# A BINOMIAL MOMENT APPROXIMATION SCHEME FOR EPIDEMIC SPREADING IN NETWORKS

Yilun SHANG[1]

*Epidemiological network models study the spread of infectious diseases through a population of individuals. In this paper, we study a moment approximation scheme for the SIS (susceptible-infected-susceptible) epidemics spreading on configuration model networks via an empirical binomial distribution with time dependent parameters describing the number of infectives during the outbreaks. Based on this assumption, the evolution equations of higher order moments are expressed in terms of lower order moments. Numerical examples are provided to illustrate the availability of our approximation method.*

**Keywords**: Kolmogorov equation, moment closure, network, epidemics.
**MSC 2010:** 92D30, 60J28.

## 1. Introduction

Complex network has emerged as a prominent field in complex system research, and network models for disease propagation in human society have been used to understand many problems in epidemiology [1, 4, 9, 13, 15, 20, 21]. The threshold of the infectivity in the paradigmatic susceptible-infected-susceptible (SIS) model, for example, exhibits distinct phenomena for different network topologies [3, 17, 24]: while regular and random networks possess a non-zero epidemic threshold, that is a critical value of transmission probability under which the disease ultimately dissipates, such threshold disappears asymptotically in scale-free networks. A common approach to describe the dynamic behavior of the epidemic dynamics is by a Kolmogorov equation (or master equation) that governs the time evolution of the joint probability function of the underlying processes and naturally leads to Markovian models [22]. However, for a network with $N$ nodes, the state space is much larger than $N$ (e.g., with $2^N$ elements for SIS epidemics). Solving this system becomes a formidable task, especially when dealing with large-scale networks.

To address this problem, pairwise-type approximate models are proposed and heavily used to capture the epidemic dynamics in networks [5, 6, 7, 8, 11]. The classic pairwise model [8] relies on a set of moments equations for the

[1] Researcher, Institute for Cyber Security, University of Texas at San Antonio, Texas 78249, USA, e-mail: shylmath@hotmail.com

expected values of individuals of different state (e.g., susceptible or infected) which depends on the expected values of individuals of different pair states with higher order moments replaced by appropriate chosen functions of lower-order moments (e.g., singles and pairs). The resulting approach is usually referred to as the moment closure/approximation method, which produces a self-contained system of ordinary differential equations (ODEs) whose solution provides approximate values for the moments of the epidemic processes. The similar thinking has been implemented earlier in a wider context of biological population processes and biochemical systems (see e.g. [2, 18, 19, 26]). Other approximate schemes include the probability generating function formalism [25] and the effective degree type models [12].

Recently, a novel moment closure is introduced in [10] based on the empirical observations that the number of infectives in SIS epidemics is well described by a binomial distribution with time dependent parameters. By using an a priori binomial distribution, the difference between the exact system from the solution of the approximate model is $O(N^{-2})$ compared to $O(N^{-1})$ obtained via classic moment closure at the level of triples used for pairwise models. Note that the population considered therein is modeled by a fully connected graph (or complete graph), which limits the application of the proposed methodology.

In this work, we investigate an SIS epidemic process on a random graph with arbitrary degree distribution. As in [10] we derive the ODE-based approximate model capturing the moments of the number of infectives at all times combined with the empirical binomial distribution with time dependent parameters. We show that the proposed model works well when the underlying network is generated by a configuration model [14] with homogeneous/heterogeneous degree distributions via numerical simulations.

The rest of the paper is organized as follows. In Section 2, we present the SIS model on a configuration model graph and its Markov chain representation. In Section 3, we derive the binomial moment closure. Finally, we discuss some possible improvements in the closing section.

## 2. The model

In this section, we introduce the configuration model graphs and describe the transmission of SIS epidemics on such graphs with a dynamical systems type approach.

A configuration model network is static with a known degree distribution [14]. We create a configuration model network with *N* nodes as follows. Suppose we are given i.i.d. random variables $d_1, \cdots, d_N$ with distribution *P(k)* (*k=0,…,N*) that represent the degrees of each node. To the node *i* are associated $d_i$ stubs (half-

edges). Once all nodes are assigned stubs, we choose two open stubs uniformly at random and pair them together to form an edge. We define

$$G(z) = \sum_{k=0}^{N} P(k) z^k, \tag{1}$$

the probability generating function of the degree distribution. So the average degree is $<k>=G'(1)$, and the average number of nodes within 2-hop distance of a node is $<k^2>=G'(1)+G''(1)$ [16]. A sample configuration model network is shown in Fig. 1.

