PDE limits of stochastic SIS epidemics on networks

Article (Accepted Version)

PDE-limits of stochastic SIS epidemics on networks

F. Di Lauro[1], J.-C. Croix[1], L. Berthouze[2], I.Z. Kiss[1]

[1] Department of Mathematics, University of Sussex, Falmer, BN1 9QH, UK
[2] Department of Informatics, University of Sussex, Falmer BN1 9QH, UK

2nd October 2020

Abstract

Stochastic epidemic models on networks are inherently high-dimensional and the resulting exact models are intractable numerically even for modest network sizes. Mean-field models provide an alternative but can only capture average quantities, thus offering little or no information about variability in the outcome of the exact process.

In this paper we conjecture and numerically demonstrate that it is possible to construct PDE-limits of the exact stochastic SIS epidemics on Regular, Erdős-Rényi, Barabási-Albert networks and lattices.

To do this we first approximate the exact stochastic process at population level by a Birth-and-Death process (BD) (with a state space of $O(N)$ rather than $O(2^N)$) whose coefficients are determined numerically from Gillespie simulations of the exact epidemic on explicit networks. We numerically demonstrate that the coefficients of the resulting BD process are density-dependent, a crucial condition for the existence of a PDE limit.

Extensive numerical tests for Regular, Erdős-Rényi, Barabási-Albert networks and lattices show excellent agreement between the outcome of simulations and the numerical solution of the Fokker-Planck equations. Apart from a significant reduction in dimensionality, the PDE also provides the means to derive the epidemic outbreak threshold linking network and disease dynamics parameters, albeit in an implicit way. Perhaps more importantly, it enables the formulation and numerical evaluation of likelihoods for epidemic and network inference as illustrated in a fully worked out example.

Keywords: Epidemics, Networks, Inference, PDEs, Fokker-Planck

1 Introduction

An epidemic is a complex phenomenon that arises from a pathogen spreading over the contact structure of a population. Similar spreading phenomena occur in various disciplines, from biology and social sciences to engineering. Unsurprisingly, much modelling effort has been put into studying spreading processes on networks, as they offer a natural framework to mimic real-life contact patterns [3] and the important heterogeneities within these. The use of networks is extremely intuitive with each individual encoded as a node, and all its contacts (to other individuals) as links. Unfortunately, the resulting exact probabilistic model does not scale well with the size of the network, N. Even when relatively simple models, such as susceptible-infected (SI) or susceptible-infected-susceptible (SIS), are considered, the exact model has 2^N equations, which quickly becomes intractable.

To address the difficulty posed by high dimensionality, mathematical descriptions often focus on population-level statistics (e.g. expected number of infected people at any given time). This has led to a number of so-called mean-field models [22, 33, 32], offering a good first approximation of the evolution of some population-level or averaged quantities. These include pairwise models based on moment closure techniques [21, 22], effective-degree [25], edge-based-compartmental [28] models and even PDE models [37]. All such mean-field models share a number of caveats [36]. For example, (i) in general the agreement between these and the exact stochastic model breaks down close to the epidemic threshold, (ii) there are very few cases where it is possible to prove mathematically that the mean-field model is the limit of the exact stochastic process (this has only been done for SIR epidemics and configuration networks [8, 19]) and (iii) they give no estimate of the variability observed in the exact process. It is also well-known that such mean-field models only work for a limited class of networks; epidemics on clustered networks are not well-understood, except for idealised clustered networks, i.e. networks with non-overlapping triangles or other clustering-inducing subgraphs.

There are ongoing efforts to try to understand and answer rigorous mathematical questions when it comes to analyse or approximate dynamical processes on networks, see [34] for a recent summary. Progress in this area is usually achieved by bringing in and combining results and techniques from different areas of mathematics. One particularly promising prospect for SIS epidemics on networks is to consider them as Birth-and-Death (BD) processes. In a recent paper [9], we conjectured and confirmed numerically that SIS epidemics are well captured by BD processes, whose rates encode characteristics of both the network structure and the epidemic dynamics. This was tested on Regular, Erdős-Rényi and Barabási-Albert networks. This choice was motivated by the intuition that epidemic spread is driven by the 'birth' of new infected nodes. However, this occurs at a rate which is proportional to the number of S-I (active) links, and these are readily observable during explicit stochastic simulations of the epidemic on networks.

In this paper we build on the above observation and take the next natural step, that is, to consider the large N limit of the BD process, i.e., the one-dimensional PDE (Fokker-Planck equation). We extend the repertoire of network models and consider Regular, Erdős-Rényi and Barabási-Albert networks and 2D lattices with periodic boundary conditions and show that the resulting rates in the BD process are density-dependent such that the limit is well defined in the sense of [20]. We compute the rates numerically and also provide a parametric form for them (with the exception of the lattice). We show that the resulting PDE agrees well with the output of explicit simulations of stochastic epidemics on networks. The existence of the PDE limit has multiple advantages. First, it reduces further the dimensionality of the system. Second, it gives us the opportunity to compute an epidemic threshold even in an implicit form. Finally, it provides the means to get a handle on the variability of the stochastic process with the solution of the PDE providing a likelihood that can computed cheaply and efficiently for inference purposes. Finally, the good agreement between the PDE and the exact process provides further evidence that the BD model may indeed serve as a valid approximation of the exact process (the relation between the exact, BD and PDE-limit model is illustrated in figure 1) and that a formal proof of this observation may be possible.

This paper is structured as follows. In Section 2 we briefly outline the Birth-and-Death approximation of SIS epidemics as in [9]. In Section 3 we numerically test and prove that the conditions for the existence of the PDE limit, as $N \to \infty$, are met for different network topologies and epidemic parameters. We then show that the solutions of the partial differential equations agree well with the empirical distributions based on simulations of the true process. In Section 4 we draw some conclusions and outline further research directions.

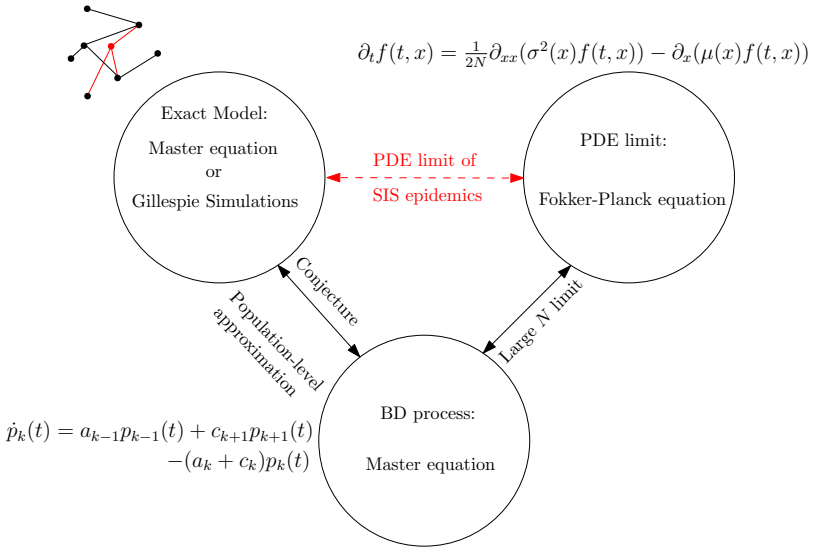

Figure 1: Schematic illustration of various approximations of the exact stochastic SIS epidemics on networks. The PDE limit comes as a result, and further confirms the validity, of the Birth-Death approximation conjectured in [9].

2 Methods

2.1 Birth-and-Death Approximation of SIS Epidemics

We briefly describe the model proposed in [9] which conjectured that exact stochastic SIS epidemics on networks can be approximated by BD processes. A standard SIS model on an undirected unweighted network $\mathcal{G}$ with N nodes is considered, where each node is either susceptible (S) or infected (I). Infected nodes spread infection to their neighbours with constant per-link rate τ and recover with rate γ (independently of the network). This stochastic process results in a continuous time Markov chain on a state space of cardinality 2^N, which forbids analysis even for relatively small values of N. Instead, we consider the population-level count of infected nodes, defined as $k(t) = \sum_{i=1}^{N} \mathcal{I}_i(t)$, where $\mathcal{I}_i$ is an indicator function equal to 1 if node i is infected at time t and 0 otherwise. $k(t) \in [0, N]$, where $k(t) = 0$ indicates the state where no infection is present in the network. Given the stochasticity of the process, $k(t)$ is itself a random variable taking values on state space of cardinality $(N + 1)$. This reduction in dimensionality makes computations much more tractable. We note that each time an infection/recovery occurs, the value of $k(t)$ changes by discrete jumps of size ± 1, respectively. This has led to the conjecture [9] that the population-level count process can be approximated by a Birth-and-Death process, governed by the following master equation:

$$\dot{p}_k(t) = a_{k-1}p_{k-1}(t) + c_{k+1}p_{k+1}(t) - (a_k + c_k)p_k(t), \tag{1}$$

where $p_k(t)$ is the probability of having k infected nodes at time t, $c_k = \gamma k$ is the global recovery rate when k nodes are infected, and a_k is the rate at which the population goes from k to $k + 1$ infected individuals.

