From Markovian to pairwise epidemic models and the performance of moment closure approximations

Michael Taylor [1*], Péter L. Simon [2], Darren M. Green [3], Thomas House [4] & Istvan Z. Kiss [1]

[1] School of Mathematical and Physical Sciences, Department of Mathematics, University of Sussex, Falmer, Brighton BN1 9QH, UK
[2] Institute of Mathematics, Eötvös Loránd University Budapest, Budapest, Hungary
[3] Institute of Aquaculture, University of Stirling, Stirling FK9 4LA, UK
[4] Department of Biological Sciences, Mathematics Institute, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK.

∗ corresponding author
email: mt264@sussex.ac.uk

Abstract Many if not all models of disease transmission on networks can be linked to the exact state-based Markovian formulation. However the large number of equations for any system of realistic size limits their applicability to small populations. As a result, most modelling work relies on simulation and pairwise models. In this paper, for a simple SIS dynamics on an arbitrary network, we formalise the link between a well known pairwise model and the exact Markovian formulation. This involves the rigorous derivation of the exact ODE model at the level of pairs in terms of the expected number of pairs and triples. The exact system is then closed using two different closures, one well established and one that has been recently proposed. A new interpretation of both closures is presented, which explains several of their previously observed properties. The closed dynamical systems are solved numerically and the results are compared to output from individual-based stochastic simulations. This is done for a range of networks with the same average degree and clustering coefficient but generated using different algorithms. It is shown that the ability of the pairwise system to accurately model an epidemic is fundamentally dependent on the underlying large-scale network structure. We show that the existing pairwise models are a good fit for certain types of network but have to be used with caution as higher-order network structures may compromise their effectiveness.
Keywords: network, epidemic, Markov chain, moment closure.

1 Introduction

The spread of diseases within a population depends not only on the nature of the pathogen but also on the way in which infectious individuals come into contact with susceptible individuals. The network of these contacts provides the supporting structure on which the disease transmission process takes place. There is a large body of research examining network epidemic models with the aim of understanding how network properties impact on disease invasion, spread and control [14]. Many different modelling approaches have been proposed, which fall into three broad classes: exact Markovian or state-based models [26], individual-based stochastic simulation or micro models [14] and deterministic ODE-based macro models [15, 22, 25, 27]. This classification is not application specific and it simply refers to the scale (e.g. individual level or population level) at which the modelling is being carried out. The links between state-based, micro and macro models are explored in detail by Gustafsson & Sternard [10].

State-based systems, given by the master equation, or Kolmogorov equation, contain information about all possible states of the system along with the associated rates of transition from one state to another. Solving the resulting set of differential equations provides a full system description with no need for simulation. This approach has typically been used for small networks [19] due to the number of equations increasing exponentially with system size (e.g. SIS type dynamics on a network with N individuals results in $2^N - 1$ equations). With significant increases in computing power, this approach provides a realistic alternative to individual-based stochastic simulation of small populations, although we are unable yet to solve a full state-based set of ODEs for realistic network sizes. For special classes of graphs however, using the lumping technique discussed by Simon et al. [26], large reductions in the system size can be achieved and the state-based models become a viable alternative even for large networks. However, for problems involving large networks with complex structure, individual-based simulation remains the most realistic approach.

The advantages offered by individual-based modelling come at the cost of little or no analytical tractability. To overcome this problem, 'moment-closure' type ODE-based models have been developed and formulated, offering faster computational time and more analytical tractability. These differ from classic compartmenl-based ODE models in that the evolution equations for the expected number of individuals involves the expected number of pairs and higher-order structures. Many such models have been derived heuristically [13, 15, 22, 25, 27] but recently their direct link to Markovian models has also been highlighted [26]. These systems are of a tractable size, however they do not form a closed set of equations since lower moments depend on higher moments. Much work has been done on deriving moment closure approximations where the expected number of triples is approximated by a combination of the expected number of pairs and individuals [13, 15, 22, 25, 27]. This closes the system and leads to a set of numerically tractable equations. Key to this type of modelling is the ability of the approximations to capture the properties of the network and as such it has been necessary to develop rigorous approaches to test their validity [13].

In this paper, we build on the work of Simon et al. [26] who rigorously derived the exact system of ODEs (i.e. Kolmogorov forward equatons) corresponding to SIS type dynamics on an arbitrary graph. Here we start from the same exact formulation and derive the system of ODEs that describe the dynamics at the level of pairs. We then provide a new justification for two pairwise closures, one well established [15, 22] and one that has been recently proposed [12], to approximate this exact system and we test the qualitative performance of these closures for networks with identical average degree and clustering but generated using different algorithms. Results from individual-based simulation are compared to those from the closed pairwise equations and finally, potential extensions to pairwise and closure models are discussed.

2 Exact formulation of the disease transmission model

The work carried out in this paper is based upon a model describing SIS type dynamics on an arbitrary network with N nodes. Each node is either infected or susceptible at any one time, and infection and recovery are modelled by two independent Poisson processes. This follows the exact same formulation used in our previous work (Simon et al. [26]) and reference should be made to that paper for a rigorous definition of the model, however, for ease of reference, Table 1 summarises all of the notations used and defines the modelling framework used in this paper.

Table 1 Notation for model formulation and parameters

N	Number of nodes in the network.
$G = (g_{ij})_{i,j=1,2,...,N} \in \{0,1\}^{N^2}$	Adjacency matrix with $g_{ij} = 1$ if node i and j are connected, $g_{ij} = 0$ otherwise. The network is bi-directional and has no self loops such that $G = G^T$ and $G_{ii} = 0$.
τ	Rate of infection per (S, I) edge.
γ	Rate of recovery.
$\mathcal{S} = \{S, I\}^N$	State space of the network, with nodes either susceptible, S, or infected, I and $\lvert S\rvert = 2^N$.
$\mathcal{S}^k = \{\mathcal{S}_1^k, \mathcal{S}_2^k, \dots \mathcal{S}_{c_k}^k\}$	The $c_k = \binom{N}{k}$ states with k infected individuals in all possible configurations, with $k = 0, 1, ..., N$.
$X_j^k(t)$	Probability of being in state $\mathcal{S}_j^k$ at time t, where $k = 0, 1, ...N$ and $j = 1, 2, ..., c_k$.
$A_{i,j}^k$	Rate of transition from $\mathcal{S}_j^{k-1}$ to $\mathcal{S}_i^k$, where $k = 0, 1, ...N$ and $i, j = 1, 2, ..., c_k$. Note that only one individual is changing (i.e. in this case an S node changes to an I through infection).
$C_{i,j}^k$	Rate of transition from $\mathcal{S}_j^{k+1}$ to $\mathcal{S}_i^k$, where $k = 0, 1, ...N$ and $i, j = 1, 2, ..., c_k$. Note that only one individual is changing (i.e. in this case an I node changes to an S through recovery).
$B_{i,j}^k$	Rate of transition out of $\mathcal{S}_j^k$, where $B_{i,j}^k = 0$ if $i \neq j$, with $k = 0, 1, ...N$ and $i, j = 1, 2, ..., c_k$.
$N_{AB}(\mathcal{S}_j^k)$	Number of (A, B) type edges in state $\mathcal{S}_j^k$, where $A, B \in \{S, I\}$, with $k = 0, 1, ...N$ and $j = 1, 2, ..., c_k$.

Using the notation outlined in Table 1, the evolution of the epidemic in the state space, $\mathcal{S} = \{S, I\}^N$, can be described by a continuous time Markov-process and the transitions between different states can be described by the Kolmogorov equations

$$\dot{X} = PX,$$

with

$$P = \begin{pmatrix} B^0 & C^0 & 0 & 0 & 0 & 0 \\ A^1 & B^1 & C^1 & 0 & 0 & 0 \\ 0 & A^2 & B^2 & C^2 & 0 & 0 \\ 0 & 0 & A^3 & B^3 & C^3 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & A^N & B^N \end{pmatrix}, \tag{1}$$

where the sub-matrices A^k, C^k and B^k capture all transitions into $\mathcal{S}^k$ via infection, into $\mathcal{S}^k$ via recovery and out of $\mathcal{S}^k$ via both infection and recovery, respectively. By utilising the block tri-diagonal form of P, the Kolmogorov equations can be written as

$$\dot{X}^k = A^k X^{k-1} + B^k X^k + C^k X^{k+1}, \qquad k = 0, 1, \ldots, N, \tag{2}$$

where A^0 and C^N are zero matrices. A full definition of the roles and derivation of the A^k, B^k and C^k matrices can be found in Simon et al. [26], with a brief description given in Table 1. When manipulating the Kolmogorov equations in later sections, the properties of the tri-diagonal matrix P are exploited and three features in particular are often used later on in this paper. These translate into three identities as follows:

$$\sum_{i=1}^{c_k} A^k_{i,j} = \tau N_{SI}(\mathcal{S}^{k-1}_j), \tag{3}$$

$$\sum_{i=1}^{c_k} C^k_{i,j} = \gamma(k+1), \tag{4}$$

$$B^k_{i,i} = -\sum_{j=1}^{c_{k+1}} A^{k+1}_{j,i} - \sum_{j=1}^{c_{k-1}} C^{k-1}_{j,i}. \tag{5}$$

