Electronic Journal of Qualitative Theory of Differential Equations
2026, No. **10**, 1–15; https://doi.org/10.14232/ejqtde.2026.1.10 www.math.u-szeged.hu/ejqtde/

The effect of network structure on the dynamics of epidemic propagation in an individual-based model

Noémi Nagy[1], **Sándor Horváth**[2] and **Péter L. Simon** ✉[3]

[1]Department of Analysis and Operations Research, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary, National Laboratory for Health Security
[2]National Laboratory for Health Security, Institute of Mathematics, Eötvös Loránd University, Budapest, Hungary
[3]Institute of Mathematics, Eötvös Loránd University, Alfréd Rényi Institute of Mathematics, HUN-REN, Budapest, Hungary

Received 15 August 2025, appeared 12 April 2026
Communicated by Mihály Kovács

Abstract. An individual-based model of SIS epidemic propagation is considered on regular graphs. This model explicitly incorporates the adjacency matrix of the network enabling us to study the effect of network structure on the dynamic of the propagation process. Our goal in this paper is to understand how the one-dimensional approximation, called also mean-field equation, performs, depending on the main characteristics of the network. It turns out that for dense random regular graphs the approximation is excellent, while for sparse graphs a significant correction term is needed to obtain the proper time dependence of the prevalence (number of infected nodes in the network) based on solving only the mean-field model. Numerical evidence is shown that this correction term depends linearly on the mean distance in the network.

Keywords: network, epidemic, low-dimensional approximation.

2020 Mathematics Subject Classification: 34C60, 92C42.

1 Introduction

Epidemic propagation has been widely studied on networks by using stochastic and deterministic models [3–5,7,8,11]. It is common in all these modelling approaches that the contact structure is represented by a graph, the nodes of which are individuals and the connecting edges are responsible for infectious contacts and each node can be in one of the epidemic states. The mostly used states are susceptible (S), infected (I) and recovered (R). The stochastic description is given by a random process during which the state of the nodes change in two ways: by infection coming from their neighbours and by recovery. Both processes are

✉Corresponding author. Emails: nagynoemi0@gmail.com (N. Nagy), dngsanyi94@gmail.com (S. Horváth), peter.simon@ttk.elte.hu (P. L. Simon)

considered typically as Markov processes. The deterministic description can be given at different levels. The most frequently used ones are individual-level and population-level and degree-based that is between the previous two. The variables of the individual-level model are the probabilities of a given node being in a given state, e.g. $I_i(t)$ denotes the probability that node i is infected at time t. The population-level model uses aggregated variables, like $I(t)$, the number (expected value) of infected nodes at time t. (The typical variable of the degree-based model is $I_k(t)$ the expected number of degree k infected nodes at time t.)

The main topic of our paper is the relation of the individual and population-level models. While both of them are mean-field like approximations of the full stochastic model, the population-level model can be considered as a low-dimensional or coarse-grained approximation of the individual-level system. The obvious goal and advantage of the low-dimensional model is that it can be investigated analytically, for example, the stability of steady states and the occurring bifurcations can be studied by using the traditional methods of dynamical system theory. Thus the accuracy of the population-level approximation is a crucial question in studying epidemic processes on networks.

From the perspective of real-world applications, dimensionality reduction plays an important role. In describing spread processes, it is essential to know which individuals are in contact with each other; however, this is difficult to measure in practice and, moreover, leads to a high-dimensional model. Thus, a key question is how accurate an approximate model is for different types of graphs. In this paper, we primarily examine regular graphs, in which each individual has the same number of connections, nevertheless, the structure of the graph itself may vary. For example, in a workplace where most of the individuals can meet each other, a random graph model may provide a good approximation. In contrast, in a small settlement where infection spreads mainly among neighbors, the connections can be described by a geometric graph with a lattice-like structure. An intermediate case between these two extremes may be a school, where within individual classes the contact network can be characterized by a random graph, while across the entire school it exhibits a more geometric structure. Here, we reveal the role played by the mean distance—which distinguishes the above networks from one another—in determining the accuracy of the low-dimensional approximation.

The accuracy of different dynamical epidemic models to stochastic simulation has been widely studied [3–5, 7]. In [10], the authors investigate how and under which conditions can the dynamics be predicted without knowing the network topology. Our goal here is to understand how the graph theoretic properties of the network determine the performance of coarse-grained approximations. The comparison of the individual-based and population-level model provides an excellent opportunity for this, because the individual-based model explicitly contains the adjacency matrix of the network, hence graph theoretical quantities can be directly determined. According to our knowledge such a systematic comparison of these models has not been carried out yet and this research enabled us to identify graph measures that play a crucial role in determining the accuracy of low-dimensional approximations.

The methodology of our approach is the following. First we derive an accuracy measure that can be generally used to compare the high-dimensional individual model to the low-dimensional approximation. Then the difference of the two models is systematically studied on two classes of graphs: random regular graphs and structured graphs (lattices). The effect of the density of the graph and the mean distance on the accuracy is numerically investigated.

The paper is structured as follows. The individual-based model and its population-level approximation are introduced in Section 2 together with the accuracy measure enabling us to compare them. Then the known results about the asymptotic behaviour, i.e. about the

stability of the steady states are summarized in Section 3. The detailed numerical study of the accuracy of this approximation for a wide range of regular graphs is presented in Sections 4 for random graphs and in Section 5 for structured graphs. The results are summarized and further research directions are shown in the Discussion in Section 6.

The main novelties of the paper are the following. Numerical evidence is shown that for dense regular random graphs the population-level equation gives an accurate approximation of the individual-level system. For sparse graphs, it is shown that the accuracy of the approximation depends on the mean distance in the graph. Moreover, numerical experiments are carried out to justify that the accuracy depends linearly on the mean distance, yielding a numerical correction term to the population-level approximation.

