Models of epidemic spreading on complex networks have attracted great attention among researchers in physics, mathematics, and epidemiology due to their success in predicting and controlling scenarios of epidemic spreading in real-world scenarios. To understand the interplay between epidemic spreading and the topology of a contact network, several outstanding theoretical approaches have been developed. An accurate theoretical approach describing the spreading dynamics must take both the network topology and dynamical correlations into consideration at the expense of increasing the complexity of the equations. In this short survey we unify the most widely used theoretical approaches for epidemic spreading on complex networks in terms of increasing complexity, including the mean-field, the heterogeneous mean-field, the quench mean-field, dynamical message-passing, link percolation, and pairwise approximation. We build connections among these approaches to provide new insights into developing an accurate theoretical approach to spreading dynamics on complex networks.

PACS numbers: 89.75.Hc, 87.19.X-, 87.23.Ge

I. INTRODUCTION

Throughout human history infectious diseases have been a constant threat to the health of society, and when diseases become epidemic they cause huge economic losses. Models that enable the prediction and control of epidemics have attracted much attention among researchers in the fields of epidemiology, biology, sociology, mathematics, and physics. Bernoulli proposed the first mathematical model for understanding the spreading of smallpox [1]. It initiated a new era in the modern mathematical modeling of infectious diseases, and many models for describing the characteristics of epidemic spreading in which the states of individuals are disaggregated by compartments have been proposed. For example, the acquired immune deficiency syndrome (AIDS) can be described using the susceptible-infected (SI) model, since infected individuals, once infected by the AIDS virus, cannot be cured. Seasonal influenza and blennorrhagia can be described using the susceptible-infected-susceptible (SIS) model, because individuals can be infected more than once. Chickenpox and measles can be modeled using the susceptible-infected-infected (SIR) model, because once infected individuals have recovered they acquire permanent immunization [2].

During the last century many researchers assumed that all individuals uniformly interact, i.e., that they interact with all other individuals with the same probability [3, 4]. Thus the internal structure or topology of the contact network through which the epidemic spreads was neglected. The main features of the topology of a contact network include its degree distribution $P(k)$, i.e., the fraction of nodes with $k$ contacts, the weights on links and nodes, degree correlations, clustering, and community structure [5]. With the availability of large-scale data in real-world contact networks, scholars have become aware of the existence of heterogeneities in the topology of networks (i.e., power-law degree distribution and different weights of contact network), the high clustering, and small-world phenomena [6–11]. The effects of contact network topology on epidemic spreading was first introduced by Pastor-Satorras and Vespignani [12] using complex networks in which the nodes represent individuals and the edges the interactions among them.

Over the past decade this pioneering work has encouraged much outstanding research in the field of network spreading dynamics [13]. Both Monte Carlo simulations [14–17] and theoretical study [18] have investigated the effects of network structures on epidemic spreading velocity [19, 20], epidemic variability [21, 22], epidemic size [23–28], and epidemic thresholds [29–34]. Both the epidemic size and threshold can indicate the probability of an epidemic occurring [32], which seeds are influential [35–38], and how to effectively control the epidemic once it begins [39, 41]. When the transmission probability is above an epidemic threshold the system is in an active epidemic state, i.e., there is a finite fraction of nodes infected by the epidemic, but when the transmission probability is below the epidemic threshold the epidemic dissipates. Near the epidemic threshold the system exhibits such interesting phenomena as the rare-region phenomenon [42–45] and the scaling behavior of the main magnitudes [46, 47]. Generally speaking, previous research has addressed how the topology of the contact networks has an affect on the macroscopic, mesoscale, and microscopic levels. Research on the macroscopic scale, which focuses primarily on the effects of degree and weight distributions [13, 48–51], has revealed
that networks with strong heterogeneous degree distributions have a vanishing epidemic threshold \[\text{[12]}\] and that the heterogeneous weight distributions increase the epidemic threshold \[\text{[24, 52, 53]}\]. Research on the mesoscopic scale, which focuses on degree-degree correlations, clustering, and communities \[\text{[54, 57]}\], has found that assortativity \[\text{[54]}\], high clustering \[\text{[57]}\], and community structure \[\text{[58]}\] enhance the epidemic outbreak and that disassortativity \[\text{[54]}\] diminishes it. Research on epidemiics from a microscopic point of view \[\text{[12, 59]}\] has discovered that high-degree nodes—hubs—are infected quickly \[\text{[19]}\] and that the epidemic is more likely to transmit through low-weight edges \[\text{[45, 60]}\].

There are many successful theoretical approaches to describing epidemic spreading on complex networks. When describing the interplay between complex network structure and the dynamics of epidemic spreading, there are two challenges. The first is describing the intricate topologies of the contact networks, since in real-world networks this involves heterogeneous degree and weight distributions \[\text{[61, 62]}\], high clustering \[\text{[57, 63, 64]}\], motif structure \[\text{[65, 66]}\], community structure \[\text{[67, 68]}\], and fractal structures \[\text{[69, 70]}\]. The second and more difficult challenge is describing the strong dynamic correlations among the states of the nodes. The dynamic correlations are produced when the epidemic being transmitted to a node from two of its neighbors are correlated \[\text{[71]}\]. When there is only one seed node and there is high clustering the dynamic correlations are obvious, since all the infection paths come from the same seed. These two challenges are not fully addressed in the existing literature, which always assumes (i) that an epidemic spreads on a large, sparse network \[\text{[23, 72–74]}\], (ii) that dynamic correlations among the neighbors do not exist \[\text{[72]}\], and (iii) that all the nodes or edges within a given class are statistically equivalent \[\text{[72, 75]}\]. These theoretical approaches can be changed by removing or adding assumptions, but a comprehensive review of the relationships among these approaches is still lacking. Here we discuss the main contributions, basic assumptions, and derivation processes of seven widely-used approaches in terms of their increasing complexity.

II. THEORETICAL APPROACHES

The two most widely used model for epidemic spreading dynamics are the reversible susceptible-infected-susceptible (SIS) and the irreversible susceptible-infected-removed (SIR) models \[\text{[2, 4]}\]. In the SIS model the nodes are either susceptible or infected. In its continuous time version, at each time step each infected node transmits its infection to all its susceptible neighbors at the same rate \(\lambda\) and returns to the susceptible state at a rate \(\gamma\). Thus the effective transmission rate, i.e., the effective transmission of infection, is \(\beta = \lambda / \gamma\). Without loss of generality we set \(\gamma = 1\). In the SIR model, unlike the SIS model, an infected node recovers and is permanently removed at the rate \(\gamma\). Figures 1(a) and 1(b) show schematically the transitions between compartments in the SIS and SIR models. The schematic temporal evolutions of the SIS and SIR models are shown in Fig. 1(c). When \(t \to \infty\), the order parameters (i.e., the epidemic sizes) of the two models overcome a second-order phase transition depending on the value of \(\beta\), as shown in Fig. 1(d). The critical threshold \(\beta_c\) divides the phase diagram into absorbing and active regions. When \(\beta < \beta_c\) there is an absorbing region and no epidemic. When \(\beta > \beta_c\) there is an active region and a global epidemic develops.

The theoretical approaches that enable the computation of the epidemic threshold and the magnitude of these models, i.e., the epidemic size, exhibit similar but distinct frameworks for the SIS and SIR models and are of two types. In the first type the only difference between the SIS and the SIR models is that in the SIR a removed state is added to the SIS model. This first type includes the mean-field, heterogeneous mean-field, dynamic message-passing, and pairwise approximation approaches. The second type includes the link percolation and edge-based compartmental approaches, which provide the valid and obvious framework for the SIR model since, unlike the SIS model, it is irreversible. Here we use the SIS model to illustrate the relations among the existing approaches of this first type. We use the SIR model to explain the approaches of the second type. To clarify the following, the definitions of the parameters are given in Table II in the Appendix.

