Epidemic spreading in networks with nonrandom long-range interactions

Ernesto Estrada,[1,2,3,4,*] Franck Kalala-Mutombo,[1] and Alba Valverde-Colmeiro[5]
[1]*Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XQ, United Kingdom*
[2]*Department of Physics, University of Strathclyde, Glasgow G1 1XQ, United Kingdom*
[3]*Institute of Complex Systems, University of Strathclyde, Glasgow G1 1XQ, United Kingdom*
[4]*SUPA, University of Strathclyde, Glasgow G1 1XQ, United Kingdom*
[5]*Department of Economic Analysis: Quantitative Economics, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
(Received 13 March 2011; revised manuscript received 7 July 2011; published 16 September 2011)

An "infection," understood here in a very broad sense, can be propagated through the network of social contacts among individuals. These social contacts include both "close" contacts and "casual" encounters among individuals in transport, leisure, shopping, etc. Knowing the first through the study of the social networks is not a difficult task, but having a clear picture of the network of casual contacts is a very hard problem in a society of increasing mobility. Here we assume, on the basis of several pieces of empirical evidence, that the casual contacts between two individuals are a function of their social distance in the network of close contacts. Then, we assume that we know the network of close contacts and infer the casual encounters by means of nonrandom long-range (LR) interactions determined by the social proximity of the two individuals. This approach is then implemented in a susceptible-infected-susceptible (SIS) model accounting for the spread of infections in complex networks. A parameter called "conductance" controls the feasibility of those casual encounters. In a zero conductance network only contagion through close contacts is allowed. As the conductance increases the probability of having casual encounters also increases. We show here that as the conductance parameter increases, the rate of propagation increases dramatically and the infection is less likely to die out. This increment is particularly marked in networks with scale-free degree distributions, where infections easily become epidemics. Our model provides a general framework for studying epidemic spreading in networks with arbitrary topology with and without casual contacts accounted for by means of LR interactions.

DOI: 10.1103/PhysRevE.84.036110 PACS number(s): 89.75.Hc, 89.75.Fb, 02.10.Ox, 87.23.Ge

I. INTRODUCTION

The main paradigm of the study of dynamic processes in complex networks is that information is transmitted through the paths that connect pairs of nodes [1–4]. Such paths are formed by sequences of nodes representing the entities of a complex system, which are connected by links representing the interactions between these entities. This paradigm is particularly useful for studying the spread of information in complex networks [4], in particular for studying how "infections" propagate and become epidemic [5–16]. One of the most important challenges of modeling the spread of epidemics is the determination of the network of social contacts that allow the spread of the infection. While in some situations, such as in the spread of sexually transmitted diseases or computer viruses, knowing the network of contacts is not very difficult, in those cases involving the transmission of airborne or close contact infections the contact network is quite hard to define [17]. This was, for instance, the case of the severe acute respiratory syndrome (SARS), which was propagated when a medical doctor from Guangzhou, China, eventually met at a hotel in Kowloon people from Singapore, Viet Nam, Canada, and Hong Kong who were not among his "close" social contacts [18]. These kinds of encounters between individuals who can facilitate the transmission of an infection are referred to as "casual" contacts in order to distinguish them from the more frequent close contacts among individuals [19–22]. These casual contacts can play a major role in a variety of phenomena, which include, for instance, imitative obesity as "it may be easier to be fat in a society that is fat" [23] or the fact that "the spread of obesity is related to the environment in which individuals live" [24] in addition to their social ties [25]. Other examples can include epidemic hysteria [26] as well as the recent growth of "binge" drinking in the United Kingdom as a "fashion-related phenomenon" [27].

Due to the importance of the social contacts, both close and casual, among individuals in our society for understanding disease and attitude spreading, there have been serious attempts to account for them on an experimental basis. The first attempt traced the route of the circulation of bank notes in the United States [28]. The second one studied the trajectory of 100 000 mobile phone users by detecting their positions during the period of half a year [29]. The results of these studies are both theoretically and practically interesting. However, we never make commercial transactions in our elevators, buses, trains, or airplanes and do not necessarily use our mobile phones all at the same time and place. Then, the problem of determining the social contacts of individuals in a society is a very difficult and challenging problem of tremendous importance. More recently, however, some attempts of quantifying all close and casual contacts among individuals have been conducted in a series of European cities [30]. This study and its implications are analyzed in the next section. The division of social contacts between close and casual is somewhat artificial, but it allows some important theoretical approaches. For instance, by knowing the network of close contacts among individuals it is possible to include some effects produced by the casual

*Corresponding author: ernesto.estrada@strath.ac.uk

 ©2011 American Physical Society

encounters among individuals. This has been generally done by considering that these casual contacts occur at random. We will analyze the implications of this assumption in the next section.

Here we propose to account for both the close and casual contacts among individuals by considering the transmission of an infection through paths and by long-range (LR) interactions in a complex network. The main paradigm used by this model is that if we know the structure of the network of close contacts, we can infer the casual contacts by means of the long-range interactions between two individuals, which here are assumed to depend on the social distance between them. Using this model, we study how an infected node can propagate an infection in random and real-world networks. We show that an epidemic expands faster and is less likely to die out when casual contacts are considered. The heterogeneity of a network also contributes significantly to the propagation of an infection when casual contacts are taken into account.

II. SOCIAL CONTACTS: EMPIRICAL EVIDENCE

In order to model effectively the way in which infections are spread in networks it is necessary to account for the pattern of human interactions [30,31]. In certain situations, e.g., sexually transmitted diseases or computer viruses, the determination of such social contacts can be done in an effective way. However, in other scenarios, such as in the transmission of airborne or close contact infections, the pattern of human interactions produced by encounters between individuals is harder to define [17]. While in the first case only close contacts produced by direct interactions among the individuals can be responsible for the transmission of the infection, in the second case the contributions of both close and casual contacts play a fundamental role. The combination of both close and casual contacts among individuals accounts for those social interactions, which include physical and nonphysical contacts occurring at different environments (home, work, school, transport, leisure, etc.) for periods of time that range between a few minutes and several hours. Only recently, some studies have been conducted that shed some light on the patterns of these social contacts. For instance, Mossong *et al.* [30] have studied 97 904 contacts among 7290 individuals in 8 different European countries. They have registered the age, sex, duration, location, frequency, and occurrence of physical contacts. Because they study the occurrence of these social contacts in places like home, work, school, leisure, transport, and others as well as the combination of them, this study accounts for both close and casual contacts among individuals. A remarkable finding of this study is that the social contacts among individuals occur preferentially among those of similar ages. This pattern is particularly pronounced among children and youngsters in the age range between 5 and 24 years. This kind of age assortativity is also observed for adults of about 40 years. It is well known that children, teenagers, and youngsters develop friendship relationships preferentially among them, observing some kind of age assortativity in their social ties. Middle-aged adults are also preferentially tied to other individuals of similar ages by mean of working relationships or other social ties. The role of assortativity in social relationships has been well documented. In social science it is also known as "homophily" and refers to the observed fact that "similarity breeds connections" or that "birds of a feather flock together." An excellent compendium on homophily in social networks is the work of McPherson *et al.* [32]. To review some of the results described in that work, it has been found in studies of close friendship that homophily (assortativity) by age is the strongest dimension controlling the relationships, with only the exception of race. For instance, about 38% of close friends among men in Detroit were found to be within two years of age and 72% within eight years. This assortativity is less marked in the people in the 60+ age group, which has been the only one group for which there was significant outbreeding [32]. These results on social friendship, or close contacts, reproduce very well those obtained for the social contacts, which include both close and causal, in the work of Mossong *et al.* [30]. Consequently, the assortativity relationship between social contacts (close and casual) and age is indicative of the relationship between social distance between individuals and social contacts. By social distance we mean the shortest path distance between two individuals in their social network. That is, two teenagers who are not friends are closer to each other than they are to some middle-aged strangers. The probability that these two teenagers frequent the same place, e.g., concerts, cinema, school, etc., is larger than that for the social contact among the teenager and the adult. By social network we understand here the social interaction between individuals that can be considered to be of relevance for epidemiological studies and which excludes those contacts that do not correlate with transmission opportunities for infections, such as links by means of only letters, telephone, emails, etc. Unfortunately, we have not found studies that provide empirical evidence of other types of homophilies in the casual contacts among individuals. However, it is highly probable that individuals with similar ethnicity, religion, education, occupation, social class, etc., who have been found to be closer in their social networks [32], live in similar geographic locations, use similar transportation, and visit similar places for leisure than individuals with less similarities, confirming the hypothesis of a correlation between casual contacts and social distance.

We are aware of the lack of empirical data about casual contacts in real social systems. Even the empirical study of Mossong *et al.* [30] does not include social contacts among individuals in a confined space or in close physical proximity in which the individuals are not talking among them, e.g., crowds at concerts. Then, we look for some empirical evidence that allow us to model casual social contacts. In some cases these casual encounters between individuals have been modeled by considering they occur at random [19–22,33]. In an analogous way epidemics are usually modeled on "small-world" networks. That is, individuals are placed at the nodes of a regular lattice whose links represent close contacts along which the infection may spread to others. Then, an infection proceeds either locally (through close contacts), within a prescribed neighborhood, or through casual contacts established at random between any two individuals [34–37]. That is, in this case also the long-range interactions among individuals are considered to be at random and not to depend at all on the social distance between individuals in the network.

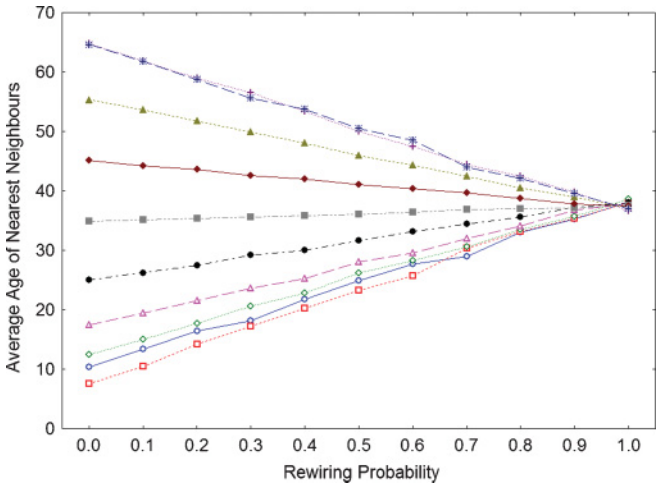

FIG. 1. (Color online) Average age of the nearest neighbor nodes in different age groups (see text) by using the WS model with node ages and a random rewiring of links. The ages are organized from top to bottom at probability 0.0 in the following groups: 0–5, 5–10, 10–15, 15–20, 20–30, 30–40, 40–50, 50–60, 60–70, and 70+.