The first result, Eq. 3, implies that the sum of the j^{th} column of sub-matrix A^k is proportional to $N_{SI}(\mathcal{S}^{k-1}_j)$, with τ being the constant of proportionality. This can be intuitively understood by considering that $A^k_{i,j}$ is the rate of transition from state $\mathcal{S}^{k-1}_j$ to state $\mathcal{S}^k_i$, which differ only at one node. The particular node that changes from S to I does it at a rate equal to τ multiplied by the number of I nodes to which it is connected, in other words, the number of (S, I) edges with this particular S node as part of the pair. Thus, the sum of the j^{th} column of A^k accounts for every possibly transition that could change $\mathcal{S}^{k-1}_j$ to a particular member of the set of states $\mathcal{S}^k$. By doing so, we are considering the rate at which every S node in state $\mathcal{S}^{k-1}_j$ can become an I. Hence every (S, I) link in $\mathcal{S}^{k-1}_j$ is accounted for, leading to the result stated in Eq. 3. The second result, Eq. 4, is easily understood by considering that there are $k+1$ infected individuals in state $\mathcal{S}^{k+1}_j$, and each recovers independently at a rate γ, leading to Eq. 4. The third result, Eq. 5, ensures that the columns of P sum to zero. This is necessary, as any transitions into a state must be balanced by transitions out of a state. Since we are dealing with a closed system, every transition out of a particular state has to have a destination state, hence Eq. 5 follows. These three properties of P give us a way to not only link the A, B and C matrices together, as in Eq. 5, but also allow us to link the matrices with some countable properties of the system, as in Eqs. 3 & 4.

The Kolmogorov equations, a simple system of linear ODEs, can be conveniently programmed into a code that will automatically generate the transition matrix, P, and provide numerical solutions for networks of reasonable size. However, the system given by Eq. (2) consists of 2^N linear differential equations that are impractical to solve, or cannot be solved, for large N. It is not always necessary, however, to determine all probability functions, and in many

situations, the expected values of the number, or proportion, of susceptible (S) and infectious (I) nodes or individuals is equally valuable. These expected values at time t are denoted by $[S](t)$ and $[I](t)$ and can be expressed as follows,

$$[I](t) = \sum_{k=0}^{N} k \sum_{j=1}^{c_k} X_j^k(t), \qquad [S](t) = \sum_{k=0}^{N} (N-k) \sum_{j=1}^{c_k} X_j^k(t). \tag{6}$$

It is now necessary to derive evolution equations for $[S]$ and $[I]$. This can be done in an exact way if Eq. (2) is used and the process is illustrated in the next sub-section. As expected, the evolution equations for the expected values at the individual level involves knowledge of the expected values at the level of pairs and above. This will require the derivation of equations for the new variables and doing this in an exact way is more challenging.

2.1 Exact equations at the level of individuals

The heuristic link between expected values at individual and pair level is well known, as is the link between the expected values at pair and triple level. Until recently, these links had not been formally shown. However Simon et al. [26], have formalized this link and used the exact model to prove the following Lemma.

Lemma 1 *The expected values $[S]$ and $[I]$ satisfy the following system*

$$[\dot{S}] = \gamma[I] - \tau[SI], \tag{7}$$
$$[\dot{I}] = \tau[SI] - \gamma[I]. \tag{8}$$

This Lemma demonstrate rigorously the direct link between the Markov-process described by Eq. (2) and the often heuristically justified system described by Eqs. (7-8). As a first approximation, this system can be closed at the level of pairs by using the simplest of closing relations $[SI] \approx [S][I]$ which is based on the statistical independence of the state of individuals. By using this relation the well known mean-field model is obtained which, for a fully connected graph, becomes exact in the limit of large N [26].

3 Exact equations at the level of pairs

By using the closure at the level of pairs, the details of the underlying network structure are lost. As seen in Eqs. (7-8) the number of individuals depends on the number of pairs (for example an S node can only become an I node if it is part of an (S, I) pair). Similarly the number of pairs in the system depends upon the arrangement of different triples (for example an (S, S) pair changes to an (S, I) pair due to infection coming from a third (infected) node acting from outside the pair, an (S, S, I) triple). The exact relation between pairs and triples can be formulated in the following Theorem:

Theorem 1 *The expected values of $[S]$, $[I]$, $[SI]$, $[II]$ and $[SS]$ satisfy the following system of differential equations*

$$[\dot{S}] = \gamma[I] - \tau[SI], \tag{9}$$
$$[\dot{I}] = \tau[SI] - \gamma[I], \tag{10}$$
$$[\dot{SI}] = \gamma([II] - [SI]) + \tau([SSI] - [ISI] - [SI]), \tag{11}$$
$$[\dot{II}] = -2\gamma[II] + 2\tau([ISI] + [SI]), \tag{12}$$
$$[\dot{SS}] = 2\gamma[SI] - 2\tau[SSI]. \tag{13}$$

This is a result that is known and has been previously derived based on heuristic arguments [15, 22]. Simon et al. [26] postulated that the statement above can be rigorously proved using arguments similar to those used in the proof of Lemma 1, but did not give a proof. Here, Theorem 1 is proved by the direct use of Kolmogorov equations (Eq. 2).

Proof of Theorem 1:

The equations at the individual levels are exact and this has been shown in [26]. This first part of the proof focuses on the derivation of Eq. (12), where $[II]$ is the expected number of (I, I) pairs at a given time and is given by

$$[II] = \sum_{k=0}^{N} N_{II}(\mathcal{S}^k)X^k, \tag{14}$$

where $N_{AB}(\mathcal{S}^k)$ is a row vector of length c_k and denotes the number of (A, B) pairs in all possible configurations with k infected individuals, i.e.

$$N_{AB}(\mathcal{S}^k) = \left(N_{AB}(\mathcal{S}_1^k), N_{AB}(\mathcal{S}_2^k), \ldots, N_{AB}(\mathcal{S}_{c_k}^k)\right).$$

Similarly, $N_{ABC}(\mathcal{S}^k)$ refers to the number of (A, B, C) triples. By differentiating Eq. (14) we obtain that

$$\begin{aligned}
[\dot{II}] &= \sum_{k=0}^{N} N_{II}(\mathcal{S}^k)\dot{X^k} = \sum_{k=0}^{N} N_{II}(\mathcal{S}^k)(A^kX^{k-1} + B^kX^k + C^kX^{k+1}) \\
&= \sum_{k=1}^{N} N_{II}(\mathcal{S}^k)A^kX^{k-1} + \sum_{k=0}^{N} N_{II}(\mathcal{S}^k)B^kX^k + \sum_{k=0}^{N-1} N_{II}(\mathcal{S}^k)C^kX^{k+1} \\
&= \sum_{k=0}^{N} \left(N_{II}(\mathcal{S}^{k+1})A^{k+1} + N_{II}(\mathcal{S}^k)B^k + N_{II}(\mathcal{S}^{k-1})C^{k-1}\right)X^k,
\end{aligned}$$

where matrices that are out of range (i.e. A^0 and C^N) are zero matrices. The term that involves B^k in the summation above can be written as

$$N_{II}(\mathcal{S}^k)B^k = (N_{II}(\mathcal{S}_1^k)B_{1,1}^k, ..., N_{II}(\mathcal{S}_{c_k}^k)B_{c_k,c_k}^k),$$

where $(N_{II}(\mathcal{S}^k)B^k)_j = N_{II}(\mathcal{S}_j^k)B_{j,j}^k$ is the j^{th} component of the $N_{II}(\mathcal{S}^k)B^k$ vector. Matrices B^k are square and diagonal and are defined in terms of A^{k+1} and C^{k-1} as given in Eq. (5). This allows us to write the j^{th} component as

$$N_{II}(\mathcal{S}_j^k)B_{j,j}^k = N_{II}(\mathcal{S}_j^k)\left(-\sum_{i=1}^{c_{k+1}} A_{i,j}^{k+1} - \sum_{i=1}^{c_{k-1}} C_{i,j}^{k-1}\right).$$

Using the definition of matrices A and C as given in Eqs. (3 & 4), the RHS of the expression above can be written as

$$N_{II}(\mathcal{S}_j^k)\left(-\tau N_{SI}(\mathcal{S}_j^k) - k\gamma\right) = -\tau\left(N_{II}(\mathcal{S}_j^k)N_{SI}(\mathcal{S}_j^k)\right) - k\gamma N_{II}(\mathcal{S}_j^k).$$

