BIOMATH
https://biomath.math.bas.bg/biomath/index.php/biomath
Biomath Forum
ORIGINAL ARTICLE

Equivalence of mass action and Poisson network SIR epidemic models

Grzegorz A. Rempała

Division of Biostatistics, The Ohio State University,
Columbus, OH 43210, United States
rempala.3@osu.edu 0000-0002-6307-4555

Received: October 12, 2023, Accepted: November 23, 2023, Published: December 15, 2023

Abstract: **This brief note highlights a largely overlooked similarity between the SIR ordinary differential equations used for epidemics on the configuration model of a Poisson network and the classical mass-action SIR equations introduced nearly a century ago by Kermack and McKendrick. We demonstrate that the decline pattern in susceptibles is identical for both models. This equivalence carries practical implications: the susceptibles decay curve, often referred to as the epidemic or incidence curve, is frequently used in empirical studies to forecast epidemic dynamics. Although the curves for susceptibles align perfectly, those for infections do differ. Yet, the infection curves tend to converge and become almost indistinguishable in high-degree networks. In summary, our analysis suggests that under many practical scenarios, it's acceptable to use the classical SIR model as a close approximation to the Poisson SIR network model.**

Keywords: **Configuration model, SIR epidemic equations, Kermack-McKendrick model**

I. Introduction

In many practical situations, especially during the recent COVID-19 pandemic, it has been suggested that the classical susceptible-infected-recovered (SIR) equations are robust in the sense that they can accurately depict a wide range of epidemic scenarios, accounting for both heterogeneity of contacts and variations in the infection process. In this brief note, we will consider one of the surprising examples of this, analyzing an interesting and apparently little-known relation between the classical SIR models and the network versions of SIR epidemics. As it turns out, in certain situations, the classical SIR equations are also valid when describing dynamics on specific networks. This fact is somewhat surprising and, at least to some extent, could explain the notable resilience of the classical SIR models, even when they are considered outside the assumptions they were originally based on or in the areas seemingly far removed from mathematical epidemiology, like political science and chemical kinetics [1, 2]. Even though the analysis considered here can be extended more broadly, for simplicity we specifically focus on an example involving a relatively simple network known as the configuration model graph with a given degree distribution. For a more comprehensive review of such models in the context of epidemics and more, refer to the monographs [3] or [4]. In the remainder of this section, we briefly recall the basic notions relevant to our discussion. The most important results can be found in Section 2, where we establish the equivalence between non-network and network SIR models.

A. Classical SIR model

The SIR model, proposed famously by Kermack and McKendrick in [5], offers a foundational mathematical approach to understand the dynamics of infectious diseases. Within this model, individuals in a population are grouped into three compartments: susceptible, infected and recovered, with their temporal proportions denoted usually by $S(t)$, $I(t)$, and $R(t)$, respectively. The time

Citation: Grzegorz A. Rempała, Equivalence of mass action and Poisson network SIR epidemic models, Biomath 12 (2023), 2311237, https://doi.org/10.55630/j.biomath.2023.11.237

progression of these compartments is detailed by the subsequent system of ordinary differential equations:

$$\dot{S} = -\beta SI, \tag{1}$$

$$\dot{I} = \beta SI - \gamma I, \tag{2}$$

$$\dot{R} = \gamma I. \tag{3}$$

Herein, β denotes the rate of infection, representing the frequency at which susceptible individuals become infected upon encountering infected ones. Conversely, γ denotes the recovery rate, defining the pace at which infected individuals recuperate or succumb, subsequently departing from the infected category. The initial conditions for the above system are taken as

$$\begin{aligned} S(0) &= 1, \\ I(0) &= \rho, \\ R(0) &= 0. \end{aligned} \tag{4}$$

A pivotal parameter in epidemiology is the basic reproduction number, $\mathcal{R}_0$, defined as:

$$\mathcal{R}_0 = \frac{\beta}{\gamma}. \tag{5}$$

This value serves as an indicator of the potential for an outbreak. When $\mathcal{R}_0 > 1$, it suggests that the epidemic can spread in the population. It's important to note that these equations inherently assume the *law of mass action*, without a specific contact structure or, in other words, homogeneous interactions among all individuals. This implies that an individual who is infectious can potentially infect any other susceptible individual.

