Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 229 (2004) 1–11
Journal of Theoretical Biology
www.elsevier.com/locate/jtbi

Cluster approximations for epidemic processes: a systematic description of correlations beyond the pair level

Thomas Petermann*, Paolo De Los Rios

Institut de Physique Théorique, Université de Lausanne, CH-1015, Lausanne, Switzerland

Received 8 August 2003; received in revised form 15 February 2004; accepted 17 February 2004

Abstract

The spread of a virus is an example of a dynamic process occurring on a discrete spatial arrangement. While the mean-field approximation reasonably reproduces the spreading behaviour for topologies where the number of connections per node is either high or strongly fluctuating and for those that show small-world features, it is inaccurate for lattice structured populations. Various improvements upon the ordinary pair approximation based on a number of assumptions concerning the higher-order correlations have been proposed. To find approaches that allow for a derivation of their dynamics remains a great challenge. By representing the population with its connectivity patterns as a homogeneous network, we propose a systematic methodology for the description of the epidemic dynamics that takes into account spatial correlations up to a desired range. The equations that the dynamical correlations are subject to are derived in a straightforward way, and they are solved very efficiently due to their binary character. The method embeds very naturally spatial patterns such as the presence of loops characterizing the square lattice or the tree-like structure ubiquitous in random networks, providing an improved description of the steady state as well as the invasion dynamics.

Keywords: Epidemic spreading; Heterogeneous mixing; Lattice model; Network model; Pair approximation; Correlation equations; Dynamical higher-order correlations

1. Introduction

The spreading dynamics of an infectious disease is determined by the connectivity patterns which underlie the population. Followed by the renewed interest in graph theory witnessed by statistical physics in recent years (Albert and Barabási, 2002; Dorogovtsev and Mendes, 2002), substantial progress has been achieved in the field of epidemiology. Possible contact networks such as the Internet or the web of human sexual contacts obey a scale-free degree distribution, meaning that the number of connections per node (i.e. the degrees) are distributed according to a power law (Faloutsos et al., 1999; Liljeros et al., 2001). Pastor-Satorras and Vespignani (2001) showed that it is this topological property which accounts for the absence of a finite epidemic threshold in the corresponding spreading phenomenon, and only targeted immunization causes this value to be non-zero (Barabási et al., 2003).

The just mentioned epidemiological insights were gained at the mean-field level, a description which does not take into account dynamical correlations. If the connectivity fluctuations of the underlying network are large, this being the case in the above examples, or for topological arrangements characterized by a high average degree, correlations of this type are indeed very weak. Furthermore, this approximation yields fairly reasonable predictions for small-world networks, in which any two nodes are only a few links apart from each other (Watts and Strogatz, 1998). Quite often however, e.g. in a population arranged on a lattice, spatial correlations can no longer be ignored. Matsuda et al. (1992) first used the ordinary pair approximation for the treatment of a population biological problem, and the resulting improvements are considerable. In addition to the analytical tractability of this approximation, simple estimates for the epidemic threshold can be obtained in terms of a "dyad heuristic": a condition for the location of the critical point is elaborated by looking

*Corresponding author. Present address: Laboratoire de Biophysique Statistique, ITP - FSB, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. Tel.: +41-21-693-0520; fax: +41-21-693-0523.

E-mail address: Thomas.Petermann@alumni.ethz.ch (T. Petermann).

0022-5193/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2004.02.017

at two neighbouring infected sites, comparing its recovery with the infections it gives rise to (Durrett and Levin, 1996).

Various extensions upon the standard pair approximation have been proposed. In order to analyse the propagation of a wave-like invasion in a lattice structured population, a remarkable improvement is brought about if the region occupied by the infected individuals is described by the ordinary pair approximation, and the leading edge of the wave front is modelled by the quasi-steady-state pair approximation (Ellner et al., 1998). Bauch and Rand (2000) developed a pair model for the situation where the population's underlying connectivity patterns are of a dynamic nature.

The grid-like structure or the presence of triangles are topological properties which both the mean-field and the standard pair approximations do not account for. An approach pursued by different authors is to use parameters that characterize the topology (such as the density of triangles) and to make a number of assumptions about the corresponding higher-order correlations, which leads to improved pair models. Van Baalen (2000) illustrates this method for the triangular and square lattices, if the higher-order correlations are set to 1. The invasion dynamics is reproduced very accurately. The same strategy can be explored for less homogeneous networks (Keeling et al., 1997), and the consequences regarding epidemiological invasions have been discussed in detail (Keeling, 1999). The improved pair approximation (Sato et al., 1994; Sato and Iwasa, 2000) takes into account the clustering property of lattice models more precisely. Its key ingredient is to make less restrictive assumptions about the higher-order correlations, e.g. they can be set to a value not equal to 1.

Morris (1997) derived the dynamics of higher-order correlations in the usual vein, i.e. an equation which determines the time evolution of an average quantity is used as point of departure.

In this paper, we introduce a novel method for the description of the epidemic dynamics which takes into account spatial correlations up to a desired range. The methodology is illustrated for the case where the population is arranged on a homogeneous network. Concerning the local contact process, we use a susceptible-infected-susceptible model involving transition rates between the two possible states (e.g. Diekmann and Heesterbeek, 2000; Durrett and Levin, 1994). The formalism is elaborated in discrete time, and the continuous-time dynamics arises as a limiting case. This limit has been performed in order to allow for a comparison with the above sketched approaches.

The paper is organized as follows. In Section 2, the adopted model involving the local dynamics as well as the selected geometries are described in detail. Section 3 reviews the mean-field and pair approximations, introduces our formalism and explains how these approximations are recovered. In Section 4, the method is illustrated for a random homogeneous network, the triangular and square lattices. Section 5 offers a discussion of the results as well as some suggestions for further investigations.

2. The model

Our approach conceives the population as a network, with connections between individuals that do not change in the course of time. Each node of the network represents an individual, and every link symbolizes a relationship between individuals that involves repeated contacts, and therefore the transmission of an infective agent proceeds along connections. As the aim of this paper is the introduction of a methodology that systematically takes into account higher-order correlations, we adopt the simple susceptible-infected-susceptible model and focus on networks where every node has the same number of nearest neighbours. Despite the homogeneity of these graphs, there exist several classes of such networks differing in topological properties beyond the degree distribution. We shall oppose the regular square lattice to the case where the underlying contact structure is fully random, furthermore our approximation scheme is illustrated for a triangular lattice. The generalization to the SIR- or SEIR-models, where the individuals can be in 3 or even 4 possible states, is straightforward.