This per-component identity can be written in vector form to give

$$[\dot{II}] = \sum_{k=0}^{N}\left(N_{II}(\mathcal{S}^{k+1})A^{k+1} - \tau\left(N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k)\right) - \gamma k N_{II}(\mathcal{S}^k) + N_{II}(\mathcal{S}^{k-1})C^{k-1}\right)X^k \tag{15}$$

where the $*$ operator stands for the per component multiplication of two vectors as exemplified below

$$N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k) = \left(N_{II}(\mathcal{S}_1^k)N_{SI}(\mathcal{S}_1^k), ..., N_{II}(\mathcal{S}_{c_k}^k)N_{SI}(\mathcal{S}_{c_k}^k)\right).$$

In Eq. (15) the first two terms govern infection and the second two terms govern recovery. These terms can be equated to the appropriate terms in Eq. (12). Indeed if Eq. (12) is rewritten as

$$[\dot{II}] = -2\gamma \sum_{k=0}^{N} N_{II}(\mathcal{S}^k)X^k + 2\tau \sum_{k=0}^{N} N_{ISI}(\mathcal{S}^k)X^k + 2\tau \sum_{k=0}^{N} N_{SI}(\mathcal{S}^k)X^k,$$

which is just using the same notation as above, then the proof for the equation governing the dynamics of $[II]$ follows from the identitites below:

$$\sum_{k=0}^{N} \left(-2\gamma N_{II}(\mathcal{S}^k) + \gamma k N_{II}(\mathcal{S}^k) - N_{II}(\mathcal{S}^{k-1})C^{k-1}\right) X^k(t) = 0,$$

$$\sum_{k=0}^{N} \left(2\tau N_{ISI}(\mathcal{S}^k) + 2\tau N_{SI}(\mathcal{S}^k) - N_{II}(\mathcal{S}^{k+1})A^{k+1} + \left(N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k)\right)\right) X^k(t) = 0.$$

These have to hold for any $t > 0$ and for all $X^k(t)$s. This is equivalent to showing that the coefficients of all terms involving $X^k(t)$ are zero. Upon removing the summation, the above identities become equalities between vectors of the same size. This means that the equality must hold for each of the c_k elements of the vectors. Thus the following Lemma has to be verified to complete the proof of the Theorem.

Lemma 2 *For any $k = 0, 1, \ldots, N$ and $j = 1, 2, \ldots, c_k$ the following identities hold,*

$$(k-2)\gamma N_{II}(\mathcal{S}_j^k) = \left(N_{II}(\mathcal{S}^{k-1})C^{k-1}\right)_j, \tag{16}$$

$$2\tau N_{ISI}(\mathcal{S}_j^k) + 2\tau N_{SI}(\mathcal{S}_j^k) = \left(N_{II}(\mathcal{S}^{k+1})A^{k+1}\right)_j - \left(N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k)\right)_j \tag{17}$$

The proof of this Lemma needs the following two auxiliary Propositions that will be stated and proved first.

Proposition 1 *For any $k = 0, 1, \ldots, N$ and $j = 1, 2, \ldots, c_k$ we have*

$$A:\ A_{j,i}^k \neq 0 \Rightarrow C_{i,j}^{k-1} \neq 0, \quad \text{for all } i = 1, 2, \ldots, c_{k-1}, \tag{18}$$

$$B:\ \sum_{i=1}^{c_{k-1}} A_{j,i}^k = \tau N_{II}(\mathcal{S}_j^k), \tag{19}$$

$$C:\ N_{ISI}(\mathcal{S}_j^k) = \frac{1}{\tau^2} \sum_{i=1}^{c_{k+1}} A_{i,j}^{k+1}(A_{i,j}^{k+1} - \tau). \tag{20}$$

Proof:

(A) In Eq. (18), $A_{j,i}^k$ is the rate of moving from the i^{th} arrangement of $\mathcal{S}^{k-1}$ to the j^{th} arrangement of $\mathcal{S}^k$. This means that $\mathcal{S}_i^{k-1}$ and $\mathcal{S}_j^k$ differ only at one position, say l, where $\mathcal{S}_j^k(l) = I$ and $\mathcal{S}_i^{k-1}(l) = S$ and hence the reverse process is also possible, that is a transition from $\mathcal{S}_j^k$ to $\mathcal{S}_i^{k-1}$ captured by $C_{i,j}^k$. Hence, $A_{j,i}^k$ and $C_{i,j}^{k-1}$ are reverse processes leading into and out of the same state through infection and recovery. Therefore, if infection is possible then also recovery can happen and Eq. (18) holds.

(B) Eq. (19) gives the total rate of entering $\mathcal{S}_j^k$ by infection. This means that for each of the k infected individuals in $\mathcal{S}_j^k$, there is a corresponding state $\mathcal{S}_i^{k-1}$ that differs only at one position, say at position l, where $\mathcal{S}_j^k(l) = I$ and $\mathcal{S}_i^{k-1}(l) = S$. The rate at which such a new infection happens is equal to the number of I neighbours of the S at position l multiplied by the individual transmission rate τ. If there are q such neighbours, giving q (S, I) pairs given that the S is at position l, then once the infection has taken place there will be q new (I, I) pairs. If this is taken into account for all new infections leading to $\mathcal{S}_j^k$, then the sum of these transition rates gives $\tau N_{II}(\mathcal{S}_j^k)$ and Eq. (19) holds.

(C) Eq. (20) gives the relationship between matrices A that govern the infection process and the number of (I, S, I) triples. If a node l is in state S and has q neighbours that are infected then there are $q(q-1)$ (I, S, I) triples centered around this S. As $A_{i,j}^{k+1}$ captures the infection rate from $\mathcal{S}_j^k$ to $\mathcal{S}_i^{k+1}$, this is equal to τ multiplied by the number of I neighbours connected to the susceptible node S that is being infected at position l. Hence, the number of (I, S, I) triples centered around this node l is given by $\frac{1}{\tau^2} A_{i,j}^{k+1}(A_{i,j}^{k+1} - \tau)$. This allows us to count $N_{ISI}(\mathcal{S}_j^k)$ by summing over all the c_{k+1} possible $\mathcal{S}_j^k$ to $\mathcal{S}_i^{k+1}$ transitions, giving Eq. (20).

□

Furthermore, based on the arguments above it is straightforward to derive a relation between the number of (I, I) pairs in a particular state and the number in a preceding or succeeding state. Namely, the following two identities hold.

Proposition 2 *For any $k = 0, 1, \ldots, N$ and $j = 1, 2, \ldots, c_k$ we have*

$$N_{II}(\mathcal{S}_i^{k+1})A_{i,j}^{k+1} = \left(N_{II}(\mathcal{S}_j^k) + \frac{2}{\tau}A_{i,j}^{k+1}\right) A_{i,j}^{k+1}, \quad \text{for all } i = 1, 2, \ldots, c_{k+1}, \tag{21}$$

$$N_{II}(\mathcal{S}_i^{k-1})C_{i,j}^{k-1} = \left(N_{II}(\mathcal{S}_j^k) - \frac{2}{\tau}A_{j,i}^{k}\right) C_{i,j}^{k-1}, \quad \text{for all } i = 1, 2, \ldots, c_{k-1}. \tag{22}$$

Proof:

We prove the first one, the proof of the second is similar. If the states $\mathcal{S}_i^{k+1}$ and $\mathcal{S}_j^k$ differ at more than one position, then $A_{i,j}^{k+1} = 0$, hence the statements trivially holds. In the case when the states $\mathcal{S}_i^{k+1}$ and $\mathcal{S}_j^k$ differ at one position, then there is a position l, such that $\mathcal{S}_i^{k+1}(l) = I$, $\mathcal{S}_j^k(l) = S$, and $\mathcal{S}_i^{k+1}(m) = \mathcal{S}_j^k(m)$ for $\forall\ m \neq l$. Moreover, we also require that there $\exists\ r \neq l$ such that $\mathcal{S}_j^k(r) = I$ and $g_{lr} = 1$ (i.e. there is an(S, I) type edge between nodes labelled l and by r) to ensure that a transition between the two states via infection is possible. In this case one can prove that

$$N_{II}(\mathcal{S}_i^{k+1}) = N_{II}(\mathcal{S}_j^k) + \frac{2}{\tau}A_{i,j}^{k+1}, \quad \text{for all } i = 1, 2, \ldots, c_{k+1}.$$

This identity can be understood by considering the transition from $\mathcal{S}_j^k \to \mathcal{S}_i^{k+1}$ (i.e. a single susceptible node becoming infected). If a single infection has occurred then the increase in the number of (II) pairs can be calculated by examining the number of Is or (SI) links centered around the newly infected node, just before becoming infected. The number of such (SI) pairs is proportional to A_{ij}^{k+1} and is given by $\frac{1}{\tau}A_{ij}^{k+1}$. Since, all the (SI) links, upon infection, become (II) links, and taking into account that (II) pairs must be counted twice, the identity follows immediately.

□

These Propositions allow us to prove Lemma 2.