2 The individual-based and population-level models

In this section we present the individual-based model of SIS epidemic propagation on networks, then show a general method for deriving low-dimensional approximation. This method is applied to the epidemic model to derive the population level approximation. Finally, a measure of accuracy of the coarse-grained model is introduced.

2.1 The individual-based epidemic model

As it was already mentioned in the Introduction, the main variables of the individual-based model are the quantities I_i representing the probabilities that node i is infected at time t. Their evolution is described by the following system of differential equations.

$$\langle \dot{I}_i \rangle = \tau \sum_{j=1}^{n} a_{ij} \langle S_i I_j \rangle - \gamma \langle I_i \rangle, \quad i = 1, 2, \dots, n, \tag{2.1}$$

where τ is the rate of infection across an edge and γ is the rate of recovery of a node, n is the number of nodes, $A = (a_{ij})$ is the adjacency matrix of the graph, $\langle I_i \rangle(t)$ is the probability that node i is infected at time t and $\langle S_i I_j \rangle(t)$ is the joint probability that node i is susceptible and node j is infected at time t. The derivation of this system from the full stochastic model can be found e.g. in Section 3 of the book [5] or in the paper [11]. It is important to observe that these differential equations are not closed in the sense that further differential or algebraic equations are needed for the joint probabilities in terms of the marginals. The widely used approximation is $\langle S_i I_j \rangle \approx \langle S_i \rangle \langle I_j \rangle = (1 - \langle I_i \rangle)\langle I_j \rangle$ expressing the independence of the infectedness of node j and susceptibility of node i (which is certainly only an approximation). Applying this approximation in equation (2.1) leads to the equation

$$\dot{x}_i = \tau(1 - x_i)(Ax)_i - \gamma x_i, \quad i = 1, \dots n, \tag{2.2}$$

where $x_i(t)$ approximates $\langle I_i \rangle(t)$, the probability that node i is infected at time t. We will refer to this differential equation as individual-based or NIMFA model, abbreviating N-intertwined mean-field approximation [13]. We note that the notion "mean-field" is used in several contexts, hence we prefer here to use "individual-based model" for (2.2). Moreover, we will use the notation

$$I(t) = \sum_{i=1}^{n} x_i(t). \tag{2.3}$$

for the main observed quantity, the expected number of infected nodes (in other words, the prevalence) obtained from the individual-based model. The theory of formulating closed

systems of ODEs describing epidemic spread on networks is well developed, see e.g. [3–5, 7, 8], where second order approximations (containing differential equations for the joint probabilities as well) are also dealt with. We chose the individual-based model (2.2) because it enables us to study the effect of network structure directly starting from the adjacency matrix.

2.2 Lumping: a general method for deriving low-dimensional approximations

Let us now turn to the derivation of the population-level approximation. First, we present the general idea of deriving low-dimensional approximations of large systems of differential equations and then we apply this to system (2.2).

Assume that we are given the high-dimensional system

$$\dot{x}(t) = f(x(t)), \tag{2.4}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$, and n is large. Furthermore, we have an observation $y(t) \in \mathbb{R}^m$ that is significantly lower dimensional, than the whole state space $\mathbb{R}^n$, that is

$$y(t) = h(x(t)), \tag{2.5}$$

where $h : \mathbb{R}^n \to \mathbb{R}^m$, and $m < n$, in fact, m is significantly smaller, than n. The problem is to find a function $g : \mathbb{R}^m \to \mathbb{R}^m$, for which the solution of system

$$\dot{z}(t) = g(z(t)), \tag{2.6}$$

is a good approximation of y, that is the error function $|y(t) - z(t)|$ is small. The low-dimensional system is sometimes called the **lumped** system. If the error is zero, then it is called exact lumping, while for non-zero error it is an approximate lumping.

The main idea of deriving an exact lumping is based on differentiating the identity $y(t) = h(x(t))$ and use the differential equation of x as follows:

$$\dot{y}(t) = h'(x(t)) \cdot \dot{x}(t) = h'(x(t)) \cdot f(x(t)).$$

In the case of exact lumping, this leads to

$$h' \cdot f = g \circ h. \tag{2.7}$$

Namely, the above equation can be written as:

$$\dot{y}(t) = h'(x(t)) \cdot f(x(t)) = g(h(x(t))) = g(y(t)),$$

thus the observation y can be obtained directly from the low-dimensional (coarse-grained) system $\dot{y} = g(y)$.

In the case of approximate lumping, we would like to show that

$$h' \cdot f - g \circ h$$

is small in some norm and we expect that the solution of $\dot{z} = g(z)$ is an accurate approximation of y.

In the next subsection, we show how this general approach can be realised in the case of the individual-based model of epidemic propagation on a network.

2.3 The population-level epidemic propagation model

Now, we apply the approach presented in the previous subsection to system (2.2) with the one-dimensional observation $I(t)$ given in (2.3). This will correspond to (2.4) when

$$f(x) = \tau(\underline{e} - x) * Ax - \gamma x$$

with $\underline{e} = (1, 1, \ldots, 1)^T$ the vector with all its coordinates equal to 1 and the product $*$ denotes multiplication coordinate-wise.

Thus the left hand side of (2.7) takes the form $h'(x(t))f(x(t)) = \underline{e}^T f(x(t)) = \dot{I}(t)$, that can be written as

$$\dot{I} = \sum_{i=1}^{n} \dot{x}_i = \tau \sum_{i=1}^{n} (Ax)_i - \tau \sum_{i=1}^{n} x_i (Ax)_i - \gamma \sum_{i=1}^{n} x_i = \tau \underline{e}^T Ax - \tau x^T Ax - \gamma I \tag{2.8}$$

Assume (as a simple special case) that the graph is d-regular, then $A\underline{e} = d\underline{e}$ and $\underline{e}^T A = d\underline{e}^T$, where $\underline{e} \in \mathbb{R}^n$ is the first eigenvector of A with largest eigenvalue d.

