

Role of clustering and gridlike ordering in epidemic spreading

Thomas Petermann* and Paolo De Los Rios

Laboratoire de Biophysique Statistique, ITP-FSB, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

(Received 18 December 2003; published 4 June 2004)

The spreading of an epidemic is determined by the connectivity patterns which underlie the population. While it has been noted that a virus spreads more easily on a network in which global distances are small, it remains a great challenge to find approaches that unravel the precise role of local interconnectedness. Such topological properties enter very naturally in the framework of our two-time-step description, also providing an approach to track a probabilistic system. The method is elaborated for SIS-type epidemic processes, leading to a quantitative interpretation of the role of loops up to length 4 in the onset of an epidemic.

DOI: 10.1103/PhysRevE.69.066116

PACS number(s): 89.75.-k, 05.10.-a, 87.23.-n

I. INTRODUCTION

Almost all of us have already met someone far from home, who turns out to share a friend with us. Milgram first put this social phenomenon on a firm basis, finding that any two individuals are separated, on average, by six acquaintances [1]. Furthermore, our society is composed of groups within which, if individual A is acquainted with persons B and C, then B is likely to know C. In the language of social networks, this means that the individuals A, B, and C are arranged on a triangle. These two facts tell us that in the social universe, global distances are small, and local interconnectedness, i.e., *clustering*, is high. A network that possesses both of these topological properties is called a *small world*. In order to find a model that accounts for both of these properties, one could choose a regular lattice whose nodes are indeed locally highly interconnected. But since global distances are large in this type of network, it is an ineligible candidate for a small world model. In a random graph [2], which is a set of nodes (e.g., laid out on a virtual circle) with connections between them established at random, a link is more likely to point to a far away node than to one close by. (Here we refer to the distance on the virtual circle's circumference.) As a consequence, the presence of the very many long-range links accounts for the desired global property, but the degree of clustering is low. Watts and Strogatz combined these two insights and proposed a model that interpolates between a regular lattice and a random graph, thus capturing both the global and local topological properties mentioned above [3]. But the small world property is not merely exhibited by social networks. Yet, through the increased availability of data, it was found that the simultaneous occurrence of high local *and* global interconnectedness is prominent to a much wider class of systems, notably the Internet [4,5], the World Wide Web [6], metabolic networks [7], and food webs [8].

In addition to the small world property, there is another important fact when it comes to characterizing the topology of a complex network: not all the nodes have the same number of edges. The corresponding measure is the *degree dis-*

tribution $P(k)$ which gives the probability that a randomly chosen node has degree k , that is, k edges. For most of the just mentioned examples, this distribution was found to be $P(k) \sim k^{-\gamma}$, $2 < \gamma \leq 3$, implying the absence of a characteristic (degree) scale, hence the name *scale-free* network. This emergence of scaling can be understood, for example, in terms of *growing* networks. Starting from a small core graph, at each time step a node is added together with a certain number of edges that are connected to existing nodes, the latter being chosen *preferentially*, that is, a link is more likely established to a high-degree node [9]. This fat-tailed degree distribution implies the presence of *hubs* (high-degree nodes) which hold together the network and play a crucial role in issues such as robustness or fragility [10].

This explosion of research activity in the field of complex networks has also shed light on epidemic spreading since the latter can be regarded as a dynamical process occurring on a complex network: a computer virus spreads on the Internet, and HIV (as a biological example) propagates on top of the web of human sexual contacts. Also these two examples fall into the category of scale-free networks [11,12], and it is this topological property that accounts for the absence of a finite epidemic threshold in the corresponding spreading phenomenon [13].

The degree distribution is only a first way to characterize the degree related topology of a complex network. Indeed, by analyzing scientific collaboration networks, researchers working with many others (high-degree nodes) tend to collaborate with other “hubs.” This means that there exist *degree correlations*, and the just mentioned property has been called *assortative mixing*, holding generally for social networks [14]. On the other hand, in the Internet (at the autonomous system level), high-degree nodes are more likely connected to low-degree ones, thus exhibiting *disassortative* mixing. The influence of such degree correlations on the spreading of an epidemic was investigated in detail [15], finding that neither assortative nor disassortative scale-free networks exhibit a finite epidemic threshold [16]. The insights about the role of these degree related topological properties in epidemic spreading have been gained at the mean-field level.

Besides degree correlations, triangles are ubiquitous in complex networks as outlined in the first paragraph, and

*Electronic address: Thomas.Petermann@alumni.ethz.ch

more generally, many loops of short length were found in these systems [17]. Motivated by Watts and Strogatz' model [3] which uses a regular lattice, i.e., an ordered network possessing many loops, as starting point, we shall also use the concept of *local ordering* when referring to the loop structure of a complex network. More information can be extracted if the triangles are sorted according to the degrees of their corners, finding that mainly the low-degree nodes account for the high level of clustering [18,19]. This suggests the presence of interesting modular organizations, and similar results were obtained from an analysis of loops of length 4 [17]. Local ordering plays a crucial role in the function of a metabolic network (with scale-free topology). Indeed, the loop structure is much richer than in the scale-free model based on growth and preferential attachment [20]. This work also investigates the number of triangles as a function of the system size for this model, a result which was generalized to loops up to length 5, yielding robust scaling relations [21].

Obviously the presence of loops has an effect on the spreading behavior since, with respect to a treelike topology, there exist many more paths along which the virus can propagate. Different strategies in order to gain insights about the role of local ordering properties have been proposed. An interpretation of how clustering influences the stationary spreading behavior was obtained by mapping the epidemic process onto bond percolation [22]. Another approach is to abandon the mean-field level and take into account spatial correlations which govern the epidemic dynamics. Matsuda *et al.* first used the ordinary pair approximation in order to study a population dynamical problem [23]. This approximation, as its name anticipates, accounts for pair correlations and lies at the basis of improved pair models [24,25] which uncover the role of local ordering in a rather indirect way: clustering enters by making a number of assumptions about the open (\sphericalangle) and closed (\triangle) triple correlations. In cluster approximations, a time-dependent probability is assigned to each configuration of the fundamental cluster whose choice is guided by the network topology [26]: for investigating the spreading dynamics on a triangular lattice one uses a triangle as fundamental cluster whereas the star is the appropriate choice for a random network. Higher-order correlations are therefore embedded very explicitly. Moreover the systematic improvability of this method makes it a powerful tool to study probabilistic systems.

In order to understand the role of local ordering properties, the exploration of temporal correlations seems to be an even more natural approach: for example, within a two-step description, it matters whether the local topology is treelike or if loops of short length are present. The method is illustrated for the susceptible-infected-susceptible model (see, for example, Ref. [27]), homogeneous networks are used as starting point, and analytical estimates are obtained also for disordered graphs obeying $P(k) = \delta_{k,K}$, K being arbitrary.

The paper is organized as follows. Section II describes the adopted model consisting of the contact network as well as the local dynamics. In Sec. III, we introduce the formalism, from which the two-step description is derived. For completeness, the necessary ingredients in order to arrive at the one-step site approximation, i.e., the common mean-field level, are also shown. Section IV explores the implications of

the double-step approach for networks of degree 4, the generalization to arbitrary degree is done in the following section. The major conclusions are drawn in Sec. VI.

II. THE MODEL INGREDIENTS

The dynamical laws that describe the spreading of an infectious disease are determined by the contact structure which underlies the population. We therefore model the epidemic as a dynamical process on top of a given network that does not change in time. The nodes of the network represent individuals, and the links correspond to relationships between individuals along which an infective agent can propagate.