B. Network SIR models

In real-world populations, interactions among individuals are not homogeneous, a simplifying assumption inherent to classical SIR models. In reality, individuals form complex interaction patterns, often described best as networks. These networks, where nodes represent individuals and edges denote interactions or contacts, embody the diverse and intricate structure of relationships within populations. For instance, some individuals, often called 'super-spreaders', have a disproportionately high number of connections, making them more likely to spread infections. Others may have limited interactions and are consequently less exposed. This variability in connectivity is not represent in the system (1)–(3). By extending the classical SIR model to networks, we can capture these heterogeneities, providing a more accurate representation of disease spread. This network-based perspective has given rise to network versions of the SIR differential equations, offering more nuanced insights into outbreak dynamics in structured populations. Notable among these models are the pairwise (PW) model [6, 7], the Volz model [8], and the dynamical survival analysis (DSA) model [9, 10].

The PW model, as detailed in [6, 7], provides equations that predict the expected number of susceptible nodes and infected nodes, along with the anticipated number of $S-I$ and $S-S$ pairs. This model employs a closure technique that predicts the expected number of triples using singles and pairs, eliminating the need to rely on higher-order moments.

On the other hand, the Volz model [8] relies on a differential equation system described via the probability generating function (PGF) of the degree distribution. Unlike other models, it focuses on edge-centric metrics, like the count of edges connecting nodes in specific states, instead of node-centric metrics such as counts of infected or susceptible nodes. This model's results often align closely with ones based on realistic simulations or real data [11]. Furthermore, Decreusefond and colleagues have formally demonstrated that if N is the number of network nodes, the Volz model represents the large N limit for a stochastic SIR epidemic on a configuration model network [12].

C. DSA model and Poisson-type networks

In a more recent study, Jacobsen and colleagues introduced an alternative approach for deriving the mean-field limit of a stochastic SIR model on a network, which is often referred to the dynamical survival analysis (DSA) approach [9,13]. Although the variables analyzed in this approach differ from those in the Volz model, the two limiting network dynamics were shown to be equivalent [13]. Moreover, the DSA framework offers a fresh perspective on epidemics by allowing us to see them through a statistical lens, for instance, by estimating the likelihood that a standard node, which was susceptible at time $t = 0$, remains susceptible at any time $t > 0$.

It turns out that all three models (PW, Volz and DSA) take a particularly simple and mutually equivalent form when the network degree distributions belong to the so-called Poisson-type class of distributions consisting of the Poisson, negative binomial, and binomial families of distributions [9,13,14]. For that class, the large network equations may be described in terms of the limiting proportions of S-type nodes (x_S), $S-I$ pairs (x_{SI}), and the infection density $x_D = x_{SI}/x_S$ [11].

The essence of the simplification for the Poisson-type networks is that the limiting equations involve only

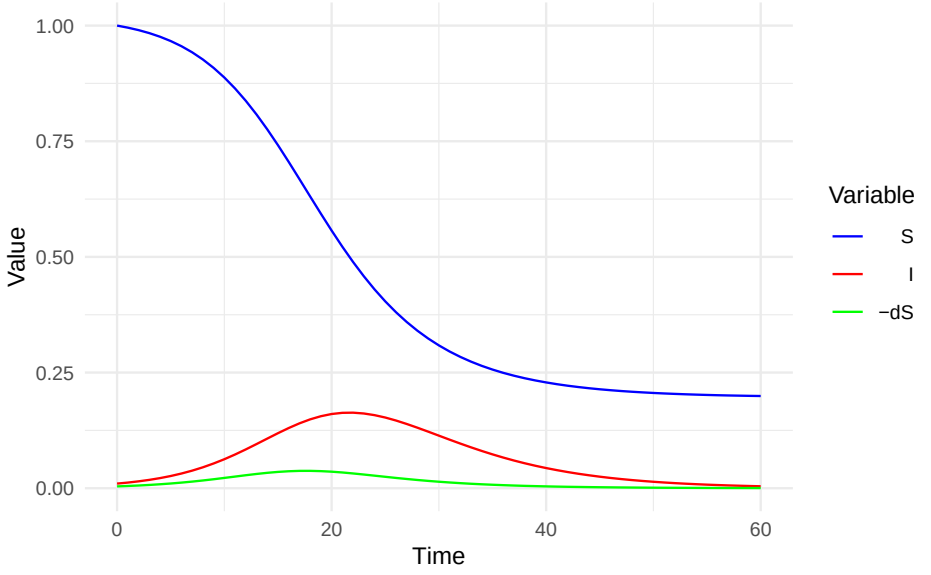

Fig. 1: SIR Curves. SIR model curves from (1)-(3) for the set of parameters $\beta = 0.4$, $\gamma = 0.2$, and $\rho = 0.01$. The lowest curve is the epidemic curve $-\dot{S}$ given in (14).