The approximation is exact in the case of complete or fully connected networks, where the a_k rates are given by the expression $a_k = \tau k(N - k)$. In the general case, the a_k's are random variables themselves, since the rates at which infections happen are the product of τ times the total number of $S - I$ links in the network, a random variable itself that reflects the topology of the network and the way in which the epidemic positions the k infected nodes on the network. This means that the

epidemic at population-level is not Markovian, making an exact treatment much more difficult and still out of reach.

However, by using the master equation, we can recast this process as a Markovian one using a suitable approximation of each rate a_k. A natural proposal is a quantity that captures the average rate of infection, weighted by the time spent in the observed states, that is:

$$a_k = \frac{\tau \sum_{\xi_k} \xi_k t_{\xi_k}}{\sum_{\xi_k} t_{\xi_k}}, \tag{2}$$

where ξ_k are the observed counts of the number of S-I links on the network when k infected nodes are present and and t_{ξ_k} is the lifetime of this particular state. This quantity is responsible for driving the epidemic: The higher the number of S-I links, the larger the rate of generating more infected nodes. The a_k's can be obtained by averaging over many realisations of the epidemic on different realisations of networks from the same family. This can also be interpreted as averaging out stochasticity at link-level and transferring it to population-level. Hence, the variability in epidemic paths will be due to the stochasticity of the master equation itself, guaranteeing the Markov property of the Birth-Death process. The solution of equation (1) with these proposed rates has been shown to be in excellent agreement with the average from many simulations for various network models and epidemic parameters [9].

2.2 Fokker Planck equation as a limit of the Birth-Death process

Master equation (1) can be used as a starting point to build its continuous (in space) limit, i.e., the Fokker-Planck equation [14, 22]. The idea is to approximate the solution $p_k(t)$ by considering it as a discretisation of a continuous function $f(t, x)$ in the interval $[0, 1]$, defined as

$$f\left(t, x = \frac{k}{N}\right) = p_k(t).$$

For the large N limit to exist, it is known [22, 24, 11, 31, 4] that the rates of the master equation need to satisfy the following density-dependence condition (with a slight abuse of notation):

$$\frac{a(k)}{N} = A\left(\frac{k}{N}\right), \qquad \frac{c(k)}{N} = C\left(\frac{k}{N}\right), \tag{3}$$

where A and C are not necessarily the same functions as a and c. It is worth noting that condition (3) is not guaranteed to hold for every network model, and must therefore be validated on networks of interest.

In the density-dependent case, it can be shown [22, 31, 4] that $f(t, x)$ satisfies the following forward Fokker-Planck equation:

$$\partial_t f(t,x) = \frac{1}{2N}\partial_{xx}\big(\sigma^2(x) f(t,x)\big) - \partial_x\big(\mu(x) f(t,x)\big), \tag{4}$$

with initial condition $f(0, x) = \delta(x - x_0)$, where the diffusion coefficient $\sigma^2(x)$ and the drift $\mu(x)$ are related to the a_k and c_k rates via:

$$\begin{aligned} \sigma^2(x) &= \frac{1}{N}(A(x) + C(x)), \\ \mu(x) &= A(x) - C(x). \end{aligned} \tag{5}$$

Boundary conditions are naturally emerging from two considerations: (1) if the process hits $k = 0$ at some time (disease-free state) it will stay there forever, and (2) the number of infected nodes cannot be greater than N at any given time. In this framework, such physical constraints translate naturally into Dirichlet and Robin boundary conditions:

$$\begin{cases} f(t, x = 0) = 0, & \text{absorption in } x = 0, \\ \frac{1}{2}\partial_x(\sigma^2(x)f(t,x))|_{x=1} - (\mu(1)f(t,1) = 0, & \text{reflection in } x = 1. \end{cases}$$

Fokker-Planck equations of this kind have been extensively studied numerically, especially in the biological context of population random genetic drifts [10, 6, 2, 5], as well as analytically [13, 38, 23]. In particular, in [23], this equation is studied in the limit of large t to characterise the so-called quasi-steady state [27, 7] (where the only steady state possible is absorption at 0), whereas in [5, 6] various numerical schemes to solve such equations are employed and compared in terms of numerical instabilities and performance. In Appendix 6 we describe our numerical scheme of choice, which is an adaptation of a finite volume method (FVM) scheme already described in [6].

Networks	$\langle k \rangle$	τ	γ	R_0
Regular	9	1	6	1.28
	7	2.5	8	1.66
	8	3.5	7	2.65
Erdős-Rényi	8	1	5	1.50
	10	1	4.5	2.44
	7	4	7	2.90
Barabási-Albert	10	0.9	3.5	2.12
	4	2	5	3.72
	18	0.55	6.2	5.38
2D lattice	4	1.2	2	1.12
	4	2	2	1.5
	4	8	2	2.4

Table 1: Values of network and epidemic parameters for the benchmark scenarios chosen to test the PDE limit of large networks. R_0 has been computed on networks of size $N = 10000$ via the formula $R_0 = \frac{\langle k^2 - k \rangle}{\langle k \rangle} \frac{\tau}{\tau + \gamma}$, as described in [22].

3 Results

3.1 Validation of the density dependence condition

In order to use eq. (4) we need to verify that the rates of the BD process satisfy condition (3). Recoveries are independent of the network, therefore, the condition is automatically satisfied as the expression for these rates is $c_k = \gamma k$. Infection rates, instead, need to be inspected more closely, as their values are dependent on the topology of the network. As an example, even fully-connected networks with $a_k = \tau k(N - k)$ violate condition (3). This can be corrected by requiring that τ scales as $\tau/N = ct$ in the limit of large N. This case is well-known in the literature [17, 1], albeit in a SDE formulation, so we limit our treatment of it to reporting the exact Fokker-Planck equation for the fully connected network, i.e.

$$\partial_t f(t,x) = \frac{1}{2N}\partial_{xx}\left[(\beta x(1-x) + \gamma x)\, f(t,x)\right] - \partial_x\Big[\,(\beta x(1-x) - \gamma x)\, f(t,x)\Big],$$