Since the aim of this paper is the investigation of the role of loops of short length, we adopt a rather simple epidemiological model where the individuals can be only in two possible states, namely, infected (I) or susceptible (S). Because the nodes repeatedly run through the cycle susceptible \rightarrow infected \rightarrow susceptible, it is called SIS model. In the physics community, it has recently been formulated as follows [13]: A node susceptible to the disease gets infected with probability $\nu\Delta t$ if it is connected to at least one infected nearest neighbor. On the other hand, infected nodes recover spontaneously with probability $\delta\Delta t$. This version of the SIS model is formally advantageous with respect to its conventional formulation, where infected nodes can infect neighboring susceptible vertices with probability $\nu\Delta t$ [27]. In the latter case, susceptible nodes become infected with probability $1 - (1 - \nu\Delta t)^{k_{\text{inf}}}$, k_{inf} being the number of infected nearest neighbors. In this paper, we will use the former version. By rescaling the time unit, we can reduce the number of parameters to one: the time evolution is determined by the effective spreading rate $\lambda \equiv \nu/\delta$, and the recovery rate is set to 1. The quantitative details of the behavior of the system still depend on the choice of Δt . In particular, the effect of the loops is of higher order in Δt , such that their influence is not seen in the continuous-time limit ($\Delta t \rightarrow 0$). As long as $\Delta t > 0$, we set this quantity to 1 without lack of generality.

The other model constituent concerns the underlying network. We shall not attempt to examine the combined effect of the degree distribution, degree correlations, and the loop structure. While the role of these degree related connectivity patterns in epidemic spreading is rather well understood, little attention has been given to the effect of local ordering. That is why we focus on this topological property. We therefore use networks in which every node has the same number of nearest neighbors (fixed degree). The adopted strategy is to start with strictly homogeneous graphs (networks in which identical connectivity patterns are "seen" from every node) differing in the detailed loop structure (Fig. 1). This leads us to an understanding of the role of these distinct local ordering properties, even for networks that are no longer strictly homogeneous but which still obey $P(k) = \delta_{k,K}$, K being the constant degree.

In the case $K=4$, examples of purely homogeneous networks are the random homogeneous network [Fig. 1(a)] constructed according to the Molloy-Reed algorithm [28,29], the square lattice [Fig. 1(b)], the Kagomé lattice [usually used in

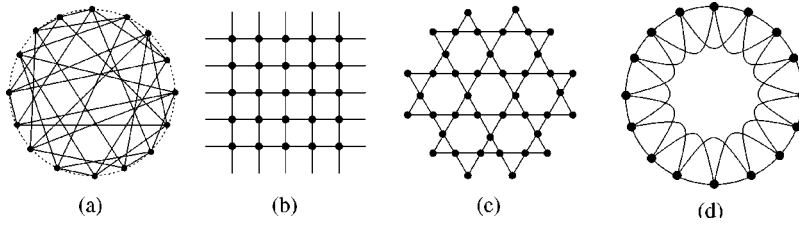


FIG. 1. Homogeneous networks of degree 4. In (a) the nodes are connected at random under the restriction that $K=4$ links emanate from every vertex, leading to a treelike topology. The role of triangles and loops of length 4 is studied by means of the square lattice (b) and the Kagomé lattice (c) where they appear separately, as well as the ring-type network (d).

condensed matter physics as this geometry represents one of the most frustrated antiferromagnetic systems, Fig. 1(c)] and a ring where nodes two units apart from each other are also directly connected [Fig. 1(d)]. For the random network, the homogeneity definition given above only holds approximately. While loops of short length do not occur in the limit of large network size, long ones of various lengths may exist. Therefore different nodes may not “see” identical connectivity patterns. Also the Cayley tree lacks completely in loops. But as the last generation of nodes has degree 1, it does not fall into the class of homogeneous networks. Indeed, since the number of nodes belonging to generation N is $K(K-1)^{N-1}$, most of the nodes are even comprised within the last generation. Clearly, all the networks depicted in Fig. 1 are characterized by the degree distribution $P(k)=\delta_{k4}$, but they differ in the way the second neighbors are arranged, that is, the loop structure becomes richer going from (a) to (d). While the random network has no short loops at all, the ones of length 4 are a fingerprint of the square lattice. In the Kagomé lattice, every node is a corner of two distinct triangles. Finally, in the ring of degree 4, triangles and two different kinds of loops of length 4 are found, namely, plaquettes (as in the square lattice) and the so-called primary quadrilaterals (two adjacent triangles), see Sec. IV B.

In order to study the epidemic process on disordered “homogeneous” topologies, we subject the networks in Figs. 1(b)–1(d) to the following rearrangement algorithm (Fig. 2) [30].

(1) Choose randomly two links (link 1 connecting node A_1 with B_1 and connection 2 linking vertex A_2 with B_2) that do not share a common node.

(2) Remove these two links and establish two new connections between A_1 and B_2 as well as A_2 and B_1 .

Repeating this rewiring procedure a certain number of times leads to locally varying numbers of loops while the degree distribution remains unaltered.

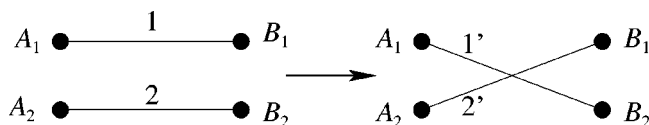


FIG. 2. Rewiring procedure not affecting the degree distribution: the end vertices of two arbitrarily chosen links are exchanged.

III. THE FORMALISM

In this section, we shall introduce the formalism that is used in order to describe the dynamics on top of the networks specified in the preceding section. For completeness, we then derive the (single step) mean-field approximation which could also be obtained heuristically. In the last paragraph, we derive a two-step description that allows us to gain insight about the loop structure in epidemic spreading.

On an exact level, we shall describe the epidemic dynamics by assigning a probability $\mathcal{P}_t(\mathbf{x})$ to each configuration \mathbf{x} at every instant of time t . The vector \mathbf{x} contains the states x_i of all the nodes i of the network, x_i being either 0 (susceptible) or 1 (infected). The system probabilities satisfy at any time t ,

$$\sum_{\mathbf{x}} \mathcal{P}_t(\mathbf{x}) = 1,$$

and evolve in time according to

$$\mathcal{P}_{t+1}(\mathbf{x}) = \sum_{\mathbf{y}} \mathcal{W}_{\mathbf{y} \rightarrow \mathbf{x}} \mathcal{P}_t(\mathbf{y}). \quad (1)$$

The transition matrix of the system $\mathcal{W}_{\mathbf{y} \rightarrow \mathbf{x}}$ is obtained from the matrix $W_{y_l \rightarrow x_l}^l$ which shall denote the probability that the state of the arbitrary site l changes from y_l to x_l , through

$$\mathcal{W}_{\mathbf{y} \rightarrow \mathbf{x}} = \prod_{l=1}^N W_{y_l \rightarrow x_l}^l,$$

N being the total number of nodes in the network. The matrix elements representing the probabilities for the possible events at the site l are given by our version of the SIS model, namely,

$$W_{1 \rightarrow 0}^l = 1, \quad W_{0 \rightarrow 1}^l = \lambda \left[1 - \prod_{j \neq l} (1 - y_j) \right],$$

$$W_{1 \rightarrow 1}^l = 0, \quad W_{0 \rightarrow 0}^l = 1 - \lambda \left[1 - \prod_{j \neq l} (1 - y_j) \right],$$

or in a more compact form

$$W_{y_l \rightarrow x_l}^l = 1 - x_l + \lambda(2x_l - 1)(1 - y_l) \left[1 - \prod_{j \text{nn}l} (1 - y_j) \right].$$

The products in the above expressions have to be taken over all the nearest neighbors j of node l . The factor $1 - \prod_{j \text{nn}l} (1 - y_j)$ is 1 if at least one $y_j = 1$ and 0 otherwise.

