Unified mean-field framework for susceptible-infected-susceptible epidemics on networks, based on graph partitioning and the isoperimetric inequality

Devriendt, K.; Van Mieghem, P.

DOI
10.1103/PhysRevE.96.052314

Publication date
2017

Document Version
Final published version

Published in
Physical Review E

Unified mean-field framework for susceptible-infected-susceptible epidemics on networks, based on graph partitioning and the isoperimetric inequality

K. Devriendt and P. Van Mieghem
Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, the Netherlands
(Received 14 July 2017; published 27 November 2017)

We propose an approximation framework that unifies and generalizes a number of existing mean-field approximation methods for the susceptible-infected-susceptible (SIS) epidemic model on complex networks. We derive the framework, which we call the unified mean-field framework (UMFF), as a set of approximations of the exact Markovian SIS equations. Our main novelty is that we describe the mean-field approximations from the perspective of the isoperimetric problem, which results in bounds on the UMFF approximation error. These new bounds provide insight in the accuracy of existing mean-field methods, such as the N-intertwined mean-field approximation and heterogeneous mean-field method, which are contained by UMFF. Additionally, the isoperimetric inequality relates the UMFF approximation accuracy to the regularity notions of Szemerédi's regularity lemma.

DOI: 10.1103/PhysRevE.96.052314

I. INTRODUCTION

Epidemic spread on complex networks is a widely studied topic in the field of network science [1], covering many applications ranging from diseases and computer viruses, to the spreading of ideas and emotions. While the mathematical study of epidemics dates back to the work of Bernoulli in the 18th century, the focus on the role of the network topology only started at the end of the 20th century with the work of Kephart and White [2]. With the recent observations that networks seem ubiquitous in both natural and man-made systems, a better understanding of the interplay between dynamic processes and network topology has become an important pursuit.

In the theory of epidemics on complex networks, the compartmental model of Kermack and McKendrick [3] from 1927 is regarded as a landmark. In compartmental models, each entity in the population is assumed to be in a certain state, for instance healthy, contagious, immune, aware of the disease, exposed or others. The state of each entity, from now on called "node," can change based on the current state of the node itself and its neighboring nodes. By these local interactions the disease can spread, die out or show other behaviors depending on the model. A general overview of the basic models and current progress in the field of epidemics on complex networks is given in Ref. [1].

Like many network-epidemic studies, we will focus on one specific compartmental model: the susceptible-infected-susceptible (SIS) model. The SIS model is attractive and simple enough for a deep theoretical study, while still complex enough to exhibit global behavior that is nontrivially coupled to the small-scale process and the topology of the underlying network. In the SIS model, each node in the network can be in either of two states: susceptible (S) or infected (I). These states can change over time when an infected node is cured, or when a susceptible node is infected by a sick neighbor. These curing and infection events are stochastic processes that determine the dynamics of the epidemic. For a given initial distribution of infected nodes, basic questions are: What is the evolution of the state of the nodes in the network? How many nodes are infected in the metastable state? Does the disease die out before reaching a significant fraction of the population?

To address these questions, further assumptions need to be made about the dynamics of the SIS process. In the continuous-time Markovian SIS model, on which this article focuses, the infection and healing events are modeled as independent Poisson processes. More general distributions are possible [4], but in Poisson processes, the waiting times for infection or healing events are exponentially distributed, which means that they satisfy the memoryless property. As a result, the transitions between different disease states in the network become equivalent to state transitions in a Markov chain, which allows the Markovian SIS model to be exactly described [5]. However, since the number of possible states grows exponentially with the number of nodes, this exact description is infeasible for large networks. Consequently, several approximate methods have been proposed. The N-intertwined mean-field approximation (NIMFA) [5] and the heterogeneous mean-field method (HMF) [6,7] are two widely used approximation methods, which are contained in our new framework. An overview of these two methods and other SIS approximation methods can be found in Refs. [1] and [8].

In this article, we present the unified mean-field framework (UMFF), which consists of two general approximation steps. First, a topological approximation leads to a coarse-grained description of the SIS process. Second, a moment-closure approximation further simplifies the SIS process description by omitting higher-order correlations. UMFF contains a number of existing mean-field methods, like NIMFA and HMF as mentioned earlier, and additionally extends the range of known SIS approximation methods. Linear stability analysis of the resulting UMFF equations also leads to the formulation of an epidemic threshold that depends on the choice of approximation steps. Apart from the unification and generalization, our main results are based on the close connection between the infection process in SIS epidemics and the well-studied isoperimetric problem [9,10]. This connection provides novel insights, e.g., about the scaling behavior of the SIS process on large graphs, and allows us to deduce powerful bounds on the UMFF approximations. To our knowledge, such

general bounds for mean-field approximations on graphs are derived for the first time.

Section II starts by defining the SIS epidemic model on networks and elaborates on the feasibility of the exact SIS description. In Sec. III, we present the unified mean-field framework, which consists of the UMFF topological approximation, the UMFF moment-closure approximation, the resulting set of UMFF equations and the UMFF epidemic threshold. In Sec. IV, we show how the UMFF equations follow from the exact SIS equations by subsequently introducing the two general UMFF approximations. Section V describes how existing mean-field methods are contained by UMFF and in particular, we show how UMFF encompasses both NIMFA and HMF. Section VI introduces the isoperimetric problem and describes its analogy with the infection process. This analogy leads to the topological UMFF approximation and bounds. In Sec. VII, we discuss the relation between UMFF and Szemerédi's regularity lemma and explore the implications of this relation for the SIS process on large graphs. Section VIII overviews some related work. Finally, Sec. IX concludes by summarizing the main properties of UMFF and by suggesting future research directions.

II. BACKGROUND: THE SIS EPIDEMIC MODEL

The compartmental SIS epidemic model describes the spread of an epidemic on a network, which we represent by a graph $G(\mathcal{N},\mathcal{L})$. Here, $\mathcal{N}$ is the set of N nodes and $\mathcal{L}$ the set of L undirected, unweighted links between pairs of nodes. The graph's topology is conveniently represented by an $N \times N$ adjacency matrix A, with elements:

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{L} \\ 0 & \text{otherwise} \end{cases}.$$

Since we consider undirected and unweighted graphs, the adjacency matrix A is real and symmetric, possessing the following eigendecomposition [11]:

$$A = X\Lambda X^T = \sum_{i=1}^{N} \lambda_i x_i x_i^T,$$

where X is the orthogonal eigenmatrix with eigenvectors x_i as columns and $\Lambda = \mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_N)$ is the diagonal matrix with eigenvalues on its diagonal. Because the adjacency matrix A is real and symmetric, all eigenvalues are real and can be ordered as $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N$. Another graph-related matrix is the Laplacian matrix Q, defined as

$$Q = \Delta - A,$$

where Δ is the diagonal matrix containing the node degrees. Since the Laplacian Q is also a real and symmetric matrix, we can write the eigendecomposition:

$$Q = ZMZ^T = \sum_{i=1}^{N} \mu_i z_i z_i^T,$$

where Z is the orthogonal eigenmatrix with eigenvectors z_i as columns and $M = \mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_N)$. Since all rows of Q sum to zero, it holds that $Qu = 0$, where u is the all-one vector. The eigenvalue equation $Qu = \mu_N u$ with $\mu_N = 0$ illustrates that Q has at least one zero eigenvalue, belonging to the eigenvector $\frac{u}{\sqrt{N}}$. The Laplacian Q is positive semidefinite, which means that all eigenvalues are non-negative, i.e., $\mu_i \geqslant 0$ for all $i \leqslant N$. Additionally, the multiplicity of the zero eigenvalue μ_N is known to be one for connected graphs [11]. Hence, for any connected graph, we can write the ordered sequence of Laplacian eigenvalues: $0 = \mu_N < \mu_{N-1} \leqslant \cdots \leqslant \mu_1$.

The disease state of each node $n \in \mathcal{N}$ at a given time t, is captured by a Bernoulli random variable $W_n(t)$; the expression $W_n(t) = 0$ means that node n is healthy, but susceptible (S) to the disease, while $W_n(t) = 1$ means that the node is infected (I) and contagious. The evolution of the disease states over time is governed by the disease dynamics

$$S \to I \to S,$$

which means that susceptible nodes can become infected nodes, which in turn can become susceptible. The $S \to I$ transition is called *infection* and can occur when a susceptible node n has an infected neighbor j in the network. The $I \to S$ transition is called *curing* and captures the process where each infected node has the possibility to cure. To make the SIS dynamics tractable, the infection and curing events are assumed to be independent Poisson processes. In particular, for the curing process,

$$\Pr[W_n(t+h) = 0 | W_n(t) = 1] = \delta e^{-\delta h} \tag{1}$$

means that, disregarding all other processes, the waiting time for the $I \to S$ transition is exponentially distributed with rate δ. In general, each node n can have a different, time-dependent rate $\delta_n(t)$, but further in this work we confine ourselves to a fixed and time-independent rate δ. If we consider just one link between a susceptible node n and an infected node j, which we will call an *infective link*, then the infection process obeys

$$\Pr[W_n(t+h) = 1 | W_n(t) = 0] = \beta e^{-\beta h}, \tag{2}$$

where we assume that the infected neighbor node j stays infected and does not cure, i.e., $W_j(t+s) = 1$ for any time $s \in [0,h]$. Again, each link $(n,j) \in \mathcal{L}$ can have a specific rate $\beta_{nj}(t)$, but for simplicity we assume a fixed and time-independent rate β. For notational purposes, we will often omit the time reference t in time-dependent variables by writing W_n instead of $W_n(t)$ and similarly for other time-dependent variables. The infection and curing events can be modeled as more general renewal processes [12], which results in different distributions for the waiting times (1) and (2). The corresponding SIS process is described in Ref. [4].

By assuming the infection and curing processes to be independent Poisson processes, the SIS process constitutes a Markov process [12]. For the continuous-time SIS Markov process, the infinitesimal generator can be deduced from the SIS dynamics [5], which allows for an exact description of the SIS process. Unfortunately, there are 2^N possible states for an SIS process on an N node network, which means that for roughly $N > 20$, finding a solution of the 2^N linear equations becomes infeasible on current computers. The computational complexity of representing all possible disease states on a network is the main problem of SIS epidemics on networks. To resolve this complexity problem of the exact SIS equations, it is necessary to introduce approximations.

TABLE I. Overview of node-level and partition-level variables according to a specific partitioning. $\mathbb{1}$ is the indicator function for which $\mathbb{1}_{\{S\}} = 1$ if statement S is true and zero otherwise.

	Single node	Partition (π)
Node or partition indicator	Node $i \in \{0,1,\ldots,N\}$	Partition $k \in \{0,1,\ldots,K\}$
Indicator vector	$e_i \in \mathbb{R}^N$	$\tilde{e}_k \in \mathbb{R}^K$
	$(e_i)_j = \mathbb{1}_{\{i=j\}}$	$(\tilde{e}_k)_m = \mathbb{1}_{\{k=m\}}$
Partition sum vector	$s_k \in \mathbb{R}^N$	N.A.
	$(s_k)_i = \mathbb{1}_{\{i\in\mathcal{N}_k\}}$	
All-one vector	$u = (1,1,\ldots,1)^T$	$\tilde{u} = (N_1,N_2,\ldots,N_K)^T$
State vector	$w = (w_1,w_2,\ldots,w_N)^T$	$\tilde{w} = (\tilde{w}_1,\tilde{w}_2,\ldots,\tilde{w}_k)^T$
	$w_i = \mathbb{1}_{\{\text{node i is infected}\}}$	$\tilde{w}_k = s_k^T w$
Adjacency matrix	$A \in \mathbb{R}^{N\times N}$	$\widetilde{A} \in \mathbb{R}^{K\times K}$
	$a_{ij} = \mathbb{1}_{\{(i,j)\in\mathcal{L}\}}$	$\widetilde{a}_{km} = \frac{s_k^T A s_m}{N_k N_m} = \frac{L_{km}}{N_k N_m}$
Submatrix	$A^{(km)}$	$\widetilde{A}^{(km)}$
	$a_{ij}^{(km)} = a_{ij}\mathbb{1}_{\{i\in\mathcal{N}_k \text{ and } j\in\mathcal{N}_m\}}$	$\tilde{a}_{ij}^{(km)} = \tilde{a}_{ij}\mathbb{1}_{\{i=k \text{ and } j=m\}}$

One family of approximation methods are the mean-field SIS approximations, which we unify and generalize below with our proposed framework.

III. DEFINITION OF THE UNIFIED MEAN-FIELD FRAMEWORK

Before describing the unified mean-field framework (UMFF), we introduce definitions and notations.

As introduced earlier, we denote the stochastic disease state of node n at time t by the Bernoulli random variable $W_n(t)$. Because the disease state of each node in the network is a result of the Poissonian infection and curing processes, the disease probabilities of different nodes are not independent. This means that the disease probabilities should be described by the joint probability distribution $\Pr[W_1(t) = w_1, W_2(t) = w_2, \ldots, W_N(t) = w_N]$, where $W_n(t)$ is the Bernoulli random variable representing the disease state of node n, and $w_n \in \{0,1\}$ represents a specific outcome of this random variable. If we concatenate the random disease states into the $N \times 1$ random vector $W(t) = (W_1(t), W_2(t), \ldots, W_N(t))^T$, and the disease state realizations into the $N \times 1$ disease-state vector $w = (w_1, w_2, \ldots, w_N)^T$, then the probability of the network to be in a specific disease state w, can be written compactly as $\Pr[W(t) = w]$. The exact description of the SIS process consists of the probabilities $\Pr[W(t) = w]$ for all possible disease states $w \in \{0,1\}^N$, which shows, indeed, that the exact SIS description consists of 2^N states.

In order to address the complexity problem of the exact SIS description, UMFF relies on partitions in a graph:

Definition 1 (Partitioning). A partitioning π of graph G defines a partitioning of the node set $\mathcal{N}$ of G into K nonempty, disjoints partitions $\mathcal{N}_k \subseteq \mathcal{N}$ such that $\bigcup_{k=1}^K \mathcal{N}_k = \mathcal{N}$.

