PHYSICAL REVIEW LETTERS

Generalization of Pairwise Models to non-Markovian Epidemics on Networks

Istvan Z. Kiss[*]

*Department of Mathematics, School of Mathematical and Physical Sciences, University of Sussex,
Falmer, Brighton BN1 9QH, United Kingdom*

Gergely Röst[†] and Zsolt Vizi[‡]

Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged 6720, Hungary

In this Letter, a generalization of pairwise models to non-Markovian epidemics on networks is presented. For the case of infectious periods of fixed length, the resulting pairwise model is a system of delay differential equations, which shows excellent agreement with results based on stochastic simulations. Furthermore, we analytically compute a new $\mathcal{R}_0$-like threshold quantity and an analytical relation between this and the final epidemic size. Additionally, we show that the pairwise model and the analytic results can be generalized to an arbitrary distribution of the infectious times, using integro-differential equations, and this leads to a general expression for the final epidemic size. By showing the rigorous link between non-Markovian dynamics and pairwise delay differential equations, we provide the framework for a more systematic understanding of non-Markovian dynamics.

Networks have provided a step change in modeling complex systems [1–3]. The study of disease transmission on networks has greatly benefitted from this modeling paradigm by uncovering the role and impact of contact heterogeneity [3]. While networks provide a clear departure from classic compartmental models, the role of mean-field models remains crucial. These offer us a reliable way to obtain analytical results, such as epidemic threshold [4,5] and final epidemic size [6], which in turn can be used to uncover the interplay between network properties and dynamic processes on networks.

Probably the most well-known mean-field model for network epidemics is the degree-based or heterogeneous mean-field model [3,4]. Similarly, pairwise models [6–9] continue to provide a fruitful framework for modeling adaptive networks involving epidemics [8,10], social interactions [11] and ecological systems [9]. Mapping out the full spectrum of possible system behaviors and analytical results are made possible by such models.

However, there is renewed interest in non-Markovian processes, such as epidemics on networks [12–18], random walks [19], and temporal networks [20]. For example, Min *et al.* [13] consider the *SIR* model with fixed recovery and an infectious process with heavy-tail distribution. Combining renewal theory with branching processes they show that as the exponent of the power law tends to 2, only disease without recovery can spread. The non-Markovian *SIS* model is considered in Refs. [12,15]. In Ref. [15], the authors consider Poisson recovery and a general infectious process and show strong variations in epidemic prevalence

and threshold despite keeping the mean of the distribution equal. In Ref. [12], non-exponential distributions of infection and recovery times lead to the same functional form of the prevalence in the quasi-steady state as for the Markovian *SIS* model as long as the spreading rate is replaced by the average number of infection attempts per recovery time.

This recent burst of activity is motivated by empirical observations, where for many real world systems, the Markovian framework is not satisfactory in describing temporal statistics, such as time intervals between discrete, consecutive events. Examples include intertrade durations in financial markets [21], socionetworks [22], or contacts between individuals being dynamic [20]. In the context of epidemiology, the period of infectiousness has a key role [23,24]. The empirical distribution of infectious periods of various diseases is often approximated by log-normal and gamma (smallpox [25,26]), fixed-length (measles [27]) or Weibull distributions (ebola [28]). Unfortunately, the reliable tools and mathematical machinery of Markovian theory do not translate directly to modeling and analysis of non-Markovian systems, and this leads to many significant challenges.

In this Letter, we present the first analog of pairwise models for non-Markovian epidemics, and show that this is equivalent to a set of delay differential equations (DDEs) which (a) shows excellent agreement with simulation and (b) allows us to define a new $\mathcal{R}_0$-like quantity and to derive an implicit analytic expression for the final epidemic size. We consider an undirected and unweighted

network with N nodes and an average degree n. Each node can be susceptible (S), infected (I), and recovered (R). For Markovian epidemics, with transmission rate τ and recovery rate γ, the epidemic is well approximated by the pairwise model [6] given below:

$$[\dot{S}] = -\tau[SI]; \qquad [\dot{I}] = \tau[SI] - \gamma[I]; \qquad [\dot{SS}] = -2\tau[SSI];$$

$$[\dot{SI}] = \tau([SSI] - [ISI] - [SI]) - \gamma[SI],$$

where $[X]$, $[XY]$ and $[XYZ]$ are the expected number of nodes in state X, links in state X-Y and triples in state X-Y-Z, respectively. The dependence on higher order moments can be broken by using that $[XSY] = (\xi[XS][SY])/[S]$ [6], where $\xi = (n-1)/n$. Applying this leads to the following self-consistent system

$$[\dot{S}] = -\tau[SI]; \qquad [\dot{I}] = \tau[SI] - \gamma[I];$$

$$[\dot{SS}] = -2\tau\xi\frac{[SS][SI]}{[S]};$$

$$[\dot{SI}] = \tau\xi\left(\frac{[SS][SI]}{[S]} - \frac{[SI][SI]}{[S]}\right) - (\tau + \gamma)[SI]. \quad (1)$$

Closing at the level of pairs, $[XY] = (n[X][Y])/N$, this system reduces to the classic compartmental SIR model,

$$\dot{S} = -\tau\frac{n}{N}SI; \qquad \dot{I} = \tau\frac{n}{N}SI - \gamma I. \quad (2)$$

We wish to apply the previous approach to nonexponentially distributed recovery times. First, a fixed infectious period, denoted by σ, is considered, and the derivation of the pairwise model from first principles is illustrated. We show that non-Markovian dynamics can be described by a system of DDEs. The infection process is assumed to be Markovian; thus, the equation for $[S]$ is the same as before, i.e., $[\dot{S}](t) = -\tau[SI](t)$. The number of infected nodes at time t is replenished by $\tau[SI](t)$ and is depleted by $\tau[SI](t-\sigma)$, and this yields $[\dot{I}](t) = \tau[SI](t) - \tau[SI](t-\sigma)$. The equation for the number of $S - S$ links is the same because the infection process is Markovian, see Eq. (1). In a similar manner, the number of $S - I$ links is replenished by $\tau\xi[SS](t)[SI](t)/[S](t)$, which is the rate of depletion of $S - S$ links. Furthermore, depletion occurs due to the infection within $S-I$ pairs, $\tau[SI](t)$, and due to the infection of the S node from outside the pair, $\tau\xi[SI](t)[SI](t)/[S](t)$. On the other hand, there are $S - I$ links, which survive for time σ, but will be removed due to the recovery of the I node.