In order to illustrate the lack of age homophily when casual contacts are considered as random we modify the Watts-Strogatz (WS) model [38] in order to account for the age of individuals. We start from a cycle graph of 100 nodes, and then we connect every node to its second nearest neighbors. This lattice, which is a circulant graph, is known as the WS graph for rewiring probability $p = 0.0$. We assign an age to every node staring from the node labeled as 1. The ages are assigned starting from 0 years with a clockwise increment of 0.757 years. The node labeled as 100 is then 75 years old. Thus, the nearest neighbors have similar ages, reproducing the observed age homophily in real social networks [30,32]. In addition, the youngest and older nodes are also linked together as a consequence of the circular nature of the lattice. This characteristic has also been observed in real-world social relationships [30,32]. For the age-WS graph we proceed with the typical rewiring of links with probability $p > 0.0$. Then, we calculate the average age of the nearest neighbors of each node for a given rewiring probability and report the average for the groups of ages 0–5, 5–10, 10–15, 15–20, 20–30, 30–40, 40–50, 50–60, 60–70, and 70+ as in the work of Mossong *et al.* [30]. The results are illustrated in Fig. 1.

As can be seen in Fig. 1 the classical WS model is unable to reproduce the age assortativity observed in social contacts. As $p \to 1.0$, all nodes tend to have neighbors of the same average age. The age assortativity disappears even for small values of p. For instance, for nodes in the group 10–15, which are known to display high age homophily in real networks [30], the average age of nearest neighbors is almost doubled from 12.5 for $p = 0.0$ to 22.8 for $p = 0.4$. For the nodes with ages between 5 and 24 for which Mossong *et al.* [30] have found strong age assortativity, the average age of nearest neighbors is almost duplicated for $p = 0.5$ when a random rewiring is used in the age-WS model. In closing, the randomness of casual contacts is not able to explain the age assortativity observed by Mossong *et al.* [30] and other possibly existing homophilies in the social contacts among individuals in eight different European countries.

In a different approach Cohen *et al.* [39] considered the spread of tuberculosis by accounting for both close and casual contacts among individuals. They considered a parameter D that controls the relative probability of creating shorter- versus longer-distance connections in a network. If the parameter D is small, individuals separated at short distances are preferentially linked; when D is large, the long-range interactions are favored. It is remarkable that the number of casual social contacts is considered to be a product of these long-range interactions. The authors have claimed that "in areas where a substantial proportion of transmission is due to 'casual' contacts, a network with higher D value would better represent the contact structure." That is, casual contacts are proportional to long-range interactions among nodes in a network. In the work of Cohen *et al.* [39], however, the proximity between individuals is considered to be the Euclidean distance between nodes placed randomly at a plane, which in some way tries to capture their "geographical" separation. Then, we need to provide some of the empirical evidence about the interrelation existing between geographic and social proximity.

The concept of proximity is widely used in social sciences, in particular in innovation studies, organization science, and regional science [40]. In many cases "proximity" refers to "geographical proximity," such as territorial, spatial, local, or physical proximity. However, there are other forms of proximity that are also in use in the same contexts, such as institutional, organizational, cultural, technological, and social proximity. Social proximity in particular refers to actors that belong to the same space of social relations [40]. In some cases it has been observed that geographical proximity is subordinate to the social one. For instance, for the transmission of knowledge Agrawal *et al.* [41] have concluded that "geographical proximity matters most in the absence of social proximity that may otherwise facilitate access to knowledge." However, it is difficult in many cases to disentangle social and geographical proximities. In fact, it has been stated that "the dichotomy between spatial and aspatial indices is somewhat a false one, since both types of measures incorporate implicit notions of social distance" [42]. We then assume here that in general the concept of social proximity encloses important information about other types of proximities, such as the geographical and cultural ones.

Finally, there is a group of empirical evidence that is important for the development of the current approach. This refers to the way in which individuals establish their links in social networks. We remark that the number of close and casual contacts among individuals has been claimed to be proportional to the probability of creating links among them [39]. This probability has been considered to be proportional to the gain that these two individuals will obtain from the new link [43]. Similarly, Sorenson and his coauthors have assumed that the new social links are created on the basis of the "expectations of the value" of those relationships [44]. An illuminating piece of evidence of the use of a "value motivation" for the establishment of social relations was obtained by Manson [45] in the study of the primates rhesus macaques. Manson [45] observed that "a young female may gain by investing in a friendship with a low-ranking male who

(1) is not presently sought by many females as a Friend and (2) will stay in the group and achieve high rank and thus have high protective ability in the future." So far, it is evident that the creation of new social ties is seen as an investment in which the "future value" of the relation is more important than the "present value" that the establishment of this link represents.

In summary, we have seen that there is empirical evidence that support the following claims.

(i) Social contacts among individuals are somewhat determined by their social distance. They account for an amalgam of proximities including social and geographical ones.

(ii) The number of close and casual contacts is somewhat determined by the probability of linking pairs of individuals by means of short- and long-range interactions.

(iii) The probability of linking pairs of individuals in a network depends on the future value that such a new link will bring to both individuals.

III. ACCOUNTING FOR THE SOCIAL CONTACTS AMONG NODES

Here we combine the empirical evidence analyzed in the previous section and summarized in the three points at the end of it into a mathematical model. Our aim is to account for the probability that two individuals have social contacts that are relevant for the transmission of an infection. We start by considering the existence of a social network among a group of individuals that is represented by a graph $G = (V,E)$. We assume that the social relationship between node i and node j in a network, which is represented by $(i,j) \in E$, corresponds to one that conveys close contacts of relevance to the transmission of the type of infection under consideration. In doing so we are assuming that if $(i,j) \in E$, the probability ε_{ij} that the two nodes have close social contacts is equal to 1, i.e., $\varepsilon_{ij} \equiv 1$.

Our next assumption is that if $(i,j) \notin E$, the probability that both nodes have social contacts is not necessarily equal to zero, but $0 < \varepsilon_{ij} < 1$. This means that both nodes can "eventually meet" in the same place and time by means of some kind of casual contacts, such as in transport, leisure, the supermarket, etc. Using empirical evidence points (i) and (ii), we will assume that this probability is determined by the structure of the social network, in particular by the probability of establishing a new link between both nodes. Now we are going to use point (iii) to determine the probability ε_{ij} for non-nearest neighbors. That is, we assume that the two nodes will have casual contacts in a way that is proportional to the establishment of a new link between them. They will see the establishment of this new link as an investment in which its future value will determine their decision of forming a new tie. We consider such a process like the one in which the time value of money, in particular the future value of a growing annuity, is determined in quantitative finance. Here, instead of money, we generalize the process by considering that a node lends a piece of information to another node. This information has a future value F_V that is determined, according to the quantitative finance theory, by its present value P_V, the interest rate r, and the number of time periods t at which the information is lent. That is [46],

$$F_V = P_V(1+r)^t. \tag{1}$$

Here we assume that if the node i lends some information to node j, the information flows through the shortest path connecting both nodes (or one of them if more than one exists) according to empirical evidence point (ii). The information is passed using a discrete time in which every step in the path is considered to have a unit time. That is, the number of periods for which the information is borrowed is assumed here to be equal to the shortest path separation of the two nodes. Then, let us consider the shortest paths between the two nodes as a directed chain from the lender to the borrower. We assume that the chain has length l and that the nodes are numbered in consecutive order starting with 1. In a process of lending information from node v_1 to node v_{l+1}, the information is first transferred to node v_2 with a value A and an interest rate r. The present value of the information in the hands of node v_2 is $A/(1+r)$. Then node v_2 enriched this information by a given value g, which we will designate as the *growth rate* of the information [46]. When node v_2 lends this information to node v_3 with the same interest and growth rates, the information will have a value $A(1+g)/(1+r)^2$ in the hands of node v_3. As every node in the chain lends the information to its nearest neighbor with interest r and growth rate g, the information in the hands of borrower node v_{l+1} will have a value of $A(1+g)^{l-1}/(1+r)^l$. The cumulative present value of the information in this process is given by the sum of all the values at the nodes of the chain [46]:

$$P_V = A/(1+r) + A(1+g)/(1+r)^2 + \cdots + A(1+g)^{l-1}/(1+r)^l. \tag{2}$$

If the growth and interest rates are the same, i.e., $g = r$, the present value of the information is simplified to

$$P_V = Al/(1+r). \tag{3}$$

Then, the future value of the information is given as [46]

$$F_V = Al(1+r)^{l-1}. \tag{4}$$

We will consider here that $A \equiv 1$ for the sake of simplicity. Then, because in a connected network any two nodes i and j are separated by a shortest-path distance d_{ij}, the expression for the future value of the information transmitted from i to j is given by

$$F_V = d_{ij}x^{d_{ij}-1}, \tag{5}$$

where $x = 1 + r = 1 + g$.

We consider here that lending information is carried out with a negative gain g. That is, every node will appropriate some part of the information they receive before lending to its nearest neighbor. Due to this, a node prefers to lend information to its closest neighbors than to far strangers. We also consider that the maximum benefit that a node can have is by lending information to its nearest neighbor. As a consequence, the values of the parameter x are bounded as $0 \leqslant x < 0.5$. The value $x = 0$ represents the situation in which no long-range transmission is allowed, which corresponds to the case in which only close contacts take place. When $x \rightarrow$

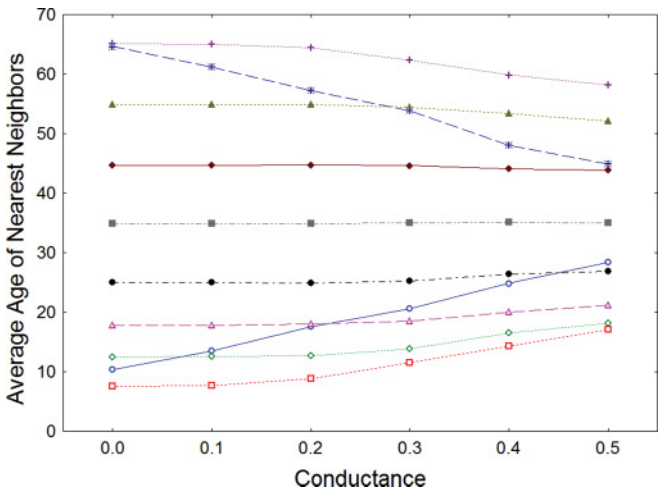

FIG. 2. (Color online) Average age of the nearest neighbor nodes in different age groups (see text) by using the WS model with node ages in which the rewiring probability depends explicitly on the internode distance. The ages are organized in groups as in Fig. 1.

0.5, the nodes are allowed to transmit information directly to their non-nearest neighbors. This situation represents scenarios in which a substantial proportion of transmission is due to casual contacts. Note that the bound is strictly smaller than 0.5 to avoid the future value of nodes separated at distance 2 becoming equal to unity.