Before deriving our two-time-step description, we show how the conventional mean-field approximation is retrieved through this formalism. At that level, the sites are considered independently from each other, and we write for the system probability

$$\mathcal{P}_t(\mathbf{x}) = \prod_{l=1}^N P_t(x_l), \quad (2)$$

i.e., the system is described by the single variable $P_t(1)$ [the probability of being susceptible is $P_t(0) = 1 - P_t(1)$]. Its dynamics is obtained from Eq. (1) by summing it over all possible configurations \mathbf{x} , x_0 held fixed

$$\sum_{\{x_j\}_{j \neq 0}} \mathcal{P}_{t+1}(\mathbf{x}) = \sum_{\mathbf{y}} \mathcal{P}_t(\mathbf{y}) \underbrace{\sum_{\{x_j\}_{j \neq 0}} W_{\mathbf{y} \rightarrow \mathbf{x}}}_{W_{y_0 \rightarrow x_0}^0},$$

where the node 0 can be chosen in an arbitrary way. The left-hand side of the above equation corresponds to the probability that node 0 is in state x_0 at time $t+1$. With the ansatz (2), the time evolution is

$$P_{t+1}(1) = \lambda P_t(0) [1 - P_t(0)^K].$$

With $P_t(1) \equiv P_t$ and $P_t(0) = 1 - P_t$, by consequence, we obtain for small values of P_t ,

$$P_{t+1} = \lambda K (1 - P_t) P_t. \quad (3)$$

The stationary-state condition ($P_{t+1} = P_t$) leads to a value for the epidemic threshold

$$\lambda_c = \frac{1}{K}. \quad (4)$$

Therefore all the networks depicted in Fig. 1 are treated identically at this level of description. The only topological property that determines the critical value of the effective spread-

ing rate λ is the number of nearest neighbors K .

Improvements upon the mean-field description can be obtained by taking into account spatial correlations. Based on the ordinary pair approximation [23], the presence of triangles enters by establishing a number of hypotheses about the open and closed triple correlations [24,25]. This approach therefore embeds this topological property in a rather implicit way. In cluster approximations [26], the local topology and its associated spatial correlations translate directly into the choice of the cluster and a set of probabilities for its possible configurations.

Another strategy that serves to incorporate local ordering properties is to take into account temporal correlations. Thus by performing two time steps exactly, we expect that the way the second neighbors are arranged, enters very naturally into the description. For example, the cases where two nearest neighbors of an arbitrary node are also directly connected (presence of a triangle), where they are linked via a second neighbor (giving rise to a loop of length 4) or where the only path goes through the original node (treelike structure), lead to different results. In the remaining part of this section, we derive the general equation, special cases are then looked at within the following section.

As outlined above, we iterate Eq. (1) once

$$\mathcal{P}_{t+1}(\mathbf{x}) = \sum_{\mathbf{y}} \left[W_{\mathbf{y} \rightarrow \mathbf{x}} \sum_{\mathbf{z}} W_{\mathbf{z} \rightarrow \mathbf{y}} \mathcal{P}_{t-1}(\mathbf{z}) \right].$$

We now again pass to a site approximation. By summing the above equation over all possible configurations \mathbf{x} , x_0 held fixed, we get

$$P_{t+1}(x_0) = \sum_{\mathbf{z}} \left[\mathcal{P}_{t-1}(\mathbf{z}) \sum_{\{y_l=0,1,\dots,K\}} \left(W_{y_0 \rightarrow x_0} \prod_{j=0}^K W_{z_j \rightarrow y_j} \right) \right], \quad (5)$$

0 again being an arbitrarily chosen node. Furthermore, the system probability $\mathcal{P}_{t-1}(\mathbf{z})$ is given by Eq. (2), and the nodes $1, 2, \dots, K$ denote the nearest neighbors of the arbitrarily chosen node 0. As only z states associated with the vertex 0, its nearest and second neighbors appear in the W factors, the sum over the z variables associated with nodes more than two links away from vertex 0, is carried out trivially. A tour de force calculation leads to

$$\begin{aligned} P_{t+1}(1) = & \lambda \left[\lambda \sum_{\alpha_1=1}^K \langle f_{\alpha_1} \rangle_{t-1} - \lambda^2 \left(\sum_{\alpha_1=1}^K \sum_{\alpha_2=\alpha_1+1}^K \langle f_{\alpha_1} f_{\alpha_2} \rangle_{t-1} + \sum_{\alpha_1=1}^K \langle f_0 f_{\alpha_1} \rangle_{t-1} \right) \right. \\ & + \lambda^3 \left(\sum_{\alpha_1=1}^K \sum_{\alpha_2=\alpha_1+1}^K \sum_{\alpha_3=\alpha_2+1}^K \langle f_{\alpha_1} f_{\alpha_2} f_{\alpha_3} \rangle_{t-1} + \sum_{\alpha_1=1}^K \sum_{\alpha_2=\alpha_1+1}^K \langle f_0 f_{\alpha_1} f_{\alpha_2} \rangle_{t-1} \right) \cdots \\ & \left. - (-\lambda)^K \left(\sum_{\alpha_1=1}^K \sum_{\alpha_2=\alpha_1+1}^K \cdots \sum_{\alpha_K=\alpha_{K-1}+1}^K \langle f_{\alpha_1} f_{\alpha_2} \cdots f_{\alpha_K} \rangle_{t-1} + \sum_{\alpha_1=1}^K \sum_{\alpha_2=\alpha_1+1}^K \cdots \sum_{\alpha_{K-1}=\alpha_{K-2}+1}^K \langle f_0 f_{\alpha_1} f_{\alpha_2} \cdots f_{\alpha_{K-1}} \rangle_{t-1} \right) \right]. \quad (6) \end{aligned}$$

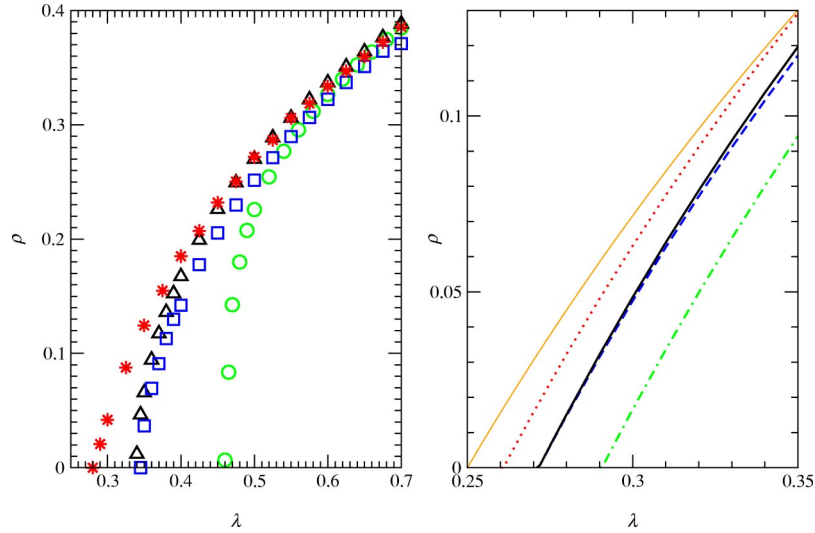


FIG. 3. Equilibrium prevalences for the epidemic process on top of different networks characterized by $P(k)=\delta_{k4}$. Left: Simulation results for the homogeneous random graph (*), the Kagomé lattice (Δ), the square lattice (\square), and the ring-type network (\circ). Right: The one-step site approximation (thin solid) ignores (local) ordering properties and yields $\lambda_c=0.25$ for any homogeneous network of degree 4. For the Molloy-Reed network (dotted), the Kagomé lattice (solid), the square lattice (dashed), and the ring (dotted dashed), different steady-state prevalences are obtained at the two-step level.