Next, we need to account for the removal of $S - I$ links which were created precisely σ times ago. Naively, one would believe that this term is simply proportional to $\tau\xi[SS](t-\sigma)[SI](t-\sigma)/[S](t-\sigma)$. However, one must take into consideration that, in the time interval $(t-\sigma, t)$, an $S - I$ link could have been destroyed either due to pair infection within or by infection of the S node from outside. Hence, a discount factor needs to be determined to capture

this effect. To calculate this factor, $S - I$ links, that are created at the same time, are considered as a cohort denoted by x, and we model infection within and from outside by writing down the following evolution equation:

$$\dot{x}(t) = -\frac{\tau\xi}{[S](t)}[SI](t)x(t) - \tau x(t), \quad (3)$$

where, the first term denotes the "outer" infection of the S node, while the second term stands for "inner" infection of the S node. We note that the outside infection is simply proportional to the probability that an S node with an already engaged link has a further infected neighbor, $\xi[SI]/[S]$. The solution of Eq. (3) in $[t-\sigma, t]$ is

$$x(t) = x(t-\sigma)e^{-\int_{t-\sigma}^{t}\left(\frac{\tau\xi}{[S](u)}[SI](u)+\tau\right)du},$$

and this provides the depletion or discount rate of $S - I$ links. In this case, $x(t-\sigma) = \tau\xi[SS](t-\sigma)[SI](t-\sigma)/[S](t-\sigma)$, which is the replenishment of $S - I$ links. Therefore, summarizing all the above, the pairwise DDE system with discrete and distributed delays for the non-Markovian case is

$$[\dot{S}](t) = -\tau[SI](t); \qquad [\dot{I}](t) = \tau[SI](t) - \tau[SI](t-\sigma)$$

$$[\dot{SS}](t) = -2\tau\xi\frac{[SS](t)[SI](t)}{[S](t)}; \qquad [\dot{SI}](t) = -\tau[SI](t)$$

$$- \tau\xi\left(\frac{[SI](t)[SI](t)}{[S](t)} - \frac{[SS](t)[SI](t)}{[S](t)}\right)$$

$$- \tau\xi\frac{[SS](t-\sigma)[SI](t-\sigma)}{[S](t-\sigma)}e^{-\int_{t-\sigma}^{t}\left([SI](u)\frac{\tau\xi}{[S](u)}+\tau\right)du}.$$

$$(4)$$

This system is now the main subject of our investigation from an analytical and numerical point of view. Similarly to the Markovian case, the non-Markovian mean-field model for the fixed infectious period is

$$\dot{S}(t) = -\tau\frac{n}{N}S(t)I(t),$$

$$\dot{I}(t) = \tau\frac{n}{N}S(t)I(t) - \tau\frac{n}{N}S(t-\sigma)I(t-\sigma). \quad (5)$$

The most important results for SIR models are the explicit formula of basic reproduction number, $\mathcal{R}_0$, and an implicit equation for the final epidemic size. In what follows, we introduce a general concept for the reproduction number associated with the pairwise model, and we refer to this as the *pairwise reproduction number*. Using this concept, the final size relations for the above mean-field, classic pairwise, and DDE-based pairwise models are derived. 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 nodes (individuals). On the other hand, the pairwise model is written at the level of links and describes the dynamics of susceptible ($S-S$) and infected ($S-I$) links. This leads to the definition of a new type of reproduction numbers, which we call the *pairwise reproduction number* and denote it by $\mathcal{R}_0^p$. 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, and (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.

The expected life time of an infectious node is the expected value of a random variable X corresponding to the distribution of the length of infectious periods. In contrast, an $S-I$ link can be removed either due to the recovery of the I node or the infection of the S node. Therefore, the expected lifetime of the $S-I$ link is the expected value of the minimum of two random variables. If we assume that the process of infection along such a link has density function f_i with survival function ξ_i, and the process of recovery has density function f_r with survival function ξ_r, then, denoting by Z the random variable defined by the lifetime of an $S-I$ link, we have

$$\mathbb{E}(Z) = \int_0^\infty t[f_i(t)\xi_r(t) + f_r(t)\xi_i(t)]dt. \quad (6)$$

From the assumption that the infection time along $S-I$ links is exponentially distributed (i.e., $f_i(t) = \tau e^{-\tau t}$, $\xi_i(t) = e^{-\tau t}$), the number of newly infected nodes per unit time in the mean-field and pairwise model are $n\tau[S]_0/N$ and $\tau\xi[SS]_0/[S]_0 = \tau(n-1)[S]_0/N$, respectively, where we used the approximation $[SS]_0 = n[S]_0^2/N$.