A. General frameworks for models of epidemic spreading

Mean-field (MF) approach. The simplest framework for describing these models—the mean-field (MF) approach—assumes that the population is fully mixed, i.e., that all nodes in the population are statistically equivalent and thus the inter-
action probabilities between any two individuals are the same. As a consequence the topology of the contact network is neglected. This approach was widely used in the last century [3]. The MF approach also assumes that there are no dynamic correlations among the states of a node and its neighbors. The time evolution of the density of infected nodes in the SIS model is given by

\[
\frac{d\rho(t)}{dt} = -\rho(t) + \beta\langle k \rangle \rho(t)[1 - \rho(t)],
\]

where \(\rho(t)\) and \(1 - \rho(t) = s(t)\) are the fractions of infected and susceptible nodes at time \(t\), respectively, and \(\langle k \rangle\) is the average contact capacity of the nodes, i.e., the average degree of the network. The first term on the right hand side of Eq. (1) is the fraction of infected nodes that returns to the susceptible, and the second term denotes the fraction of susceptible nodes that are infected by infected neighbors. In the steady state, i.e., \(dp(t)/dt = 0\), we have

\[
\rho(\infty) - \beta\langle k \rangle [1 - \rho(\infty)]\rho(\infty) = 0,
\]

where \(\rho(\infty) = \rho(t \to \infty)\) is the fraction of infected nodes in the steady state, i.e., the relative epidemic size. Equation (2) has two roots with a trivial solution \(\rho(\infty) = 0\), and a non-trivial solution \(\rho(\infty) > 0\) that exists only when the effective transmission rate is greater than an epidemic threshold

\[
\beta_c^{MF} = \frac{1}{\langle k \rangle}.
\]

The epidemic threshold \(\beta_c^{MF}\) divides the solutions into absorbing and active regions. When \(\beta \leq \beta_c^{MF}\) there is an absorbing region and no epidemic. When \(\beta > \beta_c^{MF}\) there is an active region and a global epidemic develops. The simplest MF approach neglects the internal structure of the contact networks and the dynamic correlations among the states of the neighbors. This oversimplified approach produces qualitatively analytical results, such as the existence of an epidemic threshold and the scaling relation of critical phenomena [4]. For networks with a homogeneous degree distribution (e.g., those with a well-mixed population and those that are random regular networks), the MF approach accurately predicts epidemic size and threshold. Performing the Taylor expansion of Eq. (2) at the epidemic threshold \(\beta = \beta_c^{MF}\), Moreno et al. describe the epidemic prevalence behavior as \(\rho(\infty) \sim (\beta - \beta_c^{MF})\) near the epidemic threshold [72]. Unfortunately the MF approach can be inaccurate in some situations, e.g., when networks have a heterogeneous degree distribution, since it neglects both network topology and dynamical correlations. For example, Ferreira et al. numerically demonstrate that there is no epidemic threshold for scale-free networks with a degree exponent \(\nu \leq 3\) [14], but the MF approach predicts a finite value.

**Heterogeneous mean-field (HMF) approach.** To more accurately capture network structure, Pastor-Satorras and Vespignani have improved the MF approach for the SIS model by creating the heterogeneous mean-field (HMF) approach [12] in which nodes with the same degree are equivalent. In the HMF approach the fraction of nodes in the infected state \(\rho(t)\) is split by the degree \(k\) of the nodes. Thus the primary magnitude is \(\rho_k(t)\), which is the fraction of infected nodes with degree \(k\) at time \(t\). The total fraction of infected nodes is \(\rho(t) = \sum_k P(k)\rho_k(t)\), where \(P(k)\) is the degree distribution of the network. In uncorrelated degree networks, a susceptible node is connected to an infected neighbor with a probability

\[
\Theta(t) = \frac{1}{\langle k \rangle} \sum_{k} \max_{r} P(k)k\rho_k(t),
\]

where \(\max_{r}\) is the maximum degree. The time evolution of \(\rho_k(t)\) is given by

\[
\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \beta k[1 - \rho_k(t)]\Theta(t).
\]

Similar to Eq. (1), the first (second) term in the right hand side of Eq. (5) is the fraction of infected (susceptible) nodes with degree \(k\) that recover (are infected by infected neighbors) at time \(t\).

To obtain the epidemic threshold, Eq. (5) is linearized around the initial conditions \(\rho_k(0) \to 0\) and then expressed in a matrix form

\[
\frac{d\bar{\rho}(t)}{dt} = C\bar{\rho}(t),
\]

where \(\bar{\rho}(t) = [\rho_1(t), \cdots, \rho_{\max}(t)]^T\). The Jacobian matrix \(C = \{C_{kk'}\}\) is given by

\[
C_{kk'} = \beta k' P(k') (k) - \delta_{k,k'},
\]

where \(\delta_{k,k'}\) is a Dirac delta function. The system has a global epidemic—an active region—in which \(\rho(t) = \sum_k P(k)\rho_k(t)\) grows exponentially, which mathematically means that the largest eigenvalue of \(C\), \(\beta \langle k^2 \rangle /\langle k \rangle - 1\), is greater than zero. Thus the epidemic threshold can be expressed

\[
\beta_c^{HMF} = \frac{\langle k \rangle}{\langle k^2 \rangle},
\]

where \(\langle k \rangle\) and \(\langle k^2 \rangle\) are the first and second moments of the degree distribution, respectively. For homogeneous networks, e.g., ER networks, the epidemic threshold is \(\beta_c^{HMF} = 1/\langle k \rangle + 1\), which for small \(\langle k \rangle\) is different from that predicted using the MF approach [See Eq. (3)]. For heterogeneous networks with a power-law degree distribution \(P(k) \sim k^{-\nu}\), in the thermodynamic limit, i.e., \(N \to \infty\), the epidemic threshold is zero for degree exponent \(\nu \leq 3\) due to the divergence of \(\langle k^2 \rangle\). When \(\nu > 3\) there is a finite epidemic threshold. The HMF approach has been highly successful in describing the dynamics of epidemic spreading for two reasons, i.e., (i) we only need to know the degree distribution and (ii) HMF is able to uncover how topological heterogeneity affects epidemic spreading, e.g., there is no epidemic threshold when the degree distributions are highly heterogeneous. Researchers have generalized the HMF approach to investigate the effects of weight distribution [74], degree-degree correlations [77], and multiplicity [75, 78, 79]. For example, Wang et
al. [73][79] generalized the HMF theory to study the effect of asymmetrically-interacting spreading dynamics on complex layered networks and found that the epidemic outbreak on the contact layer can induce an outbreak of information on the communication layer, and that the information spreading can effectively raise the epidemic threshold.

The HMF approach is usually effective when the networks have an infinite topological dimension, i.e., when the number of nodes in a neighborhood grows proportionately to network size $N$ with the topological distance from an arbitrary origin [81]. Although random networks above the percolation threshold indeed have an infinite dimension, some researchers find that the HMF approach can fail because two important factors are not taken into consideration [81][84]. First, because the HMF approach describes the network topology using degree distribution as the only input parameter, the quenched connections among the nodes are neglected. Second, the dynamical correlations among the states of neighbors are neglected, since Eq. [5] assumes that the states of neighbors are independent [81]. This simplified assumption allows the HMF approach to accurately capture spreading dynamics on complex networks [85]. When epidemics spread on quenched networks, the HMF description of the dynamics is only qualitative [82].