Thereby the connectivity embedding factor

$$f_\alpha \equiv (1 - z_\alpha) \left[1 - \prod_{\sigma \text{nn}\alpha} (1 - z_\sigma) \right]$$

$$= \begin{cases} 1 & \text{if } z_\alpha = 0 \text{ and at least one } z_\sigma = 1 \\ 0 & \text{otherwise} \end{cases}$$

and the expectation value of a function of the states of node 0, its nearest and second neighbors, these vertices collectively being denoted by \mathcal{N}^2 ,

$$\langle g(\{z_k\}_{k \in \mathcal{N}^2}) \rangle_t \equiv \sum_{\{z_k\}_{k \in \mathcal{N}^2}} \left[\prod_{l \in \mathcal{N}^2} P_l(z_l) g(\{z_m\}_{m \in \mathcal{N}^2}) \right] \quad (7)$$

were introduced for notational convenience. In the following, an expectation value of a product of n f factors will be referred to as a term of n th order although it is proportional to λ^n . As is illustrated in detail in the following section, every term of Eq. (6) corresponds to a subgraph of the graph composed of the nodes \mathcal{N}^2 and whose links are according to the network under investigation. It can already be anticipated that the first term accounts for the degree distribution only, whereas the contributions of higher order will give insight into the role of the loop structure.

IV. NETWORKS OF DEGREE 4

In this section, we elaborate the implications of our two-step description, the analysis being restricted to the stationary state.

In the left part of Fig. 3, we show the simulation results for the graphs introduced in Sec. II. The ring-type network exhibits the largest epidemic threshold. For both the square and Kagomé lattices, the critical value is $\lambda_c \approx 0.34$. We there-

fore anticipate that either four plaquettes or two triangles (per node) lead to the same effect in the regime of low prevalences. Finally, the lowest epidemic threshold is found if the population is arranged on a homogeneous random network (of degree 4). The last result is very intuitive since in such a graph, global distances are small, making it more easy for a virus to spread. Therefore, even if the effective spreading rate is rather low, a finite fraction of the population will be infected in the stationary state, hence the small value for the location of the onset of the epidemic. In summary, these results indicate that the poorer the loop structure, the lower the corresponding epidemic threshold.

In the right part of Fig. 3, the one-step and two-step site approximations are reported. The former corresponds to the steady-state solution of Eq. (3) for which $\rho=0$ at $\lambda_c=1/4$ according to Eq. (4). All the networks in question are therefore treated identically, the loop structure being ignored at this level of description. Yet, the two-step solutions [Eq. (5)] are diverse for the different graphs. Going from right to left, the curves correspond to the ring, the square lattice, the Kagomé lattice, and the Molloy-Reed network, that is, they appear in the same sequence as at the level of simulation. Furthermore, the curves corresponding to the Kagomé and square lattice also meet the x axis at the same value of λ . These findings confirm our intuitive arguments given in the preceding section. It has to be noted that the two-step estimates for the threshold values are still considerably inaccurate especially for the ring and lattices, but this just highlights the presence of higher-order spatiotemporal correlations. However, the important point is that the degeneracy associated with the one-step description disappears at the two-step level.

On the basis of Eq. (6), we shall now analytically study the effect of local ordering properties upon the epidemic

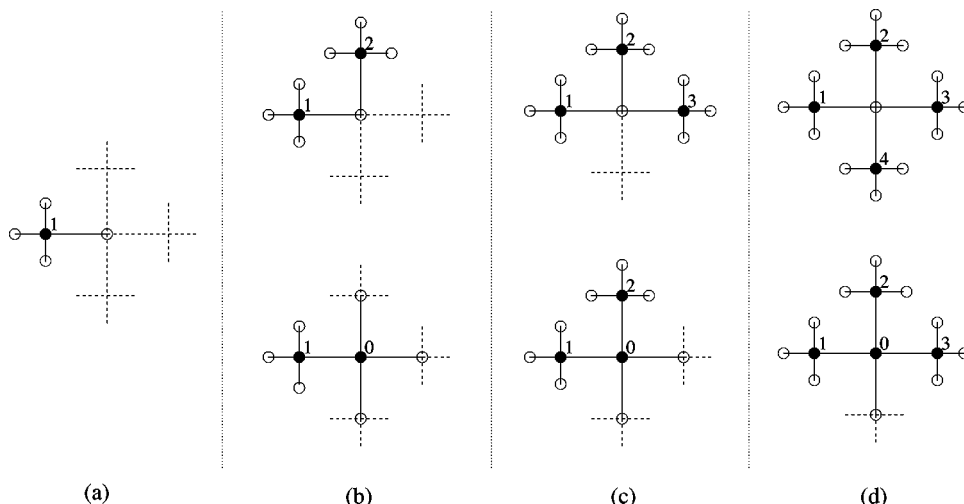


FIG. 4. Pictorial representation of the terms of Eq. (6) for a treelike topology. The order of a specific subgraph is given by the number of filled circles, (a) corresponding to $\langle f_1 \rangle$, (b) to $\langle f_1 f_2 \rangle$ and $\langle f_0 f_1 \rangle$, and so forth. In this vein, the indices of the f factors appearing in the expectation values correspond to filled circles whereas empty circles represent their nearest neighbors. The dashed lines are the links of the complete graph which are not contained in a specific subgraph.

spreading, leading to a quantitative understanding of the threshold value.

A. Random network

We shall now evaluate all the terms of Eq. (6) for a locally treelike topology. Figure 4 shows the subgraphs representing the terms in Eq. (6), in increasing order. Thereby the correspondence is as follows: Given the term $\langle f_\alpha f_\beta \rangle$, the nodes α and β are represented by filled circles whereas their nearest neighbors are drawn by empty circles. The links which enter at the level of the subgraph in question, are represented by solid lines whereas the ignored ones are dashed. If we denote the second neighbors of the central vertex 0 by l_1, l_2, l_3 for $l=1, 2, 3, 4$ and follow Eq. (7), the first-order contribution for $\alpha_1=1$ [subgraph in Fig. 4(a)] is

$$\begin{aligned} \langle f_1 \rangle &= \sum_{z_1} \sum_{z_0} \sum_{z_{11} \dots z_{13}} \{P(z_1)P(z_0)P(z_{11})P(z_{12})P(z_{13}) \\ &\quad \times (1-z_1)[1-(1-z_0)(1-z_{11})(1-z_{12})(1-z_{13})]\} \\ &= \underbrace{\dots}_{f_1} \\ &= 4P + O(P^2) \end{aligned}$$

where the sum over the z variables to which no circles are associated, has been carried out trivially. Furthermore, we again have set $P \equiv P(1)$ in the third line (this is also done below), and the time index was omitted since we are only interested in the steady state. This term appears with multiplicity 4 (due to $\sum_{\alpha_1=1}^K$), giving the contribution $16P$ to first order in P .