We illustrate Eq. (6) for infectious periods of fixed length (σ). In this case, the survival function is $\xi_r(t) = 1$ if $0 \le t < \sigma$ and $\xi_r(t) = 0$ if $t \ge \sigma$, and the density function $f_r(t)$ is the Dirac delta $\delta(t - \sigma)$. Using fundamental properties of the delta function, we have

$$\mathbb{E}(Z) = \left(-\sigma e^{-\tau\sigma} + \frac{1-e^{-\tau\sigma}}{\tau}\right) + \sigma e^{-\tau\sigma},$$

and multiplying this by the number of newly generated $S-I$ links, the formula in Table I for $\mathcal{R}_0^p$ follows. It is noteworthy to highlight that $\mathbb{E}(Z)$, in the case of a Markovian infection process, reduces to evaluating the Laplace transform of the density of the recovery time, see Supplemental Material [29]. This provides a general result, which in many cases leads to an analytical result for $\mathcal{R}_0^p$, see Table I.

For the standard Markovian mean-field model, the process of calculating the final epidemic size is well known. From Eq. (2), we evaluate dI/dS and integrate it to obtain

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

	$\mathcal{R}_0$	$\mathcal{R}_0^p$
Markovian	$(n/N)(\tau/\gamma)S_0$	$[(n-1)/N][\tau/(\tau+\gamma)][S]_0$
Fixed	$(n/N)\tau\sigma S_0$	$[(n-1)/N](1-e^{-\tau\sigma})[S]_0$
General	$(n/N)\tau\mathbb{E}(X)S_0$	$[(n-1)/N][1-\mathcal{L}[f_r](\tau)][S]_0$

$$\ln(S_\infty/S_0) = \mathcal{R}_0(S_\infty/S_0 - 1), \quad (7)$$

where $\mathcal{R}_0 = (\tau n[S]_0)/(\gamma N)$. The final epidemic size can be easily computed by using $R_\infty = N - S_\infty$. In the non-Markovian case, the calculations (which are included in the Supplemental Material [29]) are rather different and the resulting final size relation is

$$\ln(S_\infty/S_0) = (\tau n\sigma[S]_0/N)(S_\infty/S_0 - 1). \quad (8)$$

As in this case $\mathcal{R}_0 = \tau n\sigma[S]_0/N$, the final size relation given by Eq. (8) shows the "standard" form of Eq. (7). The dynamical systems, Eqs. (1) and (4), can be manipulated conveniently to derive an analytic relation between $\mathcal{R}_0^p$ and R_∞. This is known for the Markovian case but it is a new result for the non-Markovian one. While the full derivation for the non-Markovian case is given in the Supplemental Material [29], the main steps of the calculations are: (a) find an invariant to reduce the dimensionality of the system, (b) integrate the equation for $[SI](t)$, (c) integrate the equation for $[S](t)$ on $[0, \infty)$, and (d) employ algebraic manipulations to obtain the final size relation. This procedure yields

$$\frac{s_\infty^{1/n} - 1}{\frac{1}{n-1}} = \frac{n-1}{N}(1-e^{-\tau\sigma})[S]_0\left(s_\infty^{(n-1)/n} - 1\right), \quad (9)$$

where $s_\infty = ([S]_\infty/[S]_0)$ and the attack rate is simply $1 - s_\infty$. The same technique for the Markovian case leads to

$$\frac{s_\infty^{1/n} - 1}{\frac{1}{n-1}} = \frac{n-1}{N}\frac{\tau}{\tau+\gamma}[S]_0\left(s_\infty^{(n-1)/n} - 1\right). \quad (10)$$

Upon inspecting the two relations above, the following important observations can be made. First, the implicit relation between final size and $\mathcal{R}_0^p$ is conserved between the Markovian and non-Markovian model. Moreover, upon using the values of $\mathcal{R}_0^p$ as given in Table I, Eqs. (9) and (10) can be cast in the following general form:

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

$$= \frac{n-1}{N}[1-\mathcal{L}[f_r](\tau)][S]_0\left(s_\infty^{(n-1)/n} - 1\right). \quad (11)$$

