# Pairwise models for non-Markovian epidemics on networks

# Abstract of Ph.D. thesis

Zsolt Vizi

Supervisors:

János Karsai, Ph.D.
associate professor

Gergely Röst, Ph.D.
research associate professor

Doctoral School in Mathematics
and Computer Science
University of Szeged, Bolyai Institute

Szeged

2016.

# Introduction

Networks (or graphs) offer a flexible framework to explicitly incorporate various heterogeneities in how individuals interact within a population [1], [9]. This framework has led to a number of models where the strong assumptions of random mixing of the classical compartmental models can be relaxed [7], [14]. Due to the flexibility of the network approach, nodes can represent not only single individuals but also groups of individuals or locations. Similarly, links can represent contacts between individuals along which diseases can spread, or interactions between groups such as flight routes between different locations.

Most $SIR$ (susceptible-infected-recovered) models on networks assume that both the disease transmission and recovery process are Markovian [16]. The assumption of Markovianity is a strong simplifying assumption, as especially in the context of epidemiology, the period of infectiousness has paramount importance, and often this is approximated from the empirical distribution of observed infectious periods of various diseases by non-exponential distributions. However, there is renewed interest in modelling non-Markovian processes, such as epidemics on networks [5], [12], [17], [2]. A possible modelling approach involves mean-field approximations, which are based on the classical compartmental principles and pairwise models, which have been very successful in capturing the average behaviour of a stochastic epidemics on networks [15].

This thesis aims to extend the pairwise model from Markovian to non-Markovian epidemic dynamics where the infection process remains Markovian but the infectious period is taken from an arbitrary distribution. In addition, we want to perform the full mathematical analysis of the resulting systems, with focus on the positivity of solutions, associated reproduction numbers and the implicit relation concerning the final epidemic size and implement explicit stochastic simulations and numerical solvers to test the validity of these models.

This thesis is based on the following publications of the author:

- Kiss, I.Z., Röst, G. and Vizi, Z., 2015. Generalization of pairwise models to non-Markovian epidemics on networks. *Physical review letters*, 115(7), p.078701. http://dx.doi.org/10.1103/PhysRevLett.115.078701

- Röst, G., Vizi, Z. and Kiss, I.Z., 2015. Impact of non-Markovian recovery on network epidemics. In *Biomat 2015: Proceedings of the International Symposium on Mathematical and Computational Biology*

- Röst, G., Vizi, Z. and Kiss, I.Z., 2016. Pairwise approximation for SIR type network epidemics with non-Markovian recovery. *arXiv preprint* arXiv:1605.02933.

## Reproduction numbers and final size relations

The mean-field approximation for an $SIR$ type disease on homogeneous network with Markovian infection and recovery is described by the following system

$$\begin{aligned} [\dot{S}](t) &= -\tau \frac{n}{N}[S][I](t) \\ [\dot{I}](t) &= \tau \frac{n}{N}[S][I](t) - \gamma[I](t) \\ [\dot{R}](t) &= \gamma[I](t), \end{aligned} \tag{1}$$

where the network has $N$ nodes and uniform degree distribution $\langle k \rangle = n$, $\tau$ denotes the transmission rate, $\gamma$ is the recovery rate and expected number/proportion of susceptible, infected or recovered nodes at time $t$ are denoted by $[S](t), [I](t)$ and $[R](t)$ respectively. Introducing the notation $[XY](t)$ and $[XYZ](t)$ for the expected number of $X-Y$ links and $X-Y-Z$ triplets, respectively, we can write down the pairwise model as follows:

$$\begin{aligned} [\dot{S}](t) &= -\tau[SI](t) \\ [\dot{I}](t) &= \tau[SI](t) - \gamma[I](t) \\ [\dot{SS}](t) &= -2\tau \frac{n-1}{n} \frac{[SS](t)[SI](t)}{[S](t)}, \\ [\dot{SI}](t) &= \tau \frac{n-1}{n} \frac{[SS](t)[SI](t)}{[S](t)} - \tau \frac{n-1}{n} \frac{[SI](t)[SI](t)}{[S](t)} - \tau[SI](t) - \gamma[SI](t). \end{aligned} \tag{2}$$

An epidemic, which acts on a short temporal scale, may be described as a sudden outbreak of a disease that infects a substantial portion of the population in a region before it disappears. Epidemics usually leave many members untouched. The number of untouched individuals appears in the final size relation, that gives a relationship between the size of the epidemic (number of members of the population who are infected over the course of the epidemic) and the associated reproduction number.