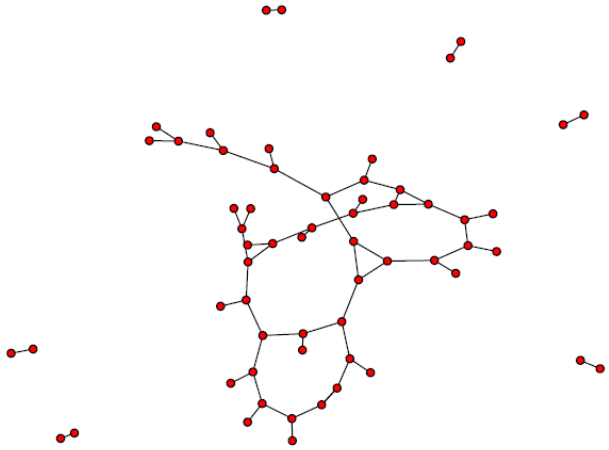

Fig. 1. A sample configuration model network with 60 nodes. The degrees are chosen using $P(1)=P(3)=0.5$. Thus, $G(z)=(z^3+z)/2$.

In this context, the population for the prototypical SIS model consists of $N$ nodes, whose states can be either susceptible or infected. This model is customarily used to describe the progression of infectious diseases conferring temporary immunity, such as common cold. An infected node spreads the disease to each one of its susceptible contacts at rate $\beta$, while it heals at a rate $\gamma$ with all events occurring independently of each other. Denote by $S$ and $I$ the sizes of the set of susceptible and infected nodes, respectively. Then $N=S+I$ at any given time. The number of infectives for the model is approximately described by a continuous time Markov chain on the state space

$$\begin{aligned} k &\longrightarrow k+1 \quad \text{at rate} \quad b_k, \\ k &\longrightarrow k-1 \quad \text{at rate} \quad d_k, \end{aligned} \tag{2}$$

where we propose to use

$$b_k = \frac{\beta k(N-k)\langle k^2\rangle}{\langle k\rangle(N-1)} \quad \text{and} \quad d_k = \gamma k, \tag{3}$$

for $k=0,\ldots,N$. While the $k \to k-1$transition is itself a spontaneous process, the transition from $k \to k+1$depends on the structure of the population and the contact patterns of individuals. Note that if the underlying contact network is a fully connected graph (i.e., $p_{N-1}=1$), it gives rise to $b_k = \beta k(N-k)$, which is the situation studied in [10, 23]. According to the above comments, the factor $\langle k^2\rangle/(\langle k\rangle(N-1))$ in (3) encodes a density dependent ``expansion'' property averaged over the ensemble and the infectives and susceptibles are assumed to be randomly distributed on the network.

Although quite straightforward, this extension provides an avenue to address epidemic spreading in general network settings, such as scale-free degree distributions, community structure and small-world phenomenon, which set them apart from simpler networks such as fully connected graphs. For example, the scale-free degree distribution is characteristic of many real-world networks, including social and computer networks on which human diseases and computer viruses propagate [15].

The computational efficiency of our approach will be shown in the next section. Let $p_k(t)$ be the probability that the system is in state $k$ at time $t$. The Kolmogorov forward equation for this process is

$$\dot{p}_k(t) = b_{k-1}p_{k-1} - (b_k + d_k)p_k + d_{k+1}p_{k+1}, \tag{4}$$

with ``birth'' rate $b_k$ and ``death'' rate $d_k$ given by (3), and additionally, $b_{-1}=d_{N+1}=0$.

### 3. Binomial moment approximation

In this section, we derive the moment equations and close them based on the empirical observation that $p_k(t)$ is well described by a binomial distribution [10].

For an integer $i \geq 1$, define

$$X_i(t) = \sum_{k=0}^{N} k^i p_k(t) \tag{5}$$

be the $i$th moment associated with the above process (2). By using the Kolmogorov equation (4), the equation for the first moment can be derived as

$$\begin{aligned}\dot{X}_1(t) &= \sum_{k=0}^{N}(kb_{k-1}p_{k-1} - kb_kp_k - kd_kp_k + kd_{k+1}p_{k+1}) \\ &= \sum_{k=0}^{N}(b_k - d_k)p_k \\ &= \sum_{k=0}^{N}\left(\frac{\langle k^2\rangle}{\langle k\rangle(N-1)}(\beta kN - \beta k^2) - \gamma k\right)p_k \\ &= \left(\frac{\beta\langle k^2\rangle N}{\langle k\rangle(N-1)} - \gamma\right)X_1 - \frac{\beta\langle k^2\rangle}{\langle k\rangle(N-1)}X_2.\end{aligned} \tag{6}$$

Similarly, the equation for the second moment is given by

$$\begin{aligned}\dot{X}_2(t) &= \sum_{k=0}^{N}((2k+1)b_k - (2k-1)d_k)p_k \\ &= \left(\frac{\beta\langle k^2\rangle(2N-1)}{\langle k\rangle(N-1)} - 2\gamma\right)X_2 - \frac{2\beta\langle k^2\rangle}{\langle k\rangle(N-1)}X_3 \\ &\quad + \left(\frac{\beta\langle k^2\rangle N}{\langle k\rangle(N-1)} + \gamma\right)X_1.\end{aligned} \tag{7}$$