The equation above holds true for fixed delay and the Markovian case and this has been shown analytically.
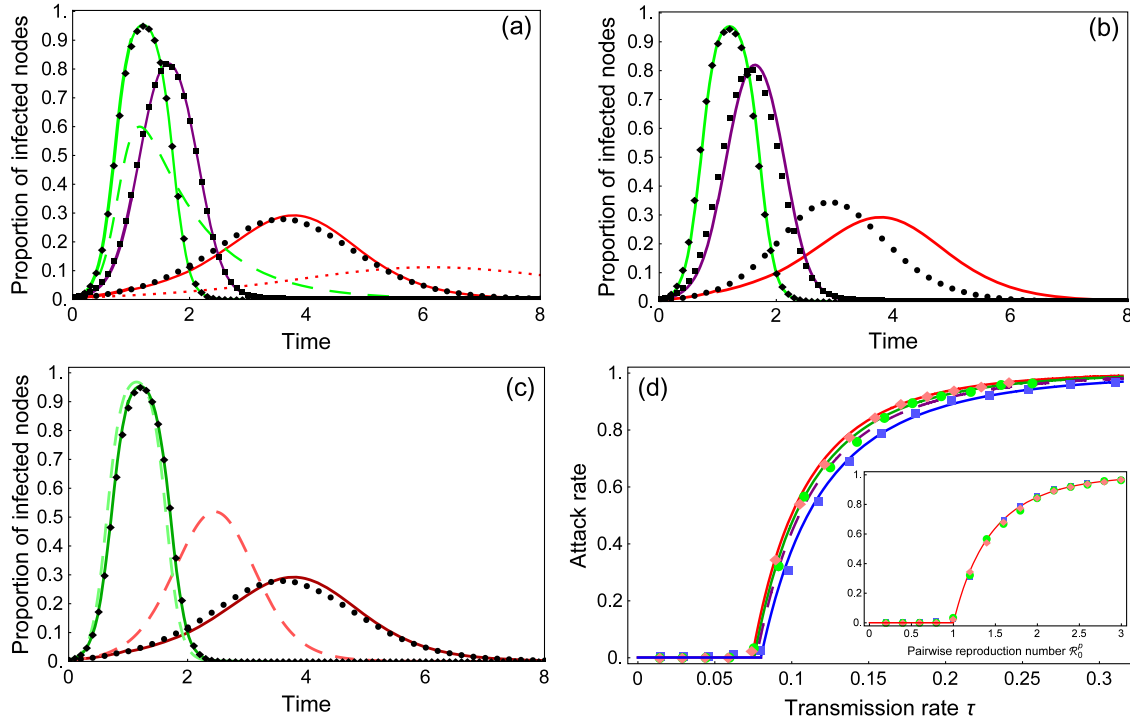
FIG. 1 (color online).   Simulations of non-Markovian epidemics on networks with $N = 1000$ nodes: (a) solid lines show the solution of
Eq. (4) and the circles, squares, and diamonds correspond to simulations for homogeneous (random regular) graphs with $\langle k \rangle = 5, 10, 15$,
respectively; dotted line, $\langle k \rangle = 5$, and dashed line, $\langle k \rangle = 15$, lines correspond to purely Markovian epidemics given by Eq. (1); (b) the
same as before but for Erdős-Rényi random graphs with $\langle k \rangle = 5, 10, 15$; (c) the solid and dashed lines show the solution of pairwise
[Eq. (4)] and mean-field [Eq. (5)] models, respectively and, for regular random graphs with $\langle k \rangle = 5$ and $\langle k \rangle = 15$. For (a), (b), and
(c) the transmission rate is $\tau = 0.55$ and the infectious period is fixed, $\sigma = 1$. Finally, (d) the diamonds, circles, and squares correspond
to simulations using regular random graphs with $\langle k \rangle = 15$ and using fixed and two different but gamma distributed infectious periods
(circle shape $\alpha = 2$, scale $\beta = \frac{1}{2}$, square shape $\alpha = \frac{1}{2}$, scale $\beta = 2$), respectively. The solid lines correspond to the analytical final size for
fixed [Eq. (9)] and general [Eq. (11)] infectious periods, with the dashed line denoting the purely Markovian case. The inset shows the
analytical and the simulated final epidemic sizes plotted against the reproduction number.

Moreover, numerical simulations for Gamma-distributed
infectious periods strongly suggest that this formula will
hold true for pairwise models with arbitrary recovery times,
see Fig. 1(d). The second observation is that taking the
limit of $n \to \infty$ in Eq. (11) gives rise to $\ln(s_\infty) = \mathcal{R}_0^p(s_\infty - 1)$, which is equivalent to the "standard" form
of Eq. (7).

To test the validity of our model we simulated the non-
Markovian *SIR* process with arbitrary recovery times. We
used an event-based simulation where waiting times for all
possible events are generated from appropriate distribu-
tions. During an update the event with the smallest waiting
time is executed followed by the necessary update of the
waiting times of events affected by the most recent change.
Bogũná et al. [18] give an alternative efficient simulation
method. In Figs. 1(a) and (b) homogenous (or regular
random) and Erdős-Rényi random networks are consid-
ered, respectively. Here, the mean of 100 simulations is
compared to the solution of system Eq. (4). The agreement
is excellent for homogenous networks, even for low
degrees. Despite the pairwise model not explicitly account-
ing for degree heterogenity, the agreement is surprisingly

good for relatively dense Erdős-Rényi networks. The figure
also shows that the fixed infectious period significantly
accelerates the growth and turnover of the epidemic
compared to the purely Markovian case.

In Fig. 1(c), the differences between simulations, mean-
field, and pairwise models for the non-Markovian case are
compared. For denser networks, $\langle k \rangle = 15$, both models
perform well with the pairwise model yielding a better
agreement. However, the difference is striking for sparser
networks, $\langle k \rangle = 5$, where the mean-field approximation
performs poorly, while the pairwise DDE model leads to
good agreement with simulation, even in this case.

In Fig. 1(d), analytic final size relations are tested against
simulation results for a range of different infectious period
distributions, all sharing the same mean. Surprisingly, the
final epidemic size can vary by as much as 15%, see
$\tau \sim 0.083$, simply due to the recovery time distributions.
The inset in Fig. 1(d) shows that the same value of $\mathcal{R}_0^p$
produces the same attack rate, regardless of the distribution
from where it originates from, in accordance with our
formula, Eq. (11). Based on Table I, the analytical
expressions for $\mathcal{R}_0^p$ are

$$\mathcal{R}^p_{0,\Gamma(\frac{1}{2},2)} = c\left(1 - \frac{1}{\sqrt{1+2\tau}}\right), \qquad \mathcal{R}^p_{0,\exp(1)} = c\left(\frac{\tau}{\tau+1}\right),$$

$$\mathcal{R}^p_{0,\Gamma(2,\frac{1}{2})} = c\left(1 - \frac{4}{(2+\tau)^2}\right), \qquad \mathcal{R}^p_{0,\mathrm{fixed}(1)} = c(1 - e^{-\tau}),$$

where $c = (n-1)[S]_0/N$, $\Gamma(a,b)$ denotes Gamma-distribution with parameters a and b, and satisfy the following inequality $\mathcal{R}^p_{0,\Gamma(\frac{1}{2},2)} \leq \mathcal{R}^p_{0,\exp(1)} \leq \mathcal{R}^p_{0,\Gamma(2,\frac{1}{2})} \leq \mathcal{R}^p_{0,\mathrm{fixed}(1)}$. We note that (a) all recovery time distributions have the same mean 1 and (b) the variances satisfy the converse inequality, with higher variance in recovery time (i.e., 2, 1, 1/2 and 0) giving a smaller $\mathcal{R}^p_0$ value, despite τ being fixed. Nonexponential recovery times have a significant impact and highlight the necessity to correctly model the recovery time distribution in order to avoid under or over estimation of reproduction number and final size. The excellent agreement between analytic results and the stochastic simulation, see Fig. 1, confirms the validity of our final size relations.