By $N_k = |\mathcal{N}_k|$, we denote the number of nodes in partition k, and by L_{km}, the number of links between nodes from partition k and m (and twice the number of links if $k = m$, see Table I). For the specific case of $K = N$ partitions, each node is in a separate partition, $N_k = 1$ for each partition and $L_{km} = a_{km}$. Based on a graph partitioning, the disease information of nodes within a same partition can be grouped, which results in a lower-dimensional description of the disease state and thus of the SIS process. Specifically, for any partitioning π, UMFF considers the $K \times 1$ *reduced-state vector* $\tilde{w}$ instead of the $N \times 1$ full-state vector w. In the reduced-state vector $\tilde{w}$, the vector component $\tilde{w}_k$ captures how many nodes are infected in partition k and equals

$$\tilde{w}_k = \sum_{i\in\mathcal{N}_k} w_i \quad \text{for any partition} k = 1,2,\ldots,K \tag{3}$$

so that $\tilde{w}_k$ is an integer bounded by $0 \leqslant \tilde{w}_k \leqslant N_k$. The reduced-state vector $\tilde{w}$ contains less information about the disease state than w, and is a coarser description. In other words, one reduced state $\tilde{w}$ can correspond to a number of different full states w (see also Appendix B 2). The $K \times 1$ reduced-state random vector $\widetilde{W}(t)$ is defined similarly as a simplified description of the random state vector $W(t)$, i.e., $\widetilde{W}_k = \sum_{i\in\mathcal{N}_k} W_i$. With this notation, the probability of the network to be in a certain reduced state $\tilde{w}$ is denoted by $\Pr[\widetilde{W} = \tilde{w}]$. Furthermore, the expected number of infected nodes in a partition k is given by $\mathbf{E}[\widetilde{W}_k]$. The variables that follow from grouping the disease states, together with some additional definitions, are summarized in Table I. Another important characteristic of the SIS process is the number of infective links, which are links with one end node infected and the other healthy, since the infection rate of a healthy node is proportional to the number of infected neighbors of this node. The infective links form the cut set between healthy and infected nodes in the graph [13]. In a partitioning π of the graph, the relations

$$\begin{aligned}(u-w)^T A^{(km)} w &= \sum_{i=1}^N \sum_{j=1}^N a_{ij}^{(km)}(1-w_i)w_j \\ &= \sum_{i\sim j} \mathbb{1}_{\{(1-w_{i\in\mathcal{N}_k})w_{j\in\mathcal{N}_m}\}}\end{aligned}$$

show that $(u-w)^T A^{(km)} w$ equals the number of infective links between susceptible nodes in partition k and infected nodes in partition m.

Based on the notion of a reduced state $\tilde{w}$, we present UMFF as

Definition 2 (Unified mean-field framework). Consider a graph $G(\mathcal{N},\mathcal{L})$, an SIS epidemic process with rates (β,δ) and a graph partitioning π of the nodes into K partitions. The

UMFF equations are approximate equations for the expected number of infected nodes in partition k:

$$\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} \approx -\delta \mathbf{E}[\widetilde{W}_k] + \beta \sum_{m=1}^{K} \tilde{a}_{km}(N_k - \mathbf{E}[\widetilde{W}_k])\mathbf{E}[\widetilde{W}_m]. \tag{4}$$

The UMFF equations follow from simplifying the exact SIS process description, using two approximations:

Approximation 1 (Topological approximation). The number of infective links between susceptible nodes in partition k and infected nodes in partition m are approximated by

$$(u-w)^T A^{(km)} w \approx (\tilde{u}-\tilde{w})^T \widetilde{A}^{(km)} \tilde{w} = \tilde{a}_{km}(N_k - \tilde{w}_k)\tilde{w}_m. \tag{5}$$

Approximation 2 (Moment-closure approximation). The covariance between the random variables $\widetilde{W}_k$ and $\widetilde{W}_m$ is approximated by zero:

$$\mathrm{Cov}[\widetilde{W}_k, \widetilde{W}_m] \approx 0 \Rightarrow \mathbf{E}[\widetilde{W}_k \widetilde{W}_m] \approx \mathbf{E}[\widetilde{W}_k]\mathbf{E}[\widetilde{W}_m]. \tag{6}$$

In the next section, we show how the UMFF equations are deduced from the exact SIS process description subject to approximations (5) and (6). The idea behind the topological approximation is further discussed in Sec. VI, while the moment-closure approximation is addressed in Appendix C. Finally, the UMFF equations (4) also lead to the formulation of an epidemic threshold:

Definition 3 (UMFF epidemic threshold). For a graph $G(\mathcal{N},\mathcal{L})$ and a partitioning π, the UMFF epidemic threshold τ_π obeys

$$\tau_\pi^{-1} = \lambda_{\max}(\widetilde{A}\,\mathrm{diag}(N_1, N_2, \ldots N_K)), \tag{7}$$

where $\lambda_{\max}(M)$ denotes the largest eigenvalue or spectral radius of the matrix M. Furthermore, the UMFF epidemic threshold is lower bounded by $\tau_\pi \geqslant \lambda_1^{-1}$ for any choice of partitioning the graph.

Equation (7) for the UMFF epidemic threshold and the lower bound are derived in Appendix A based on a linear stability analysis of the UMFF equations. The UMFF epidemic threshold τ_π specifies the phase transition (after sufficiently long time) between the dying-out regime when $\frac{\beta}{\delta} < \tau_\pi$, and the very long survival regime when $\frac{\beta}{\delta} > \tau_\pi$. While the description of UMFF does not distinguish between different choices of partitioning the graph, which implies that any epidemic threshold τ_π is equally valid, certain choices of partitioning are expected to lead to closer approximations of the real dynamics, and thus yield more reliable predictions for the epidemic threshold, than others. The accuracy of the UMFF approximations is further discussed in Sec. VI, where bounds are given for the topological approximation. Since every epidemic threshold is lower bounded by $\tau_\pi \geqslant \lambda_1^{-1}$, which equals the N-intertwined mean-field approximation (NIMFA) threshold [5], every UMFF threshold will predict that, for $\frac{\beta}{\delta} < \lambda_1^{-1}$, the disease dies out exponentially fast for sufficiently large time [13]. Hence, the NIMFA epidemic threshold provides a safe criterion for $\frac{\beta}{\delta}$ to ensure that the epidemic dies out.

IV. DERIVATION OF THE UNIFIED MEAN-FIELD FRAMEWORK

Figure 1 overviews the variables and approximations involved in UMFF, and how the UMFF equations are derived from the exact SIS equations. Additionally, Fig. 1 shows for which particular choices of partitioning, UMFF is equivalent to existing mean-field methods (see also Sec. V). In the next sections, we follow the variables in Fig. 1 from left to right to derive the UMFF equations.

A. Exact SIS equations

The UMFF approximation of the SIS process is based on two process variables: the reduced-state probability $\Pr[\widetilde{W}(t) = \tilde{w}]$ for each reduced state $\tilde{w}$, and the expected number of infected nodes $\mathbf{E}[\widetilde{W}_k(t)]$ for each partition k. In Appendix B 2, the reduced-state probabilities are derived based on the birth-death equations as

$$\begin{aligned}\frac{d\Pr[\widetilde{W}=\tilde{w}]}{dt} &= -\delta \sum_{k=1}^{K} \tilde{w}_k \Pr[\widetilde{W}=\tilde{w}] \\ &\quad + \delta \sum_{k=1}^{K} (\tilde{w}_k + 1)\Pr[\widetilde{W}=\tilde{w}+\tilde{e}_k] \\ &\quad - \beta \sum_{k=1}^{K}\sum_{m=1}^{K} \sum_{w \in \mathcal{W}^k_{\tilde{w}_k} \cap \mathcal{W}^m_{\tilde{w}_m}} (u-w)^T A^{(km)} w \\ &\quad \times \Pr[W=w] \\ &\quad + \beta \sum_{k=1}^{K}\sum_{m=1}^{K} \sum_{w \in \mathcal{W}^k_{(\tilde{w}_k-1)} \cap \mathcal{W}^m_{\tilde{w}_m}} (u-w)^T A^{(km)} w \\ &\quad \times \Pr[W=w]\end{aligned} \tag{8}$$

for any reduced-state vector $\tilde{w}$, where $\mathcal{W}^k_x = \{w \in \{0,1\}^N | \; w^T s_k = x\}$ is the set of all full states w with x infected nodes in partition k.

For each partition k, Appendix B 3 derives the governing equation for the expected number of infected nodes as

$$\begin{aligned}\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} &= -\delta \mathbf{E}[\widetilde{W}_k] + \beta \sum_{m=1}^{K} \sum_{\tilde{w}_k=0}^{N_k} \sum_{\tilde{w}_m=0}^{N_m} \\ &\quad \times \sum_{w \in \mathcal{W}^k_{\tilde{w}_k} \cap \mathcal{W}^m_{\tilde{w}_m}} (u-w)^T A^{(km)} w \Pr[W=w].\end{aligned} \tag{9}$$

B. Birth-death process

Appendix B 2 illustrates that the rate of the infection transitions $\tilde{w} \to \tilde{w} + \tilde{e}_k$ and $\tilde{w} - \tilde{e}_k \to \tilde{w}$, which correspond to a node in partition k being infected, depends on the number of *infective links*. The consequence is that the reduced-state probability $\Pr[\widetilde{W}=\tilde{w}]$ in (8) depends on the full-state probability $\Pr[W=w]$, which means that Eqs. (8) are not closed. This closure problem is solved by invoking the UMFF topological approximation (5)

$$(u-w)^T A^{(km)} w \approx (\tilde{u}-\tilde{w})^T \widetilde{A}^{(km)} \tilde{w},$$

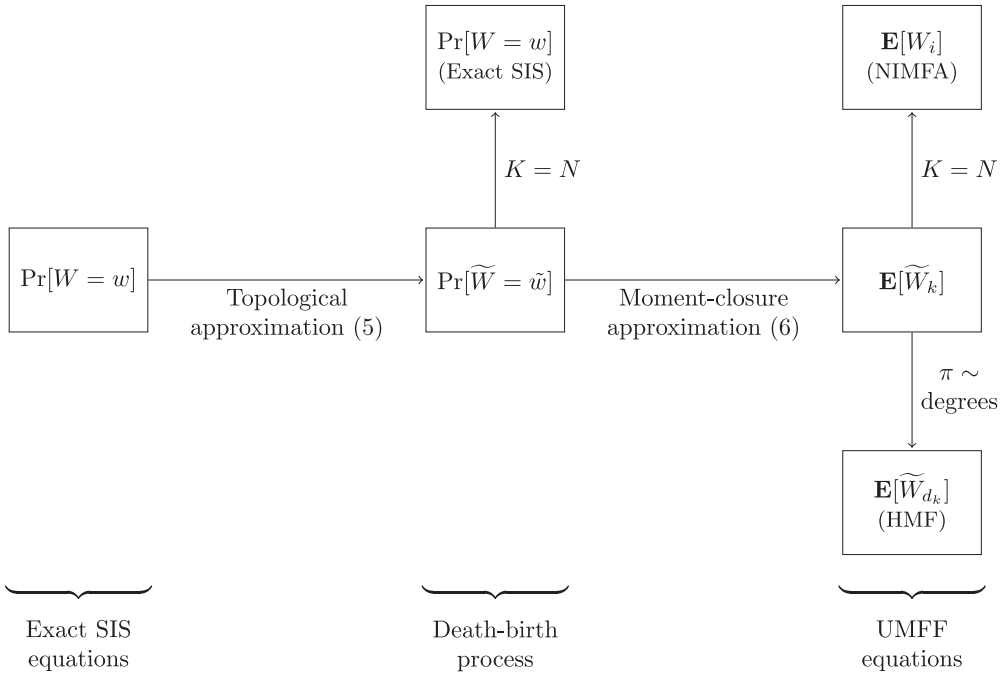


FIG. 1. Schematic representation of the relationship between the different variables involved in the UMFF approximation steps.

which enables the simplifications

$$\sum_{w\in\mathcal{W}^k_{\tilde{w}_k}\cap\mathcal{W}^m_{\tilde{w}_m}} (u-w)^T A^{(km)} w \Pr[W=w] \approx (\tilde{u}-\tilde{w})^T \widetilde{A}^{(km)} \tilde{w} \Pr[\widetilde{W}_k = \tilde{w}_k, \widetilde{W}_m = \tilde{w}_m]$$

$$\sum_{w\in\mathcal{W}^k_{(\tilde{w}_k-1)}\cap\mathcal{W}^m_{\tilde{w}_m}} (u-w)^T A^{(km)} w \Pr[W=w] \approx (\tilde{u}-(\tilde{w}-\tilde{e}_k))^T \widetilde{A}^{(km)} \tilde{w} \Pr[\widetilde{W}_k = \tilde{w}_k - 1, \widetilde{W}_m = \tilde{w}_m]. \quad (10)$$

Substituting (10) in the exact Eqs. (8) yields

$$\begin{aligned}\frac{d\Pr[\widetilde{W}=\tilde{w}]}{dt} \approx & -\delta\sum_{k=1}^{K}\tilde{w}_k \Pr[\widetilde{W}=\tilde{w}] + \delta\sum_{k=1}^{K}(\tilde{w}_k+1)\Pr[\widetilde{W}=\tilde{w}+\tilde{e}_k] \\ & -\beta\sum_{k=1}^{K}\sum_{m=1}^{K}(\tilde{u}-\tilde{w})\widetilde{A}^{(km)}\tilde{w}\Pr[\widetilde{W}_k=\tilde{w}_k,\widetilde{W}_m=\tilde{w}_m] \\ & +\beta\sum_{k=1}^{K}\sum_{m=1}^{K}(\tilde{u}-(\tilde{w}-\tilde{e}_k))\widetilde{A}^{(km)}\tilde{w}\Pr[\widetilde{W}_k=\tilde{w}_k-1,\widetilde{W}_m=\tilde{w}_m], \quad (11)\end{aligned}$$

which no longer depends on the full-state probability $\Pr[W = w]$. Although cumbersome, (11) is a closed set of equations that completely characterizes $\Pr[\tilde{W}(t) = \tilde{w}]$ for a given initial distribution $\Pr[\tilde{W}(0) = \tilde{w}]$.

