Epidemic spreading on interconnected networks

Anna Saumell-Mendiola, M. Ángeles Serrano, and Marián Boguñá

Departament de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain

Many real networks are not isolated from each other but form networks of networks, often interrelated in nontrivial ways. Here, we analyze an epidemic spreading process taking place on top of two interconnected complex networks. We develop a heterogeneous mean-field approach that allows us to calculate the conditions for the emergence of an endemic state. Interestingly, a global endemic state may arise in the coupled system even though the epidemics is not able to propagate on each network separately and even when the number of coupling connections is small. Our analytic results are successfully confronted against large-scale numerical simulations.

Epidemic spreading is one of the most successful application areas of the new science of networks [1,2]. Indeed, the general acceptance within the scientific community that many diseases, like sexually transmitted diseases or the H1N1 virus, spread over networked systems represents a major step toward their understanding and control [3–5]. From a physics perspective, epidemic processes have been widely studied as a paradigm of nonequilibrium phase transitions with absorbing states [6]. When applied to complex networks, these processes have become a source of new and striking phenomena that do not have a counterpart in regular lattices. Germane examples are the absence of epidemic and percolation thresholds in scale-free networks with a power law degree distribution $P(k) \sim k^{-\gamma}$, with $\gamma \in (2,3]$, and an anomalous critical behavior when $\gamma \in (3,4)$ [7–9].

We currently have a solid understanding of epidemic processes when they take place on single isolated networks. In contrast, our comprehension is very limited when epidemics happen on coupled interconnected networks. For example, sexually transmitted diseases can propagate both in heterosexual and homosexual networks of sexual contacts [4]. These two networks are not completely isolated due to the existence of bisexual individuals, which act as an effective coupling between the two networks and potentially affect their epidemic properties [10]. Apart from a few previous works [11,12], a theory describing these types of systems has not yet been fully developed.

In this paper, we present a rigorous heterogeneous mean-field study of the susceptible-infected-susceptible (SIS) model taking place on two interconnected complex networks. Our analysis reveals a highly nontrivial behavior of the epidemic process depending on the strength and nature of the coupling between the networks. We calculate the global epidemic threshold of the process, which turns out to be smaller than the epidemic thresholds of the two networks separately under certain conditions. This implies that *an endemic state may arise even if the epidemics is not endemic in any of the two networks separately*, as we prove analytically and with large-scale computer simulations. Analogous results apply also to the susceptible-infected-recovered (SIR) model.

To begin our analysis, we have to specify the topological properties of our networks. Let A and B be two interconnected random networks with given local properties and given two-point correlations. Each node in network A (respectively B) is characterized by a vector degree $\mathbf{k}_a \equiv (k_{aa}, k_{ab})$ representing the number of its internal connections k_{aa} to other nodes in A and the number of external connections k_{ab} with nodes of network B. The analogy to the degree distribution in single networks is then the probability that a randomly chosen node of network A has a vector degree $\mathbf{k}_a$, $P_A(\mathbf{k}_a) = P_A(k_{aa}, k_{ab})$, where, in general, k_{aa} and k_{ab} may be correlated. A similar definition applies to $P_B(\mathbf{k}_b)$ for network B. These two distributions are arbitrary except for the consistency condition $N_A/N_B = \langle k_{ba} \rangle / \langle k_{ab} \rangle$, where N_A/N_B is the ratio between the sizes of the two networks and the right-hand side is the ratio between their external average degrees. This condition states that the total number of edges leaving network A towards network B must be the same going from network B to network A. Two-point correlations, on the other hand, are encoded by the transition probabilities $P_{AA}(\mathbf{k}'_a|\mathbf{k}_a)$, $P_{AB}(\mathbf{k}'_b|\mathbf{k}_a)$, $P_{BA}(\mathbf{k}'_a|\mathbf{k}_b)$, and $P_{BB}(\mathbf{k}'_b|\mathbf{k}_b)$. For instance, $P_{AB}(\mathbf{k}'_b|\mathbf{k}_a)$ is the probability that being in a node of network A with vector degree $\mathbf{k}_a$ a randomly chosen neighbor in network B has vector degree $\mathbf{k}'_b$. Similar definitions apply to the rest of the transition probabilities.