In closing,

$$\varepsilon_{ij} = F_V = d_{ij} x^{d_{ij}-1},$$

which means that the maximum social contacts is obtained for the nearest neighbors and it decreases with the increase of the social distance between the two nodes. This simple model agrees with the three groups of empirical evidence analyzed in the previous section. In addition it also agrees with the empirical observation that most of the transmission usually occurs through close contacts rather than through casual ones [39]. At this point it is straightforward to propose a matrix representation for the social contacts among all nodes in a network. We designate this by $\mathbf{\Gamma}(G,x)$ [47], which is a square matrix whose entries are defined by

$$\mathbf{\Gamma}(G,x) = \begin{cases} 1 & \text{if } i \sim j, \\ d_{ij} x^{d_{ij}-1} & \text{if } i \neq j \quad \text{and} \quad i \not\sim j, \\ 0 & \text{if } i = j. \end{cases}$$

This matrix, of course, generalizes the adjacency matrix $\mathbf{A}(G)$ of the network. Note that $\mathbf{\Gamma}(G,x=0) = \mathbf{A}(G)$ [47]. From here on we will call the parameter x the conductance as it controls the way in which casual contacts are allowed in a network. In a zero conductance network only close contacts are allowed, which can be the case for the transmission in sexually transmitted diseases or computer viruses.

One expected characteristic of the current model is that it reproduces the age assortativity observed by Mossong *et al.* [30] in the social contacts in real networks. Then, we consider a modification of the age-WS model used in the previous section in which casual contacts depend on the social distance between individuals. Then, instead of considering a random rewiring such as in the age-WS model we consider that for nodes i and j, with $d_{ij} > 1$, the probability that i has a link with j is given by $p_{ij} = d_{ij} x^{dij-1}$. That is, the rewiring is carried out here on the basis of the "social distance" that separates two individuals. The details of this process will be given in a future presentation [48]. Using exactly the same age assignation to nodes as in the previous section, we obtained the results illustrated in Fig. 2. As can be seen, the age-WS model with distance-based rewiring displays strong age assortativity for all values of the conductance. In the case of the two extreme age groups, i.e., 0–5 and 70+, there is a larger difference in the average age of the nearest neighbors between $x = 0.0$ and $x = 0.5$, which is about 18 years. We remark again that in these groups it has been observed "experimentally" that there is a larger outbreeding than in the rest of age groups [30]. For the nodes in each of the other groups of ages analyzed in the previous section the increase of age does not exceed 10 years even for a high conductivity of $x = 0.5$, in complete agreement with the empirical evidence provided by the work of Mossong *et al.* [30]. For instance, for the same age group analyzed previously (10–15 years), the average age of nearest neighbors changes from 12.5 at $x = 0.0$ to 18.1 at $x = 0.5$. For the nodes with ages between 5 and 24 years analyzed by Mossong *et al.* [30] the average age of nearest neighbors changed only by 6 years when the conductance changed from 0.0 to 0.5. These results clearly point out the fact that the consideration of social distance as a director for casual social contacts is of relevance for studying the spread of infections in the real world. This characteristic has not long been reproduced by existing models that account for casual contacts as random long-range interactions among individuals.

IV. INFECTION SPREADING IN NETWORKS WITH LR INTERACTIONS

Here we use our model of social contacts to extend the idea of a discrete-time formulation of the problem of contact-based epidemic spreading in networks. Our approach consists in extending a model independently developed by Chakrabarti *et al.* [49] and Gómez *et al.* [50,51] on the basis of susceptible-infected-susceptible (SIS) epidemic models. However, our model of social contacts can be implemented in any other epidemic spreading strategy. The model considered was first developed by Chakrabarti *et al.* [49] as a nonlinear dynamical system (NLDS) in order to explain the propagation of viruses in computer networks. More recently, Gómez *et al.* [50,51] have proposed a similar microscopic Markov-chain approach (MMCA), which uses the same principles as in NLDS but has concentrated on the probability of infection of individuals rather than on the common heterogeneous mean-field approach. We clarify beforehand that our interest here, as in the precedent papers [49–51], is in networks of large size with any kind of topology. For obvious reasons we propose to call this model NLDS-MMCA.

In the NLDS-MMCA model [49–51], a node i remains healthy at time t in a network if it does not receive an infection from its nearest neighbors at a previous time step, $t-1$. In addition, if the node has been infected, it can recover and become healthy again with a certain probability. Then,

the probability $1 - p_{i,t}$ that the node remains healthy in the network is given by

$$1 - p_{i,t} = (1 - p_{i,t-1})\xi_{i,t} + \delta p_{i,t-1}\xi_{i,t}, \tag{6}$$

where $p_{i,t}$ is the probability that node i is infected at time t, $\xi_{i,t}$ is the probability that it does not receive an infection from its nearest neighbors at time t, and δ is the rate at which it can recover from infection. The probability that a node does not receive an infection from its nearest neighbors is assumed to be determined by the product of individual probabilities (independence assumption), which is determined by $p_{j,t-1}$ for all nearest neighbors of i and by the universal birth rate of the infection β [49–51]. That is,

$$\xi_{i,t} = \prod_{j\sim i}(1 - \beta p_{j,t-1}), \tag{7}$$

where $j \sim i$ indicates that j is directly connected to i. As shown in [49], the epidemic threshold τ for an undirected network is

$$\tau = \frac{1}{\lambda_{1,A}}, \tag{8}$$

where $\lambda_{1,A}$ is the largest eigenvalue of the adjacency matrix $\mathbf{A}$ of the network.

The size of the infected population in the network is given by

$$\eta(t) = \sum_{i=1}^{n} p_{i,t}. \tag{9}$$

Here we are interested in considering the case in which an infection is transmitted from one infected node to its close and casual contacts with certain probability. Obviously, the nearest neighbors, representing close contacts in the social network, of that infected node are at the highest risk to be infected. However, we consider here that every node in that network can be infected directly from that infected node as they can be casually proximal to it at a certain stage. Our assumption is that these casual contacts depend on the shortest-path distance at which these nodes are from the infected node. Then, $\mathbf{\Gamma}(G,x)$ is a natural substitution for the adjacency matrix in NLDS-MMCA, which transforms this model into a generalized one (GNLDS-MMCA), where the probability $1 - p_{i,t}$ that the node remains healthy in the network is given by

$$1 - p_{i,t} = (1 - p_{i,t-1})\xi^G_{i,t} + \delta p_{i,t-1}\xi^G_{i,t}, \tag{10}$$

and now the generalized probability that a node does not receive an infection at time t, $\xi^G_{i,t}$, is given by

$$\xi^G_{i,t} = \prod_{j\sim i}(1 - \beta p_{j,t-1})\prod_{j\not\sim i}(1 - d_{ij}x^{d_{ij}-1}\beta p_{j,t-1}). \tag{11}$$

The first term in this expression, which corresponds to $\xi_{i,t}$, represents the probability that a node is not infected by close contacts in the social network. The second term accounts for the probability that a node does not receive an infection from its casual contacts. Obviously, as $x \to 0$, $\xi^G_{i,t} \to \xi_{i,t}$ and $\mathbf{\Gamma}(G,x = 0) \to \mathbf{A}(G)$, which means that GNLDS-MMCA is reduced to the NLDS-MMCA model. In this context, the parameter x controls the feasibility that an infected node can transmit an infection in only one step to others that are not its close contacts.

For the GNLDS-MMCA model the epidemic threshold τ for an undirected network is determined [as in (8)] by the largest eigenvalue of the generalized network matrix $\mathbf{\Gamma}(G,x)$, that is, $\tau = \frac{1}{\lambda_{1,\mathbf{\Gamma}(x)}}$ (see the Appendix).

V. APPLICATIONS OF THE GNLDS-MMCA MODEL

We start by analyzing the accuracy of the GNLDS-MMCA model by comparing it with the results obtained from simulations in random networks. We generate random networks with Poisson degree distribution using the Erdős-Rényi (ER) model [52] as well as random networks with power-law degree distributions according to the Barabási-Albert (BA) model [53]. These networks were generated by using NETWORKX [54] in PYTHON. In addition, we also generate scale-free networks having power-law degree distribution of the form $p(k) \sim k^{-\gamma}$ for a given power exponent $1.89 \leqslant \gamma \leqslant 3$, which were generated using the algorithm described in [55]. For the simulations we use the average of at least 100 individual runs that begins with a set of randomly chosen infected nodes (usually between 2.0% and 2.5% of the total number of nodes) and fixed values of parameters β and δ and the conductance x. Simulations evolve in steps of one time unit. During each step, an infected node i attempts to infect its nearest neighbors j with probability β and also nodes that are far away with probability $d_{ij}x^{d_{ij}-1}\beta$. Every infected node is cured with probability δ. An infection attempt on an already infected node has no effect. Simulations as well GNLDS-MMCA were implemented in C, and the programs are available on request. In Fig. 3 we illustrate the results of the exact GNLDS-MMCA and the simulations for the ER and BA random networks at different values of the conductance.

First of all, it can be seen that the GNLDS-MMCA reproduces very well the results obtained by simulating the infection spread in both types of networks. The second, and more important observation, is related to the relationship between the structure of these networks and the dynamics of the epidemic spreading. By keeping all other topological parameters identical, we can compare the effect of the degree distribution of a network on the propagation of an infection. It can be seen that the initial rate of propagation is faster in scale-free BA networks than in Poissonian ER ones. This is true for any values of the conductance studied. For instance, let $t_{1/2}$ be the time needed by an infection to infect 50% of the population in a given network. Then, for $x = 0$, $t_{1/2} \approx 50$ for ER networks with 1000 nodes and $\bar{k} = 6$, while it is only $t_{1/2} \approx 35$ for BA networks of the same size and average degree. As soon as we allow for casual contacts, the time needed to infect 50% of the population reduces dramatically, and it continues to be smaller for BA than for ER networks. For instance, a very small increase in the conductance to $x = 0.03$ drops this time to $t_{1/2} \approx 30$ in ER networks and to $t_{1/2} \approx 15$ in

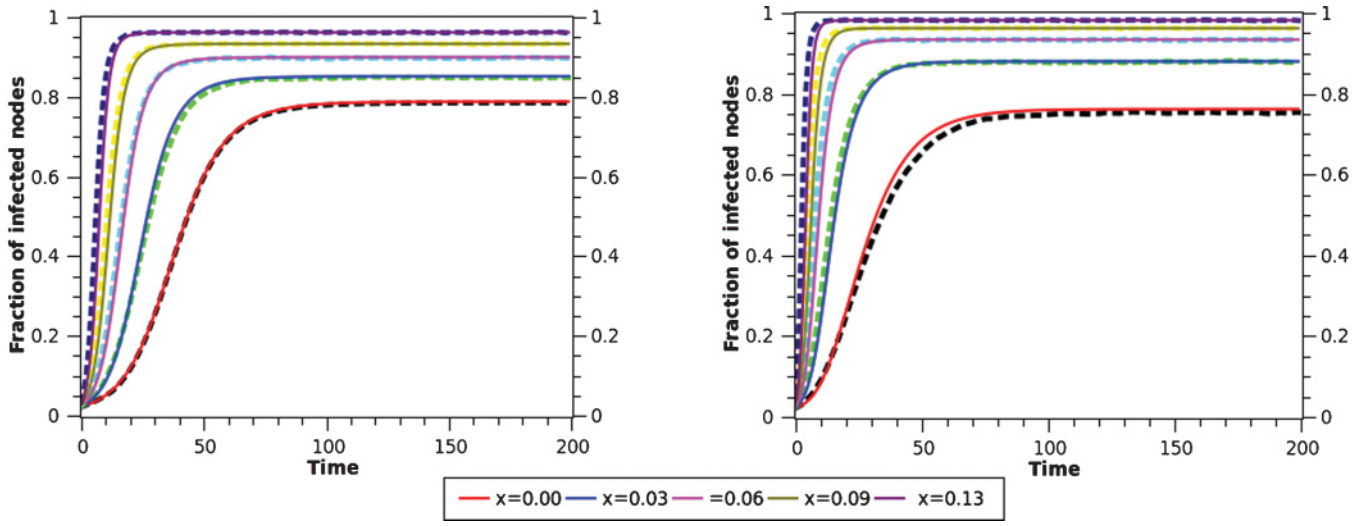

FIG. 3. (Color online) Results of the simulations (dashed lines) and the exact GNLDS-MMCA (solid lines) for (left) ER and (right) BA networks having $n = 1000$ nodes, $\bar{k} \approx 6$, with $\beta = 0.02$ and $\delta = 0.12$, and for different values of the parameter x. The results are the average of 250 realizations. The values of conductivity parameters are, from bottom to top, 0.0, 0.03, 0.06, 0.09, and 0.13.