Figure 4(b) shows the subgraphs representing $\langle f_1 f_2 \rangle$ (upper part) and $\langle f_0 f_1 \rangle$ (lower part). Their contributions are

$$\begin{aligned} \langle f_1 f_2 \rangle &= \sum_{z_0} \sum_{z_1} \sum_{z_2} \sum_{z_{11} \dots z_{13}} \sum_{z_{21} \dots z_{23}} \{P(z_0)P(z_1)P(z_2) \\ &\quad \times P(z_{11})P(z_{12})P(z_{13})P(z_{21})P(z_{22})P(z_{23})f_1 f_2\} \\ &= P + O(P^2), \end{aligned}$$

occurring $\binom{4}{2}=6$ times and

$$\begin{aligned} \langle f_0 f_1 \rangle &= \sum_{z_0} \sum_{z_1} \sum_{z_2 \dots z_4} \sum_{z_{11} \dots z_{13}} \{P(z_0)P(z_1)P(z_2)P(z_3)P(z_4) \\ &\quad \times P(z_{11})P(z_{12})P(z_{13})f_0 f_1\} = 9P^2 + O(P^3), \end{aligned}$$

thus not giving a contribution to first order in P . As a consequence, the second-order term (that is the one proportional to $\lambda \lambda^2$) is $6P$.

As the procedure should now be clear, we only give the results for the remaining orders. The upper subgraph of Fig. 4(c) represents the term $\langle f_1 f_2 f_3 \rangle$. Its contribution is $P + O(P^2)$. The lower subgraph corresponds to $\langle f_0 f_1 f_2 \rangle$, yielding $18P^3 + O(P^4)$. As the former term has multiplicity $\binom{4}{3}=4$, the total third-order contribution is $4P$.

As far as the fourth order is concerned [Fig. 4(d)], the subgraph involving node 0 as filled circle neither gives a

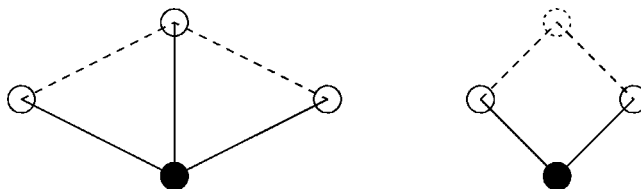


FIG. 5. Loops of length 4 in a network. Primary quadrilaterals (left) are two adjacent triangles and therefore involve only nearest neighbors (solid empty circles) of the central node (filled circle). A secondary quadrilateral (right) involves a second neighbor (dotted empty circle) as well.

TABLE I. Loop properties for the simple non-tree-like networks described in Sec. II.

Network	E	Q_1	Q_2
Square lattice	0	0	4
Kagomé lattice	2	0	0
Ring	3	2	2

contribution whereas the term $\langle f_1 f_2 f_3 f_4 \rangle$ having multiplicity 1 also gives $P + O(P^2)$, thus totally yielding $\lambda \lambda^4 P$.

Collecting these findings, we obtain the following condition that determines the epidemic threshold for a treelike topology

$$1 = \lambda(16\lambda - 6\lambda^2 + 4\lambda^3 - \lambda^4), \quad (8)$$

which is satisfied by $\lambda_c \approx 0.2609$. This is the value that can be extracted from the right panel of Fig. 3 (second curve from the left).

B. Graphs with loops

It is easy to imagine that the preceding analysis yields different results when triangles and loops of length 4 are present. Caldarelli *et al.* [17] have classified loops of length 4 in a complex network into *primary* and *secondary quadrilaterals* (Fig. 5). In the former case, the external vertices which the loop is composed of are all nearest neighbors whereas secondary quadrilaterals are plaquettes, the external nodes being two nearest and a second neighbor. With these concepts, the loop structure of a strictly homogeneous graph can quantitatively be characterized as follows: By choosing an arbitrary node, the number of edges between its nearest neighbors is denoted by E . Q_1 and Q_2 shall refer to the number of primary and secondary quadrilaterals. For the networks depicted in Figs. 1(b)–1(d), we report the corresponding values in Table I.

Let us now look at the subgraph development for the square lattice whereby we focus on the important changes with respect to the treelike case. The reader interested in the full elaborations is referred to the Appendix. We have already noticed that the first-order term is fully determined by

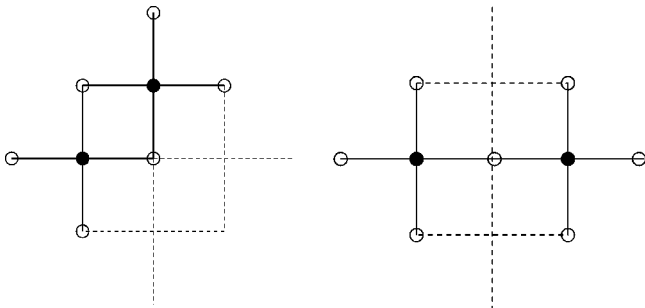


FIG. 6. Second-order subgraphs not involving the central node as filled circle for the square lattice. To first order in P , the left subgraph yields the contribution $2P$, the right one P . For further explanations, see Fig. 4.

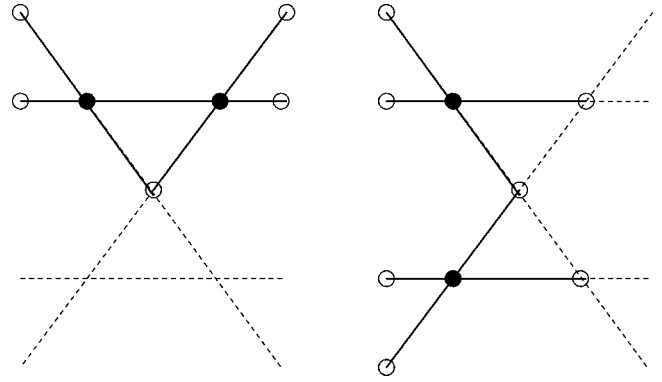


FIG. 7. Subgraphs of second order for the Kagomé lattice, both contribute with $P + O(P^2)$. See Fig. 4 for details regarding line and circle styles.

the degree distribution, therefore the $\lambda \lambda$ coefficient is 16, as in the treelike case. At order 2, the term $\langle f_1 f_2 \rangle$ [upper subgraph of Fig. 4(b)] splits into two subgraphs in the presence of plaquettes (Fig. 6). The right subgraph is the same as in the treelike case, yet the left yields a contribution $2P + O(P^2)$. Their multiplicities are 4 (left) and 2 (right) summing up to $\binom{4}{2} = 6$. The resulting $\lambda \lambda^2$ coefficient is therefore -10 . Although different subgraphs enter into the development also at the orders ≥ 3 , the coefficients appearing in the equation determining the epidemic threshold do not change.

The second-order subgraphs for the Kagomé lattice are depicted in Fig. 7. Both's contributions are $P + O(P^2)$. The one involving a triangle appears six times whereas the right subgraph has multiplicity 4. We therefore obtain the same second-order coefficient as for the square lattice. An analysis for the higher-order subgraphs yields no difference with respect to the square lattice. These two cases are therefore equivalent at the two-step level for $P \ll 1$.

