J. Math. Biol. (2014) 69:1627–1660
DOI 10.1007/s00285-013-0744-9

Mathematical Biology

Spreading dynamics on complex networks: a general stochastic approach

Pierre-André Noël · Antoine Allard · Laurent Hébert-Dufresne · Vincent Marceau · Louis J. Dubé

Received: 16 August 2013 / Published online: 24 December 2013

Abstract Dynamics on networks is considered from the perspective of Markov stochastic processes. We partially describe the state of the system through network motifs and infer any missing data using the available information. This versatile approach is especially well adapted for modelling spreading processes and/or population dynamics. In particular, the generality of our framework and the fact that its assumptions are explicitly stated suggests that it could be used as a common ground for comparing existing epidemics models too complex for direct comparison, such as agent-based computer simulations. We provide many examples for the special cases of susceptible-infectious-susceptible and susceptible-infectious-removed dynamics (*e.g.*, epidemics propagation) and we observe multiple situations where accurate results may be obtained at low computational cost. Our perspective reveals a subtle balance between the complex requirements of a realistic model and its basic assumptions.

Keywords Spreading dynamics · Complex networks · Stochastic processes · Contact networks · Epidemics · Markov processes

Mathematics Subject Classification (2000) 93A30 · 05C82 · 60J28 · 92D25 · 92D30

P.-A. Noël
University of California, Davis, CA 95616, USA
e-mail: noel.pierre.andre@gmail.com

A. Allard · L. Hébert-Dufresne · V. Marceau · L. J. Dubé (✉)
Département de Physique, de Génie Physique et d'Optique, Université Laval,
Québec, Québec G1V 0A6, Canada
e-mail: ljd@phy.ulaval.ca

1 Introduction

Mathematical modelling has proven a valuable tool when addressing population dynamics problems, both for public health and ecological issues. The increased availability of powerful computer resources has facilitated the use of agent-based models and other complex modelling approaches, all accounting for numerous parameters and assumptions (Auchincloss and Diez Roux 2008; McLane et al. 2011; Broeck et al. 2011). Our confidence in these models may increase when they are shown to agree with empirical observations and/or with previously accepted models. However, when discrepancies appear, the complexity of these computer programs may conceal the effects of underlying assumptions, making it difficult to isolate the source of disagreement. While analytical approaches offer more insights on the underlying assumptions, their use is often restricted to simpler interaction structures and/or dynamics.

The purpose of this paper is to systematically model the global behaviour of stochastic systems composed of numerous elements interacting in a complex way. "Complex" here implies that interactions among the elements follow some nontrivial patterns that are neither perfectly regular nor completely random, as often seen in real-world systems. "Stochastic" implies that the system may not be completely predictable to an observer and that a probabilistic solution is sought.

We present (Sect. 2) a general modelling scheme where network theory (Barrat et al. 2008; Boccaletti et al. 2006; Durrett 2007; Newman 2010) accounts for the interactions between the elements of the system and where a birth-death Markov process (Gardiner 2004) models the stochastic dynamics. Since a tremendous amount of information may be required to store the state of the whole system, we seek the part of this information that is important for the problem at hand and then approximate the dynamics by tracking only this limited subset. Part of the discarded data may still affect, albeit weakly, the behaviour of the system. We fill this knowledge gap by inferring the missing information such that it is consistent both with the information we follow and any other prior information that is available to us.

An important part of this paper (Sect. 3) provides explicit examples to these general ideas. Keeping in mind that the generality of our approach is in the mathematical description and not in the studied cases, we focus our study to spreading processes such as the propagation of infectious diseases on contact networks (Keeling and Eames 2005; Bansal et al. 2007; Danon et al. 2011; House and Keeling 2011; Sharkey 2011; Taylor et al. 2012). Specifically, our examples correspond to either one of two standard epidemic models, the susceptible-infectious-susceptible (SIS) and the susceptible-infectious-removed (SIR) dynamics, for which we also adopt the relevant vocabulary. While our first examples study simpler cases, facilitating the understanding of our systematic method, the later models show how the same approach applies to more complex interaction structures.

We then compare and analyse the results of these examples (Sect. 4). This reveals some general considerations for both the accuracy and the complexity of our modelling approach. We find that treating the inferences of missing information explicitly helps systematize the model development and highlights numerous possibilities for future developments. An important simplification occurs for SIR spreading processes and related dynamics, leading to an *exact* model with a small number of dynamical variables.

We conclude (Sect. 5) on how our general approach may be applied to population dynamics in general, as well as to other spreading processes such as the cascading extinctions of species in food webs (Rezende et al. 2007; Bascompte and Stouffer 2009; Dunne and Williams 2009). Returning to the problem of understanding the source of discrepancies in complex models, we explain how *modelling these models* with our method could help identifying important assumptions and isolating the source of disagreement. Mathematical details and further generalizations are presented in two separate Appendices.

2 General modelling scheme

This section is not concerned with any specific model per se, but instead seeks to answer a "meta-modelling" question: given limited resources, how can one design a simple, yet reliable, model? Specific examples are deferred to Sect. 3.

2.1 Levels of abstraction

To simplify the following discussion, we introduce three "levels" of increasing abstraction (decreasing complexity).

Level I: The *real-world system* that we desire to model. Examples include the propagation of infections in populations; and/or the formation and dissolution of groups and/or partnerships. Such systems may be very complex, and we definitely do not know everything about them. We are particularly interested in a subset of these unknowns, which we will hereafter refer to as *open questions* (*e.g.*, "How does the population structure affects the spread of infections?").

Level II: A (possibly hypothetical) agent-based Monte Carlo computer simulation, which we call *full system*. Since we do not know everything about Level I, obtaining Level II clearly requires some assumptions and/or approximations. Nonetheless, we *assume* that running this simulation on an infinitely powerful computer would provide some (perhaps partial) answers to our open questions. In this sense, Level II *approximately reproduces* the behaviour of Level I: understanding Level II provides some understanding of Level I.

Level III: A *simplified system* which requires much less computational resources than the Level II full system. In practice, the truth is that we do not have access to an infinitely powerful computer, and regular computers may not be sufficient for our needs (which may include, for instance, time-consuming sensitivity analysis for parameters estimation). To obtain Level III, we must endeavour to extract the relevant parts of Level II, and remove everything else. We say that Level III *approximately reproduces* the behaviour of Level II if it (approximately) provides the same answers to the Level I open questions.

We emphasize that our main concern is the modelling of Level II by Level III. The transition from Level I to Level II is *outside the scope of this article*, meaning that, given a complex system (Level I) to analyse, it is left to the experts in the appropriate field to provide a reliable Level II.

Another point of importance is that there is no a priori guarantee that we will figure out the correct "relevant parts" to obtain a simple Level III model that approximately reproduces the behaviour of the Level II simulations. In fact, it is not even guaranteed that such a simplification exists. Nonetheless, we may a posteriori test whether or not the approach was successful by directly comparing the outcomes predicted by the Level II simulations to those of the Level III model.

For the purpose of such comparison, the examples presented in Sect. 3 have been chosen simple enough for us to actually perform the Level II simulations. Ultimately, we would want to acquire enough confidence in our Level III models to directly address the Level I open questions, hence skipping the costly Level II simulations. We also identify some factors influencing the success of our approach in Sect. 4.

2.2 States, rules, and priors

We now present a general method for translating a Level II simulations into a Level III model. Our approach, conceptually outlined in Fig. 1, first requires to formulate both levels as Markov stochastic processes.

The Level II simulations are specified by two objects.

$Z(t)$: The *state of the full system* at time t This corresponds to any form of storage available to the Level II simulations (*e.g.,* heap, stack, registers, hard drive).[1] Further details concerning this object are delayed to Sect. 2.3.

V: The *rules of the full system* This corresponds to the program itself: given the state $Z(t)$ at time t, V tells us towards which state[2] $Z(t + \mathrm{d}\,t)$ we should update the memory after an infinitesimal time step $\mathrm{d}\,t$. It is easy to show that such a simulation respects the Markov property: knowing the current state $Z(t)$ of the computer's memory, none of the past state may further affect our probability estimate for its future state.

We may also define two similar objects for the Level III model.

$X(t)$: The *state of the simplified system* at time t. We will typically be interested in $X(t)$ that are *much smaller* than $Z(t)$. Hence, knowledge of $X(t)$ typically convey only *partial* information on $Z(t)$. As a trivial example, if $Z(t) \in \{0, 1\}^N$ is a sequence of N zeros or ones, then we could choose $X(t) \in \{0, 1, 2, \ldots, N\}$ as specifying the total number of ones in this sequence. Further details concerning this object are delayed to Sect. 2.4.

[1] Computer science terminology is used for the sake of specificity. However, the reader should keep in mind that these simulations may be hypothetical: our goal is to *replace* them by a Level III model.

[2] While V should technically be deterministic, we may also perceive it as stochastic by considering pseudo-random number generators as "actually" random. In any case, these subtleties are of no concerns for our purpose.

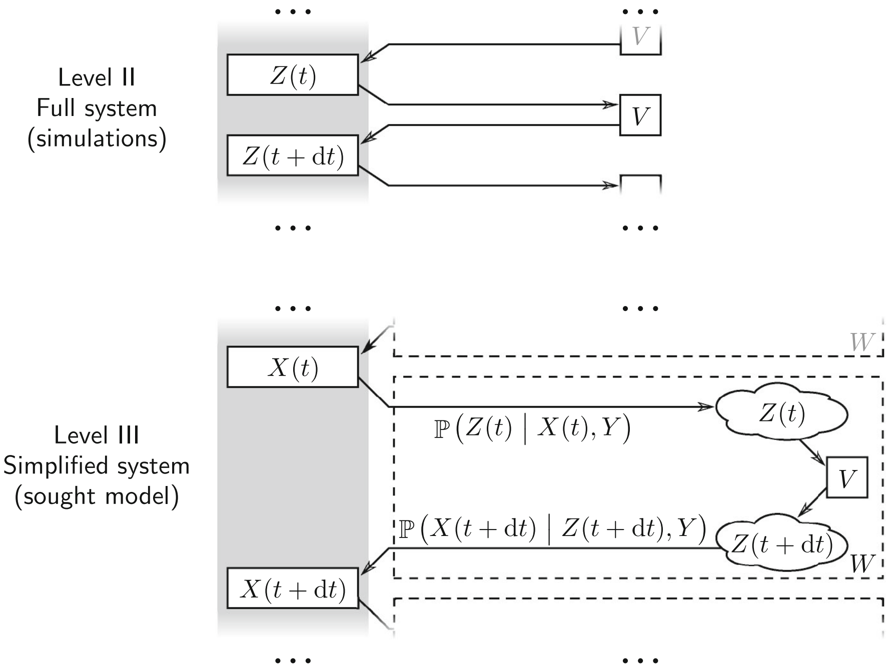


Fig. 1 Conceptual outline of our approach. The state $Z(t)$ of the Level II simulations is updated to $Z(t+\mathrm{d}\,t)$ by applying the rules V. Similarly, the state $X(t)$ of the Level III model is updated to $X(t+\mathrm{d}\,t)$ by applying the rules W (*dashed box*). However, W is a priori unknown to us: it has to be obtained from V. Since the latter acts on $Z(t)$, not $X(t)$, the translation from V to W requires a Bayesian inference accounting for the prior information Y available to us

W: The *rules of the simplified system*. Given the state $X(t)$ at time t, W provides a probability distribution for $X(t+\mathrm{d}\,t)$ after an infinitesimal time step $\mathrm{d}\,t$. Although V was *shown* to respect the Markov property, we *choose* W to also be Markovian.

Our problem may thus be restated as follow: translate the Level II rules V to some Level III rules W such that $X(t)$ approximately reproduces[3] the behaviour of $Z(t)$. Since $X(t)$ may only convey partial information on $Z(t)$, a crucial step is to assert the likelihood of each Level II states by using as much information as is available to us. For this purpose, we define one more object.

Y: The *prior information* that can be used in a Bayesian inference of $Z(t)$ from $X(t)$. Indeed, not all $Z(t)$ may be equally likely, and some constraints may even *forbid* large subsets of them. In the previous trivial example, $X(t)$ specifies the total number of ones in $Z(t) \in \{0, 1\}^N$, and our best guess may thus be to assign equal probability to every $Z(t)$ satisfying this constraint (and probability zero to the violating cases). However, complications may occur: if some mechanisms in V disfavours the formation of long sequences of 0 in $Z(t)$, then it may be necessary

[3] Remember that what we are ultimately interested in are the Level I open questions, so $X(t)$ and $Z(t)$ should be compared with respect to these concerns.

to account for this when translating V to W. Such specifications are also done through Y.

The "Bayesian inference" mentioned in the definition of Y corresponds to

$$\mathbb{P}(Z(t)|X(t), Y) = \frac{\mathbb{P}(Z(t)|Y)\,\mathbb{P}(X(t)|Y, Z(t))}{\mathbb{P}(X(t)|Y)} \tag{1}$$

where

$\mathbb{P}(Z(t)|X(t), Y)$ is the probability for the Level II system to be in state $Z(t)$ at time t, given that the Level III system is in state $X(t)$ at time t, and provided the prior information Y;

$\mathbb{P}(Z(t)|Y)$ is the probability for the Level II system to be in state $Z(t)$ at time t, given that we only know the prior information Y;

$\mathbb{P}(X(t)|Z(t), Y)$ is the probability for the Level III system to be in state $X(t)$ at time t, given that the Level II system is in state $Z(t)$ at time t, and provided the prior information Y; and

$\mathbb{P}(X(t)|Y)$ is the probability for the Level III system to be in state $X(t)$ at time t, given that we only know the prior information Y.