these quantities along with the proportion of infected x_I as follows:

$$\begin{aligned}\dot{x}_S &= -\tilde{\beta} x_D x_S,\\ \dot{x}_D &= \tilde{\beta}(1-\kappa)x_D^2 + \left(\tilde{\beta}\kappa\mu x_S^{2\kappa-1} - (\tilde{\beta}+\tilde{\gamma})\right) x_D,\\ \dot{x}_I &= \tilde{\beta} x_D x_S - \tilde{\gamma} x_I,\end{aligned} \tag{6}$$

with the set of initial conditions

$$\begin{aligned}x_S(0) &= 1,\\ x_I(0) &= \tilde{\rho},\\ x_D(0) &= \mu\tilde{\rho}.\end{aligned} \tag{7}$$

For simplicity, we omit here the dynamics of recovery, which is the same as in the classical model (3). Note that the constants $\tilde{\beta}$ and $\tilde{\gamma}$ represent the rates of infection and recovery, with $\tilde{\rho}$ representing the initial amount of infection. The quantities μ and κ are network-related and represent, respectively, the average of the network degree and the ratio of the average degree to the average *excess degree*. Recall that the excess degree distribution q_k on a configuration model random graph with degree distribution p_k and average degree $\mu < \infty$ is defined as the degree of a random neighbor node of a randomly selected node of degree at least one. That is, $q_k = (k+1)p_{k+1}/\mu$. For the network SIR model (6), the basic reproduction number (5) needs to be redefined to reflect the added heterogeneity. To avoid ambiguity, we denote this basic reproduction number by $\mathcal{R}_0^{NET}$ and the basic reproduction number given in (5) by $\mathcal{R}_0^{MA}$. The network basic reproduction number is then

$$\mathcal{R}_0^{NET} = \frac{\mu\tilde{\beta}}{\tilde{\beta}+\tilde{\gamma}} \tag{8}$$

and reflects the somewhat intuitive notion that the average number of new infections has to be proportional to the average of the degree distribution.

II. Models equivalence

Let us consider the case of a Poisson network, that is, the special case when $p_k = \exp(-\mu)\mu^k/k!$ for $k \geq 0$ are Poisson probabilities. In this case it is easy to see that we have $q_k = p_k$ and thus $\kappa = 1$. Consequently, the network equations (6) simplify to

$$\dot{x}_S = -\tilde{\beta} x_D x_S, \tag{9}$$

$$\dot{x}_D = \tilde{\beta}\mu x_D x_S - (\tilde{\beta}+\tilde{\gamma}) x_D, \tag{10}$$

$$\dot{x}_I = \tilde{\beta} x_D x_S - \tilde{\gamma} x_I, \tag{11}$$

with the same set of initial conditions (7). These equations may be manipulated in order to yield only a single equation describing the dynamics of susceptibles known as the *epidemic curve equation*.