In Table II, we summarize the coefficients for these two lattices as well as the ring [Fig. 1(d)]. The full developments are given in the Appendix.

Of course the idea is now to extend Eq. (8) such that it holds for all the investigated graphs. Our findings suggest that the local ordering properties enter in the following way into the equation determining the epidemic threshold

$$1 = \lambda[16\lambda - (6 + 2E + Q_1 + Q_2)\lambda^2 + (4 + Q_1)\lambda^3 - \lambda^4]. \quad (9)$$

However this is not the full story. What about loops of length 5? Let us argue why they do not enter in the frame-

TABLE II. Coefficients of the two-step threshold equation for our networks having in common $P(k) = \delta_{k4}$, but differing in the loop pattern.

$\lambda \lambda^n$ coeff.	$n=1$	$n=2$	$n=3$	$n=4$
Square lattice	16	-10	4	-1
Kagomé lattice	16	-10	4	-1
Ring	16	-16	6	-1

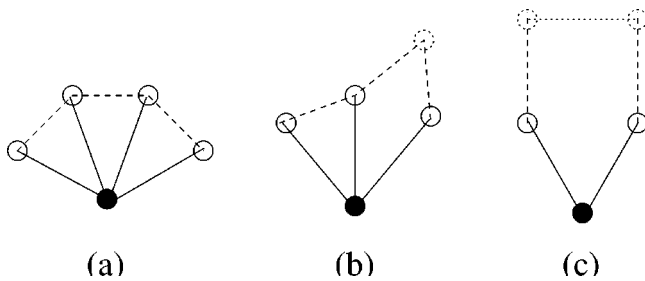


FIG. 8. Loops of length 5 involving different hierarchies of nearest neighbors. Connections emanating from the central node (filled circle) are drawn as a solid line, links going from nearest neighbors (empty solid circle) to other nearest neighbors or second neighbors (empty dotted circle) are dashed, and second neighbors are connected by a dotted line. The pentalateral (a) involves nearest neighbors only, in (b) the loop traverses a second neighbor and in the one in (c) lacking in internal connections, a link between two second neighbors serves to close it.

work of a two-step description. Although there exist such loops involving only first and second neighbors [Figs. 8(a) and 8(b)], it may also be closed only between two second neighbors [Fig. 8(c)]. Obviously such a connection is ignored at the two-step level. In the language of graph theory [31], the latter case corresponds to a *fundamental* loop whereas the former examples can be reduced to loops of length 3 and 4. But whatever the number of hierarchies of nearest neighbors involved in the formation of the loop is, the point is the following. If the central node is infected at time t , it can causally affect only vertices two links away, corresponding to a chain of four links. Obviously, it matters whether the first and the last node of this chain are identical. In this case, we have a loop of length 4. Otherwise it cannot be distinguished whether the topology is fully treelike or if loops of length greater than 4 are present. Along these lines, it has to be expected that loops up to length $2n$ enter within an n -time-step description. In contrast, the presence of higher-order quadrilaterals modifies the coefficients of Eq. (9). Figure 9(a) shows what we shall call a *tertiary* quadrilateral: the three nearest neighbors of the central node are all connected to another common node. Obviously, the presence of a tertiary quadrilateral implies $Q_2 = \binom{3}{2} = 3$ secondary quadrilaterals. In a *fourth-order* quadrilateral [Fig. 9(b)], four

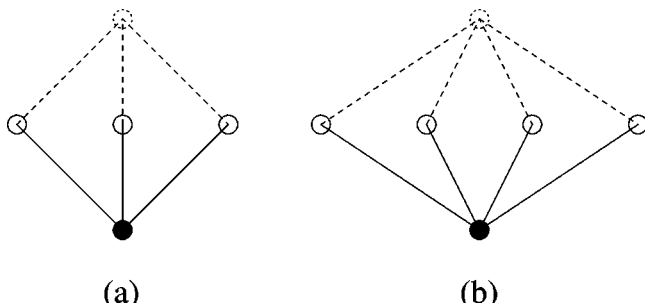


FIG. 9. Quadrilaterals of orders 3 and 4. The n (a, $n=3$; b, $n=4$) nearest neighbors (empty solid circles) of the central node (filled circle) share another vertex (dotted empty circle) as nearest neighbor.

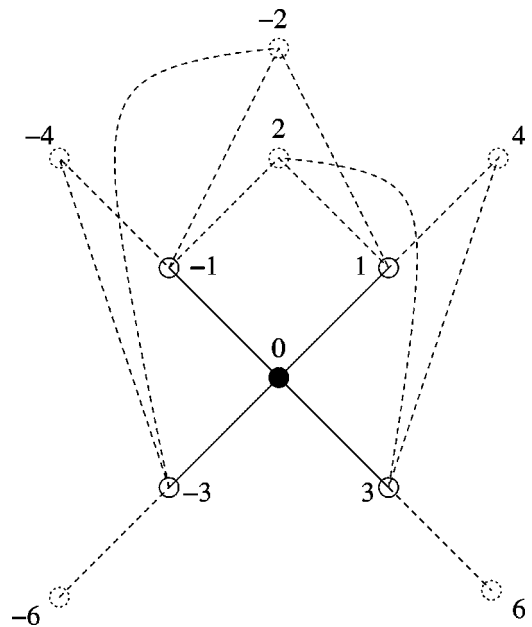


FIG. 10. This figure visualizes how the nearest (empty solid circles) and second neighbors (empty dotted circles, vertices $\{-3, -1, 1, 3\}$) of an arbitrarily chosen node (filled circle, node 0) are arranged in a one-dimensional lattice with additional connections between sites 3 units apart. The sets $\{0, \{-1, 1, 3\}, 2\}$ as well as $\{0, \{-3, -1, 1\}, -2\}$ are forming tertiary quadrilaterals.

nodes share two common vertices as nearest neighbors, implying the presence of $Q_3 = \binom{4}{3} = 4$ tertiary quadrilaterals and $Q_2 = \binom{4}{2} = 6$ secondary quadrilaterals. In a network of degree $K > 4$, quadrilaterals up to order K can, in principle, be found.

A one-dimensional lattice with additional connections between the nodes i and $i+3$ for all i [instead of $i+2$ as in the ring investigated up to now, Fig. 1(d)] possesses the neighborhood structure shown in Fig. 10, i.e., it is characterized by $E=Q_1=0$, $Q_2=8$, $Q_3=2$, and $Q_4=0$. By applying our formalism to this case and to a network that has fourth-order quadrilaterals, Eq. (9) generalizes to

$$1 = \lambda [16\lambda - (6 + 2E + Q_1 + Q_2)\lambda^2 + (4 + Q_1 + Q_3)\lambda^3 - (1 + Q_4)\lambda^4], \quad (10)$$

the coefficients of order 3 and 4 being modified only.

C. Introducing disorder

The networks considered up to now lack in the small world phenomenon, a property characterizing social networks on which the epidemic process is occurring. By starting with a ringlike network [Fig. 1(d)] and repeating the rewiring procedure described in Sec. II a certain number of times, we obtain graphs of fixed degree $K=4$ that are simultaneously highly clustered, and in which the average distance between pairs of nodes is small [3].