The proposed model provides a viable framework for a more systematic analysis of non-Markovian processes on networks with several future research directions. Similarly to the evolution of the original pairwise model for Markovian dynamics, the proposed model and new closure can be extended to networks with heterogenous degree distribution [30], clustering, or to directed and weighted networks. For example, for heterogenous networks, the critical term in the evolution equation for $[S_iI_j]$, i.e., $d[S_iI_j]/dt$, becomes

$$\tau \sum_k \frac{j-1}{j} \frac{[S_iS_j](t-\sigma)[S_jI_k](t-\sigma)}{[S_j](t-\sigma)} x_i(t),$$

where

$$x_i(t) = e^{-\int_{t-\sigma}^{t}\left(\tau\frac{i-1}{i}\frac{\Sigma_k[S_iI_k](u)}{[S_i](u)}+\tau\right)du}$$

is the discount factor to account for $[S_iI_j]$ edges which are destroyed in the time interval $(t-\sigma, t)$, see the Supplemental Material [29] for full derivation and equations. Additionally, this framework can be employed to model different dynamics, such as *SIS* epidemics and the voter model, or more complex systems, such as adaptive networks. Preliminary investigations indicate that our model can be extended to consider arbitrary recovery time distributions. Our generalization shows an important way in which the analysis of non-Markovian processes can be linked to delay and integro-differential equations, where a well-developed suite of analytical tools is available.

---

*i.z.kiss@sussex.ac.uk
†rost@math.u-szeged.hu
‡zsvizi@math.u-szeged.hu

[1] M. E. Newman, SIAM Rev. **45**, 167 (2003).
[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[3] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, arXiv:1408.2701.
[4] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[5] M. Liu, G. Röst, and G. Vas, Comput. Math. Appl. **66**, 1534 (2013).
[6] M. J. Keeling, Proc. R. Soc. B **266**, 859 (1999).
[7] D. Rand, *Advanced Ecological Theory: Principles and Applications* (John Wiley & Sons, New York, 2009), Chap. 4, pp. 100–142.
[8] T. Gross, C. J. Dommar D'Lima, and B. Blasius, Phys. Rev. Lett. **96**, 208701 (2006).
[9] L. Hébert-Dufresne, O. Patterson-Lomba, G. M. Goerg, and B. M. Althouse, Phys. Rev. Lett. **110**, 108103 (2013).
[10] A. Szabó-Solticzky, L. Berthouze, I. Z. Kiss, and P. L. Simon, J. Math. Biol. (2015).
[11] G. Demirel, F. Vazquez, G. Böhme, and T. Gross, Physica D **267**, 68 (2014).
[12] E. Cator, R. van de Bovenkamp, and P. Van Mieghem, Phys. Rev. E **87**, 062816 (2013).
[13] B. Min, K.-I. Goh, and I.-M. Kim, Europhys. Lett. **103**, 50002 (2013).
[14] F. Cooper, http://www.dtc.ox.ac.uk/people/13/cooperf/files/MA469ThesisFergusCooper.pdf (2013).
[15] P. Van Mieghem and R. van de Bovenkamp, Phys. Rev. Lett. **110**, 108701 (2013).
[16] H.-H. Jo, J. I. Perotti, K. Kaski, and J. Kertész, Phys. Rev. X **4**, 011041 (2014).
[17] E. Kenah and J. M. Robins, Phys. Rev. E **76**, 036113 (2007).
[18] M. Boguná, L. F. Lafuerza, R. Toral, and M. A. Serrano, Phys. Rev. E **90**, 042108 (2014).
[19] T. Hoffmann, M. A. Porter, and R. Lambiotte, Phys. Rev. E **86**, 046102 (2012).
[20] A. Moinet, M. Starnini, and R. Pastor-Satorras, Phys. Rev. Lett. **114**, 108701 (2015).
[21] E. Scalas, T. Kaizoji, M. Kirchler, J. Huber, and A. Tedeschi, Physica A **366**, 463 (2006).
[22] R. D. Malmgren, D. B. Stouffer, A. E. Motter, and L. A. Amaral, Proc. Natl. Acad. Sci. U.S.A. **105**, 18153 (2008).
[23] A. L. Lloyd, Theor. Popul. Biol. **60**, 59 (2001).
[24] M. J. Keeling and B. T. Grenfell, Proc. R. Soc. B **269**, 335 (2002).
[25] H. Nishiura and M. Eichner, Epidemiol. Infect. **135**, 1145 (2007).
[26] M. Eichner and K. Dietz, American Journal of Epidemiology **158**, 110 (2003).
[27] N. Bailey, Biometrika **43**, 15 (1956).
[28] G. Chowell and H. Nishiura, BMC Medicine **12**, 196 (2014).
[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.115.078701 for proof of the final size relations for the mean-field and pairwise DDEs, as well as the proof that the lifetime of an $S-I$ link for Markovian epidemics amounts to evaluating the Laplace transform of the density are given. The pairwise DDEs for heterogeneous networks are also derived.
[30] K. T. Eames and M. J. Keeling, Proc. Natl. Acad. Sci. U.S.A. **99**, 13330 (2002).