As the epidemic spreading model, we consider the SIS model that, together with the SIR model, is one of the best-studied models in epidemiology [13]. The model has a nonequilibrium phase transition between an endemic state with sustained epidemic activity and a healthy phase where the epidemics dies out. Individuals can be in two different states, either susceptible (S) or infected (I). Infected individuals decay spontaneously to the susceptible state at rate δ (which, without loss of generality, we set to $\delta = 1$), whereas susceptible ones get infected at a rate proportional to the number of infected neighbors they have at a given time. For two interconnected networks, we have to specify these processes separately. Let λ^{aa} (λ^{bb}) be the infectious rate between nodes in network A (B) and λ^{ab} (λ^{ba}) the infectious rate from a node in A (B) to a node in B (A).

The quantity of interest in the SIS dynamics is the prevalence, $\rho(t)$, defined as the fraction of infected nodes at a given time. To describe its evolution, we use the heterogeneous mean-field approximation. In this approximation, nodes are classified within classes of equivalence such that all nodes within a given class are considered as statistically equivalent. In our case, classes of equivalence are defined by the network itself (A or B) and by specific values of the vector degree.

Following these assumptions, we define the partial prevalences $\rho_{\mathbf{k}_a}^A(t)$ and $\rho_{\mathbf{k}_b}^B$ as the fraction of infected nodes with a given vector degree. The total prevalences for each network is then computed as $\rho^A(t) = \sum_{\mathbf{k}_a} P_A(\mathbf{k}_a)\rho_{\mathbf{k}_a}^A(t)$ and $\rho^B(t) = \sum_{\mathbf{k}_b} P_B(\mathbf{k}_b)\rho_{\mathbf{k}_b}^B$. Following Refs. [7,14], the time evolution of the dynamics can be written as

$$\frac{d\rho_{\mathbf{k}_a}^A}{dt} = -\rho_{\mathbf{k}_a}^A + \lambda^{aa}\left(1 - \rho_{\mathbf{k}_a}^A\right)k_{aa}\sum_{\mathbf{k}_a'}\rho_{\mathbf{k}_a'}^A P_{AA}(\mathbf{k}_a'|\mathbf{k}_a)$$
$$+ \lambda^{ba}\left(1 - \rho_{\mathbf{k}_a}^A\right)k_{ab}\sum_{\mathbf{k}_b'}\rho_{\mathbf{k}_b'}^B P_{AB}(\mathbf{k}_b'|\mathbf{k}_a). \quad (1)$$

The first line in Eq. (1) describes the standard SIS model for a single network [14], whereas the second line appears due to the coupling of network A with network B. An analogous equation can be written for network B by swapping the indices $A \to B$ and $a \to b$ in Eq. (1).

As in a single network, this process undergoes a phase transition between a healthy phase with $\rho_{\mathbf{k}_a}^A = \rho_{\mathbf{k}_b}^B = 0$ and an endemic phase with $\rho_{\mathbf{k}_a}^A \neq 0$ and $\rho_{\mathbf{k}_b}^B \neq 0$. However, a mixed phase with endemic activity in one network whereas the other is in a heathy state is not possible in the system formed by the two coupled networks, where the epidemics propagates to the whole system if it is able to propagate in one of the networks. This is due to the fact that the state $\rho_{\mathbf{k}_a}^A = 0$ and $\rho_{\mathbf{k}_b}^B \neq 0$ is not a fixed point of the dynamics in Eq. (1). The critical point separating the healthy and endemic phases can be obtained by studying the stability of the absorbing solution. This can be done by linearizing the system of Eq. (1) around $\rho_{\mathbf{k}_a}^A = \rho_{\mathbf{k}_b}^B = 0$ and studying the spectrum of the corresponding matrix. Close to the absorbing state, Eq. (1) can be written as

$$\frac{d\vec{\rho}}{dt} = -\vec{\rho} + \mathbb{C}\vec{\rho}, \quad (2)$$

where we define the vector prevalence as $\vec{\rho} \equiv (\rho_{\mathbf{k}_a}^A, \rho_{\mathbf{k}_b}^B)$ and