Reproduction numbers play a crucial role in mathematical epidemiology and are defined as the expected number of secondary infections caused by a 'typical' infected individual during its infectious period when placed in a fully susceptible population, which is a definition understood at the level of individuals.

Clearly, the mean-field model is written at the level of nodes and study the disease spread between susceptible and infected nodes. Similarly, the pairwise model is written at the level of links and describes the dynamics of susceptible $(S-S)$ and infected $(S-I)$ links. These remarks lead to the definitions of *basic* and *pairwise* reproduction numbers. More precisely, we distinguish the following two useful quantities:

(a) the *basic* reproduction number is the expected lifetime of an $I$ **node** multiplied by the number of newly infected **nodes** per unit time (denoted by $\mathcal{R}_0$);

(b) the *pairwise* reproduction number is the expected lifetime of an $S-I$ **link** multiplied by the number of newly generated $S-I$ **links** per unit time (denoted by $\mathcal{R}_0^p$).

The most important cases for $\mathcal{R}_0$ and $\mathcal{R}_0^p$ are summarised in Table 1.

| | $\mathcal{R}_0$ | $\mathcal{R}_0^p$ |
|---|---|---|
| Markovian | $\frac{n}{N}\frac{\tau}{\gamma}[S]_0$ | $\frac{n-1}{N}\frac{\tau}{\tau+\gamma}[S]_0$ |
| Fixed | $\frac{n}{N}\tau\sigma[S]_0$ | $\frac{n-1}{N}(1-e^{-\tau\sigma})[S]_0$ |
| General | $\frac{n}{N}\tau\mathbb{E}(\mathcal{I})[S]_0$ | $\frac{n-1}{N}\left(1-\mathcal{L}[f_{\mathcal{I}}](\tau)\right)[S]_0$ |

Table 1: Basic and pairwise reproduction numbers for different recovery distributions. $\mathcal{L}[f_{\mathcal{I}}](\tau)$ denotes the Laplace transform of $f_{\mathcal{I}}$, the density of the recovery process, at $\tau$.

We can reduce $d[I]/d[S]$ in Eq. (1) and integrate it to obtain

$$\ln\left(\frac{[S]_\infty}{[S]_0}\right) = \mathcal{R}_0\left(\frac{[S]_\infty}{[S]_0} - 1\right).$$

We will use the notation $s_\infty = \frac{[S]_\infty}{[S]_0}$. Clearly, attack rate is $1 - s_\infty$. Using these formulae, we have

$$\ln s_\infty = \mathcal{R}_0\left(s_\infty - 1\right). \tag{3}$$

This equation is called *final size relation* and gives an implicit equation for the proportion of remaining individuals after the disease outbreak. Clearly, larger the reproduction number, smaller the $s_\infty$ (thus larger the attack rate). For (2) as we will show for general case later, a significantly longer calculation yields

$$\frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n-1}} = \mathcal{R}_0^p\left(s_\infty^{\frac{n-1}{n}} - 1\right). \tag{4}$$

# Impact of distribution on disease spread

Note, while $\mathcal{R}_0$ depends only on the expected value, (see Table 1, case 'General'), the pairwise reproduction number $\mathcal{R}_0^p$ uses the complete density function, thus the average length of infectious period does not determine exactly the reproduction number. It implies that for an epidemic we have to know as precisely as possible the shape of the distribution. In the following, we consider some special distributions.