Moreover, since only transitions of the form $\tilde{w} \to \tilde{w} \pm \tilde{e}_k$ and $\tilde{w} \pm \tilde{e}_k \to \tilde{w}$ exist (i.e., only single nodes are infected or cured during one event), Eq. (11) is equivalent to the description of a K-dimensional birth-death process. The reduced-state vector $\tilde{w}$ can be regarded as a coordinate in the $(N_1 + 1) \times (N_2 + 1) \times \cdots \times (N_K + 1)$ lattice for this birth-death process. Figure 2 illustrates for the specific case of $K = 1$ how the state reduction $w \to \tilde{w}$ results in a lattice transition structure, i.e., a birth-death process. Furthermore, Eq. (11) indicates that the birth rates are quadratic in $\tilde{w}$ and the death rates are linear in $\tilde{w}$, which means that the SIS process is equivalent to a higher-dimensional quadratic birth-death process. The number of infected nodes in the complete graph K_N can be exactly described as a quadratic birth-death process [12,14]. While no analytical solutions exist for the quadratic birth-death process [14], the equivalence between the SIS and the quadratic birth-death process is an interesting observation and means that insights in one setting translate directly to the other (see also Sec. IX).

C. UMFF equations

The exact Eq. (9) for the expected number of infected nodes $\mathbf{E}[\widetilde{W}_k]$ is not "closed" for two reasons: the exact SIS dynamics

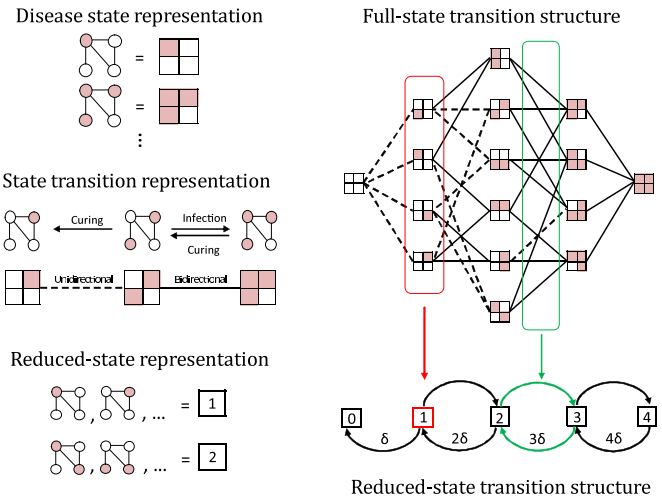


FIG. 2. Example of the transition structure of the SIS process on a four-node network. On the left, a network is shown together with its schematic representation. Sick nodes in the network are colored pink, which is encoded in the schematic representation as pink squares. State transitions are represented by connected squares. Full lines represent bidirectional transitions and dotted lines represent unidirectional transitions. The schematic representation of the reduced network disease state is simply the number of infected nodes in that state. On the right, the transition structure of the SIS process on this four-node network is shown. Already for a small network, the full-state transition structure turns out to be complex. As an illustration of the UMFF approach, the full-state transition structure is simplified to the reduced-state transition structure for a one-partition partitioning ($K = 1$). This partitioning combines all the full states with the same number of infected nodes into a single reduced state. Additionally, the reduced-state transition rates are a combination of the full-state transition rates. Appendix B 2 describes this reduction in more detail.

depend on the number of infective links (i.e., on full-state probability $\Pr[W = w]$) and on higher-order moments, i.e., the first-order moment Eqs. (9) depend on the second-order moments $\mathbf{E}[\widetilde{W}_k \widetilde{W}_m]$ (see also Appendix C). Similar to the derivation of the birth-death process, invoking the UMFF topological approximation (5) results in simplifications (10), which allows Eq. (9) to be approximated by

$$\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} = -\delta \mathbf{E}[\widetilde{W}_k] + \beta \sum_{m=1}^{K} \sum_{\tilde{w}_k=0}^{N_k} \sum_{\tilde{w}_m=0}^{N_m} (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w} \times \Pr[\widetilde{W}_k = \tilde{w}_k, \widetilde{W}_m = \tilde{w}_m]. \quad (12)$$

While the dependence on the full-state probability $\Pr[W = w]$ is solved, Eq. (12) still contains higher-order moment terms

$$\sum_{\tilde{w}_k=0}^{N_k} \sum_{\tilde{w}_m=0}^{N_m} \tilde{w}_k \tilde{w}_m \Pr[\widetilde{W}_k = \tilde{w}_k, \widetilde{W}_m = \tilde{w}_m] = \mathbf{E}[\widetilde{W}_k \widetilde{W}_m] \quad (13)$$

for partition pairs (k,m). In general, these second-order moments $\mathbf{E}[\widetilde{W}_k \widetilde{W}_m]$ cannot be determined from $\mathbf{E}[\widetilde{W}_k]$ and $\mathbf{E}[\widetilde{W}_m]$ alone. Invoking the UMFF moment-closure approximation (6)

$$\mathrm{Cov}[\widetilde{W}_k, \widetilde{W}_m] \approx 0 \Rightarrow \mathbf{E}[\widetilde{W}_k \widetilde{W}_m] \approx \mathbf{E}[\widetilde{W}_k]\mathbf{E}[\widetilde{W}_m]$$

solves this closure problem by enabling Eq. (12) to be approximated by

$$\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} \approx -\delta \mathbf{E}[\widetilde{W}_k] + \beta \sum_{m=1}^{K} \tilde{a}_{km} (N_k - \mathbf{E}[\widetilde{W}_k])\mathbf{E}[\widetilde{W}_m],$$

which are the UMFF equations (4). In Appendix C an extension of the UMFF equations for higher-order moments is described. These higher-order equations are more general, but a detailed description is beyond the focus of this article.

1. Bounds on the moment-closure approximation

For the particular case of $K = N$ partitions (for which UMFF is equivalent to NIMFA, see Sec. V), the infection probabilities of nodes are non-negatively correlated [15], i.e., $\mathrm{Cov}[\widetilde{W}_k, \widetilde{W}_m] \geqslant 0$. Based on the definition of the covariance

$$\mathrm{Cov}[\widetilde{W}_k, \widetilde{W}_m] = \mathbf{E}[\widetilde{W}_k \widetilde{W}_m] - \mathbf{E}[\widetilde{W}_k]\mathbf{E}[\widetilde{W}_m] \quad (14)$$

we can rewrite Eq. (12), which is only exact in the $K = N$ partitioning, as

$$\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} = -\delta \mathbf{E}[\widetilde{W}_k] + \beta \sum_{m=1}^{K} \tilde{a}_{km} (N_k - \mathbf{E}[\widetilde{W}_k])\mathbf{E}[\widetilde{W}_m] - \beta \sum_{m=1}^{N} \tilde{a}_{km} \mathrm{Cov}[\widetilde{W}_k, \widetilde{W}_m]. \quad (15)$$

Omitting the last, negative sum in Eq. (15) implies that for $K = N$ partitions, the moment-closure approximation is an upper bound of the true process. However, for any other partitioning ($K \neq N$), we do not know about any such results for $\mathrm{Cov}[\widetilde{W}_k, \widetilde{W}_m]$. In other words, we do not know how to bound the UMFF moment-closure approximation error.

V. EXISTING MEAN-FIELD METHODS CONTAINED BY UMFF

An important feature of UMFF is that by particular choices of graph partitioning, the UMFF equations are equivalent to existing mean-field methods. In particular, the widely used N-intertwined mean-field approximation [5] and heterogeneous mean-field approximation [6] are contained by UMFF. Additionally, by the higher-order extension of UMFF described in Appendix C, also second-order NIMFA [15] and pair quenched mean-field theory [16] are contained by (higher-order) UMFF.

A. N-intertwined mean-field approximation (NIMFA)

The N-intertwined mean-field approximation [5] incorporates the full topological information of the graph. The only approximation consists of assuming independence between the infection states of adjacent nodes, i.e., the moment-closure approximation (6). Denoting the infection probability of node k by $\rho_k = \Pr[W_k = 1]$, the NIMFA equations for $1 \leqslant k \leqslant N$ are given by [5]

$$\frac{d\rho_k}{dt} = -\delta \rho_k + \sum_{m=1}^{N} \beta a_{km} (1 - \rho_k) \rho_m. \quad (16)$$

The same NIMFA equations (16) are retrieved from UMFF with $K = N$ partitions, which corresponds to each node being

in a separate partition. The expected number of infected nodes in a partition $\mathbf{E}[\widetilde{W}_k]$ is then equal to the infection probability ρ_k of node k that constitutes that partition. The $K = N$ partitioning, where $N_k = 1$ and $\widetilde{A} = A$, illustrates that the NIMFA equations (16) are indeed a particular case of the UMFF equations (4).

B. Heterogeneous mean-field method (HMF)

Pastor-Satorras and Vespignani [6] introduced the heterogeneous mean-field method, which approximates the SIS process assuming that all nodes of a certain degree are equivalent (in their connections with other nodes). Consequently, the SIS process is described based on the degree distribution of the underlying graph. Differently from UMFF and NIMFA, HMF [6] does not assume a known graph G, but rather considers a class of graphs. Specifically, in HMF the epidemic is assumed to spread on a graph with a specified degree distribution and with the link probability between pairs of nodes independent of their degrees. For each degree $d_1 \leqslant d_k \leqslant d_K$, the probability distribution $\Pr[D = d_k]$ denotes the probability that a randomly chosen node has degree d_k. The variable $0 \leqslant \tilde{\rho}_k \leqslant 1$ reflects the expected fraction of infected nodes with degree d_k, leading to the HMF [6] equations:

$$\frac{d\tilde{\rho}_k}{dt} = -\delta\tilde{\rho}_k + \beta d_k(1-\tilde{\rho}_k)\Theta, \tag{17}$$

where Θ is the probability that a healthy node is linked to an infected node and calculated in Ref. [6] as

$$\Theta = \sum_{m=1}^{K} \tilde{\rho}_m \frac{d_m \Pr[D = d_m]}{\sum_{i=1}^{K} d_i \Pr[D = d_i]}. \tag{18}$$

Substituting expression (18) for Θ in (17) gives

$$\frac{d\tilde{\rho}_k}{dt} = -\delta\tilde{\rho}_k + \beta \sum_{m=1}^{K} \frac{d_k d_m \Pr[D = d_m]}{\sum_{i=1}^{K} d_i \Pr[D = d_i]}(1-\tilde{\rho}_k)\tilde{\rho}_m.$$

Introducing the variable $\rho_k = \Pr[D = d_k]\tilde{\rho}_k$ then yields

$$\frac{d\rho_k}{dt} = -\delta\rho_k + \beta \sum_{m=1}^{K} \frac{d_k d_m}{\sum_{i=1}^{K} d_i \Pr[D = d_i]} \times (\Pr[D = d_k] - \rho_k)\rho_m. \tag{19}$$

While Eqs. (19) are derived in HMF for a probabilistic graph, the same equations are found from UMFF for a particular graph with the same degree distribution, so that the number of nodes with degree k equals $N_k = c \Pr[D = d_k]$ for some scalar $c \in \mathbb{R}$, and with degree-uncorrelated links. For such a graph, the number of links $L_{km} = s_k^T A s_m$ between nodes of degree d_k and degree d_m obeys the consistency relation $\sum_{m=1}^{K} L_{km} = N_k d_k$ as

$$L_{km} = \frac{d_k d_m N_k N_m}{\sum_{i=1}^{K} d_i N_i},$$

from which the UMFF equations follow as

$$\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} = -\delta\mathbf{E}[\widetilde{W}_k] + \beta \sum_{m=1}^{K} \frac{d_k d_m}{\sum_{i=1}^{K} d_i N_i} \times (N_k - \mathbf{E}[\widetilde{W}_k])\mathbf{E}[\widetilde{W}_m]. \tag{20}$$

Equations (20) are equivalent to (19) for the scaling $\mathbf{E}[\widetilde{W}_k] = c\rho_k$, where c is the same scalar relating N_k to $\Pr[D = d_k]$. Hence, the HMF equations are found from the UMFF framework by considering a specific graph realization consistent with the random graph properties assumed by HMF.

Boguñá and Pastor-Satorras [7] extend the HMF model to random graphs with correlated degrees. Instead of only assuming $\Pr[D = d_k]$, also the probability $\Pr[i \sim j | i \in \mathcal{N}_k, j \in \mathcal{N}_m]$ that a node i of degree d_k links with a node j of degree d_m is assumed to be known for any pair of degrees (d_k, d_m). With these extra assumptions in the HMF methodology, the SIS process is then approximately described based on the degree distribution and the linking probabilities. If we now consider a specific graph realization with $N_k = c_1 \Pr[D = d_k]$ nodes of degree d_k and with $L_{km} = c_2 \Pr[i \sim j | i \in \mathcal{N}_k, j \in \mathcal{N}_m]$ links between nodes with degree d_k and d_m (for some scalars $c_1, c_2 \in \mathbb{R}$), then again the UMFF equations (4) are equivalent to the correlated HMF equations. In the same way that the HMF equations are fully determined by the degree distribution and the linking probabilities, also the UMFF equations are fully determined by N_k and L_{km}.

Since (correlated) HMF is a particular case of UMFF, HMF implicitly assumes the UMFF moment-closure approximation (6) with respect to the partitioning according to node degree. As discussed in Sec. IV C, this means that, apart from simulation results [17], we do not know in general whether HMF upper- or lower-bounds the infection probabilities, nor how HMF relates to the exact SIS process in general. But a consequence of the equivalence between UMFF (4) and (correlated) HMF (19) is that we can bound the topological approximation errors of HMF (with respect to a specific realization of the probabilistic graph model).

Since the partitions $\mathcal{N}_k$ do not need to correspond to node degrees specifically, UMFF enables the description of SIS dynamics for a wider range of graph classes. For any graph model, where a probability distribution $\Pr[K = k]$ of a node belonging to partition $\mathcal{N}_k$ is given, together with a linking probability $\Pr[i \sim j | i \in \mathcal{N}_k, j \in \mathcal{N}_m]$, the UMFF equations can be directly found. Such graph models are more general than graphs with degree-based partitions only and, in some settings, a specific structure in the graph might suggest a natural way to partition the nodes such that grouped nodes have a similar connectivity to the rest of the network (see also further directions in Sec. IX).