Let $x_i(t) = \sum_{k=0}^{N}(k/N)^i p_k(t)$ for $i$=1,2,…. Therefore, (6) and (7) can be rewritten in terms of $x_i$'s as

$$\dot{x}_1(t) = \left(\frac{\beta\langle k^2\rangle N}{\langle k\rangle(N-1)} - \gamma\right)x_1 - \frac{\beta\langle k^2\rangle N}{\langle k\rangle(N-1)}x_2 \tag{8}$$

and

$$\dot{x}_2(t) = \left(\frac{\beta\langle k^2\rangle(2N-1)}{\langle k\rangle(N-1)} - 2\gamma\right)x_2 - \frac{2\beta\langle k^2\rangle N}{\langle k\rangle(N-1)}x_3 + \left(\frac{\beta\langle k^2\rangle}{\langle k\rangle(N-1)} + \frac{\gamma}{N}\right)x_1 \tag{9}$$

Note that the above system is not self-contained since the second moment ($x_2$) relies on the third moment ($x_3$). In theory, the dynamics for $x_3$ can be evaluated by a differential equation similar to the ones above, which requires evaluation of higher-order moments [8]. Now we employ the empirical observation that the distribution of the infectives are given by a binomial distribution Bin($n$, $p$), and hence the parameters $n$ and $p$ can be expressed by [10]

$$p = 1 + X_1 - \frac{X_2}{X_1}, \quad \text{and} \quad n = \frac{X_1^2}{X_1 + X_1^2 - X_2}. \tag{10}$$

The third moment can then be recast in terms of the first and second moment [10]

$$\begin{aligned}X_3 &= np + 3n(n-1)p^2 + n(n-1)(n-2)p^3 \\ &= \frac{2X_2^2}{X_1} - X_2 - X_1X_2 + X_1^2\end{aligned} \tag{11}$$

which is equivalent to

$$x_3 = \frac{2x_2^2}{x_1} - x_1 x_2 + \frac{x_1^2 - x_2}{N}. \tag{12}$$

The closure (12) together with (8) and (9) provides a self-contained system, whose solution gives approximate values for the moments for the SIS process. In general, the expected number of infectives is given by

$$E[I(t)] = \sum_{k=0}^{N} k p_k(t) = X_1(t) = N x_1(t). \tag{13}$$

We demonstrate in Fig. 2 a comparison of theoretical results with Monte-Carlo stochastic simulations. The population size is taken as $N=10^3$ and 10 randomly chosen nodes are infected at the initial state for all the simulations. The results show good agreement for the approximate system (except that we have made time shifts for simulations due to stochastic effects early on in the epidemic).

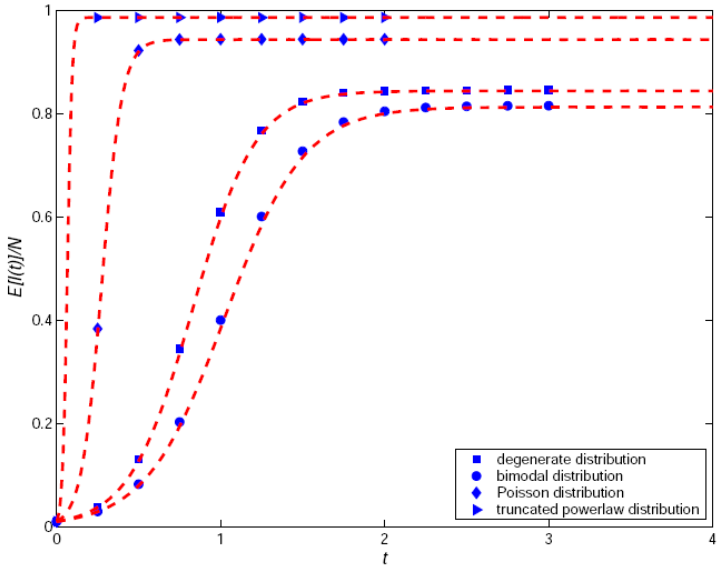


Fig. 2. Plot of fraction infected $E[I(t)]/N$ against time based on simulation and binomial closures (dashed curves) in configuration model networks of $10^3$ nodes with $\beta$=1.6 and $\gamma$ =1. Simulations are conducted for degenerate distribution: $P(4)=1$ (squares), bimodal distribution: $P(2)=P(4)=0.5$ (circles), Poisson distribution with $\langle k\rangle=10$ (diamonds), and truncated power law distribution: $P(k)=0.673k^{-2}e^{-k/30}$ for $1\leq k\leq 20$ (triangles).