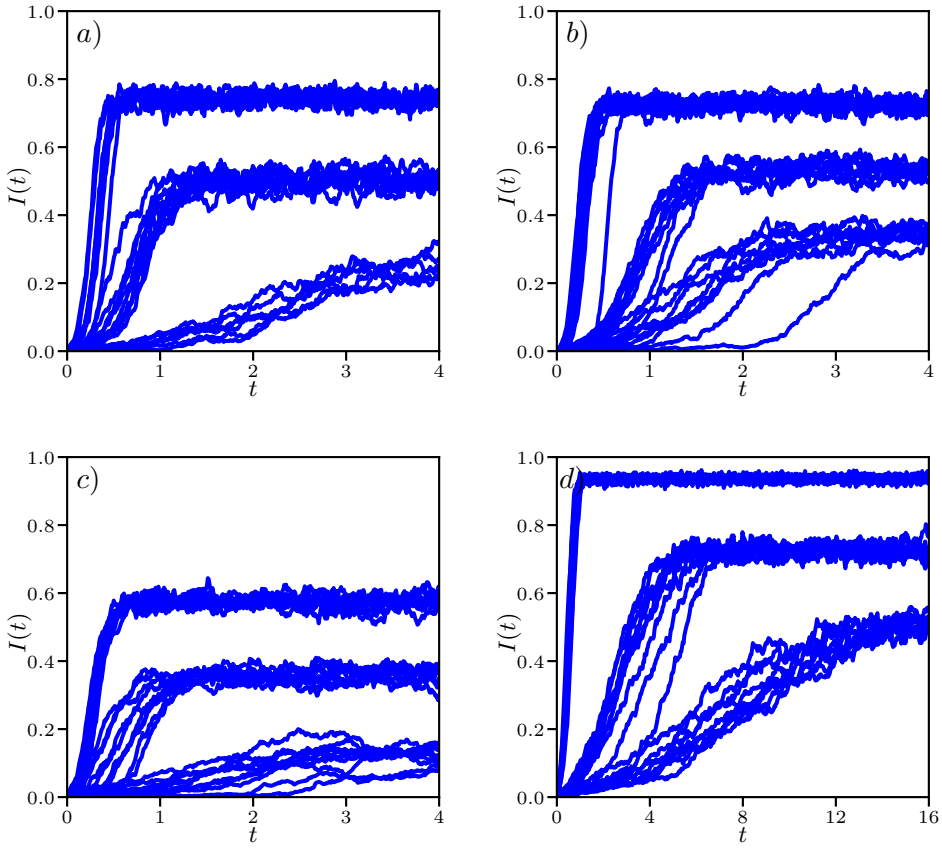

Figure 2: Typical realisations of SIS epidemics on (a) Regular, (b) Erdős-Rényi, (c) Barabási-Albert and (d) 2D lattice networks, for the parameter values shown in Table 1 and with $N = 1000$ (for the lattice, this number is 1024). In each panel 10 realisations of the epidemics are plotted,. The parameters used to generate such networks are also reported in Table [1], higher prevalence corresponds to higher values of R_0.

where $\beta = \frac{\tau}{N^2}$.

Since degree heterogeneity and higher-order stucture in networks have a marked effect on epidemics we explore Regular, Erős-Rényi and Barabási-Albert networks and lattices. To test the scaling hypothesis, the infection rates, based on eq. (2), are computed on different networks and for different values of N (typically from $N = 10^2$ to $N = 10^5$ with slight variations for lattices). The resulting (k, a_k) curves are plotted in figures 3 (Regular) and 4 (Erdős-Rényi) (see also figures 11 and 12 corresponding to Barabási-Albert networks and lattices, respectively, in Appendix 7). Using eq. (3), these rates are rescaled and plotted again in the same figures confirming that they define a universal rate.

In figure 5, the universal curves (based on the highest N explored) are compared for the four different network types in order to highlight how the topology of the network impacts the shape of the (k, a_k) curves. As expected on a lattice, the a_k's grow linearly with k. On all other networks, the curves are parabola-like but higher degree heterogeneity leads to a more pronounced left skew in the location of the maximum point of the rate curve. This is because nodes with many links are likely to be infected early on in the epidemic, meaning that even for low k values a_k can be high (if a hub is infected then this leads to many new S-I links) compared to networks with milder degree heterogeneity. Once most of the highly connected nodes are infected (typically only a small

proportion of N), the epidemic will unfold on the less well connected nodes, therefore fewer links, and thus the parabola decreases early on as shown by the left skew. We note that some variability between the scaled (k, a_k) curves emerges and this is likely due to finite size effects where stochastic variability is accentuated. However, the difference is so small that the Fokker-Planck equation and its solution appear insensitive to the exact choice of the universal rates.

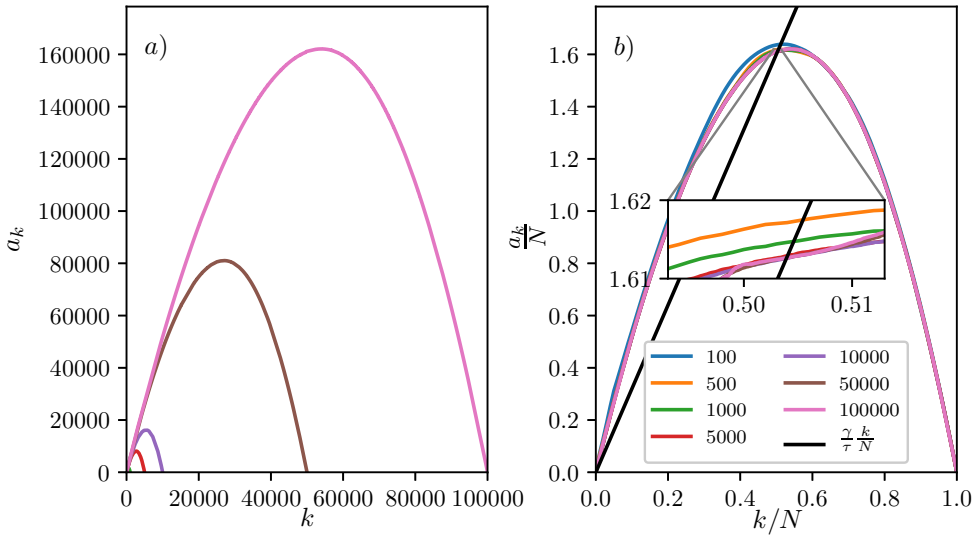

Figure 3: Scaling for regular networks using parameters given in the first row of Table (1). (a) Unscaled (k, a_k) curves for values of N ranging from $N = 100$ to $N = 100000$. Each curve is obtained by simulating 10000 realisations of the epidemic across 50 realisations of the network, half of the epidemics starting from $k_0 = 1$, the other half from $k_0 = N$. (b) Corresponding scaled rate $(k, \frac{a_k}{N})$ curves. The scaling hypothesis can be checked by noticing that the higher the values of N, the closer the scaled curves get to the limiting universal curve. As N increases, the differences between scaled rates decrease. In the inset, the small mismatch between curves with $N \geq 1000$ are highlighted using a 30x zoom. For completeness, the $(k/N, \gamma k/N)$ curve is provided (in black); it intercepts the scaled curves around the steady state.

3.2 Comparing PDE and simulations

Since the limit of large N is of interest, it is beneficial to have a continuous function that fits the discrete a_k rates (2). In [9], the following three-parameter model was proposed:

$$a_k^{(C,\alpha,p)} = Ck^p(N-k)^p\left(\alpha\left(k - \frac{N}{2}\right) + N\right). \tag{6}$$

This model can be fitted to the a_k curves via a least-square approach, by minimizing the following cost:

$$e = \sum_{k,\sum t_{\xi_k}>0} \left(a_k^{(C,\alpha,p)} - a_k\right)^2. \tag{7}$$

In [9], we showed that this approach leads to good agreement with simulations from different network classes, in particular, Regular and Erdős-Rényi networks. The fit for Barabási-Albert networks is acceptable and the fit breaks down for lattices.

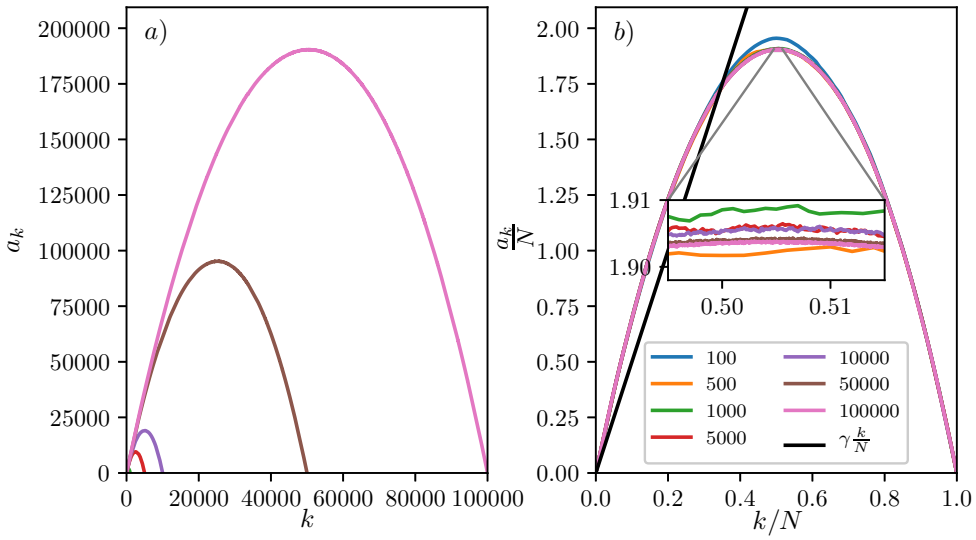

Figure 4: Same scenario as in figure 3 but for Erdős-Rényi networks using parameter values from the sixth row of Table (1). (a) Unscaled (k, a_k) curves for values of N ranging from $N = 100$ to $N = 100000$. Each curve is obtained by simulating 10000 realisations of the epidemic across 50 realisations of the network; half of the epidemics starting from $k_0 = 1$, the other half from $k_0 = N$. (b) Corresponding scaled $(k, \frac{a_k}{N})$ curves.