A homogeneous random network of degree K and size N is constructed as follows (Fig. 1). To each of the

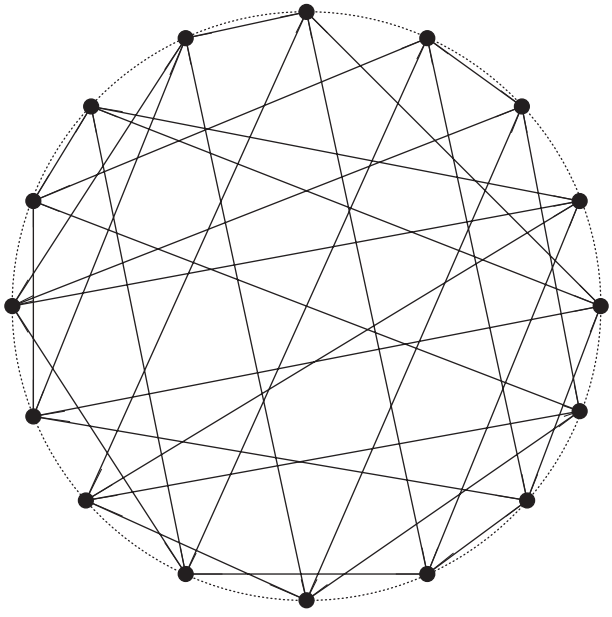

Fig. 1. Construction of a homogeneous random network of degree $K = 4$ and size $N = 16$. The nodes are connected randomly and its number of emanating edges are constrained to be $K = 4$, not allowing multiple connections and self-loops.

N vertices, K ends of edges are attached. The free ends are then connected at random, such that multiple connections and self-loops are avoided. This procedure leads to networks consisting of only one component for $K \geqslant 3$, in agreement with Newman et al. (2001).

The time evolution of the states of the vertices are given by the following rules. Infected nodes recover spontaneously at a rate δ. On the other hand, an infected individual can infect any of its K nearest neighbours at a rate v. Since what matters is the ratio of the transmission and recovery rates, we can reduce the number of parameters by rescaling the time unit. Thus without loss of generality, the local dynamics is determined by the recovery rate 1 and the effective spreading rate $\lambda = v/\delta$. In Section 3, we will elucidate how this continuous-time model is recovered as a limiting case from a more general discrete-time description.

3. Revisiting the mean-field and pair approximations

In this section, we first review the mean-field and standard pair approximations. These descriptions are obtained by a rate equation which determines the time evolution of some average quantity such as the density of infected individuals or the density of pairs of infected individuals. Up to the level of pair correlations, this is indeed a reasonable approach. But if one wants to keep track of higher-order correlations (e.g. the density of plaquettes of four infected nodes in the case of the square lattice), a more general starting point reveals itself as advantageous. In Section 3.2, we derive an exact description of the epidemic dynamics and show how the mean-field and standard pair approximations are recovered in a rather automatic way in part C of this section. The various higher-order approximations are elaborated thereafter.

3.1. Conventional approach

The rate of change of an average quantity f (such as the fraction of sites in a particular state) is described as

$$\dot{f} = \sum_{x \in X} \sum_{e_x \in E_x} r(e_x)(f_{e_x} - f), \tag{1}$$

where X is the set of all sites, and E_x represents the set of all events that can occur at x. A particular event e_x changes the average from f to f_{e_x} and occurs at rate $r(e_x)$ (Van Baalen, 2000).

The SIS-model allows for two possible states, namely susceptible (0) and infected (1). At the mean-field level, the dynamics is described in terms of the density of infected individuals ρ_1, and the fraction of susceptible nodes obeys $\rho_0 = 1 - \rho_1$. Eq. (1) translates into

$$\dot{\rho_1} = -\rho_1 + \lambda K \rho_0 \rho_1. \tag{2}$$

The first term accounts for infected nodes becoming healthy whereas the second term describes the new infections, fully ignoring pair correlations.

In the framework of the standard pair approximation (Matsuda et al., 1992), the dynamics is described in terms of the doublet densities ρ_{xy} $(x, y \in \{0, 1\})$, this quantity corresponds to the probability that a randomly chosen pair is in configuration (xy). They are related to the global densities ρ_x and local densities (conditional probabilities) $\rho_{x\,|\,y}$ by: $\rho_{xy} = \rho_{yx} = \rho_x \rho_{y\,|\,x} = \rho_y \rho_{x\,|\,y}$. The global and local densities satisfy

$$\sum_{x=0}^{1} \rho_x = 1 \quad \text{and} \quad \sum_{x=0}^{1} \rho_{x\,|\,y} = 1 \quad \text{for any} \quad y \in \{0, 1\}.$$

Eq. (1) tells that the density of infected individuals and the doublet density ρ_{11} evolve in time according to

$$\begin{aligned} \dot{\rho_1} &= -\rho_1 + \lambda K \rho_{0|1} \rho_1, \\ \dot{\rho_{11}} &= -2\rho_{11} + 2\lambda \rho_{10} + 2\lambda (K-1) \rho_{1|01} \rho_{10}. \end{aligned} \tag{3}$$

The first of Eq. (3) can also be regarded as the result of substituting ρ_0 by $\rho_{0|1}$ in Eq. (2), i.e. the susceptible node that is to be infected has to be a nearest neighbour of the vertex which will transmit the infective agent. The second of Eq. (3) includes a recovery term (the first term on the right-hand side, destruction of (11)-pairs) and transmission terms (the second and the third terms, creation of (11)-pairs). The first term describes transitions of pairs in state (11) to either (10) or (01). Both transitions occur at rate 1 (the recovery rate) and thus give rise to the factor 2. The factor 2 in the second and the third terms is needed because we do not assume any asymmetry between sites, which means $\rho_{10} = \rho_{01}$. A (11)-pair can be created from a (10)-pair either if the infective agent proceeds along the connection within that pair (second term) or if the susceptible node is infected by one of the other $K - 1$ nearest neighbours of it (third term, see also Fig. 2). This path involves the conditional probability $\rho_{1|01}$ (i.e. the probability of finding an infected node adjacent to a (01)-pair) which

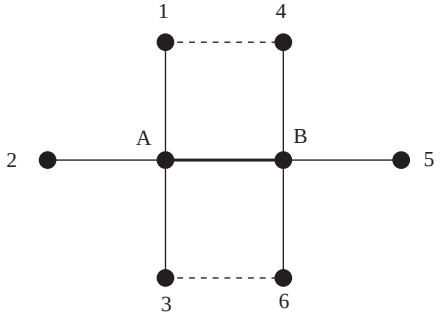

Fig. 2. An arbitrarily chosen link and its nearest neighbourhood within a homogeneous network characterized by the degree distribution $P(k) = \delta_{k4}$. The dashed lines indicate the connections which are present in the case of a square lattice.

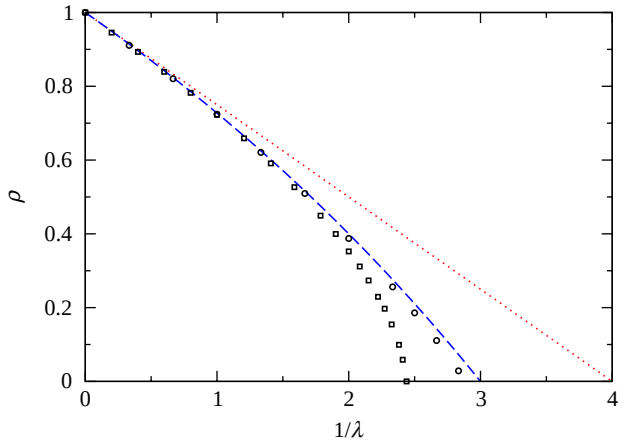

Fig. 3. Epidemic spreading on homogeneous networks of degree 4. The average number of infected individuals ρ (prevalence) as a function of the inverse infection rate $1/\lambda$ in the steady state is shown. The simulation result for the square lattice (squares) was obtained by relaxing the system (of size $N = 10^4$) into equilibrium for 10 different initial configurations, henceforth this shall be referred to as the number of iterations. In the case of the random network (circles), we further averaged over 10 realizations of networks consisting of $N = 10^5$ nodes. The adopted timestep was $\Delta t = 0.01$ for both examples. This figure shows that the simulations exhibit higher epidemic thresholds with respect to the approximations. The mean-field description (dotted line) yields $\lambda_c = \frac{1}{4}$ whereas the pair approximation (dashed line) leads to $\lambda_c = \frac{1}{3}$ for the epidemic threshold. The latter is also in better agreement with the simulation results for $1/\lambda \to 0$.

is approximated by $\rho_{1|0}$ as in the ordinary pair approximation, and only nearest-neighbour correlations are taken into account. In order to solve Eq. (3), the system has to be closed. The set $\rho_1, \rho_{1|1}$ is a suitable choice but ρ_{11}, ρ_{10} works equally well.

