



UNIVERSITY OF LEEDS

This is a repository copy of *Mean-field approximation of dynamics on networks*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/231865/>

Version: Accepted Version

Article:

Ward, J.A. orcid.org/0000-0002-2469-7768, Timar, G. and Simon, P.L. (Accepted: 2025)
Mean-field approximation of dynamics on networks. SIAM Journal on Applied
Mathematics. ISSN: 0036-1399 (In Press)

This is an author produced version of an article accepted for publication in SIAM Journal on Applied Mathematics, made available under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

MEAN-FIELD APPROXIMATION OF DYNAMICS ON NETWORKS*

JONATHAN A. WARD[†], GÁBOR TIMÁR[†], AND PÉTER L. SIMON[‡]

Abstract. Many real-world phenomena can be modelled as dynamical processes on networks, a prominent example being the spread of infectious diseases such as COVID-19. Mean-field approximations are a widely used tool to analyse such dynamical processes on networks, but these are typically derived using plausible probabilistic reasoning, introducing uncontrolled errors that may lead to invalid mathematical conclusions. In this paper we present a rigorous approach to derive mean-field approximations from the exact description of Markov chain dynamics on networks through a process of averaging called approximate lumping. We consider a general class of Markov chain dynamics on networks in which each vertex can adopt a finite number of “vertex-states” (e.g. susceptible, infected, recovered etc.), and transition rates depend on the number of neighbours of each type. Our approximate lumping is based on counting the number of each type of vertex-state in subsets of vertices, and this results in a density dependent population process. In the large graph limit, this reduces to a low dimensional system of ordinary differential equations, special cases of which are well known mean-field approximations. Our approach provides a general framework for the derivation of mean-field approximations of dynamics on networks that unifies previously disconnected approaches and highlights the sources of error.

Key words. Complex systems, network science, dynamical systems, Markov chains.

MSC codes. 37N99, 60J28, 91C99, 92D25, 92D30, 05C82.

1. Introduction. Dynamical processes on networks are important and widely studied [1, 26, 30, 23]. They have been used to study real-world phenomena, such as epidemics [23, 27], opinion dynamics [13, 32, 33], and spin systems with critical phenomena [9, 24, 8]. Many such models can be described mathematically as Markov chains [31, 36, 37], but often their state-space is so large it is impossible to use mathematical tools from the theory of Markov chains. Instead it is standard to make use of “mean-field” approximations [35, 28, 23, 15, 12, 25, 6, 19], in which aspects of network structure and dynamical correlations are ignored [17].

A class of mean-field approximations, usually referred to as dynamical mean-field, concerns dynamics on networks encoded in the form of sets of Langevin-type or deterministic equations, where the disorder (in either network structure or interaction strengths) is generally averaged out using path-integral methods [25, 6, 7]. Most mean-field approximations that attempt to provide a low-dimensional description of stochastic dynamics starting from the original Markov chain, however, tend to be based on plausible probabilistic reasoning, thus lacking a rigorous mathematical foundation [35, 28, 23, 15, 12]. Such approximations have the potential to introduce uncontrolled errors that limit the potential for mathematical analysis, since the mean-field approximation is not faithful to the original process. For example, controversy concerning the critical epidemic threshold in scale-free networks stemmed from the use of mean-field approximations [28, 18, 5, 2]. Moreover, the assumptions that underpin mean-field approximations—the absence of clustering, modularity/community structure, dynamical correlations—are routinely violated by dynamical processes on real-world complex networks [17]. Thus it is difficult to know when a mean-field ap-

*Submitted to the editors DATE.

Funding: J.A. Ward and G. Timár acknowledge funding from the Leverhulme Trust Project Grant number RPG-2023-187. P.L. Simon acknowledges support from the Hungarian Scientific Research Fund, OTKA (grant no. 135241) and from ERC Synergy Grant No. 810115 - DYNASNET.

[†]University of Leeds, UK (j.a.ward@leeds.ac.uk).

[‡]Eotvos Lorand University, Hungary

proximation will be accurate or how the error depends on the network structure or the dynamic [38]. Consequently, the quantification of approximation error has been recognised as a key challenge for network epidemic modellers [29].

In this paper, we develop a mathematical foundation for mean-field approximation that starts with the exact Markov chain description of a broad class of dynamical processes on networks where interactions are governed by vertices' local neighbourhoods. We use a technique called approximate lumping to derive a “density dependent population process” that in the large network limit converges to a relatively small set of differential equations. There are three significant advances on previous work [38]. Firstly, by basing the approximate lumping on a partition of vertices, we are able to incorporate network structure into the mean-field approximation in a very flexible way. Secondly, we combine two key techniques—approximate lumping and convergence of density dependent population processes—to connect the exponentially-large but exact micro-scale Markov chain, through to a highly reduced system of ordinary differential equations. Thirdly, we show that both degree-based and individual-based mean-field approximations can be captured in a unified way through our general framework. Crucially, our approach is rigorous, elucidates the averaging process and highlights sources of error.

We start by describing Markov chain dynamics on networks in Section 2 and approximate lumping in Section 3. In Section 4 we describe how we use a partition of vertices to define an approximate lumping. The combinatorics to derive our mean-field approximation is involved, so in Section 5 we describe the process for the case where the vertices are partitioned into two sets, before we generalise this to an arbitrary finite number of partitions in Section 6. Furthermore, we present a simple example in the corresponding sections of the paper's Supplementary Materials. Readers interested in the working details of our approach may find it helpful to read Sections SM2–SM5 of the Supplementary Materials as they read the corresponding sections of the main paper. We treat some special cases of our density dependent population process in Section 7, then derive the large network limit mean-field equations in Section 8. In Sections 9 and 10 we are then able to derive degree-based and individual-based mean-field approximations respectively, and we consider degree-based mean-field approximation of the configuration model in Section 11. Finally we discuss our findings in Section 12.

2. Mathematical background. Let $G = (V, E)$ denote a graph or network with vertex set V and edge set $E \subset V \times V$, where the number of vertices is $N = |V|$. Unless otherwise stated, we consider dynamical processes on finite simple networks (i.e. undirected, unweighted with no self-loops or multiple edges) described by continuous-time Markov chains where each vertex can be in one of a finite number M of *vertex-states* and the set of possible vertex-states is $\mathcal{W} = \{\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_M\}$.

2.1. State-space. The state-space of the Markov chain is the set of all permutations of N vertex-states chosen from \mathcal{W} with repetition. This is equivalent to $\mathcal{S} = \mathcal{W}^V$, i.e. the set of all functions from V to \mathcal{W} , and so if the network is in state $S \in \mathcal{S}$ then the vertex-state of vertex $v \in V$ is $S(v)$. From here onwards we will refer to the states $S \in \mathcal{S}$ as *microstates*, to clearly distinguish them from vertex-states and *lumped states*, which will be introduced in Section 3. The number of microstates in \mathcal{S} is M^N , where N is the number of vertices, which is extremely large for even moderate N . However, since \mathcal{S} is finite we can enumerate the microstates so that $\mathcal{S} = \{S^{[1]}, S^{[2]}, \dots, S^{[M^N]}\}$.

2.2. Transitions. We assume that changes in microstate correspond to a single vertex $v \in V$ changing its vertex-state, and the rate that this occurs is a function of only the number of v 's neighbours in each of the vertex-states. We also assume that this rate function is the same for all vertices. Thus we assume model dynamics are driven by local interactions captured via a collection of functions between vertex-states. Note that a more general class of models would allow for behaviours in which multiple vertices change vertex-states at once, for example if a vertex exports its vertex-state to its neighbours [36]. We will now give a precise definition of the network dynamics that we consider.

DEFINITION 2.1. *For a finite non-empty set of vertex-states \mathcal{W} and $\mathcal{A}, \mathcal{B} \in \mathcal{W}$, let*

$$\mathbf{R}_{\mathcal{A}, \mathcal{B}} : \mathbb{Z}_{\geq 0}^M \rightarrow \mathbb{R}_{\geq 0}.$$

A vertex-state transition matrix (VSTM) \mathbf{R} is the collection of functions $\mathbf{R}_{\mathcal{A}, \mathcal{B}}$ for each $\mathcal{A}, \mathcal{B} \in \mathcal{W}$.

In the models we consider, $\mathbf{R}_{\mathcal{A}, \mathcal{B}}(n_1, n_2, \dots, n_M) \geq 0$ gives the rate that a vertex in vertex-state \mathcal{A} changes to vertex-state \mathcal{B} if it has n_1 neighbours in vertex-state \mathcal{W}_1 , n_2 neighbours in vertex-state \mathcal{W}_2 , etc. If transitions between a pair of vertex-states $\mathcal{A}, \mathcal{B} \in \mathcal{W}$ do not occur in a particular model, then the rate $\mathbf{R}_{\mathcal{A}, \mathcal{B}}$ is identically zero.

DEFINITION 2.2. *A homogeneous Single-Vertex Transition model (SVT) is a pair $(\mathcal{W}, \mathbf{R})$.*

We think of an SVT $\mathcal{M} = (\mathcal{W}, \mathbf{R})$ as a directed graph over vertex-states where a directed edge goes from vertex-state \mathcal{A} to \mathcal{B} if $\mathbf{R}_{\mathcal{A}, \mathcal{B}}$ is not identically zero. In this paper, our main focus will be on models whose VSTMs are affine functions, so

$$(2.1) \quad \mathbf{R}_{\mathcal{A}, \mathcal{B}}(n_1, n_2, \dots, n_M) = \zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} n_m,$$

where all of the constants $\zeta_m^{\mathcal{A}, \mathcal{B}}$ are non-negative. We will refer to SVTs with affine VSTMs as affine SVTs. Most SVTs have VSTMs of this form [37], although notable exceptions include the non-zero temperature Ising-Glauber dynamics [14], the nonlinear q -voter model [4] and threshold models [39].

2.3. Kolmogorov equations: infinitesimal generator. Given the network G and model \mathcal{M} , we need to define the corresponding continuous-time Markov chain. Let $X(t) = (X_1(t), X_2(t), \dots, X_{M^N}(t))^T$ be the time-dependent Markov chain probability distribution over \mathcal{S} , where $X_i(t)$ is the probability of being in microstate $S^{[i]}$ at time t . The evolution of $X(t)$ is then given by the forward Kolmogorov or master equation [22],

$$\dot{X} = \mathbf{Q}^T X,$$

where \mathbf{Q} is the infinitesimal generator, an M^N by M^N matrix in which each off-diagonal component \mathbf{Q}_{kl} gives the transition rate from $S^{[k]}$ to $S^{[l]}$, and the diagonal components ensure that rows sum to zero so that probability is conserved. We assume that a vertex changes vertex-state instantaneously, thus transitions only occur between pairs of microstates that differ in exactly one vertex-state.

DEFINITION 2.3. *A pair of states $S^{[k]}, S^{[l]} \in \mathcal{S}$ forms a **transition pair** with transition vertex v , denoted $S^{[k]} \rightsquigarrow S^{[l]}$, if $S^{[k]}(v) \neq S^{[l]}(v)$ and $S^{[k]}(u) = S^{[l]}(u)$ for all $u \neq v$.*

For vertex v and microstate $S^{[k]}$ let $n^{[k]}(v) = (n_1^{[k]}(v), n_2^{[k]}(v), \dots, n_M^{[k]}(v))$, where $n_m^{[k]}$ is the number of neighbours of v with vertex-state $\mathcal{W}_m \in \mathcal{W}$ in microstate $S^{[k]}$. Thus for $S^{[k]}, S^{[l]} \in \mathcal{S}$ and $S^{[k]} \neq S^{[l]}$, the transition rate from $S^{[k]}$ to $S^{[l]}$ in an SVT is given by

$$\mathbf{Q}_{kl} = \begin{cases} \mathbf{R}_{S^{[k]}(v), S^{[l]}(v)}(n^{[k]}(v)) & \text{if } S^{[k]} \overset{v}{\sim} S^{[l]} \\ 0 & \text{otherwise} \end{cases},$$

where vertex v is the transition vertex (if the states $S^{[k]}$ and $S^{[l]}$ form a transition pair) that goes from vertex-state $S^{[k]}(v)$ to $S^{[l]}(v)$.

3. Coarse-graining via lumping: theoretical foundation. We consider *lumping* of Markov chains [20]. Let $\Pi_{\mathcal{S}} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n\}$ be a partition of microstate-space, so $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset$ for each $i \neq j$, and $\cup_i \mathcal{S}_i = \mathcal{S}$. An *exact lumping* is a partition of microstate-space $\Pi_{\mathcal{S}}$ that preserves the Markov property, a necessary and sufficient condition for which is that the sum of transition rates from microstate $S^{[k]} \in \mathcal{S}_i$ to microstates in the cell \mathcal{S}_j , i.e.

$$\sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl},$$

is the same for all microstates $S^{[k]}$ in the cell \mathcal{S}_i . In matrix notation [37], this is equivalent to the existence of an $n \times n$ matrix \mathbf{q} such that

$$(3.1) \quad \mathbf{Q}\mathbf{C} = \mathbf{C}\mathbf{q},$$

where $\mathbf{C} \in \{0, 1\}^{M^N \times n}$ is the *collector matrix* [3] whose kj th component is

$$(3.2) \quad \mathbf{C}_{kj} = \begin{cases} 1 & \text{if } S^{[k]} \in \mathcal{S}_j, \\ 0 & \text{otherwise.} \end{cases}$$

The collector matrix collects those microstates in a column that belong to the same cell, or in other words the same “lumped state”, in the partition.

We call (3.1) the *lumpability condition*. Note that \mathbf{q} can be given explicitly for an exact lumping by introducing the *distributor matrix* [3] $\mathbf{D} \in \mathbb{R}^{n \times M^N}$, whose il th component is

$$(3.3) \quad \mathbf{D}_{il} = \begin{cases} \frac{1}{|\mathcal{S}_i|} & \text{if } S^{[l]} \in \mathcal{S}_i, \\ 0 & \text{otherwise.} \end{cases}$$

Specifically, $\Pi_{\mathcal{S}}$ satisfies the lumpability condition when \mathbf{Q} commutes with $\mathbf{C}\mathbf{D}$ [37]. Note that $\mathbf{D}\mathbf{C} = \mathbf{I}$, the identity matrix, hence multiplying (3.1) by \mathbf{D} we get the generator \mathbf{q} of the lumped system explicitly as

$$(3.4) \quad \mathbf{q} = \mathbf{D}\mathbf{Q}\mathbf{C}.$$

We use $x(t) = (x_1(t), \dots, x_n(t))^T$ to denote the time dependent Markov chain probability distribution over $\Pi_{\mathcal{S}}$, where $x_i(t)$ is the probability of being in the lumping partition cell \mathcal{S}_i . For this reason, we will refer to \mathcal{S}_i as a *lumped state*. When the lumpability condition is satisfied, the evolution of $x(t)$ is determined by the lumped master equation

$$(3.5) \quad \dot{x} = \mathbf{q}^T x,$$

and if $x(0) = \mathbf{C}^T X(0)$ then we have $x(t) = \mathbf{C}^T X(t)$ for all t . In other words, for each $\mathcal{S}_i \in \Pi_{\mathcal{S}}$, the sum of the probabilities of being in each microstate in \mathcal{S}_i at time t is equal to $x_i(t)$.

A lumping that does not satisfy the lumpability condition is an *approximate lumping* [3]. Given a partition $\Pi_{\mathcal{S}}$ of microstate-space that does not satisfy the lumpability condition (3.1), our approach is to still use the set of lumped states $\Pi_{\mathcal{S}}$ and the corresponding generator $\mathbf{q} = \mathbf{DQC}$, then solve the lumped master equation (3.5) for $x(t)$. Note that while this defines a Markov chain, it does not directly correspond to the SVT that it has been derived from, so we do not expect $x(t)$ to equal $\mathbf{C}^T X(t)$ for all t . We will however assume that the initial condition of the approximate lumping can be chosen so that $x(0) = \mathbf{C}^T X(0)$. Using the definition of \mathbf{C} (3.2), we have

$$(\mathbf{QC})_{kj} = \sum_{l=1}^{M^N} \mathbf{Q}_{kl} \mathbf{C}_{lj} = \sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl},$$

i.e. $(\mathbf{QC})_{kj}$ is the sum of the rates out of the microstate $S^{[k]}$ into the j^{th} lumped state when $S^{[k]} \notin \mathcal{S}_j$ and minus the sum of rates out of microstate $S^{[k]}$ when $S^{[k]} \in \mathcal{S}_j$. Then using the definition of the distributor matrix (3.3) we have

$$\mathbf{q}_{ij} = (\mathbf{DQC})_{ij} = \sum_{k=1}^{M^N} \mathbf{D}_{ik} (\mathbf{QC})_{kj} = \sum_{k=1}^{M^N} \mathbf{D}_{ik} \sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl} = \frac{1}{|\mathcal{S}_i|} \sum_{S^{[k]} \in \mathcal{S}_i} \sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl}.$$

Thus \mathbf{q}_{ij} is the average of the sum of rates out of microstates in the i^{th} lumped state and into the j^{th} lumped state.

Summarising, we can say that starting from the full infinitesimal generator, \mathbf{Q} , and choosing a partition of the state-space, equation (3.7) yields the infinitesimal generator \mathbf{q} of the lumped (coarse-grained) system. Note that the partition of the microstate-space is encoded in the collector and distributor matrices, \mathbf{C} and \mathbf{D} respectively.

4. Lumping based on vertex set partitions. In the previous section we introduced the notion of lumping in general, however we have not yet described how we will determine the partition of microstate-space, on which the lumping is based. This will be dealt with in this section.