BA ones. As $x \to 0.5$, the number of infected nodes on both networks tends to saturation.

Let us now introduce some results that relate network structure to infection dynamics in general networks. The first of these results is related to the epidemic threshold. The epidemic threshold determines whether the outbreaks will die out or spread and become epidemic. Let G be a connected undirected network with generalized contact matrix $\mathbf{\Gamma}(G,x)$. Let $\lambda_1(x) \geqslant \lambda_2(x) \geqslant \cdots \geqslant \lambda_n(x)$ be the eigenvalues of $\mathbf{\Gamma}(G,x)$. Then, the epidemic threshold for this network by considering a conductance equal to x is uniquely determined by $\tau = 1/\lambda_1(x)$ (see the Appendix). If the ratio between the birth and death rates of the infection is larger than the inverse of the largest eigenvalue of the generalized matrix, $\beta/\delta > 1/\lambda_1(x)$, the infection survives and becomes epidemic. Otherwise, the infection dies out over time. In addition, we have also proved mathematically that if $\beta/\delta > 1/\lambda_1(x_c)$, the infection survives and becomes epidemic for any $x \geqslant x_c$ (see the Appendix). Obviously, if $\beta/\delta > 1/\lambda_1(x_c = 0)$, the infection will become epidemic for any value of the conductance. In a similar way, if $\beta/\delta < 1/\lambda_1(x_c)$, the infection dies out for any value of conductance $x \leqslant x_c$. That is, if $\beta/\delta < 1/\lambda_1(x_c = 0.5)$ the infection will die out for any value of conductance smaller than 0.5. In a network with a star topology having n nodes we illustrate the dependence of the epidemic threshold with the conductance [Fig. 4(a)] as well as the progress of an infection [Fig. 4(b)]. Note that for any $0 \leqslant x \leqslant 0.5$, we have $\mathbf{\Gamma}(G,x) \geqslant \Gamma(G,0) = \mathbf{A}$, and

$$\lambda_1(x) = x(n-2) + \sqrt{x^2(n-2)^2 + (n-1)} \geqslant \sqrt{n-1} = \lambda_{1,\mathrm{A}} = \lambda_1(0).$$

Thus, when $\beta/\delta > 1/\lambda_1(x = 0)$, the infection survives for any value of the conductance, and the number of infected nodes saturates for relatively small times as $x \to 0.5$ [Fig. 4(b), left]. On the other hand, when $\beta/\delta < 1/\lambda_1(x = 0.5)$ the infection dies out [right in Fig. 4(b)]. We can see that as the conductance parameter is increased, the network is more resistant to the elimination of the infection. That is, when there is no conductance, about 10% of nodes remain infected at time $t = 100$. However, for $x = 0.13$ this percentage is about 30%.

We turn now our attention to the influence of the network heterogeneity on the rate of epidemic spreading. Concretely, we consider the variation of the power-law exponent in the degree distribution of scale-free networks. That is, we consider networks with 1000 nodes having power-law degree distribution of the form $p(k) \sim k^{-\gamma}$, $1.89 \leqslant \gamma \leqslant 3$. In Fig. 5 we illustrate the results obtained for two of these networks having $\gamma = 1.89$ and $\gamma = 1.98$. We explore different values of the conductance parameter, both by using our simulation strategy and by using the GNLDS-MMCA model. As can be seen for $x = 0$, there are no significant differences between the epidemics spreading in both networks. However, even for relatively low values of the conductance the differences between the spreading in both kinds of networks are quite significant. For instance, the networks with power-law coefficient $\gamma = 1.98$ have about 20% more nodes infected for $x = 0.03$ than when $x = 0$. A small drop of the power-law exponent to $\gamma = 1.89$ almost doubles the percentage of infected nodes for $x = 0.03$ in comparison with the network having $\gamma = 1.98$.

In general, as can be seen in Fig. 6, the rate of epidemic spreading measured by $t_{1/2}$ increases very fast with the increase of the power-law exponent γ and decreases with the increases of the conductance x. The value of $t_{1/2}$ scales as a negative exponential of the parameter x and as a power law of the exponent γ: $t_{1/2} \approx 24.36\exp(-13.58x) + 0.44\gamma^{3.38} - 10.56$. In other words, an epidemic spreads much faster in a network with high heterogeneous degree distribution than in one with more regularity, i.e., for small values of γ. This rate of spreading is significantly increased if casual contacts (LR interactions) are present, in which case the rate of spreading is exponentially affected by small variations of the conductance parameter. The term $d_{ij}x^{d_{ij}-1}$influences directly

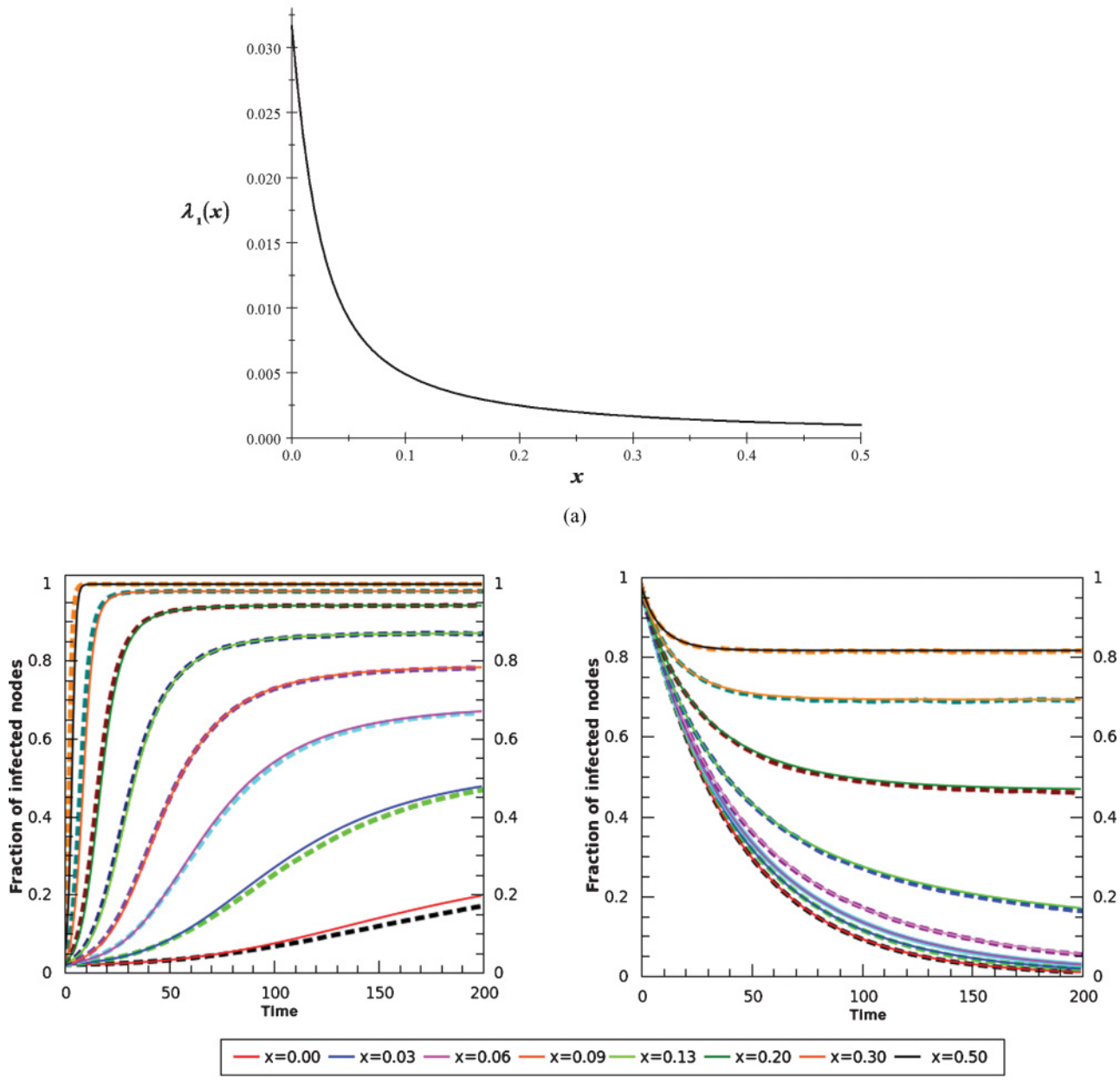

FIG. 4. (Color online) (a) Evolution of the epidemic threshold for a star graph with $n = 1000$ nodes as a function of x. (b). Results of the simulations (dashed lines) and the exact GNLDS-MMCA (solid lines) for a star graph with $n = 1000$ nodes, where $\tau = 1/\sqrt{999} \approx 0.0316$. Parameters in the left plot are $\beta = 0.002$, $\delta = 0.01$ and starting with about 20% nodes infected; since $\beta/\delta = 0.01 > \tau$, the infection will become epidemic for any value of the conductance x. Parameters in the right plot are $\beta = 0.0002$, $\delta = 0.024$; thus $\beta/\delta = 0.0083 < \tau$, and even starting with all nodes infected, the epidemic dies out for all values of the parameter x. The results are the average of 100 realizations. The values of conductivity parameter are, from bottom to top, 0.0, 0.03, 0.06, 0.09, 0.13, 0.20, 0.30, and 0.50.