C. Second-order NIMFA and pair quenched mean-field theory

As described in Appendix C, second-order NIMFA (sNIMFA) [18] is a second-order extension of NIMFA, approximating the joint probability $\Pr[W = w]$ by first- and second-order moments $\mathbf{E}[W_i]$ and $\mathbf{E}[W_i W_j]$ for all nodes $i \neq j$. Similarly, pair quenched mean-field theory (pQMF) [16] is an extension of quenched mean-field theory (QMF) [19], which is an SIS approximation method introduced to investigate the epidemic threshold. The extension QMF $\rightarrow$ pQMF is conceptually the same as NIMFA $\rightarrow$ sNIMFA, but a different moment-closure approximation approximates the third-order moments. Both sNIMFA as well as pQMF are contained by the higher-order UMFF equations (C1), for $K = N$ partitions and order $n = 2$ if the generic moment-

closure approximation is chosen as in Refs. [15] and [16], respectively.

VI. THE ISOPERIMETRIC PROBLEM IN SIS EPIDEMICS

In this section, we focus on the UMFF topological approximation (5). We first describe how the closure problem of Eqs. (8) and (9) can be related to the isoperimetric problem. Then, we show how this analogy leads to approximation (5) and bounds on the approximation error.

A. The isoperimetric problem

The isoperimetric problem is an ancient problem that has interested many mathematicians throughout history. For the most basic form of the isoperimetric problem, we refer to Blåsjö [9], who provides a broad historical and conceptual overview of the isoperimetric problem:

Problem 1 (The isoperimetric problem). Among all figures in the plane with a given perimeter P, which one encloses the greatest area A?

Theorem 1 (The isoperimetric theorem). The solution to the isoperimetric problem is the circle of perimeter P.

Theorem 2 (The isoperimetric inequality). For all figures with a given perimeter P and area A, it holds that $P^2 - 4\pi A \geqslant 0$ and equality only occurs for the circle.

While the question in problem 1 might seem simple and its solution intuitive, it took until the 20th century to rigorously prove the isoperimetric theorem. After the extensive historical study of the isoperimetric problem in the 2D plane, similar problems were studied in different geometric contexts. The basic interest in these problems always consisted of describing the relationship between the *volume* and *surface* of a certain object, leading to isoperimetric inequalities of the form

$$\theta_{\min} \leqslant f(\text{volume}) + g(\text{surface}) \leqslant \theta_{\max}. \tag{21}$$

For instance, Osserman [10] describes isoperimetric inequalities in higher dimensions, on curved surfaces and on general Riemannian manifolds. The geometric context of interest for UMFF, is the study of the isoperimetric problem on graphs (see, for instance, Ref. [20]).

B. Infective links and infected nodes: An isoperimetric analogy

The dynamics of SIS epidemics are governed by two processes: infected nodes are cured and infection takes place on infective links, i.e., the links between healthy and infected nodes. *The curing process is proportional to the number of infected nodes while the infection process is proportional to the number of infective links*. In a nontechnical way, we can associate the number of infected nodes to a volume on the graph, while the infective links accord to a surface or interface around the infected volume, as illustrated by Fig. 3. *The curing process is then proportional to the infected volume, while the infection process is proportional to the infective surface*.

To use the concepts of volume and surface adequately, we must define a unit of volume and surface in the context of graphs: we define a set of one node to have unit volume, and a set of one link to have unit surface. Other choices are possible, e.g., the volume of a node being proportional to its degree, but for the purpose of deriving and bounding the UMFF topological approximation (5), this would be a less natural choice.

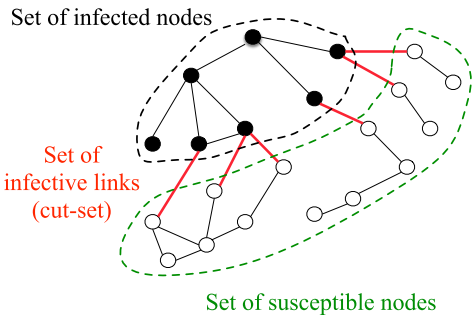


FIG. 3. Schematic of the disease state in a network. The infected and healthy nodes determine two separate partitions separated by the cut set, the set of infective links.

In the derivation of the exact reduced-state Kolmogorov equations (8), the transition rate between reduced states depends on the number of infective links. Hence, the exact equations (8) for $\Pr[\widetilde{W} = \tilde{w}]$ and (9) for $\mathbf{E}[\widetilde{W}]$ are not closed, because they contain terms of the form $(u - w)^T A^{(km)} w \Pr[W = w]$. In the language of the isoperimetric problem, this *closure problem* translates to the volume equations (8) and (9) containing terms related to the surface. The UMFF topological approximation (5) replaces the surface term by a function of volume terms and thus solves the closure problem. Now, by analogy with the isoperimetric problem, we can bound the approximation error caused by this replacement, as shown in Fig. 4, where ϵ represents the introduced error.

It remains to find the correct translation of the isoperimetric inequality into the setting of SIS epidemics. The UMFF topological approximation is defined as (5)

$$(u - w)^T A^{(km)} w \approx (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w},$$

which we can rewrite by introducing an error term $\epsilon \in \mathbb{R}$ as

$$(u - w)^T A^{(km)} w = (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w} + \epsilon, \tag{22}$$

or, by upper-bounding the error term $|\epsilon| \leqslant \theta$, as

$$|(u - w)^T A^{(km)} w - (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w}| \leqslant \theta. \tag{23}$$

In the next subsection, we specify the error bound θ based on the isoperimetric inequalities on graphs. More than just providing an error bound, the analogy with the isoperimetric problem and the mathematical techniques in the proofs (see Appendix D) also provide a motivation for the specific form of the UMFF topological approximation (5).

C. Isoperimetric inequalities for the number of infective links

The bound for the approximation error is based on the isoperimetric and discrepancy inequalities of Chung [20]:

Theorem 3 (General-graph isoperimetric inequality). For a graph $G(\mathcal{N}, \mathcal{L})$ and a partitioning π, the error of the UMFF topological approximation (5) between any two partitions k and m is bounded as

$$|(u - w)^T A^{(km)} w - (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w}| \leqslant \frac{\theta}{N} \sqrt{\tilde{w}_m (N - \tilde{w}_m)(N_k - \tilde{w}_k)[N - (N_k - \tilde{w}_k)]}, \tag{24}$$

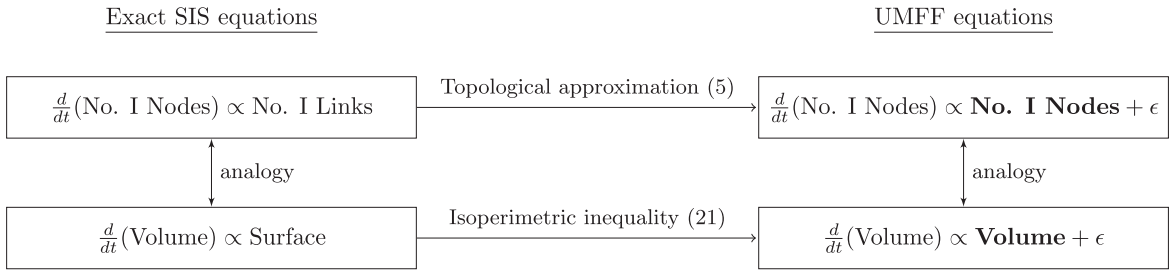


FIG. 4. Conceptual diagram depicting the analogy between the UMFF topological approximation (5) and the isoperimetric inequality (21).

where $|\tilde{a}_{km} - \mu_i| \leqslant \theta$ holds for $1 \leqslant i < N$, with μ_i the eigenvalues of the Laplacian matrix corresponding to adjacency matrix A.

For biregular graphs $A^{(km)}$, meaning that $A^{(km)}s_m = c_1 s_m$ and $s_k^T A^{(km)} = c_2 s_k^T$ for some constants $c_1, c_2 \in \mathbb{R}$, a tighter bound can be given based on interlacing techniques of Haemers [21]:

Theorem 4 (Biregular-graph isoperimetric inequality). For a graph $G(\mathcal{N},\mathcal{L})$ and a partitioning π such that $A^{(km)}$ is biregular for some partitions k and m, the error of the UMFF topological approximation (5) is bounded as

$$|(u-w)^T A^{(km)} w - (\tilde{u}-\tilde{w})^T \tilde{A}^{(km)} \tilde{w}| \leqslant \frac{\lambda_2}{\sqrt{N_k N_m}} \sqrt{\tilde{w}_k (N_k - \tilde{w}_k) \tilde{w}_m (N_m - \tilde{w}_m)}, \quad (25)$$

where λ_2 is the second-largest eigenvalue of $A_{km,r} = A^{(km)} + A^{(mk)}$.

Theorem 3 and Theorem 4, proved in Appendix D, rely heavily on proofs given by Chung [20] and Haemers [21].

VII. UMFF AND SZEMERÉDI'S REGULARITY LEMMA

The isoperimetric problem is a well-studied mathematical problem that appears in many different fields, including graph theory and network science, and thus provides a conceptual link between those fields. For instance, Szemerédi's regularity lemma (SRL) is a lemma with interesting implications for UMFF, which follows from the relation of both UMFF and SRL with the isoperimetric problem. We will discuss how SRL may indicate for which graphs the UMFF topological approximation (5) is expected to be accurate, and for which the SIS dynamics are thus well approximated by the UMFF equations.

A. Szemerédi's regularity lemma

The following definitions and interpretations are based on Diestel's [22] description of SRL. We start by defining a so-called regularity condition between pairs of partitions, which is related to the isoperimetric inequality.

Definition 4 (ϵ-regular partition pair). [22] Consider a graph $G(\mathcal{N},\mathcal{L})$ and two disjoint node partitions $\mathcal{N}_k, \mathcal{N}_m \subseteq \mathcal{N}$. If for any pair of subsets $\mathcal{N}_x \subseteq \mathcal{N}_k$ and $\mathcal{N}_y \subseteq \mathcal{N}_m$ of size N_x and N_y with $N_x \geqslant \epsilon N_k$ and $N_y \geqslant \epsilon N_m$ for some real $\epsilon > 0$, the inequality

$$\left| \frac{s_x^T A^{(km)} s_y}{N_x N_y} - \frac{s_k^T A^{(km)} s_m}{N_k N_m} \right| \leqslant \epsilon \quad (26)$$

holds, then we say that the partition pair (k,m) is ϵ-regular.

Inequality (26) can be rewritten as

$$\left| s_x^T A^{(km)} s_y - \frac{L_{km}}{N_k N_m} N_x N_y \right| \leqslant \epsilon N_x N_y, \quad (27)$$

which shows that the regularity condition (26) is related to the difference between the size of the cut set $s_x^T A^{(km)} s_y$ (for all subsets of partitions k,m with N_x, N_y nodes, respectively) and the approximate size of the cut set: $\frac{L_{km}}{N_k N_m} N_x N_y$. For lower values of ϵ, the regularity condition becomes stronger. First, because the true size of the cut set can deviate less from the approximate cut-set size if ϵ is smaller, and secondly, because the regularity condition must hold for a larger range of subsets $(\mathcal{N}_x, \mathcal{N}_y)$, since $N_x \geqslant \epsilon N_k$ is a less stringent condition if ϵ is lower (and similarly for $\mathcal{N}_y$). Based on the notion of ϵ-regular partition pairs, we define a regularity condition on a partitioning π of a graph:

Definition 5 (ϵ-regular graph partitioning). [22] Consider a graph $G(\mathcal{N},\mathcal{L})$ with a partitioning π of the nodes into $K+1$ partitions $\{\mathcal{N}_0, \mathcal{N}_1, \ldots, \mathcal{N}_K\}$. Such a graph partitioning is called ϵ-regular if it meets the following conditions:

(i) $N_0 \leqslant \epsilon N$

(ii) $N_1 = N_2 = \cdots = N_K$

(iii) All except at most ϵK^2 of the partition pairs (k,m) for $1 \leqslant k < m \leqslant K$ are ϵ-regular.

Roughly speaking, a graph partitioning is ϵ-regular if it contains K equally sized partitions (ii) such that most partition pairs are regular (iii), where one additional "small" partition is allowed to exist (i) on which conditions (ii) and (iii) do not apply. For a given K, a smaller ϵ strengthens the regularity conditions. First, because the regularity condition between partition pairs becomes stronger, second, because $N_0 \leqslant \epsilon N$ means that a lower number of nodes are allowed to make up the "leftover partition" $\mathcal{N}_0$ and, finally, because ϵK^2 becomes smaller, implying that an increasing proportion of the partition pairs need to satisfy the regularity condition (26). Since condition (iii) holds for partition pairs (k,m) with $k \neq m$, the regularity condition applies only to links between partitions and not within partitions. Based on the regularity notion of a graph partitioning, Szemerédi's regularity lemma is a statement about the possibility of finding a regular partitioning in arbitrary graphs, with a number K of partitions effectively independent of the size N of the graph.

Definition 6 (Szemerédi's regularity lemma). For every $\epsilon > 0$ and every integer $K_{\min} \geqslant 1$, there exists an integer $K_{\max}$ such that every graph on $N \geqslant K_{\min}$ nodes admits an ϵ-regular graph partitioning in K partitions, with $K_{\min} \leqslant K \leqslant K_{\max}$.

The proof of SRL can be found in Diestel [22]. We exemplify the lemma: if we take a certain ϵ and choose

$K_{\min} = 10$, then SRL states that there is an integer $K_{\max}$, such that for any graph with $N > 10$ nodes there exists an ϵ-regular partitioning of $10 \leqslant K \leqslant K_{\max}$ partitions. While for $N \leqslant K_{\max}$, the existence of an ϵ-regular partitioning automatically holds by choosing the $K = N$ partitioning, the result becomes stronger for $N > K_{\max}$. For very large graphs, i.e., $N \gg K_{\max} \geqslant K$, SRL states that it is always possible to have an ϵ-regular K-partitioning. An interesting interpretation of SRL is due to Tao [23] who states that, roughly speaking: "SRL can be viewed as a structure theorem for large dense graphs, approximating such graphs to any specified accuracy by objects, whose complexity is bounded independently of the number of nodes in the original graph." Applied to UMFF, this means that, for any large dense graph and any desired accuracy ϵ, there exists a partitioning in $K \ll N$ partitions, such that the topological approximation of UMFF between most (k,m) partition pairs ($k \neq m$) is ϵ-accurate, in the sense that (k,m) are ϵ-regular partition pairs. While a regular graph partitioning does not imply any regularity conditions on the within-partition links, Diestel [22] mentions that by choosing $K_{\min}$ large "we may increase the proportion of links running between different partition sets (rather than inside one), i.e., the proportion of links that are subject to the regularity assertion." In other words, if we take $K_{\min}$ large enough for a given ϵ, then most links will be *between* partitions (rather than *within*) and will thus satisfy the regularity conditions.