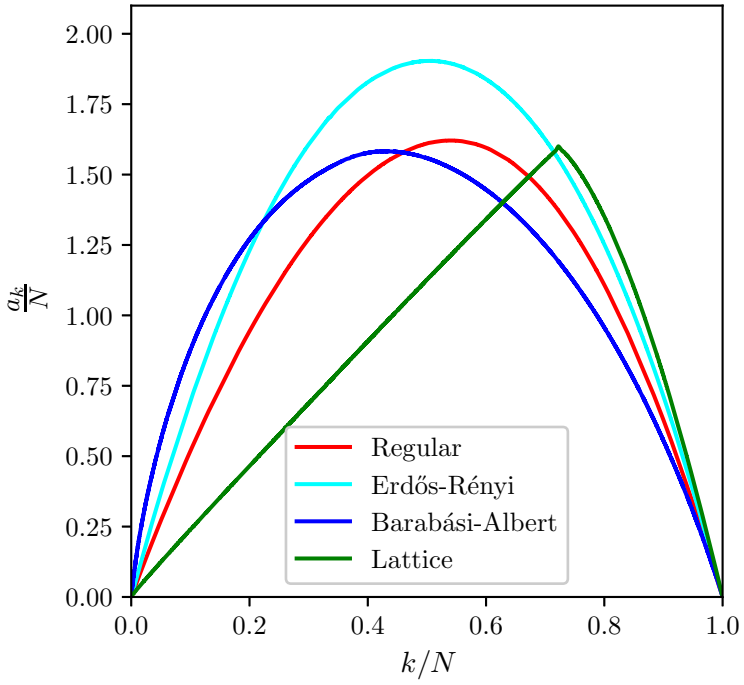

Figure 5: Comparison of different scaled (k, a_k) curves produced by different network models for large N. The (k, a_k) curves are scaled by a factor N. The parameters to generate data for each curve are provided in the second, the fourth, the eighth and the last rows of table 1, respectively. The lattice network is of size 316×316, and the other networks have $N = 10^5$.

In the following, we make use of this function to model the infection rates of master equation (1). However, using a simple function to model the complexity of the a_k rates adds an additional layer of approximation to our approach. Therefore, in addition to eq. (6) we also consider a cubic spline of the a_k rates, as it provides an even better fit to the rates based on eq. (2) and therefore yields better results. This is particularly apparent for lattices, where the (C, α, p) model fails for obvious reasons, and to some extent for Barabási-Albert networks. To summarise, the rates of infection are first found based on simulations via eq. (2). As this approach produces a discrete function that cannot be used as is in the Fokker-Planck equation, we propose two alternatives: (a) the (C, α, p) model, eq. (6), and (b) a spline. The PDE is considered with both rates (except for lattices), and the numerical solution of the PDE is computed via a Finite Volume Method (several other numerical schemes [29, 5, 6] were tested) as it guarantees that the solution remains non-negative and preserves mass, see Appendix.

To show the agreement between the Fokker-Planck equation (4) and results from simulations on networks, we selected twelve (three for each network model) combinations of network and epidemic parameters, as described in Table 1. We tuned the parameters such that for each family we could get epidemics with different characteristics, i.e., different transient and quasi-steady state. To show this, in figure (2) we illustrate a few realisations of epidemics on networks of size $N = 1000$ for each scenario (for the 2D-lattice network, the size is $N = 1024$). We also report the computation of R_0 as described in [22], see Table 1. Note that the initial condition for the PDE is always taken as in the simulations, so $k_0 = 1$. In the simulations, at every run of the epidemic we select a node at random to be the initially infected one. This, however, does not prevent setting initial conditions with a higher number of infected nodes. The initial condition in such cases should be based on measurements taken from the simulations.

Parameters were chosen so that, for each family, the three quasi-steady states showed different levels of prevalence. To find the (k, a_k) curves via minimisation of (7), we generated data as follows: for each scenario, we created 50 realisations of the network, and on each we ran 200 realisations of the epidemic, half of which started from $k_0 = 1$, the other half from $k_0 = N$. This was done in order to obtain observations over the whole range of possible values of the infected nodes. Indeed, when epidemics start from low k_0 values, they only very rarely reach a prevalence much higher than the quasi-steady-state.

The numerical solutions of equation (4) are compared with results based on Gillespie simulations [15, 16], see figures 6, 7, 8 and 9. Excellent agreement holds for all scenarios we tested, as long as the size of the network is ≥ 1000. For small networks, there is a finite-size effect that does not allow for as good a fit. Interestingly enough, although there are small differences between different a_k curves, as long as $N \geq 1000$, the exact choice of N has little impact on the numerical solutions of the PDE. This supports our conjecture that there is indeed a large N limit and, therefore, a universal scaled a_k curve which is approached as N increases. As can be seen, the spline consistently leads to a better approximation. This is simply due to a tighter fit to the discrete data compared to the fit based on the (C, α, p) model. We note, however, that the (C, α, p) model captures the trend of the epidemic and the quasi-steady state is fitted well. Of course, in the case of the lattice we only use the spline as the (C, α, p) model cannot capture the linear rise.

To realise the comparisons provided in figures 6, 7, 8 and 9 we proceeded as follows. We considered the same network realisations and epidemic parameters used to find the (k, a_k) rates. We fixed the initial condition to be $k_0 = 1$ and ran 200 simulations on each realisation. Each individual path was then sampled at regular times in order to build the empirical distribution $p(x, t)$. Note that all simulations were kept, even those that died out early. This is because the numerical scheme

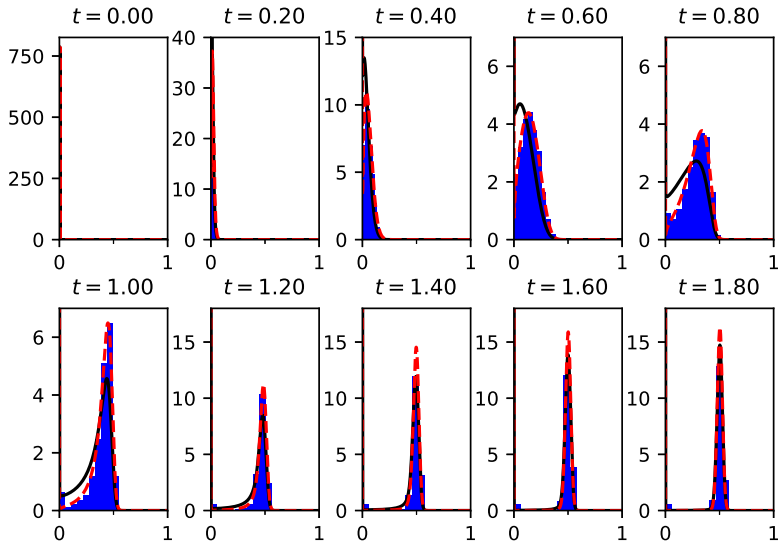

Figure 6: Temporal evolution of the probability distribution $p_{x=\frac{k}{N}}$ (blue histogram) sampled from 25000 realisations of epidemics across 100 realisations of regular networks (2nd row of Table 1), with $N = 1000$. Lines are the numerical solutions to the Fokker-Planck equation (4) computed from two different a_k rates: best (C, α, p) fit (continuous curve) and cubic spline of the raw a_k computed as in eq. (2) (dashed line). The first panel shows the initial condition ($t = 0$), which for all simulations is $k_0 = 1$, while the last panel shows the quasi-steady state distribution.

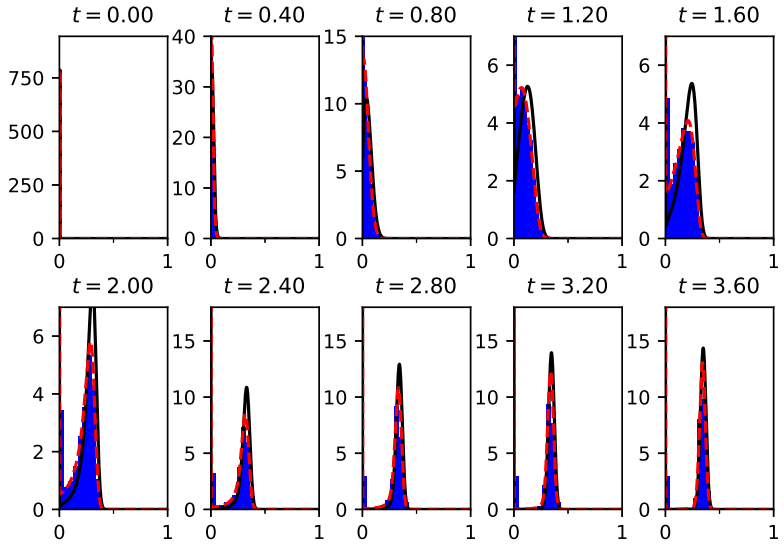

Figure 7: Same scenario as in figure 6, but using the parameters given in the sixth row of Table 1, i.e., the first parameter configuration for Erős-Rényi networks.

preserves the total probability and can account for these early extinctions.