4.1. Motivating example. As a simple example, consider an SIS epidemic on a graph with three nodes. Before we consider specific graphs, we first define some notation. We will use \mathcal{B} (for blue) to denote susceptible vertices and \mathcal{R} (for red) to denote infected vertices, and we will represent a microstate with three such letters. An important property of a useful lumping is that two microstates in the same partition cell should have the same number of infected nodes. This ensures that the total number of infected nodes (i.e. the prevalence) can be determined from the lumped system as well. The simplest lumping that preserves this property contains the following four lumping classes: $\mathcal{S}_1 = \{\mathcal{BBB}\}$, $\mathcal{S}_2 = \{\mathcal{BBR}, \mathcal{BRB}, \mathcal{RBB}\}$, $\mathcal{S}_3 = \{\mathcal{BRR}, \mathcal{RBR}, \mathcal{RRB}\}$, $\mathcal{S}_4 = \{\mathcal{RRR}\}$. Generalizing this idea to arbitrary dynamics, it has been shown that approximate lumping based on partitions of microstate-space into sets of microstates with the same number of vertices in each vertex-state result in mean-field birth-death processes for $M = 2$ and mean-field population models for $M > 2$ [38].

Returning to the example above, one can realise that the lumping does not take into account the graph structure at all. For example, having a path graph with three nodes, the degree of the central node is two, while that of the nodes at the left and right end is only one, which could be reflected by the lumping. A natural choice of lumping with a finer partition is the following: $\mathcal{S}_1 = \{\mathcal{B}\mathcal{B}\mathcal{B}\}$, $\mathcal{S}_2 = \{\mathcal{B}\mathcal{B}\mathcal{R}, \mathcal{R}\mathcal{B}\mathcal{B}\}$, $\mathcal{S}_3 = \{\mathcal{B}\mathcal{R}\mathcal{B}\}$, $\mathcal{S}_4 = \{\mathcal{B}\mathcal{R}\mathcal{R}, \mathcal{R}\mathcal{R}\mathcal{B}\}$, $\mathcal{S}_5 = \{\mathcal{R}\mathcal{B}\mathcal{R}\}$, $\mathcal{S}_6 = \{\mathcal{R}\mathcal{R}\mathcal{R}\}$. In this lumping partition, the central node plays a different role than the end nodes. This will be interpreted as follows: the nodes are divided into two groups, the central node and the end nodes. The lumping is based on the number of infected nodes in these groups. The number of infected nodes in the first group can be 0 or 1 while in the second group it can be 0, 1 or 2. Thus the lumping classes above can be encoded by the number of infected nodes in the two groups as follows: \mathcal{S}_1 by $s^{[1]} = (0, 0)$, \mathcal{S}_2 by $s^{[2]} = (0, 1)$, \mathcal{S}_3 by $s^{[3]} = (1, 0)$, \mathcal{S}_4 by $s^{[4]} = (1, 1)$, \mathcal{S}_5 by $s^{[5]} = (0, 2)$ and \mathcal{S}_6 by $s^{[6]} = (1, 2)$. (Note that our notation distinguishes between subsets of microstates in a lumping class, such as \mathcal{S}_1 , and the encoding of that lumping class according to the number of infected vertices in each group, such as $s^{[1]}$.) This encoding can be extended by the number of susceptible nodes, which in this example is redundant information, so that for example

$$\mathbf{s}^{[5]} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix},$$

where the first row of the matrix shows the number of susceptible nodes in the two groups, while the second row contains the number of infected nodes. The other lumped states can be similarly encoded by 2×2 matrices. This notation will be used below for the general case.

4.2. Microstate-space partition based on a vertex set partition. Now we generalise the above approach by considering partitions of the *vertex* set, where the lumped states are based on partitions of microstate-space into sets of microstates with the same number of vertices in each vertex-state *within* each of the cells of the *vertex* partition. This defines the lumped states that we consider, and so in the remainder of the paper we will primarily refer to lumped states in terms of these counts, rather than cells in the partition of microstate-space $\Pi_{\mathcal{S}}$ (like \mathcal{S}_i , which are subsets of microstates).

Let $\Pi_V = \{V_1, \dots, V_P\}$ be a partition of the vertex set V , such that $V_p \cap V_q = \emptyset$ for $p \neq q$ and $\cup_p V_p = V$. We now define the lumped macrostate-space and the corresponding partition of microstate-space precisely.

DEFINITION 4.1. For a network $G = (V, E)$, with $N = |V|$, and an SVT $\mathcal{M} = (\mathcal{W}, \mathbf{R})$, with $M = |\mathcal{W}|$, let Π_V be a vertex-partition, where $P = |\Pi_V|$ and $N_p = |V_p|$ for each $V_p \in \Pi_V$. The corresponding vertex-partition macrostate-space is the set of non-negative, integer valued matrices $\mathbf{s} \in \mathbb{Z}_{\geq 0}^{M \times P}$ such that

$$(4.1) \quad \sum_{m=1}^M \mathbf{s}_{m,p} = N_p.$$

Note that since $\sum_{p=1}^P N_p = N$, we have $\sum_{p=1}^P \sum_{m=1}^M \mathbf{s}_{m,p} = N$.

DEFINITION 4.2. Let $(\mathbf{s}^{[1]}, \mathbf{s}^{[2]}, \dots, \mathbf{s}^{[n]})$ be an enumeration of a vertex-partition macrostate-space. A vertex-partition lumping is a partition of microstates $\Pi_{\mathcal{S}} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n\}$ such that for $\mathcal{S}_i \in \Pi_{\mathcal{S}}$, the number of vertices in vertex-state $\mathcal{W}_m \in \mathcal{W}$ in vertex-partition cell $V_p \in \Pi_V$ is $\mathbf{s}_{m,p}^{[i]}$.

5. The lumped generator for two cell vertex-partitions. We now describe the approximate lumping approach using two vertex-partition cells for finite graphs and homogeneous SVTs. The fully general case with a finite number of vertex-partitions cells will be presented in Section 6. Let us consider the case $P = 2$ in Definition 4.1, i.e. denote the vertex partition by $\Pi_V = \{V_1, V_2\}$, where $V_1 \cap V_2 = \emptyset$ and $V_1 \cup V_2 = V$; also $N_1 = |V_1|$ and $N_2 = |V_2|$. For finite M and $P = 2$, a lumped state will be denoted by a matrix $\mathbf{s} \in \mathbb{Z}_{\geq 0}^{M \times P}$ whose m, p th entry, $\mathbf{s}_{m,p}$, is the number of vertices in vertex-state \mathcal{W}_m in the vertex partition cell V_p . Thus summing the entries in the p th column of \mathbf{s} yields N_p , as in (4.1), and summing all entries yields N . A lumped state corresponds to choosing N_1 vertices from the M possible vertex-states with repetition in the first partition and N_2 vertices from the M possible vertex-states with repetition in the second partition. Thus the total number of lumped states is

$$(5.1) \quad n = \binom{N_1 + M - 1}{N_1} \binom{N_2 + M - 1}{N_2}.$$

Consequently we can number the lumped states as $\mathbf{s}^{[1]}, \mathbf{s}^{[2]}, \dots, \mathbf{s}^{[n]}$, and the lumping partition is then $\Pi_S = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n\}$, where

$$\mathcal{S}_i = \left\{ S^{[j]} \in \mathcal{S} \mid \sum_{v \in V_p} \delta_{S^{[j]}(v), \mathcal{W}_m} = \mathbf{s}_{mp}^{[i]}, \mathcal{W}_m \in \mathcal{W}, V_p \in \Pi_V \right\},$$

and $\delta_{A,B}$ is the Kronecker delta function. Note that for a given microstate $S^{[j]}$, this simply counts the number of vertices in each vertex-state and in each vertex partition cell, and checks whether these match the corresponding values in $\mathbf{s}^{[i]}$; if they do, then microstate $S^{[j]}$ is assigned to partition cell \mathcal{S}_i . We use $\mathbf{s}_p = (\mathbf{s}_{1,p}, \mathbf{s}_{2,p}, \dots, \mathbf{s}_{M,p})^T$ to denote the p th column of \mathbf{s} . We will only use a single subscript on \mathbf{s} when referring to this vector.

5.1. The lumped generator for an arbitrary two cell vertex-partition.

We will start by considering the transition rate from an arbitrary lumped state $\mathbf{s}^{[i]}$ to another arbitrary lumped state $\mathbf{s}^{[j]} \neq \mathbf{s}^{[i]}$. In Section 3, we saw that in general this is given by

$$(5.2) \quad \mathbf{q}_{ij} = \frac{1}{|\mathcal{S}_i|} \sum_{S^{[k]} \in \mathcal{S}_i} \sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl},$$

but it turns out that rather than summing over microstates, as (5.2) suggests, it is easier to consider the possible transitions of individual vertices and sum their rates.

First note that the number of arrangements of vertex-states over vertices in the p th cell that correspond to $\mathbf{s}_p^{[i]}$ is the multinomial

$$\binom{N_p}{\mathbf{s}_p^{[i]}} := \binom{N_p}{\mathbf{s}_{1,p}^{[i]} \mathbf{s}_{2,p}^{[i]} \dots \mathbf{s}_{M,p}^{[i]}},$$

so the total number of microstates in the lumping partition cell \mathcal{S}_i (which corresponds to the lumped state $\mathbf{s}^{[i]}$) is

$$|\mathcal{S}_i| = \binom{N}{\mathbf{s}^{[i]}} := \prod_{p=1}^P \binom{N_p}{\mathbf{s}_p^{[i]}}.$$

Note we will always have a scalar on the top part of our generalised multinomial notation. When we have a vector on the bottom part, the notation corresponds to a multinomial over the entries in the vector, and when we have a matrix on the bottom, it will correspond to the *product* of the multinomials of each of the columns of the matrix. We will also assume the typical convention that if any of the terms in the multinomial are negative, or if the sum of the terms on the bottom of the multinomial are not equal to the top, then the value of the multinomial is zero.

Without loss of generality we assume that the transition rate \mathbf{q}_{ij} from the lumped state $\mathbf{s}^{[i]}$ to the lumped state $\mathbf{s}^{[j]}$ corresponds to a vertex in vertex-partition $q \in \{1, 2\}$ transitioning from vertex-state $\mathcal{A} \in \mathcal{W}$ to $\mathcal{B} \in \mathcal{W}$, where $\mathcal{A} \neq \mathcal{B}$. To understand why, suppose that $S^{[k]} \in \mathcal{S}_i$, then for there to be a non-zero rate from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ there must be a microstate $S^{[l]} \in \mathcal{S}_j$ such that $S^{[k]}$ and $S^{[l]}$ form a transition pair, whose transition vertex is v say. We are free to use V_q to label which vertex-partition cell v belongs to, and we may also use the labels $\mathcal{A}, \mathcal{B} \in \mathcal{W}$ to indicate what vertex-states v changes from and to respectively. Moreover, any other non-zero transition rate from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ must also correspond to a vertex in V_q changing from \mathcal{A} to \mathcal{B} , otherwise it would result in different counts of vertices of each type of vertex-state in each vertex-partition cell, i.e. a lumped state different to $\mathbf{s}^{[j]}$. To compute (5.2), for each $v \in V_q$ we can construct all possible microstates where v has vertex-state \mathcal{A} . If in this process we specify the vertex-states of the neighbours of v , then we can determine the rate at which v changes from \mathcal{A} to \mathcal{B} . Summing this contribution from all possible cases yields \mathbf{q}_{ij} . A proof of this will be given in Section 6. In the following paragraphs we will build up the components of this sum.

First we define some notation related to the neighbourhoods of vertices. Let d_p^v denote the number of neighbours of vertex v in the p th vertex-partition cell. The degree of vertex v is

$$(5.3) \quad d^v := \sum_{p=1}^P d_p^v.$$

We represent the neighbourhood of v using a non-negative, integer-valued matrix $\mathbf{n}^v \in \mathbb{Z}_{\geq 0}^{M \times P}$, which we call a *neighbourhood count*.

DEFINITION 5.1. *For a vertex $v \in V$, a neighbourhood count is a matrix $\mathbf{n}^v \in \mathbb{Z}_{\geq 0}^{M \times P}$ such that*

$$\sum_{m=1}^M \mathbf{n}_{m,p}^v = d_p^v,$$

for $0 \leq p \leq P$.

Note that it follows from (5.3) that $\sum_{p=1}^P \sum_{m=1}^M \mathbf{n}_{m,p}^v = d^v$. We will use a single index on this matrix to indicate a column, i.e. $\mathbf{n}_p^v \in \mathbb{Z}_{\geq 0}^M$ is the p th column of \mathbf{n}^v . The component $\mathbf{n}_{m,p}^v$ of a neighbourhood count is the number of neighbours of vertex v in the m th vertex-state and in the p th vertex-partition cell.

The number of ways that we can arrange the vertex-states of the neighbours of v in vertex-partition cell $p \neq q$ according to some \mathbf{n}_p^v , as well as the vertex-states of the other vertices in vertex-partition cell $p \neq q$ according to \mathbf{s}_p is

$$(5.4) \quad A(\mathbf{s}_p, \mathbf{n}_p^v) := \binom{d_p^v}{\mathbf{n}_p^v} \binom{N_p - d_p^v}{\mathbf{s}_p - \mathbf{n}_p^v}.$$

Recall that a vector in the bottom of the multinomial coefficient notation indicates that the elements of the vector should be in the denominator of the multinomial coefficient. While in (5.4) we have used d_p^v and N_p in the top of the multinomial coefficients, we will assume that these values are actually determined from summing the vectors in the bottom. We will also assume the standard convention that a multinomial coefficient is zero if any entry is negative.

For the vertex-partition cell q , which contains v , the number of ways that we can arrange the neighbours of v according to \mathbf{n}_q^v is

$$A(\mathbf{s}_q - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) = \binom{d_q^v}{\mathbf{n}_q^v} \binom{N_q - 1 - d_q^v}{\mathbf{s}_q - \mathbf{e}_{\mathcal{A}} - \mathbf{n}_q^v},$$

where $\mathbf{e}_{\mathcal{A}}$ is a vector of length M with a one in the entry corresponding to vertex-state \mathcal{A} and zeros elsewhere. This is to account for the fact that we assumed vertex v is in vertex-state \mathcal{A} , so there is one less \mathcal{A} vertex in \mathbf{s}_q .

For a given \mathbf{n}^v and $p \neq q$, using (5.4) there are

$$(5.5) \quad A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v)$$

microstates in \mathcal{S}_i in which vertex v in vertex partition cell q is in vertex-state \mathcal{A} , its neighbours' vertex-states correspond to \mathbf{n}^v , and the total number of vertices in each vertex-state and in each vertex-partition cell corresponds to $\mathbf{s}^{[i]}$. We also need to know the rate that vertex v will change from vertex-state \mathcal{A} to \mathcal{B} . For SVTs with affine VSTM given by (2.1), if vertex v is in vertex-state \mathcal{A} and the number of its neighbours in each of the vertex-states in each of the vertex-partition cells is given by \mathbf{n}^v , then it will transition from \mathcal{A} to \mathcal{B} with rate

$$\zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{n}_{m,r}^v.$$

Recall that we assumed, without loss of generality, that \mathbf{q}_{ij} corresponds to a vertex in vertex-partition cell V_q changing from vertex-state $\mathcal{A} \in \mathcal{W}$ to $\mathcal{B} \in \mathcal{W}$. We can now compute \mathbf{q}_{ij} by summing over all feasible realisations of the matrix \mathbf{n}^v and vertices in V_q , which yields

$$(5.6) \quad \mathbf{q}_{ij} = \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} \left(\zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{n}_{m,r}^v \right) A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v).$$

In this equation, the sums over $\mathbf{n}_1^v | d_1^v$ and $\mathbf{n}_2^v | d_2^v$ specify the $P = 2$ columns of \mathbf{n}^v .

It seems that we have swapped one difficult sum, (5.2), for another, (5.6). However, it turns out that the sum over vertex neighbourhoods can be simplified, and we will illustrate how this can be done in the next section.

5.2. Simplified form of the lumped generator. Our next step is to simplify (5.6). Crucial to achieving this is a generalisation of the Vandermonde identity and what we call the “sum-product property”, which for arbitrary v can be used to obtain

$$(5.7) \quad \binom{N}{\mathbf{s}} = \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} A(\mathbf{s}_1, \mathbf{n}_1^v) A(\mathbf{s}_2, \mathbf{n}_2^v).$$

A generalisation of this will be presented in Section 6 and further details can be found in the Supplementary Materials. First consider the $\zeta_0^{\mathcal{A}, \mathcal{B}}$ term in (5.6), then using (5.7)

we find

$$\begin{aligned} & \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} \zeta_0^{\mathcal{A}, \mathcal{B}} A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v) \\ &= \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \zeta_0^{\mathcal{A}, \mathcal{B}} \binom{N_q - 1}{\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}} \binom{N_p}{\mathbf{s}_p^{[i]}} = \sum_{v \in V_q} \zeta_0^{\mathcal{A}, \mathcal{B}} \frac{\mathbf{s}_{\mathcal{A}, q}^{[i]}}{N_q} = \zeta_0^{\mathcal{A}, \mathcal{B}} \mathbf{s}_{\mathcal{A}, q}^{[i]}, \end{aligned}$$

where we have used \mathcal{A} as an index in \mathcal{W} . This is what we might have expected: each microstate in $\mathbf{s}^{[i]}$ has $\mathbf{s}_{\mathcal{A}, q}^{[i]}$ vertices in vertex partition cell q in vertex-state \mathcal{A} , so there will be a factor of $\zeta_0^{\mathcal{A}, \mathcal{B}}$ for each of these.

