Percolation of Hierarchical Networks and Networks of Networks

Louis M. Shekhtman[1] and Shlomo Havlin[1]

[1]*Department of Physics, Bar-Ilan University, Ramat Gan, Israel*

(Dated: May 7, 2018)

Much work has been devoted to studying percolation of networks and interdependent networks under varying levels of failures. Researchers have considered many different realistic network structures, but thus far no study has incorporated the hierarchical structure of many networks. For example, infrastructure across cities will likely be distributed such that nodes are tightly connected within small neighborhoods, somewhat less connected across the whole city, and have even fewer connections between cities. Furthermore, while previous work identified interconnected nodes, those nodes with links outside their neighborhood, to be more likely to be attacked, here we have various levels of interconnections (between neighborhoods, between cities, etc.). We consider the nodes with interconnections at the highest level most likely to be attacked, followed by those with interconnections at the next level, etc. We develop an analytic solution for both single and interdependent networks of this structure and verify our theory through simulations. We find that depending on the number of levels in the hierarchy there may be multiple transitions in the giant component (fraction of interconnected nodes), as the network separates at the various levels. Our results show that these multiple jumps are a feature of hierarchical networks and can affect the vulnerability of infrastructure networks.

INTRODUCTION

The robustness of infrastructure systems can be understood through the frameworks of complex networks, percolation, and interdependent networks [1–10]. The initial research on network robustness was later expanded to include various network structures such as various degree distributions [11–13], clustering [14, 15], spatial embedding [16–18], and quite recently, community structure [19, 20]. Further, additional research has considered various types of attacks on these networks such as degree-based attacks [21, 22], localized attacks [23, 24], and attacks based on nodes linking across communities [19, 20, 25].

However, other common network structures that are likely relevant for robustness have not yet been studied. Among these is a hierarchical structure, which we will study here, where communities connect loosely with one another to form larger communities and so on [26–28] (See Fig. 1). In the context of infrastructure robustness these overlapping modules are likely described through neighborhoods overlapping to form cities, which then overlap to form states, etc. which are then interconnected among themselves.

Furthermore, in this model, the nodes at the highest level of the hierarchy (e.g. between states) are likely more vulnerable to failure or attack than those at the next highest level, which are in turn more vulnerable than those at an even lower level etc. This is because the nodes at higher levels have longer distance links between them which are more likely to fail or be attacked [29] and also have higher betweenness [19] which yields additional load on them [30, 31]. Also, recent work by da
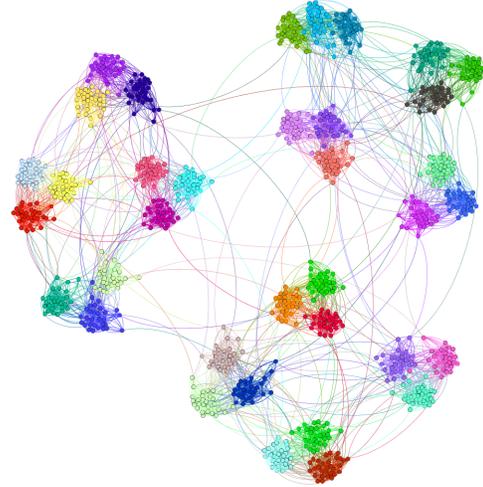

FIG. 1. **Model Illustration.** For this realization, the model has 4 hierarchical layers. At the top layer there are 3 modules, each of which is broken down into 4 modules, each of which is then broken down into 3 modules, which are not broken down further. We could describe this configuration of modules by the vector $\vec{m} = [1, 3, 12, 36]$

Cunha et al. [25] showed that attacks on these 'interconnected nodes' are an optimal form of attack on the US power grid, an infrastructure system of critical interest.

MODEL

Our model is a stochastic block model [32–35] with overlap among the various blocks.

We first define the vector $\vec{m}$ describing the number of distinct modules or communities (blocks) at each layer. At the first layer we always consider the entire network as a single community, thus $m_1 = 1$. The next layer, m_2 counts how many modules are at the top layer. Next is the total number of modules at the third layer, next is the modules at the fourth layer, etc (we also assume for simplicity that all of the m_j modules are broken down into the same fixed number of m_{j+1} modules). For example, if we take the network shown in Fig. 1, we would say $\vec{m} = [1, 3, 12, 36]$, since the top layer is a connected graph, at the next layer we have three modules, then a total of 12 modules (i.e. each of the three is broken down into four smaller modules and $3 \times 4 = 12$), and finally 36, since each of the 12 modules is broken down into three additional ones.