Fig. 3 contrasts the solutions of Eqs. (2) and (3) with the simulations for two different homogeneous networks of degree $K = 4$, i.e. the square lattice and the one introduced in Fig. 1. The pair approximation provides a rather good description of the equilibrium dynamics on top of a random homogeneous network, whereas the deviation from the simulation result is remarkable if the population is arranged on a square lattice whose topology is characterized by the presence of many loops of short length.

We shall now develop a more general formalism that will serve as a starting point in order to investigate the role of correlations beyond the pair level.

3.2. Exact description

In order to arrive at a more general point of departure which will allow us to investigate the role of higher-order spatial correlations, we shall describe the system by assigning a probability $\mathscr{P}_t(\mathbf{x})$ to every possible configuration $\mathbf{x}$ at a given time t where each of the $x_i's$ can be either 0 (susceptible) or 1 (infected). This probability is subject to

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

at every instant of time. The SIS model introduced in Section 2 implies that: (i) infected nodes recover with probability Δt and (ii) they can infect any susceptible nearest neighbour with probability $\lambda \Delta t$. Clearly, λ and Δt must be chosen such that the resulting probabilities are smaller than 1. For the possible events that can occur at an arbitrary site l, we obtain the following transition probabilities:

$$W^l_{1\to 0} = \Delta t, \quad W^l_{0\to 0} = \prod_{j\mathrm{nn}l} (1 - \lambda \Delta t y_j),$$

$$W^l_{1\to 1} = 1 - \Delta t, \quad W^l_{0\to 1} = 1 - \prod_{j\mathrm{nn}l} (1 - \lambda \Delta t y_j),$$

where the products have to be taken over the nearest neighbours of site l. By using the binary variable x_l in addition to y_l, the above expressions are summarized as

$$W^l_{y_l\to x_l} = x_l + (1 - 2x_l)\left[\Delta t y_l + (1 - y_l)\prod_{j\mathrm{nn}l} (1 - \lambda \Delta t y_j)\right]. \quad (4)$$

If the total number of nodes is denoted by N, the transition probability that the system changes from configuration $\mathbf{y}$ to $\mathbf{x}$ can be written as

$$\mathscr{W}_{\mathbf{y}\to\mathbf{x}} = \prod_{l=1}^{N} W^l_{y_l\to x_l} \quad (5)$$

and on an exact level, the epidemic dynamics is governed by

$$\mathscr{P}_{t+\Delta t}(\mathbf{x}) = \sum_{\mathbf{y}} \mathscr{W}_{\mathbf{y}\to\mathbf{x}} \mathscr{P}_t(\mathbf{y}) \quad (6)$$

with $\mathscr{W}_{\mathbf{y}\to\mathbf{x}}$ given by Eq. (5). Eq. (6) will serve as a starting point for various approximations, be it in discrete or continuous time. In the latter case, only the terms up to order 1 in Δt have to be taken into account, but this limit shall be carried out later on. As most of the existing methods are formulated in continuous time, we will elaborate the approximations for this case in order to allow for a comparison.

3.3. Derivation of the mean-field and pair approximations

Within this subsection, it is shown how approximations (2) and (3) are recovered from the exact description (6).

At the mean-field level, the dynamics is expressed in terms of the density of infected individuals. This quantity corresponds to the probability that an arbitrarily chosen site i is in state $x_i = 1$. In order to derive its time evolution, we sum Eq. (6) over all possible

configurations, x_i held fixed

$$\sum_{\{x_j\}_{j\neq i}} \mathscr{P}_{t+\Delta t}(\mathbf{x}) = \sum_{\mathbf{y}} \mathscr{P}_t(\mathbf{y}) \underbrace{\sum_{\{x_j\}_{j\neq i}} \mathscr{W}_{\mathbf{y}\to\mathbf{x}}}_{W^i_{y_i\to x_i}}. \tag{7}$$

The left-hand side of the above equation is $P_{t+\Delta t}(x_i)$, i.e. the probability that site i is in state x_i at time $t+\Delta t$. The mean-field approximation consists in considering the sites as independent from each other, i.e.

$$\mathscr{P}_t(\mathbf{y}) = \prod_{l=1}^{N} P_t(y_l), \tag{8}$$

which corresponds to the homogeneous mixing hypothesis. Performing the summations, we find for $x_i = 1$

$$P_{t+\Delta t}(1) = 1 - \Delta t P_t(1) - P_t(0)[1 - \lambda \Delta t P_t(1)]^K, \tag{9}$$

whose continuous-time limit ($\Delta t \to 0$) is

$$\dot{P}(1) = -P(1) + \lambda K P(0) P(1),$$

which is easily identified with Eq. (2) since $P(1) = \rho_1$ and $P(0) = \rho_0$.

Let us now see how the pair approximation is obtained by using our formalism. For this purpose, we sum Eq. (6) over all possible configurations, x_A and x_B held fixed, where A and B are the two sites of an arbitrarily chosen pair

$$\sum_{\{x_i\}_{i\notin\{A,B\}}} \mathscr{P}_{t+\Delta t}(\mathbf{x}) = \sum_{\mathbf{y}} \mathscr{P}_t(\mathbf{y}) \underbrace{\sum_{\{x_i\}_{i\notin\{A,B\}}} \mathscr{W}_{\mathbf{y}\to\mathbf{x}}}_{W^A_{y_A\to x_A} W^B_{y_B\to x_B}}. \tag{10}$$