If all nodes of the network has the same degree and $x_i(0)$ is the same for all i, then $x_i(t)$ remains the same for all t, that is the vector $\underline{e}$ yields an invariant subspace. We can easily obtain the differential equation in this invariant line as follows. Assuming that $x_i = u$ holds for all i, equation (2.2) takes the form

$$\dot{u} = \tau(1 - u)du - \gamma u. \tag{2.9}$$

In the d-regular case, equation (2.8) can be transformed to the simpler form

$$\dot{I} = \tau d\underline{e}^T x - \tau x^T Ax - \gamma I = \tau dI - \gamma I - \tau x^T Ax, \tag{2.10}$$

by using that $\underline{e}^T A = d\underline{e}^T$. However, this form is not the best choice for an approximate lumping (2.6), because the quadratic term can be expressed in terms of I and a remainder term as follows. If $x(t)$ is on the invariant line $\underline{e}$, then $x^T Ax$ can be expressed in terms of I since $x(t) = \frac{I(t)}{n} \cdot \underline{e}$ and

$$x^T Ax = \frac{I^2}{n^2} \underline{e}^T A\underline{e} = \frac{I^2}{n^2} \underline{e}^T d\underline{e} = \frac{I^2}{n^2} nd = \frac{I^2}{n} d.$$

Then using equation (2.10) leads to

$$\dot{I} = \tau dI - \gamma I - \tau \frac{I^2}{n} d = \tau dI \left(1 - \frac{I}{n}\right) - \gamma I,$$

which is identical to (2.9) with $I = nu$.

In the general case when it is not assumed that $x(t)$ is on the invariant line $\underline{e}$, then write (2.10) in the form

$$\dot{I} = \tau dI \left(1 - \frac{I}{n}\right) - \gamma I + \tau \left(I^2 \frac{d}{n} - x^T Ax\right). \tag{2.11}$$

Note that introducing the matrix $E = \underline{e}\underline{e}^T$ with all entries ones, we have $I^2 = x^T Ex$ and the remainder term can be written in the quadratic form as

$$H = I^2 \frac{d}{n} - x^T Ax = x^T \left(\frac{d}{n} E - A\right) x.$$

Numerical experiments show that the approximate lumping can be given by the first two terms of (2.11) and the third term can be considered as the error term. That is let the right hand side of the lumped equation (2.6) be given as

$$g(z) = \tau dz\left(1 - \frac{z}{n}\right) - \gamma z.$$

hence the lumped equation can be written as

$$\dot{\overline{I}} = \tau d\left(1 - \frac{\overline{I}}{n}\right)\overline{I} - \gamma\overline{I}, \tag{2.12}$$

where d is the average degree and $\overline{I}$ is used instead of z since this serves as an approximation of I.

Thus the goal is to derive an estimate for the difference term $|I(t) - \overline{I}(t)|$. We note that this difference can be zero, i.e. $\overline{I}(t)$ can be identical to the prevalence I under special circumstances as it is given in the proposition below.

Proposition 2.1. *If the graph is d-regular (the degree of each node is d) and the initial conditions are identical in each node (i.e. $x_i(0)$ is the same value for all i), then $x_i(t) = u(t)$ for all i, implying $x(t) = \underline{e}I(t)/n$ and hence $I(t) = \overline{I}(t)$ for all t.*

Although this statement is simple, it has limited applicability since the propagation of an epidemic starts from one or a few nodes, not from a homogeneous initial condition. Nevertheless, the low-dimensional approximation $\overline{I}(t)$ will serve as a good approximation for several networks. Our goal is to investigate the accuracy of the approximation in the case of regular graphs with different structure and with arbitrary initial conditions.

2.4 Measuring the accuracy of the low-dimensional approximation

Our goal is to study the accuracy of the one-dimensional approximation (2.12) of the individual-based model (2.2). This will be carried out by comparing the numerical solutions of (2.2) and (2.12), more precisely, by comparing the functions $I(t)$ and $\overline{I}(t)$ defined in (2.3) and (2.12). We are interested in the transient regime, hence we will do the comparison for various regular graphs, for which $I(t)$ and $\overline{I}(t)$ are identical in the steady state. Our goal is to understand how the accuracy, that is the difference $|I(t) - \overline{I}(t)|$, depends on the structure of the network.

First we simply show the graphs of the functions $I(t)$ and $\overline{I}(t)$. Then, to have a quantitative comparison, we compute the integral of their difference numerically. In order to get comparable results for graphs with different sizes and for time evolutions on time intervals with different lengths, we normalize the prevalence values to $[0, 1]$, by using $I(t)/n$ and $\overline{I}(t)/n$ and normalize the length of the time interval, by dividing the integral on the time interval $[0, T]$ by T. Thus we will use the following quantity to measure the accuracy of the approximation

$$\Delta = \frac{1}{nT}\int_0^T \left|I(t) - \overline{I}(t)\right| \mathrm{d}t. \tag{2.13}$$

Before showing numerical comparisons, we need to choose the values of the epidemic parameters, the infection rate τ and the recovery rate γ. Since we focus on the effect of network structure, we fix their values as follows. The recovery rate is fixed as $\gamma = 1$, this basically determines the time scale. The infection rate is set in such a way that the steady state

value of the number of infected nodes is 50% of the population. That is for d-regular graphs $\tau = \frac{2\gamma}{d}$, since the steady state value is $I_e = n/2$ then according to (3.1). In most of the cases we will consider graphs with $n = 400$ nodes. This size will be enough to illustrate how well the low-dimensional approximation performs.

The following network classes are studied

- Random regular graphs with high and low degrees and the complete graph.
- Structured graphs, namely lattices with different degree and with different shape.