Proof of Lemma 2:

The RHS of (16) can be expressed as

$$\left(N_{II}(\mathcal{S}^{k-1})C^{k-1}\right)_j = \sum_{i=1}^{c_{k-1}} N_{II}(\mathcal{S}_i^{k-1})C_{i,j}^{k-1}.$$

Using Eq. (22), this can be written as

$$\left(N_{II}(\mathcal{S}^{k-1})C^{k-1}\right)_j = \sum_{i=1}^{c_{k-1}} \left(N_{II}(\mathcal{S}_j^k) - \frac{2}{\tau}A_{j,i}^k\right) C_{i,j}^{k-1} = N_{II}(\mathcal{S}_j^k) \sum_{i=1}^{c_{k-1}} C_{i,j}^{k-1} - \frac{2}{\tau}\sum_{i=1}^{c_{k-1}} A_{j,i}^k C_{i,j}^{k-1}$$

From Eq. (18) it follows that every non-zero element of A^k multiplies a non-zero element of C^{k-1}. As every non-zero element in a C matrix is γ, each and every non-zero element in A^k is multiplied by γ. Hence using (19) and that $\sum_{i=1}^{c_{k-1}} C_{i,j}^{k-1} = k\gamma$ (see Eq. (4)), the RHS of Eq. (16) can be written as:

$$\left(N_{II}(\mathcal{S}^{k-1})C^{k-1}\right)_j = k\gamma N_{II}(\mathcal{S}_j^k) - 2\gamma N_{II}(\mathcal{S}_j^k) = \gamma(k-2)N_{II}(\mathcal{S}_j^k),$$

which completes the proof of the first part of Lemma 2.

In order to prove the identity given in the second part of Lemma 2, as given in Eq. (17), let us start from the first term in the RHS of (17). Using (21) and (3) it can be written as

$$\begin{aligned}\left(N_{II}(\mathcal{S}^{k+1})A^{k+1}\right)_j &= \sum_{i=1}^{c_{k+1}} N_{II}(\mathcal{S}_i^{k+1})A_{i,j}^{k+1} = \sum_{i=1}^{c_{k+1}} N_{II}(\mathcal{S}_j^k)A_{i,j}^{k+1} + \frac{2}{\tau}\sum_{i=1}^{c_{k+1}}(A_{i,j}^{k+1})^2 \\ &= N_{II}(\mathcal{S}_j^k)\sum_{i=1}^{c_{k+1}} A_{i,j}^{k+1} + \frac{2}{\tau}\sum_{i=1}^{c_{k+1}}(A_{i,j}^{k+1})^2 = \tau N_{II}(\mathcal{S}_j^k)N_{SI}(\mathcal{S}_j^k) + \frac{2}{\tau}\sum_{i=1}^{c_{k+1}}(A_{i,j}^{k+1})^2 \\ &= \tau\left(N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k)\right)_j + \frac{2}{\tau}\sum_{i=1}^{c_{k+1}}(A_{i,j}^{k+1})^2.\end{aligned}$$

This can be rearranged to give the RHS of Eq. (17):

$$\left(N_{II}(\mathcal{S}^{k+1})A^{k+1}\right)_j - \left(N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k)\right)_j = \frac{2}{\tau}\sum_{i=1}^{c_{k+1}}(A_{i,j}^{k+1})^2. \tag{23}$$

However, using Eqs. (20) and (3), Eq. (23) can be rewritten as follows:

$$\begin{aligned}\left(N_{II}(\mathcal{S}^{k+1})A^{k+1}\right)_j - \left(N_{II}(\mathcal{S}^k) * N_{SI}(\mathcal{S}^k)\right)_j &= 2\tau\left(\frac{1}{\tau^2}\sum_{i=1}^{c_{k+1}}(A_{i,j}^{k+1})^2\right) \\ &= 2\tau\left(N_{ISI}(\mathcal{S}_j^k) + \frac{1}{\tau}\sum_{i=1}^{c_{k+1}} A_{i,j}^{k+1}\right) \\ &= 2\tau N_{ISI}(\mathcal{S}_j^k) + 2\sum_{i=1}^{c_{k+1}} A_{i,j}^{k+1} = 2\tau N_{ISI}(\mathcal{S}_j^k) + 2\tau N_{SI}(\mathcal{S}_j^k),\end{aligned}$$

which completes the proof of the Lemma.

□

4 Moment closures

The exact system of Equations (9 - 13) that has been derived from the Kolmogorov equation is not closed, as the dynamics of the pairs depends upon the number of triples in the system. Either equations for the dynamics of the triples can be written down, but these would be dependent on the number of four-motifs [11]. This hierarchical dependence can be broken by approximating the number of triple as a function of pairs and singles.

We assume here that every individual has the same neighbourhood size n (in the case of networks with heterogeneous degree distribution we obtain an approximation by taking n to be the real-valued mean degree, which is not rigorously interpretable) and also define a clustering coeffieient ϕ

$$n = \sum_j G_{ij}, \forall i \ , \qquad \phi = \frac{\sum_{i,j,k} G_{ij}G_{jk}G_{ki}}{\sum_{i,j,k} G_{ij}G_{jk}(1-\delta_{ik})} \in [0,1] \ . \tag{24}$$

We will use these two real parameters n, ϕ in the formulation of different closure approximations. In the absence of clustering ($\phi = 0$) both approximations we will consider agree that

$$[ABC] \approx n(n-1)p_{A|B}p_{C|B} \ , \text{ where } \quad p_{A|B} := \frac{[AB]}{n[B]} \ . \tag{25}$$

This closure implicitly assumes that the disease states of individuals around a node in state B are given by independent trials, and so neighbourhood types are multinomially distributed [3]. However, this assumption breaks down as correlations will develop between the status of neighbouring nodes, as new infections are bound to be the neighbours of their infectors. This will especially become an issue for SIS type dynamics as, for example, if the middle node of an $[III]$ triple recovers, it is likely to quickly become infected again, leading to more $[II]$ pairs and $[III]$ triples than would be expected if the infected nodes were distributed at random. This becomes of even more importance in clustered networks, where there are highly structured cliques of individuals. These create areas of the graph where infection can become entrenched. As such, for the case of clustering, understanding the implicit assumptions behind different closure methods is much harder. Here, we present a new motivation of existing moment closure techniques, similar to the classic statistical approach to clustering of Klotz [18]. This explanation makes it much easier to see which kinds of network structure are assumed during momemt closure, and how different closures are related to each other.

Our starting point is the correlation matrix between different adjacent dynamical states, which is equal to unity for homogeneous (non-network) mixing

$$\mathcal{C}_{A,B} := \frac{N}{n}\frac{[AB]}{[A][B]} \ . \tag{26}$$

Figure 1 shows the construction used in our closure derivation. We start by noting that only triples of type $[ABI]$ need to be closed, and that this can be done in terms of the neighbourhood around each $[BI]$ pair. We consider each of the other $n-1$ neighbours of a B in such a $[BI]$ pair in turn. For each such neighbour X_i, we decide with probability ϕ whether it is connected to the I in the $[BI]$ pair or not. We then pick its dynamical state, taking into account the correlations between the proposed state and I if they are connected. Making use of equation (26), the fact that X_i state probabilities have to sum to unity, and provided the identity $\sum_a [aB] = n[B]$ is conserved by the dynamics (as it was shown to in [12]), our explicit consistent local assumption is that

$$\Pr(X_i = A) = \begin{cases} p_{A|B} \text{ with probability } (1-\phi) \ , \\ p_{A|B}\mathcal{C}_{A,I} / \left(\sum_a p_{a|B}\mathcal{C}_{a,I}\right) \text{ with probability } \phi \ . \end{cases} \tag{27}$$

Averaging over all neighbourhoods around $[BI]$ pairs gives the expected number of relevant triples in the network as

$$\begin{aligned} [ABI] &\approx (n-1)[BI]\left((1-\phi)p_{A|B} + \phi\frac{p_{A|B}\mathcal{C}_{A,I}}{\sum_a p_{a|B}\mathcal{C}_{a,I}}\right) \\ &\approx (n-1)\left((1-\phi)\frac{1}{n}\frac{[BI][BA]}{[B]} + \phi\frac{[BA][IA]/[A]}{\sum_a ([Ba][Ia]/[a])}\right) \ . \end{aligned} \tag{28}$$

This is the improved closure of [12], which was originally motivated by its satisfaction of two practical desiderata: the conservation of pair number and consistent behaviour of $[III]$-type triples. These properties can now be seen to flow from the consistent probabilistic neighbourhood-based description of the improved closure above.

The standard clustered pairwise closure of [15, 22] can be recovered by making the assumption $\left(\sum_a p_{a|B}\mathcal{C}_{a,I}\right) \approx 1$ so that

$$[ABI] \approx \frac{n-1}{n}\frac{[AB][BI]}{[B]}\left((1-\phi)+\phi\frac{N}{n}\frac{[AI]}{[A][I]}\right) \ . \tag{29}$$

This closure has the benefit of significantly reducing the complexity (and hence numerical effort in integration) of the closed set of equations, but since the approximation used is logically inconsistent there is the possibility of serious pathologies creeping into numerical results. This possibility partly motivates our comparison below of both closures against a more comprehensive set of clustered networks than has been previously considered.