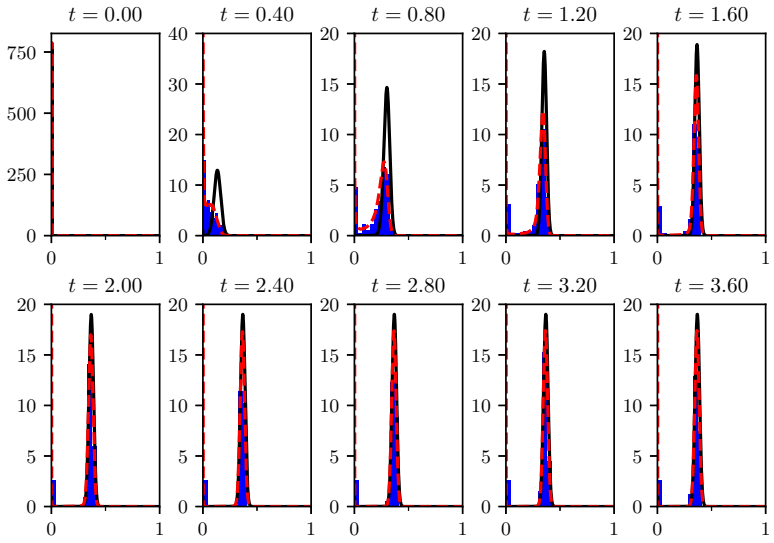

Figure 8: Same scenario as in figure 6, but using the parameters given in the eighth row of Table 1, i.e., the second parameter configuration for Barabási-Albert networks.

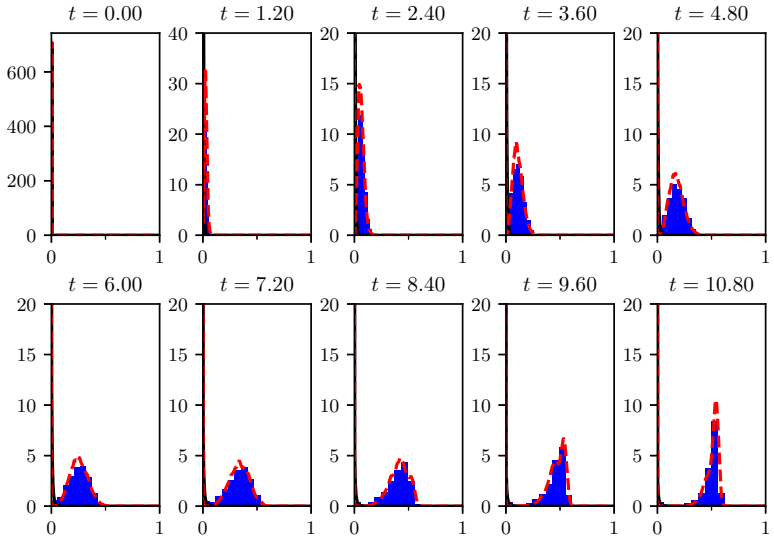

Figure 9: Same scenario as in figure 6, but on a 2D lattice with periodic boundary conditions, using the parameters given in the 10th row of Table 1, i.e., the first parameter configuration for 2D lattices networks.

The PDE with the (C, α, p) model is

$$\partial_t f(t,x) = \frac{1}{2N}\partial_{xx}\left[\left(CN^{2p}\left(x^p(1-x)^p\right)\left(\alpha\left(x-\frac{1}{2}\right)+1\right)+\gamma x\right)f(t,x)\right] + \\ -\partial_x\left[\left(CN^{2p}\left(x^p(1-x)^p\right)\left(\alpha\left(x-\frac{1}{2}\right)+1\right)-\gamma x\right)f(t,x)\right]. \tag{8}$$

Our three-parameter model, (C, α, p), can be used to derive the epidemic threshold. In terms of

the PDE, see equation (8), and as figures 6, 7, 8 and 9 show, an epidemic is supercritical when the drift term is positive. This implies that the epidemic threshold is equivalent to

$$CN^{2p}\Big(x^p(1-x)^p\Big)\left(\alpha\left(x-\frac{1}{2}\right)+1\right)-\gamma x\geq 0,$$

at the start of the epidemic, that is $x\simeq 0$. As shown in [9], for Regular and Erdős-Rényi networks, $p\simeq 1$ and this leads to

$$x\left[CN^2(1-x)\left(\alpha\Big(x-\frac{1}{2}\Big)+1\right)-\gamma\right]\geq 0,$$

taking the limit $x\to 0$ in the expression within the brackets above leads to

$$CN^2\left(1-\frac{\alpha}{2}\right)>\gamma.$$

This expression reduces to the well-known condition $R_0=\frac{\tau}{\gamma}\geq 1$ for fully connected networks. Indeed, scaled rates for such networks can be computed exactly to be $a(x)=\frac{\tau}{N^2}x(1-x)$, meaning that $C=\frac{\tau}{N^2}$, $a=0$ and $p=1$ in this case.

This equation is implicit, as, of course, both C and α depend on the network and epidemic parameters in a non-trivial way. Therefore it cannot be used as it is, but it offers an interesting interpretation since α determines whether the (k,a_k) curves are left- or right-skewed, see figures 5 and [9]. Furthermore, the topology of the underlying network plays an important role in determining the shape of this curve; for example, Barabási-Albert networks lead to (k,a_k) curves with a left skew [9]. Thus, all else being constant, networks with high degree heterogeneity are more likely to see the threshold go past the critical value.

3.3 Inference of infection rates using the Fokker-Planck approximation

In this last section, we provide a simple example of the usefulness of the Fokker-Planck approximation: inferring epidemic and network parameters given data. Specifically, we consider the case in which a single trajectory of BD (or Gillespie simulation of the epidemic on an explicit network) process is observed at discrete time-steps, i.e.:

$$y=\{(t_1,k_1),\ldots,(t_n,k_n)\},$$

where $(t_1,\ldots,t_n)\in[0,T]^n$ $(0\leq t_1<\cdots<t_n\leq T)$ and $(k_1,\ldots,k_n)\in\{0,\ldots,N\}^n$ are the sets of times and states, respectively. To set up the inference, we express the likelihood using the transition probability function of a BD process as follows (using the independence of increments and time homogeneity):

$$\mathcal{L}_{BD}(y;a,c)=\prod_{i=1}^{n-1}\mathbb{P}\left(X(t_{i+1}-t_i)=k_{i+1}|X(0)=k_i;a_k,c_k\right).$$

Unfortunately, for a large state space, these transition probabilities are numerically expensive to compute. Additionally, inferring the full set of rates a_k's and c_k's may not be efficient. Instead, we recast this problem as that of inferring the C,α,p parameters of the Fokker-Planck approximation, as in eq. (8). This is a much more tractable numerical problem, that can still provide useful information about the underlying network and epidemic, as showed in [9]. In terms of the computational complexity of solving the PDE versus solving the ODE system, we argue as follows: the system of ODEs requires to solve exactly $(N+1)$ equations for each time-step (in a naive explicit Euler scheme), meaning that the complexity scales as $O(N)$. The Finite Volume Method we are using