Supplemental material for

Generalization of pairwise models to non-Markovian epidemics on networks

Istvan Z. Kiss, Gergely Röst & Zsolt Vizi

July 1, 2015

1 Non-Markovian mean-field model

We consider the following mean-field model with fixed infectious period:

$$
\begin{aligned}
S'(t) &= -\tau \frac{n}{N} S(t)I(t), &\qquad (1)\\
I'(t) &= \tau \frac{n}{N} S(t)I(t) - \tau \frac{n}{N} S(t-\sigma)I(t-\sigma).
\end{aligned}
$$

Below we illustrate how to obtain the final size relation. From first equation of (1), we have

$$
S(t) = S_0 e^{-\tau \frac{n}{N} \int_0^t I(u)du}
$$

and

$$
S_\infty - S_0 = -\tau \frac{n}{N} \int_0^\infty S(u)I(u)du.
$$

On the other hand,

$$
I(t) = \int_0^\sigma \tau \frac{n}{N} S(t-w)I(t-w)dw,
$$

where $\tau \frac{n}{N} S(t-w)I(t-w)$ is the new infections at $t-w$. Hence

$$
\begin{aligned}
\ln\left(\frac{S_\infty}{S_0}\right) &= -\tau \frac{n}{N} \int_0^\infty I(u)du = \tau \frac{n}{N} \int_0^\infty \int_0^\sigma \tau \frac{n}{N} S(u-w)I(u-w)dwdu\\
&= \left(\tau \frac{n}{N}\right)^2 \int_0^\sigma \int_0^\infty S(u-w)I(u-w)dudw\\
&= \left(\tau \frac{n}{N}\right)^2 \int_0^\sigma \left(\int_0^\infty S(q)I(q)dq + \int_{-w}^0 \phi(q)\psi(q)dq\right)dw,
\end{aligned}
$$

where $\phi(t), \psi(t)$ are the initial functions for $S(t), I(t)$ on $[-\sigma, 0]$. By neglecting the small amount of initial infecteds, we have the approximation

$$
-\ln\left(\frac{S_\infty}{S_0}\right) = \left(\tau \frac{n}{N}\right)^2 \sigma \int_0^\infty S(q)I(q)dq = -\tau \frac{n}{N} \sigma (S_\infty - S_0).
$$

Therefore,

$$
\ln\left(\frac{S_\infty}{S_0}\right) = \tau \frac{n}{N} \sigma S_0 \left(\frac{S_\infty}{S_0} - 1\right). \qquad (2)
$$

2 Non-Markovian pairwise model

2.1 Proof of final size relation

We consider the following pairwise model with fixed infectious period:

$$\begin{aligned}
[\dot{S}](t) &= -\tau[SI](t), \\
[\dot{SS}](t) &= -2\tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)}, \\
[\dot{I}](t) &= \tau[SI](t) - \tau[SI](t-\sigma), \\
[\dot{SI}](t) &= \tau\frac{n-1}{n}\frac{[SS](t)[SI](t)}{[S](t)} - \tau\frac{n-1}{n}\frac{[SI](t)[SI](t)}{[S](t)} - \tau[SI](t) \\
&\quad -\tau\frac{n-1}{n}\frac{[SS](t-\sigma)[SI](t-\sigma)}{[S](t-\sigma)}e^{-\int_{t-\sigma}^{t}\tau\frac{n-1}{n}\frac{[SI](u)}{[S](u)}+\tau du}.
\end{aligned} \tag{3}$$

We derive the final size relation for this model. The first integral in the system is $\frac{[SS](t)}{[S]^{2\frac{n-1}{n}}(t)}$. To see this, let's divide the second by the first equation in (3):

$$\frac{d[SS]}{d[S]} = \frac{-2\tau\frac{n-1}{n}\frac{[SS][SI]}{[S]}}{-\tau[SI]} = 2\frac{n-1}{n}\frac{[SS]}{[S]}.$$

Solving this equation, we get $\frac{[SS]}{[S]^{2\frac{n-1}{n}}} = K$, where K is a constant. Thus $\frac{[SS](t)}{[S]^{2\frac{n-1}{n}}(t)}$ is an invariant quantity in the system and its value is

$$K = \frac{[SS](0)}{[S]^{2\frac{n-1}{n}}(0)} = \frac{[SS]_0}{[S]_0^{2\frac{n-1}{n}}} = \frac{n[S]_0\frac{[S]_0}{N}}{[S]_0^{2\frac{n-1}{n}}} = \frac{n}{N}[S]_0^{\frac{2}{n}}.$$

Using this result, we can reduce the four dimensional system to a two-dimensional system:

$$\begin{aligned}
[\dot{S}](t) &= -\tau[SI](t), \\
[\dot{SI}](t) &= \tau\kappa[S]^{\frac{n-2}{n}}(t)[SI](t) - \tau[SI](t) - \tau\frac{n-1}{n}\frac{[SI](t)}{[S](t)}[SI](t) \\
&\quad -\tau\kappa[S]^{\frac{n-2}{n}}(t-\sigma)[SI](t-\sigma)e^{-\int_{t-\sigma}^{t}\tau\frac{n-1}{n}\frac{[SI](u)}{[S](u)}+\tau du},
\end{aligned} \tag{4}$$

where

$$\kappa = \frac{n-1}{N}[S]_0^{\frac{2}{n}}.$$

From the equation

$$w'(t) = \text{in}(t) - \text{out}(t)w(t) - \text{in}(t-\sigma)e^{-\int_{t-\sigma}^{t}\text{out}(u)du}$$

and its solution

$$w(t) = \int_{t-\sigma}^{t} \mathrm{in}(u) e^{-\int_u^t \mathrm{out}(s)ds} du,$$

for $t \geq \sigma$, with setting

$$
\begin{aligned}
\mathrm{in}(t) &= \tau\kappa[S]^{\frac{n-2}{n}}(t)[SI](t), \\
\mathrm{out}(t) &= \tau + \tau\frac{n-1}{n}\frac{[SI](t)}{[S](t)},
\end{aligned}
$$

the equation for $[SI](t)$ is

$$[SI](t) = \int_{t-\sigma}^{t} \tau\kappa[S]^{\frac{n-2}{n}}(u)[SI](u) e^{-\int_u^t \tau + \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)} ds} du.$$