B. Implications of SRL for UMFF

We believe that SRL can be translated to a statement about the scaling behavior of the SIS process on large graphs. We will describe the conceptual idea here, realizing that a more rigorous investigation would be necessary to proof any of the claims. Since the regularity inequality (26) can be rewritten as (27), which has the same form as the isoperimetric inequality, the ϵ-regularity of a partition pair also implies that the UMFF topological approximation (5) has an ϵ-bounded approximation error (for subsets of sufficiently large size). For an ϵ-regular graph partitioning with $K + 1$ pairs, this isoperimetric interpretation then means that, for most of the partition pairs ($\geqslant \epsilon K^2$), the UMFF topological approximation error is ϵ-bounded. Finally, SRL indicates that for any chosen accuracy ϵ and sufficiently large minimum number of partitions $K_{\min}$, an integer $K_{\max}$ exists such that for any graph on $N \geqslant K_{\min}$ nodes, a partitioning can be found with $K_{\min} \leqslant K \leqslant K_{\max}$ partitions, such that most links are between partitions and most of the partition pairs have ϵ-bounded approximation errors. Applied to UMFF, this means that *for large graphs on N nodes, a partitioning in $K_{\min} < K \ll N$ partitions can always be found such that the UMFF topological approximation between most partition pairs is bounded by a chosen ϵ, where choosing a large enough $K_{\min}$ results in most links being between partitions* (by Diestel's argument). While the isoperimetric bounds on the UMFF approximations are defined for any possible partition, the relation with SRL leads to a statement about how good these isoperimetric bounds can become. An important difference between SRL regularity and the UMFF approximation error is that SRL regularity only holds for subsets of size $N_x \geqslant \epsilon N_k$, where $N_k \approx \frac{N}{K}$. Hence, the regularity weakens for growing N, because it no longer holds for cut sets between small subsets. The consequence for UMFF is that the regularity, and thus the boundedness of the topological approximation error, only holds if a sufficiently large fraction of nodes is infected in both partitions. Thus, the dynamics are well approximated by lower-dimensional dynamics, only for disease states w where enough nodes are infected between any pair of partitions, i.e., $\tilde{w}_k \geqslant \epsilon N_k$ and $\tilde{w}_m \geqslant \epsilon N_m$.

VIII. RELATED WORK

A. NIMFA on graphs with an equitable partitioning

Bonaccorsi *et al.* [24] study the NIMFA equations on graphs with an equitable partitioning. A partitioning π is equitable if the subgraph between any two (possibly the same) partitions, is biregular (regular). If a graph has such an equitable partitioning, and the initial infection probability is the same for all nodes within one partition, then the NIMFA equations for the SIS process on that graph can be exactly described by K rather than N equations [24]. This result follows from the observation that equality in the UMFF topological approximation (5) holds, i.e.,

$$(u - w)^T A^{(km)} w = (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w} = \tilde{a}_{km}(N_k - \tilde{w}_k)\tilde{w}_m,$$

when $A^{(km)}$ is biregular, and that

$$\begin{aligned}\Pr[W(0) = w] = {} & \left|\mathcal{W}^k_{\tilde{w}_k} \cap \mathcal{W}^m_{\tilde{w}_m}\right|^{-1} \\ & \times \Pr[\widetilde{W}_k(0) = \tilde{w}_k, \widetilde{W}_m(0) = \tilde{w}_m] \\ & \forall w \in \mathcal{W}^k_{\tilde{w}_k} \cap \mathcal{W}^m_{\tilde{w}_m}\end{aligned}$$

holds, when nodes from the same partition have equal initial infection probabilities. Hence, the main point of Ref. [24] is that for this specific type of graph and initial condition, the number of infective links between any two partitions only depends on the number of infected nodes in those partitions, which enables a lower-dimensional description of the SIS process (within the NIMFA approximation). This result is based on similar ideas as the UMFF framework, but from a very different perspective: UMFF describes how the topological approximation (5) applied to *any graph*, followed by a moment-closure approximation (6), results in a lower-dimensional *approximate* description of the SIS process.

B. Approximating the number of infective links in SIS

A central concept of UMFF is the description of the topological approximation (5) from the perspective of the isoperimetric problem. This approach of approximating the SIS process by approximating the number of infective links has appeared before. Ganesh *et al.* [25] find an upper bound for the epidemic threshold, by relating the infection terms in the SIS process to the isoperimetric problem. The isoperimetric or Cheeger constant [11] of a graph with adjacency matrix A is defined as

$$\eta_c(A) = \min_{w \in \{0,1\}^N} \frac{(u - w)^T A w}{w^T w},$$

which leads to a lower bound for the number of infective links as

$$(u-w)^T A w \geqslant \eta_c(A)\tilde{w} \tag{28}$$

for any $w \in \{0,1\}^N$ and where $\tilde{w} = w^T w$ is the number of infected nodes. By assuming equality in (28), the SIS process is approximated by a linear birth-death process, from which an approximate epidemic threshold is derived in Ref. [25].

Van Mieghem [13,26] also approximated the SIS process by approximating the size of the cut set. Rather than relying on the isoperimetric problem, the most dominant terms in the spectral decomposition of the quadratic form $w^T Q w$, which equals the number of infective links, approximate the cut set. Specifically, the approximation

$$(u-w)^T A w \approx \frac{\mu_{N-1}}{N}\tilde{w}(N-\tilde{w})$$

is made. If this approximation error can be bounded by a constant $\theta \in \mathbb{R}$, i.e.,

$$\left|(u-w)^T A w - \frac{\mu_{N-1}}{N}\tilde{w}(N-\tilde{w})\right| \leqslant \theta, \tag{29}$$

then the exact equation for the expected number of infected nodes can be bounded as

$$\mathbf{E}[\widetilde{W}_{-\theta}(t)] \leqslant \mathbf{E}[\widetilde{W}_{\text{exact}}(t)] \leqslant \mathbf{E}[\widetilde{W}_{+\theta}(t)], \tag{30}$$

where the bounds follow from the differential equations:

$$\begin{aligned}
\frac{d\mathbf{E}[\widetilde{W}_{+\theta}(t)]}{dt} &= -\delta\mathbf{E}[\widetilde{W}_{+\theta}] \\
&\quad + \beta\frac{\mu_{N-1}}{N}\mathbf{E}[\widetilde{W}_{+\theta}](N-\mathbf{E}[\widetilde{W}_{+\theta}]) + \theta \\
\frac{d\mathbf{E}[\widetilde{W}_{-\theta}(t)]}{dt} &= -\delta\mathbf{E}[\widetilde{W}_{-\theta}] \\
&\quad + \beta\frac{\mu_{N-1}}{N}\mathbf{E}[\widetilde{W}_{-\theta}](N-\mathbf{E}[\widetilde{W}_{-\theta}]) - \theta,
\end{aligned} \tag{31}$$

which are Riccati differential equations, whose analytic solutions are known and have a hyperbolic-tangent form [26]. In other words, the method of Refs. [26] and [13] gives bounds on the exact expected number of infected nodes $\mathbf{E}[\widetilde{W}_{\text{exact}}(t)]$, if a constant bound θ on the approximation error (29) is known.

By filling in $c = \mu_{N-1}$ in Lemma 1 from Appendix D, we can show that $\theta \leqslant \frac{N(\mu_1-\mu_{N-1})}{4} = \theta^\star$. Although not a tight bound, filling in $\theta = \theta^\star$ in Eqs. (31) gives

$$\mathbf{E}[\widetilde{W}(t)_{-\theta^\star}] \leqslant \mathbf{E}[\widetilde{W}_{\text{exact}}(t)] \leqslant \mathbf{E}[\widetilde{W}(t)_{+\theta^\star}],$$

which is a new result based on the spectral decomposition methodology of Refs. [26] and [13].

IX. SUMMARY

We have introduced a novel and unified approximation framework UMFF for the continuous-time Markovian SIS process on complex networks, whose main features are the following:

(1) UMFF unifies and generalizes a number of existing mean-field approximations for SIS epidemics on complex networks. In particular, two widely used approximations, the N-intertwined mean-field approximation [5] and the heterogeneous mean-field method [6] are shown to be contained by UMFF.

(2) The accuracy of UMFF and of all its contained methods can be assessed based on the isoperimetric analogy, which provides bounds on the error of the UMFF topological approximation (Theorem 3).

(3) UMFF conceptually describes the scaling behavior of SIS epidemics on large graphs. Since the UMFF accuracy is related to the regularity notion of Szemerédi's regularity lemma (SRL), we can translate the statements of SRL about the structural regularity of large graphs to statements about the possibility to accurately approximate SIS dynamics on large graphs by a lower-dimensional description.

A. Future directions

By providing a unified description of mean-field approximation techniques for the SIS process, UMFF offers a framework, in which the existing techniques can be compared and which enables their respective accuracy to be assessed. In principle, UMFF could prescribe which existing (or new) mean-field method is more suitable, for a certain graph and for a specific SIS process parameter of interest.

While derived for SIS epidemics, the UMFF approach is applicable to more general epidemic models, such as the generalized epidemic mean-field model (GEMF) in Ref. [27], which generalizes NIMFA to any number of compartments and with a general transition structure between different compartments. The global dynamics of GEMF follow from *node-based* compartmental transitions and *edge-based* compartmental transitions, which translates to *volume-based* transitions and *surface-based* transitions in context of the isoperimetric problem. Hence, by exploiting the same problem structure and the isoperimetric analogy, UMFF could generalize GEMF in a similar vein as UMFF generalizes NIMFA for the SIS compartmental process.

The general partitioning feature of UMFF also creates the possibility to develop new approximation techniques for the SIS process. Specifically, if nodes can be grouped in partitions based on some parameter such that similarity in that parameter corresponds to similarity in connectivity, then UMFF is expected to yield a good approximation of the SIS process. For instance, in the embedding of graphs in metric spaces as in Refs. [28] and [29], similar spatial coordinates between a pair of nodes means that their distance to other nodes is also similar. Hence, for such graph models, spatial closeness of nodes seems to provide a good partitioning criterion for UMFF, and the coarse graining of the infection state would then correspond to the intuitively attractive notion of spatial similarity.

Furthermore, the observation that both the exact and approximate Markovian SIS processes are equivalent to a higher-dimensional quadratic birth-death process opens up new perspectives on modeling the SIS process. Some questions about the epidemic process have tractable solutions if properly formulated in terms of birth-death processes. Ganesh *et al.* [25] characterized the disease die-out probability of the SIS process, based on the gambler's ruin problem [12] of a birth-death process. Conversely, the knowledge about the

epidemic process might provide valuable insights in the quadratic birth-death process, whose exact solution is still an open problem [14].

APPENDIX A: EPIDEMIC THRESHOLD

From the UMFF equations (4), the epidemic threshold τ_π can be found based on a linear stability analysis. In order to use the standard approach for mean-field equations, for instance as described by Boguñá and Pastor-Satorras in Ref. [7], we introduce a change of variables. Instead of the expected number of infected nodes $\mathbf{E}[\widetilde{W}_k]$ in a partition k, we will consider the expected fraction of infected nodes $\rho_k = \frac{\mathbf{E}[\widetilde{W}_k]}{N_k}$ in that partition. The UMFF equations (4) then reduce to

$$\frac{d\rho_k}{dt} = -\delta\rho_k + \beta \sum_{m=1}^{K} L_{km}(1-\rho_k)\rho_m \tag{A1}$$

for every partition $k \in \{1,2,\ldots,K\}$. Equation (A1) can be linearized around the all-healthy state $\boldsymbol{\rho} = \mathbf{0}$ as follows:

$$\frac{d\rho_k}{dt} \approx \sum_{m=1}^{K} J_{km}\rho_m \quad \text{with}$$

$$J_{km} = \left(\frac{\partial}{\partial \rho_m} \left(-\delta\rho_k + \beta \sum_{j=1}^{K} L_{kj}(1-\rho_k)\rho_j \right) \right)_{\boldsymbol{\rho}=\mathbf{0}}, \tag{A2}$$

where the vector $\boldsymbol{\rho} = [\rho_1,\rho_2,\ldots,\rho_K]^T$ contains the infection fraction of all partitions, and where J_{km} are matrix elements of the Jacobian J of Eqs. (A1). The linearized Eq. (A2) reads in matrix form

$$\frac{d\boldsymbol{\rho}}{dt} \approx J\boldsymbol{\rho}. \tag{A3}$$

Equation (A3) indicates that the all-healthy state is a stable point of the UMFF equations (4) if the Jacobian J has a negative largest eigenvalue, while a positive largest eigenvalue of J means that the all-healthy state is not a stable point. Translated to the setting of the SIS epidemic process, a positive largest eigenvalue $\lambda_{\max}(J) > 0$ reflects that the epidemic disease will spread over the network, while $\lambda_{\max}(J) < 0$ correspond to a die-out of the epidemic [1]. Calculating the Jacobian elements J_{km} from Eq. (A2) yields

$$J_{km} = -\delta \mathbb{1}_{\{k=m\}} + \beta \frac{L_{km}}{N_k}.$$

Using the matrix elements $\tilde{a}_{km} = \frac{L_{km}}{N_k N_m}$, the Jacobian matrix J becomes

$$J = -\delta I + \beta \widetilde{A}\,\mathrm{diag}(N_1,N_2,\ldots,N_K). \tag{A4}$$

From Eq. (A4), it follows that $\tau_\pi^{-1} = \lambda_{\max}(\widetilde{A}\,\mathrm{diag}\,(N_1,N_2,\ldots,N_K))$ determines the epidemic threshold, because $\lambda_{\max}(J) > 0 \iff \frac{\beta}{\delta} > \tau_\pi$ corresponds to the disease spreading over the network, while $\lambda_{\max}(J) < 0 \iff \frac{\beta}{\delta} < \tau_\pi$ corresponds to the disease dying out. Secondly, it is possible to lower-bound the UMFF epidemic threshold by invoking properties of the *quotient matrix* [21], which is defined as $A^{(\pi)} = \widetilde{A}\,\mathrm{diag}(N_1,N_2,\ldots,N_K)$. As discussed in more detail in Appendix D, the eigenvalues of this quotient matrix $A^{(\pi)}$ can be bounded by the eigenvalues of the corresponding adjacency matrix A (see Theorem D 1 in Appendix D). In particular, the largest eigenvalue of the quotient matrix $A^{(\pi)}$ can be bounded by

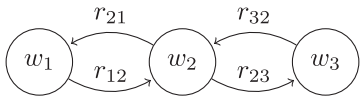
$$\lambda_{\max}(A^{(\pi)}) \leqslant \lambda_{\max}(A). \tag{A5}$$

From inequality (A5) and $\tau_\pi = \lambda_{\max}^{-1}(A^{(\pi)})$ follows that the UMFF epidemic threshold can be lower-bounded by $\tau_\pi \geqslant \lambda_1^{-1}$.