We next define the vector $\vec{k}$, which describes the average degree between nodes connected at each layer of the network. Thus, if at the highest layer there is an average of 0.1 links between each module, this will be the first entry, k_1 in $\vec{k}$. If the average degree at the next layer is 0.3 then that will be the second entry, k_2 etc. We assume that the entries of $\vec{k}$ should be strictly increasing since we expect there to be more links within communities at a lower layer than at a higher layer (e.g. neighborhoods are more tightly connected than cities).

We will carry out a targeted attack on the nodes of the network, assuming that nodes that are interconnected at the top level are most likely to fail. To do so, we must determine how many nodes are connected at each level and convert from their survival likelihood to the overall survival likelihood. We can estimate how many nodes are connected at level i as $1 - e^{-k_i}$ [19]. We then define r_i as the survival probability of interconnected nodes at layer i. For the top layer of interconnections we can convert from r_1, the survival probability of interconnected nodes at the highest level, to p, the overall survival probability, using [19]

$$r_1 = \frac{p - e^{-k_1}}{1 - e^{-k_1}}. \tag{1}$$

After we have removed all nodes with interconnections at the top level, we then begin removing those nodes with interconnections at the next level. In order to convert from the survival probability of nodes at this next level, r_2 and the new overall survival probability p, we must take into account those nodes removed at the previous layer.

We do so by first finding the value of p for which we have removed all nodes at a given layer. For example, for the first layer the cutoff for which all nodes with in-

terconnections at this layer are removed is $p_{co_1} = e^{-k_1}$. For the next layer the cutoff is given by recognizing that the number of nodes with interconnections at this next layer is e^{-k_2}, but we are already at a survival probability of only p_{co_1} and also some of these nodes also had interconnections at the previous layer. Thus taking these into account gives that the cutoff of p at the second layer, p_{co_2} is given by the inclusion-exclusion principle as

$$p_{co_2} = 1 - \left(1 - e^{-k_2} + 1 - e^{-k_1} - (1 - e^{-k_2})(1 - e^{-k_1})\right)$$
$$= e^{-k_2 - k_1}. \tag{2}$$

Thus the cutoff value of p, for a given level i is given by

$$p_{co_i} = e^{-\sum_{j=1}^{i} k_j}. \tag{3}$$

We can then convert from r_i, the survival probability in level i (after having removed all nodes in higher layers), to p using

$$r_i = \frac{p - p_{co_i}}{p_{co_{i-1}} - p_{co_i}}. \tag{4}$$

ANALYTIC THEORY FOR A SINGLE NETWORK

Having made these conversions, we can now find the size of the giant component after some fraction of nodes are removed. For the top level, we can extend previous results on modular networks by setting our average interconnected degree to the degree at the level we are attacking and setting our average intra-degree to the sum of the degrees at the lower levels [19, 20]. We thus obtain

$$P_\infty = e^{-k_1}(1 - r_1)\left(1 - e^{-\left(\sum_{i=2}^{l} k_i\right)P_\infty}\right)$$
$$+ r_1\left(1 - e^{-\left(\sum_{i=1}^{l} k_i\right)P_\infty}\right), \qquad p > p_{co_0}. \tag{5}$$