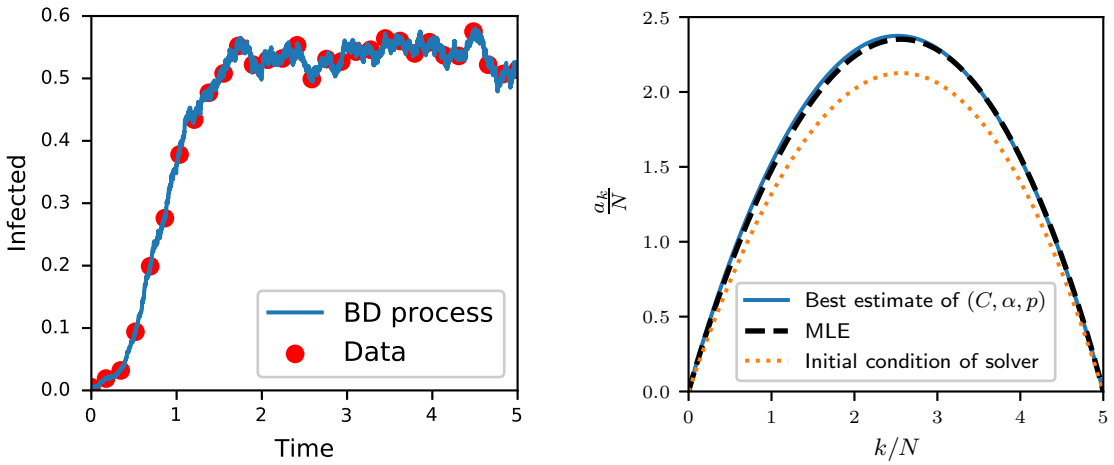

Figure 10: (left) Data generated from a single realisation of an SIS process on an Erdős-Rényi network with $N = 1000, \langle k \rangle = 10$, $\tau = 1$, $\gamma = 4.5$ via the Gillespie algorithm. the curve was sampled regularly to get 30 data-points over 5 units of time. (right) (C, α, p) function obtained by maximising the logarithm of the likelihood 9 (black dashed line) compared to the $(\hat{C}, \hat{\alpha}, \hat{p})$ function obtained by fitting the (k, a_k) curve obtained for $N = 1000$ by exploring the full curve with continuous observations of 10000 epidemics across 50 network realisations, as in figure 4 (blue continuous line). The initial condition inputed to the locally bounded gradient-descent solver is shown by the orange dotted line.

requires us to invert a matrix at each time-step. The size of the matrix is $V \times V$, where V is the size of the volume mesh. The computational complexity of this operation is $O(V^\alpha)$, $\alpha \geq 1$. The advantage of PDE and FVM is that we can choose $V = O(N^\beta)$ such that $\beta\alpha \leq 1$. Notice that a viable choice of β is $\beta = 0$, in which case the grid size is constant with respect to N. This results in a trade-off between space resolution and speed of our solver, while the ODE does not offer this degree of flexibility. This leads us to replace the previous likelihood function with the following:

$$\mathcal{L}_{FP}\left(y; C, \alpha, p\right) = \prod_{i=1}^{n-1} f(t_{i+1} - t_i, x_{i+1}; x_i, C, \alpha, p), \tag{9}$$

where $f(t, x; x', C, \alpha, p)$ is the transition probability density obtained from equation (4), the coefficients are given by the C, α, p model, $x_i = \frac{k_i}{N}$ for all $i \in [1, n]$ and the initial data is a Dirac delta at location $x' \in (0, 1)$. In this example, f is computed numerically using the finite-volume numerical scheme described in Appendix 6. To illustrate the accuracy of this approach, we consider a set of parameters from the choices of Table 1 (figure 10 shows the behaviour of the system when parameters are those on the 5th row of Table 1, i.e. $C = 1.36e-05, \alpha = 3.44e-2, p = 9.7e-1$) and generate a trajectory from a single realisation of the SIS epidemic on a Erdős-Rényi network of size 1000, via Gillespie algorithm. This dataset is shown in figure 10 and consists of $n = 30$ distinct and equally spaced data points taken from the epidemic curve. These are then scaled to $[0, 1]$ (taking $x_i = \frac{k_i}{N}$ for all $i \in [1, n]$). The dataset is then used to find a Maximum Likelihood Estimator (MLE) by simply maximizing the likelihood function from equation (9) with respect to C, α, p, that is finding

$$(\hat{C}, \hat{\alpha}, \hat{p}) = \arg\max \mathcal{L}_{FP}\left(y; C, \alpha, p\right).$$

To show that this method provides a good estimate, we simply plot the MLE rates, $(\hat{C}, \hat{\alpha}, \hat{p})$, against the rates obtained by fitting the (k, a_k) curve directly from data from continuous observations of

multiple realisations of epidemics on networks of size $N = 1000$, we call this the best estimate, as in figure 4. The best estimate and the rates based on $(\hat{C}, \hat{\alpha}, \hat{p})$ are indeed in good agreement. We repeated the inference scheme for all benchmark cases in Table 1 (not shown), with the exception of lattices. The agreement was similar to that shown in figure 10 (right panel).

It is worth noting that the goodness of inference depends on how many points the dataset contains and also how much of the transient and the quasi-steady state is captured. In the transient, the drift dominates and the process is more stochastic. On the other hand, in the quasi-steady state, the drift coefficient tends to zero and fluctuations around it are mainly due to diffusion. Hence, both regimes are needed if drift and diffusion are to be inferred correctly. Data in the transient or in the quasi-steady state alone can lead to sub-optimal inference as different parameter combinations that provide good fit can be found. The example reported in figure 10 is an illustration of how useful the PDE limit of epidemics on finite networks can be in a network and epidemic inference setting. Further, the approximation that we provide can also be used in a Bayesian approach, by first setting a prior over the parameters C, α, p for instance.

4 Conclusions

In this paper we conjectured and showed numerical evidence for the existence of PDE limits for exact SIS epidemics on Regular, Erdős-Rényi and Barabási-Albert networks and 2D lattices with periodic boundary conditions. The key to our approach is to use a BD approximation which then has a PDE limit provided that the coefficients of the BD process are density-dependent. Hence, one of the main challenges was to verify, at least numerically, that this was the case. What is common between all the networks that we considered is that simply increasing the number of nodes in the network will not change what a node experiences locally, e.g. the number or distribution of neighbours. In fact for Erdős-Rényi networks we made sure this is the case by choosing the probability of connection p such that $p = \langle k \rangle / (N-1)$. The same argument seems to hold for scale free networks where the average degree stays constant and nodes at any scale in networks of any size experience the same type of neighbourhood. Of course, this is not the case for fully connected networks since the number of neighbours of a node increases with network size. Furthermore, we note that a_k's are random variables with some distribution around a well-defined mean. This spread/variance in a_k's is in some sense due to higher-order moments in the network. The variance of these distributions is larger at the beginning of the epidemic and it decreases with time or as k increases - meaning that the system tends to reach an equilibrium where the higher order moments in the network are not significant enough to produce a real effect. Based on these arguments, we expect our method to extend readily to configuration networks [30] whose degree distribution does not depend on network size.

Of further interest will be to test and, if it works, extend our approach to clustered and/or networks with community structure. This is a difficult task as clustering and community structure can be introduced in different ways; for example clustered networks can be generated by using Big-V rewiring or by using a family of clustering inducing subgraphs (e.g. triangles, four fully connected nodes etc)p [35]. For networks with community structure, a good choice could be the stochastic block model [18]. However, as our analysis shows the (C, α, p) model struggles to capture the infection rate curves for all four network models that we considered. This suggests that a more flexible model is needed, possibly a non-parametric one, expecially when networks with more complex topologies are considered.

To solve the PDE numerically, we employed a second order in time finite volume method whose

stability was proven in [6]. We compared such numerical solutions to probability distribution sampled from the Gillespie simulation. The agreement is in general good and, as expected, it becomes excellent as N increases. The existence of the PDE limit is not surprising, given that the coefficients of the BD process are density-dependent. However, it is important to note that the agreement between the solutions of the PDEs and empirical distributions based on simulations provides strong support for the validity of the BD process, strengthening the evidence provided in [9], and thus closing the loop illustrated in figure 1.

A PDE perspective on epidemics provides several efficiency gains. The first is to do with computational efficiency and the possibility to quantify variability. More importantly perhaps, the solution of the PDE serves as a likelihood which can be very efficiently computed/evaluated and can form the basis of many networks and epidemic inference models, see Section 3.3. This is in contrast with approaches where the networks are explicitly modelled [26] and computational complexity can make inference out of reach.

At least three separate avenues of future research emerge. First, and perhaps most importantly, a theoretical justification for the Birth-and-Death approximation is still needed, if indeed that is possible. Second, there is a need to investigate the extent to which this method can be extended to other network families and epidemic dynamics. Thirdly, there is scope to consider how the approximate master equation could be used to look into the impact of the network on quantities such as time to extinction. Nevertheless, given that handling exact epidemic models on networks is still challenging even for networks of modest size, we believe that proposing new ways to approximate epidemics is worthwhile and may contribute new modelling and analysis perspectives.