3 Steady states and asymptotic behaviour

In this section we present the known results about the asymptotic behaviour of the individual-based model (2.2). The existence and stability of steady states are dealt with first, then the main result about the global behaviour of the system is presented. It turned out that the long-time behaviour is completely determined by the largest eigenvalue, Λ, of the adjacency matrix A and the full characterisation of steady states is available for any type of network.

The local behaviour of (2.2) was studied first in [14], where it was proved that the disease-free steady state, $x = 0$, is stable when $\tau\Lambda < \gamma$. The endemic steady state occurs via transcritical bifurcation as it is shown in the book [5], see Theorem 3.8, which is proved by the method presented in [2]. It is also proved in [5], see Theorems 3.5 and 3.6, that the endemic steady state is unique. The proof is based on the ideas of [6].

The global stability of the steady states is proved in the paper [1], using a Lyapunov function inspired by the well-known result of Lajmanovich and Yorke [6]. We note that the NIMFA model (2.2) is a monotone dynamical system, hence the global stability can also be studied by using the theory of monotone systems [12]. This property implies that periodic orbits cannot occur and trajectories tend to steady states.

The following theorem summarises the results about the steady states and their local and global stability.

Theorem 3.1. *System (2.2) has a disease-free steady state, $x = 0$, which is globally stable when $\tau\Lambda < \gamma$, where Λ is the largest eigenvalue of the adjacency matrix A.*

Transcritical bifurcation occurs in the system when $\tau\Lambda = \gamma$. During this bifurcation, an endemic steady state appears for $\tau\Lambda > \gamma$ and the disease-free steady state loses its stability. The endemic steady state is unique and is globally stable when $\tau\Lambda > \gamma$.

A formula for the endemic steady state in terms of the graph parameters would be desirable. This is available at the moment only for regular graphs, namely, the steady state value of the number of infected nodes, that is the endemic steady state is

$$I_e = n\left(1 - \frac{\gamma}{\tau d}\right), \tag{3.1}$$

where d is the degree of the nodes. We note that for a d-regular graph the largest eigenvalue is $\Lambda = d$. Thus in accordance with the theorem above, the endemic steady state exists if and only if $\tau d > \gamma$.

Further approximations of the endemic steady state close to the epidemic threshold are presented in [9].

4 The approximation accuracy for random regular graphs

First we consider random regular graphs with degree $d = 3$, $d = 10$, $d = 50$ and the complete graph with degree $d = 399$ (note that the number of nodes is $n = 400$). Figure 4.1 shows the time dependence of the ratio of infected nodes for these networks. The diameters of these graphs are 11, 3, 2 and 1, respectively. Regarding accuracy, it is clearly improved as the degree increases and the diameter decreases. The extreme case is the cycle graph (with degree $d = 2$ and diameter 200). Here it is more emphasized that the population-level approximation is significantly different from the individual-based model. This phenomenon can be explained as follows: according to the population-level model, the infection spreads randomly over the network, but according to the individual-based model (which is a more realistic one) the epidemic spreads along the cycle graph as a front. Even in the case of a 3-regular random graph, the disease starts from a single cluster and initially has difficulty in spreading randomly throughout the network. This property – uniform spreading – is well characterised by the graph diameter: the smaller the graph diameter, the more evenly the nodes in the graph are infected.

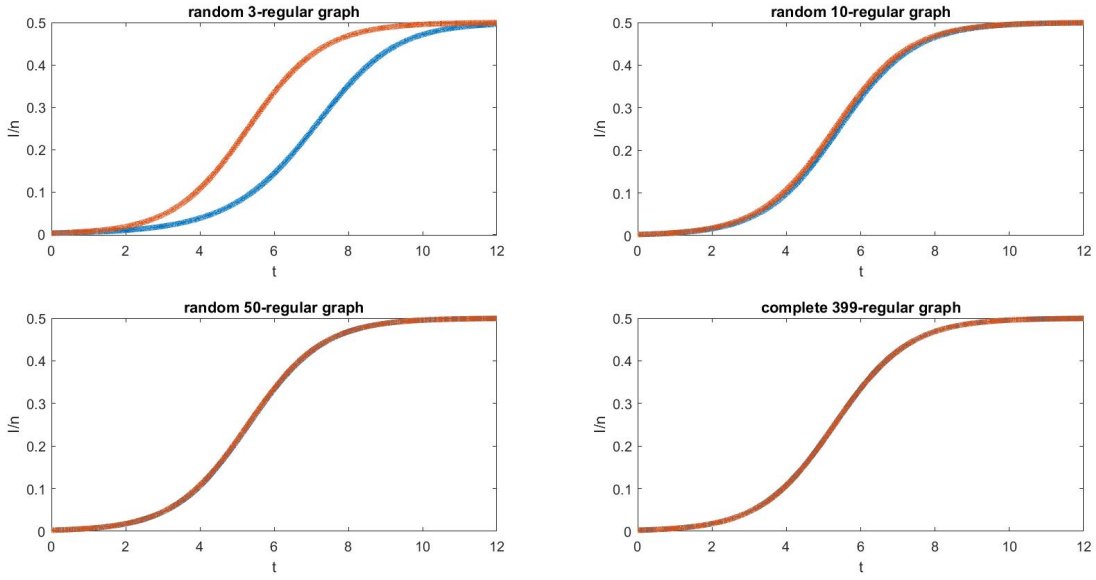

Figure 4.1: The ratio of the infected nodes obtained from the individual-based model, $\frac{I(t)}{n}$ (blue) and from the population-level approximation, $\frac{\bar{I}(t)}{n}$ (red) starting with one infected node. The comparison is shown for regular random graphs with $n = 400$ nodes, degree $d = 3, 10, 50$ and the complete graph with $d = 399$. The epidemic parameters are $\gamma = 1$, $\tau = \frac{2}{d}$. The diameters and mean distances are $(11; 6.76)$; $(3; 1.88)$, $(2; 1.75)$, and $(1; 1)$ respectively.