Quench mean-field (QMF) approach. Because neither the MF nor the HMF approach can describe the full network structure, researchers use the adjacency matrix $A$ to represent the full contact network topology. The $A_{ij}$ component values of matrix $A$ are $A_{ij} = 1$ when nodes $i$ and $j$ are connected. Incorporating the adjacency matrix, the quench mean-field (QMF) approach is widely used to study the spreading dynamics. Note that other approaches also use the adjacency matrix to describe network topology, including the discrete-time Markov chain approach [86] and the N-intertwined approach [87][88]. At time $t$ a susceptible node $i$ is infected by its neighbors with a probability $\beta \sum_{j=1}^{N} A_{ij} \rho_j(t)$, where $\rho_j(t) = 1 - s_j(t)$ is the probability that neighboring node $j$ of node $i$ is in the infected state at time $t$. Thus the evolution of $\rho_i(t)$ can be expressed as

$$\frac{d\rho_i(t)}{dt} = -\rho_i(t) + \beta [1 - \rho_i(t)] \sum_{j=1}^{N} A_{ij} \rho_j(t). \tag{9}$$

The first (second) term on the right hand side of Eq. (9) is the probability that node $i$ recovers (becomes infected by its neighbors) at time $t$. The fraction of nodes in the infected state at time $t$ is $\rho(t) = 1/N \sum_{i=1}^{N} \rho_i(t)$. Since only a vanishing small fraction of nodes are in the infected state at the beginning of the spreading, i.e., $\rho_i(0) \rightarrow 0$, we linearize Eq. (9) around $\rho_i(t) \rightarrow 0$ and rewrite it in matrix form

$$\frac{d\vec{\rho}(t)}{dt} = -\vec{\rho}(t) + \beta A \vec{\rho}(t), \tag{10}$$

where $\rho_i(t)$ is element $i$ of the vector $\vec{\rho}(t) = (\rho_1(t), \ldots, \rho_N(t))^T$. Using the same tool as that used to obtain Eq. (8), the epidemic threshold is given by

$$\beta_c^{QMF} = \frac{1}{\Lambda_A}. \tag{11}$$

where $\Lambda_A$ is the largest eigenvalue of $A$. The epidemic threshold predicted by the QMF approach is dependent only on network topology. In uncorrelated scale-free (SF) networks with a power-law degree distribution, $\beta_c^{QMF} \propto \langle k \rangle / \langle k^2 \rangle$ when $\nu < 2.5$, which produces the same threshold as Eq. (8). When $\nu > 2.5$, $\beta_c^{QMF} \propto 1/\sqrt{k_{\text{max}}}$, which indicates that there is no epidemic threshold in the thermodynamic limit [89]. This result is in contrast to the prediction from the HMF approach. The discrepancy between the HMF and QMF approaches is addressed in Refs. [18][23][33][42].

The QMF approach uses the adjacency matrix to describe network topology and to predict epidemic sizes and thresholds more accurately than those predicted by the MF and HMF approaches. However the dynamical correlations among the states of neighbors are still neglected in Eq. (9), and this produces deviations between the theoretical predictions and numerical simulations. Some researchers doubt the accuracy of the epidemic threshold value predicted by the QMF approach and some consider it completely wrong in some specific situations [14][15][33][42]. To validate the effectiveness of the QMF approach in predicting a epidemic threshold, Goltsev et al. [42] define an indicator, the inverse participation ratio (IPR), that quantifies the eigenvector localization of $\Lambda_A$. The IPR of $\Lambda_A$ is given by $v(\Lambda_A) = \sum_{i=1}^{N} f_i(\Lambda_A)^4$, where $f_i(\Lambda_A)$ is the $i$-th element of the eigenvector $\vec{\rho}(\Lambda_A)$ of $\Lambda_A$.

If $\vec{\rho}(\Lambda_A)$ is delocalized, $v(\Lambda_A) \propto O(1)$. If $\vec{\rho}(\Lambda_A)$ is localized, $v(\Lambda_A) \propto O(1)$ [18]. Goltsev et al. [42] claimed that $\vec{\rho}(\Lambda_A)$ is delocalized for $\nu < 2.5$, which implies that the epidemic size is finite when $\beta > \beta_c^{QMF}$. However $\vec{\rho}(\Lambda_A)$ is localized when $\nu > 2.5$, which means that only hubs and their neighbors are infected, and as a consequence the epidemic grows very slowly and may die out due to fluctuations. Thus they found that this localized state does not constitute a true active state and that the epidemic threshold is closer to that given by Eq. (8).

Recently Pastor-Satorras and Castellano [13] have further proven that $\vec{\rho}(\Lambda_A)$ of $\Lambda_A$ is localized on the hubs when $\nu > 5/2$ and localized on nodes with the largest index in the $K$-core decomposition [35] when $\nu < 5/2$. To explain why the epidemic threshold predicted by the QMF approach sometimes fails, we must understand the physical meaning of $\vec{\rho}(\Lambda_A)$, which can be regarded the centrality of a node (i.e., the eigenvector centrality) [15]. The eigenvector centrality assigns to each node a centrality proportional to the sum of the eigenvector centralities of its neighbors. Unfortunately the hubs with high centralities induce high centralities in their neighbors, who in turn feed the centralities back to the hub. As a result, the centrality of the hub is overestimated. Similarly we know that the probability that a node is in the infected state is also overestimated. The variable $\rho_i(t)$ grows as $\rho_j(t)$ increases, and the value of $\rho_j(t)$ also increases when $\rho_i(t)$ increases. Thus the infection is transmitted back and forth through the same edge, which results in an "echo chamber" effect, and the infection probability of susceptible nodes is overestimated [90]. We know that Eq. (9) overestimates the probability that a node is in the infected state.

Dynamical message passing (DMP) approach. To over-
come the weaknesses of the QMF approach but retain its advantages, i.e., to take into consideration the full network structure, the dynamic message-passing (DMP) approach was first proposed by Karrer and Newman \cite{91} in their study of the SIR model and generalized later by Shrestha et al. \cite{74} to describe the SIS model. The DMP approach disallows a node in the "cavity" state from transmitting an infection to its neighbors but allows it to be infected by them (see Fig. 2). This prevents the epidemic from passing back and forth through the same edge, causing an "echo chamber" \cite{90} effect in Eq. \ref{eq:theta} that decreases the overestimation of the infection probability of susceptible nodes. Some dynamic correlations among the states of the neighbors are also taken into consideration. The DMP approach is exact in tree-like networks \cite{91}, i.e., in networks with no loops. Based on the DMP approach, the time evolution of $\rho_i(t)$ can be written

$$\frac{d\rho_i(t)}{dt} = -\rho_i(t) + \beta(1 - \rho_i(t)) \sum_{j=1}^{N} A_{ij} \theta_j \rightarrow i(t),$$

\label{eq:rho}

where $\theta_j \rightarrow i(t)$ is the probability that node $j$ is infected by its neighbors at time $t$ in the absence of node $i$ (i.e., node $i$ is in the cavity state). The first term on the right hand side of Eq. \ref{eq:rho} takes into account the probability that node $i$ recovers. The second term is the probability that it will be infected by its neighbors. If node $j$ recovers from the infected state, $\theta_j \rightarrow i(t)$ will decrease, but if node $j$ is infected by its neighbors in the absence of node $i$, $\theta_j \rightarrow i(t)$ will increase with a probability $\beta(1 - \rho_j(t)) \sum_{\ell \in N(j) \setminus i} \theta_\ell \rightarrow j(t)$, where $N(j)$ is the set of neighbors of node $j$. Combining these two factors, the evolution of $\theta_j \rightarrow i(t)$ can be written

$$\frac{d\theta_j \rightarrow i(t)}{dt} = -\theta_j \rightarrow i(t) + \beta[1 - \rho_j(t)] \sum_{\ell \in N(j) \setminus i} \theta_\ell \rightarrow j(t).$$

\label{eq:theta}

Using Eqs. \ref{eq:rho}–\ref{eq:theta} we obtain the evolution of the states of the nodes. Note that there will be $2E + N$ differential equations, where $N$ and $E$ are the number of nodes and edges. Initially $\theta_j \rightarrow i(0) \rightarrow 0$, since only a vanishing small fraction of nodes are in the infected state. Thus linearizing Eq. \ref{eq:theta} around $\theta_j \rightarrow i(0) = 0$, Eq. \ref{eq:theta} can be rewritten

$$\frac{d\theta_j \rightarrow i(t)}{dt} = \mathbf{B} \theta_j \rightarrow i(t) - \theta_j \rightarrow i(t),$$

where $\mathbf{B}$ is the non-backtracking matrix \cite{92}, and $\theta_j \rightarrow i(t)$ is an element of the vector $\vec{\theta}(t)$. The element of $\mathbf{B}$ is

$$B_{j \rightarrow i, \ell \rightarrow h} = \delta_{jh}(1 - \delta_{i\ell}),$$

\label{eq:B}

where $\delta_{i\ell}$ is the Dirac delta function. The physical meaning of $B_{j \rightarrow i, \ell \rightarrow h}$ is that if $i \neq \ell$ the edge $\ell \rightarrow h$ can influence edge $j \rightarrow i$. With arguments similar to those used to obtain Eq. \ref{eq:beta}, the epidemic threshold can be expressed

$$\beta_e^{\text{DMP}} = \frac{1}{\Lambda_B},$$

\label{eq:beta_DMP}

where $\Lambda_B$ is the largest eigenvalue of the non-backtracking matrix $\mathbf{B}$.