5 Acknowledgments

All authors acknowledge support from the Leverhulme Trust for the Research Project Grant RPG-2017-370. The authors acknowledge useful discussions with Dr M. Dashti during the development of this research.

6 Appendix: numerical method for solving the PDE

In this section we detail the numerical method and algorithms used to solve eq. (4). The algorithm employed is an adaptation and modification of the finite volume method (FVM) named FVM3 in [6]. First, we write the Fokker-Planck equation in the form

$$\frac{\partial f(x,t)}{\partial t} + \frac{\partial j(x,t)}{\partial x} = 0,$$

where in our case the current term is $j(x,t) = -\frac{1}{2N}\frac{\partial \sigma^2(x) f(x,t)}{\partial x} + \mu(x) f(x,t)$, while initial and boundary conditions are:

$$\begin{cases} f(x,0) = \delta(x - x_0), & \text{initial condition,} \\ f(0,t) = 0, & \text{absorption in } x = 0, \\ \frac{\partial f(x,t)}{\partial x}|_{x=1} = 0, & \text{reflection in } x = 1. \end{cases}$$

In our case, both $\mu(x)$ and $\sigma(x)$ vanish at 0, indicating that the only possible steady state is absorption [23]. Therefore, the solution to this equation is such that $\lim_{t\to\infty} f(x,t) = \delta(x)$. Further, since the solution should provide a probability density function, we require that $f(x,t) \geq 0$ everywhere (positivity) and that $\int_0^1 f(x,t)\,dx = 1$ for any $t > 0$ (conservation of mass). Finite Volume Methods are a class of numerical methods to solve PDEs [12] in which the constraints described above are explicitly satisfied, therefore FVM is the natural candidate for this type of problems. Following notation of [6] we consider a uniform grid, with spacing $h = \frac{1}{M}$ and grid points $x_i = ih$, $0 \leq i \leq M$. Similarly for time, we consider a uniform grid with spacing τ and grid points $t_i = n\tau$, $0 \leq n \leq n_{max}$. We define j_i^n and f_i^n to be the numerical approximations of $j(x_i, t_n)$ and $f(x_a, t_n)$, respectively. The control volume $\mathcal{D}_i = \{x \; s.t. \; x_{i-\frac{1}{2}} \leq x \leq x_{i+\frac{1}{2}}\}$ is associated to each inner point x_i, whereas two control domains $\mathcal{D}_0 = \{x \; s.t. \; 0 \leq x \leq x_{\frac{1}{2}}\}$ and $\mathcal{D}_M = \{x \; s.t. \; x_{M-\frac{1}{2}} \leq x \leq 1\}$ are reserved for boundary points, and $x_{i+\frac{1}{2}}$ is defined as $(i + \frac{1}{2})h$.

The discretisation of the time derivative can be done with a first-order scheme (as in [6]) or a higher-order scheme. We opted for a second-order scheme for time, for which, in general

$$\frac{\partial f(x_i, t_n)}{\partial t} \approx \frac{3f_i^n - 4f_i^{n-1} + f_i^{n-2}}{2\tau},$$

and the first iteration is done with the first order time scheme $\frac{\partial f(x_i,t_{n+1})}{\partial t} \approx \frac{f_i^{n+1} - f_i^n}{\tau}$. The reason for this choice is that in our case the current term contains both first and second order space derivatives, so to balance out the required space precision, we matched it with a second-order discrete time derivative. This improved the stability of the solution.

The discretised Fokker-Planck equation then becomes:

$$\frac{3f_i^n - 4f_i^{n-1} + f_i^{n-2}}{2\tau} + \frac{j_{i+\frac{1}{2}}^n - j_{i-\frac{1}{2}}^n}{h} = 0 \tag{10}$$

We now define the numerical equations imposed by the boundary conditions. The stability of the numerical scheme (in particular, conservation of mass) is influenced by the boundary condition at $x = 0$. Naturally, this condition would be $f(0,t) = 0$ (absorption), as we already discussed. However, changing it to be a zero-current condition (i.e. $j(x,t) = 0$) results in a numerical solution

that is more stable. This change of condition does not influence the solution, as the discretised process is never evaluated at $x = 0$. Therefore, we use the following boundary conditions:

$$j(0,t) = j(1,t) = 0, \tag{11}$$

which translates to:

$$\frac{3f_i^n - 4f_i^{n-1} + f_i^{n-2}}{2\tau} + \frac{j_{\frac{1}{2}}}{h/2} = 0, \qquad \frac{3f_i^n - 4f_i^{n-1} + f_i^{n-2}}{2\tau} - \frac{j^n_{M-\frac{1}{2}}}{h/2} = 0. \tag{12}$$

The difference between instances of FVMs is how the current term is discretised. In [6], several different schemes are explored. In particular, the FVM that performed better was the so-called central scheme FVM, in which rather than discretise each component of the current term, we discretise the current as a whole, as follows:

$$j^n_{i+\frac{1}{2}} = -\frac{1}{2N}\frac{\sigma^2(x_{i+1})f^n_{i+1} - \sigma^2(x_i)f^n_i}{h} + \frac{\mu(x_{i+1})f^n_{i+1} - \mu(x_i)f^n_i}{2}, \tag{13}$$

where $\sigma^2(x_i) = a(x_i) + \gamma x_i$ and $\mu(x_i) = a(x_i) - \gamma x_i$.

The discretised PDE then becomes:

$$\begin{aligned}&\frac{3f_i^n - 4f_i^{n-1} + f_i^{n-2}}{2\tau} - \frac{1}{2N}\frac{\sigma^2(x_{i+1})f^n_{i+1} - \sigma^2(x_i)f^n_i}{h^2} + \frac{\mu(x_{i+1})f^n_{i+1} - \mu(x_i)f^n_i}{2h} + \\ &+\frac{1}{2N}\frac{\sigma^2(x_i)f^n_i + \sigma^2(x_{i-1})f^n_{i-1}}{h^2} + \frac{\mu(x_i)f^n_i - \mu(x_{i-1})f^n_{i-1}}{2h} = 0,\end{aligned} \tag{14}$$

with the boundary conditions (12). Note that this is an implicit scheme, that requires at each time-step to invert a matrix of size $M \times M$.

The initial condition $f(x,0) = \delta(x - x_0)$ is approximated by a Normal distribution $f(x,0) \approx \mathcal{N}(x_0, \tilde{\sigma})$ with $\tilde{\sigma} \ll 1$. The stability of the solution with respect to the variance $\tilde{\sigma}$ is discussed in [6], and we have chosen $\tilde{\sigma} = h^2$. The mismatch that can be seen in figures 6 and 7 at 0 is due to the fact that the algorithm cannot reproduce a δ in 0, and should not be considered a problem, as the mass outside of 0 is correctly computed by the numerical solver. To test whether absorption at 0 could have been a problem for the solver, we repeated the calculation allowing for a small re-infection rate at 0 $\epsilon > 0$, without noticing differences in the results. Our implementation is available online at https://github.com/Fdl1989/PDElimitofepidemics.

7 Appendix: universal infection rate curves for Barábasi-Albert networks and lattices

Below we show evidence that the (k, a_k) parameters scale favourably with network size and thus satisfy the density-dependent condition.

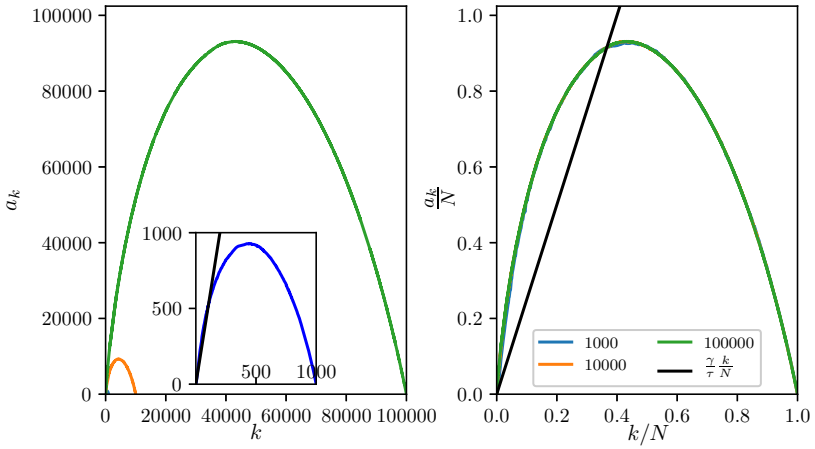

Figure 11: Same scenario as in figure 3 but for Barabási-Albert networks using parameter values from the eighth row of Table (1). (a) Unscaled (k, a_k) curves for values of $N = 10^3, 10^4$ and 10^5. Each curve is obtained by simulating 1000 realisations of the epidemic across 50 realisations of the network; half of the epidemics starting from $k_0 = 1$, the other half from $k_0 = N$. The inset shows a $30x$ zoom of the curve produced for $N = 1000$. (b) Corresponding scaled $(k, \frac{a_k}{N})$ curves.