the probability with which an infection spreads through casual contacts in a network. In order to understand the basic differences between the consideration of casual contacts in the spread of epidemics in networks with Poissonian and scale-free degree distributions we start by considering how distances are distributed in both types of networks as a function of the node degrees. In Fig. 7 we illustrate the plot of the probability of infection (infectability) $\overline{d}(k)x^{\overline{d}(k)-1}\beta$ versus k for networks with Poissonian and scale-free degree distributions. Here $\overline{d}(k)$ is the average shortest-path distance for nodes having degree k, and we use a fixed value of the parameter β. Nodes with large degree tend to have small average shortest-path distance, which means that they are closer to the rest of the nodes than nodes with low degree. Then, for a given value of $0 < x < 0.5$ the term $\overline{d}(k)x^{\overline{d}(k)-1}\beta$ is larger for smaller distances and decreases as the distance separating a pair of nodes increases. Consequently, the most connected nodes in the network display the largest infectability, which means

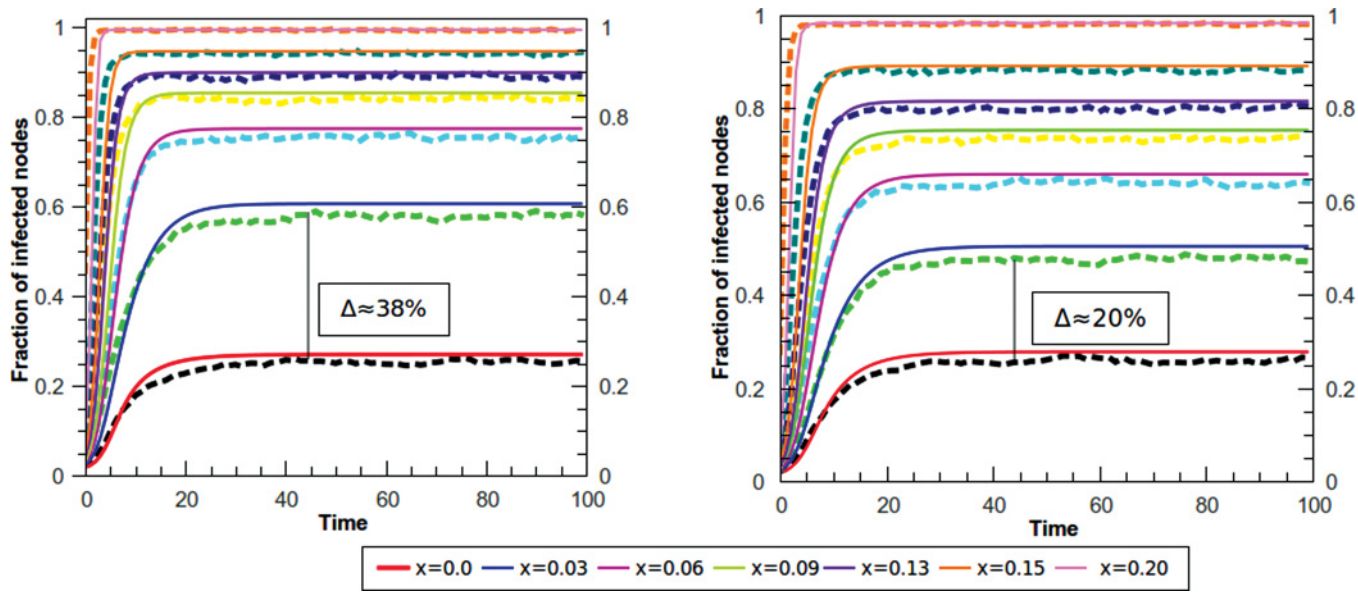

FIG. 5. (Color online) Results of the simulations (dashed lines) and the exact GNLDS-MMCA (solid lines) for networks with power-law degree distributions $p(k) \sim k^{-\gamma}$ with (left) $\gamma = 1.89$ and (right) $\gamma = 1.98$. The results are the average of 100 realizations for networks with $n = 1000$ nodes, $\beta = 0.02$, and $\delta = 0.12$ and for different values of the parameter x. The values of conductivity parameter are, from bottom to top, 0.0, 0.03, 0.06, 0.09, 0.13, 0.15, and 0.20.

that the probability that they are infected through casual contacts is very high. In networks with power-law degree distributions there are nodes with much higher degree than in Poissonian networks of the same size and density. Thus, these nodes are very susceptible to being infected through casual encounter transmission, and once they are infected, they can spread the infection in a very effective way, both by close and casual contact transmission. As the value of the conductance increases, the infectability is also increased (see Fig. 7), which explains why in scale-free networks the infection spreads so fast when the conductance increases.

The situation occurring here can be illustrated by considering the degrees of every node in a scale-free (SF) network. In Fig. 8(a) we illustrate the typical values for a BA network in which we simply represent the nodes in the abscissa and the degree of the nodes in the ordinate for $x = 0$. This plot looks like any typical distribution of the degrees in a SF network, with very few nodes of high degree and many of low degree. Now, if we explore what happens when

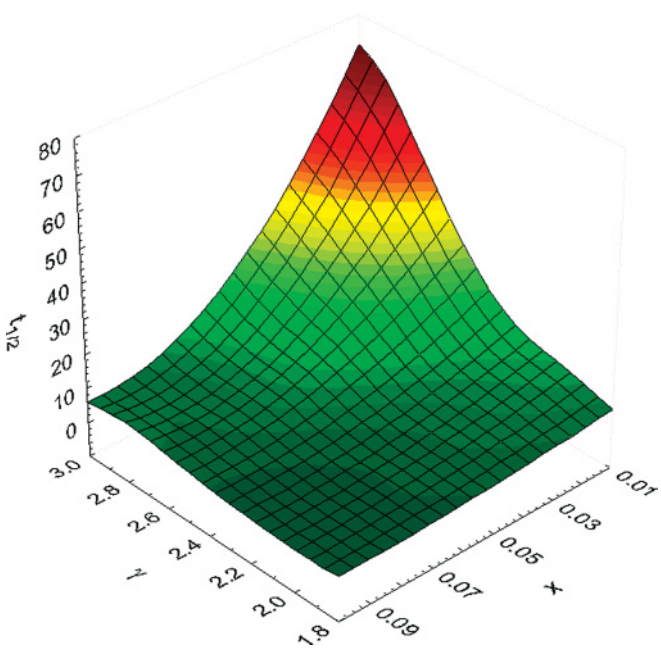

FIG. 6. (Color online) Rate of epidemic spreading measured by $t_{1/2}$, i.e., the time needed to infect 50% of the whole population, for different values of the power-law exponent γ and for different values of the conductance x.

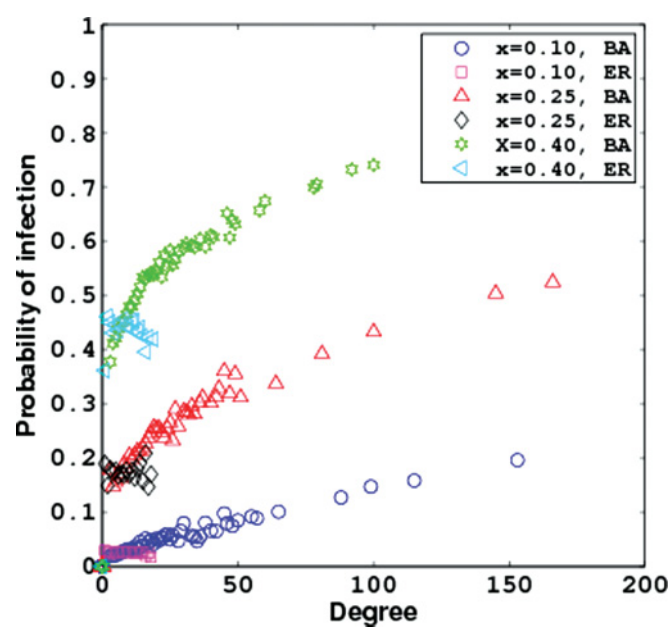

FIG. 7. (Color online) Probability of infection $\bar{d}(k)x^{\bar{d}(k)-1}$ for different values of the conductance x for BA and ER networks having the same number of nodes, $n = 1000$, and the same average degree.

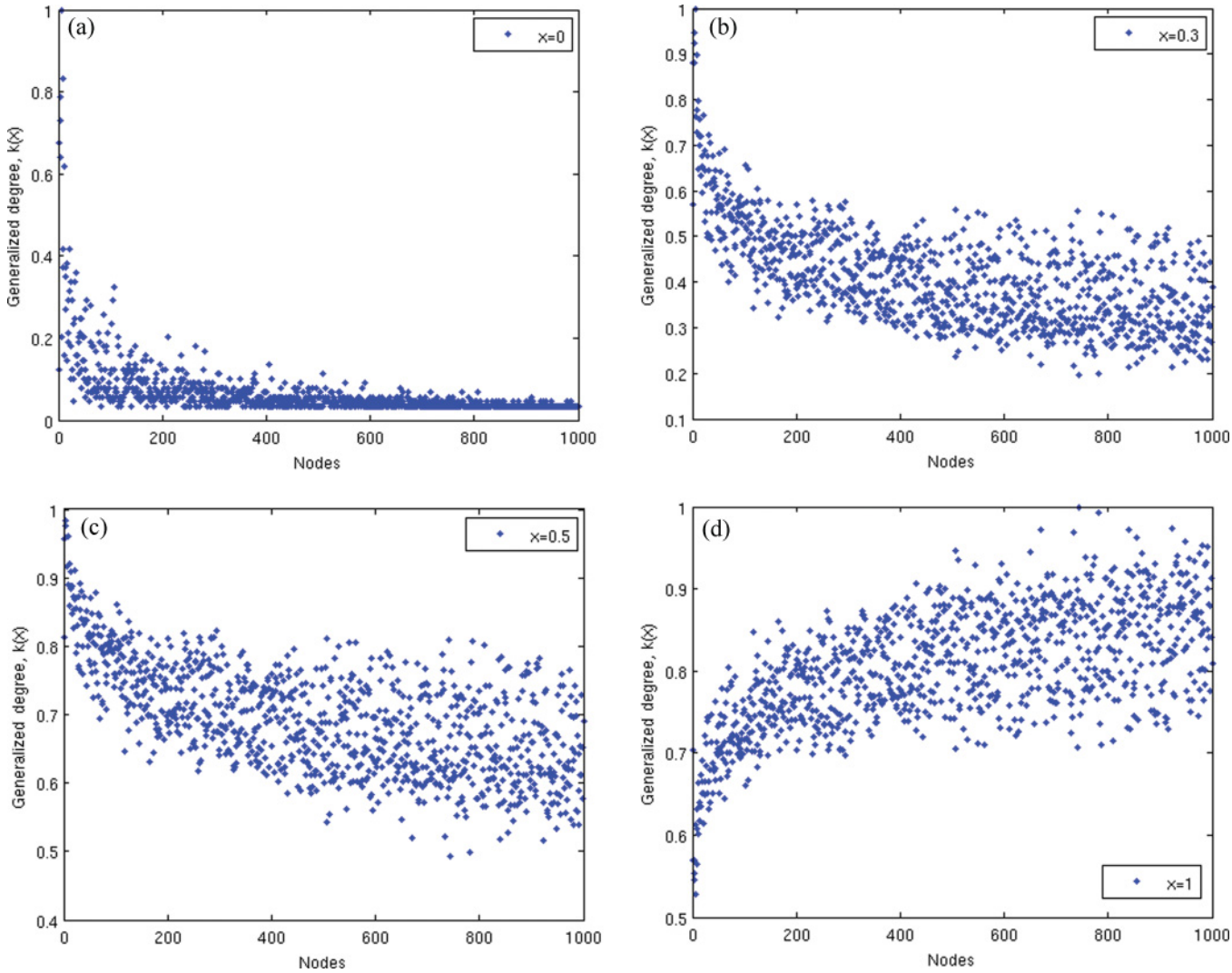

FIG. 8. (Color online) Illustration of the evolution of the degree of the nodes in a network created with the BA model as the values of the conductance change. Here for a BA network with $n = 1000$ and $\bar{k} \approx 6$ we represent the values of the generalized degree $k(x)$ for every node for different values of x. Notice that there is an inversion of the centrality of the nodes as the values of the conductance increases.