Once we have removed all interconnected nodes in the first level, we then move on to removing nodes that are interconnected at the second level. In this case, the average degree of interconnections is now k_2 and the average degree of intraconnections is $\sum_{i=3}^{l} k_i$, where l is the number of levels. Furthermore, for this case we will have removed some fraction $1 - r_2$ of nodes based on targetted attack of interconnected nodes at this layer, and also p_{co_1} nodes randomly due to the attacks at the above layer. Additionally, we must recall that the network at this stage is already split into m_2 separate modules thus we must scale
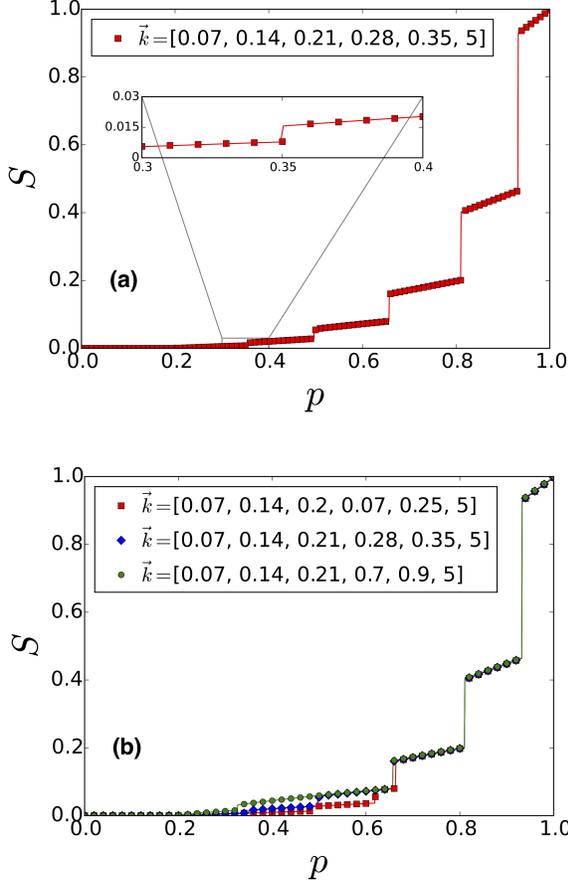
FIG. 2. **Comparison of simulation results with theory.** In both figures we have six layers, where each layer splits a module into two other modules, thus there are $\vec{m} = [1, 2, 4, 8, 16, 32]$ modules at each layer, with the average degrees between nodes at each layer given by the vector $\vec{k}$ in the legend. We have different values for the degree at each layer as given in the legends of **(a)** and **(b)**. The lines represent the theory outlined above and the points are simulations averaged over 10 runs on networks of $N = 10^6$ nodes.

P_∞ by $1/m_2$. This gives

$$m_2 P_\infty = p_{co_1}\left[e^{-k_2}(1-r_2)\left(1-e^{-\left(\sum_{i=3}^l k_i\right)m_2 P_\infty}\right)\right.$$

$$\left.+ r_2\left(1 - e^{-\left(\sum_{i=2}^l k_i\right)m_2 P_\infty}\right)\right], \qquad p_{co_1} > p > p_{co_2}. \quad (6)$$

In general for all values of p we can find the size of the largest connected component, P_∞ using

$$m_j P_\infty = p_{co_{j-1}}\left[e^{-k_j}(1-r_j)\left(1-e^{-\left(\sum_{i=j+1}^l k_i\right)m_j P_\infty}\right)\right.$$

$$\left.+ r_j\left(1 - e^{-\left(\sum_{i=j}^l k_i\right)m_j P_\infty}\right)\right], \qquad p_{co_{j-1}} > p > p_{co_j}. \quad (7)$$

We compare theory and simulations in Fig. 2, observing excellent agreement between them. Also, in the figure we observe multiple discontinuities [36, 37] as once all interconnected nodes in a particular layer are removed, the system experiences a discontinuous jump.

Number of Abrupt Jumps

We next compute the expected number of jumps that will take place using our analytic theory from above. We begin by noting that jumps will occur when all interconnected nodes at a particular layer are removed, yet there remain enough total surviving nodes such that the hierarchy of modules at the next layer remains connected. This condition can be expressed by recognizing that we need the value of p of the remaining intralinks to be lower than the value of p for which all interconnected nodes at this given level are removed. This condition is given by for a given level i

$$e^{-k_i} \geq \frac{1}{\sum_{j=i+1}^l k_j}. \quad (8)$$

Assuming that the degree at each level of hierarchy is strictly decreasing, then e^{-k_i} is strictly decreasing. Furthermore, assuming all $k_i > 0$ implies that $1/\sum_{j=i+1}^l k_j$ is strictly increasing as i increases (since the denominator must decrease as there are fewer k_j terms). Therefore, once the condition of Eq. (8) is first violated for a particular level, we know that it will continue to break down for later levels and thus we can be sure that our number of jumps is the number of levels for which Eq. (8) is valid.

We plot the two sides of Eq. (8) in Fig. 3a, where for $l \leq 5$ we see that the left-hand-side (LHS) of the equation is larger than the right-hand-side (RHS). Comparing to the number of jumps in Fig. 2a we see that the network indeed experiences 5 abrupt jumps as expected (see inset for the 5th jump).