The left-hand side of the above equation corresponds to the probability that the pair AB is in state $(x_A x_B)$ at time $t+\Delta t$, which shall be denoted by $P_{t+\Delta t}(x_A x_B)$. By adopting the enumeration introduced in Fig. 2, we obtain from Eq. (4) for the transition probability $(y_A y_B) \to (x_A x_B)$

$$\begin{aligned} W^A_{y_A\to x_A} W^B_{y_B\to x_B} &= \tau_A \tau_B + \Delta t(1-2x_A)[y_A - \lambda(1-y_A) \\ &\quad \times (y_B + y_1 + y_2 + y_3)]\tau_B \\ &\quad + \Delta t(1-2x_B)[y_B - \lambda(1-y_B) \\ &\quad \times (y_A + y_4 + y_5 + y_6)]\tau_A, \end{aligned} \tag{11}$$

where the linearization has been carried out at this point due to technical convenience and

$$\tau_i = \tau_i(x_i, y_i) \equiv x_i + (1-2x_i)(1-y_i), \tag{12}$$

an abbreviation which will also be used below. Furthermore, expression (11) only involves state variables y_i where i is either A, B or one of its nearest neighbours. The sum over the remaining y_j is therefore carried out trivially. Taking into account correlations up to range 2 only, we write for the probability that the pair AB and its nearest neighbours are in given states

$$\begin{aligned} P_t &\begin{pmatrix} & y_1 & y_4 & \\ y_2 & y_A & y_B & y_5 \\ & y_3 & y_6 & \end{pmatrix} \\ &= P_t(y_A y_B) P_t(y_1 \mid y_A) P_t(y_2 \mid y_A) P_t(y_3 \mid y_A) \\ &\quad \times P_t(y_4 \mid y_B) P_t(y_5 \mid y_B) P_t(y_6 \mid y_B). \end{aligned} \tag{13}$$

The conditional probabilities in the above ansatz are expressed as

$$P(y_i \mid y_A) = \frac{P(y_i y_A)}{P(y_A)},$$

where $P(y_A) = \sum_{x=0}^{1} P(x y_A)$. Using this ansatz and performing the remaining summations, the continuous-time-limit of Eq. (10) leads to the system (for general K)

$$\begin{aligned} \dot{P}(00) &= 2P(10)\left[1 - \lambda(K-1)\frac{P(00)}{P(0)}\right], \\ \dot{P}(10) &= P(11) - P(10) + \lambda P(10)\left[2(K-1)\frac{P(00)}{P(0)} - K\right], \\ \dot{P}(11) &= -2P(11) - 2\lambda P(10)\left[(K-1)\frac{P(00)}{P(0)} - K\right]. \end{aligned} \tag{14}$$

By identifying the pair probabilities $P(xy)$ with the doublet densities ρ_{xy} and since $\rho_{00}/\rho_0 = 1 - \rho_{10}/\rho_0$, the system of Eqs. (14) corresponds to Eqs. (3).

In summary, in the standard derivation of the mean-field and pair approximations based on Eq. (1), the rate of change of an average density is directly expressed by all different events that can alter its value in a rather heuristic way (Eqs. (2) and (3)). On the other hand, the derivation of the approximations becomes an automatic procedure involving:

- an initial summation of the system probability $\mathscr{P}_{t+\Delta t}(\mathbf{x})$ over all possible states except a few in order to obtain $P_{t+\Delta t}(x)$ or $P_{t+\Delta t}(x_A x_B)$ (Eqs. (7) and (10)),
- an ansatz corresponding to the approximation (Eqs. (8) and (13)),
- and the continuous-time limit.

However, the last step is not really imperative. Our methodology works equally well in discrete time. If Δt is set to 1, $\lambda \Delta t = \lambda$ then corresponds to a probability rather than to a rate and higher-order terms in λ appear in the equations. As an example, the discrete-time evolution at the mean-field level is governed by Eq. (9). Obviously, the results quantitatively differ from the continuous-time limit. The full advantage of this formalization will become clear in the next section.

It is also important to note that topological properties beyond the degree distribution do not enter at the level of the standard pair approximation. In the case of the square lattice, the nodes 1 and 4 as well as 3 and 6 (Fig. 2) are also connected whereas these links are missing in its random counterpart. Various improvements upon the ordinary pair approximation have been

proposed. Instead of deriving the higher-order correlations from the dynamics of the system, these pair models consist in making a number of biologically motivated assumptions involving parameters that characterize the topology of the underlying network. We shall compare our approach with these improved pair models in the next section.

4. Further systematic improvement

The difference between the simulation results and the pair approximation in Fig. 3 is rooted in the neglection of correlations of range greater than 2. Especially in the vicinity of the phase transition, where a finite fraction of the nodes starts being infected, the links should not be considered independently, and higher-order dynamical correlations have to be taken into account. In other words, the state x_i of node i at time $t+\Delta t$ is determined by all the states of its nearest neighbours, i.e. it is not the case that the states of the various nearest neighbours at time t contribute independently from each other to the state x_i at time $t+\Delta t$.

We therefore want to incorporate the longer correlation range by extending the fundamental cluster (site, pair) to a star or square, respecting the underlying network's topology. Therefore, different spatial patterns are embedded very naturally by our method. The equations, to which the dynamics of the higher-order correlations are subject to, are derived in a very straightforward way by our formalism. The binary nature of these equations allows for a very efficient solution by the computer. On the other hand, the equations can be simplified further by taking into account the underlying symmetries. This procedure will be illustrated for the triangular and square lattices. Performing this extension, we find an improved description of the steady state as well as the dynamics.

Alternatively, it is possible to derive the dynamics of triple correlations by using Eq. (1) (Morris, 1997). Although this approach has the advantage that no specific cluster must be chosen, it is a rather difficult undertaking.

4.1. Homogeneous random network

Random networks are characterized by a vanishing clustering (local interconnectedness), but the average distance between any pair of nodes only increases logarithmically with the system size: this is known as the small world phenomenon (Watts and Strogatz, 1998). It is easy to imagine that the more rapidly the epidemic spreads the smaller the underlying "world" is.

As the local topology is fully tree-like, we shall use a star as our fundamental element. In contrast to regular lattices, this extension is a unique choice. Fig. 4 shows an arbitrarily chosen node in a homogeneous random network and two hierarchies of its nearest neighbours, also introducing the notation which is adopted below.

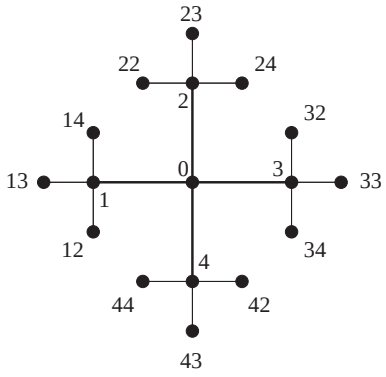

Fig. 4. An arbitrary node (denoted by 0) with its corresponding star-like fundamental cluster within a homogeneous random network of degree $K=4$.

The probability that, at time t, node 0 is in state x_0 and its nearest neighbours $1,2,3,4$ are in the states $\{x_1,x_2,x_3,x_4\}$ is denoted by

$$P_t\begin{pmatrix} & x_2 & \\ x_1 & x_0 & x_3 \\ & x_4 & \end{pmatrix}$$

and obtained by summing $\mathscr{P}_t(\mathbf{x})$ over all possible configurations $\{x_0,x_1,x_2,x_3,x_4\}$ held fixed.