the values of x increase, the results are very appealing [see Figs. 8(b)–8(d)]. As can be seen, there is an inversion in the population of high and low degree nodes in the network. After a certain value of x the original hubs of the network become the poorest connected in terms of the generalized degree $k_i(x) = \sum_{j=1}^{n} \mathbf{\Gamma}_{ij}(x)$. We allow for a while that x takes values up to 1. At the same time all nodes with low degree $k(x = 0)$ are now among the most central ones in the network according to $k(x = 1)$. This inversion is a direct consequence of the fact that $k_i(x = 1) = \sum_{j=1}^{n} d_{ij}$ is the sum of all distances from node i to the rest of the nodes in the network. Then, in Fig. 8(d) we observe the distribution of the distance sum for every node in the network. It has been shown that the node-node distribution of distances in SF network have Poisson-like shapes [56]. Then, as illustrated in Fig. 9, where we plot the cumulative generalized degree distributions for a BA network, as soon as we depart from the value of $x = 0$, the distributions become Poisson-like even for small values of the conductivity.

The results illustrated in Fig. 8 indicate that if we obtain the rank correlation between $k_i(x = 0)$ versus $k_i(x \neq 0)$ for all nodes i in the network, we will observe a point in which the initially positive correlation becomes negative. This is exactly what we observe in Fig. 10(a), where we plot the values of the rank correlation coefficient, measured by the Kendall ρ index, versus the values of the conductivity for networks with different values of the power-law exponent γ. It is interesting to note that the value of x at which the sign inversion occurs increases with the value of γ. That is, the more heterogeneous networks make the inversion of the rank correlation for smaller values of x than the more homogenous one. The value of x at which the inversion of rank correlation occurs (inversion point) changes as a sigmoid function with the power-law exponent [see Fig. 10(b)]. More exactly, it can be expressed as $x_{\text{inv}} \approx 0.099 \tanh(3.123\gamma - 7.357) + 0.736$. All in all, these results indicate that heterogeneous networks are very sensitive to the changes of the conductivity and consequently to increase dramatically the spread of infections when casual contacts are included.

We turn now our attention to the real world, where not only diseases can propagate in a network through close and

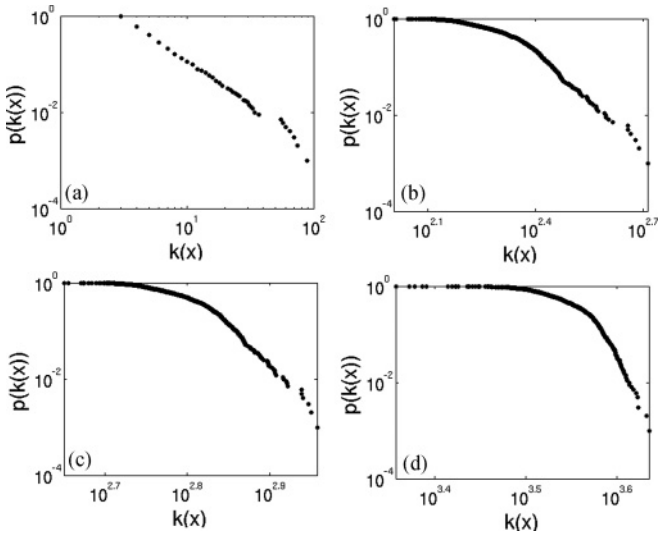

FIG. 9. Illustration of the evolution of the cumulative degree distribution of the nodes for the network studied in Fig. 8 for (a) $x = 0$, (b) $x = 0.3$, (c) $x = 0.5$, and (d) $x = 1$. Notice that there is change in the distribution from a power law at $x = 0$ to Poissonian-like distributions as x increases.

casual contacts but also attitudes, fads, fashion styles, and tendencies of a different nature can use similar mechanisms of propagation. We consider here a couple of real-world networks from two different scenarios. The first one is a network of collaboration between 1265 jazz musicians in which two nodes are linked if the respective musicians have collaborated in the same band [57]. The total number of such collaborations in this network is 32 358. The second network represents 1586 corporate directors of the 500 top corporations in the United States [58]. Here two nodes are connected if the corresponding directors share a position on the board of at least one corporation. In the first scenario we can think about the propagation of musical tendencies and styles in jazz, which can be diffused through the direct collaboration between musicians. In addition, two musicians that have not collaborated directly in a band can influence each other simply if they have listened or studied their respective music. In the second scenario we can be interested in the analysis of how strategic decisions taken in one corporation can be adopted by others. Such strategies can be transmitted by those directors who share positions on the board of more than one corporation, but they can also be propagated by casual encounters of the directors. In this case casual contacts can account for the way in which some directors analyze, copy, and adapt what other directors are doing in corporations where the first are not members of the direction board. For the network of jazz musicians, taking $\beta = 0.02$, $\delta = 0.12$, and null conductance, the infection propagates in a very fast way, infecting about 70% of the whole population for $t \geqslant 20$. This network has a large average degree, $\bar{k} = 50.6$, and a fat-tail degree distribution. Consequently, an infection propagates through close contacts in a very effective way due to the density of the network and the fact that each time one of the high degree nodes is infected, the infection is able to propagate to a large number of other nodes. The consideration of long-range interactions does not have a big impact in the infection spreading in this network. Here $t_{1/2} < 5$ for $x = 0$, and it is impossible to have a dramatic increase in the rate of propagation due to an increase in the conductance. However, as can be seen in Fig. 11, the consideration of a conductance of $x = 0.13$ increases the percentage of infected population to about 90%, and saturation is reached with small increases of this parameter. The situation is quite different for the network of the US corporate elite. First of all, the time at which 50% of the population is infected drops from $t_{1/2} \approx 40$ for $x = 0$ to $t_{1/2} \approx 7$ for $x = 0.13$. This represents a dramatic increase in the rate of propagation of attitudes among directors of the

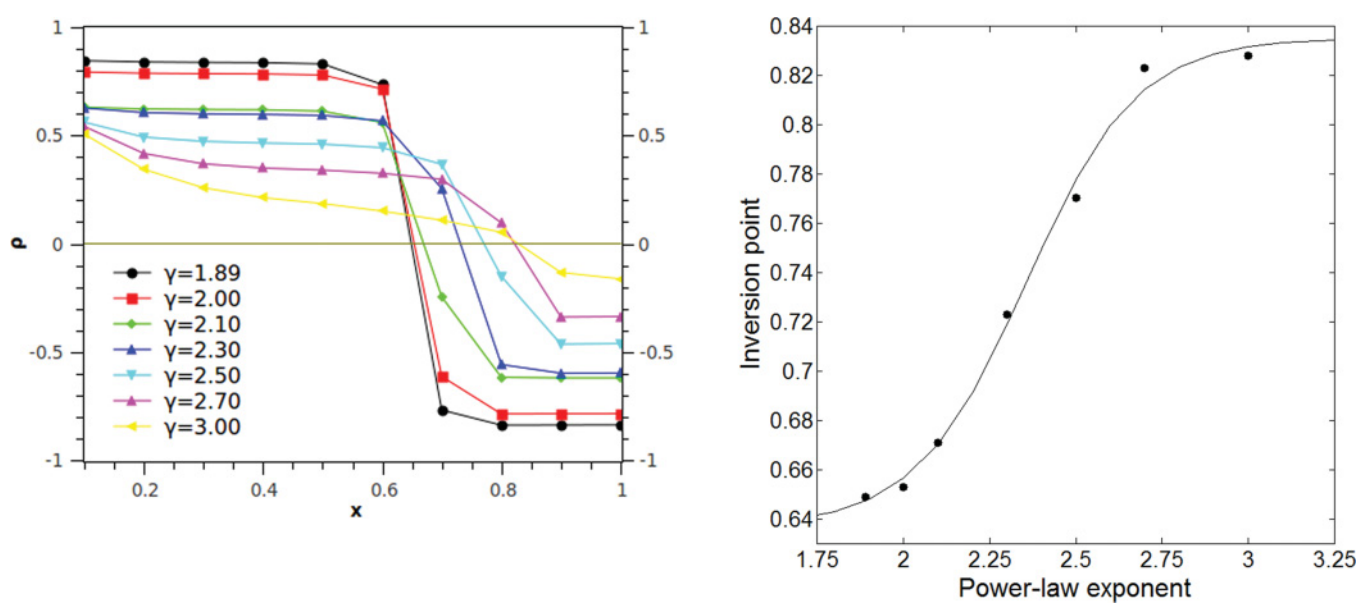

FIG. 10. (Color online) (a) Change of the rank correlation between $k_i(x = 0)$ and $k_i(x \neq 0)$ as a function of the conductance x for scale-free (SF) networks with different exponents of the power law. (b) Scaling of the conductance at which an inversion in the rank correlation occurs as a function of the exponent of the power law in SF networks (see text for explanations).