Now consider the $\zeta_m^{\mathcal{A}, \mathcal{B}}$ terms in (5.6). The sums over m and r in (5.6) do not depend on any of the others sums, so we can move these to the outside. Note that

$$\begin{aligned} \mathbf{n}_{m, p}^v A(\mathbf{s}_p, \mathbf{n}_p^v) &= \mathbf{n}_{m, p}^v \binom{d_p^v}{\mathbf{n}_p^v} \binom{N_p - d_p^v}{\mathbf{s}_p - \mathbf{n}_p^v} \\ &= d_p^v \binom{d_p^v - 1}{\mathbf{n}_p^v - \mathbf{e}_m} \binom{N_p - 1 - (d_p^v - 1)}{\mathbf{s}_p - \mathbf{e}_m - (\mathbf{n}_p^v - \mathbf{e}_m)} \\ &= d_p^v A(\mathbf{s}_p - \mathbf{e}_m, \mathbf{n}_p^v - \mathbf{e}_m). \end{aligned}$$

For the case where $r = q$, it follows that

$$\begin{aligned} & \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{n}_{m, q}^v A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v) \\ &= \zeta_m^{\mathcal{A}, \mathcal{B}} d_q^v \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}} - \mathbf{e}_m, \mathbf{n}_q^v - \mathbf{e}_m) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v) \\ &= \zeta_m^{\mathcal{A}, \mathcal{B}} d_q^v \binom{N_q - 2}{\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}} - \mathbf{e}_m} \binom{N_p}{\mathbf{s}_p^{[i]}} \end{aligned}$$

where we have used (5.7) to obtain the third line. Similarly, for the case $r = p$ we have

$$\sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{n}_{m, p}^v A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v) = \zeta_m^{\mathcal{A}, \mathcal{B}} d_p^v \binom{N_q - 1}{\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}} \binom{N_p - 1}{\mathbf{s}_p^{[i]} - \mathbf{e}_m}.$$

Thus for the $\zeta_m^{\mathcal{A}, \mathcal{B}}$ terms in (5.6) we can cancel multinomial terms to get

$$\begin{aligned} & \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \left[d_q^v \binom{N_q - 2}{\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}} - \mathbf{e}_m} \binom{N_p}{\mathbf{s}_p^{[i]}} + d_p^v \binom{N_q - 1}{\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}} \binom{N_p - 1}{\mathbf{s}_p^{[i]} - \mathbf{e}_m} \right] \\ &= \sum_{v \in V_q} \sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \left[d_q^v \frac{\mathbf{s}_{\mathcal{A}, q}^{[i]} (\mathbf{s}_{m, q}^{[i]} - \delta_{\mathcal{A}, \mathcal{W}_m})}{N_q (N_q - 1)} + d_p^v \frac{\mathbf{s}_{\mathcal{A}, q}^{[i]} \mathbf{s}_{m, p}^{[i]}}{N_q N_p} \right], \end{aligned}$$

where recall that $\delta_{\mathcal{A}, \mathcal{B}}$ is the Kronecker delta function. Using these simplifications, the expression for \mathbf{q}_{ij} becomes

$$\mathbf{q}_{ij} = \mathbf{s}_{\mathcal{A}, q}^{[i]} \left\{ \zeta_0^{\mathcal{A}, \mathcal{B}} + \frac{1}{N_q} \sum_{v \in V_q} \sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \left[d_q^v \frac{(\mathbf{s}_{m, q}^{[i]} - \delta_{\mathcal{A}, \mathcal{W}_m})}{(N_q - 1)} + d_p^v \frac{\mathbf{s}_{m, p}^{[i]}}{N_p} \right] \right\}.$$

Note that this formula enables us to compute \mathbf{q} without direct reference to \mathbf{Q} .

6. The lumped generator for arbitrary finite vertex-partitions. We can now generalise relatively easily to arbitrary vertex partitions. We consider a vertex partition $\Pi = \{V_1, \dots, V_P\}$ with $V_p \cap V_q = \emptyset$ for $p \neq q$ and $\cup_p V_p = V$. We then have $\mathbf{s} \in \mathbb{Z}_{\geq 0}^{M \times P}$, whose m, p th entry, $\mathbf{s}_{m,p}$, is the number of vertices in vertex-state \mathcal{W}_m in the vertex-partition cell V_p . The total number of microstates is

$$M^N = \sum_{\mathbf{s}} \binom{N}{\mathbf{s}} := \sum_{\mathbf{s}_1 | N_1} \sum_{\mathbf{s}_2 | N_2} \cdots \sum_{\mathbf{s}_P | N_P} \prod_{p=1}^P \binom{N_p}{\mathbf{s}_p},$$

which can be obtained from application of the multinomial theorem and interchanging the product of sums with a sum of products (see the Supplementary Information for further details). The number of lumped states is

$$(6.1) \quad n = \prod_{p=1}^P \binom{N_p + M - 1}{N_p}.$$

Again we assume, without loss of generality, that a transition from the lumped state $\mathbf{s}^{[i]}$ to the lumped state $\mathbf{s}^{[j]} \neq \mathbf{s}^{[i]}$ corresponds to a vertex in vertex-partition cell V_q transitioning from vertex-state \mathcal{A} to \mathcal{B} .

Our initial goal is to write (5.2) in terms of a sum over vertices and possible neighbourhood counts. In order to prove this, we must first consider some properties of neighbourhood counts. Since we sum over neighbourhood counts, we need to be clear about which correspond to arrangements of vertex-states in microstates in the correct lumped partition cell, and that only these contribute non-zero values to the summation. To this end, we say a neighbourhood count \mathbf{n}^v is *realisable* if there is at least one microstate in which the number of neighbours of vertex v in each vertex-partition corresponds to \mathbf{n}^v .

DEFINITION 6.1. A neighbourhood count \mathbf{n}^v is realisable in $\mathcal{S}_i \subset \mathcal{S}$ if there is a microstate $S^{[k]} \in \mathcal{S}_i$ in which the number of neighbours of vertex v in vertex-state $\mathcal{W}_m \in \mathcal{W}$ and vertex-partition cell $V_p \in \Pi_V$ is $\mathbf{n}_{m,p}^v$.

There is a simple condition that ensures a neighbourhood count is realisable.

LEMMA 6.2. For $v \in V$, a neighbourhood count \mathbf{n}^v is realisable in $\mathcal{S}_i \subset \mathcal{S}$ if and only if $\mathbf{n}_{m,p}^v \leq \mathbf{s}_{m,p}^{[i]}$ for all m, p .

Proof. Suppose that \mathbf{n}^v is realisable in $\mathcal{S}_i \subset \mathcal{S}$, then there is a microstate $S^{[k]} \in \mathcal{S}_i$ such that the number of neighbours of vertex v in vertex-state $\mathcal{W}_m \in \mathcal{W}$ and vertex-partition cell $V_p \in \Pi_V$ is $\mathbf{n}_{m,p}^v$. Since $\mathbf{s}_{m,p}^{[i]}$ is the number of vertices in vertex-state $\mathcal{W}_m \in \mathcal{W}$ and vertex-partition cell $V_p \in \Pi_V$, we must have $\mathbf{n}_{m,p}^v \leq \mathbf{s}_{m,p}^{[i]}$ for all m, p .

Suppose that $\mathbf{n}_{m,p}^v \leq \mathbf{s}_{m,p}^{[i]}$ for all m, p . Then for each partition p , we can assign a vertex-state to each of the neighbours of v such that the total number of neighbours of v in vertex-state $\mathcal{W}_m \in \mathcal{W}$ and vertex-partition cell $V_p \in \Pi_V$ is $\mathbf{n}_{m,p}^v$. Similarly we can assign vertex-states to each of the remaining vertices (including v) in each vertex-partition cell such that the number of them in vertex-state $\mathcal{W}_m \in \mathcal{W}$ and vertex-partition cell $V_p \in \Pi_V$ is $\mathbf{s}_{m,p}^{[i]} - \mathbf{n}_{m,p}^v \geq 0$. Thus the total number of vertices in vertex-state $\mathcal{W}_m \in \mathcal{W}$ and vertex-partition cell $V_p \in \Pi_V$ is $\mathbf{s}_{m,p}^{[i]}$, and hence this assignment of vertex-states corresponds to a microstate in \mathcal{S}_i . \square

Importantly, when we sum over neighbourhood counts, only those that are realisable contribute to the sum.

COROLLARY 6.3. For $v \in V$, a neighbourhood count \mathbf{n}^v is realisable in $\mathcal{S}_i \subset \mathcal{S}$ if and only if

$$(6.2) \quad \prod_{p=1}^P A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v) > 0.$$

Proof. Suppose that for $v \in V$, a neighbourhood count \mathbf{n}^v is realisable in $\mathcal{S}_i \subset \mathcal{S}$, then by Lemma 6.2 $\mathbf{s}_{m,p}^{[i]} - \mathbf{n}_{m,p}^v \geq 0$ for all m, p and hence (6.2) is true. Conversely, if (6.2) is true then each term in the product is positive and hence $\mathbf{n}_{m,p}^v \leq \mathbf{s}_{m,p}^{[i]}$ for all m, p . Then by Lemma 6.2 \mathbf{n}^v is realisable in $\mathcal{S}_i \subset \mathcal{S}$. \square

We may now prove that lumped rate (5.2) can be written in terms of a sum over vertices and possible neighbourhood counts. In the following lemma, we use $\mathbf{1}_P$ to denote a column vector with P entries, each 1.

LEMMA 6.4. Let \mathcal{S} be the microstate-space of a homogeneous SVT with affine VSTM on a network with vertex set V and let $\Pi_V = \{V_1, V_2, \dots, V_P\}$ be a partition of V . Suppose that the corresponding vertex-partition lumping macrostate-space is $(\mathbf{s}^{[1]}, \mathbf{s}^{[2]}, \dots, \mathbf{s}^{[n]})$ (Definition 4.1) and the partition of microstates is $\Pi_{\mathcal{S}} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n\}$ (Definition 4.2). If a transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to a vertex $v \in V_q$ changing from vertex-state \mathcal{A} to vertex-state \mathcal{B} , then

$$(6.3) \quad \sum_{S^{[k]} \in \mathcal{S}_i} \sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl} = \sum_{v \in V_q} \sum_{\mathbf{n}^v | d^v} \mathbf{R}_{\mathcal{A}, \mathcal{B}}(\mathbf{n}^v \mathbf{1}_P) \prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q} \mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v)$$

Proof. We construct a surjective map from transition pairs on the left of (6.3) to feasible neighbourhood counts \mathbf{n}^v on the right of (6.3), and show that the multinomial terms on the right account for the many-to-one multiplicity of the mapping. For $S^{[k]} \in \mathcal{S}_i$ and $S^{[l]} \in \mathcal{S}_j$, \mathbf{Q}_{kl} can only be non-zero if $S^{[k]} \stackrel{v}{\sim} S^{[l]}$, which identifies a unique vertex v and by assumption $v \in V_q$. Then $\mathbf{Q}_{kl} = \mathbf{R}_{S^{[k]}(v), S^{[l]}(v)}(n^{[k]}(v))$. Furthermore, the microstate $S^{[k]}$ corresponds to a unique neighbourhood count \mathbf{n}^v , which is evidently realisable. Since $n^{[k]}(v)$ is derived from $S^{[k]}$, we have $n^{[k]}(v) = \mathbf{n}^v \mathbf{1}_P$. It is assumed that a transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to a vertex $v \in V_q$ changing from vertex-state \mathcal{A} to vertex-state \mathcal{B} , thus we have $S^{[k]}(v) = \mathcal{A}$ and $S^{[l]}(v) = \mathcal{B}$. Consequently, each transition pair in the summation on the left of (6.3) corresponds to a unique pair $v \in V_q$ and $\mathbf{n}^v | d^v$ in the summation on the right with rate $\mathbf{R}_{\mathcal{A}, \mathcal{B}}(\mathbf{n}^v \mathbf{1}_P)$. Since \mathbf{n}^v is realisable and the vertex-state of $v \in V_q$ is \mathcal{A} , $\prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q} \mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v) > 0$.

For a given $v \in V_q$ and neighbourhood count \mathbf{n}_p^v , when $p \neq q$ there are

$$A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v) = \binom{d_p^v}{\mathbf{n}_p^v} \binom{N_p - d_p^v}{\mathbf{s}_p^{[i]} - \mathbf{n}_p^v}$$

ways of arranging the vertex-states of the neighbours of v in vertex-partition cell p according to \mathbf{n}_p^v , as well as the vertex-states of the other vertices in vertex-partition cell p according to $\mathbf{s}_p^{[i]}$. Similarly, for the case $p = q$ we have

$$A(\mathbf{s}_p^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v) = \binom{d_p^v}{\mathbf{n}_p^v} \binom{N_p - 1 - d_p^v}{\mathbf{s}_p^{[i]} - \mathbf{e}_{\mathcal{A}} - \mathbf{n}_p^v}.$$

Thus using Corollary 6.3, there are

$$(6.4) \quad \prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q} \mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v)$$

microstates in \mathcal{S}_i in which the vertex-state of v is \mathcal{A} and its neighbourhood count is \mathbf{n}^v . These microstates form transition pairs with the corresponding states in \mathcal{S}_j in which the vertex-state of v is \mathcal{B} . Thus the number of transition pairs on the left of (6.3) associated with each $v \in V_q$ and realisable neighbourhood count \mathbf{n}^v in the summation on the right is given by (6.4). \square

We saw in Section 5.2 that we needed a generalisation of the Vandermonde identity. This can be stated as

LEMMA 6.5. *Let \mathbf{s} be a lumped state and \mathbf{n}^v a neighbourhood count of a vertex $v \in V$, then*

$$(6.5) \quad \binom{N}{\mathbf{s}} = \sum_{\mathbf{n}^v | d^v} \prod_{p=1}^P A(\mathbf{s}_p, \mathbf{n}_p^v).$$

A detailed proof can be found in the Supplementary Materials.

We now present the main result of the paper.

THEOREM 6.6. *Let \mathcal{S} be the state-space of a homogeneous SVT with affine VSTM on a network with vertex set V and let $\Pi_V = \{\Pi_1, \Pi_2, \dots, \Pi_P\}$ be a partition of V . Suppose that $\mathbf{q} = \mathbf{DQC}$ is the lumped infinitesimal generator corresponding to the vertex-partition lumping with macrostate-space $(\mathbf{s}^{[1]}, \mathbf{s}^{[2]}, \dots, \mathbf{s}^{[n]})$. If a transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to a vertex $v \in V_q$ changing from vertex-state \mathcal{A} to vertex-state \mathcal{B} , then*

$$(6.6) \quad \mathbf{q}_{ij} = \mathbf{s}_{\mathcal{A},q}^{[i]} \left\{ \zeta_0^{\mathcal{A},\mathcal{B}} + \frac{1}{N_q} \sum_{r=1}^P \left(\frac{\sum_{v \in V_q} d_r^v}{N_r - \delta_{q,r}} \right) \left[\sum_{m=1}^M \zeta_m^{\mathcal{A},\mathcal{B}} \left(\mathbf{s}_{m,r}^{[i]} - \delta_{\mathcal{A},\mathcal{W}_m} \delta_{q,r} \right) \right] \right\}.$$

Proof. From (5.2) we have

$$\mathbf{q}_{ij} = \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{S^{[k]} \in \Pi_i} \sum_{S^{[l]} \in \Pi_j} \mathbf{Q}_{km},$$

since the number of states in Π_i is $|\Pi_i| = \binom{N}{\mathbf{s}^{[i]}}$. Using Lemma 6.4 and (2.1) it follows that

$$(6.7) \quad \mathbf{q}_{ij} = \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{\mathbf{n}^v | d^v} \left(\zeta_0^{\mathcal{A},\mathcal{B}} + \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A},\mathcal{B}} \mathbf{n}_{m,r}^v \right) \prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q} \mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v)$$

We will deal with the $\zeta_0^{\mathcal{A},\mathcal{B}}$ and $\zeta_m^{\mathcal{A},\mathcal{B}} \mathbf{n}_m$ terms separately. From Lemma 6.5, the sums around the constant term $\zeta_0^{\mathcal{A},\mathcal{B}}$ are

$$(6.8) \quad \begin{aligned} \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{\mathbf{n}^v | d^v} \zeta_0^{\mathcal{A},\mathcal{B}} \prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q} \mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v) &= \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \zeta_0^{\mathcal{A},\mathcal{B}} \prod_{p=1}^P \binom{N_p - \delta_{p,q}}{\mathbf{s}_p^{[i]} - \delta_{p,q} \mathbf{e}_{\mathcal{A}}}, \\ &= \zeta_0^{\mathcal{A},\mathcal{B}} \mathbf{s}_{\mathcal{A},q}^{[i]}. \end{aligned}$$

448 In (6.7) we are able to do the sums over m and r after we sum over v and $\mathbf{n}^v|d^v$, since
 449 they do not determine \mathbf{n}^v . Thus we can focus on an individual term $\zeta_m^{\mathcal{A},\mathcal{B}}\mathbf{n}_{m,r}$; using
 450 Lemma 6.5 we have

$$451 \sum_{\mathbf{n}^v|d^v} \mathbf{n}_{m,r}^v \prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q}\mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v) = d_r^v \prod_{p=1}^P \binom{N_p - \delta_{p,q} - \delta_{p,r}}{\mathbf{s}_p^{[i]} - \delta_{p,q}\mathbf{e}_{\mathcal{A}} - \delta_{p,r}\mathbf{e}_m}.$$

453 We then find

$$454 \begin{aligned} & \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{\mathbf{n}^v|d^v} \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A},\mathcal{B}} \mathbf{n}_{m,r}^v \prod_{p=1}^P A(\mathbf{s}_p^{[i]} - \delta_{p,q}\mathbf{e}_{\mathcal{A}}, \mathbf{n}_p^v) \\ 455 &= \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A},\mathcal{B}} d_r^v \prod_{p=1}^P \binom{N_p - \delta_{p,q} - \delta_{p,r}}{\mathbf{s}_p^{[i]} - \delta_{p,q}\mathbf{e}_{\mathcal{A}} - \delta_{p,r}\mathbf{e}_m} \\ 456 (6.9) \quad &= \sum_{v \in V_q} \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A},\mathcal{B}} d_r^v \frac{\mathbf{s}_{\mathcal{A},q}^{[i]} (\mathbf{s}_{m,r}^{[i]} - \delta_{\mathcal{A},\mathcal{W}_m} \delta_{q,r})}{N_q (N_r - \delta_{q,r})}. \end{aligned}$$

457 Substituting (6.8) and (6.9) into (6.7), after some rearranging yields (6.6). \square

459 **7. Applications of Theorem 6.6.** We now describe two special cases of The-
 460 orem 6.6 and illustrate its application on a bipartite network.