APPENDIX B: DERIVATION OF EXACT SIS EQUATIONS FOR $\widetilde{W}$ AND $\mathbf{E}[\widetilde{W}]$

1. The Kolmogorov equations for Markov Chains: Brief reminder

As a background for the further derivation of the UMFF equations (4), we start with a toy example to illustrate how the Kolmogorov equations are found for a Markov chain. Further details can be found in [12]. Consider the three-state Markov chain in $W(t)$ below:



The Markov chain has three states: w_1, w_2 and w_3, with state probabilities $\Pr[W(t) = w_i]$ and transition rates r_{ij}, for $1 \leqslant i \neq j \leqslant 3$. By the subscript "$ij$" in the rates r_{ij}, we denote the transition from state i to state j, i.e., $i \to j$. As mentioned in Sec. II, we assume that the transition processes are independent Poisson processes with exponentially distributed interevent times, for example, for the transition r_{12} this yields

$$\Pr[W(t+h) = w_2 | W(t) = w_1] = r_{12}e^{-r_{12}h}.$$

For $h \to 0$, this transition leads to

$$\frac{d\Pr[W(t)=w_2]}{dt} = r_{12}\Pr[W(t)=w_1]$$
$$\frac{d\Pr[W(t)=w_1]}{dt} = -r_{12}\Pr[W(t)=w_1].$$

Combining all transitions then leads to the Kolmogorov equations:

$$\begin{aligned}
\frac{d\Pr[W(t)=w_1]}{dt} &= -r_{12}\Pr[W(t)=w_1] \\
&\quad + r_{21}\Pr[W(t)=w_2] \\
\frac{d\Pr[W(t)=w_2]}{dt} &= r_{12}\Pr[W(t)=w_1] - (r_{23}+r_{21}) \\
&\quad \times \Pr[W(t)=w_2] + r_{32}\Pr[W(t)=w_3] \\
\frac{d\Pr[W(t)=w_3]}{dt} &= r_{23}\Pr[W(t)=w_2] - r_{32}\Pr[W(t)=w_3].
\end{aligned}$$

Hence, by identifying the state transitions and according rates, one obtains the Kolmogorov equations of a Markov chain, which completely characterize the dynamics of the process for a given initial distribution $\Pr[W(0) = w_i]$, for each possible state w.

2. State probability $\Pr[\widetilde{W}(t) = \tilde{w}]$

As described in Secs. III and IV, the reduced-state vector $\tilde{w}$ is introduced to compactly describe the disease state and to reduce the complexity of the SIS process description. Instead of describing the state of each node separately, the reduced-state vector $\tilde{w} = (\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_K)^T$ captures the number of infected nodes in each partition, by the relation $\tilde{w}_k = s_k^T w$. By $\mathcal{W}_x^k = \{w \in \{0,1\}^N | s_k^T w = x\}$ we denote the set of all full-state vectors w with x nodes infected in partition k (and with any possible number of infected nodes in the other partitions $m \neq k$). Each full-state vector $w \in \bigcap_{k=1}^K \mathcal{W}_{\tilde{w}_k}^k$ then corresponds to the reduced-state vector $\tilde{w}$, since each set $\mathcal{W}_{\tilde{w}_k}^k$ constrains the number of infected nodes in a specific partition k. Based on this notation, we can represent the coarse graining of the full states to the reduced states as

$$\bigcap_{k=1}^K \mathcal{W}_{\tilde{w}_k}^k \xrightarrow{\text{group by partitioning } \pi} \tilde{w}.$$

The full-state and reduced-state probabilities are then related as

$$\Pr[\widetilde{W} = \tilde{w}] = \sum_{w \in \bigcap_{k=1}^K \mathcal{W}_{\tilde{w}_k}^k} \Pr[W = w], \tag{B1}$$

and similarly, the rates are related as

$$r_{\tilde{w}_k(\tilde{w}_k \pm \tilde{e}_k)} \Pr[\widetilde{W}_k = \tilde{w}_k] = \sum_{w \in \bigcap_{i=1}^K \mathcal{W}_{\tilde{w}_i}^i} \sum_{j \in \mathcal{N}_k} r_{w(w \pm e_j)} \times \Pr[W = w]. \tag{B2}$$

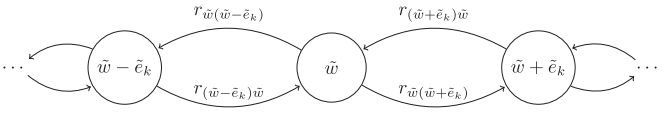
More can be said about the reduced-state transition structure: first, the entries $\tilde{w}_k$ represent the number of infected nodes in partition k, from which it follows that

$$\tilde{w} \in \{0,1,\ldots,N_1\} \times \{0,1,\ldots,N_2\} \times \cdots \times \{0,1,\ldots,N_K\},$$

and, second, since a state transition in the Markovian SIS process reflects a single infection or curing event, the possible transitions between reduced states are of the form

$$\tilde{w} \to \tilde{w} \pm \tilde{e}_k.$$

Hence, the reduced states and their transitions constitute an $(N_1 + 1) \times (N_2 + 1) \times \cdots \times (N_K + 1)$ lattice. This structure can be represented compactly by the chain below, which depicts one specific "direction" in the lattice, corresponding to one partition k:



However, since the transition rates between reduced states depend on the full states (B2), the transitions at the reduced-state level do not describe a Markov chain. Nonetheless, it is still possible to write the exact, but not-closed differential equations for the reduced-state probabilities by grouping the Kolmogorov equations according to the partitions:

$$\frac{d \Pr[\widetilde{W} = \tilde{w}]}{dt} = \sum_{w \in \bigcap_{k=1}^K \mathcal{W}_{\tilde{w}_k}^k} \frac{d \Pr[W = w]}{dt}.$$

Considering the transitions within the partitions separately enables the Kolmogorov equations at the reduced-state level to be written as

$$\begin{aligned}\frac{d \Pr[\widetilde{W}(t) = \tilde{w}]}{dt} = \sum_{k=1}^K (&- r_{\tilde{w}(\tilde{w}-\tilde{e}_k)} \Pr[\widetilde{W} = \tilde{w}] \\ &+ r_{(\tilde{w}+\tilde{e}_k)\tilde{w}} \Pr[\widetilde{W} = \tilde{w} + \tilde{e}_k] \\ &- r_{\tilde{w}(\tilde{w}+\tilde{e}_k)} \Pr[\widetilde{W} = \tilde{w}] \\ &+ r_{(\tilde{w}-\tilde{e}_k)\tilde{w}} \Pr[\widetilde{W} = \tilde{w} - \tilde{e}_k]).\end{aligned} \tag{B3}$$

The transition rates at the reduced-state level are derived below.

a. Transition rates $r_{\tilde{w}(\tilde{w}-\tilde{e}_k)}$ and $r_{(\tilde{w}+\tilde{e}_k)\tilde{w}}$: Node healing in partition k

By the grouping relation (B2) between the full states and the reduced states, the reduced-state transition rates are given by

$$r_{\tilde{w}(\tilde{w}-\tilde{e}_k)} \Pr[\widetilde{W} = \tilde{w}] = \sum_{w \in \bigcap_{i=1}^K \mathcal{W}_{\tilde{w}_i}^i} \sum_{j \in \mathcal{N}_k} r_{w(w-e_j)} \Pr[W = w]. \tag{B4}$$

The transition rate $r_{w(w-e_j)}$ in Eq. (B4) corresponds to node j healing in state w, i.e., the transition $W_j = 1 \to W_j = 0$. The healing rate in UMFF is δ for any node, hence the transition rate equals

$$r_{w(w-e_j)} = \delta w_j$$

for any full-state vector w and node j. The sum of the healing rates for all nodes in a partition k is then

$$\sum_{j \in \mathcal{N}_k} r_{w(w-e_j)} = \delta s_k^T w = \delta \tilde{w}_k. \tag{B5}$$

Substituting (B5) in the rate Eq. (B4) and invoking (B1) then yields

$$r_{\tilde{w}(\tilde{w}-\tilde{e}_k)} \Pr[\widetilde{W} = \tilde{w}] = \delta \tilde{w}_k \Pr[\widetilde{W} = \tilde{w}] \tag{B6}$$

for the reduced-state transition rate corresponding to a node healing in partition k, in state $\tilde{w}$. A similar derivation yields

$$r_{(\tilde{w}+\tilde{e}_k)\tilde{w}} \Pr[\widetilde{W} = \tilde{w} + \tilde{e}_k] = \delta(\tilde{w}_k + 1) \Pr[\widetilde{W} = \tilde{w} + \tilde{e}_k] \tag{B7}$$

for the reduced-state transition rate corresponding to a node healing in partition k, in state $\tilde{w} + \tilde{e}_k$.

b. Transition rates $r_{\tilde{w}(\tilde{w}+\tilde{e}_k)}$ and $r_{(\tilde{w}-\tilde{e}_k)\tilde{w}}$: a node in partition k is infected.

By the grouping relation (B2) between the full states and the reduced states, the reduced-state transition rates are given by

$$r_{\tilde{w}(\tilde{w}+\tilde{e}_k)} \Pr[\widetilde{W} = \tilde{w}] = \sum_{w \in \bigcap_{i=1}^K \mathcal{W}_{\tilde{w}_i}^i} \sum_{j \in \mathcal{N}_k} r_{w(w+e_j)} \Pr[W = w]. \tag{B8}$$

The transition rate $r_{w(w+e_j)}$ in Eq. (B8) corresponds to node j becoming infected in state w, i.e., the transition $W_j = 0 \to$

$W_j = 1$. Since

$$e_j^T A w = \sum_{i=1}^{N} a_{ij} w_i$$

is the number of infected neighbors of node j, and since each infected neighbor infects node j at a rate β if $w_j = 0$, the full-state transition rate

$$r_{w(w+e_j)} = \beta(1 - w_j) e_j^T A w$$

is found. The sum of infection rates for all nodes in partition k is then

$$\sum_{j \in \mathcal{N}_k} r_{w(w+e_j)} = \beta \sum_{m=1}^{K} (u - w)^T A^{(km)} w, \tag{B9}$$

where the sum over partitions $1 \leqslant m \leqslant K$ is introduced such that the block-matrix $A^{(km)}$, which naturally reflects the partition structure, can be used. Filling (B9) into the rate equation (B8) then yields

$$r_{\tilde{w}(\tilde{w}+\tilde{e}_k)} \Pr[\widetilde{W} = \tilde{w}] = \beta \sum_{m=1}^{K} \sum_{w \in \mathcal{W}^k_{\tilde{w}_k} \cap \mathcal{W}^m_{\tilde{w}_m}} (u - w)^T A^{(km)} w \Pr[W = w] \tag{B10}$$

for the reduced-state transition rate corresponding to a node becoming infected in partition k, in state $\tilde{w}$. A similar derivation yields

$$r_{(\tilde{w}-\tilde{e}_k)\tilde{w}} \Pr[\widetilde{W} = \tilde{w} - \tilde{e}_k] = \beta \sum_{m=1}^{K} \sum_{w \in \mathcal{W}^k_{(\tilde{w}_k - 1)} \cap \mathcal{W}^m_{\tilde{w}_m}} (u - w)^T A^{(km)} w \Pr[W = w] \tag{B11}$$

for the reduced-state transition rate corresponding to a node becoming infected in partition k, in state $\tilde{w} - \tilde{e}_k$.

c. *Resulting reduced-state equations*

Introducing the rates (B6), (B7), (B10), and (B11), the Kolmogorov equations (B3) establish Eq. (8).

3. Expected number of infected nodes $\mathbf{E}[\widetilde{W}_k]$

The equations for the expected number of infected nodes $\mathbf{E}[\widetilde{W}_k]$ can be derived from the reduced-state probability equations (8), based on the definition of expectation and the law of total probability. For any partition k, we can write the expected number of infected nodes as

$$\mathbf{E}[\widetilde{W}_k] = \sum_{\tilde{w}_k=0}^{N_k} \tilde{w}_k \Pr[\widetilde{W}_k = \tilde{w}_k]. \tag{B12}$$

By the law of total probability, the marginal probability can be written as

$$\Pr[\widetilde{W}_k = \tilde{w}_k] = \underbrace{\sum_{\tilde{w}_1=0}^{N_1} \cdots \sum_{\tilde{w}_l=0}^{N_l} \cdots \sum_{\tilde{w}_K=0}^{N_K}}_{\forall l \neq k} \Pr[\widetilde{W} = \tilde{w}]$$

such that (B12) equals

$$\mathbf{E}[\widetilde{W}_k] = \underbrace{\sum_{\tilde{w}_1=0}^{N_1} \cdots \sum_{\tilde{w}_l=0}^{N_l} \cdots \sum_{\tilde{w}_K=0}^{N_K}}_{\forall l} \tilde{w}_k \Pr[\widetilde{W} = \tilde{w}]. \tag{B13}$$

Differentiation with respect to time of Eq. (B13) then yields

$$\frac{d\mathbf{E}[\widetilde{W}_k]}{dt} = \underbrace{\sum_{\tilde{w}_1=0}^{N_1} \cdots \sum_{\tilde{w}_l=0}^{N_l} \cdots \sum_{\tilde{w}_K=0}^{N_K}}_{\forall l} \tilde{w}_k \frac{d \Pr[\widetilde{W} = \tilde{w}]}{dt}. \tag{B14}$$

After substitution of $\frac{d \Pr[\widetilde{W}=\tilde{w}]}{dt}$ from Eq. (8), we arrive at Eq. (9).