The p Values of the Jumps

Having found the number of jumps that the network will undergo above, we can now analyze the multiple

values of p_c, the critical thresholds at which the jumps occur. We first note that so long as the LHS of Eq. (8) is greater than the RHS of the same equation, then there will be a transition at the point of the LHS of the equation. After these i transitions, there will be one final $i+1$st transition [20] at the point r_{i+1}, which can be found by solving the below equation for r_{i+1},

$$r_{i+1}^2 \left(p_{co_i}^2 \left(\sum_{j=i+2}^{l} k_j \right) k_{i+1} e^{-k_{i+1}} \right) + r_{i+1} \left(p_{co_i} k_{i+1} + p_{co_i} \left(\sum_{j=i+2}^{l} k_j \right) - p_{co_i} \left(\sum_{j=i+2}^{l} k_j \right) e^{-k_{i+1}} - p_{co_i}^2 \left(\sum_{j=i+2}^{l} k_j \right) k_{i+1} e^{-k_{i+1}} \right)$$

$$+ \left(p_{co_i} \left(\sum_{j=i+2}^{l} k_j \right) e^{-k_{i+1}} - 1 \right) = 0. \quad (9)$$

After finding r_{i+1} we can convert it to a value of p using Eq. (4). We note a slight subtlety in this system, in that even for the case where the hierarchical network is completely isolated at the lowest level we do not precisely recover percolation on a random network since we are targeting only those nodes which began with at least one link in the construction of the network. This leads to a slight correction where we obtain $r_{i+1} = 1/k_{i+1}$ (and then convert this to a value for p_c), rather than obtaining the usual $p_c = 1/k_{i+1}$. In most cases this correction will be quite small as for any reasonable value of k at the lowest level, there will be very few nodes that do not have a single link. For example, for the case of the network referred to by the top line of the legend in Fig. 3b, the transition for the 6th layer takes place at $p_c \approx 0.201$ as opposed to $1/k_6 = 0.2$.

INTERDEPENDENT NETWORKS

Much recent research has also explored the resilience of interdependent networks where the nodes of one network depend on nodes in another network [5, 38–45]. One example is that of a communication network that is interdependent with a power grid, yet more complex interdependencies are also possible [46, 47]. Many of these interdependent networks will likely possess the hierarchical structure described above. Therefore we now extend our theory to the case of networks of interdependent networks (NON).

We will assume that each network is formed of the same hierarchical structure, i.e. there are the same number of modules at each level. Again, this is intuitive since the number of cities, neighborhoods, etc. that exist for the power grid are likely the same as those for a communications network. Further, we will assume that nodes are dependent on other nodes within their same module at the lowest level. This corresponds to the assumption that nodes are most likely dependent for resources from nodes in their same neighborhood, i.e. a power station depends on a communication tower in the same neighborhood.

In the case of interdependent networks formed of n networks with $n > 2$, the structure of the dependencies can take various shapes. Among these are both treelike structures, where networks depend on one another such that their dependencies form a tree, or looplike structures where the dependencies form loops. Here we will consider (i) treelike structures and (ii) a random-regular (RR) network of networks where each network depends on exactly w other networks.

Treelike Network Formed of Hierarchical Networks

Here we introduce the theory for a network of networks formed of n interdependent networks such that they form a tree. For the interdependent case, Eqs. (4)-(5) remain valid as for the single-network case.

For the treelike NON we assume that all nodes between pairs of interdependent networks have a dependent node. Further, we assume the no-feedback condition, meaning that if node a in network n_1 depends on node b in network n_2, then node b also depends on node a. Lastly, we will attack the nodes of only one of the networks and let the attack propagate to the other networks.