Note that the probabilities $\mathbb{P}(X(t)|Z(t), Y)$, $\mathbb{P}(Z(t)|Y)$ and $\mathbb{P}(X(t)|Y)$ are typically much easier to obtain than directly tackling $\mathbb{P}(Z(t)|X(t), Y)$.

We are now ready to specify the rules W of the Level III model (the dashed box in Fig. 1), which advances the state $X(t)$ of the simplified system to the state $X(t + \mathrm{d}\,t)$.

- Use $\mathbb{P}(Z(t)|X(t), Y)$ to infer, from the available information (*i.e.*, $X(t)$ and Y), the distribution of Level II states $Z(t)$ that is currently predicted by our Level III model.
- Use the Level II rules V to obtain the updated distribution of Level II states $Z(t+\mathrm{d}\,t)$.
- Use $\mathbb{P}(X(t + \mathrm{d}\,t)|Z(t + \mathrm{d}\,t), Y)$ to translate $Z(t + \mathrm{d}\,t)$ back to a distribution of Level III states, $X(t + \mathrm{d}\,t)$.

Note that this full procedure should be understood at a conceptual level, and somehow corresponds to a "worst case scenario". While we typically do have to infer some information missing from $X(t)$ before updating it to $X(t + \mathrm{d}\,t)$, obtaining a full-blown distribution of states $Z(t)$ is usually not required.

Up to now, our discussion has remained very general—too general in fact to be of much practical use. Since we are primarily interested in systems composed of many elements interacting through complex patterns, it is natural to reconsider the previous quantities in terms of networks.

2.3 Networks

A *network* (graph) is a collection of *nodes* (vertices) and *links* (edges) (Barrat et al. 2008; Boccaletti et al. 2006; Durrett 2007; Newman 2010). Nodes model the elements of a system; links join nodes pairwise to represent interactions between the corresponding elements. Two nodes sharing a link are said to be *neighbours* and the *degree* of a node is its number of neighbours. The part of a link that is attached to a node is called a *stub*: there are two stubs per link and each node is attached to a number of stubs equal to its degree. A link with both ends leading to the same node is called a *self-loop* and *repeated links* occur when more than one link join the same two nodes.

Links may (or may not) be directed: an *undirected* link represents a bidirectional and symmetric interaction between the two linked nodes, while a *directed* links represents interactions that are either unidirectional or asymmetrical. A network that has only undirected links is said to be *undirected*, one that has only directed links is *directed* and one that has both is *semi-directed*.

There are systems such that specifying its state $Z(t)$ exactly amounts to specifying the network structure. However, most systems are not purely structural: they are composed of elements that, by themselves, require additional information to be properly characterized. Hence, we assign to each node (resp. link) a *node state* (resp. *link state*) that specifies the intrinsic properties of the corresponding element (resp. interaction) in the system. There is a total of $\mathcal{N}$ discrete node states (resp. $\mathcal{L}$ discrete link states) and the intrinsic state ν of a given node (resp. ℓ of a given link) may change in time to any of these accessible values; the special case $\mathcal{N} = 1$ (resp. $\mathcal{L} = 1$) corresponds to the situation where every nodes (resp. links) remain intrinsically identical during the whole process. At any given time t, the state $Z(t)$ of the full system is specified by the intrinsic state of all of its components (nodes and links) in addition to the network structure governing their interactions. While each node (or link) may be in only one intrinsic state at any given time, different pieces of information may be encoded in this single state.

2.4 Motifs

Specifying the complete structure of a complex network requires a tremendous amount of information. Since we want the state $X(t)$ of a simplified system to be of manageable size, approximations have to be made. A convenient way to do so, and one that has proven to give good results in the past (House et al. 2009; Karrer and Newman 2010; Gleeson 2011; Marceau et al. 2010; Hébert-Dufresne et al. 2010; Marceau et al. 2011; Noël et al. 2012), is to specify the network structure through its building blocks.

A network *motif* is a pattern that may appear a number of times in the network. For example, two linked nodes form a *pair motif* while three nodes all neighbours of one another form a *triangle motif*. Motifs may encode intrinsic node states or other relevant information; further details and examples are provided throughout Sect. 3 as well as in Appendix A.

We define the *state vector* $\mathbf{x}(t) \in \mathbb{N}_0^d$ of a system as a vector of d natural numbers (including zero) specifying how many times different motifs appear in the network at a given time t. We may perceive these d tracked motifs as building blocks that should be attached together to form a network structure (see Fig. 2), and we now specify the simplified system in these terms. Hence, we enumerate the available building blocks with the simplified system state $X(t) = \mathbf{x}(t)$, whereas the prior information Y specifies how such blocks may be attached. There will usually be numerous valid ways to attach the blocks, some more probable than others. Given the available information, the resulting distribution is our best estimate for $\mathbb{P}(Z(t)|\mathbf{x}(t), Y)$.

By judiciously choosing the motifs enumerated in $\mathbf{x}(t)$ and by specifying informative prior information Y, one may hope for this probability distribution to be densely

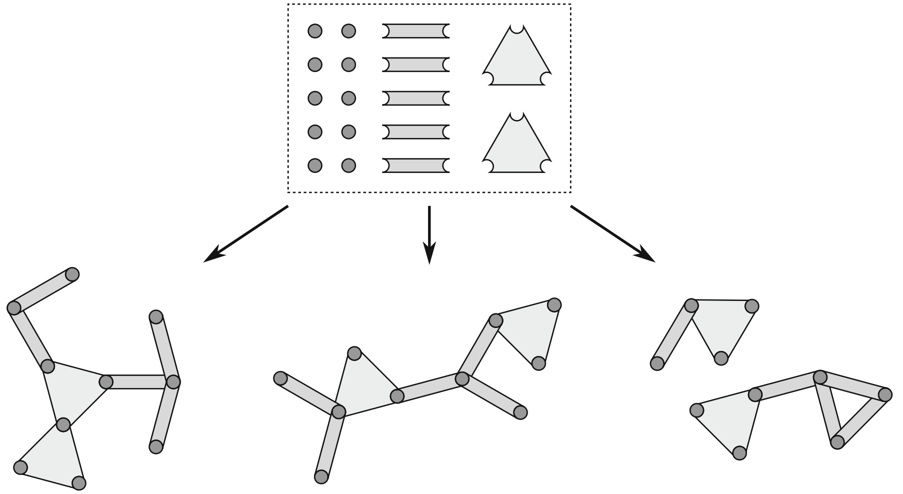


Fig. 2 Motifs may be seen as building blocks that are assembled according to some specified rules. The illustrated situation corresponds to the state vector $\mathbf{x}(t) = (10, 5, 2)$ where the entries respectively provide the number of node motifs (*circles*), of pair motifs (which must join two nodes), and of *triangle* motifs (which must joint three nodes). These motifs may be assembled to form three of the numerous valid networks defined by this state vector. From the viewpoint of the simplified model, any of these arrangements could be the “right” one

localized around the “real” value of Z in the full system. This mapping can then be used to convert the rules V of the full system to the rules W of the new simplified one. We approach this problem from the perspective of birth-death Markov processes.

2.5 Birth-death stochastic processes

In a *birth-death* process (Gardiner 2004; Van Kampen 2007), the elements composing a system may be destroyed (death) while new ones may be created (birth). It is therefore natural to state the rules W of our simplified system in those terms: any change in the state vector $\mathbf{x}(t)$ may be perceived as an event where motifs are created and/or destroyed.

Quantitatively, a *forward transition event of type* j takes the system from state $\mathbf{x}(t)$ to state $\mathbf{x}(t + \mathrm{d}\,t) = \mathbf{x}(t) + \mathbf{r}^j$ and has probability $q_j^+\big(\mathbf{x}(t), Y\big)\,\mathrm{d}\,t$ to occur during the time interval $[t, t + \mathrm{d}\,t)$. Similarly, a *backward transition event of type* j takes the system from state $\mathbf{x}(t)$ to state $\mathbf{x}(t+\mathrm{d}\,t) = \mathbf{x}(t)-\mathbf{r}^j$ and has probability $q_j^-\big(\mathbf{x}(t), Y\big)\,\mathrm{d}\,t$ to occur during the same time interval.

Specifying for each j the integer[4] elements r_i^j of the *shift vector* $\mathbf{r}^j \in \mathbb{Z}^d$ together with the *rate functions* $q_j^+\big(\mathbf{x}(t), Y\big)$ and $q_j^-\big(\mathbf{x}(t), Y\big)$ thus completely define the rules W governing the simplified system. This Markov process is summarized in the master equation

[4] The r_i^j may be positive, negative, or zero. For a given $\mathbf{x}(t)$, if $\mathbf{x}(t) \pm \mathbf{r}^j$ contains any negative entries, then we should have $q_j^{\pm}\big(\mathbf{x}(t), Y\big) = 0$ [i.e., $\mathbf{x}(t + \mathrm{d}\,t)$ has zero probability to be negative].

$$\frac{\mathrm{d}\,\mathbb{P}(\mathbf{x}(t)|Y)}{\mathrm{d}\,t}=\sum_j\Big[q_j^+\big(\mathbf{x}(t)-\mathbf{r}^j,\ Y\big)\mathbb{P}\big(\mathbf{x}(t)-\mathbf{r}^j\big|Y\big)-q_j^+\big(\mathbf{x}(t),\ Y\big)\mathbb{P}\big(\mathbf{x}(t)|Y\big)$$
$$+q_j^-\big(\mathbf{x}(t)+\mathbf{r}^j,\ Y\big)\mathbb{P}\big(\mathbf{x}(t)+\mathbf{r}^j\big|Y\big)-q_j^-\big(\mathbf{x}(t),\ Y\big)\mathbb{P}\big(\mathbf{x}(t)|Y\big)\Big] \tag{2}$$

specifying the evolution of the probability $\mathbb{P}(\mathbf{x}(t)|Y)$ to observe state $\mathbf{x}(t)$ at time t knowing the prior information Y.

Solving the model specified by W, $\mathbf{x}(t)$ and Y thus amounts to obtaining $\mathbb{P}(\mathbf{x}(t)|Y)$ by solving Eq. (2). However, such an approach is often unpractical due to the large number of accessible states for $\mathbf{x}(t)$. In the rest of this section, we consider an approximation that greatly simplifies this problem by assuming that $\mathbb{P}(\mathbf{x}(t)|Y)$ is a multi-dimensional Gaussian probability distribution. Note that the following introduces no new mathematics, and we will thus heavily rely on the textbook (Gardiner 2004) to simplify the discussion. Our starting point, Eq. (2), amounts to (Gardiner, 2004, equation (7.5.9)).

We first consider the approximation, often justified for large systems, that the elements of $\mathbf{x}(t)$ may be treated as varying continuously [i.e., $\mathbf{x}(t)\in\mathbb{R}^d$]. Specifically, we use van Kampens system size expansion truncated to order two, which gives the same result as one would obtain using the Kramers-Moyal expansion (Gardiner 2004, section 7.5.3, see also sections 7.2.2 and 7.2.3). The resulting expression may be expressed as a Fokker-Planck equation (Gardiner 2004, equations (7.5.31)–(7.5.32)) or equivalently as a stochastic differential equation (Gardiner 2004, equation (4.3.21)); we here choose to do the latter

$$\mathrm{d}\,\mathbf{x}(t)=\mathbf{a}\big(\mathbf{x}(t)\big)\,\mathrm{d}\,t+\widetilde{\mathsf{B}}\big(\mathbf{x}(t)\big)\cdot\mathrm{d}\,\mathbf{W}(t). \tag{3}$$

The *drift vector* $\mathbf{a}\big(\mathbf{x}(t)\big)$ has the same dimension as $\mathbf{x}(t)$ and is composed of the elements

$$a_i\big(\mathbf{x}(t)\big)=\sum_j r_i^j\,\Big[q_j^+\big(\mathbf{x}(t),Y\big)-q_j^-\big(\mathbf{x}(t),Y\big)\Big]. \tag{4a}$$

The *vector Wiener process* $\mathbf{W}(t)$ needs not be of the same dimension as $\mathbf{x}(t)$, so the *stochastic weight matrix* $\widetilde{\mathsf{B}}\big(\mathbf{x}(t)\big)$, of elements[5]

$$\widetilde{B}_{ij}\big(\mathbf{x}(t)\big)=r_i^j\sqrt{q_j^+\big(\mathbf{x}(t),Y\big)+q_j^-\big(\mathbf{x}(t),Y\big)}$$

needs not be square.

Taking the mathematical expectation of both sides of (3) provides the ordinary differential equation

$$\frac{\mathrm{d}\,\boldsymbol{\mu}(t)}{\mathrm{d}\,t}=\mathbf{a}\big(\boldsymbol{\mu}(t)\big) \tag{4b}$$

[5] Notice that, although $q_j^+\big(\mathbf{x}(t),Y\big)$ and $q_j^-\big(\mathbf{x}(t),Y\big)$ tend to "cancel out" in the deterministic contribution [negative sign in (4a)], the stochastic contributions of transitions happening both forward and backwards instead "accumulate" [positive sign in (4d)].

for the mean value $\boldsymbol{\mu}(t) = \sum_{\mathbf{x}(t)} \mathbf{x}(t)\mathbb{P}(\mathbf{x}(t)|Y)$ of the vector $\mathbf{x}(t)$, where $\boldsymbol{\mu}(0) = \mathbf{x}(0)$ for a deterministic initial condition.