Now we show results about the relation of the diameter of the graph and the approximation accuracy. As it was mentioned above, the accuracy is measured by the integral Δ given in (2.13). The length of the time interval T is chosen so that the transient behaviour can be fully observed.

As it was shown in Figure 4.1, as the degree gets larger, the diameter gets smaller and the approximation gets better. To illustrate this further, Table 4.1 shows how the accuracy Δ varies as a function of the degree and the diameter. These numerical results reveal that for a degrees

$d \geq 10$ the difference between the two curves is negligible. However, for smaller values of the degree ($d < 10$), the smaller the degree, the larger is the inaccuracy. Note that the table shows the mean distances as well, as it will be seen later that this parameter will also play an important role.

Degree (d)	2	3	10	50	100
Accuracy (Δ)	0.3175	0.0138	0.001	$1.2{\cdot}10^{-4}$	$1.4{\cdot}10^{-5}$
Diameter	200	11	4	3	2
Mean distance	100.25	6.76	2.84	1.88	1.75

Degree (d)	150	198	250	298	399
Accuracy (Δ)	$1.7{\cdot}10^{-5}$	$4.8{\cdot}10^{-6}$	$2.9{\cdot}10^{-6}$	$7.6{\cdot}10^{-6}$	$1.4{\cdot}10^{-5}$
Diameter	2	2	2	2	1
Mean distance	1.62	1.5	1.37	1.25	1

Table 4.1: The effect of the degree, diameter and mean distance on the accuracy Δ for d-regular random graphs. (In the case $d = 2$ it is the cycle graph and for $d = 399$ it is the complete graph.) The epidemic parameters are $\gamma = 1$, $\tau = 2/d$.

5 The approximation accuracy for lattices with low degree

In order to understand better how the structure of the graph affects the accuracy Δ, we take a closer look at graphs with lower degrees. For a fixed degree, the diameter (or mean distance) will have a primary effect on the accuracy. As we have seen that regular graphs with degree larger than 10 yield good accuracy, we will study 4- and 6-regular graphs in detail.

First, we investigate 4-regular graphs with lattice structure. A lattice graph of size $n \times m$ means a graph where a grid of $n \times m$ on a rectangle is taken and the nodes on the opposite sides are connected to each other, thus obtaining a lattice on a two-dimensional torus (see Figure 5.1 for two examples of lattice graphs).

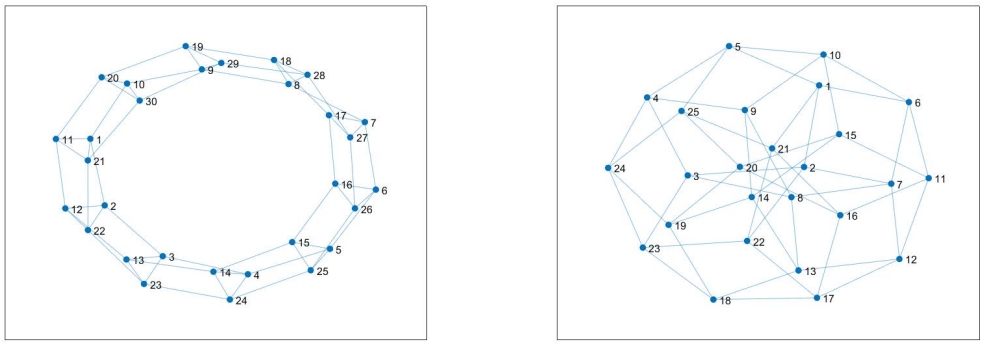

Figure 5.1: Examples of lattice graphs: 3×10 lattice (left), 5×5 lattice (right). The left one is closer to a ring, while the right one resembles a random graph.

In the case of $n = 400$ nodes, one can generate lattice graphs of size 8×50, 10×40 and 16×25. As the rectangle is getting closer to a square, the diameter of the lattice becomes

smaller. The diameter of these graphs are 29, 25 and 20, respectively. These are quite large diameters compared to a 4-regular random graph on $n = 400$ nodes, the diameter of which is only 7. The solution of the individual-based model (2.2) and that of the low-dimensional approximation (2.12) is computed numerically and the prevalences $I(t)$ and $\overline{I}(t)$ are plotted in Figure 5.2. As we can observe, the smaller the graph diameter, the better the approximation. This can be explained by the fact that the sooner the epidemic reaches to an arbitrary node of the network from the initially infected node, the better accuracy is obtained, since the spreading process is more similar to the one modelled by the population-level approximation. (Note that the spreading process on a 8×50 lattice graph is similar to the behaviour observed in the cycle graph because of the large diameter and cycle-like structure of these graphs.)