To include the effects of the interdependencies, we note that we must add an additional likelihood of failure based on the interdependence. For a treelike network of n interdependent networks with the dependencies within the neighborhoods, this term is $\left(1 - e^{-\left(\sum_{i=j}^{l} k_i \right) m_j P_\infty} \right)^n$ [20,
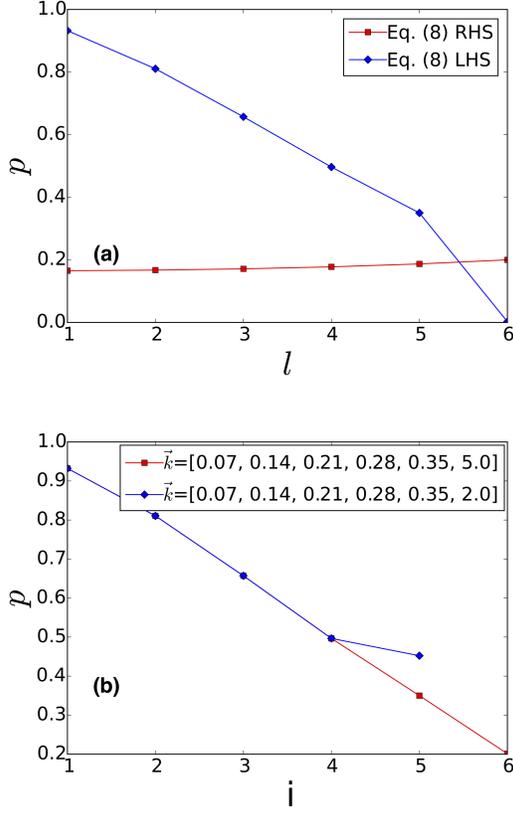
FIG. 3. **(a)** The RHS and LHS of Eq. (8) for the network of Fig. 2a. We observe that for the first $l = 5$ layers the value of p for which all interconnected nodes are removed (LHS of Eq. (8)) is greater than the value at which the network connectivity as a whole breaks down (RHS of Eq. (8)). However for the 6th layer this is no longer true and we observe the continuous percolation transition of a random network. **(b)** The values of the ith critical points for both the network described in Fig. 2a and another network which has a smaller degree at its bottom level. We note how changing the degree at the bottom level effects the number of jumps since for the network with a lowest level average degree of 2, before removing all interconnected nodes at the 5th layer, the network already breaks apart.

43]. We then multiply this term with the other terms of Eq. (7) to obtain

$$
\begin{aligned}
m_j P_\infty = p_{co_{j-1}} \Bigg[& e^{-k_j}(1 - r_j)\left(1 - e^{-\left(\sum_{i=j+1}^{l} k_i\right)m_j P_\infty}\right) \\
& + r_j\left(1 - e^{-\left(\sum_{i=j}^{l} k_i\right)m_j P_\infty}\right)\Bigg]\left(1 - e^{-\left(\sum_{i=j}^{l} k_i\right)m_j P_\infty}\right)^{n-1},
\end{aligned}
$$

$$
p_{co_{j-1}} > p > p_{co_j}. \quad (10)
$$

We note that numerical simulations show excellent agreement with the theory (Fig. 4). We also note that in the case of interdependent networks, the final transition is now discontinuous due to the interdependence [5, 38, 43].
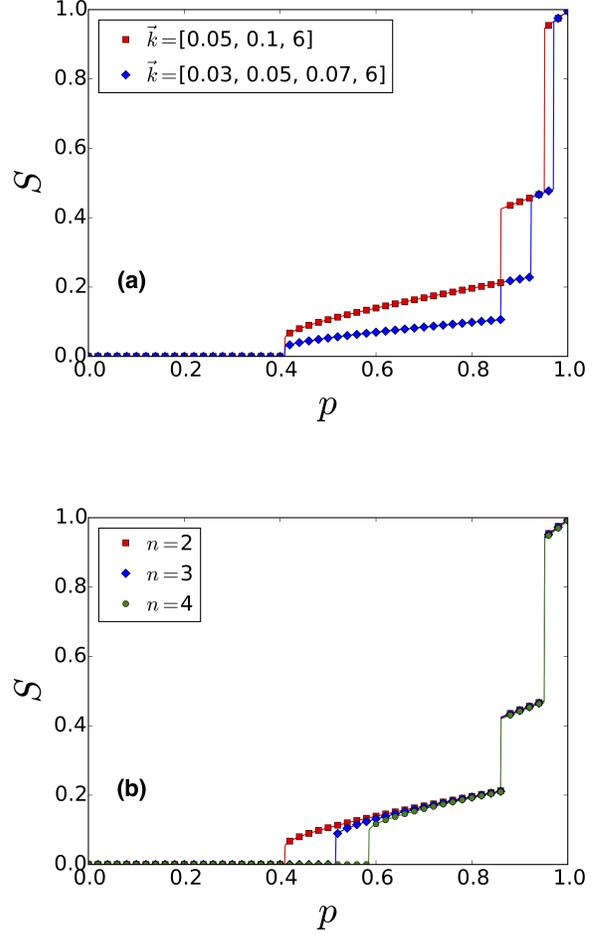

FIG. 4. **(a).** The case of two interdependent networks with the values of degree at each layer as given in the legend. Points are simulations averaged over 10 realizations of networks with $N = 10^6$ nodes and lines are theory. **(b).** Varying the number of networks with the degree vector fixed to $\vec{k} = (0.05, 0.1, 6)$.