If the probability distribution $\mathbb{P}(\mathbf{x}(t)|Y)$ remains concentrated around its mean value $\boldsymbol{\mu}(t)$, we may assume that the "noise term" $\widetilde{\mathsf{B}}\big(\mathbf{x}(t)\big) \cdot \mathrm{d}\,\mathbf{W}(t)$ of Eq. (3) is much smaller than its "deterministic" one $\mathbf{a}\big(\mathbf{x}(t)\big)\,\mathrm{d}\,t$. Linearizing this stochastic differential equation thus provides the time-dependent Ornstein-Uhlenbeck process

$$\mathrm{d}\,\mathbf{x}(t) = \mathsf{A}\big(\boldsymbol{\mu}(t)\big) \cdot \mathbf{x}(t)\,\mathrm{d}\,t + \widetilde{\mathsf{B}}\big(\boldsymbol{\mu}(t)\big) \cdot \mathrm{d}\,\mathbf{W}(t),$$

where we defined the *evolution matrix* $\mathsf{A}(t, t')$ and the *diffusion matrix* $\mathsf{B}\big(\mathbf{x}(t)\big)$ $\big[$of elements $B_{ii'}\big(\mathbf{x}(t)\big)\big]$

$$\mathsf{A}(t, t') = \exp\left[\int_{t'}^{t} \mathsf{J}_{\mathbf{a}}\big(\boldsymbol{\mu}(t'')\big)\,\mathrm{d}\,t''\right] \tag{4c}$$

$$B_{ii'}\big(\mathbf{x}(t)\big) = \sum_j r_i^j r_{i'}^j \Big[q_j^+\big(\mathbf{x}(t), Y\big) + q_j^-\big(\mathbf{x}(t), Y\big)\Big], \tag{4d}$$

both matrices being square. Note that $\mathsf{J}_{\mathbf{a}}\big(\mathbf{x}(t)\big)$ is the Jacobian matrix of $\mathbf{a}$ evaluated at $\mathbf{x}(t)$, and that $\mathsf{B}\big(\mathbf{x}(t)\big) = \widetilde{\mathsf{B}}\big(\mathbf{x}(t)\big) \cdot \widetilde{\mathsf{B}}^{\top}\big(\mathbf{x}(t)\big)$. Because this Ornstein-Uhlenbeck process has the same form as (Gardiner, 2004, equation (4.4.79)), the resulting *covariance matrix* $\mathsf{C}(t)$ of $\mathbf{x}(t)$ for a deterministic initial condition is provided by (Gardiner, 2004, equation (4.4.85))

$$\mathsf{C}(t) = \int_0^t \mathsf{A}(t, t') \cdot \mathsf{B}\big(\boldsymbol{\mu}(t')\big) \cdot \mathsf{A}(t, t')^{\top}\,\mathrm{d}\,t'. \tag{4e}$$

Recalling that d is the size of the vector $\mathbf{x}(t)$, the probability distribution may be approximated by a d-dimensional Gaussian

$$\mathbb{P}(\mathbf{x}(t)|Y) = \frac{\exp\left\{-\frac{1}{2}\big[\mathbf{x}(t) - \boldsymbol{\mu}(t)\big]^{\top} \cdot \mathsf{C}(t)^{-1} \cdot \big[\mathbf{x}(t) - \boldsymbol{\mu}(t)\big]\right\}}{\sqrt{(2\pi)^d \big|\mathsf{C}(t)\big|}} \tag{4f}$$

where $\big|\mathsf{C}(t)\big|$ is the determinant of $\mathsf{C}(t)$. The approximation provided by Eqs. (4a)–(4f) should be valid in large systems when the *actual* solution of (2) is close to a d-dimensional Gaussian.

Although many other tools are available for the analysis of stochastic systems, the simplicity, the generality and the straightforwardness of the Gaussian approximation (4) makes it an instrument of choice that will be used extensively in this article.

3 Application to spreading dynamics

Without prejudice to the generality of Sect. 2, we now focus our study to spreading processes. An epidemiological terminology is used: whatever propagates among neighbouring nodes, be it desirable or not, is called an *infection*. We find that the basic SIS and SIR epidemiological models, both to be defined shortly, require little prior knowledge from the part of the reader while being sufficiently complex for the needs of the present study. For similar reasons, we assume that links are undirected and intrinsically identical (*i.e.*, $\mathcal{L} = 1$); see Appendix A for discussions concerning directed links and more than one intrinsic link state.

At a given time t, the intrinsic state of each node of an *SIS model* may be either of $\mathcal{N} = 2$ accessible intrinsic node states: *Susceptible* (not carrying the infection) or *Infectious* (carrying the infection). The full system state $Z(t)$ hence specifies each node's intrinsic state together with the complete structure of the network. The rules V are simple: during any time interval $[t, t + \mathrm{d}\,t)$, each infectious node may recover (*i.e.*, it becomes susceptible) with probability $\alpha\ \mathrm{d}\,t$ and, for each of its susceptible neighbours, has probability $\beta\ \mathrm{d}\,t$ to transmit the infection (*i.e.*, the neighbour becomes infectious).

In addition to the susceptible and infectious intrinsic states, the nodes of an *SIR model* may also be *Removed* (once had the infection and can neither acquire nor transmit it ever again) and there are thus $\mathcal{N} = 3$ accessible intrinsic node states. The rules V are the same than for the SIS model with respect to infection (*i.e.*, infectious nodes transmit to their susceptible neighbour with probability $\beta\ \mathrm{d}\,t$), but recovery is replaced by removal (*i.e.*, infectious nodes become removed with probability $\alpha\ \mathrm{d}\,t$).

The remainder of this section studies how different choices of state vector $\mathbf{x}(t)$ and prior information Y translate in the rules W of the simplified system. Each case corresponds to a different model where W is defined through a set of equations whose tags all share the same numeral [*e.g.*, (5a)–(5f)]. We thus define numerous different $\mathbf{x}(t)$, $\mathbf{r}\big(\mathbf{x}(t)\big)$, $q_j^+\big(\mathbf{x}(t), Y\big)$, $q_j^-\big(\mathbf{x}(t), Y\big)$ etc., which must all be understood within the scope of their respective equation tag numeral. Although figures present results concomitantly with the specification of the corresponding models, all discussions are delayed to Sect. 4.

3.1 Pair-based SIS model

Section 2.4 defined a pair motif as two linked nodes. Since the nodes of a SIS model are either susceptible or infectious, there are three possibilities for pair motifs: two linked susceptible nodes (noted $S{-}S$), two linked infectious nodes (noted $I{-}I$) and a susceptible node linked to an infectious one (noted $S{-}I$). Two nodes involved in a pair motif may have other neighbours.

Pair motifs are often used in conjunction with *node motifs*: the trivial structure that is one node. In the SIS model, there are two possibilities for a node motif: susceptible nodes (noted S) and infectious nodes (noted I). A state vector $\mathbf{x}(t)$ based on both node and pair motifs would thus be composed of five elements enumerating the amount of times each motif appears in the network: $x_S(t)$, $x_I(t)$, $x_{S-S}(t)$, $x_{S-I}(t)$ and $x_{I-I}(t)$. However, additional assumptions about the structure of the network may cause some of these quantities to be redundant.

3.1.1 Degree-regular random network

We first consider the simple case where the network is known to be a κ*-regular random network* of size N: there are N nodes in the network which all have κ neighbours (degree κ), and each such neighbours are chosen randomly.[6] Such a network must respect the structural constraints $x_S(t) = N - x_I(t)$, $x_{S-S}(t) = \frac{1}{2}\big[\kappa x_S(t) - x_{S-I}(t)\big]$ and $x_{I-I}(t) = \frac{1}{2}\big[\kappa x_I(t) - x_{S-I}(t)\big]$. Hence, with the prior information Y specifying N and κ, the state vector

$$\mathbf{x}(t) = \big(x_I(t), x_{S-I}(t)\big) \tag{5a}$$

suffices to obtain all the five node and pair motifs.

In those terms, the rules V specify that an infection has probability $\beta\, x_{S-I}(t)\ \mathrm{d}\,t$ to occur during the time interval $[t, t + \mathrm{d}\,t)$ while a recovery has probability $\alpha\, x_I(t)\ \mathrm{d}\,t$ to occur. Clearly, an infection translates to the destruction of a S motif and the creation of a new I one, and a recovery corresponds to the inverse process. However, pair motifs are also affected by such transitions since the affected node had neighbours. Hence, the effect on $\mathbf{x}(t)$ of the infection or recovery of a node depends on some information that is not directly tracked by $\mathbf{x}(t)$—*i.e.*, what is the state of the infected or recovered node's neighbours—and we thus have to *infer* this information from the available data.

In order to facilitate this inference, we define the *first neighbourhood motif* $S\Gamma_1(k_S, k_I)$ as a susceptible node that has k_S susceptible neighbours and k_I infectious neighbours. Similarly, the motif $I\Gamma_1(k_S, k_I)$ corresponds to an infectious node with k_S susceptible neighbours and k_I infectious ones. In both cases, we qualify as *central* the node whose neighbours are explicitly stated. The other nodes of the first neighbourhood motif, *i.e.*, the neighbours of the central node, may have other neighbours of their own.

We can now define a forward transition event of type $j \in \{0, 1, \ldots, n\}$ as the infection of the central node of a $S\Gamma_1(\kappa - j, j)$ motif. In terms of node and pair motifs, this implies the destruction of one of the S motifs, of $\kappa - j$ of the $S{-}S$ motifs and of j of the $S{-}I$ motifs together with the creation of one new I motif, of $\kappa - j$ new $S{-}I$ motifs and of j new $I{-}I$ motifs. Since only $x_I(t)$ and $x_{S-I}(t)$ are tracked, the shift vectors are[7]

$$\mathbf{r}^j = \big(1, \kappa - 2j\big). \tag{5b}$$

[6] This is thus a special case of the configuration model, defined later, for which $n_\kappa = N$ and $n_{\kappa'} = 0\ \forall \kappa' \neq \kappa$.

[7] Compare (5b) with (5a): if a forward transition of type j occurs at time t, we get $\mathbf{x}(t + \mathrm{d}\,t) = \mathbf{x}(t) + \mathbf{r}^j$, so $x_I(t + \mathrm{d}\,t) = x_I(t) + 1$ and $x_{S-I}(t + \mathrm{d}\,t) = x_{S-I}(t) + \kappa - 2j$. This event has probability $q_j^+\big(\mathbf{x}(t), Y\big)\,\mathrm{d}\,t$ to occur during the time interval $[t, t + \mathrm{d}\,t)$.

This same vector also defines the backward transition events $j \in \{0, 1, \ldots, \kappa\}$ which correspond to the recovery of the central node of a $I\Gamma_1(\kappa - j, j)$ motif.[8]

Looking back at the rules V, the corresponding forward and backward transition rate functions are

$$q_j^+\big(\mathbf{x}(t), Y\big) = \beta\, x_{S-I}(t)\, \mathbb{P}(S\Gamma_1(\kappa - j, j)|S, j \geq 1, \mathbf{x}(t), Y) \tag{5c}$$

$$q_j^-\big(\mathbf{x}(t), Y\big) = \alpha\, x_I(t)\, \mathbb{P}(I\Gamma_1(\kappa - j, j)|I, \mathbf{x}(t), Y) \tag{5d}$$

where two inference terms have been defined.

The inference term of (5d) gives the probability for a motif to be a $I\Gamma_1(\kappa - j, j)$ knowing that it has an infectious node at its center and that the current state vector is $\mathbf{x}(t)$ with prior information Y. For a sufficiently large network,[9] this approximately corresponds to randomly drawing the neighbors (i.e., κ independent sampling) of the central infectious node among the pair motifs $S{-}I$ and $I{-}I$ {i.e., each sampled neighbor has probability $[x_{S-I}(t)]/[\kappa\, x_I(t)]$ to be susceptible}

$$\mathbb{P}(I\Gamma_1(\kappa - j, j)|I, \mathbf{x}(t), Y) = \binom{\kappa}{j} \left(\frac{x_{S-I}(t)}{\kappa\, x_I(t)}\right)^{\kappa - j} \left(1 - \frac{x_{S-I}(t)}{\kappa\, x_I(t)}\right)^{j}. \tag{5e}$$

The inference term of (5c) is very similar except that the central susceptible node is known to have at least one infectious neighbours since it acquired the infection through a $S{-}I$ motif

$$\begin{aligned}&\mathbb{P}(S\Gamma_1(\kappa - j, j)|S, j \geq 1, \mathbf{x}(t), Y)\\ &\quad= \binom{\kappa - 1}{j} \left(\frac{x_{S-I}(t)}{\kappa\,(N - x_I(t))}\right)^{j} \left(1 - \frac{x_{S-I}(t)}{\kappa\,(N - x_I(t))}\right)^{\kappa - 1 - j}.\end{aligned} \tag{5f}$$

Together, Eqs. (5a)–(5f) form the rules W for the pair-based SIS model on a κ-regular network of N nodes.

One could obtain the probability distribution for $\mathbf{x}(t)$ by directly solving Eq. (2). However, using the approximation given by the system (4) greatly simplifies this task. Figure 4 compares the results produced by this simplified model [defined by W, $\mathbf{x}(t)$ and Y; solved using the system (4)] to the corresponding full one [defined by V and $Z(t)$; solved by Monte Carlo simulations]. Figure 5 shows the probability distribution for the same data. Note that, although presented differently, this model corresponds to the one presented in Dangerfield et al. (2009); Fig. 3 is provided for comparison with (Dangerfield et al., 2009, Fig. 2(c)).

[8] Recall that a backward transition of type j occurring at time t has the effect $\mathbf{x}(t + \mathrm{d}\,t) = \mathbf{x}(t) - \mathbf{r}^j$. This event has probability $q_j^-\big(\mathbf{x}(t), Y\big)\,\mathrm{d}\,t$ to occur during the time interval $[t, t + \mathrm{d}\,t)$.

[9] The actual distribution should be obtained by sampling *without* replacement, but for large networks we may make the approximation that the sampling is done *with* replacement, hence resulting in a binomial distribution.