APPENDIX C: HIGHER-ORDER UMFF

The UMFF equations can be extended to higher-order moments, in order to better capture the dynamic correlations of the SIS process. For the case of $K = N$ partitions, Cator *et al.* [18] and Mata *et al.* [16] have described how the NIMFA [5] and quenched mean-field (QMF) [19] equations can be extended to nth-order moments:

$$\mathbf{E}[W_i] \rightarrow \mathbf{E}[W_i], \mathbf{E}[W_i W_j], \ldots, \mathbf{E}[\underbrace{W_i W_j \ldots W_l}_{n}]$$

based on the exact SIS dynamics. In order to have a closed set of equations for order n, the $(n+1)$th-order moments must be approximated by lower-order moments, i.e., an approximation of the form

$$\mathbf{E}[\underbrace{W_i W_j \ldots W_l}_{n+1}] \approx f\left(\mathbf{E}[\underbrace{W_i W_j \ldots W_l}_{\forall m \leqslant n}]\right),$$

where different choices for the moment-closure approximation f are given in Refs. [18] and [16]. Similarly, we can define the higher-order UMFF as:

Definition 7 (Higher-order UMFF). Consider a graph $G(\mathcal{N}, \mathcal{L})$, an epidemic process with rates (β, δ) and a graph partitioning π. For any integer $n \leqslant \sum_{k=1}^{K} N_k$, the nth-order UMFF equations are given by

$$\frac{d\mathbf{E}[\prod_{k=1}^{K} \widetilde{W}_k^{p_k}]}{dt} = \underbrace{\sum_{\tilde{w}_1=0}^{N_1} \cdots \sum_{\tilde{w}_k=0}^{N_k} \cdots \sum_{\tilde{w}_K=0}^{N_K}}_{\forall k} \prod_{k=1}^{K} \times \tilde{w}_k^{p_k} \frac{d \Pr[\widetilde{W} = \tilde{w}]}{dt} \tag{C1}$$

for all vectors $p \in \{p \in \mathbb{N}^K | 0 \leqslant p_k \leqslant N_k, \forall k \text{ and } u^T p \leqslant n\}$ and with $\frac{d\Pr[\widetilde{W}=\tilde{w}]}{dt}$ given by Eq. (11). The $(n+1)$th-order moments appearing in the higher-order UMFF equations are approximated by

$$\mathbf{E}\left[\prod_{k=1}^{K} \widetilde{W}_k^{p_k}\right] \approx f\left(\left\{\mathbf{E}\left[\prod_{k=1}^{K} \widetilde{W}_k^{q_k}\right]\right\}_{\forall q \in \mathcal{Q}}\right) \tag{C2}$$

for all vectors $p \in \{p \in \mathbb{N}^K | 0 \leqslant p_k \leqslant N_k, \forall k \text{ and } u^T p = n+1\}$ and with $\mathcal{Q} = \{q \in \mathbb{N}^K | 0 \leqslant q_k \leqslant p_k, \forall k \text{ and } u^T q \leqslant n\}$. The function f represents a generic moment-closure approximation.

Remark 1. The higher-order UMFF equations are found from the definition of expectation and the law of total probability, similar to the derivation of the first-order moments in Appendix B 3.

Remark 2. For a certain partition k, only the moments $\mathbf{E}[\dots \widetilde{W}_k^{p_k} \dots]$ for values $p_k \in \{1, \dots, N_k\}$ are considered. Since $\tilde{w}_k \in \{0, 1, \dots, N_k\}$ has (N_k+1) possible values, the probability distribution $\Pr[\widetilde{W}_k = \tilde{w}_k]$ is fully determined by the first N_k moments. Hence the set $\{p \in \mathbb{N}^K | 0 \leqslant p_k \leqslant N_k, \forall k \text{ and } u^T p \leqslant n\}$ represents the set of powers of all nth-order moments.

APPENDIX D: PROOF OF ISOPERIMETRIC INEQUALITIES

In this section, we prove the isoperimetric inequalities (24) and (25) of Theorem 3 and Theorem 4. We start by introducing some definitions and notations, based on the work of Haemers [21]. We then state and prove Lemma 1, from which Theorem 3 follows. Finally, we prove Theorem 4 based on the specific structure of biregular graphs.

1. Interlacing and quotient matrices

The following definitions are given in Refs. [21] and [11]:

Definition 8 (Interlacing sequences). Consider two sequences of real numbers: $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_N$ and $\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_K$ with $K \leqslant N$. The second sequence is said to interlace the first whenever

$$\alpha_i \geqslant \gamma_i \geqslant \alpha_{N-K+i} \quad \text{for } i = 1, \dots, K.$$

Definition 9 (Quotient matrix). The quotient matrix $A^{(\pi)}$ of an adjacency matrix A according to a partitioning π, is the matrix whose entries are the average row sums of the blocks of A. More precisely, $a_{km}^{(\pi)}$ is the entry in the quotient matrix according to the submatrix of A between nodes of $\mathcal{N}_k$ and $\mathcal{N}_m$ with value

$$a_{km}^{(\pi)} = \frac{1}{N_k} s_k^T A s_m.$$

These concepts can be combined by the interlacing theorem [21]:

Theorem 5 (Interlacing theorem). Suppose $A^{(\pi)}$ is the quotient matrix of a matrix A, then the eigenvalues of $A^{(\pi)}$ interlace the eigenvalues of A.

The interlacing theorem is crucial for the proof of the isoperimetric inequality as will become clear in the proof of Lemma 1.

2. General isoperimetric inequality

We start by proving Lemma 1 below:

Lemma 1. Consider a graph $G(\mathcal{N}, \mathcal{L})$ with N nodes. For any $c \in \mathbb{R}$ and any pair of Bernoulli vectors $w_x, w_y \in \{0,1\}^N$, with $N_x = u^T w_x$ and $N_y = u^T w_y$ ones, respectively, and with $w_x^T w_y = 0$, the following inequality holds:

$$\left| w_x^T A w_y - \frac{c}{N} N_x N_y \right| \leqslant \frac{\theta}{N} \sqrt{N_x(N-N_x) N_y (N-N_y)}, \tag{D1}$$

where $|c - \mu_i| \leqslant \theta$ for $1 \leqslant i < N$ holds.

A first proof of Lemma 1 is given by Chung [20] in the context of isoperimetric inequalities and discrepancy inequalities on graphs. The proof is mainly based on algebraic manipulations of the term $w_x^T A w_y$ and the eigendecomposition of the Laplacian matrix Q. As mentioned in Sec. II, the Laplacian Q is a positive semidefinite matrix possessing the eigendecomposition:

$$Q = Z M Z^T,$$

where Z is the orthogonal eigenmatrix with eigenvectors z_i as columns, and $M = \text{diag}(\mu_1, \mu_2, \dots, \mu_N)$, the diagonal matrix containing the eigenvalues. These eigenvalues can be ordered as $0 = \mu_N < \mu_{N-1} \leqslant \cdots \leqslant \mu_1$, where the 0 eigenvalue corresponds to the all-one eigenvector $z_N = \frac{u}{\sqrt{N}}$. Now, if we denote by $\widetilde{Z}$ the $N \times (N-1)$ matrix with z_N removed, and by $\widetilde{M}$ the $(N-1) \times (N-1)$ diagonal matrix $\widetilde{M} = \text{diag}(\mu_1, \dots, \mu_{N-1})$, then we can also write

$$Q = \widetilde{Z} \widetilde{M} \widetilde{Z}^T.$$

If we further denote by $Q_K = NI - uu^T$ the Laplacian matrix of the complete graph, then we can write $\widetilde{Z}\widetilde{Z}^T = \frac{1}{N} Q_K$, which holds for $\widetilde{Z}$ of any Laplacian matrix.

Proof A: We start by rewriting $w_x^T A w_y = w_x^T (\Delta - Q) w_y$. Due to the condition that $w_x^T w_y = 0$, we have $w_x^T \Delta w_y = 0$ and thus $w_x^T A w_y = -w_x^T Q w_y$. We then introduce the value $c \in \mathbb{R}$ as follows:

$$\begin{aligned} w_x^T A w_y &= -w_x^T Q w_y + \frac{c}{N} w_x^T Q_K w_y - \frac{c}{N} w_x^T Q_K w_y \\ &= w_x^T \left(\frac{c}{N} Q_K - Q\right) w_y - \frac{c}{N} w_x^T Q_K w_y. \end{aligned}$$

Since $w_x^T Q_K w_y = -N_x N_y$, and using the eigendecomposition of Q and Q_K, we obtain

$$w_x^T A w_y = w_x^T \widetilde{Z} (cI - \widetilde{M}) \widetilde{Z}^T w_y + \frac{c}{N} N_x N_y$$

or

$$w_x^T A w_y - \frac{c}{N} N_x N_y = w_x^T \widetilde{Z} (cI - \widetilde{M}) \widetilde{Z}^T w_y.$$

By introducing the variables $\alpha_i = w_x^T z_i$ and $\beta_i = w_y^T z_i$, we can write

$$w_x^T A w_y - \frac{c}{N} N_x N_y = \sum_{i=1}^{N-1} \alpha_i \beta_i (c - \mu_i). \tag{D2}$$

We can upper-bound the right-hand side as $|\sum_{i=1}^{N-1} \alpha_i \beta_i (c-\mu_i)| \leqslant \theta \sum_{i=1}^{N-1} |\alpha_i \beta_i|$, where we introduce θ with $|c - \mu_i| \leqslant \theta, \forall i \neq N$ as an upper bound. Equation (D2)

can then be written as

$$\left| w_x^T A w_y - \frac{c}{N} N_x N_y \right| \leqslant \theta \sum_{i=1}^{N-1} |\alpha_i \beta_i|.$$

Now, invoking the Cauchy-Schwartz inequality on the right-hand side of the equation, and replacing α_i, β_i by their original values yields

$$\left| w_x^T A w_y - \frac{c}{N} N_x N_y \right| \leqslant \theta \sqrt{\sum_{i=1}^{N-1} \alpha_i^2 \sum_{i=1}^{N-1} \beta_i^2} \leqslant \theta \sqrt{\left(w_x^T \widetilde{Z} \widetilde{Z}^T w_x\right)\left(w_y^T \widetilde{Z} \widetilde{Z}^T w_y\right)},$$

which by $\widetilde{Z}\widetilde{Z}^T = I - \frac{uu^T}{N}$ and $w_x^T(NI - uu^T)w_x = N_x(N - N_x)$ proves (D1). ■

A second proof for Lemma 1 can be formulated based on Haemers' interlacing theorem and applications [21]. Haemers ingeniously describes how quotient matrix constructions combined with the interlacing theorem can lead to algebraic expressions (i.e., involving Laplacian eigenvalues) for combinatorial quantities (i.e. possible number of links between subsets of nodes in a graph).

Proof B: Haemers defines the block matrix B

$$B = \begin{bmatrix} 0 & Q + cI \\ Q + cI & 0 \end{bmatrix} \tag{D3}$$

for some graph Laplacian Q, and any scalar $c \in \mathbb{R}$. By the antidiagonal blockform of B, we know that each eigenvalue μ_j of the Laplacian Q corresponds to two eigenvalues $\tilde{\lambda}_i = \mu_j + c$ and $\tilde{\lambda}_{2N-i} = -(\mu_j + c)$ of B. We consider a specific partitioning π of the rows of B (nodes in the combined graph), for which we can explicitly write the quotient matrix. For the Laplacian in the upper-right block, we partition the nodes $\mathcal{N}$ into a subset $\mathcal{N}_x$ of size N_x, and a remainder set $\mathcal{N}_{rx}$. For the Laplacian in the lower-left block, we partition the nodes $\mathcal{N}$ into a subset $\mathcal{N}_y$ of size N_y, where $\mathcal{N}_y$ is nonoverlapping with the N_x-size block of the other Laplacian, and a remainder set $\mathcal{N}_{ry}$. Overall, this results in the partitioning $\{\mathcal{N},\mathcal{N}\} \to \{\mathcal{N}_x,\mathcal{N}_{rx},\mathcal{N}_y,\mathcal{N}_{ry}\}$ for matrix B. For this partitioning, and because $Bu = cu$ due to $Qu = 0$, we can write the quotient matrix $B^{(\pi)}$ explicitly as

$$B^{(\pi)} = \begin{bmatrix} \frac{1}{N_x} & 0 & 0 & 0 \\ 0 & \frac{1}{N-N_x} & 0 & 0 \\ 0 & 0 & \frac{1}{N-N_y} & 0 \\ 0 & 0 & 0 & \frac{1}{N_y} \end{bmatrix} \begin{bmatrix} 0 & 0 & cN_x + m & -m \\ 0 & 0 & c(N - N_x - N_y) - m & cN_y + m \\ cN_x + m & c(N - N_x - N_y) - m & 0 & 0 \\ -m & cN_y + m & 0 & 0 \end{bmatrix}, \tag{D4}$$

where m is the number of links between subsets $\mathcal{N}_x$ and $\mathcal{N}_y$, i.e., $w_x^T A w_y$ in Lemma 1.

We can write the determinant of $B^{(\pi)}$ in two ways: an equality involving m and an inequality involving the eigenvalues of the Laplacian Q. Combining both expressions for the determinant then yields the isoperimetric inequality (D1) of Lemma 1.