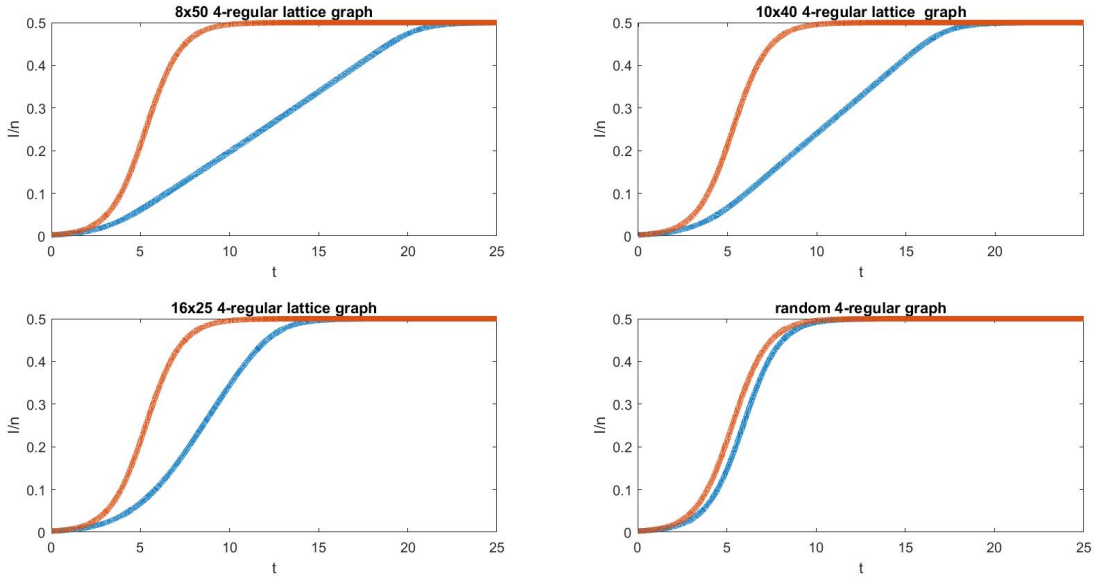

Figure 5.2: The ratio of the infected nodes obtained from the individual-based model, $\frac{I(t)}{n}$ (blue) and from the population-level approximation, $\frac{\overline{I}(t)}{n}$ (red) starting with one infected node on the following 4-regular graphs: 8×50, 10×40 and 16×25 lattice graphs and a random graph. The epidemic parameters are $\gamma = 1$, $\tau = \frac{1}{2}$. The diameters and mean distances are $(29; 14.54)$, $(25; 12.53)$, $(20; 10.27)$ and $(7; 4.8)$, respectively.

Now, we turn to the case of 6-regular graphs and study the effect of the graph structure. A 6-regular lattice is represented by a three-dimensional, $n \times m \times k$ grid on a box and the corresponding nodes on opposite sides are identified, leading to a lattice on a 3-dimensional torus. In the case of $n = 400$ nodes, one can generate 6-regular lattice graphs of size $4 \times 5 \times 20$, $5 \times 5 \times 16$ and $4 \times 10 \times 10$. As the box is getting closer to a cube, the diameter of the lattice becomes smaller. The diameter of these graphs are 14, 12 and 12, respectively. These are quite large diameters compared to a 6-regular random graph on $n = 400$ nodes, the diameter of which is only 6. The two lattices with diameter 12 can be distinguished by their mean distance. This quantity is defined as the average of the lengths of smallest distances between any two nodes of the graph. The mean distances of these graphs are 7.22, 6.42 and 6.02 while that of the random graph is 3.61.

The solution of the individual-based model (2.2) and that of the population-level approximation (2.12) is computed again numerically. The prevalence values $I(t)$ and $\overline{I}(t)$ are plotted

in Figure 5.3. As we can observe, the smaller the mean distance, the better the approximation.

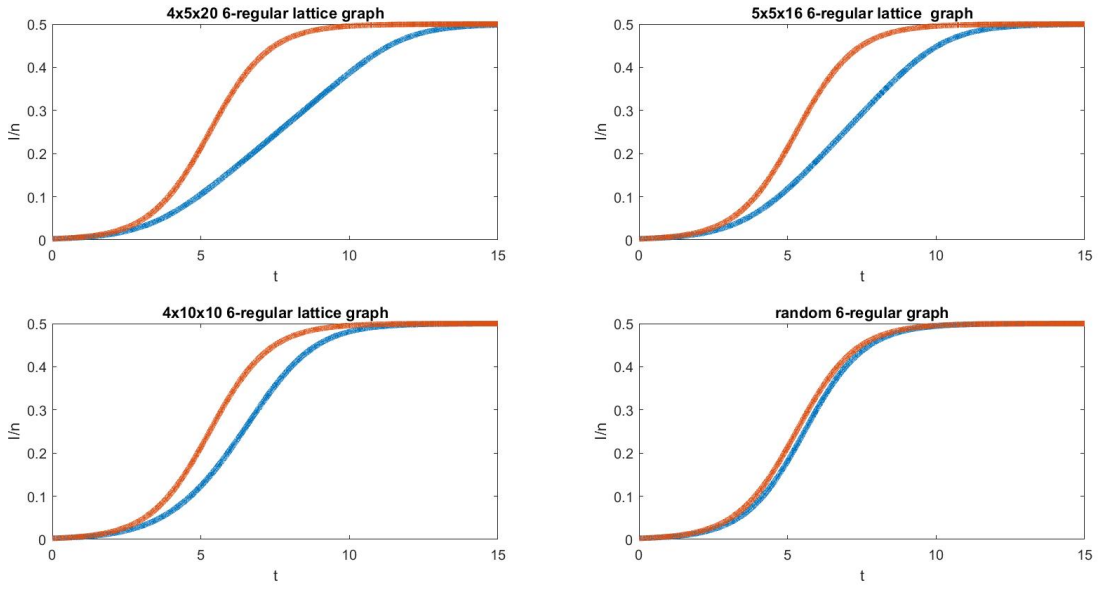

Figure 5.3: The ratio of the infected nodes obtained from the individual-based model, $\frac{I(t)}{n}$ (blue) and from the population-level approximation, $\frac{\bar{I}(t)}{n}$ (red) starting with one infected node on the following 6-regular graphs: $4 \times 5 \times 20$, $5 \times 5 \times 16$, $4 \times 10 \times 10$ lattice graphs and a random graph. The epidemic parameters are $\gamma = 1$, $\tau = \frac{1}{3}$. The diameters and mean distances are $(14; 7.22)$, $(12; 6.42)$, $(12; 6.02)$ and $(6; 3.61)$, respectively.

Figure 5.3 suggests that the accuracy depends on the mean distance of the graph. To reveal this relation we computed the accuracy for several values of the mean distance. In order to create graphs with different mean distance we started from a $4 \times 5 \times 20$ lattice graph, and then reconnect the nodes randomly while keeping the regularity of the graph, i.e. the degree of each node remains 6. The reconnection is performed by eliminating two arbitrarily selected edges starting from different nodes in the mesh and establishing two new edges between the endpoints of the deleted edges, taking into account that no multiple edges are created. Thus, the new graphs become closer to a random one and their mean distance will decrease as the rewiring process is progressing by choosing newer edges to delete and create. Figure 5.4 shows the accuracy Δ as a function of the mean distance. It can be observed that the smaller the mean distance of a graph, the more accurate the population-level approximation.