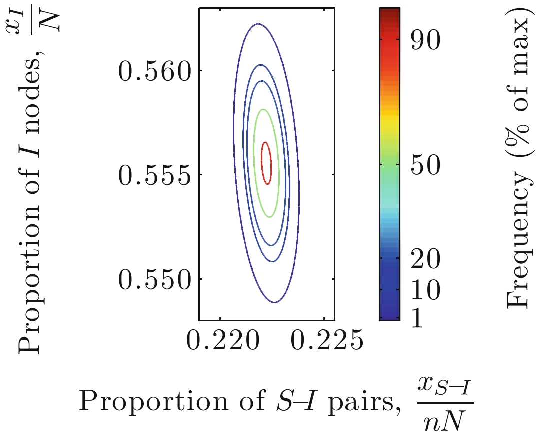


Fig. 3 (Online version in colour.) Distribution of post-transient ($t \to \infty$) outcomes as predicted by the SIS model on regular network [Eq. 5] using the approximations of equations (4). The axes (proportion of I among node motifs versus proportion of $S{-}I$ among pair motifs), network structure ($N = 10^5$ nodes, each of degree $\kappa = 5$) and parameters ($\alpha = 0.1$ (Keeling confirmed us that $\alpha = 0.01$, noted γ in Dangerfield et al. 2009, is a typo.) and $\beta = 0.05$) are the same as for (Dangerfield et al., 2009, Fig. 2(c)). Frequencies (in percent) are used to facilitate comparison with Dangerfield et al. (2009); they are simply obtained from the probability densities of (4f) multiplied by 100

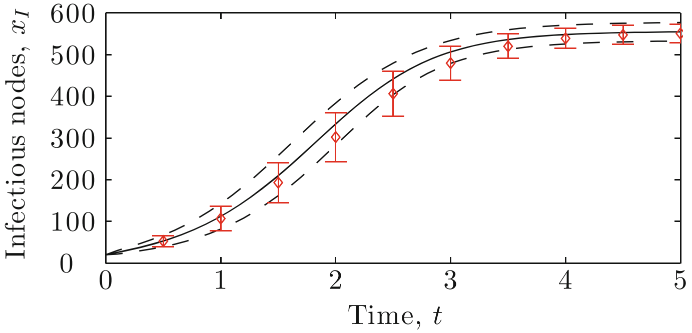


Fig. 4 (Online version in colour.) Time evolution of the number of infectious nodes x_I for SIS dynamics ($\alpha = 2$ and $\beta = 1$) on a regular network of $N = 10^3$ nodes (20 initially infectious) of degree $n = 5$. *Curves* results for the simplified system (5) approximated with Eq. (4). The continuous curve shows the mean value while the dashed curves delimit the range of one standard deviation above and below the mean. *Symbols* averaged results of 10^5 Monte Carlo simulations of the full system. The parameters α and β correspond to those of Fig. 3 after rescaling the time unit

3.1.2 Erdős-Rényi network

We now consider the case where the network is an Erdős-Rényi network: there are N nodes in the networks and M links are randomly assigned. This knowledge constrains two of the five node and pair motifs [*i.e.*, $x_S(t) = N - x_I(t)$ and $x_{S-S}(t) = M - x_{S-I}(t) - x_{I-I}(t)$] and a state vector of three elements suffices

$$\mathbf{x}(t) = \big(x_I(t), x_{S-I}(t), x_{I-I}(t)\big). \tag{6a}$$

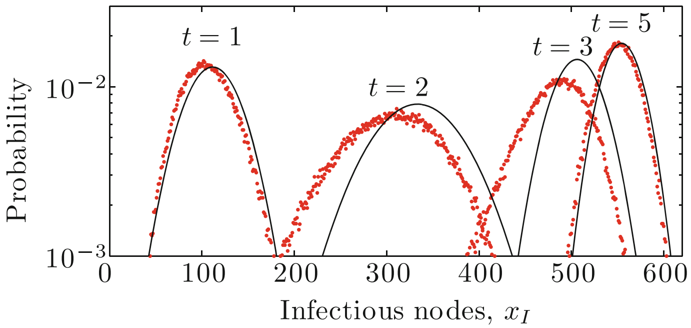


Fig. 5 (Online version in colour.) Probability distribution at different times for the number of infectious nodes $x_I(t)$. All parameters are the same as in Fig. 4. *Curves* numerical integration of (4) for the simplified system. *Symbols* binned results of 10^5 Monte Carlo simulations of the full system

The method used in Sect. 3.1.1 has to be adapted since the degree of each node is not constrained to a single value. Indeed, a susceptible node that gets infected may a priori be the center of any of the $S\Gamma_1(k_S, k_I)$ motifs. Still, we could design a bijective mapping between the vector of integers $\mathbf{k} = (k_S, k_I)$ and an event type j.

The details of the chosen mapping do not matter: we simply define the forward transition event of type $\mathbf{k}$ as the infection of the central node of a $S\Gamma_1(k_S, k_I)$ motif, which may conveniently be noted $S\Gamma_1(\mathbf{k})$ instead. Similarly, the backward transition event of type $\mathbf{k}$ is defined as the recovery of the central node to a $I\Gamma_1(\mathbf{k})$ motif. The corresponding shift vector and rate functions are

$$\mathbf{r}^{\mathbf{k}} = \big(1, k_S - k_I, k_I\big) \tag{6b}$$

$$q_{\mathbf{k}}^{+}\big(\mathbf{x}(t), Y\big) = \beta\, x_{S-I}(t)\, \mathbb{P}(S\Gamma_1(\mathbf{k})|S, k_I \geq 1, \mathbf{x}(t), Y) \tag{6c}$$

$$q_{\mathbf{k}}^{-}\big(\mathbf{x}(t), Y\big) = \alpha\, x_I(t)\, \mathbb{P}(I\Gamma_1(\mathbf{k})|I, \mathbf{x}(t), Y) \tag{6d}$$

where the inference terms bear the same meaning as their previous counterpart. Again assuming large network size, these inference terms are

$$\begin{aligned}\mathbb{P}(S\Gamma_1(\mathbf{k})|S, l \geq 1, \mathbf{x}(t), Y) = &\binom{2\,x_{S-S}(t)}{k_S} x_S(t)^{-k_S}\big(1 - x_S(t)^{-1}\big)^{2\,x_{S-S}(t) - k_S} \\ &\times \binom{x_{S-I}(t) - 1}{k_I - 1} x_S(t)^{-(k_I - 1)}\big(1 - x_S(t)^{-1}\big)^{x_{S-I}(t) - k_I}\end{aligned} \tag{6e}$$

$$\begin{aligned}\mathbb{P}(I\Gamma_1(\mathbf{k})|I, \mathbf{x}(t), Y) = &\binom{x_{S-I}(t)}{k_S} x_I(t)^{-k_S}\big(1 - x_I(t)^{-1}\big)^{x_{S-I}(t) - k_S} \\ &\times \binom{2\,x_{I-I}(t)}{k_I} x_I(t)^{-k_I}\big(1 - x_I(t)^{-1}\big)^{2\,x_{I-I}(t) - k_I}.\end{aligned} \tag{6f}$$

These products of binomial distributions evaluate the probability for each pair motif to include the considered central node. Taking the example of the first half of equa-

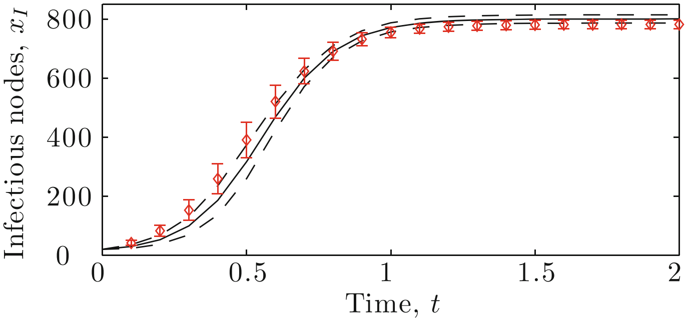


Fig. 6 (Online version in colour.) Time evolution of the number of infectious nodes $x_I(t)$ for SIS dynamics ($\alpha = 2$ and $\beta = 1$) on an Erdős-Rényi network of $N = 10^3$ nodes (20 initially infectious) and $M = 5 \cdot 10^3$ links. The mean and range of one standard deviation above and bellow the mean are shown. *Curves* numerical integration of (4) for the simplified system. *Symbols* full system (10^5 Monte Carlo simulations)

tion (6f), each $S{-}I$ pair motif contains one infectious node, so each one of them has probability $x_I(t)^{-1}$ to be the considered central node (so the number of "successes" corresponds to k_S). Together, Eqs. (6a)–(6f) form the rules W for the pair-based SIS model on a Erdős-Rényi network of size N with M links.

Figure 6 compares the results produced by this simplified model to the corresponding full one. Sect. 4.2 discusses these results and provides further details concerning pair-based models.

3.2 First neighbourhood SIS model

We consider a full model (V and $Z(t)$) for SIS dynamics on a *configuration model* (CM) network: given a sequence $\{n_0, n_1, n_2, \ldots\}$, links are randomly assigned between nodes such that, for each degree κ, there are n_κ nodes of degree κ. In a computer simulation, we create n_κ nodes with κ stubs for each possible κ and then randomly pair stubs to form links. No particular mechanism is used to prevent the formation of repeated links and self-loops: this simplifies the analytical treatment and has little effect when the network size is sufficiently large.

For networks with heterogeneous degree distributions (i.e., there are nodes with degree much higher and/or lower than the average), basic node and pair motifs usually do not appropriately capture the state of the system.[10] Although one could generalize the principle of pair motifs to account for the degree of a node (see Appendix A.5), we here prefer to handle the heterogeneity in node degree by enumerating every possible first neighbourhood motifs in the state vector

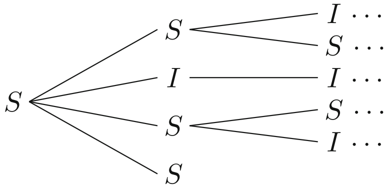
$$\mathbf{x}(t) = \big(x_{S\Gamma_1(0,0)}(t), x_{S\Gamma_1(1,0)}(t), \ldots, x_{I\Gamma_1(0,0)}(t), \ldots\big). \tag{7a}$$

[10] For example, a node of high degree is much more likely to acquire the infection, and is much more dangerous once it has acquired it, than a low degree node.

Although this vector should be infinite in the general case, it is not the case when, *e.g.*, the prior information Y states that no node has a degree superior to $\mathcal{K}$.

For the same reasons that models tracking node and pair motifs (Sect. 3.1) had their transition events defined in terms of first neighbourhood motifs, the transition events are here defined in terms of *second neighbourhood motifs*: a central node, its neighbours and the neighbours of those neighbours. In the same way that we note $\nu\Gamma_1(\mathbf{k})$ the first neighbourhood motif formed by a state ν central node with neighbourhood specified by $\mathbf{k}$, we note $\nu\Gamma_2(\mathbf{K})$ the second neighbourhood motif formed by a state ν central node with neighbourhood specified by $\mathbf{K}$.

The elements of $\mathbf{K}$ may be indexed with first neighbourhood motifs: the central node has $K_{\nu'\Gamma_1(\mathbf{k}')}$ state ν' neighbours whose other neighbours (*i.e.*, excluding the central node) are specified by $\mathbf{k}'$. Hence, the second neighbourhood motif



(second neighborhood example)

is noted $S\Gamma_2(\mathbf{K})$ with all elements of $\mathbf{K}$ zero except for $K_{S\Gamma_1(0,0)} = 1,\ K_{I\Gamma_1(0,1)} = 1$ and $K_{S\Gamma_1(1,1)} = 2$. Note that the central node of this second neighbourhood motif is also the central node of the first neighbourhood motif $S\Gamma_1(3, 1)$. In general, we note $\nu\widetilde{\Gamma}_1(\mathbf{K})$ the first neighbourhood motif that shares the same central node as the second neighbourhood motif $\nu\Gamma_2(\mathbf{K})$.

We digress further to introduce the *unit vector* notation $\widehat{\mathbf{e}}_m$ where m represents a motif; all the elements of this vector are zero except for the mth, which is one. The total number of elements in $\widehat{\mathbf{e}}_m$ should be clear from the context. As a concrete example, the right hand side of (5b) could be noted $\widehat{\mathbf{e}}_I + (n - 2j)\widehat{\mathbf{e}}_{S-I}$.

Similarly to Sect. 3.1.2, we define the forward transition event of type $\mathbf{K}$ to be the infection of the central node of a $S\Gamma_2(\mathbf{K})$ motif and the backward transition event of type $\mathbf{K}$ as the recovery of the central node of a $I\Gamma_2(\mathbf{K})$ motif. The corresponding shift vector is

$$\mathbf{r}^{\mathbf{K}} = \widehat{\mathbf{e}}_{I\widetilde{\Gamma}_1(\mathbf{K})} - \widehat{\mathbf{e}}_{S\widetilde{\Gamma}_1(\mathbf{K})} + \sum_{\nu}\sum_{\mathbf{k}} K_{\nu\Gamma_1(\mathbf{k})}\left(\widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_I)} - \widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_S)}\right) \tag{7b}$$

(Examples in the following assume the infection of the node at the center of the second neighborhood motif depicted on page 17.) The first two terms show the direct effect of a change of state in the central node [example: one less $S\Gamma_1(3, 1)$ and one more $I\Gamma_1(3, 1)$] while the sums handle the "collateral effect" on its immediate neighbours [example: one less $S\Gamma_1(2, 1)$ and one more $S\Gamma_1(3, 0)$; one less $I\Gamma_1(1, 1)$ and one more $I\Gamma_1(0, 2)$; one less $S\Gamma_1(2, 1)$ and one more $S\Gamma_1(1, 2)$; and one less $S\Gamma_1(1, 0)$ and one more $S\Gamma_1(0, 1)$]. Here the unit vector $\widehat{\mathbf{e}}_\nu$ has the same dimension as $\mathbf{k}$ (*i.e.*, two) while $\widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k})}$ has the same dimension as $\mathbf{x}$. Sums are taken over all the accessible values of ν and $\mathbf{k}$.