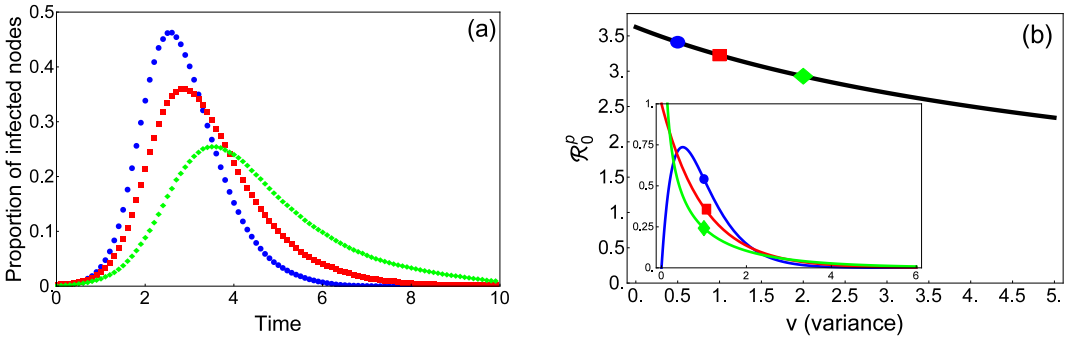


Figure 1: **(a)** Epidemic curves as averages of explicit stochastic simulations for non-Markovian epidemics, where the transmission rate is $\tau = 0.3$ and the initial number of susceptibles is $[S]_0 = 999$ on a homogeneous network with $N = 1000$ nodes and degree $n = 15$. The circles/squares/diamonds correspond to simulations for gamma distributed recovery time with parameters $(a, b) = (2, 0.5)/(1, 1)/(0.5, 2)$, respectively. **(b)** The solid curve shows the reproduction number $\mathcal{R}_0^p$ as a function of variance $v$ for fixed $m = 1$, and the circle/square/diamond represent the cases simulated in Fig. (a). In the inset figure, the shapes of the three corresponding probability density functions are presented.

The gamma distribution is one of the most commonly used distributions in the epidemiology literature to approximate empirically observed latent periods and infectious periods. It is applied in a wide spectrum of models, because of its flexibility and the possibility of incorporating it into ordinary differential equation models by the method of stages. If the infectious period $\mathcal{I}$ is gamma distributed with shape parameter $a$ and scale parameter $b$, that is $\mathcal{I} \sim \text{Gamma}(a, b)$, the following proposition can be proved:

**Proposition 4.1.1.** *Consider two random variables* $\mathcal{I}_1 \sim \text{Gamma}(a_1, b_1)$ *and* $\mathcal{I}_2 \sim \text{Gamma}(a_2, b_2)$ *such that* $\mathbb{E}(\mathcal{I}_1) = \mathbb{E}(\mathcal{I}_2)$ *and* $\text{Var}(\mathcal{I}_1) \leq \text{Var}(\mathcal{I}_2)$. *If* $\mathcal{I}_1$ *and* $\mathcal{I}_2$ *represent the recovery time distribution, then for the corresponding reproduction numbers the relation* $\mathcal{R}_{0,\mathcal{I}_1}^p \geq \mathcal{R}_{0,\mathcal{I}_2}^p$ *holds (i.e. for gamma distributions with a given mean, the pairwise reproduction number is monotonically decreasing with respect to the variance).*

The monotonicity of the reproduction number in the variance is depicted in Fig. 1(b). For a fixed mean but different variances of the gamma distribution, we can observe different epidemic curves in Fig. 1(a).

Since its simplicity allows us to make explicit calculations, we outline how the reproduction number and the disease dynamics behave when the recovery time follows uniform distribution. Uniformly distributed incubation and infectious periods have been used for example, in the modelling of avian influenza. Let $\text{Uniform}(a, b)$ denote a uniform distribution corresponding to the interval $[a, b]$, where $a \geq 0, b > a$. The

same monotonicity property can be shown for uniform distribution, which is stated in the following proposition:

**Proposition 4.2.1.** *Consider two random variables* $\mathcal{I}_1 \sim \text{Uniform}(a_1, b_1)$ *and* $\mathcal{I}_2 \sim \text{Uniform}(a_2, b_2)$ *such that* $\mathbb{E}(\mathcal{I}_1) = \mathbb{E}(\mathcal{I}_2)$ *and* $\text{Var}(\mathcal{I}_1) \leq \text{Var}(\mathcal{I}_2)$*. If* $\mathcal{I}_1$ *and* $\mathcal{I}_2$ *represent the recovery time distribution, then for the corresponding reproduction numbers the relation* $\mathcal{R}^p_{0,\mathcal{I}_1} \geq \mathcal{R}^p_{0,\mathcal{I}_2}$ *holds (i.e. for uniform distributions with a given mean, the pairwise reproduction number is monotonically decreasing with respect to the variance).*

In general, we consider a random variable $\mathcal{I}$ corresponding to recovery times with probability density functions $f_{\mathcal{I}}(t)$, cumulative distribution function $F_{\mathcal{I}}(t) = \int_0^t f_{\mathcal{I}}(s)ds$ and integral function of CDF $\mathcal{F}_{\mathcal{I}}(t) := \int_0^t F_{\mathcal{I}}(s)ds$. The following theorem gives a sufficient condition for monotonicity of pairwise reproductio number $\mathcal{R}^p_0$ in variance.

