum_{t\geq 1}\phi(t)x^t\right)^k && \text{(C6)} \\ &= x\sum_{k\geq 1} Q(k)(H_1(x))^k && \text{(C7)} \end{aligned}$$

The result of the calculation above is the generating function of a vulnerable component of size one or greater. However, the possibility that the initially chosen node is not vulnerable, hence the random choice selected an empty vulnerable component of size zero must also be taken into account. This probability is given by $1 - \sum_{k>0} Q(k)$, hence the generating function for the probability that a randomly chosen node is part of a vulnerable component is given by

$$H_0(x) = 1 - G_0(1) + xG_0(H_1(x)). \tag{C8}$$

The generating function for $H_1(x)$ can be established in a similar way and is given by

$$H_1(x) = 1 - G_1(1) + xG_1(H_1(x)). \tag{C9}$$

[1] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg, Nature **411**, 907 (2001).

[2] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli, Scientific reports **2** (2012).

[3] T. Roukny, H. Bersini, H. Pirotte, G. Caldarelli, and S. Battiston, Scientific Reports **3** (2013).

[4] S. Cauchemez, A. Bhattarai, T. L. Marchbanks, R. P. Fagan, S. Ostroff, N. M. Ferguson, D. Swerdlow, and the Pennsylvania H1N1 working group, Proceedings of the National Academy of Sciences **108**, 2825 (2011).

[5] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).

[6] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[7] M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001).

[8] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000).

[9] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[10] M. A. Serrano and M. Boguñá, Phys. Rev. Lett. **97**, 088701 (2006).

[11] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature **464**, 1025 (2010).

[12] M. Dickison, S. Havlin, and H. E. Stanley, Phys. Rev. E **85**, 066109 (2012).

[13] D. Zhou, J. Gao, H. E. Stanley, and S. Havlin, Phys. Rev. E **87**, 052812 (2013).

[14] Y. Hu, D. Zhou, R. Zhang, Z. Han, C. Rozenblat, and S. Havlin, Phys. Rev. E **88**, 052805 (2013).

[15] C. Granell, S. Gómez, and A. Arenas, Phys. Rev. Lett. **111**, 128701 (2013).

[16] X.-L. Peng, X.-J. Xu, X. Fu, and T. Zhou, Phys. Rev. E **87**, 022813 (2013).

[17] G. Caldarelli, A. Capocci, P. De Los Rios, and M. Muñoz, Physical review letters **89**, 258702 (2002).

[18] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **68**, 036112 (2003).

[19] V. D. Servedio, G. Caldarelli, and P. Butta, Physical Review E **70**, 056126 (2004).

[20] I. E. Smolyarenko, K. Hoppe, and G. J. Rodgers, Phys. Rev. E **88**, 012805 (2013).

[21] K. Hoppe and G. J. Rodgers, Phys. Rev. E **88**, 042804 (2013).

[22] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[23] G. Bianconi and A.-L. Barabási, EPL (Europhysics Letters) **54**, 436 (2001).

[24] G. De Masi, G. Iori, and G. Caldarelli, Phys. Rev. E **74**, 066112 (2006).

[25] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 188701 (2004).

[26] D. Garlaschelli and M. I. Loffredo, Physica A: Statistical Mechanics and its Applications **355**, 138 (2005).

[27] D. J. Watts, Proceedings of the National Academy of Sciences **99**, 5766 (2002).

[28] C. Moore and M. E. J. Newman, Phys. Rev. E **62**, 7059 (2000).

[29] M. E. J. Newman, Phys. Rev. E **76**, 045101 (2007).

[30] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).

[31] A. Clauset, C. R. Shalizi, and M. E. Newman, SIAM review **51**, 661 (2009).

[32] A. Drăgulescu and V. Yakovenko, Physica A: Statistical Mechanics and its Applications **299**, 213 (2001).