7.1. Recovering the infinitesimal generator. In the case where each vertex is in its own partition, we expect to recover the full infinitesimal generator \mathbf{Q} . Recall that (6.6) corresponds to a vertex v in vertex partition cell q changing from vertex-state \mathcal{A} to \mathcal{B} . For each p we have $N_p = 1$, and since we have assumed the network is simple, i.e. there are no self-loops or multiple edges, it follows that $d_q^v = 0$. We also have $d_p^v = 1$ if vertex v and the vertex in cell p are neighbours, and $d_p^v = 0$ if they are not. Since vertex v is the only vertex in cell q and it is in vertex-state \mathcal{A} , we have $\mathbf{s}_{\mathcal{A},q}^{[i]} = 1$. We also have

$$\sum_{p \neq q} \left(\sum_{v \in V_q} \frac{d_p^v}{N_p} \right) \left[\sum_{m=1}^M \zeta_m^{\mathcal{A},\mathcal{B}} \mathbf{s}_{m,p}^{[i]} \right] = \sum_{m=1}^M \zeta_m^{\mathcal{A},\mathcal{B}} n_m^v,$$

where n_m^v is the number of neighbours of v that are in vertex-state \mathcal{W}_m . This follows from the fact that the sum over p is effectively a sum over neighbours, since d_p^v is zero otherwise, and $\mathbf{s}_{m,p}^{[i]} = 1$ if the neighbouring vertex in cell p is in vertex-state \mathcal{W}_m and zero otherwise. Thus when each vertex is in its own partition, (6.6) reduces to

$$\mathbf{q}_{ij} = \zeta_0^{\mathcal{A},\mathcal{B}} + \sum_{m=1}^M \zeta_m^{\mathcal{A},\mathcal{B}} n_m^v,$$

461 i.e. our definition of \mathbf{Q}_{ij} .

7.2. Single vertex partition: population model. In the case $P = 1$, Theorem 6.6 reduces to our result published in [38]. In this case, the lumped state can be represented by a vector, i.e. the number of vertices in each of the possible vertex-states, and so we will not use a bold font to represent lumped states. In the $P = 1$

case, (6.6) becomes

$$\mathbf{q}_{ij} = s_{\mathcal{A}}^{[i]} \left\{ \zeta_0^{\mathcal{A}, \mathcal{B}} + \frac{z}{N-1} \left[\sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \left(s_m^{[i]} - \delta_{\mathcal{A}, \mathcal{W}_m} \right) \right] \right\},$$

where $z = \sum_{v \in V} d^v / N$ is the mean degree of the network and $s_m^{[i]}$ ($s_{\mathcal{A}}^{[i]}$) is the number of vertices in lumped state i that are in vertex-state \mathcal{W}_m (\mathcal{A}). This corresponds to the population model equation derived in [38].

7.3. SIS epidemic on a complete bipartite graph. For the case of SIS dynamics, the lumped states \mathbf{s} are matrices of size $2 \times P$, where $m = 1$ corresponds to susceptible vertices and $m = 2$ corresponds to infected vertices. For the SIS model, we have $\zeta_2^{1,2} = \beta$, $\zeta_0^{2,1} = \gamma$ and all other $\zeta_m^{\mathcal{A}, \mathcal{B}} = 0$. Thus if the transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to an infection in the q th partition, then

$$\mathbf{q}_{ij} = \beta \frac{s_{1,q}^{[i]}}{N_q} \left\{ \left(\sum_{v \in V_q} \frac{d_q^v}{N_q - 1} \right) s_{2,q}^{[i]} + \sum_{p \neq q} \left(\sum_{v \in V_q} \frac{d_p^v}{N_p} \right) s_{2,p}^{[i]} \right\}.$$

Similarly, if the transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to a recovery in the q th partition, then

$$(7.1) \quad \mathbf{q}_{ij} = \gamma s_{2,q}^{[i]}.$$

Note that this is the exact recovery rate for state $\mathbf{s}^{[i]}$, regardless of the choice of vertex-partition, since there are $s_{2,q}^{[i]}$ infected vertices in $\mathbf{s}^{[i]}$.

We'll now use this to write out the lumped Markov chain equations for the SIS model on a complete bipartite graph, where the vertex-partition corresponds to the bipartite partition. For $N_1 \neq N_2$, the automorphism group of a complete bipartite graph is $S_{N_1} \times S_{N_2}$, i.e. all possible pairs of permutations consisting of a permutation of the vertices in V_1 and a permutation of the vertices in V_2 . Thus the microstates in the exact lumping correspond to all possible counts of the number of infected vertices in each of the vertex partition cells, which is exactly what the approximate lumping uses. Consequently, we expect the approximate lumping formula to recover the exact lumping. We have $P = 2$, so the lumped states \mathbf{s} are two-by-two matrices. For $v \in V_1$ we have $d_1^v = 0$ and $d_2^v = N_2$; similarly for $v \in V_2$ we have $d_1^v = N_1$ and $d_2^v = 0$. Note that

$$\sum_{v \in V_q} d_q^v = 0, \quad \text{and} \quad \sum_{v \in V_q} d_p^v = N_q N_p,$$

where p is the alternative partition to q . Thus if the transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to an infection in the q th partition, then

$$\mathbf{q}_{ij} = \beta s_{1,q}^{[i]} s_{2,p}^{[i]}.$$

In state $\mathbf{s}^{[i]}$, there are $s_{1,q}^{[i]}$ susceptible vertices in V_q and each of these has $s_{2,p}^{[i]}$ infected neighbours, thus we obtain the exact lumped transition rate. The recovery rates (7.1) are also exact.

We can also write out the lumped master equation for this case. Note that for the lumped state \mathbf{s} , we have $s_{1,1} + s_{2,1} = N_1$ and $s_{1,2} + s_{2,2} = N_2$, thus we can write \mathbf{s} in terms of just two numbers, $k_1 = s_{2,1}$ and $k_2 = s_{2,2}$, so k_p is the number of infected

vertices in the p th partition. Thus we write the probability of being in the lumped state \mathbf{s} as x_{k_1, k_2} , i.e. the probability of there being k_1 infected nodes in partition cell 1 and k_2 infected nodes in partition cell 2. There are four possible transitions, corresponding to an infection or a recovery in each of the two vertex-partition cells. Consequently, if we assume that $x_{k_1, k_2} = 0$ for $k_1, k_2 < 0$ and $k_1, k_2 > N$ then summing over the four possible transitions yields

$$\begin{aligned} \dot{x}_{k_1, k_2} = & \beta(N_1 - k_1 + 1)k_2 x_{k_1-1, k_2} + \beta k_1(N_2 - k_2 + 1)x_{k_1, k_2-1} \\ & + \gamma(k_1 + 1)x_{k_1+1, k_2} + \gamma(k_2 + 1)x_{k_1, k_2+1} \\ & - [\beta(N_1 - k_1)k_2 + \beta k_1(N_2 - k_2) + \gamma(k_1 + k_2)] x_{k_1, k_2}. \end{aligned}$$

We can simplify the bipartite graph case further and assume that $N_2 = 1$ (hence $N_1 = N - 1$), which corresponds to a star graph. In this case, k_2 can only be 0 or 1, so we have

$$\begin{aligned} \dot{x}_{k_1, 0} = & \gamma(k_1 + 1)x_{k_1+1, 0} + \gamma x_{k_1, 1} - [\beta k_1 + \gamma k_1] x_{k_1, 0}, \\ \dot{x}_{k_1, 1} = & \beta(N - k_1)x_{k_1-1, 1} + \beta k_1 x_{k_1, 0} + \gamma(k_1 + 1)x_{k_1+1, 1} \\ & - [\beta(N - 1 - k_1) + \gamma(k_1 + 1)] x_{k_1, 1}. \end{aligned}$$

7.4. General procedure for applying Theorem 6.6. Now, we briefly present the wide applicability of Theorem 6.6, explaining the procedure by which it can be applied to a broad range of node dynamics and an arbitrary choice of the vertex-set partition. Concerning node dynamics, the following processes can be handled among many others.

- Beyond the widely used SIS and SIR dynamics one can use SEIR dynamic when the exposed compartment E is also taken into account, as well as all other similar variants. For example, introducing a vaccination state V and applying a contact tracing state T lead to SIVS, SITR etc. Considering the parallel propagation of two infectious diseases, more complicated models also fall within our framework.
- Information spread on networks leads to many different node-dynamics. One of them is rumour spreading when node states are ignorant (I), spreader (S) and stifler (R), which resembles SIR epidemic but where the transition from S to R depends also on the number of neighbours in the S and R states. Information spread is also modelled by using the node states: “Unknown”, the individual has not yet come into contact with the information, “Known”, the individual has received the information, but is not willing to propagate it, “Accepted”, the individual accepts the information and then propagates it, and “Exhausted”, after propagating the information to their neighbours, the individual will lose his interests in it. The concurrent propagation of two types of information is modelled by the node states S , I_1 and I_2 , where the effect of one information to the other can also be accounted for by the appropriate choice of the rate functions.
- The concurrent spread of epidemic and information can be described by the node states susceptible and aware (S_a), susceptible and not aware (S_{na}), infected and aware (I_a), infected and not aware (I_{na}). This can also be extended with a treatment class (T).
- Propagation of neuronal activity can be modelled by the node states quiescent (Q) and active (A) with both excitatory and inhibitory neurons. The effect of different neurons to each other can be described by the rate functions.

See [16, 30, 23, 36, 37] for more information about, and examples of, models within our framework. To translate such examples into our framework, it is necessary to associate the corresponding model rate constants with the set of functions that constitute the VSTM in 2.1. Typical models have far fewer vertex-state transitions and rate constants than the general case, simplifying what needs to be considered.

Concerning the network structure one can specify a vertex partition, for which we list a few possibilities below.

- If the nodes play a similar role in the network, then choosing a single partition ($P = 1$) is a reasonable choice and corresponds to the ‘well mixed’ case where network structure is essentially ignored, with the rates scaled by the network density. This case was derived in Section 7.2.
- In some cases, nodes can be divided into two groups, for example highly and weakly connected nodes, then choosing two partitions, $P = 2$, is natural.
- If the network is given by a bipartite graph, then the two node groups lead again to $P = 2$.
- The case of k -partite graphs can be handled with $P = k$ partitions. We note that in the case of complete k -partite graphs the lumping is exact.
- A natural choice for vertex partitions is based on the node degrees, i.e. two nodes are in the same vertex partition if their degree is equal.

Once the node dynamics is specified through the rate functions (2.1) and the vertex partitions are given according to Definitions 4.1 and 4.2, one can determine the generator (6.6) as follows. The macro-states $\mathbf{s}^{[i]}$ are $M \times P$ matrices defined in Definition 4.2 by specifying the number of vertices in a given state being in a given partition. The coefficients $\zeta_m^{\mathcal{A},\mathcal{B}}$ are determined by the transition functions given in (2.1). Considering only the few examples of node-dynamics and partitions listed above, several dozens of models can be derived based on our main Theorem 6.6 since each node dynamics can be combined with each partition.

Note that the generator in (6.6) is an $n \times n$ matrix, where n is given in (6.1). For example, when we have only one partition, $P = 1$ and two node-states, $M = 2$, then $n = N + 1$ which is significantly smaller than the full system size 2^N . In the case of $M = 3$ node-states we have $n = \mathcal{O}(N^2)$ which can be large but still much smaller than the full system size 3^N . In the case of $P = 2$ partitions and $M = 2$ node-states one has $n = \mathcal{O}(N^2)$, while for $M = 3$ node-states it is $n = \mathcal{O}(N^4)$. Thus the size of the lumped system is polynomial in N compared to exponential for the full system. We will get a further significant decrease in system size in the next section. Note however that using the generator (6.6) provides information about the probability distribution over macrostate-space, and in systems with absorbing states, one can estimate the probability of and time to absorption [37].

8. Large N limit of density dependent population processes. The approximate lumping process derived in Section 6 significantly reduces the number of equations that need to be considered, while introducing some approximation error. However, the number of equations will often still be very large, see the calculations at the end of the previous section. However, our approximation (6.6) is a “density-dependent population process” [11], which in the large N limit converges to a smaller system of $M \times P$ differential equations. We will now describe this in more detail.

8.1. Mean-field limit of a general density dependent process. For positive integer N , $\xi \in \mathbb{Z}^d$, $E \subset \mathbb{R}^d$, and a collection of non-negative functions $\lambda_\xi : E \rightarrow \mathbb{R}_{\geq 0}$, let

$$E_N = E \cap \{k/N \mid k \in \mathbb{Z}^d\},$$

and assume that $y \in E_N$ and $\lambda_\xi(y) > 0$ imply that $y + \xi/N \in E_N$. Then a density dependent family corresponding to λ_ξ is a sequence $\{Y_N\}$ of Markov jump processes such that Y_N has state-space E_N and transition intensities

$$q_{x,y}^{(N)} = N \left[\lambda_{N(y-x)}(x) + \mathcal{O}\left(\frac{1}{N}\right) \right], \quad x, y \in E_N.$$

Let

$$F(y) = \sum_{\xi \in \mathbb{Z}^d} \xi \lambda_\xi(y),$$

then provided that for each compact $K \subset E$,

$$\sum_{\xi \in \mathbb{Z}^d} |\xi| \sup_{y \in K} \lambda_\xi(y) < \infty,$$

and there exists an $M_K > 0$ such that

$$|F(x) - F(y)| \leq M_K |x - y|, \quad x, y \in K,$$

then in the limit $N \rightarrow \infty$ there is almost sure convergence between the Markov chain jump process $Y_N(t)$ and $y(t)$, where $y(t)$ is the solution to system of differential equations

$$\dot{y} = F(y).$$

569 A more precise statement can be found in Ethier and Kurtz's book [11].

As a simple example of a density dependent family, consider the SIS model with population N . Suppose that if there are i susceptible individuals and $j = N - i$ infected individuals then the infection rate is

$$q_{(i,j)(i-1,j+1)} = N\beta \frac{i}{N} \frac{j}{N},$$

and the recovery rate is

$$q_{(i,j)(i+1,j-1)} = N\gamma \frac{j}{N}.$$

570 This corresponds to the standard SIS stochastic compartmental model birth-death
571 process. The possible values of ξ are $(-1, +1)$ and $(+1, -1)$, so for $y = (y_1, y_2)$ we
572 have $\lambda_{(-1,+1)}(y) = \beta y_1 y_2$ and $\lambda_{(+1,-1)}(y) = \gamma y_2$. Consequently

$$573 \quad F(y) = \sum_{\xi} \xi \lambda_\xi(y) = (-\beta y_1 y_2 + \gamma y_2, +\beta y_1 y_2 - \gamma y_2). \\ 574$$

575 Thus the familiar compartmental SIS model ODEs are the large N limit of the cor-
576 responding stochastic birth-death process.

577 **8.2. Limiting equations of vertex-partition lumping.** To connect the ap-
578 proach presented in the previous subsection to vertex-partition lumping, we first in-
579 troduce some notation. Let $\mathbf{e}(\mathcal{W}_m, p) \in \{0, 1\}^{M \times P}$ be a matrix whose m, p th entry is
580 1, and all other entries are zero. Let $\xi_q^{\mathcal{A}, \mathcal{B}} = \mathbf{e}(\mathcal{B}, q) - \mathbf{e}(\mathcal{A}, q)$, so a transition from \mathbf{s}
581 to $\mathbf{s} + \xi_q^{\mathcal{A}, \mathcal{B}}$ corresponds to a vertex in partition V_q changing from vertex-state \mathcal{A} to
582 \mathcal{B} . Then we can write the transition rate (6.6) as

(8.1)

$$583 \quad \mathbf{q}_{\mathbf{s}, \mathbf{s} + \xi_q^{\mathcal{A}, \mathcal{B}}} = N \frac{\mathbf{s}_{\mathcal{A}, q}^{[i]}}{N} \left\{ \zeta_0^{\mathcal{A}, \mathcal{B}} + \frac{N}{N_q} \sum_{p=1}^P \left(\frac{\sum_{v \in V_q} d_p^v}{N_r - \delta_{q,p}} \right) \left[\sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \frac{\left(\mathbf{s}_{m,p}^{[i]} - \delta_{\mathcal{A}, \mathcal{W}_m} \delta_{q,p} \right)}{N} \right] \right\}. \quad \blacksquare \\ 584$$

585 We consider the density $\mathbf{y} = \mathbf{s}/N$ and we assume that

$$586 \quad (8.2) \quad \lim_{N \rightarrow \infty} \frac{N_p}{N} = \theta_p > 0 \quad \text{and} \quad \lim_{N \rightarrow \infty} \sum_{v \in V_q} \frac{d_p^v}{N_p} = z_{p,q} \geq 0,$$

587 with θ_p and $z_{p,q}$ finite. Note that θ_p is the fraction of vertices in partition p , so for this
 588 to be finite in the large N limit, all vertex partitions must scale with N . Also, $z_{p,q}$
 589 is the average number of edges between partitions p and q , averaged over the number
 590 of vertices in vertex partition p . We also have

$$591 \quad \sum_{v \in V_q} \frac{d_p^v}{N_q - 1} = \left(\sum_{v \in V_q} \frac{d_p^v}{N_q} \right) \frac{N_q}{N_q - 1} = \left(\sum_{v \in V_q} \frac{d_p^v}{N_q} \right) \left(1 + \frac{1}{N \left[\frac{N_q}{N} - \frac{1}{N} \right]} \right).$$

593 Thus we can write (8.1) as

$$594 \quad \mathbf{q}_{\mathbf{s}, \mathbf{s} + \xi_q^{\mathcal{A}, \mathcal{B}}} = N \left[\lambda_{\xi_q^{\mathcal{A}, \mathcal{B}}} \left(\frac{\mathbf{s}}{N} \right) + \mathcal{O} \left(\frac{1}{N} \right) \right],$$

595 where

$$596 \quad (8.3) \quad \lambda_{\xi_q^{\mathcal{A}, \mathcal{B}}}(\mathbf{y}) = \mathbf{y}_{\mathcal{A}, q} \left(\zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{p=1}^P \frac{z_{p,q}}{\theta_q} \sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{y}_{m,p} \right).$$

597 Note that $\lambda_{\xi_q^{\mathcal{A}, \mathcal{B}}}$ returns a scalar but $\xi_q^{\mathcal{A}, \mathcal{B}}$ is a matrix and it identifies the $\mathcal{A}, \mathcal{B} \in \mathcal{W}$
 598 and $0 \leq q \leq P$ used in (8.3). In the large N limit, we have a system of matrix
 599 differential equations for \mathbf{y} , given by

$$600 \quad (8.4) \quad \dot{\mathbf{y}} = \sum_{\mathcal{A} \in \mathcal{W}} \sum_{\mathcal{B} \neq \mathcal{A}} \sum_{q=1}^P \xi_q^{\mathcal{A}, \mathcal{B}} \lambda_{\xi_q^{\mathcal{A}, \mathcal{B}}}(\mathbf{y}).$$

601 **8.3. Application of the mean-field-limit approach.** Here we present the
 602 applicability of the mean-field limit derivation shown in the previous subsection. Using
 603 our approach one can derive the mean-field limit equation (8.4) for each node-dynamic
 604 and vertex partition listed in Section 7.4. The model dynamics specify the transitions
 605 $\xi_q^{\mathcal{A}, \mathcal{B}}$ and the corresponding transition rates $\zeta_m^{\mathcal{A}, \mathcal{B}}$. The vertex-set partition can also
 606 be chosen quite generally, subject to the conditions (8.2), which assume that as N
 607 tends to infinity, the proportion of vertices in each partition, θ_p , and the proportion
 608 of edges between partitions, $z_{p,q}$, tend to constant values. Once these constants have
 609 been determined, one can formulate transition functions $\lambda_{\xi_q^{\mathcal{A}, \mathcal{B}}}$ for each transition
 610 $\xi_q^{\mathcal{A}, \mathcal{B}}$ using (8.3). The dependent variable in the mean-field equations (8.4) is the time-
 611 dependent matrix \mathbf{y} , which is the scaled version of the lumped variables, i.e. $\mathbf{y} = \mathbf{s}/N$,
 612 meaning that $\mathbf{y}_{m,p}$ is the proportion of vertices in vertex-state \mathcal{W}_m and in partition
 613 cell V_p . Thus the number of differential equations in (8.4) is $M \times P$. Considering
 614 only the few examples of node-dynamics and partitions listed in Section 7.4, several
 615 dozens of mean-field equations can be derived based on our approach since each node
 616 dynamic can be combined with each partition.