**Theorem 4.4.1.** *Consider two random variables* $\mathcal{I}_1$ *and* $\mathcal{I}_2$ *such that*

$$\mathbb{E}(\mathcal{I}_1) = \mathbb{E}(\mathcal{I}_2) < \infty, \tag{5}$$

*and*

$$\text{Var}(\mathcal{I}_1) < \text{Var}(\mathcal{I}_2) < \infty. \tag{6}$$

*Let us assume, that*

$$\lim_{t\to\infty} t^3 f_{\mathcal{I}}(t) = 0 \tag{7}$$

*and for all* $t > 0$*,*

$$\mathcal{F}_{\mathcal{I}_1}(t) \neq \mathcal{F}_{\mathcal{I}_2}(t). \tag{8}$$

*holds. If* $\mathcal{I}_1$ *and* $\mathcal{I}_2$ *represent the recovery time distribution, then for the corresponding reproduction numbers the relation* $\mathcal{R}^p_{0,\mathcal{I}_1} > \mathcal{R}^p_{0,\mathcal{I}_2}$ *holds.*

# Models with fixed recovery time

We consider a non-Markovian epidemic process with fixed recovery time denoted by $\sigma$. If the infection process is assumed to be Markovian, the equations for $[\dot{S}](t)$ and $[\dot{SS}](t)$ will be the same as in Eq. (2). Through this section we assume, that all initial infected nodes are newborn at $t = 0$, thus there is no recovery for $0 \leq t < \sigma$ and we have different equations describing changing of $[I](t)$ and $[SI](t)$ for $0 \leq t < \sigma$

and for $t > \sigma$. After some calculations, the following pairwise model can be derived for fixed recovery time:

$$[\dot{S}](t) = -\tau[SI](t), \tag{9a}$$
$$[\dot{I}](t) = \tau[SI](t), \tag{9b}$$
$$[\dot{SS}](t) = -2\tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)}, \tag{9c}$$
$$[\dot{SI}](t) = \tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)} - \tau[SI](t) - \tau\frac{n-1}{n}\frac{[SI](t)[SI](t)}{[S](t)} \tag{9d}$$

holds for $0 \leq t < \sigma$ and

$$[\dot{S}](t) = -\tau[SI](t), \tag{10a}$$
$$[\dot{I}](t) = \tau[SI](t) - \tau[SI](t-\sigma) \tag{10b}$$
$$[\dot{SS}](t) = -2\tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)}, \tag{10c}$$
$$\begin{aligned}[\dot{SI}](t) = {} & \tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)} - \tau[SI](t) - \tau\frac{n-1}{n}\frac{[SI](t)[SI](t)}{[S](t)} \\ & - \tau\frac{n-1}{n}\frac{[SS](t-\sigma)[SI](t-\sigma)}{[S](t-\sigma)}e^{-\int_{t-\sigma}^{t}\tau\frac{n-1}{n}\frac{[SI](u)}{[S](u)}+\tau du}.\end{aligned} \tag{10d}$$

satisfies for $t > \sigma$. The Eq. (9) is a system of ordinary differential equations, given initial values $[S]_0, [I]_0$, $[SS]_0$ and $[SI]_0$ at $t = 0$ are sufficient to guarantee a unique solution. Let us denote the solution of (9) on the time interval $[0, \sigma]$ by

$$X^*(t) = ([S]^*(t), [I]^*(t), [SS]^*(t), [SI]^*(t)).$$

At time $t = \sigma$, the initial infected nodes recover 'instantly', thus a discontinuity appears and obviously, the solution for $t > \sigma$ starts from

$$\widetilde{X} = ([S]^*(\sigma), [I]^*(\sigma) - [I]_0, [SS]^*(\sigma), [SI]^*(\sigma) - [SI]_0).$$

Similarly to Markovian case, the non-Markovian mean-field model for fixed infectious period is

$$[\dot{S}](t) = -\tau\frac{n}{N}[S](t)[I](t), \tag{11a}$$
$$[\dot{I}](t) = \tau\frac{n}{N}[S](t)[I](t), \tag{11b}$$

for $0 \leq t < \sigma$ and

$$[\dot{S}](t) = -\tau\frac{n}{N}[S](t)[I](t), \tag{12a}$$
$$[\dot{I}](t) = \tau\frac{n}{N}[S](t)[I](t) - \tau\frac{n}{N}[S](t-\sigma)[I](t-\sigma), \tag{12b}$$

for $t > \sigma$. Here, if we denote the solution of (11) for initial values $[S]_0, [I]_0$ and for time interval $t \in [0, \sigma]$ by $X_m^*(t) = ([S]^*(t), [I]^*(t))$, the initial function associated to (12) is $X_m^*(t)$ for $0 \leq t < \sigma$ and $([S]^*(\sigma), [I]^*(\sigma) - [I]_0)$ at $t = \sigma$.

First, we can find a first integral of the pairwise model (9)-(10), which allows us to reduce the dimensionality.