A. Epidemic curve equations

Dividing (10) by (9) we obtain

$$\frac{dx_D}{dx_S} = -\mu + \frac{\mu}{\mathcal{R}_0^{NET}}\frac{1}{x_S},$$

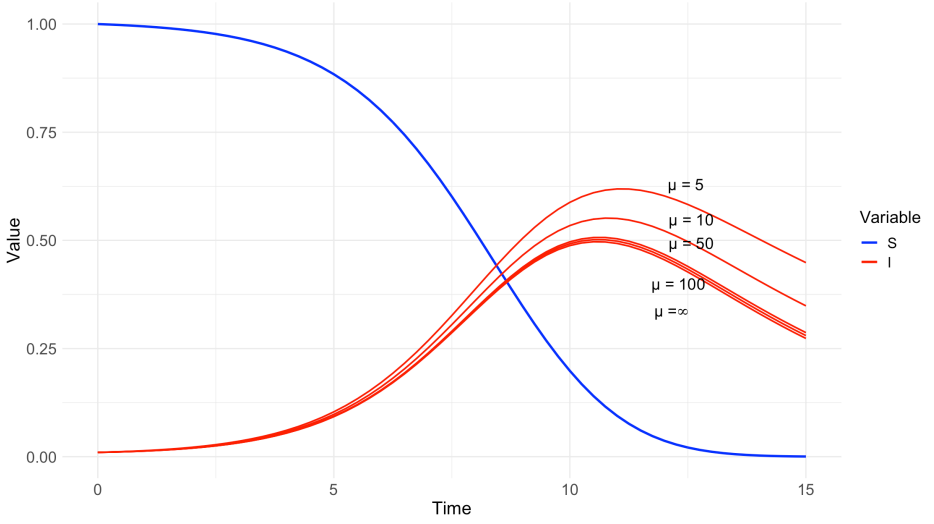

Fig. 2: Approximating Network SIR Model. Poisson network SIR model curves for infected (red) from (9)-(11) plotted for several different mean degrees (μ). The remaining parameter values are as in Figure 1, with the network parameters matched as in Theorem 1. It is seen that for large μ (at least 50) the network curves for infected get close to the one corresponding to the classical SIR model, where $I = x_I$ and the correction factor in (15) is $\theta = 1$ ($\mu = \infty$). At the same time according to Theorem 1 the susceptible curve (blue) is the same for both models ($S = x_S$).

and therefore (applying the initial conditions (7)) we get

$$x_D = \mu \left(1 + \tilde{\rho} - x_S - \ln x_S / \mathcal{R}_0^{NET}\right). \tag{12}$$

Substituting this expression into (9) we obtain the following differential equation which involves only x_S

$$-\dot{x}_S = \mu\tilde{\beta} \left(1 + \tilde{\rho} - x_S - \ln x_S / \mathcal{R}_0^{NET}\right) x_S. \tag{13}$$

The equation above inherits the initial condition $x_S(0) = 1$ from (6) and is seen as describing the rate of decay of the susceptible population in terms of the number of susceptibles only. It is especially relevant to analyzing new infections or the *incidence rate* of an epidemic. The quantity $-\dot{x}_S$ is therefore often referred to as the *epidemic curve*, see, e.g., [11,15]. If we think about the quantity $x_S(t)$ as the survival probability for susceptibles (that is, the probability of a randomly selected initially susceptible avoiding infection by time $t > 0$, see, e.g., [13]) then the epidemic curve may be also viewed as the *density of infection*, see, e.g., [14]. An example of the survival curve and the corresponding density of infection (or epidemic curve) is presented in Figure 1.

The equation (13) can be also used as a convenient way of comparing the dynamics of the Poisson network SIR model (9)-(11) with that of the classical (mass action) SIR model (1)-(3). Indeed, note that since the algebraic structures of the right hand sides of the differential equations (9)-(10) and (1)-(2) are identical, the same manipulation as for the former may also be used for the latter, leading to the mass action epidemic curve equation

$$-\dot{S} = \beta \left(1 + \rho - S - \ln x_S / \mathcal{R}_0^{MA}\right) S. \tag{14}$$

The equations (14) and (13) both describe the epidemic curve, they use however different assumptions on the population contact structure, namely the mass action and Poisson degree distribution, respectively. The two equations are seen to be equivalent if and only if the appropriate parameter values coincide. Note also that in this case $S = x_S$ but $I \neq x_I$. Let us formulate this as follows.

Theorem 1. *Assume that we wish to approximate the dynamics of a Poisson network SIR epidemic given by* (9)-(11) *and* (7) *using the classical SIR mass action model* (1)-(3) *and* (4). *The two models respective epidemic curve equations coincide iff* $\beta = \mu\tilde{\beta}$, $\rho = \tilde{\rho}$, *and* $\gamma = \tilde{\gamma} + \tilde{\beta}$. *In this case* $S = x_S$, $I = x_D/\mu$, *and* $\mathcal{R}_0^{MA} = \mathcal{R}_0^{NET} = \mathcal{R}_0$. *Moreover, if* $\mathcal{R}_0 < \mu$ *the true amount of infection under the network model* (x_I) *satisfies*

$$\dot{x}_I = -\dot{x}_S - \gamma\theta x_I \tag{15}$$

where the correction factor θ *is given by*

$$\theta = 1 - \mathcal{R}_0/\mu.$$

The result above indicates that it is reasonable to analyze the epidemic curve on the Poisson network using the classical SIR equations (1)-(3). This will lead to the correct dynamics of the incidence (infection density) given by x_S curve. However, in order to correctly calculate the number of infections x_I, the classical equation (2) needs to be replaced by the corrected one as given in (15). The discrepancy will depend on the values of the parameters but may be considerable as illustrated in Figure 2. However, for large μ we see that $\theta \approx 1$ and the discrepancy becomes negligible in practice.