9. Degree-based mean-field. We now relate the large N limit of our vertex-
 partition lumping to the well known degree-based mean-field. Consider the vertex

partition where for each $k > 0$, we have that V_k is the set of vertices with degree k . Thus N_k is the number of vertices with degree k and $p_k = N_k/N$ is the fraction of vertices with degree k , i.e. the degree distribution. We assume that $p_k > 0$ for all k (although the same approach could be applied to cases where there are k such that $p_k = 0$) and that there is a maximum degree k_{\max} , so that the vertex-set is partitioned into a finite number of cells. We will call a partition based on vertex degree “degree-based mean-field” and we will derive the ODEs for the SIS model. Let \mathcal{W}_1 correspond to susceptible nodes and \mathcal{W}_2 correspond to infected nodes. Recall that for the SIS model we have $\zeta_2^{1,2} = \beta$ and $\zeta_0^{2,1} = \gamma$, and all other $\zeta_m^{A,B} = 0$. Since $\mathbf{y}_{1,k} + \mathbf{y}_{2,k} = p_k$, we will refer to the fraction of infected nodes as $y_k = \mathbf{y}_{2,k}$, from which we can infer $\mathbf{y}_{1,k} = p_k - y_k$. Thus we only need to write the differential equations for y_k , and we will write $y = (y_1, y_2, \dots, y_{k_{\max}})$. Consequently we only need the second column of the matrices $\xi_q^{A,B}$. Hence let ξ_k be a vector of zeros with a one in the k th entry, then ξ_k corresponds to an infection and $-\xi_k$ corresponds to a recovery. Using (8.3), the large N infection rate is,

$$\lambda_{\xi_k}(y) = \beta(p_k - y_k) \sum_{k'} \frac{z_{k',k}}{p_k} y_{k'},$$

the large N recovery rate is

$$\lambda_{-\xi_k}(y) = \gamma y_k,$$

617 and hence the evolution equations are

$$618 \quad (9.1) \quad \dot{y}_k = -\gamma y_k + \beta(p_k - y_k) \sum_{k'} \frac{z_{k',k}}{p_k} y_{k'}.$$

We will now show that (9.1) is equivalent to the Eames and Keeling [10] and Pastor-Satorras and Vespignani [28] degree-based mean-field approximations. Eames and Keeling use $[I^k]$ to denote the number of infected nodes with k neighbours and $[S^k I^{k'}]$ to denote the number of partnerships between a susceptible node with k partners and an infected node with k' partners. The dynamics is then described by

$$\frac{d[I^k]}{dt} = -\gamma[I^k] + \beta \sum_{k'} [S^k I^{k'}],$$

and this is closed with the approximation

$$[S^k I^{k'}] \approx \frac{[S^k]}{N_k} \times \frac{[I^{k'}]}{N_{k'}} \times [kk'],$$

619 where $[kk']$ is the number of partnerships between individuals with k and k' partners.
620 Thus the Eames and Keeling degree-based mean-field is

$$621 \quad (9.2) \quad \frac{d[I^k]}{dt} = -\gamma[I^k] + \beta \sum_{k'} \frac{[S^k]}{N_k} \frac{[I^{k'}]}{N_{k'}} [kk'].$$

Let $y_k = [I^k]/N$, and noting that since $[S^k] + [I^k] = N_k$ we have $[S^k]/N = p_k - y_k$, then dividing (9.2) through by N yields

$$\dot{y}_k = -\gamma y_k + \beta(p_k - y_k) \sum_{k'} \frac{[kk']}{p_k N_{k'}} y_{k'}.$$

Since $[kk']$ is the number of edges between degree k and degree k' vertices, we have

$$[kk'] = \sum_{v \in V_k} d_{k'}^v,$$

and consequently, using (8.2), the Eames and Keeling degree-based mean-field is equivalent to (9.1). It is also easy to show that the Eames and Keeling degree-based mean-field is equivalent to the Pastor-Satorras and Vespignani [28] degree-based mean-field, and we provide details of this in the Supplementary Information.

10. Individual-based mean-field. We will now show how our approach can be used to derive individual-based mean-field approximations. Consider a graph with N vertices and N_e isomorphic copies of this graph. We will call the graph that is copied the base graph, and thus the collection of all copies of the base graph consists of NN_e vertices. We will apply approximate lumping to this collection of graphs by choosing the partition of vertices, Π_V , so that each partition cell V_i has the corresponding vertex in each of the N_e copies of the graph. Thus there are $P = N$ partition cells, and we can increase the total number of vertices by increasing N_e , the number of copies of the base graph. Applying our approximate lumping approach produces an ensemble average over the collection of isomorphic base graphs.

To apply our approximate lumping approach, we need to compute the fraction of vertices in each vertex partition cell, θ_i , and the mean number of edges between vertex partition cells i and j , $z_{i,j}$, as defined in (8.2). Note that our notation is slightly modified here, where we are taking the limit $N_e \rightarrow \infty$ (and hence the total number of vertices in the collection of graphs) rather than N . Since there are N_e vertices in each partition cell and NN_e vertices in total, we have that $\theta_i = 1/N$. Let \mathbf{A} be the adjacency matrix of the base graph, so the component \mathbf{A}_{ij} is one if vertices i and j are connected in the base graph and zero otherwise. Thus if $V_i \in \Pi_V$, then the number of neighbours of vertex $v \in V_i$ that are in the partition cell V_j is $d_j^v = \mathbf{A}_{ij}$ (i.e. one if i and j are connected in the base graph, and zero otherwise). We will assume for $v \in V_i$ that $d_i^v = 0$, i.e. there are no self-loops in the base graph. Consequently, since $N_i = N_e$ for all $V_i \in \Pi_V$, we have

$$z_{i,j} = \lim_{N_e \rightarrow \infty} \sum_{v \in V_i} \frac{d_j^v}{N_e} = \mathbf{A}_{ij}.$$

We are now in a position to make use of (8.3), but note that following the approach described in Section 8, we would use the density variable $\mathbf{y} = \mathbf{s}/(NN_e)$, i.e. the total fraction of vertices in each vertex-state and vertex partition. However, individual-based mean-field approximations are typically based on the ‘probability’ that a vertex is in a given vertex-state. We can obtain a similar quantity here by instead using $\mathbf{y} = \mathbf{s}/N_e$, i.e. the fraction of base graphs in which each vertex is in each vertex-state. We again let ξ correspond to a vertex in V_i changing from \mathcal{A} to \mathcal{B} and so (8.1) becomes

$$\mathbf{q}_{\mathbf{s}, \mathbf{s} + \xi_i^{\mathcal{A}, \mathcal{B}}} = N_e \frac{\mathbf{s}_{\mathcal{A}, i}}{N_e} \left(\zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{j \neq i} \mathbf{A}_{ij} \left[\sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \frac{\mathbf{s}_{m, j}}{N_e} \right] + \mathcal{O}(1/N_e) \right).$$

Consequently, using $\mathbf{y} = \mathbf{s}/N_e$, the large N_e transition rate is

$$\lambda_{\xi_i^{\mathcal{A}, \mathcal{B}}}(\mathbf{y}) = \mathbf{y}_{\mathcal{A}, i} \left(\zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{j \neq i} \mathbf{A}_{ij} \left[\sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{y}_{m, j} \right] \right).$$

From this we recognise the transition rates that appear in standard individual-based mean-field equations, in which pairwise interaction terms are included according to the graph's adjacency matrix. For example, for the SIS model the evolution equation for the 'probability' y_i that vertex i is infected is [34, 23]

$$\dot{y}_i = -\gamma y_i + \beta(1 - y_i) \sum_{j \neq i} \mathbf{A}_{ij} y_j.$$

11. The configuration model. In the previous section, we saw how our approximate lumping approach could be applied to isomorphic copies of a given graph to obtain individual-based mean-field approximations. In this section we consider how this approach can be extended to derive mean-field approximations for dynamics on families of graphs, specifically configuration models.

Suppose that the vertices V of a graph are labelled $1, 2, \dots, N$, then a degree sequence is a sequence of integers d_1, d_2, \dots, d_N , such that $\sum_i d_i = 2M$ and for each i , $d_i \leq N - 1$. A random graph can be constructed from a given degree sequence by allocating each node i with d_i 'stubs' and then picking pairs of stubs at random without replacement from the collection of all unpaired stubs to form edges in the graph. Such a graph may have self-edges and multi-edges. While the theory we have developed assumed simple networks, we conjecture that extending this approach to graphs with self- and multi-edges only introduces $\mathcal{O}(1/N)$ corrections to (6.6) in the large N limit. Here we consider the family of all such graphs for a given degree sequence, and we will refer to these as configurations (of the associated degree sequence). The number of configurations can be computed fairly easily by considering the combinations of ways that pairs of stubs can be chosen, while accounting for the fact that the order in which pairs of stubs are selected is not important, which yields

$$N_{\text{CM}}(M) = \frac{1}{M!} \binom{2M}{2} \binom{2M-2}{2} \cdots \binom{2}{2} = \frac{2M!}{2^M M!}.$$

Thus there are $NN_{\text{CM}}(M)$ vertices in the collection of configurations.

It is also easy to show that for a pair of vertices i, j , the average number of edges between i and j across all configurations is

$$\frac{d_i d_j}{2M - 1}.$$

To see this, note that there are $d_i d_j$ ways to match each of the d_i stubs to each of the d_j stubs. Having used a pair of stubs to do this, there are then $2M - 2$ stubs that remain to be matched, and there are $N_{\text{CM}}(M - 1)$ ways to do this. Note that some of these remaining pairs of stubs may also lead to edges between i and j , but this exactly accounts for all possible multi-edges between i and j . Thus

$$d_i d_j \frac{N_{\text{CM}}(M - 1)}{N_{\text{CM}}(M)},$$

and cancellation leads to the result.

To apply the vertex-partition lumping to the configuration model, we need to consider the large N limit. Thus we suppose that for a given N , the degree sequence is sampled at random and the degrees of the N vertices are $d_1^N, d_2^N, \dots, d_N^N$, where there is an M_N such that $\sum_i d_i^N = 2M_N$. Furthermore, we assume that for each $0 < k \leq k_{\text{max}}$, the number of vertices of degree k is n_k and

$$\lim_{N \rightarrow \infty} \frac{n_k}{N} = p_k,$$

where $p_k > 0$ for each k and $\sum_k p_k = 1$. Thus in the large N limit, the mean degree of the network will converge to $z = \sum_k k p_k$.

We again consider a vertex-partition based on degree for the family of configurations of N vertices. Thus for finite k_{\max} the vertex partition is $\Pi_V = \{V_1, V_2, \dots, V_{k_{\max}}\}$. As described already, there are $NN_{\text{CM}}(M_N)$ vertices across the family of configurations and hence there are $N_k = |V_k| = n_k N_{\text{CM}}(M_N)$ vertices of degree k . In order to apply our general formula (8.3), it remains to compute $z_{k',k}$, the average number of edges between vertices of degree k and vertices of degree k' , with respect to the number of vertices of degree k' . We have seen that the average number of edges between nodes with degree k and k' is $kk'/(2M_N - 1)$, so the total number of edges between nodes of degree k and k' across all configurations is this multiplied by the total number of vertices of degree k and the number of vertices of degree k' in a single configuration. Thus it follows that

$$(11.1) \quad z_{k',k} = \lim_{N \rightarrow \infty} \sum_{v \in V_k} \frac{d_{k'}^v}{N_{k'}} = \lim_{N \rightarrow \infty} \frac{kk'}{2M_N - 1} \frac{N_k n_{k'}}{N_{k'}} = \frac{kk' p_k}{z}.$$

Focusing on the SIS model and using the notation from Section 9, we can substitute (11.1) into (9.1), which yields

$$\dot{y}_k = -\gamma y_k + \beta(p_k - y_k) \sum_{k'} \frac{kk'}{z} y_{k'}.$$

Furthermore, substituting $y_k = \rho_k p_k$, rearranging and rescaling recovers the uncorrelated degree-based mean-field [28].

12. Discussion. In this paper we have derived mean-field approximations for a broad class of dynamical processes on networks directly from their exact Markov chain description. We have done this using the method of approximate lumping, where macrostates are defined in terms of the number of vertices in each vertex-state in subsets of vertices that form a partition of the vertex set. We have proved that this approach results in a density dependent population process, from which the large N limiting behaviour can be described in terms of a relatively small number of equations, specifically $M \times P$ equations, where M is the number of vertex states and P is the size of the vertex partition. We have shown how this approach can be used to derive degree-based and individual-based mean-field approximations.

Given how involved the direct calculations are, it is surprising and impressive that we recover exactly the well-known degree-based and individual-based mean-field approximations. However, we emphasise that the use of approximate lumping means that not only is the averaging process clear, it tells us under what circumstances the approximation would be exact, i.e. the Markov chain that corresponds to the density dependent population process. Our approach also highlights that there are two sources of error. The main, and uncontrolled, source of error arises from the choice of vertex-partition that defines the partition of microstate-space in the approximate lumping. The second arises from the large N limit, but this vanishes as N becomes large.

Our methodology formalises the process of obtaining mean-field equations for a given dynamical model, eliminating the need to rely on intuitive probabilistic arguments. This approach also explicitly reveals the type of averaging represented by mean-field approximations on networks. Our results apply in a broad range of cases, allowing researchers working with new models—within the class of single-vertex tran-

sition models with local, affine transition functions considered here—to easily obtain a range of mean-field approximations, in a controlled manner.

The mean-field approximations that we have derived using our approximate lumping approach ignore dynamical correlations between the states of neighbouring vertices. Such correlations are taken into account in higher-order, edge-based mean-field approximations [23], that are typically derived using moment closure arguments. In a follow-up paper we will show how our approach can also be adapted to derive such edge-based mean-field approximations, extending its generality. A high-accuracy form of mean-field approximation are “approximate master equations” [15, 16], which are based on the number of susceptible/infected nodes of degree k with m infected neighbours. While it may seem that these might be derived from our approach, we have been unable to do so because it is not clear how the counts change when vertices recover or become infected. In particular, it seems that one must know the degree and number of infected neighbours of the neighbours of a vertex that changes vertex-state.

Our ambition is to quantify mean-field approximation error in terms of the dynamics and network structure. Our approach in this paper highlights that the main source of error results from the choice of vertex-partition and how far the corresponding lumping is from being exact. Establishing quantitative estimates for the resulting error is a critical open challenge. An attempt to do this is described in [38] for the case of the SISa model and where vertices are not partitioned, but in this case the error could not be entirely unravelled from the full Markov chain. However, it may be that hierarchies of approximations can be constructed in which the error decays monotonically [21]. In this paper we have shown how our approach can be applied to families of graphs, and this suggests that it may be also possible to use our approach to derive mean-field approximations for models in which the network co-evolves with the vertex-state dynamic. We would also like to generalise our approach to models with nonlinear VSTMs.