The DMP approach is widely used in such network science topics as spreading dynamics \cite{74, 41}, percolation \cite{93–95}, and cascading \cite{96–98}. It is widely applicable because (i) it describes the complete network structure by using the non-backtracking matrix, and (ii) it captures some dynamical correlations among the states of neighbors by assuming that "cavity" nodes cannot transmit messages. This means that the DMP approach produces exact results in networks that are tree-like. Through extensive numerical simulations, Shrestha et al. \cite{74} found that the DMP approach accurately predicts the SIS model in many real-world networks. Similar results were found for the SIR model in Refs. \cite{99, 100}.

Note that the DMP approach has two drawbacks, (i) the equations are highly complex, and (ii) it is inaccurate in non-local tree-like networks. We recall that it is difficult to analytically solve Eqs. \ref{eq:rho}–\ref{eq:theta} because there are $2E + N$ differential equations. To resolve drawback (i), we simplify the DMP approach by assuming that each edge has the same probability of connecting to infected neighbors. This simplified DMP (SDMP) approach and can be applied only to uncorrelated local tree-like networks (e.g., to uncorrelated configuration networks). Thus a susceptible node connects to an infected neighbor with a probability

$$\Theta(t) = \frac{1}{2E} \sum_{j \rightarrow i} \theta_j \rightarrow i(t),$$

\label{eq:Theta}

When we classify nodes according to their degree, for uncorrelated networks Eq. \ref{eq:Theta} can be rewritten

$$\Theta(t) = \frac{1}{(k)} \sum_{k} (k - 1) P(k) \rho_k(t),$$

\label{eq:Theta_k}

which was first derived by Barthélemy and his collaborators \cite{19, 53}. Inserting Eq. \ref{eq:Theta_k} into Eq. \ref{eq:beta}, Barthélemy et al. predicted the velocity and hierarchical structure of the epidemic spreading on scale-free networks. In this approach the epidemic threshold is

$$\beta_e^{\text{SDMP}} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}.$$

\label{eq:beta_SDMP}

To resolve drawback (ii) we decrease the "echo chamber" effect caused by finite loops. Radicchi and Castellano introduce a more complicated DMP approach to excluding redundant paths caused by triangles and obtain more accurate predictions on both artificial and real-world networks \cite{90}. We
still need to develop more accurate approaches to describing the dynamics in real-world networks with degree correlations, motifs, and community structures.

**Pairwise approximation (PA) approach.** The DMP approach cannot accurately capture the dynamic correlations in non-tree-like networks. The pairwise approximation (PA) approach best captures the dynamic correlations \[101,102\] by considering the evolution of the pair node states, instead of the evolution of the nodes. Denote \(\psi_{x,i}(t)\) as the probability that nodes \(i\) and \(j\) are in the \(x\) and \(y\) states and \(x \in \{S, I\}\), the following relationships are fulfilled, \(\psi_{I,I}(t) + \psi_{S,I}(t) = \rho_I(t), \psi_{S,S}(t) + \psi_{I,S}(t) = 1 - \rho_I(t), \) and \(\psi_{S,S}(t) + \psi_{I,S}(t) = 1 - \rho_I(t).\) With these equations in mind, Eq. (9) can be written \[103\]

\[
\frac{d\rho_I(t)}{dt} = -\rho_I(t) + \sum_{j=0}^{N} A_{ij} \psi_{S,I}(t). \tag{20}
\]

The first term is the probability that node \(i\) recovers from the infected state, and the second term is the probability that node \(i\) becomes infected by its neighbors. If we neglect the dynamical correlations between neighbors, i.e., \(\psi_{S,I}(t) = s_i(t) \rho_I(t),\) Eq. (20) reduces to Eq. (9).

In discussing the evolution of \(\psi_{S,I}(t)\), three events cause \(\psi_{S,I}(t)\) to decrease, i.e., (i) nodes \(j\) are infected from the infected state, (ii) node \(i\) is infected by neighboring node \(j\) with a probability \(\beta \psi_{S,I}(t),\) and (iii) the susceptible node \(i\) is infected by another neighbor node \(\ell\) with a probability \(\beta \sum_{\ell \in N(i) \setminus j} \phi_{I,I}(t),\) where \(\phi_{I,I}(t)\) is the probability that node \(i\) becomes infected by another neighbor node \(\ell\) with a probability \(\beta \sum_{\ell \in N(i) \setminus j} \phi_{S,S}(t) = 1 - \rho_I(t)\) and \(N(i)\) is the susceptible neighbor set of node \(i\). There are two events that cause \(\psi_{S,I}(t)\) to increase, (i) node \(i\) recovers from the infected state with a probability \(\psi_{I,I}(t),\) and (ii) the susceptible node \(j\) is infected by another neighbor node \(\ell\) with a probability \(\beta \sum_{\ell \in N(j) \setminus i} \phi_{S,S}(t)\), where \(\phi_{S,S}(t)\) is the probability that node \(i\) recovers from the infected state, and \(N(j)\) is the susceptible neighbor set of node \(j\). Based on this, the evolution of \(\psi_{S,I}(t)\) is given by

\[
\frac{d\psi_{S,I}(t)}{dt} = -\psi_{S,I}(t) - \beta \psi_{S,I}(t) + \beta \sum_{\ell \in N(i) \setminus j} \phi_{S,S}(t) + \beta \sum_{\ell \in N(j) \setminus i} \phi_{S,S}(t). \tag{21}
\]

To complete Eq. (21), we apply a pair approximation, i.e., we consider only the pair dynamic correlations as

\[
\phi_{x,x}(t) \approx \frac{\psi_{x,x}(t)}{x(t)}, \tag{22}
\]

where \(x(t)\) is the probability that node \(x\) is in the \(x\) state at time \(t\). Inserting Eq. (22) into Eq. (21) and combining it with Eq. (20), the evolution of the states of the SIS model can be described using \(N + E\) differential equations.

To obtain the epidemic threshold, we linearize Eq. (21) around the initial conditions \(\psi_{I,I}(0) \to 0\) and \(\psi_{S,S}(0) \to 1\). Using arguments similar to those for obtaining Eq. (8) we get the epidemic threshold when the largest eigenvalue of \(L\) is zero, where \(L\) is the Jacobian matrix of Eq. (21), and the elements of \(L\) are \[103\]

\[
L_{ij} = -(1 + \frac{\beta^2 k_i}{2\beta + 2}) \delta_{ij} + \frac{\beta(2 + \beta)}{2\beta + 2} A_{ij}. \tag{23}
\]

At the expense of increasing the complexity of the equations in the PA approach, we use the adjacency matrix to accurately describe the full network topology, and we capture the dynamical correlations among the states of neighbors by considering the evolution of the pair node states. Performing extensive simulations, Mata and Ferreira demonstrated that the epidemic size and threshold predictions of the PA approach are more accurate than those predicted using other methods, e.g., MF and HMF \[103\]. Although the PA approach can capture some of the dynamical correlations among the states of neighbors, solving the above equations numerically is time-consuming, which hinders its wide application. In order to reduce the number of equations researchers assume that all nodes of the same degree are statistically the same \[104,105\]. Thus we use \(k_{\text{max}}\) to describe the spreading dynamics because it decreases the complexity of the equations. Eames and Keeling, for example, used the PA approach to describe the spread of sexually transmitted diseases on heterogeneous networks, and their results are often in excellent agreement with simulations \[102\]. Gross et al. used the PA approach to capture the assortative degree correlation, oscillations, hysteresis, and first order transition when an epidemic spreads on an adaptive network \[105,107\]. Kiss et al. used the PA approach to theoretically predict non-Markovian epidemic spreading dynamics \[108\]. Recently researchers developed a generalized PA approach to study epidemic spreading on weighted complex networks \[52,109\].