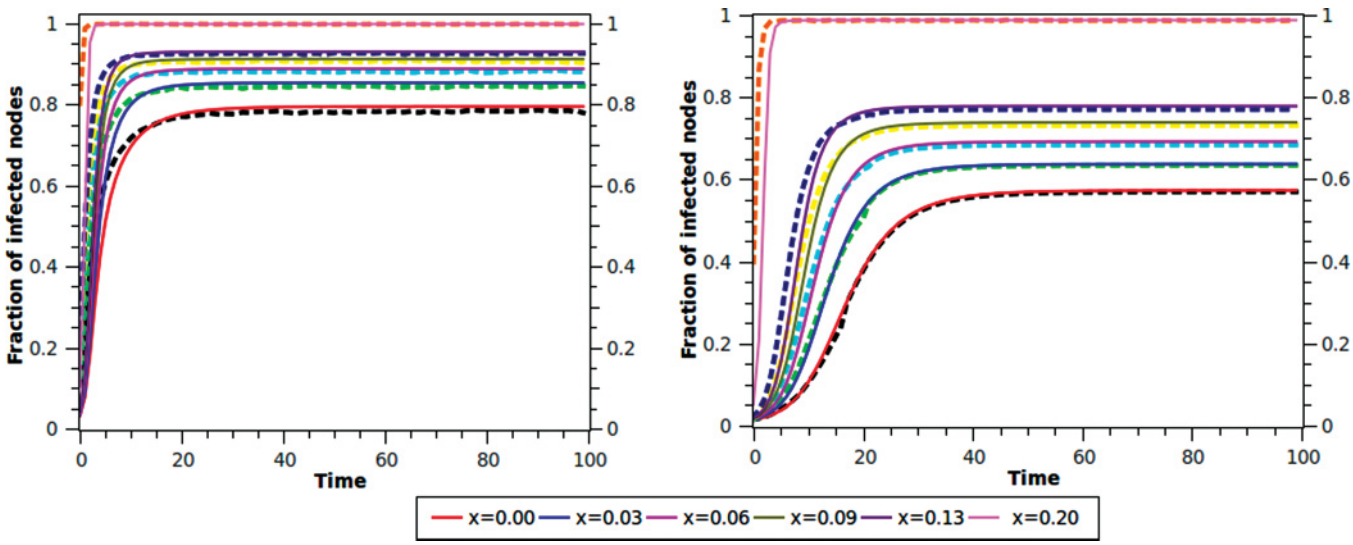

FIG. 11. (Color online) Results of the simulations and GNLDS-MMCA for the networks of collaboration among (left) jazz musicians and (right) for the corporate directors of the top 500 corporations in the United States. The results are the average of 250 realizations with $\beta = 0.02$, $\delta = 0.12$ and for different values of the conductance parameter x. The values of conductivity parameter are, from bottom to top, 0.0, 0.03, 0.06, 0.09, 0.13, and 0.20.

corporate elite in the United States if relatively small chances for casual contacts are allowed. In fact, the increase of the conductance up to $x = 0.13$ produces an increase of about 30% in the infected population in comparison with the consideration of direct contagion only.

VI. CONCLUSIONS

We have proposed here a way for accounting for the social contacts among individuals by considering that casual contacts can be inferred from the network of close contacts. We based our model on a series of empirical observations made in the epidemiological and social science literature. We model such casual contacts by means of the probability that two nonconnected individuals in a close contact social network have of creating a new link between them. Then, we use the principle that new social ties are created on the basis of the future value of this relationship to infer the casual contacts among individuals. In this model casual contacts are created on the basis of long-range interactions as a function of the social distance between two individuals, while close contacts are assumed to be determined by the links in the social network.

This approach is then integrated in an epidemic spreading model such as the NLDS-MMCA model. In this case we observe that there are two main factors influencing the rate of propagation of an epidemic in a complex network when both close and casual interactions are considered. The first one is the conductance parameter, which controls how feasible casual contacts are by means of LR interactions. If this conductance is set to zero, there is no possibility of contagion through casual contacts, and everything happens only by means of the close contacts among individuals, such as in the case of sexually transmitted diseases or computer viruses. As the conductance parameter increases, the rate of propagation increases dramatically, and the infection is less likely to die out. In these cases the number of infected nodes saturates in relatively short times after the initiation of the propagation. The second factor influencing the propagation is the heterogeneity of the network. It has been observed that epidemics are propagated much faster in scale-free networks than in more regular ones. Furthermore, in scale-free networks the influence of the conductance parameter to the propagation is significantly more marked than in networks with Poissonian degree distributions. All in all, an infection propagates very fast in heterogeneous networks when the number of casual contacts is large, making the infections easily become epidemics with high resistance to dying out. As we have shown here, GNLDS-MMCA can be a useful tool for understanding important problems in modern societies, ranging from viral epidemics to the propagation of attitudes and consumer styles.

ACKNOWLEDGMENTS

E.E. is thankful for the useful discussions about MMCA with A. Arenas and S. Gómez (Universitat Rovira i Virgili) as well as partial financial support from the New Professor's Fund, University of Strathclyde. F.K.M. acknowledges the Overseas Research Students Award Scheme (ORSAS) and the University of Strathclyde (Mathematics and Statistics Department) for funding and support. We thank both referees for constructive comments that improved considerably the presentation of this work.

APPENDIX: ANALYTIC RESULTS ON GNLDS-MMCA

Theorem 1. Let G be a connected undirected network with generalized matrix $\mathbf{\Gamma}(G,x)$. Let $\lambda_1(x) \geqslant \lambda_2(x) \geqslant \cdots \geqslant \lambda_n(x)$ be the eigenvalues of $\mathbf{\Gamma}(G,x)$. Then, the epidemic threshold for this network by considering a conductance equal to x ($0 \leqslant x \leqslant 0.5$) is uniquely determined by $\tau = 1/\lambda_1(x)$.

Proof. The proof follows similar lines as in Ref. [49]. In the GNLDS-MMCA model, the probability $1 - p_{i,t}$ that the node remains healthy in the network is given by

$$1 - p_{i,t} = (1 - p_{i,t-1})\xi_{i,t}^{G} + \delta p_{i,t-1}\xi_{i,t}^{G}, \qquad \text{(A1)}$$

where the generalized probability that a node does not receive infection at time t, $\xi_{i,t}^G$, is given by

$$\xi_{i,t}^G = \prod_{j\sim i}(1-\beta p_{j,t-1})\prod_{j\not\sim i}(1-d_{ij}x^{d_{ij}-1}\beta p_{j,t-1}). \quad \text{(A2)}$$

Note that (A1) can be expressed as

$$\mathbf{p_t} = F(\mathbf{p_{t-1}}),$$

with

$$F_i(\mathbf{p_{t-1}}) = \mathbf{1} - (\mathbf{1}-\mathbf{p_{i,t-1}})\xi_{\mathbf{i,t}}^{\mathbf{G}} + \delta\mathbf{p_{i,t-1}}\xi_{\mathbf{i,t}}^{\mathbf{G}}. \quad \text{(A3)}$$

The infection dies out when $p_i = 0$ for all i. If $p_{i,t-1} = 0$ for all i, it follows from (A1) and (A2) that $p_{i,t} = 0$ and therefore $\mathbf{p} = \mathbf{0}$ is a *fixed point* of the system (A3). Thus, we need to see if $\mathbf{p} = \mathbf{0}$ is a fixed point *asymptotically stable* in the dynamical system (A3) and, proceeding as in [49], we will use the following lemma.

Lemma 1 (asymptotic stability). The system is asymptotically stable at $\mathbf{p} = \mathbf{0}$ if the eigenvalues of $\nabla F(\mathbf{0})$ are less than 1 in absolute value, where $[\nabla F(\mathbf{0})]_{i,j} = \frac{\partial F_i}{\partial p_j}|_{p=0}$.

From (A3) we have

$$[\nabla F(\mathbf{0})]_{i,j} = \begin{cases} \beta & \text{for } d_{ij} = 1, \\ \beta d_{ij}x^{d_{ij}-1} & \text{for } j \neq i, d_{ij} \neq 1, \\ 1-\delta & \text{for } j = i, \end{cases} \quad \text{(A4)}$$

and by the definition of the generalized network matrix $\mathbf{\Gamma}(G,x)$ in section 2,

$$[\nabla F(\mathbf{0})]_{i,j} = \begin{cases} \beta\Gamma_{ij} & \text{for } j \neq i, \\ 1-\delta & \text{for } j = i. \end{cases} \quad \text{(A5)}$$

Thus, we can obtain the system matrix $\mathbf{S}$

$$\mathbf{S}(x) = \nabla F(\mathbf{0}) = (1-\delta)\mathbf{I} + \beta\mathbf{\Gamma}(G,x), \quad \text{(A6)}$$

which describes the behavior of the virus when it is about to die. As in Lemma 2 in [49], it can be easily shown that the eigenvalues of $\mathbf{S}(x)$ are given by

$$\lambda_{i,\mathbf{S}(x)} = 1-\delta+\beta\lambda_{i,\mathbf{\Gamma}(x)}, \quad \text{(A7)}$$

where $\lambda_{i,\Gamma(x)}$ denote the eigenvalues of $\mathbf{\Gamma}(G,x)$, and the eigenvectors of $\mathbf{S}(x)$ are the same as those for $\mathbf{\Gamma}(G,x)$. Hence, by Lemma 1, the system is asymptotically stable when

$$|\lambda_{i,\mathbf{S}(x)}| < 1 \quad \forall i, \forall x \quad \text{(A8)}$$

Now, since $\mathbf{\Gamma}(G,x)$ is a real symmetric matrix, its eigenvalues are real, and by (A7), the eigenvalues of $\mathbf{S}(x)$ are real too. Also, since G is connected, $\mathbf{\Gamma}(G,x)$ represents the adjacency matrix of a weighted complete graph, and therefore it is irreducible. Thus, $\mathbf{\Gamma}(G,x)$ is a real, symmetric, non-negative, irreducible, and square matrix. Under these conditions, the Perron-Frobenius theorem states that the largest eigenvalue $\lambda_{1,\Gamma(x)}$ [called the Perron root of $\mathbf{\Gamma}(G,x)$] is positive and simple. Thus

$$\lambda_{1,\mathbf{\Gamma}(x)} = |\lambda_{1,\mathbf{\Gamma}(x)}| > \lambda_{i,\mathbf{\Gamma}(x)} \quad \forall i > 1, \quad \text{(A9)}$$

and

$$\lambda_{1,\mathbf{S}(x)} = |\lambda_{1,\mathbf{S}(x)}| > |\lambda_{i,\mathbf{S}(x)}| \quad \forall i > 1, \forall x. \quad \text{(A10)}$$

Using (A7) and (A8),

$$\lambda_{1,\mathbf{S}(x)} = 1-\delta+\beta\lambda_{1,\mathbf{\Gamma}(x)} < 1. \quad \text{(A11)}$$

Thus, if the epidemic dies out, we must have

$$\frac{\beta}{\delta} < \tau = \frac{1}{\lambda_{1,\mathbf{\Gamma}(x)}}. \quad \text{(A12)}$$

In order to complete the proof of Theorem 1, we need to see that if $\frac{\beta}{\delta} < \tau = \frac{1}{\lambda_{1,\mathbf{\Gamma}(x)}}$, then the epidemic will die out over time (sufficiency of epidemic threshold).