REFERENCES

- [1] A. BARRAT, M. BARTHELEMY, AND A. VESPIGNANI, *Dynamical processes on complex networks*, Cambridge University Press, 2008.
- [2] M. BOGUNÁ, C. CASTELLANO, AND R. PASTOR-SATORRAS, *Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks*, Physical Review Letters, 111 (2013), p. 068701.
- [3] P. BUCHHOLZ, *Exact and ordinary lumpability in finite markov chains*, Journal of Applied Probability, 31 (1994), pp. 59–75.
- [4] C. CASTELLANO, M. A. MUÑOZ, AND R. PASTOR-SATORRAS, *Nonlinear q-voter model*, Physical Review E, 80 (2009), p. 041129.
- [5] S. CHATTERJEE AND R. DURRETT, *Contact processes on random graphs with power law degree distributions have critical value 0*, The Annals of Probability, 37 (2009), pp. 2332–2356.
- [6] A. CRISANTI AND H. SOMPOLINSKY, *Path integral approach to random neural networks*, Physical Review E, 98 (2018), p. 062120.
- [7] E. DE GIULI AND C. SCALLIET, *Dynamical mean-field theory: from ecosystems to reaction networks*, Journal of Physics A: Mathematical and Theoretical, 55 (2022), p. 474002.
- [8] S. N. DOROGVTSEV, A. V. GOLTSEV, AND J. F. MENDES, *Critical phenomena in complex networks*, Reviews of Modern Physics, 80 (2008), pp. 1275–1335.
- [9] S. N. DOROGVTSEV, A. V. GOLTSEV, AND J. F. F. MENDES, *Ising model on networks with an arbitrary distribution of connections*, Physical Review E, 66 (2002), p. 016104.
- [10] K. T. EAMES AND M. J. KEELING, *Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 13330–13335.
- [11] S. N. ETHIER AND T. G. KURTZ, *Markov processes: characterization and convergence*, vol. 282, John Wiley & Sons, 2009.
- [12] P. G. FENNELL AND J. P. GLEESON, *Multistate dynamical processes on networks: analysis*

- through degree-based approximation frameworks, *SIAM Review*, 61 (2019), pp. 92–118.
- [13] S. GALAM, *Minority opinion spreading in random geometry*, *The European Physical Journal B*, 25 (2002), pp. 403–406.
- [14] R. J. GLAUBER, *Time-dependent statistics of the ising model*, *Journal of Mathematical Physics*, 4 (1963), pp. 294–307.
- [15] J. P. GLEESON, *High-accuracy approximation of binary-state dynamics on networks*, *Physical Review Letters*, 107 (2011), p. 68701.
- [16] J. P. GLEESON, *Binary-state dynamics on complex networks: Pair approximation and beyond*, *Physical Review X*, 3 (2013), p. 021004.
- [17] J. P. GLEESON, S. MELNIK, J. A. WARD, M. A. PORTER, AND P. J. MUCHA, *Accuracy of mean-field theory for dynamics on real-world networks*, *Physical Review E*, 85 (2012), p. 026106.
- [18] J. GOMEZ-GARDENES, V. LATORA, Y. MORENO, AND E. PROFUMO, *Spreading of sexually transmitted diseases in heterosexual populations*, *Proceedings of the National Academy of Sciences*, 105 (2008), pp. 1399–1404.
- [19] J. A. HERTZ, Y. ROUDI, AND P. SOLLICH, *Path integral methods for the dynamics of stochastic and disordered systems*, *Journal of Physics A: Mathematical and Theoretical*, 50 (2016), p. 033001.
- [20] J. G. KEMENY, J. L. SNELL, ET AL., *Finite Markov Chains*, vol. 356, van Nostrand Princeton, NJ, 1960.
- [21] W. R. KHUDABUKHSH, A. AUDDY, Y. DISSER, AND H. KOEPL, *Approximate lumpability for Markovian agent-based models using local symmetries*, *Journal of Applied Probability*, 56 (2019), pp. 647–671.
- [22] M. KIJIMA, *Markov processes for stochastic modeling*, vol. 6, CRC Press, 1997.
- [23] I. Z. KISS, J. C. MILLER, AND P. L. SIMON, *Mathematics of Epidemics on Networks*, Springer, 2017.
- [24] M. LEONE, A. VÁZQUEZ, A. VESPIGNANI, AND R. ZECCHINA, *Ferromagnetic ordering in graphs with arbitrary degree distribution*, *The European Physical Journal B-Condensed Matter and Complex Systems*, 28 (2002), pp. 191–197.
- [25] F. L. METZ, *Dynamical mean-field theory of complex systems on sparse directed networks*, *Physical Review Letters*, 134 (2025), p. 037401.
- [26] M. NEWMAN, *Networks*, Oxford University Press, 2018.
- [27] R. PASTOR-SATORRAS, C. CASTELLANO, P. VAN MIEGHEM, AND A. VESPIGNANI, *Epidemic processes in complex networks*, *Reviews of Modern Physics*, 87 (2015), p. 925.
- [28] R. PASTOR-SATORRAS AND A. VESPIGNANI, *Epidemic spreading in scale-free networks*, *Physical Review Letters*, 86 (2001), p. 3200.
- [29] L. PELLIS, F. BALL, S. BANSAL, K. EAMES, T. HOUSE, V. ISHAM, AND P. TRAPMAN, *Eight challenges for network epidemic models*, *Epidemics*, 10 (2015), pp. 58–62.
- [30] M. A. PORTER AND J. P. GLEESON, *Dynamical systems on networks*, *Frontiers in Applied Dynamical Systems: Reviews and Tutorials*, 4 (2016), p. 29.
- [31] P. L. SIMON, M. TAYLOR, AND I. Z. KISS, *Exact epidemic models on graphs using graph-automorphism driven lumping*, *Journal of Mathematical Biology*, 62 (2011), pp. 479–508.
- [32] V. SOOD AND S. REDNER, *Voter model on heterogeneous graphs*, *Physical Review Letters*, 94 (2005), p. 178701.
- [33] K. SZNAJD-WERON AND J. SZNAJD, *Opinion evolution in closed community*, *International Journal of Modern Physics C*, 11 (2000), pp. 1157–1165.
- [34] P. VAN MIEGHEM, *The N-intertwined SIS epidemic network model*, *Computing*, 93 (2011), pp. 147–169.
- [35] F. VÁZQUEZ AND V. M. EGUÍLIZ, *Analytical solution of the voter model on uncorrelated networks*, *New Journal of Physics*, 10 (2008), p. 063011.
- [36] J. A. WARD AND J. EVANS, *A general model of dynamics on networks with graph automorphism lumping*, in *International Conference on Complex Networks and their Applications*, Springer, 2018, pp. 445–456.
- [37] J. A. WARD AND M. LÓPEZ-GARCÍA, *Exact analysis of summary statistics for continuous-time discrete-state markov processes on networks using graph-automorphism lumping*, *Applied Network Science*, 4 (2019), p. 108.
- [38] J. A. WARD, A. TAPPER, P. L. SIMON, AND R. P. MANN, *Micro-scale foundation with error quantification for the approximation of dynamics on networks*, *Communications Physics*, 5 (2022), p. 71.
- [39] D. J. WATTS, *A simple model of global cascades on random networks*, *Proceedings of the National Academy of Sciences*, 99 (2002), pp. 5766–5771.

SUPPLEMENTARY MATERIALS: MEAN-FIELD APPROXIMATION OF DYNAMICS ON NETWORKS*

JONATHAN A. WARD[†], GÁBOR TIMÁR[†], AND PÉTER L. SIMON[‡]

SM1. Introduction. Sections SM2 to SM5 of this document are intended to supplement Sections 2 to 5 in the main paper with an illustrative example. The sections here could be read alongside those in the main paper or separately, although the general theory presented here is highly abbreviated.

SM2. Mathematical background. Let $G = (V, E)$ denote a network with vertex set V and edge set $E \subset V \times V$, where the number of vertices is $N = |V|$. Unless otherwise stated, we consider dynamical processes on finite simple networks (i.e. undirected, unweighted with no self-loops or multiple edges) described by continuous-time Markov chains where each vertex can be in one of a finite number M of *vertex-states* and the set of possible vertex-states is $\mathcal{W} = \{\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_M\}$.

SM2.1. State-space. The state-space of the Markov chain is the set of all permutations of N vertex-states chosen from \mathcal{W} with repetition. This is equivalent to $\mathcal{S} = \mathcal{W}^V$, i.e. the set of all functions from V to \mathcal{W} , and so if the network is in state $S \in \mathcal{S}$ then the vertex-state of vertex $v \in V$ is $S(v)$. We refer to the states $S \in \mathcal{S}$ as *microstates* and the number of microstates in \mathcal{S} is M^N . We enumerate the microstates as $\mathcal{S} = \{S^{[1]}, S^{[2]}, \dots, S^{[M^N]}\}$.

As an illustration of the theory developed in the main paper, a simple example will be presented here, specifically for dynamics on a square/two-by-two lattice/four cycle with vertices $V = \{1, 2, 3, 4\}$, illustrated in Figure SM1. We will refer to this as the square example. In this example, we consider SISa dynamics in which vertices are either susceptible, which corresponds to $\mathcal{W}_1 = \mathcal{B}$ and the colour blue, or infected, which corresponds to $\mathcal{W}_2 = \mathcal{R}$ and the colour red. The total number of microstates is $M^N = 2^4 = 16$, we number them as illustrated in Figure SM2.

SM2.2. Transitions. We assume that changes in microstate correspond to a single vertex $v \in V$ changing its vertex-state, and the rate that this occurs is a function of only the number of v 's neighbours in each of the vertex-states. We also assume that this rate function is the same for all vertices. In the models we consider, $\mathbf{R}_{\mathcal{A}, \mathcal{B}}(n_1, n_2, \dots, n_M) \geq 0$ gives the rate that a vertex in vertex-state \mathcal{A} changes to vertex-state \mathcal{B} if it has n_1 neighbours in vertex-state \mathcal{W}_1 , n_2 neighbours in vertex-state \mathcal{W}_2 , etc. If transitions between a pair of vertex-states $\mathcal{A}, \mathcal{B} \in \mathcal{W}$ do not occur in a particular model, then the rate $\mathbf{R}_{\mathcal{A}, \mathcal{B}}$ is identically zero. We consider models where $\mathbf{R}_{\mathcal{A}, \mathcal{B}}$ is an affine function, i.e.

$$(SM2.1) \quad \mathbf{R}_{\mathcal{A}, \mathcal{B}}(n_1, n_2, \dots, n_M) = \zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{m=1}^M \zeta_m^{\mathcal{A}, \mathcal{B}} n_m,$$

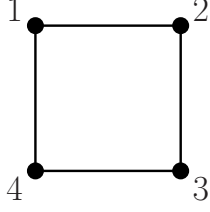
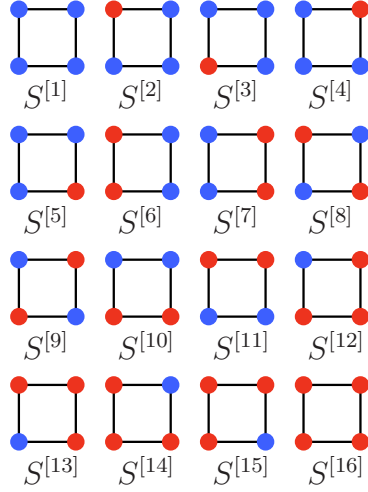
To illustrate the transitions between microstates, consider the example of SISa dynamics. A susceptible vertex v with n_1 susceptible neighbours and n_2 infected

*Submitted to the editors DATE.

Funding: J.A. Ward and G. Timár acknowledge funding from the Leverhulme Trust Project Grant number RPG-2023-187. P.L. Simon acknowledges support from the Hungarian Scientific Research Fund, OTKA (grant no. 135241) and from ERC Synergy Grant No. 810115 - DYNASNET.

[†]University of Leeds, UK (j.a.ward@leeds.ac.uk).

[‡]Eotvos Lorand University, Hungary

FIG. SM1. *Four cycle with vertex labelling.*FIG. SM2. *Labelling of microstate-space, blue corresponds to susceptible and red to infected vertices.*

neighbours becomes infected at a rate $\alpha + \beta n_2$, where $\alpha, \beta > 0$, so

$$\mathbf{R}_{\mathcal{B},\mathcal{R}}(n_1, n_2) = \alpha + \beta n_2.$$

Thus $\zeta_0^{\mathcal{B},\mathcal{R}} = \alpha$, $\zeta_1^{\mathcal{B},\mathcal{R}} = 0$ and $\zeta_2^{\mathcal{B},\mathcal{R}} = \beta$. An infected vertex with n_1 susceptible neighbours and n_2 infected neighbours becomes susceptible at a rate $\gamma > 0$, so

$$\mathbf{R}_{\mathcal{R},\mathcal{B}}(n_1, n_2) = \gamma.$$

37 Thus $\zeta_0^{\mathcal{R},\mathcal{B}} = \gamma$ and $\zeta_1^{\mathcal{R},\mathcal{B}} = \zeta_2^{\mathcal{R},\mathcal{B}} = 0$.

SM2.3. Kolmogorov equations: infinitesimal generator. Let

$$X(t) = (X_1(t), X_2(t), \dots, X_{M^N}(t))^T$$

be the time-dependent Markov chain probability distribution over \mathcal{S} , where $X_i(t)$ is the probability of being in microstate $S^{[i]}$ at time t . The evolution of $X(t)$ is then given by the forward Kolmogorov or master equation [SM4],

$$\dot{X} = \mathbf{Q}^T X,$$

where \mathbf{Q} is the infinitesimal generator, an M^N by M^N matrix in which each off-diagonal component \mathbf{Q}_{kl} gives the transition rate from $S^{[k]}$ to $S^{[l]}$, and the diagonal components ensure that rows sum to zero so that probability is conserved. We assume that a vertex changes vertex-state instantaneously, thus transitions only occur between pairs of microstates that differ in exactly one vertex-state.

Considering the SISa dynamics on the square again, we can easily determine the entries of the matrix \mathbf{Q} . For example, $\mathbf{Q}_{9,12} = \alpha + 2\beta$, because the transition from $S^{[9]}$ to $S^{[12]}$ means that node 3 becomes infected. (See the numbering of nodes and states in Figures SM1 and SM2.) This can happen at rate α spontaneously and at rate 2β by infection along edges from nodes 2 and 4. Calculating the rates for all possible transition pairs, the infinitesimal generator of the small example is a 16×16 matrix that can be written as

$$(SM2.2) \quad \mathbf{Q} = \begin{pmatrix} \Delta_0 & \mathbf{A}_0 & 0 & 0 & 0 \\ \mathbf{B}_1 & \Delta_1 & \mathbf{A}_1 & 0 & 0 \\ 0 & \mathbf{B}_2 & \Delta_2 & \mathbf{A}_2 & 0 \\ 0 & 0 & \mathbf{B}_3 & \Delta_3 & \mathbf{A}_3 \\ 0 & 0 & 0 & \mathbf{B}_4 & \Delta_4 \end{pmatrix}$$

where the matrix \mathbf{A}_k corresponds to infection from states with k infected nodes, matrix \mathbf{B}_k corresponds to recovery from states with k infected nodes and the matrix Δ_k is diagonal with negative entries in the diagonal, determined in such a way that the sum of entries in a row of \mathbf{Q} is zero. These matrices take the following form:

$$\mathbf{A}_0 = \begin{pmatrix} \alpha & \alpha & \alpha & \alpha \end{pmatrix}, \quad \mathbf{A}_1 = \begin{pmatrix} \alpha + \beta & 0 & \alpha & 0 & 0 & \alpha + \beta \\ \alpha + \beta & 0 & 0 & \alpha & \alpha + \beta & 0 \\ 0 & \alpha + \beta & 0 & \alpha & 0 & \alpha + \beta \\ 0 & \alpha + \beta & \alpha & 0 & \alpha + \beta & 0 \end{pmatrix},$$

$$\mathbf{A}_2 = \begin{pmatrix} 0 & 0 & \alpha + \beta & \alpha + \beta \\ \alpha + \beta & \alpha + \beta & 0 & 0 \\ 0 & \alpha + 2\beta & \alpha + 2\beta & 0 \\ \alpha + 2\beta & 0 & 0 & \alpha + 2\beta \\ \alpha + \beta & 0 & \alpha + \beta & 0 \\ 0 & \alpha + \beta & 0 & \alpha + \beta \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} \alpha + 2\beta \\ \alpha + 2\beta \\ \alpha + 2\beta \\ \alpha + 2\beta \end{pmatrix},$$

$$\mathbf{B}_1 = \begin{pmatrix} \gamma \\ \gamma \\ \gamma \\ \gamma \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} \gamma & \gamma & 0 & 0 \\ 0 & 0 & \gamma & \gamma \\ \gamma & 0 & 0 & \gamma \\ 0 & \gamma & \gamma & 0 \\ 0 & \gamma & 0 & \gamma \\ \gamma & 0 & \gamma & 0 \end{pmatrix}, \quad \mathbf{B}_3 = \begin{pmatrix} 0 & \gamma & 0 & \gamma & \gamma & 0 \\ 0 & \gamma & \gamma & 0 & 0 & \gamma \\ \gamma & 0 & \gamma & 0 & \gamma & 0 \\ \gamma & 0 & 0 & \gamma & 0 & \gamma \end{pmatrix},$$

$$\mathbf{B}_4 = \begin{pmatrix} \gamma & \gamma & \gamma & \gamma \end{pmatrix}.$$

SM3. Coarse-graining via lumping: theoretical foundation. We consider *lumping* of Markov chains [SM3]. Let $\Pi_{\mathcal{S}} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n\}$ be a partition of microstate-space, so $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset$ for each $i \neq j$, and $\cup_i \mathcal{S}_i = \mathcal{S}$. We will refer to \mathcal{S}_i as a *lumped state*. An *exact lumping* is a partition of microstate-space $\Pi_{\mathcal{S}}$ that preserves the Markov property, a necessary and sufficient condition for which is that the sum of transition rates out of a microstate $S^{[k]} \in \mathcal{S}_i$ into the cell \mathcal{S}_j is the same

for all microstates in the cell \mathcal{S}_i . In matrix notation [SM6], this is equivalent to the existence of an $n \times n$ matrix \mathbf{q} such that

$$(SM3.1) \quad \mathbf{Q}\mathbf{C} = \mathbf{C}\mathbf{q},$$

where $\mathbf{C} \in \{0, 1\}^{M^N \times n}$ is the *collector matrix* [SM2] whose kj th component is

$$(SM3.2) \quad \mathbf{C}_{kj} = \begin{cases} 1 & \text{if } S^{[k]} \in \mathcal{S}_j, \\ 0 & \text{otherwise,} \end{cases}$$

that is the collector matrix collects those microstates in a column that belong to the same cell, or in other words, the same macro-state, in the partition. We call (SM3.1) the *lumpability condition*.