**B. Specific approaches using the SIR model**

**Link percolation (LP) approach.** In contrast to the reversible SIS model, the irreversible SIR model allows us to examine the final state of the epidemic at which an individual is either susceptible or recovered. The most commonly used approach is link percolation (LP) approach, and the most studied version is the time-continuous Kermack-McKendrick \[110\] formulation in which an infected individual transmits the disease to a susceptible neighbor at a rate \(\lambda\) and recovers at a rate \(\gamma\). This SIR version has been widely studied in the epidemiological literature, but unfortunately it allows some individuals to recover immediately after being infected, which is unrealistic since any real-world disease has a characteristic average recovery time. To overcome this shortcoming, many studies use the discrete Reed-Frost model \[111\] in which an infected individual transmits the disease to a susceptible neighbor with a probability \(\lambda\) and recovers \(t_r\) steps following the time of infection. In the discrete updating method the transmissibility \(\beta\) is the probability that an individual will infect one susceptible neighbor before recovery, and it is given
by
\[ \beta = \sum_{u=1}^{\tau} \lambda (1 - \lambda)^{u-1} = 1 - (1 - \lambda)^{\tau}. \quad (24) \]

Note that in the continuous time updating approach \( \beta \approx 1 - e^{-\lambda/\gamma} \approx \lambda/\gamma \) is used in Ref. [23].

The order parameter \( M_R(\beta) = M_R \), which is the final fraction of recovered nodes, overcomes a second-order phase transition at the epidemic threshold \( \beta_c^{LP} \), which is determined by the network structure. Note that the Reed-Frost model can be mapped into a link percolation process [115,118]. Heuristically, the relation between SIR and link percolation is sustained because the probability \( \beta \) that a link is traversed by the disease is equivalent to the occupancy probability \( p \) in link percolation. Thus both processes have the same threshold and belong to the same universality class. In addition, each realization of the SIR model corresponds to a single cluster of link percolation. This feature is relevant when mapping between the order parameters \( g(p=\beta) = g \) of link percolation and \( M_R \) for epidemics, as we will explain below. In a SIR realization, only one infected cluster emerges for any value of \( \beta \). In contrast, in a percolation process when \( p < 1 \) many clusters with a cluster size distribution are generated [115]. Thus we need criteria to distinguish between epidemics (the giant connected cluster in percolation) and outbreaks (finite clusters). The cluster size distribution over many realizations of the SIR process, close to but above criticality, has a gap between small clusters (no epidemics) and large clusters (epidemics). Thus when defining a cutoff \( s_c \) of the cluster size as the minimum value before the gap interval, the cluster sizes below \( s_c \) are not considered epidemics but those above \( s_c \) are (see Fig. 3(a)).

Note that \( s_c \) depends on \( N \). Then averaging those SIR realizations larger than the cutoff \( s_c \) we find that the fraction of recovered individuals \( M_R \) maps exactly with \( g \) (see Fig. 3(b)). In our simulations, we use \( s_c = 200 \) for \( N = 10^5 \).

When we use a cutoff close to criticality, all the exponents that characterize the transition are the same for link percolation and epidemic spreading [114,118]. Above but close to \( \beta_c^{LP} \) we have
\[ M_R(\beta) \sim (\beta - \beta_c^{LP})^{\alpha}; \quad g \sim (p - p_c)^{\alpha}, \quad (25) \]
with \( [119] \)
\[ \alpha = \begin{cases} 1 \quad &\text{for SF with } \nu \geq 4 \text{ and ER networks, for } 3 < \nu < 4, \\ \frac{1}{\nu - 3} \quad &\text{for } 3 < \nu < 4. \end{cases} \quad (27) \]

The exponent \( \tau \) of the finite cluster size distribution in percolation close to criticality is given by
\[ \tau = \begin{cases} \frac{5}{2} \quad &\text{for SF with } \nu \geq 4 \text{ and ER networks;} \\ \frac{1}{\nu - 2} + 2 \quad &\text{for } 2 < \nu < 4. \end{cases} \quad (28) \]

Near criticality the probability of a cluster of size \( s \), \( P(s) \), has an exponent \( \tau - 1 \) in which \( \tau \) is given by Eq. (23) (see Fig. 3). For SF networks with \( \nu \leq 3 \) in the thermodynamic limit, the threshold is zero and there is no percolation phase transition. In addition, all the exponents take mean-field values when \( \nu \geq 4 \) for SF networks and ER networks.

In an uncorrelated network with a degree distribution \( P(k) \), the probability of reaching a node with a degree \( k \) by following a randomly chosen link on the graph is \( kP(k)/\langle k \rangle \), where \( \langle k \rangle \) is the average degree. This is because the probability of reaching a given node by following a randomly chosen link is proportional to the number of links \( k \) connected to that node, and \( \langle k \rangle \) is needed for normalization. Note that when we arrive to a node with a degree \( k \) by following a random chosen link, the total number of outgoing links or branches of that node is \( k - 1 \). Thus the probability of arriving at a node with \( k - 1 \) outgoing branches by following a randomly chosen link is also \( kP(k)/\langle k \rangle \). This value is the excess degree probability [120,121].

To obtain the threshold of link percolation, we consider a randomly-chosen occupied link. To compute the probability that through this link an infinite cluster will not be reached we assume, for simplicity, that we have a Cayley tree with a given degree distribution. Note that link percolation can be thought of as many realizations of a Cayley tree with an occupancy

FIG. 3: (Color online) Effects of the cutoff \( s_c \) on the mapping between the SIR model and link percolation for a Poisson degree distribution network (ER) with \( \langle k \rangle = 4 (\beta_c^{LP} = 0.25), N = 10^5 \). In (a) we show the probability \( P(s) \) of a cluster of size \( s \) (including the size of the giant component) in the SIR model for \( \beta = 0.27 (\bigcirc) \) and \( \beta = 0.40 (\square) \). We can see that the gap between the epidemic sizes and the distribution of outbreaks increases with \( \beta \). As \( \beta \) is at the threshold the slope is \( \tau - 1 = 3/2 \) marked by dashed lines and \( s_c \approx 200 \). In Fig. (b) we show the simulation results for \( M_R \) for \( s_c = 1 (\bigcirc) \) and \( s_c = 200 (\square) \). Note that when \( s_c = 200 \), we average the final size of infected clusters only over epidemic realizations. Considering only the conditional averages, we can see that \( M_R \) maps with \( g \) (solid line) for \( s_c = 200 \). The simulations are averaged over \( 10^7 \) realizations.
probability $p$ that gives rise to many clusters. The probability that when starting from an occupied link we will not reach shell $n$ through a path of occupied links is given by

\[
Q_n(p) = \sum_{k=1}^{\infty} k P(k) \frac{(1 - p) + p Q_{n-1}(p))^{k-1}}{\langle k \rangle},
\]  

(29)

\[
= G_1[(1 - p) + p Q_{n-1}(p)),
\]  

(30)

where $G_1(x) = \sum_{k=1}^{\infty} k P(k) / \langle k \rangle x^{k-1}$ is the generating function of the excess degree distribution. As $n$ increases, $Q_n \approx Q_{n-1} = u$ and the probability that we will not reach an infinite cluster is

\[
Q_n(p) = G_1[(1 - p) + p u],
\]  

(31)

Thus the probability that the starting link connects to an infinite cluster is $f_\infty(p) = 1 - u$. From Eq. (31), $f_\infty(p)$ is given by

\[
f_\infty(p) = 1 - G_1[1 - p f_\infty(p)].
\]  

(32)

The solution to Eq. (32) can be geometrically understood as the intersection of the identity line $y = x$ and $y = 1 - G_1(1 - px)$, which has at least one solution at the origin. $x = f_\infty(p) = 0$, for any value of $p$. If the derivative of the right hand side of Eq. (32) with respect to $x$ is $[1 - G_1(1 - px)]'_{x=0} = p G'_1(1) > 1$, we will have another solution in $0 < x \leq 1$. This solution $x = f_\infty(p)$ is the probability that a randomly-selected occupied link is connected to an infinite cluster for a given value of $p$. The criticality corresponds to the value of $p = p_c$, at which the curve $1 - G_1(1 - px)$ has a slope equal to one. Thus $p_c$ is given by

\[
p_c = \frac{1}{G'_1(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle},
\]  