Applying $[S]'(t) = -\tau[SI](t)$, we obtain

$$
\begin{aligned}
[SI](t) &= \int_{t-\sigma}^{t} \tau\kappa[S]^{\frac{n-2}{n}}(u)[SI](u) e^{-\int_u^t \tau + \tau\frac{n-1}{n}\frac{[SI](s)}{[S](s)} ds} du \\
&= -\int_{t-\sigma}^{t} \kappa[S]'(u)[S]^{\frac{n-2}{n}}(u) e^{-\tau(t-u)} e^{\int_u^t \frac{n-1}{n}\frac{[S]'(t)}{[S](s)} ds} du \\
&= -\int_{t-\sigma}^{t} \kappa[S]'(u)[S]^{\frac{n-2}{n}}(u) e^{-\tau(t-u)} e^{\ln\left([S]^{\frac{n-1}{n}}(t)\right) - \ln\left([S]^{\frac{n-1}{n}}(u)\right)} du \\
&= -\kappa[S]^{\frac{n-1}{n}}(t) \int_{t-\sigma}^{t} [S]^{-\frac{1}{n}}(u)[S]'(u) e^{-\tau(t-u)} du.
\end{aligned}
$$

Substituting back to the first equation of (3), we get

$$[S]'(t) = \tau\kappa[S]^{\frac{n-1}{n}}(t) \int_{t-\sigma}^{t} [S]^{-\frac{1}{n}}(u)[S]'(u) e^{-\tau(t-u)} du.$$

Solving this scalar equation leads to

$$[S]^{1-\frac{n-1}{n}}(s) = [S]_0^{1-\frac{n-1}{n}} + \tau\kappa\left(1 - \frac{n-1}{n}\right) \int_0^s \int_{t-\sigma}^{t} [S]^{-\frac{1}{n}}(u)[S]'(u) e^{-\tau(t-u)} du dt.$$

For the final size relation, we need to consider the following equation:

$$[S]_\infty^{\frac{1}{n}} = [S]_0^{\frac{1}{n}} + \kappa\frac{\tau}{n} \int_0^\infty e^{-\tau t} \int_{t-\sigma}^{t} [S]^{-\frac{1}{n}}(u)[S]'(u) e^{\tau u} du dt. \tag{5}$$

First, we compute the double integral:

$$
\begin{aligned}
\int_0^\infty e^{-\tau t} \int_{t-\sigma}^t [S]^{-\frac{1}{n}}(u)[S]'(u)e^{\tau u}\,du\,dt
&= \int_0^\infty [S]^{-\frac{1}{n}}(u)[S]'(u)e^{\tau u}\int_u^{u+\sigma} e^{-\tau t}\,dt\,du \\
&= \int_0^\infty [S]^{-\frac{1}{n}}(u)[S]'(u)e^{\tau u}\left[\frac{e^{-\tau t}}{-\tau}\right]_u^{u+\sigma}\,du \\
&= -\frac{1}{\tau}\int_0^\infty [S]^{-\frac{1}{n}}(u)[S]'(u)e^{\tau u}\left(e^{-\tau(u+\sigma)}-e^{-\tau u}\right)\,du \\
&= \frac{1}{\tau}\left(1-e^{-\tau\sigma}\right)\int_0^\infty [S]^{-\frac{1}{n}}(u)[S]'(u)\,du \\
&= \frac{1}{\tau}\left(1-e^{-\tau\sigma}\right)\left[\frac{[S]^{\frac{n-1}{n}}(u)}{\frac{n-1}{n}}\right]_0^\infty \\
&= \frac{1}{\tau}\left(1-e^{-\tau\sigma}\right)\frac{n}{n-1}\left([S]_\infty^{\frac{n-1}{n}}-[S]_0^{\frac{n-1}{n}}\right).
\end{aligned}
$$

Plugging into (5) we obtain

$$
[S]_\infty^{\frac{1}{n}} = [S]_0^{\frac{1}{n}} + \frac{\kappa}{n}\frac{n}{n-1}\left(1-e^{-\tau\sigma}\right)\left([S]_\infty^{\frac{n-1}{n}}-[S]_0^{\frac{n-1}{n}}\right).
$$

Thus,

$$
\frac{[S]_\infty^{\frac{1}{n}}-[S]_0^{\frac{1}{n}}}{\frac{1}{n}} = \kappa\frac{n}{n-1}\left(1-e^{-\tau\sigma}\right)\left([S]_\infty^{\frac{n-1}{n}}-[S]_0^{\frac{n-1}{n}}\right).
$$

Using the analytical expression for κ from above we have:

$$
\frac{[S]_\infty^{\frac{1}{n}}-[S]_0^{\frac{1}{n}}}{\frac{1}{n}} = \frac{n}{N}[S]_0^{\frac{2}{n}}\left(1-e^{-\tau\sigma}\right)\left([S]_\infty^{\frac{n-1}{n}}-[S]_0^{\frac{n-1}{n}}\right).
$$

Therefore, the relation

$$
\frac{[S]_\infty^{\frac{1}{n}}-[S]_0^{\frac{1}{n}}}{\frac{1}{n}} = \frac{n}{N}\left(1-e^{-\tau\sigma}\right)[S]_0^{\frac{n+1}{n}}\left(\frac{[S]_\infty^{\frac{n-1}{n}}}{[S]_0^{\frac{n-1}{n}}}-1\right) \tag{6}
$$

holds, which can be written as

$$\frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n-1}} = \frac{n-1}{N} \left(1 - e^{-\tau\sigma}\right) [S]_0 \left(s_\infty^{\frac{n-1}{n}} - 1\right),$$

where $s_\infty = \frac{[S]_\infty}{[S]_0}$.