In both the standard and improved closures, the now-explicit assumption made about clustering is that each transitive link exists with independent probability ϕ, and so we would expect networks where transitive links are themselves clustered together into cliques as in [1] or unclustered as in [28] where no triangles overlap not to give good dynamical agreement with the proposed closures.

5 Simulation versus different moment closures approximations

The pairwise models aim to capture the local or small scale network properties such as n, the average node degree, and ϕ, the clustering coefficient by using different moment closures. However, as shown by Green & Kiss [8], it is possible to produce networks that differ vastly in large scale properties even though the local properties are identical. In this section, we explore how much of the underlying network structure is captured by n and ϕ on networks generated by different algorithms. This is done by examining how well the pairwise models given in Eqs. 29 & 28 agree with results from stochastic simulation.

The networks used here were generated following algorithms explored by Green and Kiss [8]. In each case, 100 distinct networks were produced using each algorithm and simulations were averaged over 5 repetitions on each network, giving 500 repetitions in total. For all simulations, $\gamma = 1$ and $\tau = 0.5$ which ensures that an epidemic will break out. Each network has $N = 10,000$, the timestep used in the simulations is $\delta t = 0.01$ with synchronous updating used and initially 5 nodes are chosen at random to be infected. In fact, apart from the networks themselves, the only parameters that are varied in the simulations are n and ϕ, with $n \in \{5, 10\}$ and $\phi \in \{0, 0.4\}$.

The first networks explored were generated using the spatial algorithm [23] whereby nodes are distributed uniformly at random across a $\sqrt{N} \times \sqrt{N}$ square with the probability of two nodes being connected given by some normal-like connectivity kernel depending on the distance between them. A second set of networks were generated by re-wiring the spatial networks using big-V rewiring [8, 13]. This preserves the clustering coefficient and node degree but removes other forms of structure. A third set of networks were generated using the group-based algorithm [21] which is based on a bipartite of groups and nodes where group size and group membership distributions are varied to obtain networks with different desired properties. A fourth set of networks comes from applying big-V rewiring to the group-based algorithm. A fifth set comes from an iterative algorithm proposed by Eames [5] where triple generation is followed by triangle generation to obtain the right level of clustering. A final set is generated by unclustering the networks generated by the iterative algorithm. Further details of all these algorithms can be found in [8]. The rewirings used are shown in Figure 2, where (a) shows the unclustering rewiring and (b) shows the big-V rewiring.

The different combinations of n and ϕ values leads to four versions of each network type. For each of these, the prevalence of infection over time is compared with the approximations from the two pairwise closures as well as the mean-field model. Figure 3 shows results for both the spatial networks and their reclustered counterparts, Figure 4 show results for both the group-based networks and their reclustered versions and Figure 5 shows results for the iterative networks and their unclustered versions. Both of the pair approximations agree best with simulation results when $n = 10$ and $\phi = 0$ regardless of the network and this is in line with previous findings [15, 26]. In this case, even the mean-field approximation provides a good fit even though $n << N$. The figures show that reducing n or

increasing ϕ reduce the goodness of fit between the moment closure approximations and the simulation results for all networks, although the amount they differ depends on the network generating algorithm.

The networks that show the worst fit for the moment closures are the group-based networks when n is low and ϕ is high, as seen in Figure 4(b). Here the epidemic picks up much quicker than predicted even by the mean-field approximation but the final epidemic size is much smaller. Another interesting feature of the group-based networks is the limited impact of the big-V rewiring on the underlying structure. This is a result of the algorithm producing high degree heterogeneity with many nodes remaining unconnected for higher values of ϕ, as shown by Kiss & Green [8]. Indeed for all four combinations of n and ϕ there is negligible difference between the results on the standard group-based networks and those on their reclustered counterparts. For these group-based networks ϕ has a more important role than n in determining whether or not the pairwise approximations are in good agreement with simulations.

The somewhat simpler and more straightforward iterative algorithm shown in Figure 5 shows good qualitative agreement between the simulations and ODEs, with the ODEs overpredicting the initial growth in all cases but with agreement becoming good for increasing n. Unclustering these networks returns simulation results that are similar to those seen in the cases where $\phi = 0$ and this provides a good check for the effectiveness of the unclustering algorithm.

The most interesting case is for networks generated by the spatial algorithm as shown in Figure 3. Here the ODEs overpredict final epidemic size in all cases, but not by a significant amount. For the cases with $\phi > 0$ the initial growth of the epidemic is much slower than predicted by the ODEs, especially when $n = 5$. This type of network, however, responds very well to being reclustered using big-V rewiring. The difference in agreement for the standard networks and their reclustered counterparts when $\phi > 0$ is very large, showing that there is much underlying network structure generated by the spatial algorithm that is not captured by n and ϕ. There are remarkable differences between the results from these two networks, as seen in Figures 3(b) and 3(d), and these show the importance of large scale spatial structure in determining how an epidemic can invade. The goodness of agreement is directly linked to the capability of the pairwise models to correctly describe the pair dynamics. This is captured in Figure 6 that shows the correlations, C_{SI}, and the normalised expected number of $[II]$ and $[SI]$ pairs. It can be seen that the rewired networks display much better agreement in both the growth of pairs and C_{SI} when comparing pairwise approximations to simulation results. This shows that the big-V rewiring has successfully removed higher level structure, and in this case, the pairwise closures correctly capture the evolution of correlations and pairs.

The performance of pairwise models does not depend solely on the structure of the network but also on the dynamics (e.g. SIS, SIR, $SITS$, where T stands for tracing triggering individuals) that unfold on the network. For example, House & Keeling [12], show that for contact tracing models, the OPA does not account correctly for the evolution of triangles with three infected individuals and this can lead to skewed outcomes and they went on to address this via the IPA. In Figure 7 we show that for simple SIS dynamics the IPA captures the initial growth of $[III]$ triples much better than the OPA. This difference is emphasised when the values of the pairs and singles that constitute the two different closures are taken directly from the simulation. Figure 7(a) shows this difference most dramatically, with the OPA predicting a much more rapid early growth of these $[III]$ triples, whereas the IPA describes their actual evolution with a much higher degree of accuracy. However, the failure of the OPA to correctly capture the evolution of $[III]$ triples does not translate to a significant difference in the time evolution of the prevalence (see Figure 3) but as House & Keeling [12] point out, the evolution of these triples becomes important when more complicated dynamics, e.g. $SITS$, are playing out on the networks. The evolution of all other triples is much better approximated by both the OPA and IPA and moreover, as noted before, the big-V rewiring successfully removes higher level network structure and leads to an overall better agreement between all approaches. Thus, these observations highlight the importance of considering both network structure and the particular dynamics when deriving macro-ODE models.

6 Discussion

The main result of the paper is the identification of the direct link between the Markovian formulation and the ODE system that governs the dynamics of pairs. The key to the proof was provided by the special tri-diagonal structure and properties of the transition matrices from the Kolmogorov equations. The resulting system of ODEs is well known,

having been derived heuristically before but here the link to the state based Markovian formulation is emphasized. This macro model is exact at the level of pairs but is not a closed system. We have used two different moment closures approximations and compared results from these to those from the micro modelling process of individual-based simulations. The results reveal that moment closures heavily rely on exploiting local network structure and for networks where higher order structure is present the agreement breaks down. We have shown that using a special rewiring technique, that removes higher-order or large-scale structure while keeping local properties unchanged, improves the agreement between micro and macro models significantly for certain types of network but has little effect on certain others.

Our analysis shows that large-scale network structure plays a crucial role when deriving moment closure approximations. Existing closures often rely on local network properties alone and perform poorly in accounting for larger scale features. However, much progress has been made in accounting for properties such as node degree and degree distribution heterogeneity [6, 17], preferential mixing [4, 20], clustering [7] and even directed or weighted edges [24]. While most of these models rely on some form of pairwise model with a corresponding closing relation, there is scope for better understanding and justification of approximation models as well as working towards a unifying framework for such approximation models.

This paper shows that it is worthwhile to consider alternative custom-made closures for different networks, and to do this, it may be necessary to incorporate non-local network properties especially when local network metrics such as n and ϕ fail to capture the key network features. For example, Green & Kiss [8] use network measures other than n and ϕ to investigate the correlation between simulation results and non-local network metrics. An alternative approach may rely on higher order motifs, such as quadruples [11], and accounting for these could lead to improved approximation models. One of the key challenges for generating valid and accurate moment closure approximations is to account for the dependence of the state of a node on the state of neighbouring nodes, and the correlations that arise as a result. Indeed these dependencies will act on a larger scale than simply a node's immediate neighbours, so the challenge of finding appropriate closures may be very difficult. Such approaches, at least in the initial stages of development, will rely on generating theoretical toy networks that can be used to assess the goodness of the approximations. If these novel approximations will only yield satisfactory results for specific networks, their practical benefits will be small but could form the building blocks to develop models that are valid for larger or more realistic sets of networks. It is clear that only by understanding the fundamental structures that underpin any network can moment closure techniques be confidently used to model dynamical processes on networks.