The corresponding rate functions are

$$q^{+}_{\mathbf{K}}\big(\mathbf{x}(t),Y\big)=\beta x_{S\widetilde{\Gamma}_1(\mathbf{K})}(t)\Big(\sum_{\mathbf{k}}K_{I\Gamma_1(\mathbf{k})}\Big)\mathbb{P}\big(S\Gamma_2(\mathbf{K})|S\widetilde{\Gamma}_1(\mathbf{K}),\mathbf{x}(t),Y\big) \tag{7c}$$

$$q^{-}_{\mathbf{K}}\big(\mathbf{x}(t),Y\big)=\alpha\, x_{I\widetilde{\Gamma}_1(\mathbf{K})}(t)\,\mathbb{P}\big(I\Gamma_2(\mathbf{K})|I\widetilde{\Gamma}_1(\mathbf{K}),\mathbf{x}(t),Y\big) \tag{7d}$$

[example: in Eq. (7c), $x_{S\widetilde{\Gamma}_1(\mathbf{K})}(t)$ is the number of $S\Gamma_1(3,1)$ motifs present at time t, and $\sum_{\mathbf{k}} K_{I\Gamma_1(\mathbf{k})} = 1$ because the central node of the second neighborhood pattern depicted on page 17 has only one infectious neighbor]. Note that, unlike (5c)–(5d) and (6c)–(6d), the inference terms in (7c)–(7d) have the same form: the probability for a motif to be a $\nu\Gamma_2(\mathbf{K})$ knowing [in addition to $\mathbf{x}(t)$ and Y] that its central node is also the central node of a $\nu\widetilde{\Gamma}_1(\mathbf{K})$ motif. Again assuming a large network size, they are provided by a product of multinomial distributions

$$\mathbb{P}\big(\nu\Gamma_2(\mathbf{K})|\nu\widetilde{\Gamma}_1(\mathbf{K}),\mathbf{x}(t),Y\big)$$
$$=\prod_{\nu'}\Big(\sum_{\mathbf{k}}K_{\nu'\Gamma_1(\mathbf{k})}\Big)!\prod_{\mathbf{k}}\frac{1}{(K_{\nu'\Gamma_1(\mathbf{k})})!}\underbrace{\left(\frac{(k_\nu+1)x_{\nu'\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_\nu)}(t)}{\sum_{\mathbf{k}'}k'_\nu x_{\nu'\Gamma_1(\mathbf{k}')}(t)}\right)}_{*}^{K_{\nu'\Gamma_1(\mathbf{k})}}. \tag{7e}$$

The main idea behind Eq. (7e) is to independently sample each first neighbors of the central node among the first neighborhood motifs of appropriate state ν'. More precisely, $K_{\nu'\Gamma_1(\mathbf{k})}$ is incremented by 1 for each type ν' neighbor selected this way that is the center of a $\nu'\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_\nu)$ motif, and the probability for this to occur is given by the term identified "$*$" in Eq. (7e). Together, Eqs. (7a)–(7e) form the rules W for the first neighbourhood SIS model.

Figure 7 compares the results produced by this simplified model and the full one. Note that this is a stochastic version of the model presented in Marceau et al. (2010), except that the network structure is here static.

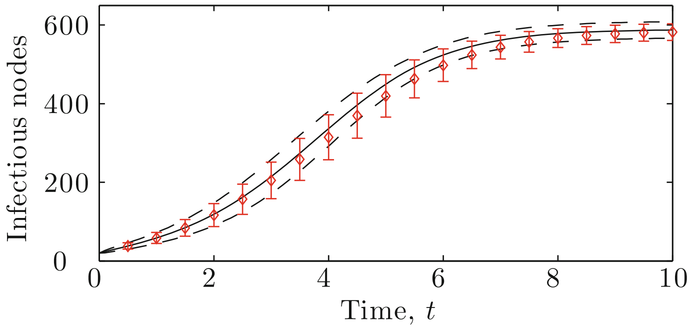


Fig. 7 (Online version in colour.) Time evolution of the number of infectious nodes for SIS dynamics ($\alpha = 1$ and $\beta = 1$) on a CM network for which the number of nodes of each degree is prescribed by the sequence {0, 50, 200, 450, 300} (total $N = 10^3$ nodes) with 2 % of the nodes of each degree initially infectious. The mean and range of one standard deviation above and bellow the mean are shown. *Curves* numerical integration of (4) for the simplified system. *Symbols* full system (10^5 Monte Carlo simulations)

3.3 First neighbourhood SIR model

As in Sect. 3.2, we consider a full network model where the network structure is specified solely by the degree of its nodes. However, this time we consider SIR epidemiological dynamics: the accessible node states are $\nu \in \{S, I, R\}$, infection is the same as in SIS but recovery is replaced by removal (see the introduction of Sect. 3 for details).

We define the forward transition event of type $\mathcal{I}\mathbf{K}$ to be the $\mathcal{I}$nfection of the central node of a $S\Gamma_2(\mathbf{K})$ motif while a forward transition event of type $\mathcal{R}\mathbf{K}$ is the $\mathcal{R}$emoval of the central node of a $I\Gamma_2(\mathbf{K})$ motif. There is no backward transition events. The model is specified by

$$\mathbf{x}(t) = \big(\ldots, x_{S\Gamma_1(\mathbf{k})}(t), \ldots, x_{I\Gamma_1(\mathbf{k})}(t), \ldots, x_{R\Gamma_1(\mathbf{k})}(t), \ldots\big) \tag{8a}$$

$$\mathbf{r}^{\mathcal{I}\mathbf{K}} = \widehat{\mathbf{e}}_{I\widetilde{\Gamma}_1(\mathbf{K})} - \widehat{\mathbf{e}}_{S\widetilde{\Gamma}_1(\mathbf{K})} + \sum_{\nu}\sum_{\mathbf{k}} K_{\nu\Gamma_1(\mathbf{k})}\big(\widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_I)} - \widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_S)}\big) \tag{8b}$$

$$\mathbf{r}^{\mathcal{R}\mathbf{K}} = \widehat{\mathbf{e}}_{R\widetilde{\Gamma}_1(\mathbf{K})} - \widehat{\mathbf{e}}_{I\widetilde{\Gamma}_1(\mathbf{K})} + \sum_{\nu}\sum_{\mathbf{k}} K_{\nu\Gamma_1(\mathbf{k})}\big(\widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_R)} - \widehat{\mathbf{e}}_{\nu\Gamma_1(\mathbf{k}+\widehat{\mathbf{e}}_I)}\big) \tag{8c}$$

$$q^+_{\mathcal{I}\mathbf{K}}\big(\mathbf{x}(t), Y\big) = \beta x_{S\widetilde{\Gamma}_1(\mathbf{K})}(t)\Big(\sum_{\mathbf{k}} K_{I\Gamma_1(\mathbf{k})}\Big)\mathbb{P}\big(S\Gamma_2(\mathbf{K})|S\widetilde{\Gamma}_1(\mathbf{K}), \mathbf{x}(t), Y\big) \tag{8d}$$

$$q^+_{\mathcal{R}\mathbf{K}}\big(\mathbf{x}(t), Y\big) = \alpha\, x_{I\widetilde{\Gamma}_1(\mathbf{K})}(t)\, \mathbb{P}\big(I\Gamma_2(\mathbf{K})|I\widetilde{\Gamma}_1(\mathbf{K}), \mathbf{x}(t), Y\big) \tag{8e}$$

$$q^-_{\mathcal{I}\mathbf{K}}\big(\mathbf{x}(t), Y\big) = q^-_{R\mathbf{K}}\big(\mathbf{x}(t), Y\big) = 0 \tag{8f}$$

where the inference terms are the same as in (7e).

3.4 First neighbourhood on-the-fly SIR model

We take a different perspective to the problem considered in Sect. 3.3 which requires to track much less elements in the state vector. Instead of considering "complete" first neighbourhood motifs, such as $\nu\Gamma_1(\mathbf{k})$, that specify the state of each of the central node's neighbours, we define the $\nu\Lambda_1(\kappa)$ motif as a central node of state ν for which we know that it has κ neighbours *unknown to us*. This last statement is important: were we to learn the state of one of these neighbours, this would cease to be a $\nu\Lambda_1(\kappa)$ motif and instead become a $\nu\Lambda_1(\kappa - 1)$ one. As usual, the state vector tracks the number of such motifs

$$\mathbf{x}(t) = \big(\ldots, x_{S\Lambda_1(\kappa)}(t), \ldots, x_{I\Lambda_1(\kappa)}(t), \ldots, x_{R\Lambda_1(\kappa)}(t), \ldots\big)\,. \tag{9a}$$

We recall from Sect. 3.2 how a CM network is built in a computer simulation: for each κ, n_κ nodes with κ stubs are created and the stubs are then randomly paired to form links. From this perspective, $\nu\Lambda_1(\kappa)$ may be reinterpreted as a ν state node with κ unpaired stubs: as stubs are removed once they are paired in the computer simulation, neighbours that were unknown are removed from these motifs once they become known to us. Hence,

$$\frac{\kappa x_{\nu\Lambda_1(\kappa)}(t) - \delta_{\nu\nu'}\delta_{\kappa\kappa'}}{\sum_{\nu''}\sum_{\kappa''}\kappa'' x_{\nu''\Lambda_1(\kappa'')}(t) - 1}$$

exactly gives the probability for an unknown neighbours of the central node of $\nu'\Lambda_1(\kappa')$ to be the central node of $\nu\Lambda_1(\kappa)$. Note that the Kronecker deltas $\left(\delta_{ii'} = \begin{cases} 1 & i = i' \\ 0 & i \neq i' \end{cases}\right)$ in the numerator and the -1 in the denominator both account for the stub of $\nu'\Lambda_1(\kappa')$ that we are pairing with a random stub.

A typical computer simulation would first build the network and then perform the SIR propagation dynamics on this network. However, we do not want to have to store the network structure for later consultation, which would require additional space in $\mathbf{x}(t)$. Instead, we delay the network construction, leaving the stubs unpaired, and start the propagation dynamics right away. Just when the state of an unknown neighbour is required do we pair the corresponding stub with a randomly selected one, hence building the network *on-the-fly*. Since the knowledge of stubs being matched will be lost in the future, this information must only be required at the very moment it is obtained if we want the resulting dynamics to *exactly* reproduce the behaviour of the full system.

We thus take a different, although equivalent, perspective on the infection dynamics where each link is "probed" at most once. Instead of considering a probability $\beta\, dt$ of infection for each *susceptible* neighbours of infectious nodes, we consider the same probability for each of their *unknown* neighbours. Only when this probability returns true do we wonder about the state of the neighbour, whose state changes to infectious if and only if it was previously susceptible. In any case, we learned who were the neighbours of two nodes (*i.e.*, the infectious and its neighbour) and we must update the state vector accordingly.

Hence, we define the $\mathcal{I}$nfection transition event $\mathcal{I}\nu\kappa\kappa'$ such that an infectious at the center of a $I\Lambda_1(\kappa')$ motif attempts to infect the ν-state node at the center of a $\nu\Lambda_1(\kappa)$ motif. Of course, only $\mathcal{I}S\kappa\kappa'$ transition events result in real infections. The more traditional transition event $\mathcal{R}\kappa$ corresponds to the $\mathcal{R}$emoval of the infectious node at the center of a $I\Lambda_1(\kappa)$ motif, thus becoming $R\Lambda_1(\kappa)$. The model is specified by

$$\mathbf{r}^{\mathcal{I}S\kappa\kappa'} = \widehat{\mathbf{e}}_{I\Lambda_1(\kappa-1)} - \widehat{\mathbf{e}}_{S\Lambda_1(\kappa)} + \widehat{\mathbf{e}}_{I\Lambda_1(\kappa'-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa')} \tag{9b}$$

$$\mathbf{r}^{\mathcal{I}I\kappa\kappa'} = \widehat{\mathbf{e}}_{I\Lambda_1(\kappa-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa)} + \widehat{\mathbf{e}}_{I\Lambda_1(\kappa'-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa')} \tag{9c}$$

$$\mathbf{r}^{\mathcal{I}R\kappa\kappa'} = \widehat{\mathbf{e}}_{R\Lambda_1(\kappa-1)} - \widehat{\mathbf{e}}_{R\Lambda_1(\kappa)} + \widehat{\mathbf{e}}_{I\Lambda_1(\kappa'-1)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa')} \tag{9d}$$

$$\mathbf{r}^{\mathcal{R}\kappa} = \widehat{\mathbf{e}}_{R\Lambda_1(\kappa)} - \widehat{\mathbf{e}}_{I\Lambda_1(\kappa)} \tag{9e}$$

$$q^{+}_{\mathcal{I}\nu\kappa\kappa'}\big(\mathbf{x}(t), Y\big) = \beta\,\kappa'\, x_{I\Lambda_1(\kappa')}(t)\frac{\kappa x_{\nu\Lambda_1(\kappa)}(t) - \delta_{I\nu}\delta_{\kappa\kappa'}}{\sum_{\nu''}\sum_{\kappa''}\kappa'' x_{\nu''\Lambda_1(\kappa'')}(t) - 1} \tag{9f}$$

$$q^{+}_{\mathcal{R}\kappa}\big(\mathbf{x}(t), Y\big) = \alpha x_{I\Lambda_1(\kappa)}(t) \tag{9g}$$