$$\mathbb{C} = \begin{pmatrix} \lambda^{aa}k_{aa}P_{AA}(\mathbf{k}_a'|\mathbf{k}_a) & \lambda^{ba}k_{ab}P_{AB}(\mathbf{k}_b'|\mathbf{k}_a) \\ \lambda^{ab}k_{ba}P_{BA}(\mathbf{k}_a'|\mathbf{k}_b) & \lambda^{bb}k_{bb}P_{BB}(\mathbf{k}_b'|\mathbf{k}_b) \end{pmatrix}. \quad (3)$$

The absorbing state is stable whenever the maximum eigenvalue of matrix $\mathbb{C}$ satisfies $\Lambda_m < 1$. Otherwise, the absorbing state is unstable and an endemic state takes over the system. The critical epidemic point is then defined by $\Lambda_m = 1$.

When the networks do not have two-point correlations, the transition probabilities can be written as

$$P_{AA}(\mathbf{k}_a'|\mathbf{k}_a) = \frac{k_{aa}'P(\mathbf{k}_a')}{\langle k_{aa}\rangle}; \quad P_{AB}(\mathbf{k}_b'|\mathbf{k}_a) = \frac{k_{ba}'P(\mathbf{k}_b')}{\langle k_{ba}\rangle}. \quad (4)$$

In this case, the eigenvalues of matrix $\mathbb{C}$ are the solutions of the following equation,

$$[x - \Lambda_A][x - \Lambda_B]\left[x^2 - \Lambda_{AB}^2\right]$$
$$= \alpha_{ab}\alpha_{ba} + \mu_{ab}\alpha_{ba}[x - \Lambda_A] + \mu_{ba}\alpha_{ab}[x - \Lambda_B], \quad (5)$$

where

$$\Lambda_A = \lambda^{aa}\frac{\langle k_{aa}^2\rangle}{\langle k_{aa}\rangle}, \quad \Lambda_B = \lambda^{bb}\frac{\langle k_{bb}^2\rangle}{\langle k_{bb}\rangle}, \quad \text{and} \quad (6)$$

$$\Lambda_{AB}^2 = \lambda^{ab}\frac{\langle k_{ab}^2\rangle}{\langle k_{ab}\rangle}\lambda^{ba}\frac{\langle k_{ba}^2\rangle}{\langle k_{ba}\rangle} \equiv \mu_{ab}\mu_{ba}, \quad (7)$$

and where

$$\alpha_{ab} = \lambda^{aa}\lambda^{ab}\frac{\langle k_{aa}k_{ab}\rangle^2}{\langle k_{aa}\rangle\langle k_{ab}\rangle} \quad \text{and} \quad \alpha_{ba} = \lambda^{bb}\lambda^{ba}\frac{\langle k_{bb}k_{ba}\rangle^2}{\langle k_{bb}\rangle\langle k_{ba}\rangle}. \quad (8)$$

Constants appearing on the left-hand side of Eq. (5) have a clear interpretation. Indeed, Λ_A and Λ_B are the maximum eigenvalues of networks A and B as if they were isolated. Therefore if, for instance, $\Lambda_A > 1$ then network A is able to sustain an endemic state by itself when isolated from network B. Similarly, Λ_{AB} is the maximum eigenvalue of the network AB as a pure bipartite system, that is, when all internal connections inside networks A and B are absent. Again, when $\Lambda_{AB} > 1$, the pure bipartite network AB is able to sustain an endemic state, both in the coupled system and even if there were no connections whatsoever within each individual network. Constants $\alpha_{ab}, \alpha_{ba}, \mu_{ab},$ and μ_{ba} appearing on the right-hand side of Eq. (5) contain information about the strength and nature of the coupling between the nets.