**Proposition 5.2.1.** *The function* $U(t) = \frac{[SS](t)}{[S]^{2\frac{n-1}{n}}(t)}$ *is a first integral of system (9)-(10).*

Consequently, using this first integral, we obtain

$$[SS](t) = \frac{n}{N}[S]_0^{\frac{2}{n}}[S]^{2\frac{n-1}{n}}(t). \tag{13}$$

Applying Eq.(13), we can reduce our pairwise model to a two-dimensional system:

$$\begin{aligned} [\dot{S}](t) &= -\tau[SI](t), \\ [\dot{SI}](t) &= \tau\frac{n-1}{N}[S]_0^{\frac{2}{n}}[S]^{\frac{n-2}{n}}(t)[SI](t) - \tau[SI](t) - \tau\frac{n-1}{n}\frac{[SI](t)}{[S](t)}[SI](t) \\ &\quad -\tau\frac{n-1}{N}[S]_0^{\frac{2}{n}}[S]^{\frac{n-2}{n}}(t-\sigma)[SI](t-\sigma)e^{-\int_{t-\sigma}^{t}\tau\frac{n-1}{n}\frac{[SI](u)}{[S](u)}+\tau du}. \end{aligned} \tag{14}$$

We are interested only in nonnegative solutions of system (9)-(10). The following proposition shows, that the solutions remain nonnegative provided that the initial conditions are nonnegative.

**Proposition 5.2.2.** *If initial conditions* $[S]_0, [SS]_0, [I]_0$ *and* $[SI]_0$ *for (9) and (11) are nonnegative, then* $[S](t) \geq 0$, $[SS](t) \geq 0$, $[I](t) \geq 0$ *and* $[SI](t) \geq 0$ *hold for* $t \geq 0$ *in both mean-field model (11)-(12) and pairwise model (9)-(10).*

For exploring the relation between disease outbreak and reproduction number, we start with the following definition:

**Definition 1.** *In a disease transmission model with no demographic effects,* there is no epidemic *if the equilibrium with all members of the population susceptible is (locally) asymptotically stable, and* there is an epidemic *if this equilibrium is unstable, in each case considering only perturbations of the equilibrium with positive infected initial states.*

Using this concept, we state the following theorems:

**Theorem 5.2.1.** *There is an epidemic for the model (11)-(12) if and only if* $\mathcal{R}_0 > 1$, *where the basic reproduction number is* $\mathcal{R}_0 = \tau\frac{n}{N}[S]_0\sigma$.

Similar procedure can be done for the pairwise model (9)-(10) and the result is summarised in the following theorem:

**Theorem 5.2.2.** *There is an epidemic for the model (9)-(10) if and only if* $\mathcal{R}_0^p > 1$*, where the pairwise reproduction number is* $\mathcal{R}_0 = \frac{n-1}{N}[S]_0(1 - e^{-\tau\sigma})$.

Finally, we derive final size relations that allow us to calculate the total number of infected nodes during an epidemic outbreak on the network.

**Theorem 5.2.3.** *The final size relation associated to the mean-field model (11)-(12) is*

$$\ln(s_\infty) = \mathcal{R}_0(s_\infty - 1), \tag{15}$$

*where the basic reproduction number is* $\mathcal{R}_0 = \tau\frac{n}{N}[S]_0\sigma$.

The main result for fixed recovery time is the derivation of the final-size relation for the pairwise system (9)-(10).

**Theorem 5.2.4.** *The final size relation associated to the pairwise model (9)-(10) is*

$$\frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n-1}} = \mathcal{R}_0^p\left(s_\infty^{\frac{n-1}{n}} - 1\right), \tag{16}$$

*where the pairwise reproduction number* $\mathcal{R}_0^p = \frac{n-1}{N}[S]_0(1 - e^{-\tau\sigma})$.

# General recovery time

We want to build mean-field and pairwise models for the $SIR$ type epidemic process with exponentially distributed transmission and general recovery time distribution. First, let $i(t,a)$ represent the density of infected nodes with respect to the age of infection $a$ at the current time $t$, then $[I](t) = \int_0^\infty i(t,a)da$. Similarly, $Si(t,a)$ and $ISi(t,a)$ describe the density of $S-i$ links and $I-S-i$ triplets, respectively, where the infected node $i$ has age $a$ at time $t$ and $[SI](t) = \int_0^\infty Si(t,a)da$, $[ISI](t) = \int_0^\infty ISi(t,a)da$. We assume that the infection process along $S-I$ links is Markovian with transmission rate $\tau > 0$. The recovery part is considered to be non-Markovian, with a cumulative distribution function $F_{\mathcal{I}}(a)$ and probability density function $f_{\mathcal{I}}(a)$. We use the associated survival function $\xi_{\mathcal{I}}(a) = 1 - F_{\mathcal{I}}(a)$ and hazard function $h_{\mathcal{I}}(a) = -\frac{\xi'_{\mathcal{I}}(a(a)}{\xi_{\mathcal{I}}(a)} = \frac{f_{\mathcal{I}}(a)}{\xi_{\mathcal{I}}(a)}$. Using the notations above, we arrive at the following