Random Regular Network Formed of Hierarchical Networks

Lastly, we consider the case of an RR NON where each network depends on exactly w other networks. We will assume that for each pair of interdependent networks only a fraction q of the nodes are interdependent and we

5

will allow feedback (in contrast to what was done for treelike NON). However, we will still restrict the dependencies such that they must be within the same community at the lowest level of the hierarchy.

In this case, the effects of the dependencies result in a reduction of the size of the giant component by a factor of $\left(1 - q + q m_j P_\infty\right)^w$ [43, 48]. Combining this with Eq. (7) yields,

$$P_\infty = p_{cutoff_{j-1}} \left[\frac{1}{m_j} e^{-k_j}(1 - r_j)\left(1 - e^{-\left(\sum_{i=j+1}^{n} k_i\right)m_j P_\infty}\right) \right.$$
$$\left. + r_j\left(1 - e^{-\left(\sum_{i=j}^{l} k_i\right)m_j P_\infty}\right)\right]\left(1 - q + q m_j P_\infty\right)^w,$$
$$p_{cutoff_{j-1}} > p > p_{cutoff_j}. \quad (11)$$

For the RR NON the final transition will be continuous (depending on the value of q) as for an RR NON formed of Erdős-Rényi networks [48]. We observe excellent agreement between the theory of Eq. (11) and simulations in Fig. 5 including the prediction regarding the nature of the transitions.

DISCUSSION

In this work, we have studied the robustness of networks and networks of interdependent networks with a hierarchical structure. This structure is very common for many infrastructure networks, biological networks and others. We have found analytical solutions and confirmed these solutions through simulations for isolated hierarchical networks and two different structures of interdependent hierarchical networks. The resilience of the network depends on the number of communities at each level of the hierarchy, the degree at each level of the hierarchy, the fraction of nodes removed, and also the parameters governing the interdependence (if present).

Our results show that hierarchical networks can undergo multiple abrupt transitions depending on the above parameters and that these transitions represent the separation of the network at different levels of the hierarchy. These results have potential applications in optimizing the resilience of networks in infrastructure and other fields.
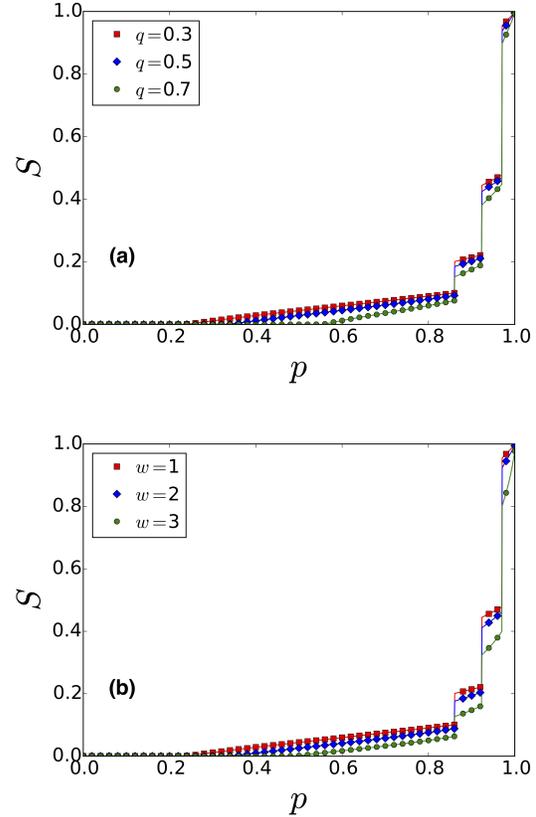
FIG. 5. A random regular network of networks where each network depends on w other networks such that they form loops. We vary both **(a)** q the level of interdependence between the networks (with $w = 1$) and **(b)** w the number of networks each network depends on (with $q = 0.3$). Symbols are simulations averaged over 10 realizations on networks with $N = 10^6$ nodes and lines are theory from Eq. (11)

[1] R. Albert, H. Jeong, and A.-L. Barabási, Nature **406**, 378 (2000).
[2] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
[3] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[4] M. Newman, *Networks: an introduction* (OUP Oxford, 2009).
[5] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, Nature Physics **8**, 40 (2012).
[6] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. **109**, 248701 (2012).
[7] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, Journal of complex networks **2**, 203 (2014).
[8] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä,

Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, Physical Review X **3**, 041022 (2013).