In (A2), since all terms β, $p_{j,t-1}$, and $d_{ij}x^{d_{ij}-1}$ are non-negative and not greater than 1,

$$\prod_{j\sim i}(1-\beta p_{j,t-1}) \geqslant 1-\beta\sum_{j\sim i}p_{j,t-1},$$

and

$$\prod_{j\not\sim i}(1-d_{ij}x^{d_{ij}-1}\beta p_{j,t-1}) \geqslant 1-\beta\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}.$$

Thus,

$$\begin{aligned}\xi_{i,t}^G &\geqslant \left(1-\beta\sum_{j\sim i}p_{j,t-1}\right)\left(1-\beta\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}\right)\\ &= 1-\beta\sum_{j\sim i}p_{j,t-1}-\beta\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}\\ &\quad+\beta^2\left(\sum_{j\sim i}p_{j,t-1}\right)\left(\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}\right),\end{aligned}$$

and since

$$\beta^2\left(\sum_{j\sim i}p_{j,t-1}\right)\left(\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}\right) \geqslant 0,$$

$$\begin{aligned}\xi_{i,t}^G &\geqslant 1-\beta\sum_{j\sim i}p_{j,t-1}-\beta\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}\\ &= 1-\beta\left(\sum_{j\sim i}p_{j,t-1}+\sum_{j\not\sim i}d_{ij}x^{d_{ij}-1}p_{j,t-1}\right)\\ &= 1-\beta\sum_j\Gamma_{ij}p_{j,t-1}. \quad \text{(A13)}\end{aligned}$$

Thus, from (A1)

$$\begin{aligned}1-p_{i,t} &= (1-p_{i,t-1})\xi_{i,t}^G+\delta p_{i,t-1}\xi_{i,t}^G \geqslant (1-(1-\delta)p_{i,t-1})\\ &\quad\times\left(1-\beta\sum_j\Gamma_{ij}p_{j,t-1}\right) = 1-(1-\delta)\,p_{i,t-1}\\ &\quad-\beta\sum_j\Gamma_{ij}p_{j,t-1}+(1-\delta)\,p_{i,t-1}\beta\sum_j\Gamma_{ij}p_{j,t-1},\end{aligned}$$

and then

$$\begin{aligned}p_{i,t} &\leqslant (1-\delta)\,p_{i,t-1}+\beta\sum_j\Gamma_{ij}p_{j,t-1}\\ &\quad-(1-\delta)\,p_{i,t-1}\beta\sum_j\Gamma_{ij}p_{j,t-1}\\ &\leqslant (1-\delta)\,p_{i,t-1}+\beta\sum_j\Gamma_{ij}p_{j,t-1}, \quad \text{(A14)}\end{aligned}$$

which can be expressed in matrix form as

$$\mathbf{p}_t \leqslant [(1-\delta)\mathbf{I} + \beta\mathbf{\Gamma}(G,x)]\mathbf{p}_{t-1}, \tag{A15}$$

which uses the same system matrix as (A6),

$$\mathbf{p}_t \leqslant \mathbf{S}(x)\,\mathbf{p}_{t-1} \leqslant \mathbf{S}^2(x)\,\mathbf{p}_{t-2} \leqslant \cdots \leqslant \mathbf{S}^t(x)\,\mathbf{p_0}$$

$$\leqslant \sum_i \lambda^t_{i,\mathbf{S}(x)}\mathbf{u}_{i,\mathbf{S}}\mathbf{u}'_{i,\mathbf{S}}\mathbf{p}_0,$$

where the last inequality is the spectral decomposition of $\mathbf{S}^t$. By (A7), when $\frac{\beta}{\delta} < \frac{1}{\lambda_{1,\mathbf{\Gamma}(x)}}$, then $\lambda_{i,\mathbf{S}(x)} < 1$ and $\lambda^t_{i,\mathbf{S}(x)} \approx 0$ for all i and large t, which makes $\mathbf{p}_t \approx 0$ as t increases, implying that the infection dies out over time.

Finally, we make use of the following Theorem which states the monotonicity property of the Perron root for nonnegative and irreducible square matrices (see Theorem 1.3 in [59]).

Theorem 2. Let $\mathbf{A}$ and $\mathbf{B}$ be non-negative matrices of order $n \geqslant 1$. If $\mathbf{A} \leqslant \mathbf{B}$, then the Perron roots of $\mathbf{A}$ and $\mathbf{B}$ satisfy the inequality $\lambda_{1,A} \leqslant \lambda_{1,B}$. Furthermore, if $\mathbf{B}$ is irreducible and $\mathbf{A} \neq \mathbf{B}$, then the inequality holds strictly: $\lambda_{1,\mathbf{A}} < \lambda_{1,\mathbf{B}}$.

Corollary. If $\beta/\delta > 1/\lambda_{1,\mathbf{\Gamma}(x_c)}$, the infection survives and becomes epidemic for any $x \geqslant x_c$. And if $\beta/\delta < 1/\lambda_{1,\mathbf{\Gamma}(x_c)}$, the infection dies out for any value of the conductance $x \leqslant x_c$.

Proof. Let $0 \leqslant x_1 \leqslant x_2$; if $d_{ij} = 1$, then $\Gamma_{ij}(x_1) = \Gamma_{ij}(x_2) = 1$, and if $d_{ij} > 1$, then $\Gamma_{ij}(x_1) = d_{ij}x_1^{d_{ij}-1} < \Gamma_{ij}(x_2) = d_{ij}x_2^{d_{ij}-1}$. Thus $\Gamma_{ij}(x_1) \leqslant \Gamma_{ij}(x_2)$ and $\mathbf{\Gamma}(G,x_1) \leqslant \mathbf{\Gamma}(G,x_2)$, and the result is a direct consequence of Theorem 2.

[1] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[2] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[3] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[4] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).
[5] N. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications* (Griffin, London, 1975).
[6] A. L. Lloyd and R. M. May, Science **292**, 1316 (2001).
[7] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[8] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2001).
[9] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 2002).
[10] S. Eubank, H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang, Nature (London) **429**, 180 (2004).
[11] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).
[12] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, Proc. Natl. Acad. Sci. USA **103**, 2015 (2006).
[13] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, Bull. Math. Biol. **68**, 1893 (2006).
[14] S. Riley, Science **316**, 1298 (2007).
[15] S. Meloni, A. Arenas, and Y. Moreno, Proc. Natl. Acad. Sci. USA **106**, 16897 (2009).
[16] D. Balcan, V. Colizza, B. Goncalves, H. Hu, J. J. Ramasco, and A. Vespignani, Proc. Natl. Acad. Sci. USA **106**, 21484 (2009).
[17] J. M. Read, K. T. D. Eames, and W. J. Edmunds, J. R. Soc. Interface **5**, 1001 (2008).
[18] D. M. Skowronski, C. Astell, R. C. Brunham, D. E. Low, M. Petric, R. L. Roper, P. J. Talbot, T. Tam, and L. Babiuk, Annu. Rev. Med. **56**, 357 (2005).
[19] I. M. Longini, Jr., J. S. Koopman, A. S. Monto, and J. P. Fox, Am. J. Epidemiol. **115**, 736 (1982).
[20] F. Ball and P. Neal, Math. Biosci. **180**, 73 (2002).
[21] H.-P. Duerr, M. Schwehm, C. C. Leary, S. J. De Vlas, and M. Eichner, Epidemiol. Infect. **135**, 1124 (2007).
[22] F. Ball and P. Neal, Math. Biosci. **212**, 69 (2008).
[23] D. G. Blanchflower, A. J. Oswald, and B. Landeghem, J. Eur. Econ. Assoc. **7**, 528 (2009).
[24] E. Cohen-Cole and J. M. Fletcher, J. Health Econ. **27**, 1382 (2008).
[25] N. A. Christakis and J. H. Fowler, N. Engl. J. Med. **9**, 357 (2007).
[26] L. P. Boss, Epidemiol. Rev. **19**, 233 (1997).
[27] P. Ormerod and G. Wiltshire, Mind Soc. **8**, 135 (2009).
[28] D. Brockmann, L. Hufnagel, and T. Geisel, Nature (London) **439**, 462 (2006).
[29] M. C. González, C. A. Hidalgo, and A.-L. Barabási, Nature (London) **453**, 779 (2008).
[30] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. Scalia Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinka, and W. J. Edmunds, PloS Med. **5**, e74 (2008).
[31] M. J. Keeling and K. T. D. Eames, J. R. Soc. Interface **2**, 295 (2005).
[32] M. McPherson, L. Smith-Lovin, and J. M. Cook, Annu. Rev. Sociol. **27**, 415 (2001).
[33] B. Song, C. Castillo-Chavez, and J. P. Aparicio, Math. Biosci. **180**, 187 (2002).
[34] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000).
[35] M. E. J. Newman, I. Jensen, and R. M. Ziff, Phys. Rev. E **65**, 021904 (2002).
[36] M. Boots and A. Sasaki, Proc. R. Soc. London, Ser. B **266**, 1933 (1999).
[37] J. Verdasca, *in BioMat 2005: Proceedings of the International Symposium on Mathematical and Computational Biology*, edited by R. P. Mondani and R. Dilão (World Scientific, Singapore, 2006), p. 171.
[38] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 409 (1998).
[39] T. Cohen, C. Colijn, B. Finklea, and M. Murray, J. R. Soc. Interface **4**, 523 (2007).
[40] J. Knoben and L. A. G. Oerlemans, Int. J. Manage. Rev. **8**, 71 (2006).

[41] A. Agrawal, D. Kapur, and J. McHale, J. Urban Econ. **64**, 258 (2008).
[42] S. F. Reardon and G. Firebaugh, Sociol. Methodol. **32**, 85 (2002).
[43] M. O. Jackson and A. Wolinsky, J. Econ. Theory **71**, 44 (1996).
[44] M. D. Ryall and O. Sorenson, Manage. Sci. **53**, 566 (2007).
[45] J. H. Manson, Primates **35**, 417 (1994).
[46] H. Bierman, Jr., *An Introduction to Accounting and Managerial Finance: A Merger of Equals* (World Scientific, Singapore, 2010), Chap. 2.
[47] E. Estrada, Chem. Phys. Lett. **336**, 248 (2001).
[48] E. Estrada and F. Kalala-Mutombo (unpublished).
[49] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, ACM Trans. Inf. Syst. Secur. **10**, 13 (2008).
[50] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Europhys. Lett. **89**, 38009 (2010).
[51] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Int. J. Complex Syst. Sci. **1**, 47 (2011).
[52] P. Erdős and A. Rényi, Publ. Math. **6**, 290 (1959).
[53] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[54] A. A. Hagberg, D. A. Schult, and P. J. Swart, in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, edited by G. Varoquaux, T. Vaught, and J. Millman (Pasadena, CA, USA, 2008), pp. 11–15.
[55] F. Viger and M. Latapy, in *The Eleventh International Computing and Combinatorics Conference* (Springer, Berlin, 2005), pp. 440–449.
[56] K. Malarz, J. Karpinska, A. Kardas, and K. Kulakowski, e-print arXiv:cond-mat/0309255v2.
[57] P. Gleiser and L. Danon, Adv. Complex Syst. **6**, 565 (2003).
[58] G. F. Davis, M. Yoo, and W. E. Baker, Strategic Organ. **1**, 302 (2003).
[59] Y. A. Al'pin and L. Y. Kolotilina, J. Math. Sci. **141**, 1576 (2007).