model

$$[\dot{S}](t) = -\tau[SI](t), \tag{17a}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) i(t,a) = -h_{\mathcal{I}}(a) i(t,a), \tag{17b}$$

$$[\dot{SS}](t) = -2\tau[SSI](t), \tag{17c}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) Si(t,a) = -\tau ISi(t,a) - (\tau + h_{\mathcal{I}}(a)) Si(t,a), \tag{17d}$$

subject to the boundary conditions

$$i(t,0) = \tau[SI](t), \tag{18a}$$

$$Si(t,0) = \tau[SSI](t), \tag{18b}$$

and initial conditions

$$[S](0) = [S]_0, [SS](0) = [SS]_0, i(0,a) = \varphi(a), \tag{19a}$$

$$Si(0,a) = \chi(a) \approx \frac{n}{N}[S]_0 i(0,a) = \frac{n}{N}[S]_0 \varphi(a). \tag{19b}$$

We shall use the biologically feasible assumption $\lim_{a\to\infty} \varphi(a) = 0$. To break the dependence on higher order moments, we apply the closure approximation formula

$$[XYZ] = \frac{n-1}{n}\frac{[XY][YZ]}{[Y]}. \tag{20}$$

for $ISi(t,a)$ in the form

$$ISi(t,a) = \frac{n-1}{n}\frac{[SI](t) Si(t,a)}{[S](t)}. \tag{21}$$

To obtain a self-consistent system for classical network variables $[S]$, $[SS]$, $[I]$ and $[SI]$, further calculations are needed. The resulting pairwise system is the following integro-differential equation:

$$[\dot{S}](t) = -\tau[SI](t) \tag{22a}$$

$$[\dot{SS}](t) = -2\tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)} \tag{22b}$$

$$[\dot{I}](t) = \tau[SI](t) - \int_0^t \tau[SI](t-a) f_{\mathcal{I}}(a) da - \int_t^\infty \varphi(a-t)\frac{f_{\mathcal{I}}(a)}{\xi_{\mathcal{I}}(a-t)} da \tag{22c}$$

$$\begin{aligned}[\dot{SI}](t) = {} & \tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)} - \tau\frac{n-1}{n}\frac{[SI](t)}{[S](t)}[SI](t) - \tau[SI](t) \\ & - \int_0^t \tau\frac{n-1}{n}\frac{[SS](t-a)[SI](t-a)}{[S](t-a)} e^{-\int_{t-a}^t \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)} + \tau ds} f_{\mathcal{I}}(a) da \\ & - \int_t^\infty \frac{n}{N}[S]_0 \varphi(a-t) e^{-\int_0^t \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)} + \tau ds} \frac{f_{\mathcal{I}}(a)}{\xi_{\mathcal{I}}(a-t)} da.\end{aligned} \tag{22d}$$

From Eq.(22), the associated mean-field model can be easily deduced by using the closure approximation formula for homogeneous networks

$$[XY](t) = \frac{n}{N}[X](t)[Y](t), \tag{23}$$

thus the node-level system becomes

$$[\dot{S}](t) = -\tau\frac{n}{N}[S](t)[I](t) \tag{24a}$$

$$[\dot{I}](t) = \tau\frac{n}{N}[S](t)[I](t) - \int_0^t \tau\frac{n}{N}[S](t-a)[I](t-a)f_{\mathcal{I}}(a)da - \int_t^\infty \varphi(a-t)\frac{f_{\mathcal{I}}(a)}{\xi_{\mathcal{I}}(a-t)}da. \tag{24b}$$