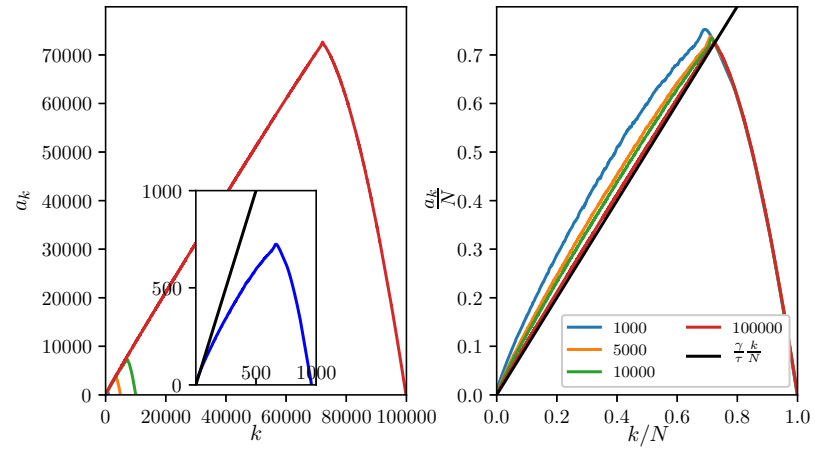

Figure 12: Same scenario as in figure 3 but for 2D lattice with periodic boundary conditions, using parameter values from the tenth row of Table (1). (a) Unscaled (k, a_k) curves for values of $N = 1024, 5041, 10000, 100489$. Each curve is obtained by simulating 1000 realisations of the epidemic across 50 realisations of the network; half of the epidemics starting from $k_0 = 1$, the other half from $k_0 = N$. The inset shows a $30x$ zoom of curve produced for $N = 1000$. (b) Corresponding scaled $(k, \frac{a_k}{N})$ curves.

References

[1] L. Allen. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modelling*, 2, 03 2017.

[2] R. Applegate. An introduction to population genetics: Theory and application. *The Canadian Field-Naturalist*, 127:281, 11 2013.

[3] F. Brauer and C. Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. Springer, 2nd edition, 2012.

[4] A. Bátkai, Á. Havasi, R. Horváth, D. Kunszenti-Kovács, and P. Simon. Pde approximation of large systems of differential equations. *Operators and Matrices*, 9, 03 2013.

[5] E. Cacio, S. E. Cohn, and R. Spigler. Numerical treatment of degenerate diffusion equations via Feller's boundary classification, and applications. *Numerical Methods for Partial Differential Equations*, 28(3):807–833, may 2012.

[6] M. Chen, C. Liu, S. Xu, and X. Yue. Behavior of different numerical schemes for population genetic drift problems. Technical report, 2016.

[7] P. Collet, S. Martínez, and J. San Martín. *Quasi-Stationary Distributions: General Results*. 2013.

[8] L. Decreusefond, J.-S. Dhersin, P. Moyal, and V. C. Tran. Large Graph Limit for an {SIR} Process in Random Network with Heterogeneous Connectivity. *The Annals of Applied Probability*, 22(2):541–575, 2012.

[9] F. Di Lauro, J.-C. Croix, M. Dashti, L. Berthouze, and I. Z. Kiss. Network Inference from Population-Level Observation of Epidemics. Technical report, 2019.

[10] C. Duan, C. Liu, C. Wang, and X. Yue. Numerical complete solution for random genetic drift by energetic variational approach. *ESAIM: Mathematical Modelling and Numerical Analysis*, 03 2018.

[11] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons, 2009.

[12] R. Eymard, T. Gallouët, and R.èle Herbin. Finite Volume Methods. In J. L. Lions and Philippe Ciarlet, editors, *Solution of Equation in Rn (Part 3), Techniques of Scientific Computing (Part 3)*, volume 7 of *Handbook of Numerical Analysis*, pages 713–1020. Elsevier, 2000.

[13] W. Feller. Diffusion processes in one dimension. *Transactions of the American Mathematical Society*, 77(1):1–31, 1954.

[14] C. W. Gardiner. *Handbook of stochastic methods for physics, chemistry and the natural sciences*, volume 13 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, third edition, 2004.

[15] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403 – 434, 1976.

[16] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[17] A. Gray, D. Greenhalgh, L. Hu, X. Mao, and J. Pan. A Stochastic Differential Equation SIS Epidemic Model. *SIAM Journal of Applied Mathematics*, 71(3):876–902, 2011.

[18] P. W. Holland, K. B. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social networks*, 5(2):109–137, 1983.

[19] S. Janson, M. Luczak, and P. Windridge. Law of large numbers for the {SIR} epidemic on a random graph with given degrees. *Random Structures & Algorithms*, 2014.

[20] NG Van Kampen. *Stochastic processes in physics and chemistry*. North Holland, 2007.

[21] M. J. Keeling. The effects of local spatial structure on epidemiological invasions. *Proceedings of Biological Science*, 266, 1999.

[22] I. Z. Kiss, J. C. Miller, and P. L. Simon. *Mathematics of Epidemics on Networks: from exact to approximate models*. Springer, 2017.

[23] R. M. Kovacevic. Stochastic contagion models without immunity: their long term behaviour and the optimal level of treatment. *Central European Journal of Operations Research*, 26(2):395–421, jun 2018.

[24] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump {M}arkov processes. *Journal of applied Probability*, 7(1):49–58, 1970.

[25] J. Lindquist, J. Ma, P. van den Driessche, and F. H. Willeboordse. Effective degree network disease models. *Journal of Mathematical Biology*, 62(2):143–164, 2011.

[26] L. Ma, Q. Liu, and P. Van Mieghem. Inferring network properties based on the epidemic prevalence. *Applied Network Science*, 4(1):93, 2019.

[27] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probability Surveys*, 9(1):340–410, 2012.

[28] J. C. Miller, A. C. Slim, and E. M. Volz. Edge-Based Compartmental Modelling for Infectious Disease Spread. *Journal of the Royal Society Interface*, 9(70):890–906, 2012.

[29] M. Mohammadi and A. Borzì. Analysis of the chang–cooper discretization scheme for a class of fokker–planck equations. *Journal of Numerical Mathematics*, 23, 09 2015.

[30] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics probability and computing*, 7(3):295–305, 1998.

[31] N. Nagy, I. Z. Kiss, and P. L. Simon. Approximate Master Equations for Dynamical Processes on Graphs. *Mathematical Modelling for Natural Phenomena*, 9(2):43–57, 2014.

[32] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925, 2015.

[33] M. A. Porter and J. P. Gleeson. Dynamical systems on networks. *Frontiers in Applied Dynamical Systems: Reviews and Tutorials*, 4, 2016.

[34] G. Rempala, H. Koeppl, R. Durrett, and H. van der Hofstad. 19w5071 - scaling limits of dynamical processes on random graphs. *https://www.birs.ca/events/2019/5-day-workshops/19w5071*, 2019.

[35] M. Ritchie, L. Berthouze, and I. Z. Kiss. Generation and analysis of networks with a prescribed degree sequence and subgraph family: higher-order structure matters. *Journal of complex networks*, 5(1):1–31, 2017.

[36] M. Roberts, V. Andreasen, A. Lloyd, and L. Pellis. Nine challenges for deterministic epidemic models. *Epidemics*, 10:49–53, 2015.

[37] H. Silk, G. Demirel, M. Homer, and T. Gross. Exploring the adaptive voter model dynamics with a mathematical triple jump. *New Journal of Physics*, 16(9):093051, 2014.

[38] F. Trabelsi and N. Naouara. Boundary classification and simulation of one-dimensional diffusion processes. *International Journal of Mathematics in Operational Research*, 11:107–138, 08 2017.