The probability that the nearest and second-nearest neighbours of node 0 are in given states is given by the ansatz (the sum over the remaining y-states is again performed trivially)

$$P_t(\{y_j\}_{j\in\mathscr{N}}) = P_t\begin{pmatrix} & y_2 & \\ y_1 & y_0 & y_3 \\ & y_4 & \end{pmatrix}\prod_{l=1}^{4} P_t(y_{l2}y_{l3}y_{l4}\,|\,y_l y_0), \tag{15}$$

where $\mathscr{N}$ represents the set of nodes depicted in Fig. 4 and the conditional probabilities are

$$P_t(y_{l2}y_{l3}y_{l4}\,|\,y_l y_0) = \frac{P_t\begin{pmatrix} & y_{l4} & \\ y_{l3} & y_l & y_0 \\ & y_{l2} & \end{pmatrix}}{P_t(y_l y_0)}.$$

The pair probability appearing in the above expression is extracted from the corresponding star probabilities by

$$P_t(y_l y_0) = \sum_{y_{l2}=0}^{1}\sum_{y_{l3}=0}^{1}\sum_{y_{l4}=0}^{1} P_t\begin{pmatrix} & y_{l4} & \\ y_{l3} & y_l & y_0 \\ & y_{l2} & \end{pmatrix}.$$

With these ingredients, the continuous-time limit of Eq. (6) reads

$$\dot{P}\begin{pmatrix} & x_2 & \\ x_1 & x_0 & x_3 \\ & x_4 & \end{pmatrix} = \sum_{\{y_j\}_{j\in\mathcal{N}}} \left[P(\{y_j\}_{j\in\mathcal{N}})\right. \times \left.\prod_{i=0}^{4}\left\{(1-2x_i)\left[y_i - \lambda(1-y_i)\sum_{j\mathrm{nn}i} y_j\right]\prod_{k\neq i}\tau_k\right\}\right] \quad (16)$$

with $P(\{y_j\}_{j\in\mathcal{N}})$ given by Eq. (15). The binary character of this system of $2^5 = 32$ equations permits a very efficient numerical implementation. On the other hand, if one takes into account the symmetries of the problem, the degrees of freedom can be reduced to 10; this procedure will be shown for the regular lattices. The left panel of Fig. 5 shows the striking agreement of the star approximation with the simulation result, all along from a high effective spreading rate to its threshold value, for the equilibrium situation. Its right part opposes the various approximations to the stochastic simulation for the case of the invasion of an infective agent, the initial prevalence being set to 0.01. Whereas the steady state is reached rather quickly in the mean-field description, the slope of the star approximation is in remarkable agreement with the simulation. As correlations of a greater range are taken into account, it can also be observed that the system equilibrates more smoothly, that is $\ddot{\rho}(t\simeq 30)$ for the star approximation is considerably smaller than the rate of change of $\dot{\rho}$ at time $t\simeq 10$ at the mean-field level.

4.2. Square lattice

In contrast to random graphs, the epidemic dynamics on top of this regular network is essentially dominated by the presence of loops. In order to arrive at a level of description beyond the pair approximation, we shall use the square as our fundamental cluster. This seems to be a natural choice, although it is not unique as discussed below. In analogy to the previous subsection, the probability that the corners of the square $ABCD$ are in the states $\{x_A, x_B, x_C, x_D\}$ at time t is deduced from the system probability by

$$P_t\begin{pmatrix} x_A & x_B \\ x_D & x_C \end{pmatrix} = \sum_{\{x_i\}_{i\notin\{A,B,C,D\}}} \mathscr{P}_t(\mathbf{x}).$$

If the nearest neighbours of the vertices A, B, C and D are enumerated according to Fig. 6, we write for the

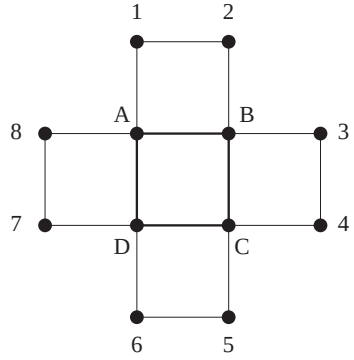

Fig. 6. An arbitrarily chosen square within a two-dimensional lattice and the denotation of the nearest neighbours of its corners. The former serves as the fundamental element within the square approximation.

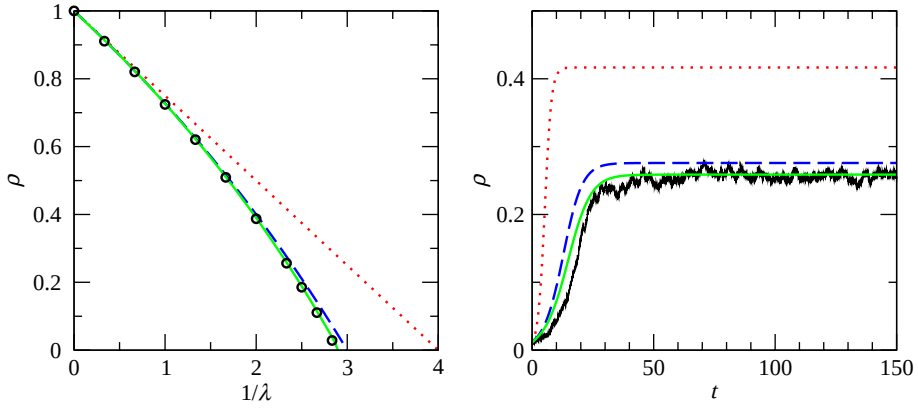

Fig. 5. Spreading behaviour for a population arranged on a random network obeying $P(k) = \delta_{k4}$. The left panel shows the simulation result (circles) and different levels of approximations for the equilibrium dynamics. The former was obtained by performing 10 iterations on 10 different realizations of networks of size $N = 10^5$, the transition probabilities being determined by $\Delta t = 0.01$. The star approximation (solid line) is in excellent agreement with the simulation result, yielding also an accurate description of the critical region. The mean-field (dotted line) and pair approximation (dashed line) have been plotted again for comparison. The right part shows the invasion of an infective agent (infection rate $\lambda = \frac{3}{7}$) on the same type of topology. At the mean-field level (dotted line), the initial prevalence of 0.01 increases to its equilibrium value during 10 time units only. The pair approximation (dashed line) provides a further improvement, and the star approximation (solid line) is in remarkable agreement with the stochastic simulation whose parameters N and Δt correspond to those already mentioned above.

probability that the nodes comprised within these 5 squares (i.e. the nearest neighbours of the central plaquette) are in given states

$$P_t(\{y_i\}_{i\in\{A,B,C,D,1,2,\ldots,8\}}) = P_t\begin{pmatrix} y_A & y_B \\ y_D & y_C \end{pmatrix} P_t(y_1y_2\,|\,y_Ay_B)P_t(y_3y_4\,|\,y_By_D) \times P_t(y_5y_6\,|\,y_Cy_D)P_t(y_7y_8\,|\,y_Ay_C) \quad (17)$$

with

$$P_t(y_1y_2\,|\,y_Ay_B) = \frac{P_t\begin{pmatrix} y_1 & y_2 \\ y_A & y_B \end{pmatrix}}{P_t(y_Ay_B)}$$

involving the pair probability

$$P_t(y_Ay_B) = \sum_{x_1=0}^{1}\sum_{x_2=0}^{1} P_t\begin{pmatrix} x_1 & x_2 \\ y_A & y_B \end{pmatrix}$$

and analogously for the other factors appearing in Eq. (17). At this point, we could again write down an equation of type (16), but we shall explicitly make use of the symmetries of the problem in order to reduce the computational load. The $2^4 = 16$ plaquette probabilities are subject to