Acknowledgements
Istvan Z. Kiss acknowledge support from EPSRC (EP/H001085/1). Michael Taylor acknowledges support from EPSRC (DTA grant). Péter L. Simon acknowledges support from OTKA (grant no. 81403). Thomas House acknowledges support from EPSRC (EP/H016139/1).

References

1. Ball F, Sirl D, Trapman P (2010) Analysis of a stochastic sir epidemic on a random network incorporating household structure. Math. Biosci., 224(2):53-73.
2. Brauer F, van den Driessche P (2009) Mathematical Epidemiology. Springer, 2008.
3. Dangerfield CE, Ross JV, and Keeling MJ (2009) Integrating stochasticity and network structure into an epidemic model. J. R. Soc. Interface, 6(38):761-774.
4. Eames KTD (2006) Contact tracing strategies in heterogeneous populations. J. R. Soc. Interface, 3:55-62.
5. Eames KTD (2007) Modelling disease spread through random and regular contacts in clustered populations. Theor. Pop. Biol., 73:104-111.
6. Eames KTD, Keeling MJ (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. Proceedings of the National Academy of Sciences of the United States of America, 99(20):13330-13335.
7. Eames KTD, Keeling MJ (2003) Contact tracing and disease control. Proc Roy Soc B, 270(1533):2565-71.
8. Green DM, Kiss IZ (2010) Large scale properties of clustered networks: implications for disease dynamics. J. Biol. Dyn., iFirst:1-15.
9. Grimmet G, Stirzaker D (2001) Probability and Random Processes (Third Edition). Oxford University Press.
10. Gustafsson L, Sternad M (2010) Consistent micro, macro and state-based population modelling. Math. Biosci., 225:94-107.
11. House T, Davies G, Danon L, Keeling MJ (2009) A motif-based approach to network epidemics. Bull. Math. Biol., 71:1693-1706.
12. House T, Keeling MJ (2010) The impact of contact tracing in clustered populations. PLoS Comput. Biol., 6(3): e1000721.
13. House T, Keeling MJ (2010) Insights from unifying modern approximations to infections on networks. J. R. Soc. Interface. doi:10.1098/rsif.2010.0179.
14. Keeling MJ, Eames KTD (2005) Networks and epidemic models. J. R. Soc. Interface, 2:295-307.
15. Keeling MJ (1999) The effects of local spatial structure on epidemiological invasions. Proc. R. Soc. Lond. B, 266:859-867.
16. Kiss IZ, Green DM (2008) Comment on properties of highly clustered networks. Phys. Rev. E, 78:048101.
17. Kiss IZ, Green DM, Kao RR (2006) The effect of network heterogeneity and multiple routes of transmission on final epidemic size. Math. Biosci., 203:124-136.
18. Klotz J (1973) Statistical inference in bernoulli trials with dependence. The Annals of Statistics, 1:373-379.
19. Keeling MJ, Ross JV (2008) On methods for studying stochastic disease dynamics. J. R. Soc. Interface, 5:171-181.
20. Kiss IZ, Simon PL, Kao RR (2009) A contact-network-based formulation of a preferential mixing model. Bulletin of Mathematical Biology, 71:888-905.
21. Newman MEJ (2003) Properties of highly clustered networks. Phys. Rev. E, 68:026121.
22. Rand DA (1999) Correlation equations and pair approximations for spatial ecologies, in: McGlade J (ed) Advanced Ecological Theory. Blackwell Science, pp 100-142.
23. Read JM, Keeling MJ (2003) Disease evolution on networks: the role of contact structure. Proc. R. Soc. Lond. B, 270:699-708.
24. Sharkey KJ (2008) Deterministic epidemiological models at the individual level. J. Math. Biol., 57:311-331.
25. Sato K, Matsuda H, Sasaki A (1994) Pathogen invasion and host extinction in lattice structured populations. J. Math. Biol., 32:251-268.
26. Simon PL, Taylor M, Kiss IZ (2010) Exact epidemic models on graphs using graph automorphism driven lumping. J. Math. Biol. doi:10.1007/s00285-010-0344-x.
27. van Baalen M (2000) Pair approximations for different spatial geometries, in: Dieckmann U et al (eds) The geometry of ecological interactions. Cambridge University Press, pp 359-387
28. Volz EM (2010) Dynamics of infectious disease in clustered networks with arbitrary degree distributions. arXiv:1006.0970v1.

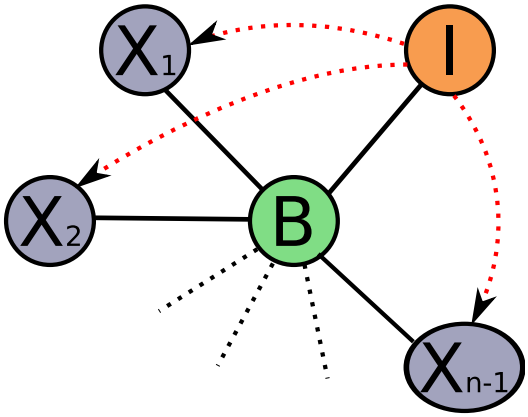
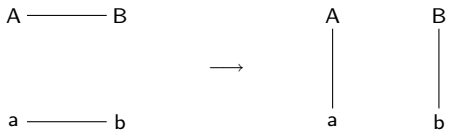
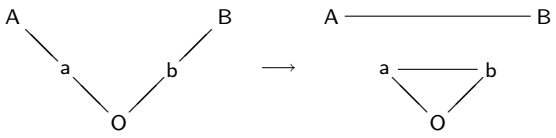


Fig. 1 Motivation for improved pairwise closure. Triples of type [ABI] are counted by consideration of the neighbourhood around a typical [BI] pair. The states $X_1, \ldots, X_{n-1}$ are chosen in order, with the presence or absence of potential transitive links (red dotted lines) decided simultaneously as outlined in Equation (27) of the main text.

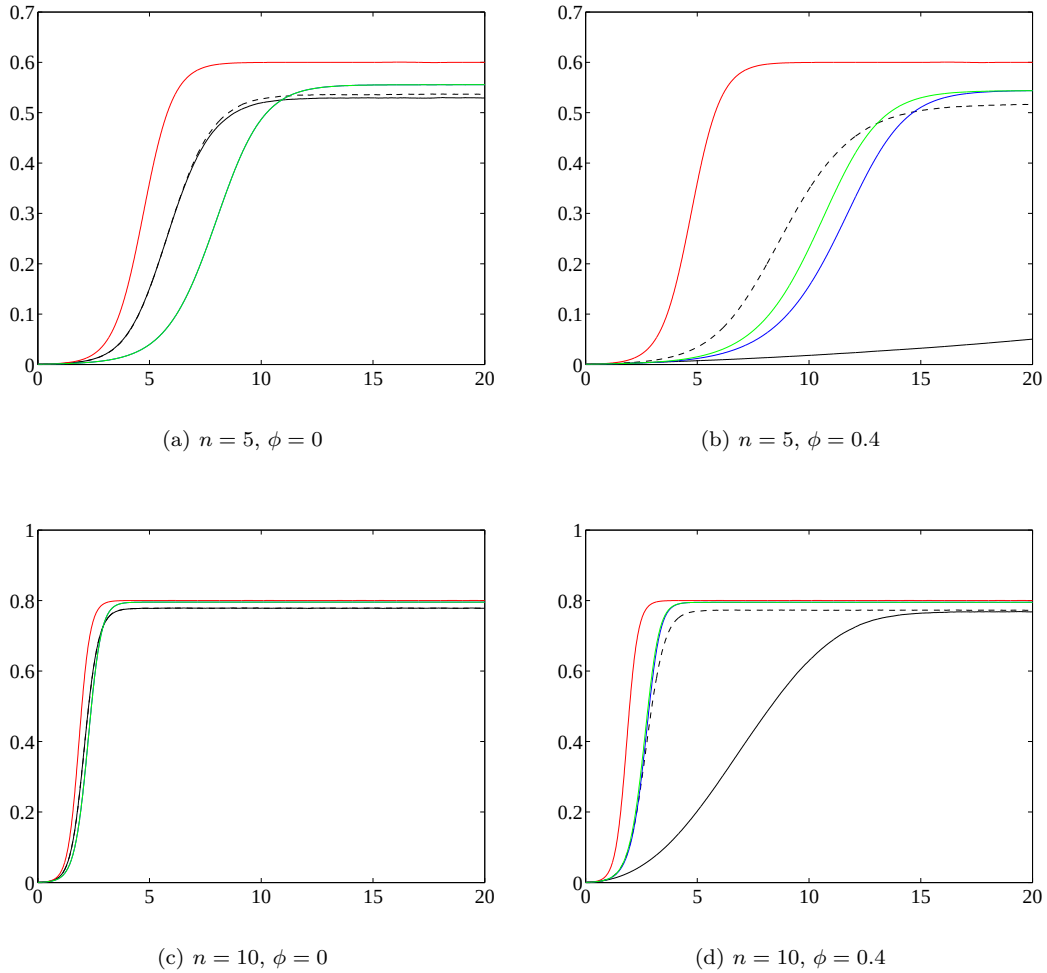


(a) Unclustering rewiring



(b) Big-V rewiring

Fig. 2 Rewiring methods. (a) shows the unclustering rewiring, which will evolve a network towards an unclustered configuration-model network of the same degree distribution. (b) shows the big-V rewiring, which generates clustering of the type expected to agree with pairwise moment closure.