(33)

which is the same epidemic threshold as that obtained using the SDMP approach [see Eq. (19)]. On the other hand we can obtain the order parameter of link percolation $g$, which represents the fraction of nodes that belongs to the giant cluster when a fraction $p$ of links are occupied. The probability that a node with degree $k$ does not belong to the giant component is given by the probability that none of its links connect the node to the giant connected cluster (GCC), i.e., $[1 - p f_\infty(p)]^k$. Thus the fraction of nodes that belong to the GCC is $g = 1 - \sum_{k=0}^{\infty} P(k) [1 - p f_\infty(p)]^k$. Since the relative epidemic size in the SIR model maps exactly with the relative size of the giant connected cluster, we find that

\[
M_R = g = 1 - G_0[1 - p f_\infty(p)],
\]  

(34)

where $G_0(x) = \sum_{k=0}^{\infty} P(k) x^k$ is the generating function of the degree distribution and $f_\infty(p)$ is the non-trivial solution to Eq. (32) for $p > p_c$. It is straightforward to show that in ER networks $G_0(x) = G_1(x) = \exp[-\langle k \rangle (1 - x)]$ and thus $f_\infty(p) = M_R$. In pure SF networks with $1 \leq k < \infty$ the generating function of the excess degree distribution is proportional to the poly-logarithm function $G_1(x) = Lt_\lambda(x)/\xi(\lambda)$, in which $\xi(\lambda)$ is the Riemann function [121].

The LP approach assumes that every link does not connect to the GCC with the same probability $u$. Note that when we calculate $u$ only the outgoing branches are considered in Eq. (31) and some dynamical correlations are thus captured. In addition, the LP approach uses the degree distribution to describe the network topology. Thus the LP approach can predict the final epidemic size and threshold on networks with an infinite uncorrelated local tree-like configuration [23]. Note that SIR spreading is a dynamical infection process with an intricate interplay between complex network structure and dynamical correlations, and this differs from the static link percolation model. Thus the network topology, the time evolutions, and the dynamical correlations among the states of neighbors cannot be described using the classical LP approach, especially near the critical point [123]. For epidemic spreading on finite size networks or when there is a non-uniform infectious time distribution, research has shown that the final state of the SIR model differs from that of the link percolation model, including in particular the epidemic threshold, mean epidemic size, and epidemic size distribution [114, 123]. A number of advanced LP approaches have been developed to address specific configurations. For example, Miller et al. [124] and Allard et al. [125] generalized the LP approach to study nodes with different levels of infection and susceptibility, Noël et al. developed the LP approach for finite uncorrelated tree-like networks [126], and Marder generalized the LP approach to obtain the time evolution of the SIR model [127]. Recently some researchers also developed LP approaches to address the effects of clustering [57, 128], degree-degree correlations [129], community structure [130], and multiplexity [131, 132] in the SIR model. In addition, Newman derived the LP approach for multiple epidemic spreading dynamics, analyzed the interactions between the epidemics [133, 135], and found a co-infection condition in the interacting epidemics. Parshani et al. developed a modified LP approach to predicting the threshold of the SIS model [32].

Message passing (MP) approach. In the LP approach the probability of reaching the infinite cluster by following a randomly chosen link on the graph is assumed to be the same for all links. This assumption is true for uncorrelated tree-like networks but is not valid for real-world networks, which may have, e.g., degree correlation, clustering, and communities. To take these into account, Karrer and Newman developed the message passing (MP) approach [92] to study the final state of the SIR model, which differs from the description supplied above. They assumed that $z_{j \rightarrow i}$ is the probability that node $j$ is infected by its neighbors in the absence of node $i$, i.e., node $i$ is in the cavity state. When there is a vanishing small fraction of initially infected nodes, $z_{j \rightarrow i}$ satisfies the relationship

\[
z_{j \rightarrow i} = 1 - \prod_{\ell \in N(j) \setminus i} (1 - \beta z_{\ell \rightarrow j}),
\]  

(35)

where $\prod_{\ell \in N(j) \setminus i} (1 - \beta z_{\ell \rightarrow j})$ is the probability that node $j$ is not infected by any neighbors in the absence of node $i$. Node
Thus the relative epidemic size is given by

$$f_i = 1 - \prod_{j \in N(i)} (1 - \beta z_{j \to i}). \quad (36)$$

Thus the relative epidemic size is given by

$$g = \frac{1}{N} \sum_{i=1}^{N} f_i. \quad (37)$$

To obtain the value of \(z_{j \to i}\) we iterate Eq. (35) from a random initial value and substitute the results into Eq. (36) and, using Eq. (37), we obtain the relative epidemic size of the SIR model. From Eq. (36),

$$\ln(1 + z_{j \to i}) = \sum_{\ell} A_{j\ell} \ln(1 - \beta z_{\ell \to i}) - A_{ji} \ln(1 - z_{j \to i}). \quad (38)$$

Defining the vectors \(\vec{u}\) and \(\vec{v}\) whose \((j \to i)\)-th components are \(u_{j \to i} = \ln(1 + z_{j \to i})\) and \(v_{j \to i} = \ln(1 - \beta z_{j \to i})\), respectively, Eq. (38) can be written

$$\vec{u} = B \vec{v}, \quad (39)$$

where \(B\) is the non-backtracking matrix of the network. If a global epidemic breaks out, Eq. (39) will have a nontrivial solution. Thus the epidemic threshold is the inverse of the largest eigenvalue of matrix \(B\), which is the same as that described in Eq. (16) for the SIS model predicted by the DMP approach.

The advantages and drawbacks of the MP approach are the same as those in the DMP approach (see details above). Unlike the DMP approach, which describes the time evolution of the spreading dynamics, the MP approach uses different formulas and considers only the final state of the SIR spreading dynamics. Since the MP approach uses a non-backtracking matrix that allows a description of the full structure of the network but disallows nodes in the ‘cavity’ state to transmit the epidemic, it accurately predicts the epidemic size and threshold in artificial and in some real-world networks [100]. Recently the MP approach has been used to control the spread of an epidemic [71], to identify patient zero [134, 137], and to locate the most influential seeds [138, 139]. Using the MP approach, Morone and Makse studied influence maximization in complex networks through optimal percolation and found that the low-connected nodes play an important role in influencing maximization problems [138]. Hu et al. discovered that the influence maximization problem is a local optimization problem, not a global one [139].

**Edge-based compartmental (EBC) approach.** Because the LP and MP approaches are static, they are usually used to address the final state of the SIR model. To investigate the time evolution of the SIR model, the edge-based compartmental (EBC) approach was developed [73, 140, 143]. It is based on the cavity theory (i.e., the MP approach) in which a node \(i\) in the cavity state cannot transmit the infection to its neighbors but can be infected by its neighbors. Unlike the MP approach in which each edge has a different probability of transmitting the infection to its neighbors, the EBC approach makes the same assumption as the LP approach, i.e., the probability of infection transmitted through each link is the same. The EBC approach is based on a generating function formalism widely applied to branching and percolation processes in complex networks. The fraction of susceptible, infected, and recovered individuals at time \(t\) are denoted \(s(t)\), \(\rho(t)\), and \(r(t)\), respectively. The EBC approach describes the evolution of the probability that a denoted root will be susceptible. To compute this probability, an edge is randomly chosen and a direction in which a node \(j\) on the target of the arrow is the root, and the base is one of its neighbors. Disallowing the root \(j\) to infect its neighbors, \(\Phi(t)\) is the probability that neighbor \(i\) does not transmit the disease to root \(j\), with \(\Phi(t)\) given by

$$\Phi(t) = \xi_S(t) + \xi(I(t) + \xi_R(t)), \quad (40)$$

where \(\xi_S(t)\), \(\xi_R(t)\), and \(\xi_I(t)\) are the probabilities that the neighbor is susceptible, recovered, or infected but has not yet transmitted the disease to the root node \(j\) [see Fig. (4a)]. The probability that node \(j\) with connectivity \(k\) is susceptible is thus \(\Phi(t)^k\), and the fraction of susceptible nodes is given by