$$P_t\begin{pmatrix} 00 \\ 00 \end{pmatrix} \equiv q_{0,t},$$

$$P_t\begin{pmatrix} 10 \\ 00 \end{pmatrix} = P_t\begin{pmatrix} 01 \\ 00 \end{pmatrix} = P_t\begin{pmatrix} 00 \\ 10 \end{pmatrix} = P_t\begin{pmatrix} 00 \\ 01 \end{pmatrix} \equiv q_{1,t},$$

$$P_t\begin{pmatrix} 11 \\ 00 \end{pmatrix} = P_t\begin{pmatrix} 01 \\ 01 \end{pmatrix} = P_t\begin{pmatrix} 00 \\ 11 \end{pmatrix} = P_t\begin{pmatrix} 10 \\ 10 \end{pmatrix} \equiv q^A_{2,t},$$

$$P_t\begin{pmatrix} 10 \\ 01 \end{pmatrix} = P_t\begin{pmatrix} 01 \\ 10 \end{pmatrix} \equiv q^C_{2,t},$$

$$P_t\begin{pmatrix} 11 \\ 10 \end{pmatrix} = P_t\begin{pmatrix} 11 \\ 01 \end{pmatrix} = P_t\begin{pmatrix} 10 \\ 11 \end{pmatrix} = P_t\begin{pmatrix} 01 \\ 11 \end{pmatrix} \equiv q_{3,t},$$

$$P_t\begin{pmatrix} 11 \\ 11 \end{pmatrix} \equiv q_{4,t}.$$

The exact description (6) leads to the following continuous-time dynamics for these quantities:

$$\dot q_0 = 4q_1 - 8\lambda T_1 q_0,$$

$$\dot q_1 = -q_1 + 2q_2^A + q_2^C + \lambda[-2q_1(1+2T_1+T_2) + 2T_1q_0],$$

$$\dot q_2^A = -2q_2^A + 2q_3 + \lambda[-4q_2^A + 2T_1(q_0+3q_1) + 2T_2(q_1-q_2^A)],$$

$$\dot q_2^C = -2q_2^C + 2q_3 + \lambda(-4q_2^C + 4T_1q_1 - 4T_2q_2^C),$$

$$\dot q_3 = -3q_3 + q_4 + \lambda[2q_1 + 4q_2^A - 4q_3 - 2T_1(q_0+2q_1) + 2T_2(q_1+2q_2^A+2q_2^C)],$$

$$\dot q_4 = -4q_4 + \lambda[8q_2^C + 16q_3 - 8T_2(q_1+q_2^A+q_2^C)], \quad (18)$$

where

$$T_1 = \frac{t_1^A}{p_0^A} \quad \text{and} \quad T_2 = \frac{t_2^C}{p_1^A}$$

involving the following triplet- and pair probabilities given by the square probabilities through

$$t_1^A = P\begin{pmatrix} 0 & 0 \\ 1 & \end{pmatrix} = q_1 + q_2^A,$$

$$t_2^C = P\begin{pmatrix} 0 & 1 \\ 1 & \end{pmatrix} = q_2^C + q_3,$$

$$p_0^A = P(00) = q_0 + 2q_1 + q_2^A \quad \text{and}$$

$$p_1^A = P(10) = q_1 + q_2^A + q_2^C + q_3.$$

Since

$$q_0 + 4q_1 + 4q_2^A + 2q_2^C + 4q_3 + q_4 = 1,$$

the square approximation in form (18) represents a dynamical system of 5 degrees of freedom, in contrast to 16 if the symmetries were not exploited.

The left part of Fig. 7 shows the systematic improvement brought about by the square and bisquare approximations in dynamic equilibrium. The latter is a description whose fundamental cluster is composed of two squares. Its prediction of the epidemic threshold ($\lambda_c \simeq 0.38$) is still lower than the simulation result ($\lambda_c \simeq 0.41$): this highlights the crucial role of the higher-order spatial correlations in lattice structured populations. The right panel of Fig. 7 represents the improvements upon the dynamics. Note that from a certain characteristic time the simulation lags behind all the approximations as a direct consequence of the stochasticity that is particularly important at low prevalences. However, this characteristic time is shifted to the right as higher-order correlations of a greater range are taken into account.

An improvement upon the standard pair approximation can also be obtained as follows (Van Baalen, 2000). Instead of deriving the square probabilities from the dynamics of the system, one can write it as

$$P\begin{pmatrix} x_i & x_a \\ x_j & x_b \end{pmatrix} = P(x_i)P(x_a)P(x_b)P(x_j)C_{ia}C_{ab}C_{bj}C_{ji}T_{\square iabj}$$

involving the relative pair and square correlation factors C_{xy} and $T_{\square iabj}$. For a straight triple, it is supposed

$$P(x_ix_ax_b) = P(x_i)P(x_a)P(x_b)C_{ia}C_{ab}T_{\angle iab}.$$

By setting the relative correlation factors $T_{\square iabj}$ and $T_{\angle iab}$ to 1 and using the fact that on the square lattice $\frac{1}{3}$ of the triples are straight and $\frac{2}{3}$ form part of a square, one obtains an improvement for $P(x_i|x_ax_b)$. In other words, the conditional probability $P(x_i|x_ax_b)$ is not simply set to $P(x_i|x_a)$ as it is done in the ordinary pair

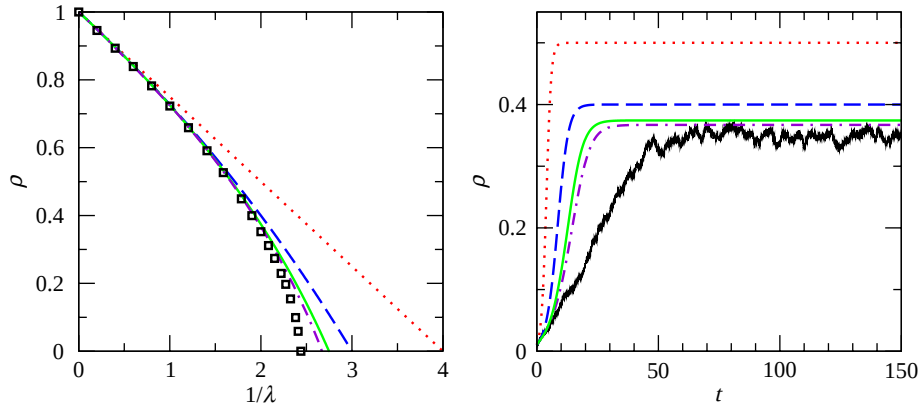

Fig. 7. Prevalences for the epidemic process on a square-lattice structured population. The left side illustrates the result for the equilibrium state. The mean-field (dotted line) and pair approximations (dashed line) are levels of description at which topological properties beyond the degree distribution do not enter. The approximations involving the square (solid line) and a rectangle composed of two squares (dashed–dotted line) as fundamental units systematically approach the steady-state behaviour, as predicted by the simulations (squares, for its details see Fig. 3). The right panel reports on the dynamics for $\lambda = 0.5$. By taking into account correlations of a greater range, the slope during the transient time decreases as a comparison of the mean-field (dotted line), pair (dashed line), square (solid line) and bisquare approximations (dashed–dotted line) shows. The difference between the simulation result (for $N = 10^4$, $\Delta t = 0.01$) and the bisquare approximation remains significant during the invasion period due to the considerable effect of random events at overall low prevalence.

approximation, but rather the loop structure is incorporated while still using pairs as building blocks.