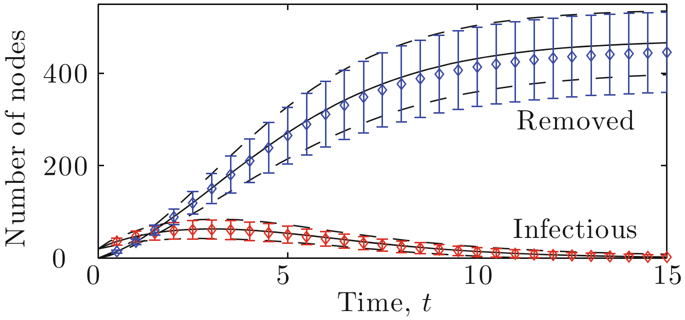


Fig. 8 (Online version in colour.) Time evolution of the number of infectious and removed nodes for SIR dynamics ($\alpha = 1$ and $\beta = 1$) on a CM network for which the number of nodes of each degree is prescribed by the sequence {0, 50, 200, 450, 300} (total $N = 10^3$ nodes) with 2 % of the nodes of each degree initially infectious (all others are susceptible). The mean and range of one standard deviation above and bellow the mean are shown. *Curves* numerical integration of (4) for the simplified system. *Symbols* full system (10^5 Monte Carlo simulations)

$$q^{-}_{\mathcal{I}\nu\kappa\kappa'}\big(\mathbf{x}(t), Y\big) = q^{-}_{\mathcal{R}\kappa}\big(\mathbf{x}(t), Y\big) = 0. \tag{9h}$$

The system (9) *exactly* reproduces the behaviour of the full system through the solution of (2). Since the Eq. (4) are only approximations of (2), results obtained through these relationships are only approximative (Figs. 8, 9). This model may be solved analytically for the mean value (see Appendix B) and the results are in agreement with Volz (2008), Miller (2010). This is a generalization to the case $\alpha \neq 0$ of the model presented in Noël et al. (2012). Further details are discussed in Sect. 4.4. We note that a conceptually similar ideas are presented in a deterministic context (Ball and Neal 2008), and more recently in Decreusefond et al. (2012) as a tool for a mathematically rigorous proof that a specific heterogeneous mean field model (Volz 2008) holds in the limit of large network size.

4 Discussion

We now take a retrospective look at the results presented in Sect. 3 and obtain from these special examples general considerations concerning our modelling approach.

4.1 Accuracy of the results

One of the aims of this paper is to obtain simplified models that accurately reproduce the behaviour of complex systems. Since approximations are usually involved, it is to be expected that the results of the simplified model only agree with those of the full system over some range of parameters, where the approximations were valid.

The parameters used in Figs. 4, 5, 6, 7, 8, 9 were chosen in order to investigate the limits of our approximations: while there is no perfect correspondence between the results of the full and simplified systems, their agreement is probably sufficient for both qualitative and quantitative applications. We distinguish between two categories

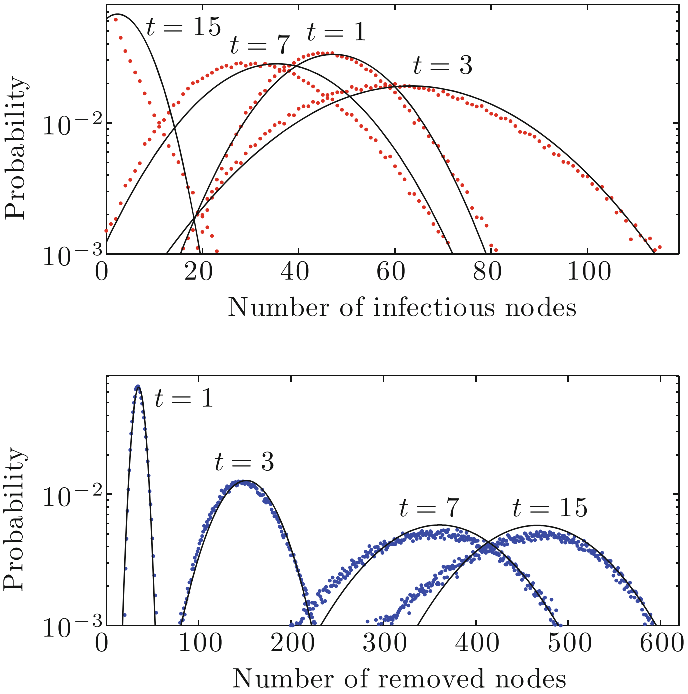


Fig. 9 (Online version in colour.) Probability distribution at different times for the number of infectious and removed nodes. The parameters are the same as in Fig. 8. *Curves* numerical integration of (4) for the simplified system. *Symbols* full system (10^5 Monte Carlo simulations)

of approximations: those inherent to the use of Eq. (4) and those due to the imperfect representation of $Z(t)$ through $\mathbf{x}(t)$ and Y.

4.1.1 Gaussian approximation

Since (2) and (9) define a system that *exactly* reproduces the behaviour of the corresponding full system, any discrepancy in Fig. 8 must originate from the use of the Gaussian approximation (4). An important requirement for this approximation to be valid is that the size N of the system must be large.

Figures 4, 5, 6, 7, 8, 9 all use networks of size $N = 1000$. As a rule of thumb, we found that (4) perform better for networks of at least a few hundred nodes, which is the case of many relevant real-world systems. Note that, for very small systems (tens of nodes), one could also directly and completely solve (2).

While a large network size N is required to justify treating the elements of $\mathbf{x}(t)$ as real numbers, other phenomena may affect the validity of this approximation. For example, when the initial conditions are such that there is a single infectious node, the continuous approximation fails at considering the probability for that node to recover (or to be removed) before transmitting the infection to one of its neighbours. Figures 4, 5, 6, 7, 8, 9 circumvent this problem by using an initial condition with 20 infectious nodes: the probability for all of them to recover (or to be removed) before transmitting the infection is very low.

It is worth noting that the plateaux seen on Figs. 4, 6, 7 and 8 reflect different dynamical behaviours for the SIS and SIR systems. Indeed, while the total number of removed nodes reaches a maximum in the SIR system because there are no infectious left to recover, the steady state observed at the later times for our SIS models corresponds to a constant flow of recovery and new infections. In the former case, the approximation errors performed at earlier times accumulate. In the later, the exact path taken to attain equilibrium is of lesser importance and errors do not accumulate the same way.

4.1.2 Representation approximation

In general, the simplified system will not exactly reproduce the behaviour of the full system, even when using (2) instead of (4). This is the case of all our SIS models: while some of the discrepancy seen in Figs. 4, 5, 6, 7 is explained by the Gaussian approximation, the imperfect representation of Z also contributes to the error.

Part of the problem can be understood as our failure to consider the correlation between the neighbours of a node and the time elapsed since this node has been in its present intrinsic state. For example, the neighbours of a susceptible node that has just recovered (*i.e.*, it was infectious a moment ago) may be much different than those of a susceptible node that has recovered a long time ago, while being similar to those of a node that is still infectious. Hence, one could hope to improve these SIS models through changes in Y alone (*i.e.*, with the same $\mathbf{x}(t)$): first estimate the probability distribution for the time since when each node has last changed state and then infer the neighbourhoods accordingly. An alternative that could be simpler to implement, at the cost of increasing the size of $\mathbf{x}(t)$, would consist in tracking more exhaustive motifs (*e.g.*, second neighbourhoods instead of first ones in Sect. 3.2).

However, there are more intricate consequences to the recovery of infectious nodes on a structure that is fixed in time: if at some point all the nodes of the same component (*i.e.*, a connected subnetwork that is disconnected from the rest of the network) are susceptible at the same time, then none of them may ever become infectious again. The connectivity of a network is strongly affected by the average degree of its nodes: our parameters correspond to an average degree of 5 for Figs. 4, 5, 6 (average degree of a neighbour also 5) and of 3 for Fig. 7 (average degree of a neighbour ≈ 3.23). When using smaller parameter values, this components-induced discrepancy becomes much larger since the simplified model then overestimates the number of infectious nodes. One could take the components into account by solving independent systems for each component (and merge the results afterwards) or by a clever adaptation of the inference process (see Sect. 4.6 for possible directions). Note that these effects are usually much less important when the network structure changes over time.

4.2 Pair-based models

Compared to the other models presented in Sect. 3, the two pair-based models of Sect. 3.1 use very small state vectors (*i.e.*, two or three elements). This is an important advantage of pair-based models in general: there are usually much less pair and node motifs than, *e.g.*, first neighbourhood ones, and tracking them thus requires much smaller $\mathbf{x}(t)$.

Although we limited our study of pair-based models to regular and Erdős-Rényi networks, more complex network structures could also be considered. In the same way that (5) and (6) differ mostly by their inference terms, obtaining good inference from the little information stored in $\mathbf{x}(t)$ is probably the principal challenge behind general and accurate pair-based stochastic models.

However, non-stochastic pair-based models are already possible on nontrivial network structures for SIR dynamics or, more generally, for processes such that a change in the state of one neighbour of a node can be treated as independent of that of another neighbour (SIS fails this assumption) (Miller et al. 2011). Knowing (in Y) that a system behaves in this manner greatly simplifies the inference process, and this is the main reason for the success of the SIR pair-based model for the evolution of mean values on CM networks that is presented in Volz (2008), Miller (2010). Whether or not the same approach may be used to obtain stochastic results is an open question.

4.3 First (and higher) neighbourhood models

By opposition, sufficiently accurate inference terms for first neighbourhood models are often straightforward to obtain. Although (7e) may be difficult to appreciate at first sight, it is the only inference term used in both Sects. 3.2 and 3.3. In fact, (7e) may well be the only inference term needed for generic first neighbourhood models for CM network structures.

Although first neighbourhood motifs are a “natural language” for expressing dynamics taking place on CM networks, they could also be used in the presence of other complex structures. This may be done through changes in $\mathbf{x}(t)$ and/or Y; see Appendix A for details.

The generality and ease of design of first-neighbourhood models comes at a cost: the state vector $\mathbf{x}(t)$ is typically much larger than it would be in an equivalent pair-model. How large is $\mathbf{x}(t)$ strongly depends on the maximal node degree present in the network and on the total number of accessible intrinsic node states (see Appendix A for details). For typical values of these quantities, this does not cause major problems for the evaluation of the mean: numerically solving (4b) requires an acceptable amount of resources even for an $\mathbf{x}(t)$ of dimension 10^6 and (4a) may often be simplified (*i.e.*, summed analytically).

However, evaluating the covariance matrix using (4e) may cause problems: unless analytical simplifications are possible, solving this system scales as the square of the number of elements in $\mathbf{x}(t)$. Future developments may decrease this bottleneck effect of the covariance matrix; see Sect. 4.5 for details. In any case, the size of $\mathbf{x}(t)$ may be decreased by “coarse graining” the number of links between the central node and its neighbours; see Appendix A.7 for details.

4.4 On-the-fly models

The on-the-fly model presented in Sect. 3.4 for SIR dynamics on CM networks *exactly* reproduces the behaviour of the full system. This is even more remarkable when one considers that the size of the state vector in the on-the-fly model is much smaller than in the alternative first neighbourhood model of Sect. 3.3. The reasons behind the success of the on-the-fly approach are similar to those discussed in Sect. 4.2 for the pair models presented in Miller et al. (2011), Volz (2008), Miller (2010): it is encoded in Y that, for each link, we *at most once* need to simultaneously know the state of the two nodes joined by that link (Noël et al. 2012).

The inference term (7e) is of "general purpose" in the sense that its Y does not provide information on the dynamical properties of the system, but only on how the motifs in $\mathbf{x}(t)$ may be interconnected. This is why both (7) and (8) rely on (7e).

However, the inference terms of (9) have a specific character: Y contains information about (9) itself. Any change to the dynamics implies changes in the inference terms, with no guarantee that an acceptable solution exists. In fact, (9) was *designed* with this problem in mind. In other words, we obtained a simple and reliable model at the cost of "pre-computations" in the design process. Of all the possibilities in model-space, the information acquired by pointing at this specific one is what replaces the reduced size of the state vector. The same could be said of the deterministic SIR pair-based model on CM networks presented in Volz (2008), Miller (2010).

By contrast with the case discussed in Sect. 4.3, the small size of the state vector here allows for evaluations of the covariance matrix through (4e), even when relatively high degree nodes are present. Alternatively, one may take advantage of the fact that, even for more complicated dynamics, the state vector of on-the-fly models can remain of manageable size for mean values calculations; see the introduction to Appendix A for the concrete example of Marceau et al. (2011).

4.5 Complicated states versus complex assumptions

Section 4.4 revealed an unexpected depth to Y: one may achieve models of similar levels of accuracy by trading off complexity in the assumptions for a reduction in the size of the state vector $\mathbf{x}(t)$. As an extreme example, if Y already gives the full behaviour of the system, then there is no need for tracking any information in $\mathbf{x}(t)$. Without reaching such extremes, our on-the-fly model and the deterministic SIR pair-based model presented in Volz (2008), Miller (2010) both demonstrate the benefits of investing some time in the assumptions of our models.

While these examples required case-by-case analysis, one may benefit from the same realization in a general context: a first simplified model (W, $\mathbf{x}(t)$ and Y) may generate the assumptions Y' to a different simplified model (W', $\mathbf{x}'(t)$ and Y'). For example, when some dynamical process (*e.g.*, SIS or SIR) occurs on a network whose structure changes in time independently from this dynamics, one could obtain a first model for the structure alone and then feed the results to the second model, handling the remaining dynamics. Even more generally, one could compensate for the higher computational requirements of (4e) by first solving (4b) on an elaborate model then feeding the resulting mean values to a simpler model for the sole purpose of estimating the covariance matrix.

4.6 Additional inference tools

While we introduced Y as a direct application of Bayes' rule, we have now seen that useful assumptions may be obtained by other means, including the solution of another system of the form (2). The next step in this direction would be to improve our inference process using alternative tools and models available to network science.