Note that \mathbf{q} can be given explicitly for an exact lumping by introducing the *distributor matrix* [SM2] $\mathbf{D} \in \mathbb{R}^{n \times M^N}$, whose il th component is

$$(SM3.3) \quad \mathbf{D}_{il} = \begin{cases} \frac{1}{|\mathcal{S}_i|} & \text{if } S^{[l]} \in \mathcal{S}_i, \\ 0 & \text{otherwise.} \end{cases}$$

Specifically, $\Pi_{\mathcal{S}}$ satisfies the lumpability condition when \mathbf{Q} commutes with $\mathbf{C}\mathbf{D}$ [SM6]. Note that $\mathbf{D}\mathbf{C} = \mathbf{I}$, the identity matrix, hence multiplying (SM3.1) by \mathbf{D} we get the generator \mathbf{q} of the lumped system explicitly as

$$(SM3.4) \quad \mathbf{q} = \mathbf{D}\mathbf{Q}\mathbf{C}.$$

A lumping that does not satisfy the lumpability condition is an *approximate lumping* [SM2]. Given a partition $\Pi_{\mathcal{S}}$ of microstate-space that does not satisfy the lumpability condition (SM3.1), our approach is to still use the set of lumped states $\Pi_{\mathcal{S}}$ and the corresponding generator $\mathbf{q} = \mathbf{D}\mathbf{Q}\mathbf{C}$. Summarising, we can say that starting from the full infinitesimal generator, \mathbf{Q} , and choosing a partition of the state-space, equation (??) yields the infinitesimal generator \mathbf{q} of the lumped (coarse-grained) system. Note that the partition of the state-space is encoded in the collector and distributor matrices, \mathbf{C} and \mathbf{D} respectively.

SM4. Lumping based on vertex set partitions. In the previous section we introduced the notion of lumping in general. In this section we illustrate how a partition of vertices is used to partition the microstate-space.

SM4.1. State-space partition based on a vertex set partition. Given a partition of the *vertex* set, we consider lumped states based on a partition of microstate-space into sets of microstates with the same number of vertices in each vertex-state *within* each of the cells of the *vertex* partition.

Let us consider an SISa epidemic propagating on a square graph. In this example we consider a two cell partition of vertices $\Pi_V = \{V_1, V_2\}$, where $V_1 = \{1, 4\}$ and $V_2 = \{2, 3\}$. Crucially, this choice of vertex partition allows us to illustrate an example of an approximate lumping¹. We use $N_1 = |V_1|$ and $N_2 = |V_2|$ to denote the number of vertices in V_1 and V_2 respectively, so $N_1 = N_2 = 2$, and we use $P = 2$ to denote the number of vertex-partition cells. We will keep our notation general in this example, e.g. by using P rather than 2, where it does not add excessive complexity. Our approximate lumping will be based on the number of susceptible and infected vertices

¹Note however that in this example it is possible to obtain an exact lumping by choosing vertices in opposite corners to be in the same cell of the vertex partition.

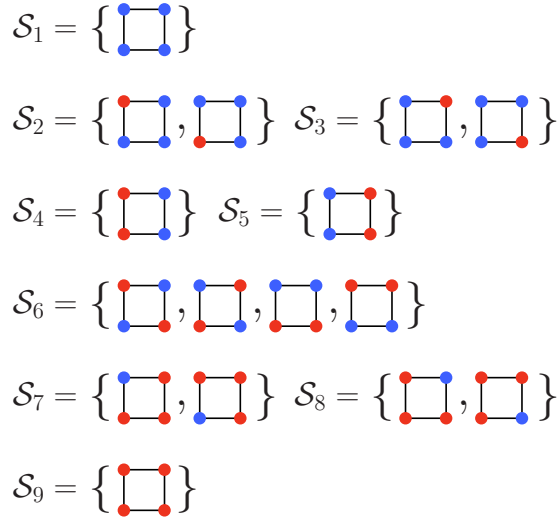


FIG. SM3. Approximate lumping partition of microstate-space for the square example.

in each of the cells in the vertex-partition. We will thus use a two-by-two matrix to represent lumped states, where the rows correspond to vertex-states (susceptible, infected) and the columns correspond to the partition cells (V_1 and V_2). Specifically, lumped states will be denoted by $\mathbf{s} \in \mathbb{Z}_{\geq 0}^{M \times P}$, a two-by-two matrix of non-negative integers whose m, p th entry, $\mathbf{s}_{m,p}$, is the number of vertices in vertex-state \mathcal{W}_m in the vertex partition cell V_p . Note that $\mathbf{s}_{1,p} + \mathbf{s}_{2,p} = N_p$ for $p = 1, 2$ and adding these for $p = 1$ and $p = 2$ we get that the sum of the entries in a matrix \mathbf{s} is $N_1 + N_2 = N = 4$.

SM4.2. The size of the lumped state-space and the sizes of the partition cells. Since there are $N_1 = 2$ vertices in the first partition cell and $M = 2$ vertex states, the possible values for $\mathbf{s}_{1,1}$ are 0, 1 and 2. The corresponding values for $\mathbf{s}_{2,1}$ are 2, 1 and 0 respectively, since $\mathbf{s}_{1,1} + \mathbf{s}_{2,1} = N_p = 2$. These three possibilities arise because a lumped state corresponds to choosing N_1 vertices from the M possible vertex-states with repetition, which is $\binom{N_1+M-1}{N_1}$. A similar argument applies to the second partition cell and hence the total number of lumped states is

$$(SM4.1) \quad n = \binom{N_1 + M - 1}{N_1} \binom{N_2 + M - 1}{N_2} = \binom{3}{2} \binom{3}{2} = 9,$$

and we label these as follows:

$$\begin{aligned} \mathbf{s}^{[1]} &= \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}, \mathbf{s}^{[2]} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}, \mathbf{s}^{[3]} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \\ \mathbf{s}^{[4]} &= \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}, \mathbf{s}^{[5]} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \mathbf{s}^{[6]} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \\ \mathbf{s}^{[7]} &= \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}, \mathbf{s}^{[8]} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}, \mathbf{s}^{[9]} = \begin{pmatrix} 0 & 0 \\ 2 & 2 \end{pmatrix}. \end{aligned}$$

The corresponding vertex-partition approximate lumping $\Pi_{\mathcal{S}} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_9\}$ is illustrated in Figure SM3. For example, there is one microstate in $\mathcal{S}^{[1]}$ (which corresponds to $\mathbf{s}^{[1]}$) which has 2 susceptible (blue) vertices in both V_1 (vertices on the

left of the square) and V_2 (vertices on the right of the square), and no infected (red) vertices in either vertex partition cell. Similarly, microstates in $\mathcal{S}^{[8]}$ have two infected vertices in V_1 and both a susceptible and an infected vertex in V_2 .

The number of microstates that correspond to $\mathbf{s}^{[i]}$ is equivalent to the number of ways to choose $\mathbf{s}_{1,1}^{[i]}$ susceptible vertices in the first partition, multiplied by the number of ways to choose $\mathbf{s}_{1,2}^{[i]}$ susceptible vertices in the second partition. As a notational short-hand, we will use

$$\binom{N}{\mathbf{s}^{[i]}} := \binom{N_1}{\mathbf{s}_{1,1}^{[i]}} \binom{N_2}{\mathbf{s}_{1,2}^{[i]}} = \frac{N_1!}{\mathbf{s}_{1,1}^{[i]}! \mathbf{s}_{2,1}^{[i]}!} \times \frac{N_2!}{\mathbf{s}_{2,1}^{[i]}! \mathbf{s}_{2,2}^{[i]}!},$$

thus for our square example we have

$$\begin{aligned} \binom{N}{\mathbf{s}^{[1]}} &= \binom{2}{2} \binom{2}{2} = 1, & \binom{N}{\mathbf{s}^{[2]}} &= \binom{N}{\mathbf{s}^{[3]}} = \binom{2}{1} \binom{2}{2} = 2, \\ \binom{N}{\mathbf{s}^{[4]}} &= \binom{N}{\mathbf{s}^{[5]}} = \binom{2}{0} \binom{2}{2} = 1, & \binom{N}{\mathbf{s}^{[6]}} &= \binom{2}{1} \binom{2}{1} = 4, \\ \binom{N}{\mathbf{s}^{[7]}} &= \binom{N}{\mathbf{s}^{[8]}} = \binom{2}{1} \binom{2}{2} = 2, & \text{and } \binom{N}{\mathbf{s}^{[9]}} &= \binom{2}{0} \binom{2}{0} = 1. \end{aligned}$$

SM4.3. Lumped generator for the square SISa example. Now we compute the infinitesimal generator for the example with SISa epidemic on a square network based on equation (SM3.4).

The collector matrix, \mathbf{C} , for our approximate lumping of the square is a matrix of size 16×9 , with most of the entries zeros and one entry in each row which is 1. The element in the k -th row is $\mathbf{C}_{k,j}$ if the microstate $\mathcal{S}^{[k]}$ belongs to the lumped state \mathcal{S}_j . Based on the labelling of the states in Figure SM2 and those of the lumping classes given in Figure SM3, the non-zero entries of the collector matrix are

$$\begin{aligned} &\mathbf{C}_{1,1}, \mathbf{C}_{2,2}, \mathbf{C}_{3,2}, \mathbf{C}_{4,3}, \mathbf{C}_{5,3}, \mathbf{C}_{6,4}, \mathbf{C}_{7,5}, \mathbf{C}_{8,6}, \mathbf{C}_{9,6}, \\ &\mathbf{C}_{10,6}, \mathbf{C}_{11,6}, \mathbf{C}_{12,7}, \mathbf{C}_{13,7}, \mathbf{C}_{14,8}, \mathbf{C}_{15,8}, \mathbf{C}_{16,9}. \end{aligned}$$

The matrix \mathbf{C} has a block diagonal form

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{C}_1 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{C}_2 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_3 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{C}_4 \end{pmatrix}$$

with $\mathbf{C}_0 = (1) = \mathbf{C}_4$ and

$$\mathbf{C}_1 = \mathbf{C}_3 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{C}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that this choice of blocks does not correspond to the lumped states, rather it relates to the number of infected vertices and the block tri-diagonal structure of \mathbf{Q}

seen in (SM2.2). Multiplying the infinitesimal generator \mathbf{Q} by \mathbf{C} we get the block tri-diagonal form

$$(SM4.2) \quad \mathbf{QC} = \begin{pmatrix} \Delta_0 \mathbf{C}_0 & \mathbf{A}_0 \mathbf{C}_1 & 0 & 0 & 0 \\ \mathbf{B}_1 \mathbf{C}_0 & \Delta_1 \mathbf{C}_1 & \mathbf{A}_1 \mathbf{C}_2 & 0 & 0 \\ 0 & \mathbf{B}_2 \mathbf{C}_1 & \Delta_2 \mathbf{C}_2 & \mathbf{A}_2 \mathbf{C}_3 & 0 \\ 0 & 0 & \mathbf{B}_3 \mathbf{C}_2 & \Delta_3 \mathbf{C}_3 & \mathbf{A}_3 \mathbf{C}_4 \\ 0 & 0 & 0 & \mathbf{B}_4 \mathbf{C}_3 & \Delta_4 \mathbf{C}_4 \end{pmatrix}$$

Each block can be easily calculated, we show only

$$\mathbf{A}_2 \mathbf{C}_3 = \begin{pmatrix} 0 & 2\alpha + \beta \\ 2\alpha + \beta & 0 \\ \alpha + 2\beta & \alpha + 2\beta \\ \alpha + 2\beta & \alpha + 2\beta \\ \alpha + \beta & \alpha + \beta \\ \alpha + \beta & \alpha + \beta \end{pmatrix}$$

which shows that this lumping is not exact. Namely, recall that a lumping is exact if the total rate from a microstate $S^{[k]} \in \mathcal{S}_i$ into a lumped state \mathcal{S}_j is the same for all microstates in \mathcal{S}_i . In the expression for \mathbf{QC} above, there are two lumped transitions where this is not the case, namely from \mathcal{S}_6 to \mathcal{S}_7 (rows 3 to 6 in the first column of the matrix $\mathbf{A}_2 \mathbf{C}_3$), and from \mathcal{S}_6 to \mathcal{S}_8 (rows 3 to 6 in the second column of this matrix). In both cases, there are two microstates where the rate is $\alpha + 2\beta$ and two where it is $\alpha + \beta$. This confirms that our partition of vertices is an approximate lumping.

We will now compute the lumped transition rate. The distributor matrix has the block diagonal form

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{D}_1 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{D}_2 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{D}_3 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{D}_4 \end{pmatrix}$$

with $\mathbf{D}_0 = (1) = \mathbf{D}_4$ and

$$\mathbf{D}_1 = \mathbf{D}_3 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \mathbf{D}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Again, this choice of blocks does not reflect the lumping partition. Multiplying \mathbf{QC} , given in (SM4.2), by \mathbf{D} we get \mathbf{q} in the block tri-diagonal form (SM4.3)

$$(SM4.3) \quad \mathbf{q} = \mathbf{DQC} = \begin{pmatrix} \mathbf{D}_0 \Delta_0 \mathbf{C}_0 & \mathbf{D}_0 \mathbf{A}_0 \mathbf{C}_1 & 0 & 0 & 0 \\ \mathbf{D}_1 \mathbf{B}_1 \mathbf{C}_0 & \mathbf{D}_1 \Delta_1 \mathbf{C}_1 & \mathbf{D}_1 \mathbf{A}_1 \mathbf{C}_2 & 0 & 0 \\ 0 & \mathbf{D}_2 \mathbf{B}_2 \mathbf{C}_1 & \mathbf{D}_2 \Delta_2 \mathbf{C}_2 & \mathbf{D}_2 \mathbf{A}_2 \mathbf{C}_3 & 0 \\ 0 & 0 & \mathbf{D}_3 \mathbf{B}_3 \mathbf{C}_2 & \mathbf{D}_3 \Delta_3 \mathbf{C}_3 & \mathbf{D}_3 \mathbf{A}_3 \mathbf{C}_4 \\ 0 & 0 & 0 & \mathbf{D}_4 \mathbf{B}_4 \mathbf{C}_3 & \mathbf{D}_4 \Delta_4 \mathbf{C}_4 \end{pmatrix}.$$

Each block can be easily calculated, and hence the coarsegrained approximate lumping

153 infinitesimal generator is
 (SM4.4)

$$154 \quad \mathbf{q} = \begin{pmatrix} \delta_1 & 2\alpha & 2\alpha & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma & \delta_2 & 0 & \alpha + \beta & 0 & 2\alpha + \beta & 0 & 0 & 0 \\ \gamma & 0 & \delta_3 & 0 & \alpha + \beta & 2\alpha + \beta & 0 & 0 & 0 \\ 0 & 2\gamma & 0 & \delta_4 & 0 & 0 & 0 & 2(\alpha + \beta) & 0 \\ 0 & 0 & 2\gamma & 0 & \delta_5 & 0 & 2(\alpha + \beta) & 0 & 0 \\ 0 & \gamma & \gamma & 0 & 0 & \delta_6 & \alpha + \frac{3}{2}\beta & \alpha + \frac{3}{2}\beta & 0 \\ 0 & 0 & 0 & 0 & \gamma & 2\gamma & \delta_7 & 0 & \alpha + 2\beta \\ 0 & 0 & 0 & \gamma & 0 & 2\gamma & 0 & \delta_8 & \alpha + 2\beta \\ 0 & 0 & 0 & 0 & 0 & 0 & 2\gamma & 2\gamma & \delta_9 \end{pmatrix},$$

155 where the lines correspond to the blocks in (SM4.3). Here δ_i is the negative of the
 156 sum of the off diagonal elements in the i th row of the matrix, so for example $\delta_2 =$
 157 $-(3\alpha + 2\beta + \gamma)$.

158 **SM5. Lumped generator for two cell vertex-partitions.** While it was easy
 159 to compute \mathbf{q} in our example using \mathbf{Q} , \mathbf{C} and \mathbf{D} , this will not be the case for large
 160 systems of interest. Thus we need to consider how we could compute \mathbf{q} without using
 161 \mathbf{Q} , \mathbf{C} and \mathbf{D} directly. We will use the square example to illustrate this process. For
 162 finite M and $P = 2$, a lumped state will be denoted by a matrix $\mathbf{s} \in \mathbb{Z}_{\geq 0}^{M \times P}$ whose
 163 m, p th entry, $\mathbf{s}_{m,p}$, is the number of vertices in vertex-state \mathcal{W}_m in the vertex partition
 164 cell V_p .

165 **SM5.1. Lumped generator for an arbitrary two cell vertex-partition.**
 166 We will start by considering the transition rate from an arbitrary lumped state $\mathbf{s}^{[i]}$ to
 167 another arbitrary lumped state $\mathbf{s}^{[j]} \neq \mathbf{s}^{[i]}$. In general this is given by

$$168 \quad (\text{SM5.1}) \quad \mathbf{q}_{ij} = \frac{1}{|\mathcal{S}_i|} \sum_{S^{[k]} \in \mathcal{S}_i} \sum_{S^{[l]} \in \mathcal{S}_j} \mathbf{Q}_{kl},$$

169 but it turns out that rather than summing over microstates, as (SM5.1) suggests, it is
 170 easier to consider the possible transitions of individual vertices and sum their rates.

171 Let d_p^v denote the number of neighbours of vertex v in the p th vertex-partition
 172 cell. In our example, all vertices have one neighbour in their own cell and one in the
 173 other cell, so $d_1^v = d_2^v = 1$ for all vertices v . The degree of vertex v is

$$174 \quad (\text{SM5.2}) \quad d^v := \sum_{p=1}^P d_p^v.$$

We represent the neighbourhood of v using a two-by-two, non-negative, integer-valued
 matrix $\mathbf{n}^v \in \mathbb{Z}_{\geq 0}^{M \times P}$, whose component $\mathbf{n}_{m,p}^v$ is the number of neighbours of vertex v in
 the m th vertex-state and in the p th vertex-partition cell. We call \mathbf{n}^v a *neighbourhood*
count and it must satisfy

$$\sum_{m=1}^M \mathbf{n}_{m,p}^v = d_p^v,$$

175 for $0 \leq p \leq P$. Note that

$$176 \quad (\text{SM5.3}) \quad \sum_{p=1}^P \sum_{m=1}^M \mathbf{n}_{m,p}^v = d^v.$$

177 We will use a single index on this matrix to indicate a column, i.e. $\mathbf{n}_p^v \in \mathbb{Z}_{\geq 0}^M$ is the
 178 p th column of \mathbf{n}^v .