From Eq. (5), it is easy to see that the maximum eigenvalue of matrix $\mathbb{C}$, Λ_m, is always larger than $\max(\Lambda_A, \Lambda_B, \Lambda_{AB})$. It is therefore possible to find endemic states with $\Lambda_m > 1$ but where $\Lambda_A < 1$, $\Lambda_B < 1$, and $\Lambda_{AB} < 1$, that is, situations where neither networks A and B isolated nor the pure bipartite network AB are able to sustain the endemic state and yet the epidemics pervades in the coupled system. This effect is more or less important depending on the strength of the coupling, i.e., the number of links between the two networks—quantified by $\langle k_{ab}\rangle$ and $\langle k_{ba}\rangle$—and the specific correlations between internal and external degrees—measured by the factors $\langle k_{aa}k_{ab}\rangle$ and $\langle k_{ba}k_{bb}\rangle$.

In the rest of the paper, we focus on the interesting case $\Lambda_A < 1$, $\Lambda_B < 1$, and $\Lambda_{AB} < 1$, and ask under which condition the endemic state exists, that is, $\Lambda_m > 1$. From Eq. (5), we see that this happens when the right-hand side of Eq. (5) evaluated at $x = 1$ is larger than the left-hand side evaluated at the same point, which after some algebra yields

$$[\alpha_{ab} + \mu_{ab}(1 - \Lambda_A)][\alpha_{ba} + \mu_{ba}(1 - \Lambda_B)]$$
$$> (1 - \Lambda_A)(1 - \Lambda_B). \quad (9)$$

This equation is one of the main results of our paper. It allows us to evaluate the conditions for the emergence of the endemic state in many different situations.

In real networks of sexual contacts, the most promiscuous individuals in one network are also the ones with the largest number of sexual partners in the other network [10]. This fact suggests that there exists a positive correlation between the internal and external degree for a given node. We model these correlations by taking the vector degree distribution to be $P_A(k_{aa}, k_{ab}) = P_A(k_{aa})g(k_{ab}|k_{aa})$, where $g(k_{ab}|k_{aa})$ is a Poisson distribution with mean $\bar{k}_{ab}(k_{aa}) = \langle k_{ab}\rangle k_{aa}^\beta/\langle k_{aa}^\beta\rangle$, and analogously for $P_B(k_{bb}, k_{bb})$. This choice allows us to interpolate between a random distribution of links between the two networks, $\beta = 0$, and positive correlations, where high degree nodes in both networks concentrate the majority of the coupling links, $\beta > 0$. Inserting this assumption in Eq. (9), we obtain the critical lines that define the phase diagram in the hyperplane $\Lambda_A, \Lambda_B, \alpha_{ab},$ and α_{ba}.
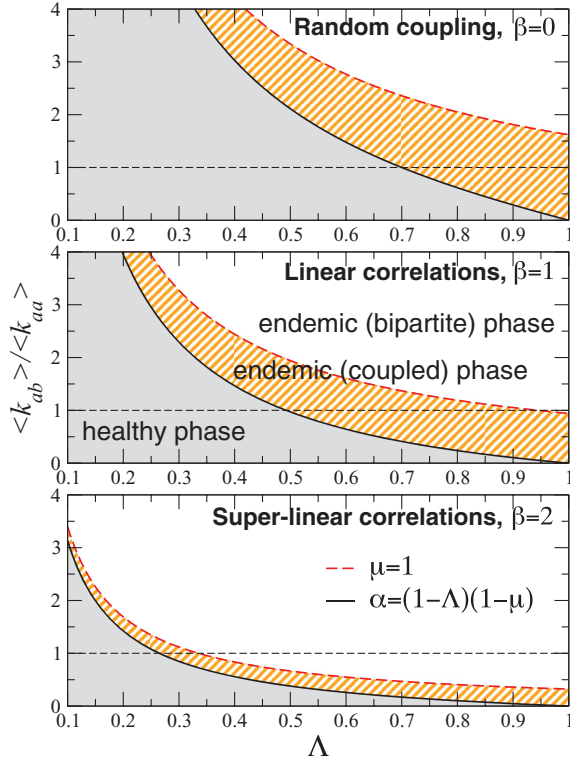
FIG. 1. (Color online) Phase diagram showing the healthy phase and the endemic phase for the case of a random coupling (top), linear correlations (middle), and superlinear correlations between external and internal degrees (bottom). In all cases, networks A and B are identical, with an exponential degree distribution with $\langle k_{aa} \rangle = \langle k_{bb} \rangle = 10$ and minimum degree 2. The area depicted in solid gray corresponds to the healthy phase and the transition to the endemic state is represented by the black curve. The dashed orange area corresponds to the endemic state that needs the contribution of both internal and external links and the white area to the case when the pure bipartite network alone is able to sustain the endemic state. The horizontal dashed line marks the point where external connections outnumber internal ones. Notice that when β increases, nodes with high internal degrees also have high external degrees with high probability. Because of this, these nodes are excellent spreaders. As a consequence, the network needs smaller values of Λ and/or interconnectivity to be in the endemic phase.