Note, that Prop. 5.2.1.. holds for the pairwise system, thus we can reduce the model (22) to the following two-dimensional system:

$$\begin{aligned}
[\dot{S}](t) &= -\tau[SI](t), \\
[\dot{SI}](t) &= \tau\frac{n-1}{N}[S]_0^{\frac{2}{n}}[S]^{\frac{n-2}{n}}(t)[SI](t) - \tau[SI](t) - \tau\frac{n-1}{n}\frac{[SI](t)}{[S](t)}[SI](t) \\
&\quad - \int_0^t \tau\frac{n-1}{N}[S]_0^{\frac{2}{n}}[S]^{\frac{n-2}{n}}(t-a)[SI](t-a)e^{-\int_{t-a}^t \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)}+\tau ds}f_{\mathcal{I}}(a)da \\
&\quad - \int_t^\infty \frac{n}{N}[S]_0\varphi(a-t)e^{-\int_0^t \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)}+\tau ds}\frac{f_{\mathcal{I}}(a)}{\xi_{\mathcal{I}}(a-t)}da.
\end{aligned} \tag{25}$$

The first proposition of this section states, that the solutions remain nonnegative provided that the initial conditions are nonnegative.

**Proposition 6.2.1.** *If initial conditions $[S]_0$, $[SS]_0$ are nonnegative and $\varphi(a) \geq 0$ for $a \geq 0$, then $[S](t) \geq 0$, $[SS](t) \geq 0$, $[I](t) \geq 0$ and $[SI](t) \geq 0$ hold for $t \geq 0$.*

We can prove, that the functional forms (15) and (16) hold for arbitrary recovery time distribution.

**Theorem 6.2.1.** *The final size relation associated to the mean-field model (24) is*

$$\ln(s_\infty) = \mathcal{R}_0(s_\infty - 1), \tag{26}$$

*where the basic reproduction number $\mathcal{R}_0 = \frac{n}{N}\tau[S]_0\mathbb{E}(\mathcal{I})$.*

Finally, a lengthy calculation gives the final-size relation for the pairwise system (22):

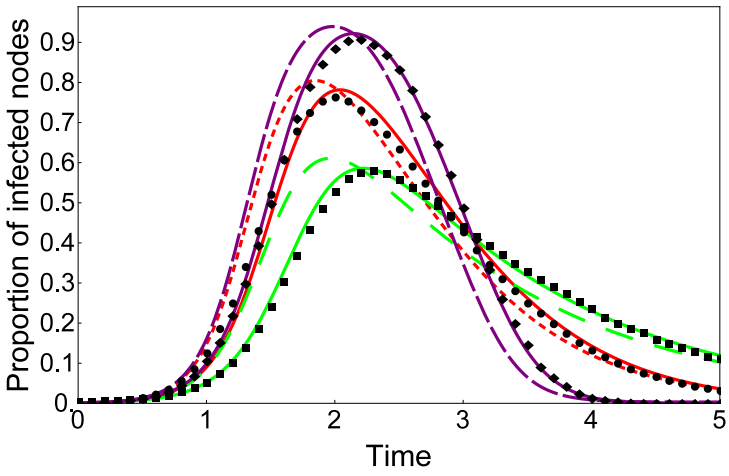


Figure 2: Stochastic and numerical experiments for non-Markovian epidemic with various recovery time distributions on homogeneous networks with $N = 1000$ nodes and infection rate $\tau = 0.35$. Squares, circles, diamonds show the mean of 100 simulations on random regular graphs with average degree $\langle k \rangle = 15$ for exponential distribution with parameter $\lambda = \frac{2}{3}$ (mean $= \frac{3}{2}$,variance $= \frac{9}{4}$), gamma distribution with shape $\alpha = 3$ and rate $\beta = 2$ (mean $= \frac{3}{2}$,variance $= \frac{3}{4}$), uniform distribution on interval $[a, b] = [1, 2]$ (mean $= \frac{3}{2}$,variance $= \frac{1}{12}$), respectively. Dashed and solid lines correspond to the numerical solution of the mean-field (24) and parwise (22) models, respectively.

**Theorem 6.2.2.** *The final size relation associated to the pairwise model (22) is*

$$\frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n-1}} = \mathcal{R}_0^p \left( s_\infty^{\frac{n-1}{n}} - 1 \right),$$

*where the pairwise reproduction number is* $\mathcal{R}_0^p = \frac{n-1}{N} \left(1 - \mathcal{L}[f_{\mathcal{I}}](\tau)\right) [S]_0$.

For the numerical solution of integro-differential equations (22) and (24), we developed a numerical scheme based on collocation method. The numerical methods in [3] were adapted to the mean-field model and the reduced, but highly nonlinear pairwise system. We implemented the developed recursive algorithm and solved the Eqs. (24) and (22) with it. In Fig. 2, homogeneous (or regular random) networks were considered and the average of 100 simulations is compared to the numerical solutions of mean-field (24) and pairwise (22) models.