179 For our square example, since $d_p^v = 1$ for all v and p , the only possibly neighbour-
 180 hoods in vertex-partition cell p are

$$181 \quad (\text{SM5.4}) \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

182 i.e. either a susceptible or an infected neighbour respectively. Thus the possible
 183 neighbourhood counts \mathbf{n}^v are

$$184 \quad (\text{SM5.5}) \quad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

185 The first of these matrices corresponds to having a susceptible neighbour in both par-
 186 titions, the second corresponds to having an infected neighbour in V_1 and a susceptible
 187 neighbour in V_2 , and so on.

188 Without loss of generality we assume that the transition from the lumped state
 189 $\mathbf{s}^{[i]}$ to the lumped state $\mathbf{s}^{[j]}$ corresponds to a vertex in vertex-partition $q \in \{1, 2\}$
 190 transitioning from vertex-state $\mathcal{A} \in \mathcal{W}$ to $\mathcal{B} \in \mathcal{W}$, where $\mathcal{A} \neq \mathcal{B}$. To compute (SM5.1),
 191 for each $v \in V_q$ we can construct all possible microstates where v has vertex-state \mathcal{A} .
 192 If in this process we specify the vertex-states of the neighbours of v , then we can
 193 determine the rate at which v changes from \mathcal{A} to \mathcal{B} . Summing this contribution from
 194 all possible cases yields \mathbf{q}_{ij} . A proof of this is given in Section 6 of the main paper.

195 The number of ways that we can arrange the vertex-states of the neighbours of v
 196 in vertex-partition cell $p \neq q$ according to some \mathbf{n}_p^v , as well as the vertex-states of the
 197 other vertices in vertex-partition cell $p \neq q$ according to \mathbf{s}_p is

$$198 \quad (\text{SM5.6}) \quad A(\mathbf{s}_p, \mathbf{n}_p^v) := \binom{\sum_{p=1}^P \mathbf{n}_{m,p}}{\mathbf{n}_p^v} \binom{\sum_{p=1}^P \mathbf{s}_{m,p} - \mathbf{n}_{m,p}^v}{\mathbf{s}_p - \mathbf{n}_p^v}.$$

Note that we have used a vector in the bottom of the multinomial coefficient nota-
 tion to indicate that the elements of the vector should be in the denominator of the
 multinomial coefficient, i.e.

$$\binom{d_p^v}{\mathbf{n}_p^v} := \frac{d_p^v!}{\mathbf{n}_{1,p}^v! \mathbf{n}_{2,p}^v! \dots \mathbf{n}_{m,p}^v!}.$$

199 We will also assume the standard convention that a multinomial coefficient is zero if
 200 any entry is negative. For the vertex-partition cell q , which contains v , the number
 201 of ways that we can arrange the neighbours of v according to \mathbf{n}_q^v is

$$202 \quad A(\mathbf{s}_q - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) = \binom{d_q^v}{\mathbf{n}_q^v} \binom{N_q - 1 - d_q^v}{\mathbf{s}_q - \mathbf{e}_{\mathcal{A}} - \mathbf{n}_q^v},$$

203 where $\mathbf{e}_{\mathcal{A}}$ is a vector of length M with a one in the entry corresponding to vertex-state
 204 \mathcal{A} and zeros elsewhere. This is to account for the fact that we assumed vertex v is in
 205 vertex-state \mathcal{A} , so there is one less \mathcal{A} vertex in \mathbf{s}_q . Then for a given \mathbf{n}^v and $p \neq q$,
 206 using (SM5.6) there are

$$207 \quad (\text{SM5.7}) \quad A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[j]}, \mathbf{n}_p^v)$$

microstates in \mathcal{S}_i in which vertex v in vertex partition cell q is in vertex-state \mathcal{A} , its neighbours' vertex-states correspond to \mathbf{n}^v , and the total number of vertices in each vertex-state and in each vertex-partition cell corresponds to $\mathbf{s}^{[i]}$. Thus we get \mathbf{q}_{ij} by summing over all feasible realisations of the matrix \mathbf{n}^v and vertices in V_q , which yields (SM5.8)

$$\mathbf{q}_{ij} = \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_q} \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} \left(\zeta_0^{\mathcal{A}, \mathcal{B}} + \sum_{m=1}^M \sum_{r=1}^P \zeta_m^{\mathcal{A}, \mathcal{B}} \mathbf{n}_{m,r}^v \right) A(\mathbf{s}_q^{[i]} - \mathbf{e}_{\mathcal{A}}, \mathbf{n}_q^v) A(\mathbf{s}_p^{[i]}, \mathbf{n}_p^v).$$

In this equation, the sums over $\mathbf{n}_1^v | d_1^v$ and $\mathbf{n}_2^v | d_2^v$ specify the $P = 2$ columns of \mathbf{n}^v .

For our square example and the case where a transition from $\mathbf{s}^{[i]}$ to $\mathbf{s}^{[j]}$ corresponds to a susceptible vertex in vertex-partition cell V_1 becoming infected, (SM5.8) becomes

$$(SM5.9) \quad \mathbf{q}_{ij} = \frac{1}{\binom{N}{\mathbf{s}^{[i]}}} \sum_{v \in V_1} \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} [\alpha + \beta (\mathbf{n}_{2,1}^v + \mathbf{n}_{2,2}^v)] A(\mathbf{s}_1 - \mathbf{e}_1, \mathbf{n}_1^v) A(\mathbf{s}_2, \mathbf{n}_2^v).$$

Here $\mathbf{n}_1^v | d_1^v$ corresponds to a sum over non-negative vectors \mathbf{n}_1^v whose elements sum to d_1^v , and similarly for $\mathbf{n}_2^v | d_2^v$ —in both cases these are the vectors listed in (SM5.4). The lumped states from which a susceptible vertex in V_1 can become infected are $\mathbf{S}^{[2]}$, $\mathbf{S}^{[3]}$, $\mathbf{S}^{[5]}$, $\mathbf{S}^{[6]}$ and $\mathbf{S}^{[7]}$. We will use (SM5.9) to compute the lumped transition rate for two examples. First consider the rate from $\mathbf{s}^{[3]}$ to $\mathbf{s}^{[6]}$. For either vertex in V_1 , expanding the sums over neighbourhoods yields

$$\begin{aligned} & \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} [\alpha + \beta (\mathbf{n}_{2,1}^v + \mathbf{n}_{2,2}^v)] A(\mathbf{s}_1 - \mathbf{e}_1, \mathbf{n}_1^v) A(\mathbf{s}_2, \mathbf{n}_2^v) \\ &= [\alpha + \beta(0+0)] A\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) A\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \\ &+ [\alpha + \beta(0+1)] A\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) A\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &+ [\alpha + \beta(1+0)] A\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) A\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \\ &+ [\alpha + \beta(1+1)] A\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) A\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right), \\ &= \alpha \cdot 1 \cdot 1 + [\alpha + \beta] \cdot 1 \cdot 1 + [\alpha + \beta] \cdot 0 \cdot 1 + [\alpha + 2\beta] \cdot 0 \cdot 1, \\ &= 2\alpha + \beta. \end{aligned}$$

Since there are two vertices in V_1 , we get a contribution of $2\alpha + \beta$ from each, but the averaging constant is

$$\binom{N}{\mathbf{s}^{[3]}} := \binom{N}{\mathbf{s}_1^{[3]}} \binom{N}{\mathbf{s}_2^{[3]}} = 2,$$

thus we have $\mathbf{q}_{3,6} = 2\alpha + \beta$, which agrees with the corresponding entry in (SM4.4).

A similar calculation for the transition from $\mathbf{s}^{[6]}$ to $\mathbf{s}^{[8]}$ yields

$$\begin{aligned} \mathbf{q}_{6,8} &= \frac{1}{4} \times 2 \times \{[\alpha + \beta(0+0)] \cdot 0 \cdot 1 + [\alpha + \beta(0+1)] \cdot 0 \cdot 1 \\ &+ [\alpha + \beta(1+0)] \cdot 1 \cdot 1 + [\alpha + \beta(1+1)] \cdot 1 \cdot 1\} = \alpha + \frac{3}{2}\beta, \end{aligned}$$

which again agrees with the corresponding entry in (SM4.4).

SM6. Sum-product property.

THEOREM SM6.1. *The sum-product property is that*

$$\prod_{p=1}^P \sum_{i_p=1}^{r_p} a_{i_p}^p = \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \cdots \sum_{i_P=1}^{r_P} \prod_{p=1}^P a_{i_p}^p.$$

Proof. We prove this by induction. This is trivially true for $P = 1$. We assume it is true for P , then for $P + 1$ we have

$$\begin{aligned} \prod_{p=1}^{P+1} \sum_{i_p=1}^{r_p} a_{i_p}^p &= \left(\prod_{p=1}^P \sum_{i_p=1}^{r_p} a_{i_p}^p \right) \left(\sum_{i_{P+1}=1}^{r_{P+1}} a_{i_{P+1}}^{P+1} \right), \\ &= \left(\sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \cdots \sum_{i_P=1}^{r_P} \prod_{p=1}^P a_{i_p}^p \right) \left(\sum_{i_{P+1}=1}^{r_{P+1}} a_{i_{P+1}}^{P+1} \right), \\ &= \sum_{i_{P+1}=1}^{r_{P+1}} a_{i_{P+1}}^{P+1} \left(\sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \cdots \sum_{i_P=1}^{r_P} \prod_{p=1}^P a_{i_p}^p \right), \\ &= \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \cdots \sum_{i_{P+1}=1}^{r_{P+1}} \prod_{p=1}^{P+1} a_{i_p}^p. \end{aligned}$$

□

As a consequence of the sum-product property, for lumped state \mathbf{s} and neighbourhood count \mathbf{n}^v of a vertex $v \in V$, we have

$$\begin{aligned} \prod_{p=1}^P \left[\sum_{\mathbf{n}_p^v | d_p^v} A(\mathbf{s}_p, \mathbf{n}_p^v) \right] &= \sum_{\mathbf{n}_1^v | d_1^v} \sum_{\mathbf{n}_2^v | d_2^v} \cdots \sum_{\mathbf{n}_P^v | d_P^v} \prod_{p=1}^P A(\mathbf{s}_p, \mathbf{n}_p^v) \\ &= \sum_{\mathbf{n}^v | d^v} \prod_{p=1}^P A(\mathbf{s}_p, \mathbf{n}_p^v). \end{aligned}$$

SM7. Generalised Vandermonde identity. The Vandermonde identity can be generalised to multinomials. To see this, consider the multinomial theorem,

$$(x_1 + x_2 + \cdots + x_M)^N = \sum_{s_1 + s_2 + \cdots + s_M = N} \binom{N}{s_1, s_2, \dots, s_M} \prod_{m=1}^M x_m^{s_m},$$

where $s_1, s_2, \dots, s_M \geq 0$. Thus for $d < N$ we have

$$\begin{aligned} &(x_1 + x_2 + \cdots + x_M)^{N-d} (x_1 + x_2 + \cdots + x_M)^d \\ &= \left[\sum_{l_1 + l_2 + \cdots + l_M = N-d} \binom{N-d}{l_1, l_2, \dots, l_M} \prod_{m=1}^M x_m^{l_m} \right] \\ &\quad \times \left[\sum_{n_1 + n_2 + \cdots + n_M = d} \binom{d}{n_1, n_2, \dots, n_M} \prod_{m=1}^M x_m^{n_m} \right], \\ &= \sum_{l_1 + l_2 + \cdots + l_M = N-d} \sum_{n_1 + n_2 + \cdots + n_M = d} \binom{d}{n_1, n_2, \dots, n_M} \binom{N-d}{l_1, l_2, \dots, l_M} \prod_{m=1}^M x_m^{l_m + n_m}, \end{aligned}$$

where $l_m, n_m \geq 0$. Thus instead of summing over $l_1 + l_2 + \dots + l_M = N - d$ we can sum over $s_1 + s_2 + \dots + s_M = N$ and use $l_m = s_m - n_m$ with the convention that the multinomial coefficients are zero if $s_m - n_m < 0$ or $\sum_m (s_m - n_m) \neq N - d$. Thus we have

$$\sum_{s_1+s_2+\dots+s_M=N} \binom{N}{s_1, s_2, \dots, s_M} \prod_{m=1}^M x_m^{s_m} = \sum_{s_1+s_2+\dots+s_M=N} \left[\sum_{n_1+n_2+\dots+n_M=d} \binom{d}{n_1, n_2, \dots, n_M} \binom{N-d}{s_1-n_1, s_2-n_2, \dots, s_M-n_M} \right] \times \prod_{m=1}^M x_m^{s_m},$$

and the equality of the polynomials on the two sides implies that the coefficients of corresponding terms are identical, hence

$$\binom{N}{s_1, s_2, \dots, s_M} = \sum_{n_1+n_2+\dots+n_M=d} \binom{d}{n_1, n_2, \dots, n_M} \binom{N-d}{s_1-n_1, s_2-n_2, \dots, s_M-n_M}.$$

LEMMA SM7.1. Let \mathbf{s} be a lumped state and \mathbf{n}^v a neighbourhood count of a vertex $v \in V$, then

$$(SM7.1) \quad \binom{N}{\mathbf{s}} = \sum_{\mathbf{n}^v | d^v} \prod_{p=1}^P A(\mathbf{s}_p, \mathbf{n}_p^v).$$

Proof. From the Vandermonde property of multinomials and (SM5.6), we have

$$\binom{N_p}{\mathbf{s}_p} = \sum_{\mathbf{n}_p^v | d_p^v} A(\mathbf{s}_p, \mathbf{n}_p^v).$$

Thus it follows that

$$\begin{aligned} \binom{N}{\mathbf{s}} &= \prod_{p=1}^P \binom{N_p}{\mathbf{s}_p} \\ &= \prod_{p=1}^P \left(\sum_{\mathbf{n}_p^v | d_p^v} A(\mathbf{s}_p, \mathbf{n}_p^v) \right). \end{aligned}$$

The result then follows from application of the sum product property. \square

SM8. Degree-based mean-field. In this section we show that the Eames and Keeling degree-based mean-field is equivalent to the Pastor-Satorras and Vespignani [SM5] degree-based mean-field. The Eames and Keeling degree-based mean-field is

$$(SM8.1) \quad \frac{d[I^k]}{dt} = -\gamma[I^k] + \beta \sum_{k'} \frac{[S^k]}{N_k} \frac{[I^{k'}]}{N_{k'}} [kk'],$$

where $[I^k]$ ($[S^k]$) denotes the number of infected (susceptible) nodes with k neighbours, N_k is the number of nodes with degree k , $[kk']$ is the number of partnerships between

individuals with k and k' partners, γ is the recovery rate and β is the infection rate. Pastor-Satorras and Vespignani use ρ_k to denote the “average density” of infected nodes of degree k . The average fraction of infected nodes is then

$$\rho = \sum_k p_k \rho_k,$$

which suggests that $\rho_k = [I^k]/N_k$. The evolution of ρ_k is described by

$$(SM8.2) \quad \dot{\rho}_k = -\rho_k + \lambda k [1 - \rho_k] \frac{\sum_{k'} k' p_{k'} \rho_{k'}}{z},$$

where $\lambda = \beta/\gamma$ is the effective spreading rate, and $z = \sum_k k p_k$ is the mean degree. Boguñá et al [SM1] extend this to account for degree correlations, where they define $P(k'|k)$ to be the conditional probability that a node of degree k is connected to a node of degree k' . The evolution of ρ_k is then given by

$$(SM8.3) \quad \dot{\rho}_k = -\rho_k + \lambda k [1 - \rho_k] \sum_{k'} P(k'|k) \rho_{k'}.$$

Now starting from the Ealing and Keemes degree-based mean field, we set $\rho_k = [I^k]/N_k$, so with $N_k = [S^k] + [I^k]$ we have $(1 - \rho_k) = [S^k]/N_k$, then dividing (SM8.1) through by N_k yields

$$\dot{\rho}_k = -\gamma \rho_k + \beta [1 - \rho_k] \sum_{k'} \frac{[kk']}{N_k} \rho_{k'}.$$

Since kN_k is the total number of edges connected to nodes of degree k , we define

$$P(k'|k) = \frac{[kk']}{kN_k},$$

and hence

$$\dot{\rho}_k = -\gamma \rho_k + \beta k [1 - \rho_k] \sum_{k'} P(k'|k) \rho_{k'},$$

which after time-rescaling yields (SM8.3). In the absence of degree correlations we have

$$P(k'|k) = \frac{k' N_{k'}}{z N} = \frac{k' p_{k'}}{z},$$

and hence we obtain (SM8.2).

REFERENCES

- [SM1] M. BOGUÑÁ, R. PASTOR-SATORRAS, AND A. VESPIGNANI, *Absence of epidemic threshold in scale-free networks with degree correlations*, Physical review letters, 90 (2003), p. 028701.
- [SM2] P. BUCHHOLZ, *Exact and ordinary lumpability in finite markov chains*, Journal of Applied Probability, 31 (1994), pp. 59–75.
- [SM3] J. G. KEMENY, J. L. SNELL, ET AL., *Finite Markov Chains*, vol. 356, van Nostrand Princeton, NJ, 1960.
- [SM4] M. KIJIMA, *Markov processes for stochastic modeling*, vol. 6, CRC Press, 1997.
- [SM5] R. PASTOR-SATORRAS AND A. VESPIGNANI, *Epidemic spreading in scale-free networks*, Physical Review Letters, 86 (2001), p. 3200.
- [SM6] J. A. WARD AND M. LÓPEZ-GARCÍA, *Exact analysis of summary statistics for continuous-time discrete-state markov processes on networks using graph-automorphism lumping*, Applied Network Science, 4 (2019), p. 108.