2.2 Limit cases

Using L'Hospital's rule, it is easy to see, that the limit of the left-hand side as $n \to \infty$ is

$$
\begin{aligned}
\lim_{n\to\infty} \frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n-1}} &= \lim_{n\to\infty} \frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n}} \frac{n-1}{n} = \lim_{n\to\infty} \frac{s_\infty^{\frac{1}{n}} - 1}{\frac{1}{n}} \\
&= \lim_{m\to0} \frac{s_\infty^m - 1}{m} = \lim_{m\to0} \frac{s_\infty^m \ln s_\infty}{1} \\
&= \ln s_\infty.
\end{aligned}
$$

2.3 Markovian infectious process leads to the Laplace transform of PDF

To calculate the expected lifetime of an $S - I$ link, if the infection is Markovian and the recovery is arbitrary with density function $f_r(x)$ and survival function $\xi_r(x)$, we integrate by parts and we obtain

$$
\begin{aligned}
\mathbb{E}(Z) &= \int_0^\infty t \left(f_i(t)\xi_r(t) + f_r(t)\xi_i(t)\right) dt = \int_0^\infty t \left(\tau e^{-\tau t}\xi_r(t) + e^{-\tau t}f_r(t)\right) dt \\
&= \int_0^\infty t\tau e^{-\tau t}\xi_r(t)dt + \int_0^\infty te^{-\tau t}f_r(t)dt \\
&= \left[\left(-te^{-\tau t} - \frac{e^{-\tau t}}{\tau}\right)\xi_r(t)\right]_0^\infty - \int_0^\infty \left(te^{-\tau t} + \frac{e^{-\tau t}}{\tau}\right) f_r(t)dt \\
&\quad + \int_0^\infty te^{-\tau t}f_r(t)dt \\
&= \frac{1}{\tau} - \frac{1}{\tau}\int_0^\infty e^{-\tau t}f_r(t)dt = \frac{1 - \mathcal{L}[f_r](\tau)}{\tau},
\end{aligned}
\tag{7}
$$

where $\mathcal{L}[f_r](\tau)$ denotes the Laplace transform of f_r at τ. Multiplying this formula with the expected number of newly generated $S - I$ links $\tau\frac{n-1}{N}[S]_0$, we have the general formula for pairwise reproduction number.

6

3 Derivation of the pairwise model for networks with heterogeneous degree distribution

Here, we show how the current pairwise equations, as given in the letter, extend naturally to heterogeneous networks. In this case variables, such as

1. $[S_i](t)$ - expected number of susceptible nodes of degree i,

2. $[I_i](t)$ - expected number of infected nodes of degree i,

3. $[S_iS_j](t)$ - expected number of $S - S$ links, where S and S have degrees i and j, respectively,

4. $[S_iI_j](t)$ - expected number of $S - I$ links, where S and I have degrees i and j, respectively,

need to be considered, where $i, j \in \{k_{min}, k_{min} + 1, \ldots, k_{max}\}$ represent the various degrees in the networks.

The slightly more technical part is to replicate Eq. (3) from our letter to degree dependent $[SI]$ pairs. This can be done as follows. Let $x(t)$ denote the factor by which $[S_iI_j]$ links needs to be discounted by. The equation for $x(t)$ is given by

$$\dot{x}(t) = -\tau \frac{i-1}{i} \frac{\sum_k [S_iI_k]}{[S_i]} x(t) - \tau x(t), \tag{8}$$

where in fact the factor x only depends on the degree of the susceptible node so it could be denoted by x_i. It is worth noting that

$$\frac{\sum_k [S_iI_k]}{i[S_i]}$$

gives the probability that a stub emanating from a susceptible nodes with i links will connect to an infected node, and $(i-1)$ stands for the remaining stubs emanating from an S_i node which is already connected to another node, in this case an infected node.

This can be integrated as in the letter and the non-Markovian pairwise system for heterogenous networks yields

$$
\begin{aligned}
[\dot{S}_i](t) &= -\tau \sum_j [S_i I_j](t), \\
[\dot{I}_i](t) &= \tau \sum_j [S_i I_j](t) - \tau \sum_j [S_i I_j](t-\sigma), \\
[\dot{S_i S_j}](t) &= -\tau \frac{j-1}{j} \frac{[S_i S_j](t)}{[S_j](t)} \sum_k [S_j I_k](t) - \tau \frac{i-1}{i} \frac{[S_i S_j](t)}{[S_i](t)} \sum_k [S_i I_k](t), \\
[\dot{S_i I_j}](t) &= -\tau [S_i I_j](t) + \tau \frac{j-1}{j} \frac{[S_i S_j](t)}{[S_j](t)} \sum_k [S_j I_k](t) - \tau \frac{i-1}{i} \frac{[S_i I_j](t)}{[S_i](t)} \sum_k [I_k S_i](t) \\
&\quad -\tau \sum_k \frac{j-1}{j} \frac{[S_i S_j](t-\sigma)[S_j I_k](t-\sigma)}{[S_j](t-\sigma)} e^{-\int_{t-\sigma}^{t} \left(\tau \frac{i-1}{i} \frac{\sum_k [S_i I_k](u)}{[S_i](u)} + \tau \right) du}. \quad (9)
\end{aligned}
$$

where $i, j, k \in \{k_{min}, k_{min}+1, \ldots, k_{max}\}$.