For example, branching processes (Newman et al. 2001) may be used to infer information concerning the connectivity and the components of the network structure. As discussed in Sect. 4.1.2, this point was a major shortfall of SIS models. This approach is even more interesting for the recently developed tools (Allard et al. 2009; Karrer and Newman 2010; Allard et al. 2012) that are particularly compatible with the motifs and intrinsic node state approach presented in this paper.

Another tool of considerable interest are exponential random networks (Park and Newman 2004). Indeed, these maximum entropy methods can simplify inferences that would have otherwise been prohibitively complex. Once again, this approach may be generalized to different kind of motifs and intrinsic node states (Rogers 2011).

5 Conclusion: general applicability

Although the examples of Sect. 3 focus on simple SIS and SIR dynamics, any specificity that could be modelled through a standard epidemiological compartmental model may a priori be considered by our approach: genders, age groups, vaccination, incubation period, disease phases, *etc*. Furthermore, population dynamics considerations may be accounted for in a straightforward manner. Assuming first neighbourhood motifs, births and deaths of individuals correspond to events adding and removing motifs, respectively. Similarly, changes in interaction patterns amount to events replacing the affected motifs by new ones. In fact, from the model's perspective, there is no important distinction between a change in the interaction structure of the population and a change in the node states: both are events affecting motifs.

Beyond the propagation of infections or parasites, an additional class of spreading processes is particularly relevant to population biology: the cascading extinctions of species in food webs (Rezende et al. 2007; Bascompte and Stouffer 2009; Dunne and Williams 2009). Our formalism is applicable to such problems by representing each species as a node and by using links to indicate feeding connections between species. While such cascades may require different rules than those of SIS or SIR dynamics, the general approach may still be adapted to this specific application. This illustrates how our philosophy has the potential to be applicable to any dynamics as long as the relevant information can be encoded through the structure of a network and the intrinsic state of its components.

The generality of our systematic approach and the fact that its assumptions are explicitly stated suggests that it could be used as a common ground for comparing existing models too complex for direct comparison. Indeed, by considering such an existing model as the full system (specified by V_1 and $Z_1(t)$), one may seek a simplified system (specified by W_1, $X_1(t)$ and Y_1) approximately reproducing the original model (over a sufficient range of parameters).

If some transition event (in W_1) appears essential, this may reveal an important feature of the original model; the same holds true for motifs (in $X_1(t)$) and prior knowledge (in Y_1). Moreover, assuming that this procedure has been done for a second existing model (specified by V_2 and $Z_2(t)$), one may directly compare their simplified version in a common framework, which will help identify the assumptions required for their description. Note that this perspective is similar to a commutation diagram

$$\begin{tikzcd}
(V_1, Z_1(t)) \arrow[r, leftrightarrow, "\text{Difficult to compare}", "/" marking] \arrow[d, Leftrightarrow, "\text{Equivalent behaviour}"'] & (V_2, Z_2(t)) \arrow[d, Leftrightarrow, "\text{Equivalent behaviour}"] \\
(W_1, X_1(t), Y_1) \arrow[r, leftrightarrow, "\text{Comparable}"] & (W_2, X_2(t), Y_2) \quad .
\end{tikzcd}$$

For example, if $X_1(t)$ and $X_2(t)$ encode the same information and if $Y_1 = Y_2$, then we know that the discrepancies between the two original models is imputable to the difference in the transition events. Finding a minimal set of changes to W_1 and/or W_2 causing both models to agree may then help identify the very cause of the discrepancies.

Acknowledgments The research team acknowledges to the Canadian Institutes of Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Fonds de recherche du Québec—Nature et technologies (FRQ–NT) for financial support. We are grateful to the anonymous referees that have helped improve our presentation.

Appendix A: More on motifs

We introduce new motifs and generalize those presented in the main text for intrinsic link types and directed (or semi-directed) networks; Table 1 summarizes the total number of motifs in each of these classes. We define $\mathcal{K}$ as the highest accessible node degree and we set $\mathcal{D} = 1$ for undirected, $\mathcal{D} = 2$ for directed and $\mathcal{D} = 3$ for semi-directed networks.

Of high practical importance is the fact that the entries of this table differ widely in their scaling behaviours. For example, the on-the-fly model (Marceau et al. 2011) for two interacting SIR dynamics each propagating on their own network structure uses $\mathcal{N} = 9$ (Cartesian product of the sub-states of two SIR dynamics), $\mathcal{L} = 3$ (links of the first network alone, links of the second network alone and overlapping links) and $\mathcal{D} = 1$ (undirected network). Looking up in Table 1, we see that this requires of the order of $\mathcal{K}^3$ on-the-fly (degreed node) motifs, where $\mathcal{K}$ denotes the highest accessible node degree. By opposition, implementing a first neighbourhood version of this model would require of the order of $\mathcal{K}^{27}$ motifs!

A.1 Pair motifs

We note $\nu \overset{\ell}{-} \nu'$ (resp. $\nu \xrightarrow{\ell} \nu'$) an undirected (resp. directed) pair motif formed of a state ν node linked through a state ℓ link to a state ν' node. In the case of directed motifs, the direction of the arrow usually specifies the "strongest causal effect" of this asymmetric interaction (although this needs not be the case). All the undirected and

Table 1 Number of motifs in selected classes

Motifs	Number
Node	$\mathcal{N}$
Pair	$\frac{1}{2}\mathcal{L}\mathcal{N}\big(\mathcal{D}\mathcal{N}+(\mathcal{D}-2)^2\big)$
Triple	$\frac{1}{2}\mathcal{D}\mathcal{L}\mathcal{N}^2(\mathcal{D}\mathcal{L}\mathcal{N}+1)$
Triangle	$\frac{1}{3}\mathcal{D}\mathcal{L}\mathcal{N}+\frac{1}{2}\mathcal{D}\big((\mathcal{D}-2)\mathcal{L}\mathcal{N}\big)^2+\frac{1}{6}(\mathcal{D}\mathcal{L}\mathcal{N})^3$
Degreed node	$\mathcal{N}\binom{\mathcal{K}+\mathcal{D}\mathcal{L}}{\mathcal{K}}$
Degreed pair	$\frac{\mathcal{L}\mathcal{N}}{2}\binom{\mathcal{K}-1+\mathcal{D}\mathcal{L}}{\mathcal{K}-1}\Big(\mathcal{D}\mathcal{N}\binom{\mathcal{K}-1+\mathcal{D}\mathcal{L}}{\mathcal{K}-1}+(\mathcal{D}-2)^2\Big)$
First neighbourhood (node)	$\mathcal{N}\binom{\mathcal{K}+\mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K}}$
First neighbourhood pair	$\frac{\mathcal{L}\mathcal{N}}{2}\binom{\mathcal{K}-1+\mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K}-1}\Big(\mathcal{D}\mathcal{N}\binom{\mathcal{K}-1+\mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K}-1}+(\mathcal{D}-2)^2\Big)$
Second neighbourhood (node)	$\mathcal{N}\binom{\mathcal{K}+\mathcal{D}\mathcal{L}\mathcal{N}\binom{\mathcal{K}-1+\mathcal{D}\mathcal{L}\mathcal{N}}{\mathcal{K}-1}}{\mathcal{K}}$

directed motifs are possible in a semi-directed network. We may omit the index ℓ over the links when $\mathcal{L}=1$.

A.2 Triple motifs

We note $\nu \overset{\ell}{-} \nu' \overset{\ell'}{-} \nu''$ an undirected triple motif formed of a state ν node linked through a state ℓ link to a state ν' node itself linked through a state ℓ' link to a state ν'' node. As for pair motifs, each of these nodes may have other neighbours than those that are explicitly specified. The notation directly generalizes to directed (*e.g.*, $\nu \overset{\ell}{\rightarrow} \nu' \overset{\ell'}{\leftarrow} \nu''$) and semi-directed (*e.g.*, $\nu \overset{\ell}{-} \nu' \overset{\ell'}{\rightarrow} \nu''$) triple motifs.

The term “2-star” is often used to refer to a triple motif for which both extremities (*e.g.*, the nodes of state ν and ν'' in the motif $\nu \overset{\ell}{-} \nu' \overset{\ell'}{-} \nu''$) are explicitly *forbidden* to be neighbours. In models that also use triangle motifs, 2-star motifs may explicit the absence of the last link that would form a triangle. Another common use of triple motifs comes in the inference process of (usually deterministic) pair based models.

A.3 Triangle motifs and other small subnetworks

Three nodes that are all neighbours of each other form a triangle motif. An horizontal bracket represents the additional link that would be missing in the analogous triple motif, *e.g.*,

$$\underbrace{\nu \overset{\ell}{-} \nu' \overset{\ell'}{-} \nu''}_{\ell''}$$

for an undirected network.

Triangle motifs are usually considered in models that should account for clustering. Their number may either directly be tracked in the state vector $\mathbf{x}(t)$ (House et al. 2009) or their implicit presence (stated in Y, *e.g.*, through a clustering coefficient) may be accounted for in the inference process (Keeling et al. 1997).

The same notation may be generalized to other motifs consisting of small subnetworks, *e.g.*, square motifs (House et al. 2009).

A.4 Clique motifs

A vague definition of a clique motif is "a subgroup of nodes that share more links among themselves than what could be expected otherwise for the same number of randomly selected nodes". In applications, one may refine this definition according to the specificities of the problem at hand, *e.g.*, "an Erdős-Rényi subnetwork (link probability p) of n_S susceptible nodes and n_I infectious nodes". Clique motifs are usually considered in models that should account for community structure (Hébert-Dufresne et al. 2010).

A.5 Degreed motifs

A degreed motif is a motif for which we know the degree of all the nodes forming the motif: a degreed node motif is a node of specified state and degree; a degreed link motif is two nodes of specified state and degree that are known to be neighbours; *etc*. The in-degree and out-degree are both specified in directed and semi-directed networks; the latter cases also specify undirected degrees. Likewise, degrees pertaining to different types of links are specified independently.

In the same way that pair motifs are usually combined with node motifs, degreed pair motifs are usually combined with degreed node motifs (Eames and Keeling 2002). Note that the on-the-fly motifs $\nu\Lambda_1(\kappa)$ presented Sect. 3.4 can be understood as a special case of degreed node motifs where the degree is replaced by the "degree to unknown nodes".

A.6 *n*th neighbourhood motifs

Similarly to degreed motifs, an nth neighbourhood motif is a motif for which we specify the state of all the nth neighbours of the nodes forming the original motif. Hence, the notation $\underline{\nu}\Gamma_1(\mathbf{k})$ $\big[$resp. $\underline{\nu}\Gamma_2(\mathbf{K})\big]$of the main text corresponds to a first (resp. second) neighbourhood *node* motif. These concepts are directly generalizable to types of links and to directed or semi-directed networks.

First neighbourhood node motifs can be understood as tracking the correlation between the state of a node and the state of all its neighbours. By opposition, degreed pair motifs track the correlation between the state and degree of two neighbouring nodes. While similar information may be obtained from both motif classes, a model based on one may perform better than a model based on the other depending on the characteristics of the full system to be modelled.

A.7 Coarse-grained degree and/or neighbourhood

Not all entries of Table 1 depend on the maximal degree $\mathcal{K}$, but those that do quickly increase for large $\mathcal{K}$. This is problematic since many real-world systems contain high degree nodes.

However, one may overcome this limitation by coarse-graining degrees through ranges: the range containing the degree of a node is specified instead of the degree itself. For example, given the ranges

$$\underbrace{[0,0]}_{\text{range 0}},\ \underbrace{[1,1]}_{\text{range 1}},\ \underbrace{[2,3]}_{\text{range 2}},\ \underbrace{[4,7]}_{\text{range 3}},\ \underbrace{[8,15]}_{\text{range 4}},\ \underbrace{[16,31]}_{\text{range 5}}\ \text{and}\ \underbrace{[32,63]}_{\text{range 6}},$$

we would say of a degree 23 node that its degree lies within range 5. Hence, for the purpose of evaluating the number of motifs in Table 1, one should here use $\mathcal{K} = 6$ (one less than the total number of ranges) instead of $\mathcal{K} = 63$ (highest representable degree).

While the previous example used powers of 2 for simplicity, a slower increase is probably desirable in most applications. However, since the number of neighbourhood and degreed motifs strongly depends on $\mathcal{K}$, even the slightest reduction in this number may be significant. Note that this coarse-graining method is of particular interest when the real-world data used to calibrate the model is already coarse-grained, which is commonly the case for census data.