III. Conclusions

Network-based epidemic models often employ random graphs with Poisson degree distributions. This choice is grounded in the widely used Poisson approximation of the binomial distribution, as applied to the basic Erdös-Rényi graph model. Personal network analysis based on such random graphs has gained prominence in modeling disease spread, a relevance accentuated by the recent COVID-19 pandemic. In this note, we pointed out to the surprising equivalence between the classical Kermack and McKendrick SIR model and its Poisson network counterpart. Our result stated in Theorem 1 indicates that SIR epidemics for Poisson networks, particularly those with a high degree, can be effectively represented using the traditional mass-action SIR dynamics. This compatibility is evident in epidemic curves and infection magnitude analyses as illustrated in our Theorem 1. The observed congruence between the classical and network-based SIR models offers a valuable perspective for models simplification especially for the high-degree Poisson networks that became especially relevant in assessing proximity-based transmission during COVID-19 pandemic. While the insights presented here are directly relevant to epidemiology, they also have implications in other fields, including social networks analysis and modeling dynamics of experimental systems in chemical physics.

References

[1] A. Volkening, D. F. Linder, M. A. Porter, G. A. Rempala, Forecasting elections using compartmental models of infection, *SIAM Review*, 62(4):837–865, 2020.

[2] H.-W. Kang, W. R. KhudaBukhsh, H. Koeppl, G. A. Rempała, Quasi-steady-state approximations derived from the stochastic model of enzyme kinetics, *Bulletin of Mathematical Biology*, 81:1303–1336, 2019.

[3] R. Van Der Hofstad, *Random graphs and complex networks*, Vol. 43, Cambridge University Press, 2016.

[4] I. Z. Kiss, J. C. Miller, P. L. Simon, *Mathematics of epidemics on networks*, Springer, Cham, 2017.

[5] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London, Series A*, 115:700–721, 1927.

[6] D. Rand, Chapter 4: Correlation equations and pair approximations for spatial ecologies, *Advanced Ecological Theory: Principles and Applications*, John Wiley and Sons, 1999.

[7] M. J. Keeling, The effects of local spatial structure on epidemiological invasions, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266:859–867, 1999.

[8] E. M. Volz, SIR dynamics in random networks with heterogeneous connectivity, *Journal of Mathematical Biology*, 56:293–310, 2008.

[9] K. A. Jacobsen, M. G. Burch, J. H. Tien, G. A. Rempała, The large graph limit of a stochastic epidemic model on a dynamic multilayer network, *Journal of Biological Dynamics*, 12:746–788, 2018.

[10] W. R. KhudaBukhsh, B. Choi, E. Kenah, G. A. Rempała, Survival dynamical systems: individual-level survival analysis from population-level epidemic models, *Interface Focus*, 10:20190048, 2020.

[11] W. R. KhudaBukhsh, C. D. Bastian, M. Wascher, C. Klaus, S. Y. Sahai, M. H. Weir, E. Kenah, E. Root, J. H. Tien, G. A. Rempała, Projecting COVID-19 cases and hospital burden in Ohio, *Journal of Theoretical Biology*, 561:111404, 2023.

[12] L. Decreusefond, J.-S. Dhersin, P. Moyal, V. C. Tran, Large graph limit for an SIR process in random network with heterogeneous connectivity, *The Annals of Applied Probability*, 22:541–575, 2012.

[13] I. Z. Kiss, E. Kenah, G. A. Rempała, Necessary and sufficient conditions for exact closures of epidemic equations on configuration model networks, *Journal of Mathematical Biology*, 87:36, 2023.

[14] G. A. Rempała, W. R. KhudaBukhsh, Dynamical survival analysis for epidemic modeling, *Handbook of Visual, Experimental and Computational Mathematics: Bridges through Data*, Springer, Cham, 2023.

[15] H. Vossler, P. Akilimali, Y. Pan, W. R. KhudaBukhsh, E. Kenah, G. A. Rempała, Analysis of individual-level data from 2018–2020 Ebola outbreak in Democratic Republic of the Congo, *Scientific Reports*, 12:5534, 2022.