The left part of Fig. 11 reports the simulation result for the equilibrium prevalence of the epidemic process on the disordered ring. Systems of size $N=10^4$ were used, and the rewiring

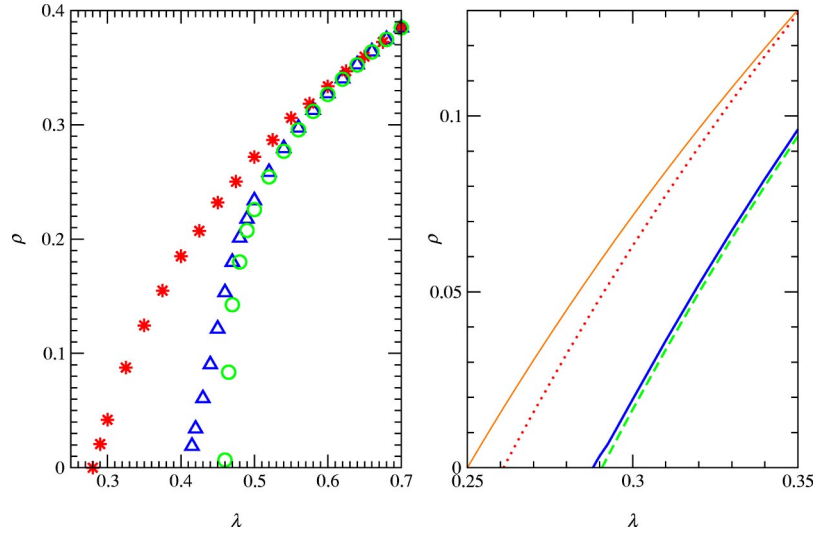


FIG. 11. Steady-state prevalence for the ring-type network having undergone different degrees of randomization. Left: Simulation result for the original network (\circ), its partially rewired version (\triangle), and the entirely random network ($*$). Right: The single-step site approximation (thin solid) treats all the networks in question identically whereas at the two-step level, the Molloy-Reed network (dotted), the partially randomized ring (solid), and the fully ordered ring (dashed) appear in the same sequence as in the left panel.

ing procedure was repeated $n=100$ times. For completeness, the two limiting cases (fully ordered ring and random network) are also depicted. This panel shows the considerable effect on the steady state spreading behavior of the rather small number of rewirings.

The right panel of this figure depicts the one-step site approximation (predicting $\lambda_c=0.25$ for all cases) and the numerical solutions of the two-step description (5). It has to be noted that for the partially rewired ring lacking in strict homogeneity, Eq. (5) was solved at every node, therefore involving the set $P_i(x_i)$, $i=1, 2, \dots, N$, the resulting prevalence being given by $1/N \sum_{i=1}^N P_i(1)$. Again at the double-step level, the networks in question are treated differently, as it could already be observed in Fig. 3. Here the curve corresponding to the partially randomized ring lies closer to the original network in proportion to the simulation result. This is due to the small world property which can have a considerable effect on the location of the onset of the epidemic. Obviously, the simulation result uncovers the real effect of this global topological property whereas at the two-step level, it is the slightly poorer loop structure that accounts for the corresponding shift in the epidemic threshold.

Of course the quantities E , Q_1 , Q_2 , Q_3 , and Q_4 are no longer reasonable for a partially randomized network due to the lack of strict homogeneity, but rather its local ordering properties can be quantified by averaging these values over the entire network. The emerging topological parameters \bar{E} and \bar{Q}_i ($i=1, 2, 3, 4$) are essentially the *clustering coefficients* [up to the factor $K(K-1)/2=6$] and *grid coefficients*, i.e., the densities of triangles and loops of length 4 [3,17]. We may therefore replace E and the number of quadrilaterals (of the different orders) in Eq. (10) by its mean values, yielding the following estimate for the epidemic threshold condition

$$1 = \lambda [16\lambda - (6 + 2\bar{E} + \bar{Q}_1 + \bar{Q}_2)\lambda^2 + (4 + \bar{Q}_1 + \bar{Q}_3)\lambda^3 - (1 + \bar{Q}_4)\lambda^4]. \quad (11)$$

For our partially randomized ring, we have $\bar{E}=2.883$, $\bar{Q}_1=1.886$, $\bar{Q}_2=1.958$, and $\bar{Q}_3=\bar{Q}_4=0$, leading to $\lambda_c \approx 0.2892$. This value corresponds approximately to where the corresponding curve in the right part of Fig. 11 meets the x axis.

V. ARBITRARY DEGREE

The implications of our two-step description have been illustrated for homogeneous networks of degree 4 in the preceding section. This was a convenient choice as there exists a number of familiar simple graphs obeying $P(k)=\delta_{k4}$, differently ordered. Of course our formalism enables us to generalize the obtained threshold condition (10) to an arbitrary degree K , which is the subject of this section.

Let us again look at a fully treelike network, using Eqs. (6) and (8) as guidelines. The λ^2 coefficient 16 incorporating the degree distribution is simply 4×4 since $\langle f_\alpha \rangle = 4P + O(P^2)$ and α runs from 1 to 4. The remaining coefficients -6 , 4 , and -1 correspond to the binomial coefficients $-\binom{4}{2}$, $\binom{4}{3}$, and $-\binom{4}{4}$. Indeed the threshold equation for a treelike network of degree K derived by Eq. (6) is

$$1 = \lambda \left[\lambda K^2 - \sum_{\kappa=2}^K \lambda^\kappa \binom{K}{\kappa} \right]. \quad (12)$$

Repeating the graph developments for homogeneous networks characterized by different values of K and varying loop structures reveals that the very same correction terms enter into Eq. (12), yielding

TABLE III. Subgraph development for the square lattice. All the terms in Eq. (6) are symbolized by a specific subgraph, its order being given by the number of filled circles. The $\lambda\lambda^2$ coefficient -10 , as an example, is obtained by summing the various $O(P)$ contributions, that is, $2 \times 4 + 1 \times 2 + 0 \times 4 = 10$, and the negative sign comes from Eq. (6). The $\lambda\lambda^3$ and $\lambda\lambda^4$ coefficients are unaltered with respect to the treelike topology although other subgraphs enter into the development.

order (n)	1	2	3	4
subgraph				
contribution	$4P + \mathcal{O}(P^2)$	$2P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$
multiplicity	4	4	4	1
subgraph				
contribution		$P + \mathcal{O}(P^2)$	$2P^2 + \mathcal{O}(P^3)$	$5P^3 + \mathcal{O}(P^4)$
multiplicity		2	4	4
subgraph				
contribution		$9P^2 + \mathcal{O}(P^3)$	$18P^3 + \mathcal{O}(P^4)$	
multiplicity		4	2	
$\lambda \cdot \lambda^n$ -coeff.	16	-10	4	-1

TABLE IV. The subgraphs of all the orders for the Kagomé lattice. See Table III for how the coefficients are obtained and as far as further details are concerned.

order (n)	1	2	3	4
subgraph				
contribution	$4P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$
multiplicity	4	6	4	1
subgraph				
contribution		$P + \mathcal{O}(P^2)$	$5P^2 + \mathcal{O}(P^3)$	$8P^4 + \mathcal{O}(P^5)$
multiplicity		4	4	4
subgraph				
contribution			$8P^3 + \mathcal{O}(P^4)$	
multiplicity			2	
$\lambda \cdot \lambda^n$ -coeff.	16	-10	4	-1

TABLE V. The full subgraph development for the ring-type network. See Table III for the derivation of the $\lambda\lambda^n$ coefficients. With respect to the two lattices treated above, the $\lambda\lambda^2$ and $\lambda\lambda^3$ coefficients are -16 and 6 .