$$s(t) = \sum_k P(k)\Phi(t)^k = G_0(\Phi(t)). \quad (41)$$

**Figure 4b** shows a schematic of this model. We next solve \(\xi_S(t), \xi_I(t), \) and \(\xi_R(t)\). A neighbor node \(i\) of the root node \(j\) can only be infected by neighbors other than \(j\). Then node \(i\) is susceptible with a probability

$$\xi_S(t) = \sum_k P(k)\Phi(t)^k \xi_R(t)^{k-1} \langle k \rangle = G_1(\Phi(t)), \quad (42)$$

where \(P(k)\langle k \rangle\) is the probability that an edge connects a node with degree \(k\) in an uncorrelated network [see Fig. (4c)]. In the discrete updating method there are two conditions that allow the increase of \(\xi_R(t)\), i.e., (i) the infected node has not transmitted the infection to \(j\) with a probability \(1 - \beta\), and (ii) the infected node is removed with a probability \(1\). Taking these two events into consideration, the evolution of \(\xi_R(t)\) is given by

$$\frac{d\xi_R(t)}{dt} = (1 - \beta)\xi_I(t). \quad (43)$$

At time \(t\) the rate of change in the probability that a random edge has not transmitted the infection is equal to the rate at which the infected neighbors transmit the infection to their susceptible neighboring nodes through edges. Thus

$$\frac{d\Phi(t)}{dt} = -\beta\xi_I(t). \quad (44)$$

Combining Eqs. (43) and (44) with initial conditions \(\Phi(0) = 1\) and \(\xi_R(0) = 0\), we obtain \(\xi_R(t) = [1 - \Phi(t)](1 - \beta)/\beta\), which together with Eq. (43) and Eq. (40) allows us to obtain the evolution of \(\Phi(t)\),

$$\frac{d\Phi(t)}{dt} = -\beta\Phi(t) + \beta G_1(\Phi(t)) + [1 - \Phi(t)](1 - \beta). \quad (45)$$
Thus we have more accurately predicted the epidemic size and threshold than as the only input parameter to describe network topology, it though the EBC approach also uses the degree distribution the time evolutions of SIR spreading into consideration. Al-

\[ \beta_c^{EBC} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \]

In the continuous updating method Eq. (43) is rewritten

\[ \frac{d\xi_R(t)}{dt} = \xi_I(t). \]

Thus we have \( \xi_R(t) = \frac{1 - \Phi(t)}{\beta} \) and obtain the epidemic threshold

\[ \beta_c^{EBC} = \frac{\langle k \rangle}{\langle k^2 \rangle - 2\langle k \rangle}. \]

Unlike the LP and MP approaches, the EBC approach takes the time evolutions of SIR spreading into consideration. Although the EBC approach also uses the degree distribution as the only input parameter to describe network topology, it more accurately predicts the epidemic size and threshold than the HMF approach. The EBC approach is based on the cavity theory in which a node in the cavity state cannot transmit infection to its neighbors but can be infected by its neighbors. Thus the EBC approach can capture some of the dynamical correlations among the states of neighbors. Researchers have found that the EBC approach is exact for the SIR model on infinite uncorrelated local tree-like networks [24, 144–147] not only in reproducing the dynamics but also in determining the final state of the model. For example, Wang et al. generalized the EBC approach to study epidemic spreading on weighted networks, and found that increasing the heterogeneity of the weight distribution decreases the size of the epidemic and increases the threshold [24]. Recently these same authors developed the EBC approach for a non-Markovian social contagion and found a transition in which the final adoption size depends on such key parameters as the transmission probability, which can change from discontinuous to continuous [148]. The transition can be triggered by such parameters and structural perturbations to the system as decreasing the adoption threshold of individuals, decreasing the heterogeneity of the adoption threshold, increasing the initial seed size or contact capacity, or enhancing network heterogeneity [148–150].

### III. DISCUSSION AND OUTLOOK

We have illustrated seven widely used approaches to the dynamics of epidemic spreading, including the MF, HMF, QMF, DMP, LP, EBC, and PA. Other not widely used approaches, such as master equations, are described in Refs. [151–153]. Table II shows which characteristic behaviors are described by each approach. Note that all seven approaches can be used in the irreversible SIR model, but that prior to now the EBC has not been used in the reversible SIS model. Unfortunately none of these approaches can adequately describe both the full topology of a network and its dynamic correlations. Table III shows that in order to more accurately capture both network topology and dynamic correlations, the number of required equations increases and they become increasingly complex.

To describe network topology, we use the adjacency matrix and non-backtracking matrix as illustrated in the QMF and DMP approaches, respectively. To capture the dynamical correlations, we use cavity theory to prevent the ‘echo chamber’
effect or the evolution of pair node states.

To capture both network topology and dynamical correlations, we adopt the continuous-time Markov (CTM) approach and find exact results for epidemic spreading. The CTM approach uses the adjacency matrix to describe the network topology, and uses the transform matrix generator $Q_{q^N \times q^N}$ to describe the evolution of the epidemic spreading and the dynamic correlations. Once the value of $Q$ is obtained, the probability that a node will be in each state can be computed. Although the CTM approach provides an exact description it is not widely used in the field of spreading dynamics because the generator $Q_{q^N \times q^N}$ is difficult to obtain, and also it is difficult to solve the complicated equations, especially for large scale networks. The CTM can be used, however, to obtain exact solutions in a few specific scenarios of the SIS model [156].

For a given epidemic spreading dynamics, these theoretical approaches sometimes yield different epidemic sizes and thresholds [84, 157]. Wang et al. classified the theoretical approaches into three categories according to the topological information used [100]. The first is the mean-field like (MFL) approach, which uses the degree distribution as the sole input parameter. This category includes the HMF, the LP, EBC, and PA approaches. The second type is the quenched mean-field (QMF) approach, which describes the topology of each network using the adjacency matrix. Examples include the discrete-time Markov chain and the $N$-intertwined approach [86, 87]. The third type is the dynamic message passing (DMP) approach, which describes network topology in terms of the non-backtracking matrix. Wang et al. determined the effectiveness of these three approaches using extensive numerical simulations of the SIR model on artificial and real-world networks [104]. For configuration networks they found that the MFL and the DMP approaches perform better than the QMF approach [see Figs. 5(a) and 5(b)]. For real-world networks, the DMP approach performs well in most situations [see Figs. 5(c)].

In summary, we began by describing in terms of increasing complexity the seven most popular theoretical approaches. We explain their main ideas and basic assumptions, and we describe the relationships among them. These approaches have also been widely used in studying the dynamics of social contagions [148–150, 158–162]. As network science has developed and expanded, many of the existing theoretical approaches have been challenged, and we now must take into consideration numerous intricate mechanisms and network topologies when we build epidemic spreading models.

The first challenge is how to describe epidemic spreading on complex real-world networks that are, for example, multilayer or temporal. With the availability of real-world data, many researchers believe that treating real-world networks as single or static networks is no longer a viable approach, and that one must utilize multilayer and temporal networks [8, 9]. Although we may want to adopt the tensor formalism to describe the topology of multilayer and temporal networks, the network topology and dynamic correlations are difficult to capture with this formalism. In addition, different spreading mechanisms such as preference spreading [60] and layer-switching [166] are produced, and the intricate network topology further increases the difficulty of describing the spreading of an epidemic.

The second challenge is how to describe epidemic spreading once human behavior is included. Human behavior will markedly affect epidemic spreading dynamics [31, 167], due to burst, memory [168, 171], and mobility effects [172, 173]. These features induce a non-Markovian effect in the spreading dynamics that causes strong dynamic correlations that are difficult to describe [71]. The existing theoretical approaches can address only some specific situations, and a general framework for non-Markovian spreading dynamics is still lacking.