[9] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, Nature Physics **12**, 901 (2016).

[10] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin, Physics Reports **544**, 1 (2014).

[11] D. Zhou, J. Gao, H. E. Stanley, and S. Havlin, Physical Review E **87**, 052812 (2013).

[12] R. Cohen, S. Havlin, and D. Ben-Avraham, Handbook of graphs and networks (2003).

[13] A. V. Goltsev, S. N. Dorogovtsev, and J. Mendes, Physical Review E **78**, 051105 (2008).

[14] M. E. Newman, Physical Review E **68**, 026121 (2003).

[15] S. Shao, X. Huang, H. E. Stanley, and S. Havlin, Physical Review E **89**, 032812 (2014).

[16] A. Bashan, Y. Berezin, S. V. Buldyrev, and S. Havlin, Nature Physics **9**, 667 (2013).

[17] W. Li, A. Bashan, S. V. Buldyrev, H. E. Stanley, and S. Havlin, Phys. Rev. Lett. **108**, 228702 (2012).

[18] L. M. Shekhtman, Y. Berezin, M. M. Danziger, and S. Havlin, Physical Review E **90**, 012809 (2014).

[19] S. Shai, D. Y. Kenett, Y. N. Kenett, M. Faust, S. Dobson, and S. Havlin, Phys. Rev. E **92**, 062805 (2015).

[20] L. M. Shekhtman, S. Shai, and S. Havlin, New Journal of Physics **17**, 123007 (2015).

[21] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).

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

[23] Y. Berezin, A. Bashan, M. M. Danziger, D. Li, and S. Havlin, Scientific reports **5** (2015).

[24] S. Shao, X. Huang, H. E. Stanley, and S. Havlin, New Journal of Physics **17**, 023049 (2015).

[25] B. R. da Cunha, J. C. González-Avella, and S. Gonçalves, PloS one **10**, e0142824 (2015).

[26] A. Lancichinetti, S. Fortunato, and J. Kertész, New Journal of Physics **11**, 033015 (2009).

[27] A. Clauset, C. Moore, and M. E. Newman, Nature **453**, 98 (2008).

[28] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, Nature **435**, 814 (2005).

[29] T. C. McAndrew, C. M. Danforth, and J. P. Bagrow, Physical Review E **91**, 042813 (2015).

[30] J. Zhao, D. Li, H. Sanhedrai, R. Cohen, and S. Havlin, Nature communications **7**, 10094 (2016).

[31] A. E. Motter and Y.-C. Lai, Physical Review E **66**, 065102 (2002).

[32] T. P. Peixoto, Physical Review E **85**, 056122 (2012).

[33] T. P. Peixoto, Physical Review X **4**, 011047 (2014).

[34] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Physical Review E **84**, 066106 (2011).

[35] B. Karrer and M. E. Newman, Physical review E **83**, 016107 (2011).

[36] G. Bianconi and S. N. Dorogovtsev, Physical Review E **89**, 062814 (2014).

[37] C. Wu, S. Ji, R. Zhang, L. Chen, J. Chen, X. Li, and Y. Hu, EPL (Europhysics Letters) **107**, 48001 (2014).

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

[39] G. Bianconi, Physical Review E **87**, 062806 (2013).

[40] G. Bianconi and F. Radicchi, Physical Review E **94**, 060301 (2016).

[41] C. D. Brummitt, K.-M. Lee, and K.-I. Goh, Physical Review E **85**, 045102 (2012).

[42] P. Van Mieghem, Physical Review E **93**, 042305 (2016).

[43] J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **107**, 195701 (2011).

[44] B. Min, S. Do Yi, K.-M. Lee, and K.-I. Goh, Physical Review E **89**, 042811 (2014).

[45] D. Cellai, E. López, J. Zhou, J. P. Gleeson, and G. Bianconi, Phys. Rev. E **88**, 052811 (2013).

[46] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. D. Porcellinis, and R. Setola, International Journal of Critical Infrastructures **4**, 63 (2008).

[47] S. Rinaldi, J. Peerenboom, and T. Kelly, Control Systems, IEEE **21**, 11 (2001).

[48] J. Gao, S. V. Buldyrev, H. E. Stanley, X. Xu, and S. Havlin, Phys. Rev. E **88**, 062816 (2013).