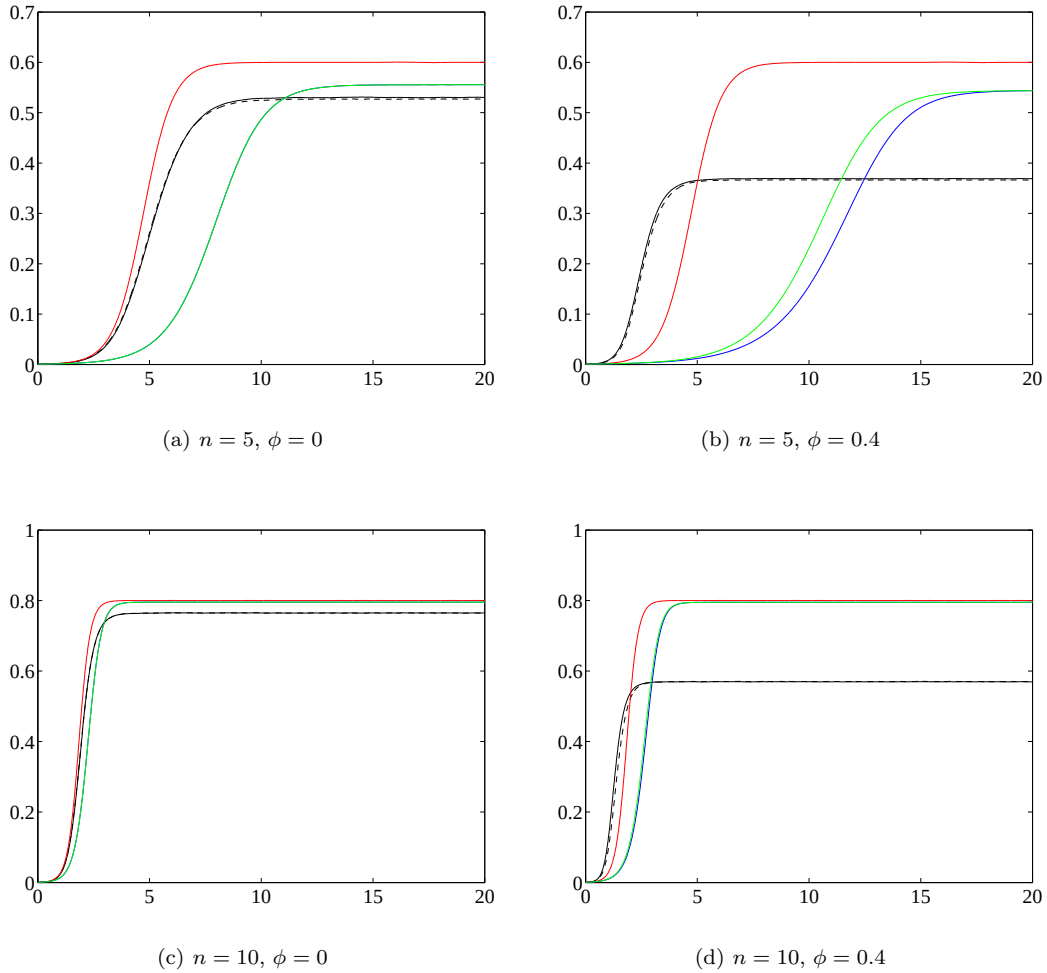
(a) $n = 5$, $\phi = 0$

(b) $n = 5$, $\phi = 0.4$

(c) $n = 10$, $\phi = 0$

(d) $n = 10$, $\phi = 0.4$

Fig. 3 Infection prevalence time series for networks generated using the spatial algorithm along with results given by moment closure equations. In all plots $N = 10,000$, $\tau = 0.5$, $\gamma = 1$ with varying n and ϕ. The solid black line is for simulation results from the spatial algorithm and the black dashed line is for simulation results from the same spatial algorithm but with the networks reclustered using big-V rewiring. The red line is the mean-field approximation, the blue line is the ordinary pairwise approximation (OPA) and the green line is for the improved pairwise approximation (IPA). In the cases where $\phi = 0$ the OPA and IPA coincide.



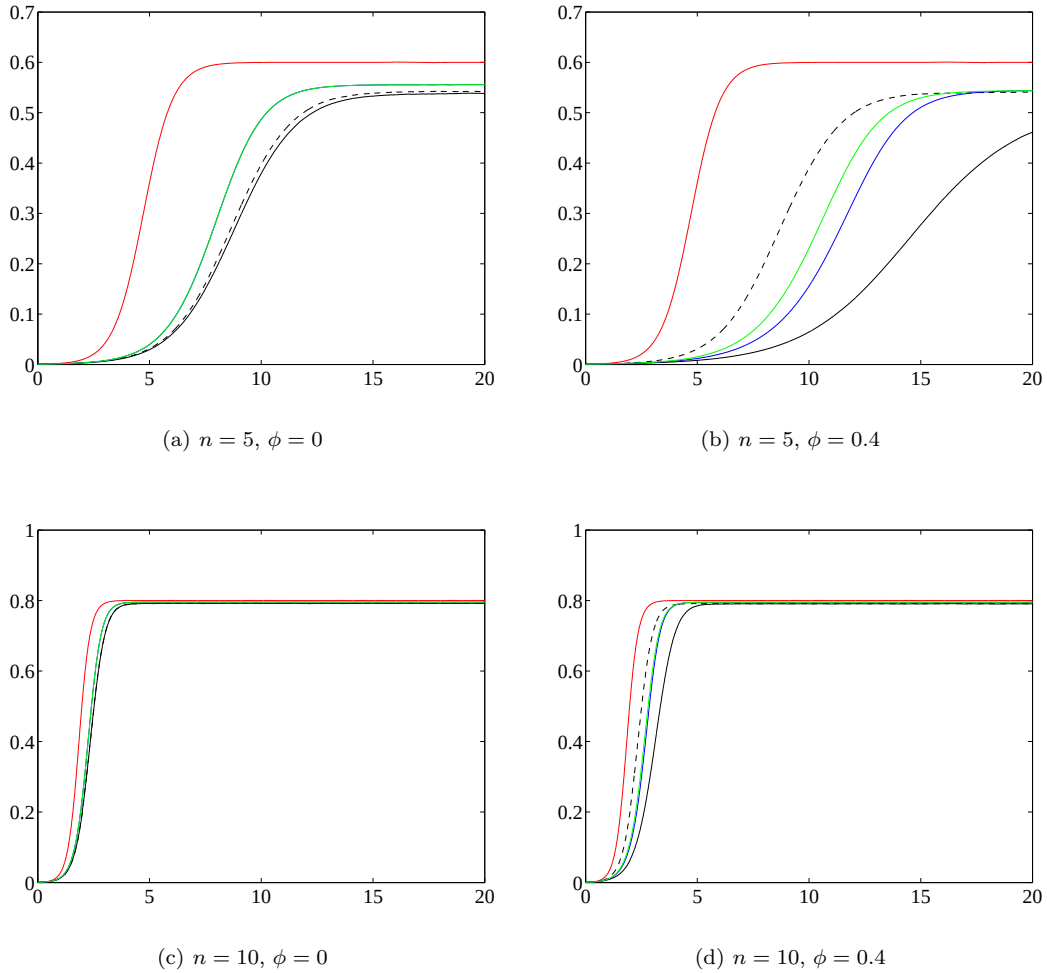
(a) $n = 5$, $\phi = 0$

(b) $n = 5$, $\phi = 0.4$

(c) $n = 10$, $\phi = 0$

(d) $n = 10$, $\phi = 0.4$

Fig. 4 Infection prevalence time series for networks generated using the group-based algorithm along with the results given by moment closure equations. In all plots $N = 10,000$, $\tau = 0.5$, $\gamma = 1$ with varying n and ϕ. The solid black line is for simulation results from the group-based algorithm and the black dashed line is for simulation results from the same group-based algorithm but with the networks reclustered using big-V rewiring. The red line is the mean-field approximation, the blue line is the ordinary pairwise approximation (OPA) and the green line is for the improved pairwise approximation (IPA). In the cases where $\phi = 0$ the OPA and IPA coincide.