Appendix B: Deterministic solution of on-the-fly SIR model

This section provides the deterministic solution of (9). We first rewrite (4b) for the specific case of (9)

$$\frac{\mathrm{d}\,\boldsymbol{\mu}(t)}{\mathrm{d}\,t} = \sum_{\nu}\sum_{\kappa}\sum_{\kappa'} \mathbf{r}^{\mathcal{I}\nu\kappa\kappa'} q^{+}_{\mathcal{I}\nu\kappa\kappa'}\big(\boldsymbol{\mu}(t), Y\big) + \sum_{\kappa} \mathbf{r}^{\mathcal{R}\kappa} q^{+}_{\mathcal{R}\kappa}\big(\boldsymbol{\mu}(t), Y\big) \tag{10}$$

then collect the contributions to $\mu_{S\Lambda_1(\kappa)}$, $\mu_{I\Lambda_1(\kappa)}$ and $\mu_{R\Lambda_1(\kappa)}$ (dropping the functional dependencies for brevity)

$$\frac{\mathrm{d}\,\mu_{S\Lambda_1(\kappa)}}{\mathrm{d}\,t} = -\sum_{\kappa'} q^{+}_{\mathcal{I}S\kappa\kappa'} \tag{11a}$$

$$\begin{aligned}\frac{\mathrm{d}\,\mu_{I\Lambda_1(\kappa)}}{\mathrm{d}\,t} &= \sum_{\kappa'} \left(q^{+}_{\mathcal{I}S(\kappa+1)\kappa'} + q^{+}_{\mathcal{I}I(\kappa+1)\kappa'} - q^{+}_{\mathcal{I}I\kappa\kappa'}\right) \\ &\quad + \sum_{\nu}\sum_{\kappa'} \left(q^{+}_{\mathcal{I}\nu\kappa'(\kappa+1)} - q^{+}_{\mathcal{I}I\kappa'\kappa}\right) - q^{+}_{\mathcal{R}\kappa}\end{aligned} \tag{11b}$$

$$\frac{\mathrm{d}\,\mu_{R\Lambda_1(\kappa)}}{\mathrm{d}\,t} = \sum_{\kappa'} \left(q^{+}_{\mathcal{I}R(\kappa+1)\kappa'} - q^{+}_{\mathcal{I}R\kappa\kappa'}\right) + q^{+}_{\mathcal{R}\kappa}. \tag{11c}$$

Using the definitions

$$\lambda = \sum_{\kappa} \kappa \mu_{I\Lambda_1(\kappa)} \quad \text{and} \quad \omega = \sum_{\nu} \sum_{\kappa} \kappa \mu_{\nu\Lambda_1(\kappa)} \tag{12}$$

where λ is the total number of stubs belonging to infectious nodes and ω is the total number of stubs in the system, (11) becomes

$$\frac{\mathrm{d}\,\mu_{S\Lambda_1(\kappa)}}{\mathrm{d}\,t} = -\frac{\beta\lambda\kappa\mu_{S\Lambda_1(\kappa)}}{\omega} \tag{13a}$$

$$\begin{aligned}\frac{\mathrm{d}\,\mu_{I\Lambda_1(\kappa)}}{\mathrm{d}\,t} &= \frac{\beta\lambda(\kappa+1)\mu_{S\Lambda_1(\kappa+1)}}{\omega} - \alpha\mu_{I\Lambda_1(\kappa)} \\ &\quad + \beta\left(1+\frac{\lambda}{\omega}\right)\Big((\kappa+1)\mu_{I\Lambda_1(\kappa+1)} - \kappa\mu_{I\Lambda_1(\kappa)}\Big)\end{aligned} \tag{13b}$$

$$\frac{\mathrm{d}\,\mu_{R\Lambda_1(\kappa)}}{\mathrm{d}\,t} = \frac{\beta\lambda}{\omega}\Big((\kappa+1)\mu_{R\Lambda_1(\kappa+1)} - \kappa\mu_{R\Lambda_1(\kappa)}\Big) + \alpha\mu_{I\Lambda_1(\kappa)}. \tag{13c}$$

Note that (9f) has been approximated by dropping the Kronecker delta in the numerator and the -1 in the denominator.

We now consider the evolution of the total number of stubs ω by summing the contributions from (13)

$$\frac{\mathrm{d}\,\omega}{\mathrm{d}\,t} = \sum_{\nu}\sum_{\kappa} \kappa \frac{\mathrm{d}\,\mu_{\nu\Lambda_1(\kappa)}}{\mathrm{d}\,t} = -2\beta\lambda. \tag{14}$$

One may understand (14) as "during the time interval $[t, t+\mathrm{d}\,t)$, each one of the λ stubs belonging to infectious nodes have probability $\beta\,\mathrm{d}\,t$ to be paired to another stub, thus causing a decrease by 2 of ω. Noting $\omega_0 = \omega(0)$ the total number of stubs in the initial condition, we introduce the change of variable

$$\theta = \sqrt{\frac{\omega}{\omega_0}} \quad \text{such that} \quad \frac{\mathrm{d}\,\theta}{\mathrm{d}\,t} = -\frac{\beta\lambda}{\theta\omega_0} \tag{15}$$

Notice that $t=0$ corresponds to $\theta = 1$ and that θ *decreases* with time. Using this change of variable in (13a) gives

$$\frac{\mathrm{d}\,\mu_{S\Lambda_1(\kappa)}}{\mathrm{d}\,\theta} = \frac{\kappa\mu_{S\Lambda_1(\kappa)}}{\theta} \tag{16}$$

which has the solution

$$\mu_{S\Lambda_1(\kappa)}(t) = x_{S\Lambda_1(\kappa)}(0)\big(\theta(t)\big)^{\kappa} \tag{17}$$

using the initial condition $\mu_{S\Lambda_1(\kappa)}(0) = x_{S\Lambda_1(\kappa)}(0)$.

For convenience, we define

$$f(\theta) = \sum_{\kappa} x_{S\Lambda_1(\kappa)}(0)\theta^{\kappa}. \tag{18}$$

Noticing that

$$\sum_{\kappa} \kappa(\kappa+1)\mu_{S\Lambda_1(\kappa+1)} = \theta^2 \sum_{\kappa}(\kappa-1)\kappa x_{S\Lambda_1(\kappa)}(0)\theta^{\kappa-2} = \theta^2 f''(\theta) \quad , \tag{19}$$

we obtain the evolution of the total number λ of stubs belonging to infectious nodes by summing the contributions (13b)

$$\frac{\mathrm{d}\,\lambda}{\mathrm{d}\,t} = \sum_{\kappa} \kappa \frac{\mathrm{d}\,\mu_{I\Lambda_1(\kappa)}}{\mathrm{d}\,t} = \frac{\beta\lambda\theta^2 f''(\theta)}{\omega} - \beta\lambda\left(1+\frac{\lambda}{\omega}\right) - \alpha\lambda. \tag{20}$$

Again using the change of variable (15), we get

$$\frac{\mathrm{d}\,\lambda}{\mathrm{d}\,\theta} = \frac{\lambda}{\theta} + \theta\omega_0\left(1+\frac{\alpha}{\beta}\right) - \theta f''(\theta) \tag{21}$$

which has the solution

$$\lambda = \theta^2\omega_0\left(1+\frac{\alpha}{\beta}\right) - \theta\omega_0\frac{\alpha}{\beta} - \theta f'(\theta) \tag{22}$$

for an initial condition without removed nodes $\left[i.e., \lambda(0) = \omega_0 - f'(1)\right]$.

Using this solution in (15) gives

$$\frac{\mathrm{d}\,\theta}{\mathrm{d}\,t} = -\beta\theta + \alpha\,(1-\theta) + \beta\frac{f'(\theta)}{\omega_0} \tag{23}$$

whose solution provides $\theta(t)$. Using

$$S(t) = f\big(\theta(t)\big) \tag{24a}$$
$$I(t) = N - S(t) - R(t) \tag{24b}$$
$$\frac{\mathrm{d}\,R(t)}{\mathrm{d}\,t} = \alpha I(t), \tag{24c}$$

we finally obtain the total number $S = \sum_{\kappa} \mu_{S\Lambda_1(\kappa)}$ of susceptibles, $I = \sum_{\kappa} \mu_{I\Lambda_1(\kappa)}$ of infectious and $R = \sum_{\kappa} \mu_{R\Lambda_1(\kappa)}$ of removed nodes at any given time t. A direct application of (17)–(18) provides (24a), conservation of the nodes provides (24b) and using the definitions of $I(t)$ and $R(t)$ in (13c) provides (24c). Although obtained differently, this solution corresponds to that of the pair-based SIR model presented in Volz (2008) and Miller (2010).

References

Allard A, Noël PA, Dubé LJ, Pourbohloul B (2009) Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics. Phys Rev E 79(036):113. doi:10.1103/PhysRevE.79.036113

Allard A, Hébert-Dufresne L, Noël PA, Marceau V, Dubé LJ (2012) Bond percolation on a class of correlated and clustered random graphs. J Phys A 45(405):005. doi:10.1088/1751-8113/45/40/405005

Auchincloss AH, Diez Roux AV (2008) A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. Am J Epidemiol 168:1–8. doi:10.1093/aje/kwn118

Ball F, Neal P (2008) Network epidemic models with two levels of mixing. Math Biosci 212:69–87

Bansal S, Grenfell BT, Meyers LA (2007) When individual behaviour matters: homogeneous and network models in epidemiology. J R Soc Interface 4:879–891. doi:10.1098/rsif.2007.1100

Barrat A, Barthélemy M, Vespignani A (2008) Dynamical processes on complex networks. Cambridge University Press, New York

Bascompte J, Stouffer DB (2009) The assembly and disassembly of ecological networks. Philos Trans R Soc Lond B 364:1781–1787. doi:10.1098/rstb.2008.0226

Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: structure and dynamics. Phys Rep 424:175–308. doi:10.1016/j.physrep.2005.10.009

Broeck W, Gioannini C, Goncalves B, Quaggiotto M, Colizza V, Vespignani A (2011) The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. BMC Infect Dis 11(1):37. doi:10.1186/1471-2334-11-37

Dangerfield CE, Ross JV, Keeling MJ (2009) Integrating stochasticity and network structure in an epidemic model. J R Soc Interface 6:761–774. doi:10.1098/rsif.2008.0410

Danon L, Ford A, House T, Jewell CP, Keeling MJ, Roberts GO, Ross JV, Vernon MC (2011) Networks and the epidemiology of infectious disease. Interdiscip Perspect Infect Dis 284:909:1-28. doi:10.1155/2011/284909

Decreusefond L, Dhersin JS, Moyal P, Tran VC (2012) Large graph limit for an sir process in random network with heterogeneous connectivity. Ann Appl Probab 22:541–575

Dunne JA, Williams RJ (2009) Cascading extinctions and community collapse in model food webs. Philos Trans R Soc Lond B 364:1711–1723. doi:10.1098/rstb.2008.0219

Durrett R (2007) Random graph dynamics

Eames KTD, Keeling MJ (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. PNAS 99:13,330–13,335. doi:10.1073/pnas.202244299

Gardiner CW (2004) Handbook of stochastic methods for physics. Chemistry and the natural sciences. Springer, Berlin

Gleeson JP (2011) High-accuracy approximation of binary-state dynamics on networks. Phys Rev Lett 107(068):701. doi:10.1103/PhysRevLett.107.068701

Hébert-Dufresne L, Noël PA, Marceau V, Allard A, Dubé LJ (2010) Propagation dynamics on networks featuring complex topologies. Phys Rev E 82(3):036,115. doi:10.1103/PhysRevE.82.036115

House T, Keeling MJ (2011) Insights from unifying modern approximations to infections on networks. J R Soc Int 8(54):67–73. doi:10.1098/rsif2010.0179

House T, Davies G, Danon L, Keeling MJ (2009) A motif-based approach to network epidemics. Bull Math Biol 71:1693–1706. doi:10.1007/s11538-009-9420-z

Karrer B, Newman MEJ (2010) Random graphs containing arbitrary distributions of subgraphs. Phys Rev E 82(6):066118. doi:10.1103/PhysRevE.82.066118

Keeling MJ, Eames KTD (2005) Networks and epidemic models. J R Soc Interface 2(4):295–307. doi:10.1098/rsif2005.0051

Keeling MJ, Rand DA, Morris AJ (1997) Correlation models for childhood epidemics. Proc R Soc B 264(1385):1149–1156. doi:10.1098/rspb.1997.0159

Marceau V, Noël PA, Hébert-Dufresne L, Allard A, Dubé LJ (2010) Adaptive networks: coevolution of disease and topology. Phys Rev E 82(3):036116. doi:10.1103/PhysRevE.82.036116

Marceau V, Noël PA, Hébert-Dufresne L, Allard A, Dubé LJ (2011) Modeling the dynamical interaction between epidemics on overlay networks. Phys Rev E 84(2):026105. doi:10.1103/PhysRevE.84.026105

McLane AJ, Semeniuk C, McDermid GJ, Marceau DJ (2011) The role of agent-based models in wildlife ecology and management. Ecol Model 222:1544–1556. doi:10.1016/j.ecolmodel.2011.01.020

Miller JC (2010) A note on a paper by Erik Volz: SIR dynamics in random networks. J Math Biol 62(3):349–358. doi:10.1007/s00285-010-0337-9

Miller JC, Slim AC, Volz EM (2011) Edge-based compartmental modeling for infectious disease spread. J R Soc Interface. doi:10.1098/rsif.2011.0403

Newman MEJ (2010) Networks: an introduction. Oxford University Press, Oxford

Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. Phys Rev E 64(026):118. doi:10.1103/PhysRevE.64.026118

Noël PA, Allard A, Hébert-Dufresne L, Marceau V, Dubé LJ (2012) Propagation on networks: an exact alternative perspective. Phys Rev E 85(031):118. doi:10.1103/PhysRevE.85.031118

Park J, Newman MEJ (2004) Statistical mechanics of networks. Phys Rev E 70(1—-13):066117

Rezende EL, Lavabre JE, Guimarães PR, Jordano P, Bascompte JB (2007) Non-random coextinctions in phylogenetically structured mutualistic networks. Nature 448:925–928. doi:10.1038/nature05956

Rogers T (2011) Maximum-entropy moment-closure for stochastic systems on networks. J Stat Mech (05):P05007

Sharkey KJ (2011) Deterministic epidemic models on contact networks: correlations and unbiological terms. Theor Popul Biol 79(4):115–129. doi:10.1016/j.tpb.2011.01.004

Taylor M, Simon PL, Green DM, House T, Kiss IZ (2012) From Markovian to pairwise epidemic models and the performance of moment closure approximations. J Math Biol 64(6):1021–1042. doi:10.1007/s00285-011-0443-3

Van Kampen NG (2007) Stochastic processes in physics and chemistry, 3rd edn

Volz E (2008) SIR dynamics in random networks with heterogeneous connectivity. J Math Biol 56(3):293–310. doi:10.1007/s00285-007-0116-4