Figure 1 shows examples of phase diagrams for two identical networks with a symmetric coupling, that is, $\Lambda_A = \Lambda_B = \Lambda$, $\alpha_{ab} = \alpha_{ba} = \alpha$, and $\mu_{ab} = \mu_{ba} = \mu$. In general, nonsymmetric cases can be easily studied using Eq. (9), which simplifies to $\alpha > (1 - \Lambda)(1 - \mu)$ for symmetric coupling. We consider the case of networks with an exponential degree distribution and the values of $\beta = 0$, 1, and 2. In sexual contact networks, the number of contacts is limited not by the size of the network but by people's ability to engage others in sexual intercourse during the finite time scale of the epidemics. This limits the network heterogeneity and justifies our choice of an exponential degree distribution. The most interesting case corresponds to the range of parameters where both internal and external links are needed for the existence of the endemic state—depicted in orange in Fig. 1, and particularly when the internal connections outnumber the external ones (the portion

below the dashed line). This area shrinks when increasing β but, simultaneously, it also appears at lower values of the ratio of external versus internal connections for the same Λ. Therefore, keeping the two networks unchanged and for a fixed number of links between them, the epidemic can be either in the healthy phase or in the endemic phase depending on how these links are distributed among the nodes of the networks. Notice that, if correlations are strong enough, the bipartite network alone is able to sustain by itself the endemic state in the coupled system.

We checked our predictions with large-scale numerical simulations. The SIS dynamics is simulated with a continuous time dynamics as follows. During the course of the simulation, we keep track of the number of infected nodes $N_I(t)$ and the number of active links $E_A(t)$, where an active link is defined as a link connecting a susceptible node and an infected node. At each step, with probability $p_r = N_I(t)[N_I(t) + \lambda E_A(t)]^{-1}$, a randomly chosen infected node is turned susceptible whereas, with probability $1 - p_r$, an active link is chosen at random and the susceptible node attached to it is turned infected. After this procedure, time is updated as $t \rightarrow t + [N_I(t) + \lambda E_A(t)]^{-1}$. We run this algorithm on the networks used in Fig. 1 of size $N = 10^6$. Having fixed the internal network properties, we fix the value of Λ by adjusting the infectious rate λ using Eq. (6). For each value of Λ, we change the average external degree and study the temporal behavior of the prevalence. The critical point is determined as the point where $\rho(t)$ decays as a power law [6], as shown in Fig. 2.

Using this method, we compute the critical line in the plane $(\langle k_{ab} \rangle / \langle k_{aa} \rangle, \Lambda)$ for the three cases analyzed in Fig. 1. We limit our search to the domain $(\langle k_{ab} \rangle < \langle k_{aa} \rangle, \Lambda < 1)$. Simulation results are shown in Fig. 3 along with the analytic predictions given by Eq. (9). We observe a systematic shift between theory and simulations. The reason is that our mean-field approach does not consider dynamical correlations. These correlations imply that, with high probability, an infected node has its infecting neighbor still infected during some time right after the infectious event. This reduces the number of potential new infections that this node can produce. We can correct this
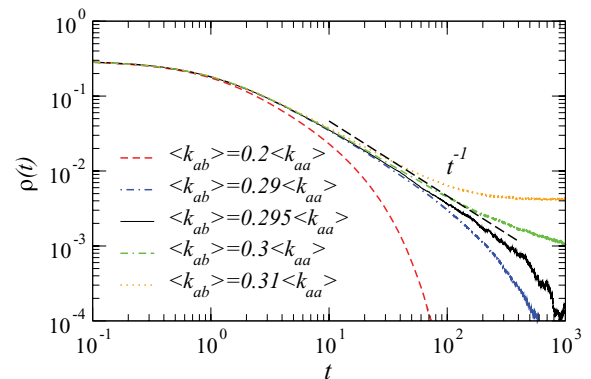