From (D4), the determinant of $B^{(\pi)}$ can be calculated as

$$\det(B^{(\pi)}) = \frac{c^2(cN_xN_y + Nm)^2}{N_x(N - N_x)N_y(N - N_y)}. \tag{D5}$$

Secondly, if we call $\delta_1 \geqslant \delta_2 \geqslant \delta_3 \geqslant \delta_4$ the eigenvalues of $B^{(\pi)}$, where $\delta_1 = -\delta_4$ and $\delta_2 = -\delta_3$ hold because of the antidiagonal blockmatrix structure, then we have a second equation for the determinant:

$$\det(B^{(\pi)}) = \delta_1\delta_2\delta_3\delta_4 = \delta_1^2\delta_2^2. \tag{D6}$$

From the definition of $B^{(\pi)}$, it follows that the all-one vector u is an eigenvector with eigenvalue c, i.e., $B^{(\pi)}u = cu$. This means that either $|\delta_1| = c$ or $|\delta_2| = c$. Additionally, because $B^{(\pi)}$ is a quotient matrix of B, we know that the eigenvalue sequence δ_i of $B^{(\pi)}$ interlaces the eigenvalue sequence $\tilde{\lambda}_i$ of B:

$$-\tilde{\lambda}_2 \leqslant \delta_1 \leqslant \tilde{\lambda}_1 \quad \text{and} \quad -\tilde{\lambda}_3 \leqslant \delta_2 \leqslant \tilde{\lambda}_2.$$

Because we know that either δ_1 or δ_2 equals $\tilde{\lambda}_i = \mu_N + c = c$, we can write

$$\det(B^{(\pi)}) = \delta_1^2\delta_2^2 \leqslant c^2\Big[\max_{\forall i \neq N} |\mu_i + c|\Big]^2. \tag{D7}$$

Combining (D5) and (D7) gives

$$\frac{c^2(cN_xN_y + Nm)^2}{N_x(N - N_x)N_y(N - N_y)} \leqslant c^2\theta^2, \tag{D8}$$

with $|c + \mu_i| \leqslant \theta, \forall i \neq N$. By taking the square root of both sides, replacing m by $w_x^T A w_y$ and c by $-c$, we find again the isoperimetric inequality (D1) in Lemma 1. ■

Remark. Proofs A and B are two different ways to arrive at the same result. Proof A, based on Chung's approach, involves two approximations that upper-bound the cut-set approximation. The first approximation is upper bounding the $(c - \mu_i)$ values by θ, i.e. $|\sum_{i=1}^{N-1} \alpha_i\beta_i(c - \mu_i)| \leqslant \theta \sum_{i=1}^{N-1} |\alpha_i\beta_i|$. The second approximation involves the Cauchy-Schwartz inequality applied to the inner product $\sum_{i=1}^{N-1} |\alpha_i\beta_i|$. Proof B based on Haemers' approach, involves one approximation step. The absolute value of the second largest eigenvalue $|\delta_2|$ of the quotient matrix $B^{(\pi)}$ is upper bounded by the second largest absolute eigenvalue $\max_{i \neq N} |\mu_i + c|$ of $Q + cI$, based on the interlacing theorem. Since both approaches lead to the same result, we can conclude that the error due to interlacing is of the same nature as the error due to upper-bounding $(c - \mu_i)$ combined with the Cauchy-Schwartz inequality, which is a nontrivial relationship.

3. Proof of Theorem 3

Theorem 3 follows from Lemma 1 by particular choices of (c, A, w_x, w_y).

Proof. First, we choose $w_x = (u - w) \circ s_k$ and $w_y = w \circ s_m$, where $(. \circ s_k)$ represents the Hadamard product (elemen-

twise product) with s_k. For this choice of w_x and w_y, which are Bernoulli vectors satisfying $w_x^T w_y = 0$, and any adjacency matrix A, we can write

$$w_x^T A w_y = (s_k \circ (u - w))^T A(w \circ s_m) = (u - w)^T A^{(km)} w.$$

Secondly, we choose the specific value $c = N\tilde{a}_{km}$ which satisfies the condition $c \in \mathbb{R}$. These choices allow us to rewrite Lemma 1 as

$$|(u - w)^T A^{(km)} w - (\tilde{u} - \tilde{w})^T \widetilde{A}^{(km)} \tilde{w}| \leqslant \frac{\theta}{N} \sqrt{\tilde{w}_m (N - \tilde{w}_m)(N_k - \tilde{w}_k)(N - (N_k - \tilde{w}_k))}$$

for any adjacency matrix A, which equals Eq. (24) and thus proves Theorem 3. ■

4. Proof of Theorem 4

Theorem 4 states that the topological approximation error can be bounded more tightly for biregular graphs $A_{km,r}$, which we prove based on Haemers' interlacing techniques [21].

Proof. Consider a biregular graph $G_{km,r}$ with partitions $\mathcal{N}_k$ and $\mathcal{N}_m$, for which the adjacency matrix has the block form:

$$A_{km,r} = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix},$$

with $Bu = d_1 u$ and $u^T B = d_2 u^T$, because the graph is biregular. The values $d_1 = \frac{L}{N_k}$ and $d_2 = \frac{L}{N_m}$ are the degrees of the partitions. Furthermore, we consider a partitioning π of the nodes of $G_{km,r}$ into four sets $\{\mathcal{N}_k^x, \mathcal{N}_k^r, \mathcal{N}_m^y, \mathcal{N}_m^r\}$ according to

$$\mathcal{N}_k^x \cup \mathcal{N}_k^r = \mathcal{N}_k, \quad \mathcal{N}_k^x \cap \mathcal{N}_k^r = \emptyset \quad \text{and } |\mathcal{N}_k^x| = N_x$$
$$\mathcal{N}_m^y \cup \mathcal{N}_m^r = \mathcal{N}_m, \quad \mathcal{N}_m^y \cap \mathcal{N}_m^r = \emptyset \quad \text{and } |\mathcal{N}_m^y| = N_y.$$

In other words, partition k is further refined into a subset of N_x nodes and a remainder subset, and similarly for partition m. For this partitioning π, the quotient matrix can be explicitly written as

$$A_{km,r}^{(\pi)} = \begin{bmatrix} \frac{1}{N_x} & 0 & 0 & 0 \\ 0 & \frac{1}{N_k - N_x} & 0 & 0 \\ 0 & 0 & \frac{1}{N_y} & 0 \\ 0 & 0 & 0 & \frac{1}{N_m - N_y} \end{bmatrix} \begin{bmatrix} 0 & 0 & m & \frac{L}{N_k} N_x - m \\ 0 & 0 & \frac{L}{N_m} N_y - m & L\left(1 - \frac{N_x}{N_k} - \frac{N_y}{N_m}\right) + m \\ m & \frac{L}{N_m} N_y - m & 0 & 0 \\ \frac{L}{N_k} N_x - m & L\left(1 - \frac{N_x}{N_k} - \frac{N_y}{N_m}\right) + m & 0 & 0 \end{bmatrix}, \quad \text{(D9)}$$

where m is the number of links between partitions $\mathcal{N}_k^x$ and $\mathcal{N}_m^y$, i.e. the cut-set size $(u - w)^T A^{(km)} w$ in Theorem 4.

We can write the determinant of $A_{km,r}^{(\pi)}$ in two ways: an expression involving m, which follows directly from the block-matrix form and secondly, an inequality involving the eigenvalues of $A_{km,r}^{(\pi)}$. Combining both expressions for the determinant yields the isoperimetric inequality of Theorem 4. From (D9), the determinant of $A_{km,r}^{(\pi)}$ can be calculated as

$$\det\left(A_{km,r}^{(\pi)}\right) = \frac{L^2\left(m - \frac{L}{N_k N_m} N_x N_y\right)^2}{N_x(N_k - N_x) N_y (N_m - N_y)}. \quad \text{(D10)}$$

Second, if we call $\delta_1 \geqslant \delta_2 \geqslant \delta_3 \geqslant \delta_4$ the eigenvalues of $A_{km,r}^{(\pi)}$, where $\delta_1 = -\delta_4$ and $\delta_2 = -\delta_3$ hold because of the antidiagonal block structure, then we have a second equation for the determinant:

$$\det\left(A_{km,r}^{(\pi)}\right) = \delta_1 \delta_2 \delta_3 \delta_4 = \delta_1^2 \delta_2^2. \quad \text{(D11)}$$

Next, two facts about the eigenvalues of $A_{km,r}^{(\pi)}$ are combined to find expression (25). First, because $A_{km,r}^{(\pi)}$ is a quotient matrix of $A_{km,r}$, we know by Theorem D 1 that the eigenvalues of the first interlace those of the latter. In other words, we can bound δ_2 by

$$\lambda_{N-K+2} \leqslant \delta_2 \leqslant \lambda_2.$$

Because $\lambda_{N-K+2} = \lambda_{N-2} = -\lambda_3 \geqslant -\lambda_2$, we find

$$\delta_2^2 \leqslant \lambda_2^2. \quad \text{(D12)}$$

The second fact we use is

$$\delta_1 = \frac{L}{\sqrt{N_k N_m}}, \quad \text{(D13)}$$

which can be verified by considering the eigenvalue equation:

$$\left(A_{km,r}^{(\pi)} - \frac{L}{\sqrt{N_k N_m}} I\right) \begin{bmatrix} \sqrt{N_m} \\ \sqrt{N_m} \\ \sqrt{N_k} \\ \sqrt{N_k} \end{bmatrix} = 0,$$

from which follows that $(\sqrt{N_m}, \sqrt{N_m}, \sqrt{N_k}, \sqrt{N_k})^T$ is the right eigenvector of $A_{km,r}^{(\pi)}$ according to eigenvalue $\delta_1 = \frac{L}{\sqrt{N_k N_m}}$. By the Perron-Frobenius theorem [11], we know that for non-negative matrices such as $A_{km,r}^{(\pi)}$, the largest (possibly nonunique) eigenvalue accords to an eigenvector with non-negative elements. This means that δ_1 is the largest eigenvalue of $A_{km,r}^{(\pi)}$ since its corresponding eigenvector is a vector with non-negative elements.

Combining (D12) and (D13) then yields an upper bound for the determinant of $A_{km,r}^{(\pi)}$ in Eq. (D11):

$$\det\left(A_{km,r}^{(\pi)}\right) \leqslant \frac{L^2}{N_k N_m} \lambda_2^2.$$

Combined with (D10) this gives

$$\frac{L^2\left(m - \frac{L}{N_k N_m} N_x N_y\right)^2}{N_x(N_k - N_x) N_y (N_m - N_y)} \leqslant \frac{L^2}{N_k N_m} \lambda_2^2.$$

Which reduces to Eq. (25) if we replace m by $(u - w)^T A^{(km)} w$, and which thus proves Theorem 4. ■

[1] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[2] J. O. Kephart and S. R. White, Directed-graph epidemiological models of computer viruses, in *Proc. IEEE Comput. Soc. Symp. Research in Security and Privacy* (IEEE, Oakland, CA, 1991), p. 343.

[3] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. London A **115**, 700 (1927).

[4] E. Cator, R. van de Bovenkamp, and P. Van Mieghem, Susceptible-infected-susceptible epidemics on networks with general infection and cure times, Phys. Rev. E **87**, 062816 (2013).

[5] P. Van Mieghem, J. Omic, and R. Kooij, Virus spread in networks, IEEE/ACM Trans. Netw. **17**, 1 (2009).

[6] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, Phys. Rev. Lett. **86**, 3200 (2001).

[7] M. Boguñá and R. Pastor-Satorras, Epidemic spreading in correlated complex networks, Phys. Rev. E **66**, 047104 (2002).

[8] W. Wang, M. Tang, H. E. Stanley, and L. A. Braunstein, Unification of theoretical approaches for epidemic spreading on complex networks, Rep. Prog. Phys. **80**, 036603 (2017).

[9] V. Blåsjö, The isoperimetric problem, Am. Math. Mon. **112**, 526 (2005).

[10] R. Osserman, The isoperimetric inequality, Bull. Am. Math. Soc. **84**, 1182 (1978).

[11] P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, Cambridge, 2011).

[12] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems* (Cambridge University Press, Cambridge, 2014).

[13] P. Van Mieghem, Approximate formula and bounds for the time-varying susceptible-infected-susceptible prevalence in networks, Phys. Rev. E **93**, 052312 (2016).

[14] P. Van Mieghem, Time evolution of SIS epidemics on the complete graph. Delft University of Technology, Report 20170405, https://www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports.html.

[15] E. Cator and P. Van Mieghem, Nodal infection in Markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated, Phys. Rev. E **89**, 052802 (2014).

[16] A. S. Mata and S. C. Ferreira, Pair quenched mean-field theory for the susceptible-infected-susceptible model on complex networks, Europhys. Lett. **103**, 48003 (2013).

[17] C. Li, R. van de Bovenkamp, and P. Van Mieghem, Susceptible-infected-susceptible model: A comparison of n-intertwined and heterogeneous mean-field approximations, Phys. Rev. E **86**, 026116 (2012).

[18] E. Cator and P. Van Mieghem, Second-order mean-field susceptible-infected-susceptible epidemic threshold, Phys. Rev. E **85**, 056111 (2012).

[19] C. Castellano and R. Pastor-Satorras, Thresholds for Epidemic Spreading in Networks, Phys. Rev. Lett. **105**, 218701 (2010).

[20] F. Chung, Discrete isoperimetric inequalities, in *Discrete Mathematics and Theoretical Computer Science* (Springer-Verlag, Singapore, 1996).

[21] W. H. Haemers, Interlacing eigenvalues and graphs, Linear Algebra Appl. **226**, 593 (1995); Honoring J. J. Seidel.

[22] R. Diestel, *Graph Theory*, 4th ed., *Graduate Texts in Mathematics*, Vol. 173 (Springer, New York, 2012).

[23] T. Tao, Szemerédi's regularity lemma revisited, Contrib. Discrete Math. **1**(1), 8 (2006).

[24] S. Bonaccorsi, S. Ottaviano, D. Mugnolo, and F. De Pellegrini, Epidemic outbreaks in networks with equitable or almost-equitable partitions, SIAM J. Appl. Math. **75**, 2421 (2015).

[25] A. Ganesh, L. Massoulie, and D. Towsley, The effect of network topology on the spread of epidemics, in *Proc. IEEE 24th Annual Joint Conf. IEEE Computer and Communications Societies* (IEEE, Miami, FL, 2005), Vol. 2, p. 1455.

[26] P. Van Mieghem, Universality of the SIS prevalence in networks, arXiv:1612.01386.

[27] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem, Generalized epidemic mean-field model for spreading processes over multilayer complex networks, IEEE/ACM Trans. Netw. **21**, 1609 (2013).

[28] M. Ángeles Serrano, D. Krioukov, and M. Boguñá, Self-Similarity of Complex Networks and Hidden Metric Spaces, Phys. Rev. Lett. **100**, 078701 (2008).

[29] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, Hyperbolic geometry of complex networks, Phys. Rev. E **82**, 036106 (2010).