4.3. Triangular lattice

In its ordinary formulation, the fact that two sites can have neighbours in common is simply ignored by the pair approximation. By means of the triangular lattice, we show how the method introduced in this paper has to be applied, i.e. what the next level of description beyond the pair approximation is.

The clue is to use the triangle as the basic element. In analogy to the previous cases, the probability that the vertices of a triangle ABC are in the states $\{x_A, x_B, x_C\}$ at time t is obtained through

$$P_t\begin{pmatrix} x_A x_B \\ x_C \end{pmatrix} = \sum_{\{x_i\}_{i\notin\{A,B,C\}}} \mathscr{P}_t(\mathbf{x}).$$

Fig. 8 shows the neighbourhood of an arbitrarily chosen triangle within this lattice. For the probability that the vertices depicted in Fig. 8 are in given states, we suppose

$$\begin{aligned} &P_t(\{y_i\}_{i\in\{A,B,C,1,2,\ldots,9\}}) \\ &\quad = P_t\begin{pmatrix} y_A y_B \\ y_C \end{pmatrix} P_t(y_1 \mid y_A y_B) P_t(y_4 \mid y_B y_C) \\ &\qquad \times P_t(y_7 \mid y_C y_A) P_t(y_8 y_9 \mid y_A) P_t(y_2 y_3 \mid y_B) P_t(y_5 y_6 \mid y_C). \end{aligned}$$

The conditional probabilities appearing in the above expression can be written as fractions involving site and pair probabilities. The latter are deduced from the triangle probabilities in analogy to previous explanations. As the triangle correlations are subject to

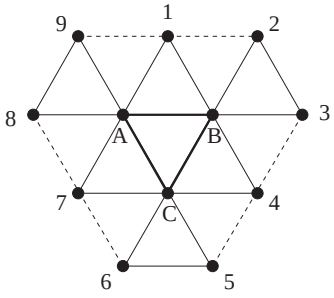

Fig. 8. An arbitrarily chosen triangle and its nearest neighbourhood. The dashed lines indicate that the corresponding links are ignored.

the symmetries

$$\begin{aligned} &P_t\begin{pmatrix} 00 \\ 0 \end{pmatrix} \equiv t_{0,t}, \\ &P_t\begin{pmatrix} 10 \\ 0 \end{pmatrix} = P_t\begin{pmatrix} 01 \\ 0 \end{pmatrix} = P_t\begin{pmatrix} 00 \\ 1 \end{pmatrix} \equiv t_{1,t}, \\ &P_t\begin{pmatrix} 11 \\ 0 \end{pmatrix} = P_t\begin{pmatrix} 10 \\ 1 \end{pmatrix} = P_t\begin{pmatrix} 01 \\ 1 \end{pmatrix} \equiv t_{2,t}, \\ &P_t\begin{pmatrix} 11 \\ 1 \end{pmatrix} \equiv t_{3,t} \end{aligned}$$

a further simplification can be performed, and finally the continuous-time triangle dynamics is governed by the equations:

$$\begin{aligned} \dot{t}_0 &= 3[t_1 - 2\lambda(A_1 + A_2)], \\ \dot{t}_1 &= -t_1 + 2t_2 + 2\lambda(-2t_1 + 3A_1 + 2A_2 - 2A_3 - A_4), \\ \dot{t}_2 &= -2t_2 + t_3 + 2\lambda(t_1 - 3t_2 - 3A_1 - A_2 + 4A_3 + 2A_4), \\ \dot{t}_3 &= 3[-t_3 + 2\lambda(t_1 + 3t_2 + A_1 - 2A_3 - A_4)], \end{aligned} \tag{19}$$

where

$$A_1 = \frac{p_1 t_0}{s_0}, \quad A_2 = \frac{t_0 t_1}{p_0},$$
$$A_3 = \frac{p_0 p_1}{s_0} \quad \text{and} \quad A_4 = \frac{t_1 t_2}{p_1}$$

depending on the pair probabilities $p_1 = P(10) = t_1 + t_2$, $p_0 = P(00) = t_0 + t_1$ and the site probability $s_0 = P(0) = t_0 + 2t_1 + t_2$. Because of the constraint

$$t_0 + 3t_1 + 3t_2 + t_3 = 1,$$

we have three degrees of freedom in the triangle approximation (19).

As far as the equilibrium prediction is concerned, the triangle approximation provides a very good description for $1/\lambda < 3$ (Fig. 9, left panel). The difference between its threshold prediction ($1/\lambda_c \simeq 4.5$) and the simulation result ($1/\lambda_c \simeq 3.9$) is of the same order of magnitude as the plaquette approximation in the case of the square lattice. Concerning the dynamics (Fig. 9, right panel), we also observe a lag between the simulation and the approximations, and the slope during the transient time is slightly improved as one goes from the pair to the triangle approximation.

The strategy outlined at the end of the last subsection can also be applied to the triangular lattice (Van Baalen, 2000). In addition to the open triplet probability, the triangle probability is written as

$$P\begin{pmatrix} x_i \\ x_a x_b \end{pmatrix} = P(x_i)P(x_a)P(x_b)C_{ia}C_{ab}T_{\triangle iab}.$$

One then obtains an analogous correction for $P(x_i \mid x_a x_b)$ involving a parameter θ denoting the fraction of triplets in closed form that is $\frac{2}{5}$ in the triangular lattice. Interestingly, the simplest elaboration of this approach ($\tau_{\triangle iab} = \tau_{\angle iab} \equiv \tau_{iab}$) reproduces the invasive period reasonably accurate if θ is chosen larger than its correct value ($\theta \simeq 0.6$). Keeling et al. (1997) and Rand (1999) also developed improved pair models based on this approach.

5. Conclusion

We have studied a dynamic model of epidemic spreading where every individual is in contact with an equal number K of nearest neighbours. Infected nodes recover spontaneously at a rate 1, on the other hand, they infect neighbouring susceptible sites at a rate λ. We have chosen this simple SIS-type model since the focus of this article is the introduction of a novel methodology that allows a rather straightforward derivation of the dynamics of higher-order correlations.