FIG. 2. (Color online) Time evolution of the prevalence $\rho(t)$ below and above the critical point for $\Lambda = 0.8$ for two identical networks interconnected with linear correlations, i.e., $\beta = 1$. Each network has an exponential degree distribution with $\langle k \rangle = 10$ and minimum degree 2. The size of each network is $N = 10^6$ and results are averaged over 100 different realizations of the process. Right at the critical point, the prevalence decays as $\rho(t) \sim t^{-1}$.
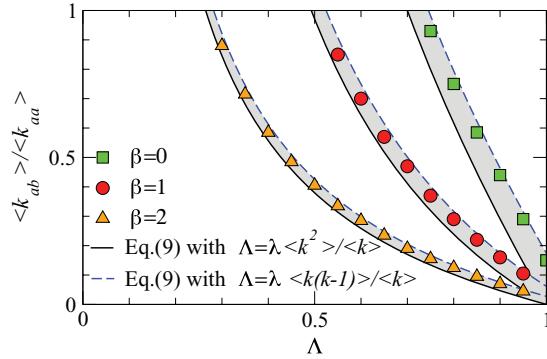
FIG. 3. (Color online) Critical lines for $\beta = 0$, 1, and 2 for the same networks as in Fig. 2. Symbols correspond to numerical simulations, black lines are the theoretical predictions according to Eq. (9), and blue dashed lines are obtained by rescaling the Λ axis as $\Lambda \to \Lambda(1 - \langle k \rangle / \langle k^2 \rangle)^{-1}$. Error bars are of the order of symbol sizes.

effect in the same way as it is done in the SIR model, just by reducing the number of contacts by one. In our equations, this is achieved by replacing the maximum eigenvalue of the networks by $\Lambda = \lambda \langle k(k-1) \rangle / \langle k \rangle$. This is, of course, a limiting case because there are cases where the infecting neighbor recovers and can be reinfected. Thus, we expect to find simulation results between these two extremes. We show this correction in Fig. 3 as the gray area. Indeed, all simulation points fall within this gray area, confirming then our intuition. In addition, after this correction is performed, our results can

be readily used to evaluate the epidemic threshold of the SIR model.

The study of interconnected and/or interdependent networks reveals new and unexpected phenomena [15]. Here, we have shown that two networks well below their respective epidemic thresholds may sustain an endemic state when coupling connections are added, even in small numbers. This may have important implications for the design of efficient control strategies. However, the effects of the coupling are highly nontrivial and may vary depending on the strength and the correlations of the interconnecting links. We foresee similar effects appearing in many different dynamics showing equilibrium and nonequilibrium phase transitions.

Note added in proof. Recently, we became aware of two recent preprints where the SIR model on interconnected networks is studied [16,17]. In Ref. [16], the authors find a mixed phase where one network propagates the epidemic while the other does not. Neither our analytic results nor our simulation results indicate the existence of such a mixing phase in our case.

[1] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).

[2] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, London, 2010).

[3] A. L. Lloyd and R. M. May, Science **292**, 1316 (2001).

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

[5] D. Balcan, H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. Broeck *et al.*, BMC Med. **7**, 45 (2009).

[6] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, UK, 1999).

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

[8] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).

[9] R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. E **66**, 036113 (2002).

[10] W. L. Jeffries, Perspect. Sex. Reprod Health **43**, 151 (2011).

[11] M. Kurant and P. Thiran, Phys. Rev. Lett. **96**, 138701 (2006).

[12] E. A. Leicht and R. D'Souza, arXiv:0907.0894.

[13] R. M. May and R. M. Anderson, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, UK, 1991).

[14] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002).

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

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

[17] O. Yagan, D. Qian, J. Zhang, and D. Cochran, arXiv:1112.4002.