Concluding this section, one can make the following observations concerning the approximation of the prevalence $I(t)$ obtained from system (2.2). This system contains as many equations as the number of nodes in the network. Our goal is to use a low dimensional (now, in fact, a one-dimensional) approximation instead, and to derive an estimate for the accuracy of the approximation. The one-dimensional approximation is (2.12) yielding $\bar{I}(t)$. Our computational findings in this section revealed the following.

- If the degree of the nodes is not small, in the case of graph size $n = 400$, not smaller than 10, then the difference of $I(t)$ and $\bar{I}(t)$ is negligible. Thus we can simply use the one-dimensional approximation instead of the large system with n equations when the graph is dense enough.

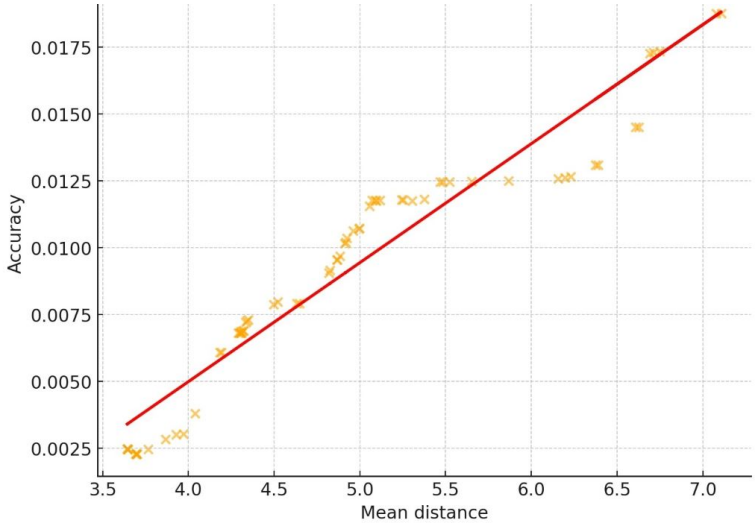

Figure 5.4: The accuracy Δ as a function of the mean distance on various 6-regular graphs that are obtained by rewiring a $4 \times 5 \times 20$ lattice graph.

- In the case of sparse graphs, the one-dimensional approximation should be amended by a remainder-like error term. This remainder term is determined basically by the mean distance of the graph. Thus we can again simply use the one-dimensional approximation instead of the large system with n equations and then add the remainder, the size of which can be computed from the mean distance. We note that an extension of the mean-field equation to a system with two equations is desirable, especially if the diameter could be incorporated into the differential equations.

6 Discussion

In this paper, we studied the individual-based model of SIS epidemic propagation on regular graphs, exploiting the fact that this system of differential equations explicitly includes the adjacency matrix of the network. This enabled us to study the effect of network structure on the time dependence of the prevalence. A one-dimensional approximation, population-level equation, was introduced. For many networks, this single differential equation yields an excellent approximation of the prevalence, despite of the fact, that exact value of the prevalence is determined by a system of n differential equations, where n, the number of nodes in the network, can be very large. It turned out that for dense random regular graphs the approximation is very good, while for sparse regular graphs a significant correction term is needed to obtain the proper time dependence of the prevalence. Thus by solving only the low-dimensional model and adding a correction term, the prevalence can be obtained with high accuracy. Numerical evidence is shown that this correction term depends linearly on the mean distance in the network.

The investigations of this paper can be extended to several directions. A two-dimensional approximation could be derived, yielding not only the prevalence but also the correction term. It would be desirable to include the mean distance as a parameter in this system, since numerical experiments show that the correction term depends linearly on this graph characteristic.

All the investigations in this paper correspond to regular graphs, hence an obvious de-

mand arises for having similar result in the cases of graphs with a more complex degree distribution. As a first step in this direction we show some numerical results for bimodal graphs that have two different degrees. The nodes are divided into two groups, in the first one there are n_1 nodes with degree d_1 and in the second group there are n_2 nodes of degree d_2. We note that this degree distribution can be realized with several different adjacency matrices, thus further properties of the graphs will certainly have effect on the performance of the approximation. The mean-field equation is again (2.12), with $n = n_1 + n_2$ and average degree

$$d = \frac{n_1 d_1 + n_2 d_2}{n_1 + n_2}.$$

Four networks were created with different values of n_1, n_2, d_1, d_2 to show the rich dynamical behaviour that one can obtain for bimodal graphs, see Figure 6.1. We note that the infection rate τ is chosen again in such a way that the endemic steady state in the population-level equation is half of the whole population. In the cases shown in the upper part of the figure, the population-level equation hits at least the endemic steady state, while in the cases belonging to the bottom figures neither the transient nor the asymptotic behaviour is captured by the one-dimensional approximation. This leads to the claim that two separate differential equations are needed for the two groups of nodes with different degrees. The extension of the results from regular to bimodal graphs is planned as a second part of this paper. The mean distances in the above examples are close to each other, highlighting that further graph characteristics will be required to capture the dynamics in the case of bimodal graphs.