order (n)	1	2	3	4
subgraph				
contribution	$4P + \mathcal{O}(P^2)$	$2P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$
multiplicity	4	4	4	1
subgraph				
contribution		$2P + \mathcal{O}(P^2)$	$P + \mathcal{O}(P^2)$	$P^2 + \mathcal{O}(P^3)$
multiplicity		2	2	2
subgraph				
contribution		$P + \mathcal{O}(P^2)$	$3P^2 + \mathcal{O}(P^3)$	$2P^2 + \mathcal{O}(P^3)$
multiplicity		1	3	2
subgraph				
contribution		$P + \mathcal{O}(P^2)$	$5P^2 + \mathcal{O}(P^3)$	
multiplicity		3	1	
$\lambda \cdot \lambda^n$ -coeff.	16	-16	6	-1

$$\begin{aligned}
 1 = \lambda & \left\{ \Theta(K-1)K^2\lambda - \Theta(K-2) \left[\binom{K}{2} + 2E + Q_1 + Q_2 \right] \lambda^2 \right. \\
 & + \Theta(K-3) \left[\binom{K}{3} + Q_1 + Q_3 \right] \lambda^3 - \sum_{\kappa=4}^K \Theta(K-\kappa) \left[\binom{K}{\kappa} \right. \\
 & \left. \left. + Q_\kappa \right] (-\lambda)^\kappa \right\}, \quad (13)
 \end{aligned}$$

where E is again the number of connections between the nearest neighbors of an arbitrarily chosen node, Q_n denotes the number of quadrilaterals of order n , and $\Theta(x)$ is the step function defined by

$$\Theta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The equation that describes the role of local ordering in a disordered network subject to $P(k) = \delta_{k,K}$ is again obtained simply by replacing the corresponding quantities by its mean values ($E \rightarrow \bar{E}, Q_i \rightarrow \bar{Q}_i$) in Eq. (13), providing an improved estimate for the epidemic threshold.

VI. CONCLUSION

The spreading of an infectious disease was modeled as a dynamical process on top of a contact network. We used a discrete-time version of the simple SIS model, which is infected nodes recover with probability Δt , and susceptible nodes become infected with probability $\lambda \Delta t$ if they are connected to at least one infected nearest neighbor. As far as the connectivity patterns underlying the population are concerned, we chose homogeneous networks as starting point and introduced an arbitrary degree of disorder by an appropriate rewiring procedure not affecting the degree distribution $P(k) = \delta_{k,K}$.

Describing the epidemic dynamics of the entire population as a Markovian process, we derived a two-step descrip-

tion that takes into account temporal correlations. This approach revealed to be very prolific if one wants to unravel the role of loops of short length in the contact network regarding epidemic spreading. Indeed it leads to a subgraph development where the complete graph involves the connectivity patterns of two hierarchies of nearest neighbors (of an arbitrarily chosen node). Within this approach serving to track a probabilistic system, the local topology, be it treelike or be loops of length 3 or 4 present, therefore enters very naturally. The analytically obtained condition for the location of the onset of the epidemic then serves as a guiding equation elucidating the role of clustering and gridlike ordering in epidemic spreading.

In principle, it is possible to apply our two-step description to more complex networks where different degrees are present, uncovering the effect of the degree-dependent densities of triangles and loops quadrilaterals on the critical value. Likewise, loops of length up to $2n$ are expected to enter within an n -time-step description, also providing a natural classification of them. However the major insight gained by the strategy of exploring temporal correlations is best illustrated as it was done in this paper.

ACKNOWLEDGMENTS

We wish to thank EC-Fet Open Project No. COSIN IST-2001-33555, and the OFES-Bern (CH) for financial support.

APPENDIX: FULL SUBGRAPH DEVELOPMENTS

Here we show the complete subgraph developments for the square lattice (Table III), the Kagomé lattice (Table IV) and the ring-type network of Fig. 1(d) (Table V). Every subgraph corresponds to a term in Eq. (6), its contribution is obtained by the procedure illustrated in Sec. IV A. The $\lambda \lambda^n$ coefficient of the threshold equation is obtained by summing all the $O(P)$ contributions (taking into account the multiplicities) of the n th order subgraphs.

-
- [1] S. Milgram, *Psychol. Today* **1**, 60 (1967).
 - [2] P. Erdős and A. Rényi, *Publ. Math., Inst. Hautes Etud. Sci.* **6**, 290 (1959).
 - [3] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
 - [4] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).
 - [5] S. Yook, H. Jeong, and A.-L. Barabási, e-print cond-mat/0107417.
 - [6] R. Albert, H. Jeong, and A.-L. Barabási, *Nature (London)* **401**, 130 (1999).
 - [7] A. Wagner and D. Fell, Santa Fe Institute Report No. 00-07-041, 2000.
 - [8] J. Montoya and R. Solé, e-print cond-mat/0011195.
 - [9] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
 - [10] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002); S. N. Dorogovster and A. F. F. Mendes, *Adv. Phys.* **51**, 1079 (2002).
 - [11] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *Comput. Commun. Rev.* **29**, 251 (1999).
 - [12] F. Liljeros, C. R. Edling, L. A.N. Amaral, H. E. Stanley, and Y. Åberg, *Nature (London)* **411**, 907 (2001).
 - [13] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).
 - [14] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
 - [15] M. Boguñá and R. Pastor-Satorras, *Phys. Rev. E* **66**, 047104 (2002).
 - [16] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, *Phys. Rev. Lett.* **90**, 028701 (2003).
 - [17] G. Caldarelli, R. Pastor-Satorras, and A. Vespignani, *Eur. Phys. J. B* (to be published), e-print cond-mat/0212026.
 - [18] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, *Phys. Rev. E* **65**, 066130 (2002).
 - [19] E. Ravasz and A.-L. Barabási, *Phys. Rev. E* **67**, 026112 (2003).

- (2003).
- [20] P. M. Gleiss, P. F. Stadler, A. Wagner, and D. A. Fell, *Adv. Complex Syst.* **4**, 207 (2001).
- [21] G. Bianconi and A. Capocci, *Phys. Rev. Lett.* **90**, 078701 (2003).
- [22] M. E. J. Newman, *Phys. Rev. E* **68**, 026121 (2003).
- [23] H. Matsuda, N. Ogita, A. Sasaki, and K. Sato, *Prog. Theor. Phys.* **88**, 1035 (1992).
- [24] M. Van Baalen, in *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, edited by U. Dieckmann, R. Law, and J. A. J. Metz (Cambridge University Press, Cambridge, 2000), pp. 359–387.
- [25] M. J. Keeling, D. A. Rand, and A. J. Morris, *Proc. R. Soc. London, Ser. B* **264**, 1149 (1997).
- [26] T. Petermann and P. De Los Rios *J. Theor. Biol.* (to be published). e-print q-bio.PE/0401028.
- [27] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases* (Wiley, Chichester, 2000).
- [28] M. Molloy and B. Reed, *Random Struct. Algorithms* **6**, 161 (1995).
- [29] M. Molloy and B. Reed, *Combinatorics, Probab. Comput.* **7**, 295 (1998).
- [30] S. Maslov and K. Sneppen, *Science* **296**, 910 (2002).
- [31] J. Gross and J. Yellen, *Graph Theory and its Applications* (CRC Press, Boca Raton, FL, 1999), p. 154.