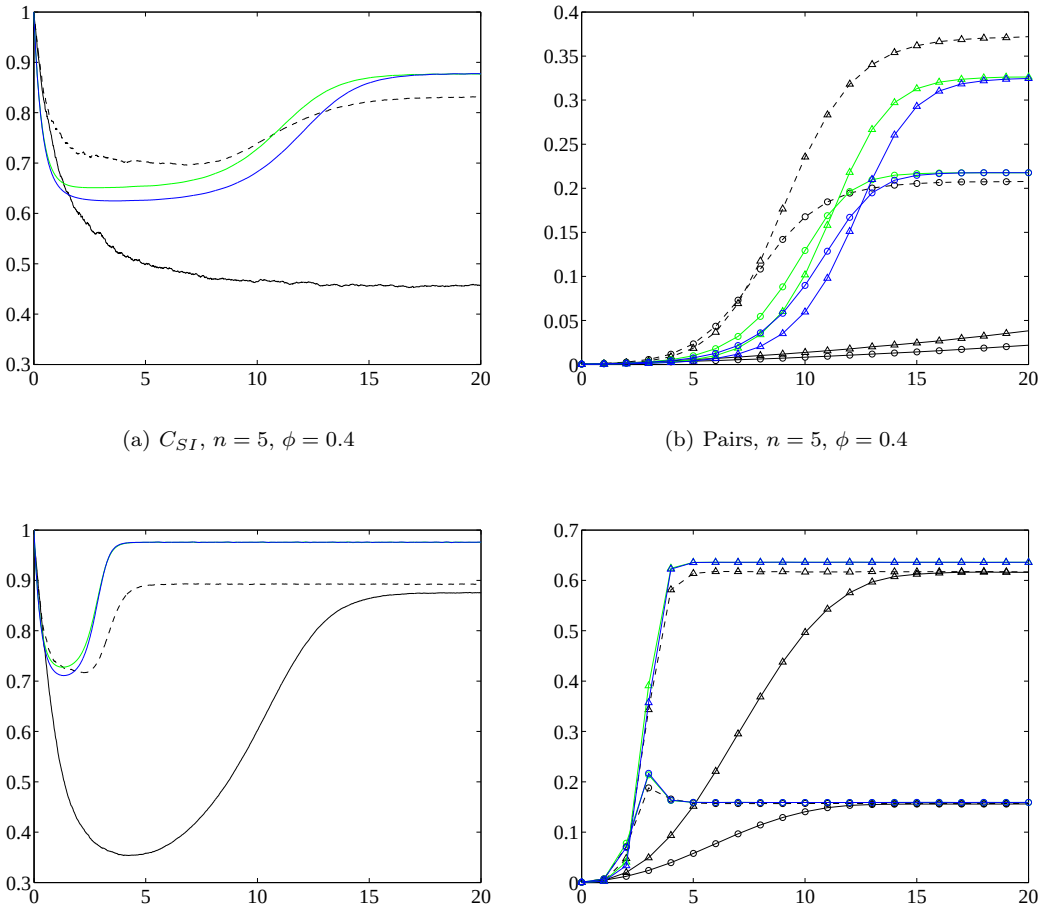
(a) $n = 5$, $\phi = 0$

(b) $n = 5$, $\phi = 0.4$

(c) $n = 10$, $\phi = 0$

(d) $n = 10$, $\phi = 0.4$

Fig. 5 Prevalence time series for networks generated using the iterative algorithm along with the results given by moment closure equations. In all plots $N = 10,000$, $\tau = 0.5$, $\gamma = 1$ with varying n and ϕ. The solid black line is for simulation results from the iterative algorithm and the black dashed line is for simulation results from the same group-based algorithm but with the networks reclustered using big-V rewiring. The red line is the mean-field approximation, the blue line is the ordinary pairwise approximation (OPA) and the green line is for the improved pairwise approximation (IPA). In the cases where $\phi = 0$ the OPA and IPA coincide.



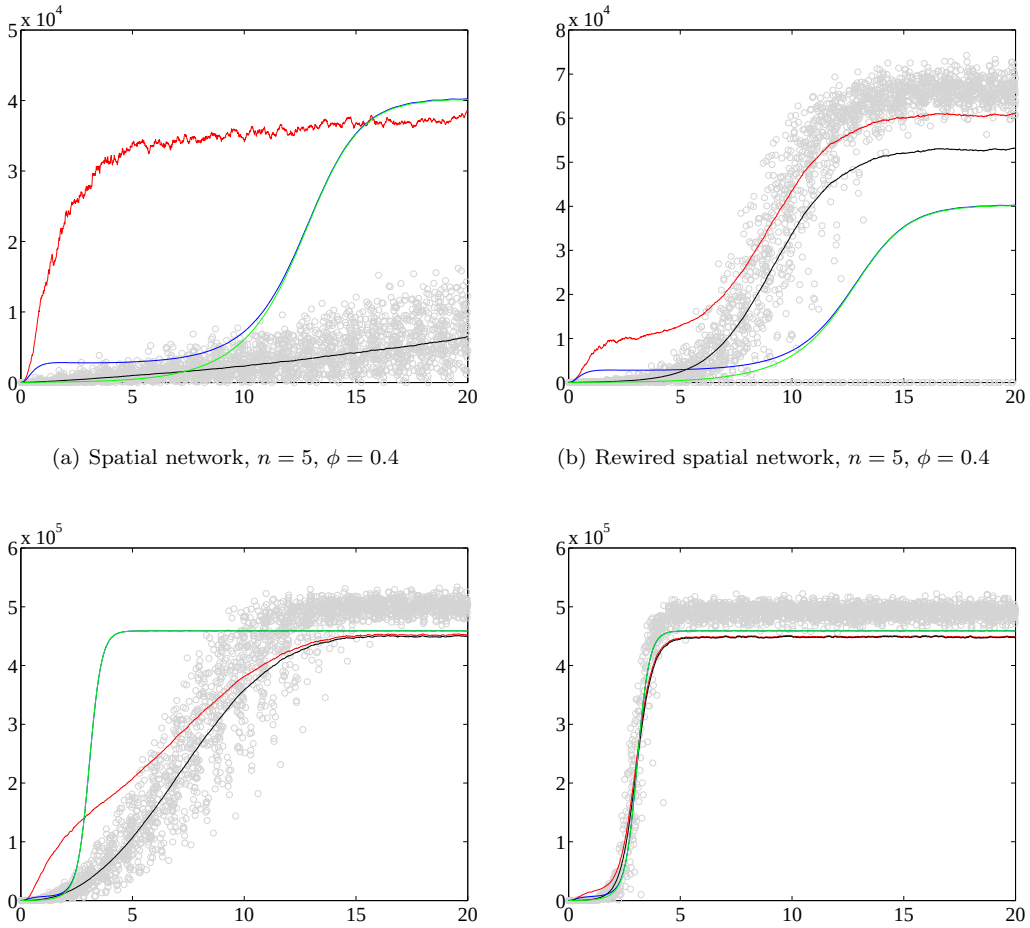
(a) C_{SI}, $n = 5$, $\phi = 0.4$

(b) Pairs, $n = 5$, $\phi = 0.4$

(c) C_{SI}, $n = 10$, $\phi = 0.4$

(d) Pairs, $n = 10$, $\phi = 0.4$

Fig. 6 Figures 6(a) & 6(c) show time series for the correlation C_{SI} as defined in Eq. (26). Figures 6(b) & 6(d) show time series for proportion of pairs that are (S, I) (circles) and (I, I) (triangles). In all plots $N = 10,000$, $\tau = 0.5$, $\gamma = 1$. The solid black line is for simulation results from the spatial algorithm and the black dashed line is for simulation results from the same spatial algorithm but with the networks reclustered using big-V rewiring. The blue line is the ordinary pairwise approximation (OPA) and the green line is for the improved pairwise approximation (IPA).



(a) Spatial network, $n = 5$, $\phi = 0.4$

(b) Rewired spatial network, $n = 5$, $\phi = 0.4$

(c) Spatial network, $n = 10$, $\phi = 0.4$

(d) Rewired spatial network, $n = 10$, $\phi = 0.4$

Fig. 7 Showing $[III]$ triples counted in different ways for $n = 5, 10$ and $\phi = 0.4$ on spatial networks and their big-V rewired counterparts. The grey cloud shows actual counts of $(I - I - I)$ triple from individual stochastic realisations. The blue and green line show [III] calculated using the OPA given by Eq. (29) and the IPA given by Eq. (28) respectively with both using pair and individual values from solving the macro ODE system (Eqs. (9 - 13)). The red and black lines show the OPA and IPA respectively, but with values for individuals and pairs as averages from the stochastic simulation.