## 4. Discussion

We have presented a framework which allows us to make some analytical headway in deriving low-dimensional approximate models applicable in network settings. The moment closure method is based on an a priori assumption about the

distribution of the infectives. Numerical results show good approximation for epidemic spreading on a range of configuration model networks.

The network models studied here have no degree correlations, namely the probability that an edge arrives at a node of degree $k$ is proportional to $kP(k)$. If a network shows assortative/disassortative mixing (i.e., a tendency for high-degree nodes to connect preferentially to high/low-degree nodes), it would be desirable to refine the above model in order to obtain an improved approximation.

## REFERENCES


[1] *F. Ball, D. Sirl, and P. Trapman*, Analysis of a stochastic SIR epidemic on a random network incorporating household structure. Math. Biosci., **224**(2010), 53-73.

[2] *B. Bolker and S. W. Pacala*, Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. Theor. Popul. Biol., **52**(1997), 179-197.

[3] *C. Castellano and R. Pastor-Satorras*, Competing activation mechanisms in epidemics on networks. Sci. Rep., **2**(2012), Art. No. 371.

[4] *R. Durrett*, Some features of the spread of epidemics and information on a random graph. Proc. Natl. Acad. Sci. USA, **107**(2010), 4491-4498.

[5] *C. S. Gillespie*, Moment-closure approximations for mass-action models. IET Syst. Biol., **3**(2009), 52-58.

[6] *K. Hausken and J. F. Moxnes*, A closure approximation technique for epidemic models. Math. Comput. Model. Dyn. Syst., **16**(2010), 555-574.

[7] *D. Hiebeler*, Moment equations and dynamics of a household SIS epidemiological model. Bull. Math. Biol., **68**(2006), 1315-1333.

[8] *M. J. Keeling*, The effects of local spatial structure on epidemiological invasions. Proc. R. Soc. Lond. B, **266**(1999), 859-867.

[9] *M. J. Keeling and K. T. D. Eames*, Networks and epidemic models. J. R. Soc. Interface, **2**(2005), 295-307.

[10] *I. Z. Kiss and P. L. Simon*, New moment closures based on a priori distributions with applications to epidemic dynamics. Bull. Math. Biol., **74**(2012), 1501-1515.

[11] *I. Krishnarajah, A. Cook, G. Marion, and G. Gibson*, Novel moment closure approximations in stochastic epidemics. Bull. Math. Biol., **67**(2005), 855-873.

[12] *J. Lindquist, J. Ma, P. van den Driessche, and F. H. Willeboordse*, Effective degree network disease models. J. Math. Biol., **62**(2011), 143-164.

[13] *S. Meloni, N. Perra, A. Arenas, S. Gomez, Y. Moreno, and A. Vespignani*, Modeling human mobility responses to the large-scale spreading of infectious diseases. Sci. Rep., **1**(2011), Art. no. 62.

[14] *M. Molloy and B. Reed*, A critical point for random graphs with a given degree sequence. Random Struct. Algor., **6**(1995), 161-180.

[15] *M. E. J. Newman*, The structure and function of complex networks. SIAM Rev., **45**(2003), 167-256.

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

[17] *R. Pastor-Satorras and A. Vespignani*, Epidemic dynamics and endemic states in complex networks. Phys. Rev. E, **63**(2001), 066117.

[18] *D. A. Rand*, Correlation equations and pair approximations for spatial ecologies. CWI Quarterly, **12**(1999), 329-368.

[19] *K. Sato, H. Matsuda, and A. Sasaki*, Pathogen invasion and host extinction in lattice structured populations. J. Math. Biol., **32**(1994), 251-268.
[20] *Y. Shang*, Modelling epidemic spread with awareness and heterogeneous transmission rates in networks. J. Biol. Phys., **39**(2013), 489-500.
[21] *Y. Shang*, Mixed SI(R) epidemic dynamics in random graphs with general degree distributions. Appl. Math. Comput., **219**(2013), 5042-5048.
[22] *Y. Shang*, A Lie algebra approach to susceptible-infected-susceptible epidemics. Electron. J. Differ. Equ., **2012**(2012), Art. no. 233.
[23] *P. L. Simon, M. Taylor, and I. Z. Kiss*, Exact epidemic models on graphs using graph-automorphism driven lumping. J. Math. Biol., **62**(2011), 479-508.
[24] *P. Van Mieghem*, Epidemic phase transition of the SIS type in networks. EPL, **97**(2012), 48004.
[25] *E. Volz*, SIR dynamics in structured populations with heterogeneous connectivity. J. Math. Biol., **56**(2008), 293-310.
[26] *P. Whittle*, On the use of the normal approximation in the treatment of stochastic processes. J. Roy. Stat. Soc. B, **19**(1957), 268-281.