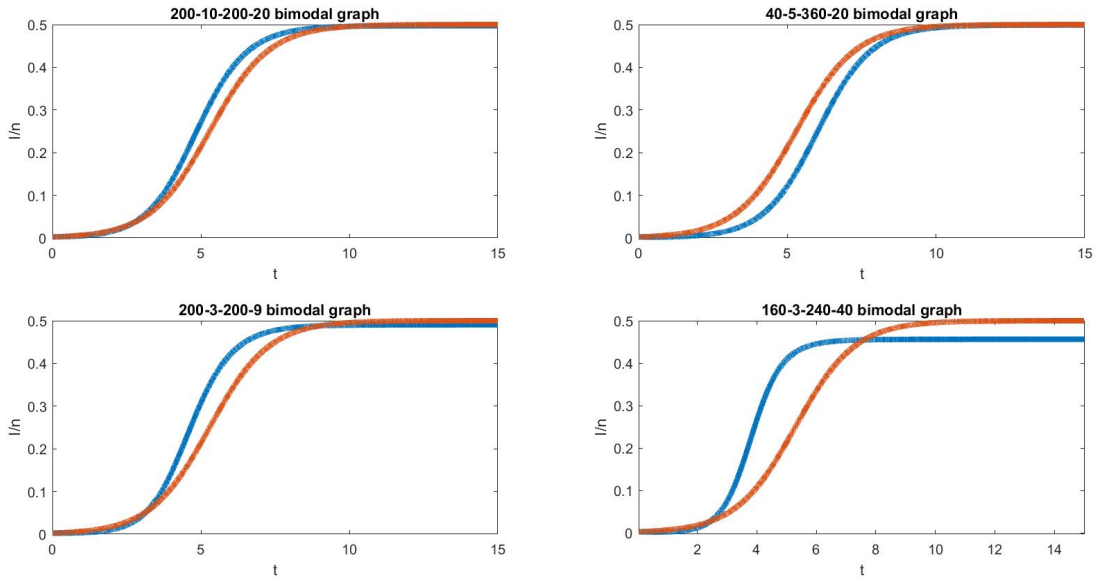

Figure 6.1: The expected value of the number of infected nodes obtained from the individual-based model, $I(t)$ (blue) and from the population-level approximation, $\bar{I}(t)$ (red) starting with one infected node on the following bimodal graphs: $n_1 = 200$, $d_1 = 10$, $n_2 = 200$, $d_2 = 20$ (top left); $n_1 = 40$, $d_1 = 5$, $n_2 = 360$, $d_2 = 20$ (top right); $n_1 = 200$, $d_1 = 3$, $n_2 = 200$, $d_2 = 9$ (bottom left) and $n_1 = 160$, $d_1 = 3$, $n_2 = 240$, $d_2 = 40$ (bottom right). The epidemic parameters are $\gamma = 1$, $\tau = \frac{2}{15}; \frac{2}{37}; \frac{1}{3}; \frac{5}{252}$ respectively. The diameters and mean distances are $(4; 2.5); (4; 2.36); (6; 3.52); (4; 2.42)$, respectively.

Acknowledgements

The study was funded by the National Research, Development and Innovation Office in Hungary (RRF-2.3.1-21-2022-00006).

P. L. Simon acknowledges support from the Hungarian Scientific Research Fund, OTKA (grant no. 135241).

References

[1] Á. Bodó, P. L. Simon, Transcritical bifurcation yielding global stability for network processes. *Nonlinear Anal.* **196**(2020), 111808. https://doi.org/10.1016/j.na.2020.111808

[2] C. Castillo-Chavez, B. Song, Dynamical models of tuberculosis and their applications. *Math. Biosci. Eng.* **1**(2004), No. 2, 361–404. https://doi.org/10.3934/mbe.2004.1.361

[3] X. Fu, M. Small, G. Chen, *Propagation dynamics on complex networks: Models, methods and stability analysis*, Wiley, Chichester/UK, 2013. https://doi.org/10.1002/9781118762783

[4] T. House, M. J. Keeling, Insights from unifying modern approximations to infections on networks, *J. Roy. Soc. Interface* **8**(2011), 67–73. https://doi.org/10.1098/rsif.2010.0179

[5] I. Z. Kiss, J. C. Miller, P. L. Simon, *Mathematics of epidemics on networks. From exact to approximate models*, Springer, Cham, 2017. https://doi.org/10.1007/978-3-319-50806-1

[6] A. Lajmanovich, J. A. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population, *Math. Biosci.* **28**(1976), 221–236. https://doi.org/10.1016/0025-5564(76)90125-5

[7] R. Pastor-Satorras, A. Vespignani, Epidemic dynamics and endemic states in complex networks, *Phys. Rev. E* **63**(2001), https://doi.org/066117.10.1103/PhysRevE.63.066117

[8] M. Porter, J. Gleeson, *Dynamical systems on networks. A tutorial*, Springer, Cham, 2016. https://doi.org/10.1007/978-3-319-23105-1

[9] B. Prasse, P. Van Mieghem, Time-dependent solution of the NIMFA equations around the epidemic threshold, *J. Math. Biol.* **81**(2020), No. 6, 1299–1355. https://doi.org/10.1007/s00285-020-01542-6

[10] B. Prasse, P. Van Mieghem, Predicting network dynamics without requiring the knowledge of the interaction graph, *Proc. Natl. Acad. Sci. USA* **119**(2022), No. 44, e2205517119. https://doi.org/10.1073/pnas.2205517119

[11] K. J. Sharkey, Deterministic epidemic models on contact networks: Correlations and unbiological terms, *Theor. Popul. Biol.* **79**2011), No. 4, 115–129. https://doi.org/10.1016/j.tpb.2011.01.004

[12] H. L. Smith, *Monotone dynamical systems. An introduction to the theory of competitive and cooperative systems*, American Mathematical Society, Providence, 2008. https://doi.org/10.1090/surv/041

[13] P. Van Mieghem, The N-intertwined SIS epidemic network model, *Computing* **93**(2011), 147–169. https://doi.org/10.1007/s00607-011-0155-y

[14] P. Van Mieghem, J. Omic, R. Kooij, Virus spread in networks, *Networking, IEEE/ACM Transactions* **17**(2009), No. 1, 1–14. https://doi.org/10.1109/TNET.2008.925623