At last, we can investigate some common choices for the recovery time. As we expect, for 'newborn' initial infecteds, if $\mathcal{I} \sim \text{Exp}(\gamma)$ (i.e. the infectious period $\mathcal{I}$ is exponentially distributed), we get back the classical Markovian models (1) and (2). In the case of fixed recovery time, the models reduce to the systems (11)-(12) and (9)-(10). We can also recover the multi-stage infection model of [13] with gamma distributed recovery time. For uniform distribution $\mathcal{I} \sim \text{Uniform}(A, B)$ we can write

down the associated equations:

$$[\dot{I}](t) = \tau[SI](t) - \int_{\max(0,t-B)}^{\max(0,t-A)} \frac{\tau[SI](u)}{B-A} du - \frac{[I]_0}{B-A}\iota_{[A,B]}(t),$$

where $\iota_{[A,B]}(t)$ is the indicator function of interval $[A, B]$. The same argument gives

$$\begin{aligned}
[\dot{SI}](t) &= \tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)} - \tau\frac{n-1}{n}\frac{[SI](t)}{[S](t)}[SI](t) - \tau[SI](t) \\
&\quad - \int_{\max(0,t-B)}^{\max(0,t-A)} \frac{\tau}{B-A}\frac{n-1}{n}\frac{[SS](u)[SI](u)}{[S](u)} e^{-\int_u^t \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)}+\tau ds} du \\
&\quad - \frac{n}{N}[S]_0 e^{-\int_0^t \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)}+\tau ds}\frac{[I]_0}{B-A}\iota_{[A,B]}(t).
\end{aligned}$$

For $t > B$ the model becomes a system of differential equations with distributed delays.

# References


[1] Barabási, A.L. and Albert, R., 1999. Emergence of scaling in random networks. *science*, 286(5439), pp.509-512.

[2] Boguñá, M., Lafuerza, L.F., Toral, R. and Serrano, M.Á., 2014. Simulating non-Markovian stochastic processes. *Physical Review E*, 90(4), p.042108.

[3] Brunner, H., 2004. Collocation methods for Volterra integral and related functional differential equations (Vol. 15). *Cambridge University Press.*

[4] Burton, T.A., 2005. *Volterra integral and differential equations* (Vol. 202). Elsevier.

[5] Cooper, F., 2013. Non-Markovian network epidemics (*MA Thesis*, 2013. www. dtc. ox. ac. uk/people/13/cooperf/files/MA469ThesisFergusCooper. pdf).

[6] Smith, H., 2010. *An introduction to delay differential equations with applications to the life sciences* (Vol. 57). Springer Science & Business Media.

[7] Keeling, M.J. and Eames, K.T., 2005. Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4), pp.295-307.

[8] Kiss, I.Z., Röst, G. and Vizi, Z., 2015. Generalization of pairwise models to non-Markovian epidemics on networks. *Physical review letters*, 115(7), p.078701.

[9] Newman, M.E., 2002. Spread of epidemic disease on networks. *Physical review E*, 66(1), p.016128.

[10] Röst, G., Vizi, Z. and Kiss, I.Z., 2015. Impact of non-Markovian recovery on network epidemics. In *Biomat 2015: Proceedings of the International Symposium on Mathematical and Computational Biology*.

[11] Röst, G., Vizi, Z. and Kiss, I.Z., 2016. Pairwise approximation for SIR type network epidemics with non-Markovian recovery. *arXiv preprint* arXiv:1605.02933. Vancouver

[12] Sharkey, K.J., 2008. Deterministic epidemiological models at the individual level. *Journal of Mathematical Biology*, 57(3), pp.311-331.

[13] Sherborne, N., Blyuss, K.B. and Kiss, I.Z., 2015. Dynamics of multi-stage infections on networks. *Bulletin of mathematical biology*, 77(10), pp.1909-1933.

[14] Simon, P.L., Taylor, M. and Kiss, I.Z., 2011. Exact epidemic models on graphs using graph-automorphism driven lumping. *Journal of mathematical biology*, 62(4), pp.479-508.

[15] M. Taylor, P. L. Simon, D. M. Green, T. House, and I. Z. Kiss. From Markovian to pairwise epidemic models and the performance of moment closure approximations. *J. Math. Biol.*, 64:1021-1042, 2012.

[16] Wilkinson, R.R. and Sharkey, K.J., 2014. Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks. *Physical Review E*, 89(2), p.022808.

[17] Wilkinson, R.R., Ball, F.G. and Sharkey, K.J., 2016. The relationships between message passing, pairwise, Kermack-McKendrick and stochastic SIR epidemic models. *arXiv preprint* arXiv:1605.03555.