The method we propose here consists in choosing a fundamental cluster composed of a certain number of nodes n as well as links connecting them. A definite probability is assigned to each possible configuration of the basic element. The size of the fundamental cluster represents the range up to which spatial correlations are exactly taken into account. At a level beyond the pair approximation, the choice of the basic element is guided by the underlying network's topology. In the case of the square lattice, clusters composed of at least one plaquette serve as the fundamental element; for random networks the local tree-like structure is incorporated by using the star as the basic unit. Spatial patterns beyond the degree distribution are therefore embedded in a very natural way by our method. Describing the epidemic dynamics of the entire population as a discrete time Markovian process, the appearing probabilities

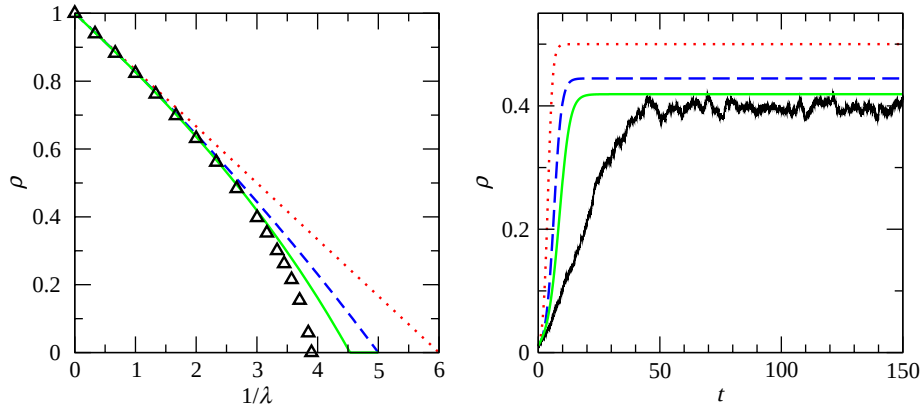

Fig. 9. Results for the SIS model on a triangular lattice. Left: steady-state prevalence as predicted by various approximations and through computer simulations (triangles) whose parameters are $N = 10^4$, $\Delta t = 0.01$ and 10 iterations. The mean-field description (dotted line) yields an epidemic threshold $\lambda_c = \frac{1}{6}$. With respect to the pair approximation (dashed line), the description based on the triangle (solid line) provides a better approximation of the simulation result. Right: invasion dynamics for an effective spreading rate $\lambda = \frac{1}{3}$. The upper two curves show the mean-field (dotted line) and pair dynamics (dashed line). The improvement brought about by the triangle approximation (solid line) still lags behind the simulation result due to the same reason as in the case of the square lattice. The latter was obtained for $N = 10^4$ and $\Delta t = 0.01$ as well.

(probability that a cluster and its nearest neighbourhood is in a given configuration) are expressed in terms of the fundamental cluster probabilities. The continuous-time dynamics emerges as a limiting case ($\Delta t \to 0$).

With respect to the ordinary (rather heuristic) derivation of the mean-field and pair approximation, these descriptions are derived with the help of our formalism by using the site or the pair, respectively, as fundamental clusters in a very automatic way. Independently of the specific choice of the cluster, the binary character of the resulting equations allows for a very efficient solution by the computer. Likewise, a further simplification can be reached if the symmetries, which the fundamental cluster probabilities are subject to, are taken into account. As soon as correlations of range greater than 2 are not ignored, our method yields improved estimates for the location of the phase transition. In the case of the random network, the star approximation already leads to an excellent description of the steady state and the transient dynamics. In the regular counterpart, many squares have to be included within the corresponding fundamental unit in order to attain the same level of accuracy. This is due to the presence of stronger correlations caused by the high local ordering. The method was also illustrated for a triangular lattice and contrasted to approaches that make a certain number of assumptions about the higher-order correlations which lead to improved pair models.

We have focused on homogeneous networks, since in this case a fundamental cluster is identified most easily. The homogeneity lies indeed at the basis of our cluster approximations since it must be possible to express the probability appearing on the right of our master equations (see e.g. Eq. (16)) entirely in terms of the fundamental cluster probabilities, such as in Eqs. (15) and (17). In principle, our method can be extended to slightly heterogeneous systems, e.g. a random network where two different degrees are present. The dynamics is then described in terms of two different star-like clusters (according to the occuring degrees), this hybridization involving the constraint that the pair probabilities derived from the two clusters must coincide.

However, the novelty of the present work lies in the formalism which essentially consists in a more general starting point and its associated systematic improvability rather than the specific results for the selected epidemiological model and geometrical examples.

Acknowledgements

We wish to thank EC-Fet Open project COSIN IST-2001-33555, and the OFES-Bern (CH) for financial support.

References

Albert, R., Barabási, A.-L., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97 doi 10.1103/RevModPhys.74.47.

Barabási, A.-L., Dezsö, Z., Ravasz, E., Yook, S.-H., Oltvai, Z., 2003. Scale-free and hierarchical structures in complex networks. AIP Conf. Proc. 661, 1–16.

Bauch, C., Rand, D.A., 2000. A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. Proc. Roy. Soc. London B 267, 2019–2027.

Diekmann, O., Heesterbeek, J.A.P., 2000. Mathematical Epidemiology of Infectious Diseases. Wiley, Chichester.

Dorogovtsev, S.N., Mendes, J.F.F., 2002. Evolution of networks. Adv. Phys. 51 (4), 1079–1187.

Durrett, R., Levin, S.A., 1994. The importance of being discrete (and spatial). Theoret. Population Biol. B 46, 363–394 doi 10.1006/tpbi.1994.1032.

Durrett, R., Levin, S.A., 1996. From individuals to epidemics. Philos. Trans. Roy. Soc. London B 351, 1615–1621.

Ellner, S.P., Sasaki, A., Haraguchi, Y., Matsuda, H., 1998. Speed of invasion in lattice population models: pair-edge approximation. J. Math. Biol. 36, 469–484.

Faloutsos, M., Faloutsos, P., Faloutsos, C., 1999. On power-law relationships of the Internet topology. Comput. Comm. Rev. 29, 251–262.

Keeling, M.J., 1999. The effects of local spatial structure on epidemiological invasions. Proc. Roy. Soc. London B 266, 859–867.

Keeling, M.J., Rand, D.A., Morris, A.J., 1997. Correlation models for childhood epidemics. Proc. Roy. Soc. London B 264, 1149–1156.

Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Åberg, Y., 2001. The web of human sexual contacts. Nature 411, 907–908 doi 10.1038/35082140.

Matsuda, H., Ogita, N., Sasaki, A., Sato, K., 1992. Statistical mechanics of population—the lattice Lotka–Volterra model. Progr. Theoret. Phys. 88, 1035–1049.

Morris, A.J., 1997. Representing spatial interactions in simple ecological models. Ph.D. Dissertation, University of Warwick, Coventry, UK.

Newman, M.E.J., Strogatz, S.H., Watts, D.J., 2001. Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E 64, 026118.

Pastor-Satorras, R., Vespignani, A., 2001. Epidemic spreading in scale-free networks. Phys. Rev. Lett. 86, 3200–3203 doi 10.1103/Phys. Rev. Lett.86.3200.

Rand, D.A., 1999. Correlation equations and pair approximations for spatial ecologies. In: McGlade, J. (Ed.), Advanced Ecological Theory: Principles and Applications. Blackwell Science, Oxford, UK, pp. 100–142.

Sato, K., Iwasa, Y., 2000. Pair approximations for lattice-based ecological models. In: Diekmann, U., Law, R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, Cambridge, UK, pp. 339–358.

Sato, K., Matsuda, H., Sasaki, A., 1994. Pathogen invasion and host extinction in lattice structured populations. J. Math. Biol. 32, 251–268.

Van Baalen, M., 2000. Pair approximations for different spatial geometries. In: Diekmann, U., Law, R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions: Simplifying Spatial Complexity. Cambridge University Press, Cambridge, UK, pp. 359–387.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of ''small-world'' networks. Nature 393, 440–442 doi 10.1038/30918.