FIG. 5: (Color online) Predicting the epidemic threshold for the SIR model on uncorrelated networks and 56 real-world networks. Theoretical predictions of $\beta_{cMFL}$ (black solid lines), $\beta_{cQMF}$ (red dashed lines), $\beta_{cDMP}$ (blue dish-dotted lines) and numerical predictions (gray squares) versus network size $N$ for power-law degree distribution $P(k) \sim k^{-\nu}$ with degree exponent $\nu = 2.1$ (a) and $\nu = 3.5$ (b). In (c), each symbol a threshold of a real-world network. $\beta_{cMFL}$, $\beta_{cQMF}$ and $\beta_{cDMP}$ are the theoretical predictions by the MFL, QMF and DMP, respectively. The value of $\beta_c$ is the numerical prediction by using a variability measure $\Delta = \sqrt{\langle r^2 \rangle - \langle r \rangle^2}$ [15].
A third challenge is how to describe coevolution spreading dynamics. In real-world systems when two strains of the same disease spread in the same population and interact through cross-immunity [91, 178, 179] or mutual reinforcement [180], the information from each competes for the limited attention-span of the participants [181], and there is an asymmetric interaction between the spread of information and the spread of the epidemic [75, 79, 182, 183]. An accurate, unified theoretical approach for coevolution dynamics is still lacking and presents great challenges because, in this case, the dynamic correlations are enhanced.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under Grant Nos. 11105025, 11575041, 61433014 and China Scholarship Council.

The Boston University work was supported by DTRA Grant HDTRA1-14-1-0017, by DOE Contract DE-AC07-05Id14517, and by NSF Grants CMMI 1125290, PHY 1505000, and CHE-1213217. LAB knowledge the support of UNMdP and FONCyT, PICT 0429/13.

References

[1] D. Bernoulli, Math. Phys. Acad. Roy. Sci. Paris 1 (1766).
[2] M. J. Keeling and P. Rohani, Modeling infectious diseases in humans and animals (Princeton University Press, 2008).
[3] W. O. Kermack and A. G. McKendrick, R. Soc. Lond. A 115, 700 (1927).
[4] R. M. Anderson, R. M. May, and B. Anderson, Infectious diseases of humans: dynamics and control, vol. 28 (Wiley Online Library, 1992).
[5] M. Newman, Networks: an introduction (Oxford university press, 2010).
[6] R. Albert and A.-L. Barabási, Reviews of modern physics 74, 47 (2002).
[7] M. E. Newman, SIAM review 45, 167 (2003).
[8] P. Holme and J. Saramäki, Physics reports 519, 97 (2012).
[9] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, Physics Reports 544, 1 (2014).
[10] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley, Physical review letters 91, 168701 (2003).
[11] Z. Wu, L. A. Braunstein, S. Havlin, and H. E. Stanley, Physical review letters 96, 148702 (2006).
[12] R. Pastor-Satorras and A. Vespignani, Physical review letters 86, 3200 (2001).
[13] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Reviews of modern physics 87, 925 (2015).
[14] S. C. Ferreira, C. Castellano, and R. Pastor-Satorras, Physical Review E 86, 041125 (2012).
[15] P. Shu, W. Wang, M. Tang, and Y. Do, Chaos: An Interdisciplinary Journal of Nonlinear Science 25, 063104 (2015).
[16] P. G. Fennell, S. Melnik, and J. P. Gleeson, arXiv preprint arXiv:1603.01132 (2016).
[17] P. Shu, W. Wang, M. Tang, P. Zhao, and Y.-C. Zhang, Chaos 26, 063108 (2016).
[18] R. Pastor-Satorras and C. Castellano, Scientific reports 6 (2016).
[19] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Physical Review Letters 92, 178701 (2004).
[20] A.-X. Cui, W. Wang, M. Tang, Y. Fu, X. Liang, and Y. Do, Chaos: An Interdisciplinary Journal of Nonlinear Science 24, 033113 (2014).
[21] P. Shu, M. Tang, K. Gong, and Y. Liu, Chaos: An Interdisciplinary Journal of Nonlinear Science 22, 043124 (2012).
[22] P. Crepey, F. P. Alvarez, and M. Barthélemy, Physical Review E 73, 046131 (2006).
[23] M. E. Newman, Physical review E 66, 016128 (2002).
[24] W. Wang, M. Tang, H.-F. Zhang, H. Gao, Y. Do, and Z.-H. Liu, Physical Review E 90, 042803 (2014).
[25] R. Yang, L. Huang, and Y.-C. Lai, Physical Review E 78, 026111 (2008).
[26] Q. Wu, X. Fu, Z. Jin, and M. Small, Physica A 419, 566 (2015).
[27] Y. Liu, Y. Deng, M. Jusup, and Z. Wang, Journal of theoretical biology 400, 92 (2016).
[28] Z. Wang, D.-W. Zhao, L. Wang, G.-Q. Sun, and Z. Jin, EPL (Europhysics Letters) 112, 48002 (2015).
[29] M. Boguñá, C. Castellano, and R. Pastor-Satorras, Physical review letters 111, 068701 (2013).
[30] C. Castellano and R. Pastor-Satorras, Physical review letters 105, 218701 (2010).
[31] P. Van Mieghem and R. Van de Boovenkamp, Physical review letters 110, 108701 (2013).
[32] R. Parshani, S. Carmi, and S. Havlin, Physical review letters 104, 258701 (2010).
[33] H. K. Lee, P.-S. Shim, and J. D. Noh, Physical Review E 87, 062812 (2013).
[34] Q. Wu, X. Fu, M. Small, and X.-J. Xu, Chaos 33, 013101 (2012).
[35] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and A. Makse, Nature physics 6, 888 (2010).
[36] S. Pei and H. A. Makse, Journal of Statistical Mechanics: Theory and Experiment 2013, P12002 (2013).
[37] L.-F. Zhong, J.-G. Liu, and M.-S. Shang, Physics Letters A 379, 2272 (2015).
[38] Y. Liu, M. Tang, T. Zhou, and Y. Do, arXiv preprint arXiv:1409.5187 (2014).
[39] R. Pastor-Satorras and A. Vespignani, Physical Review E 65, 036104 (2002).
[40] R. Cohen, S. Havlin, and D. Ben-Avraham, Physical review letters 91, 247901 (2003).
[41] H. Yang, M. Tang, and H.-F. Zhang, New Journal of Physics 14, 123017 (2012).
[42] A. V. Goltsev, S. N. Dorogovtsev, J. Oliveira, and J. F. Mendes,
TABLE II: Definitions of parameters and abbreviations.

| Parameter/Abbreviation | Definition |
|------------------------|------------|
| Mean-field approach | Mean-field approach |
| HMF approach | Heterogeneous mean-field approach |
| QMF approach | Quench mean-field approach |
| DMP approach | Dynamical message-passing approach |
| PA approach | Pairwise approximation approach |
| LP approach | Link percolation approach |
| EBC approach | Edge-based compartmental approach |
| CTM approach | Continuous-time Markov approach |
| GCC | Giant connected cluster |
| $P(k)$ | Degree distribution |
| $A$ | Adjacency matrix |
| $B$ | Non-backtracking matrix |
| $E$ | Number of edges in the network |
| $N$ | Network size |
| $p$ | Edge occupancy probability |
| $q$ | The relative size of the GCC |
| $u$ | The endpoint of a randomly selected edge is not connected to the GCC |
| $t_r$ | Recovery time |
| $f_i$ | Probability that node $i$ is infected by neighbors |
| $N(j)$ | Neighbor set of node $j$ |
| $G_0(x)$ | Generation function of the degree distribution |
| $G_1(x)$ | Generation function of the excess degree distribution |
| $M_R$ | The final fraction of recovered nodes |
| $\nu$ | Exponent of power-law degree distribution |
| $\langle k \rangle$ | Average degree |
| $\langle k^2 \rangle$ | Second moment of the degree distribution |
| $\beta_{\text{max}}$ | Maximum degree of networks |
| $\lambda$ | Infection transmission rate |
| $\gamma$ | Recovery rate |
| $\beta$ | Effective transmission rate |
| $p(t)$ | Fraction of infected node at time $t$ |
| $p'(t)$ | Vector of $p_k(t)$ |
| $p(\infty)$ | Fraction of infected node in the steady state |
| $s(t)$ | Fraction of susceptible node at time $t$ |
| $s_c$ | Cutoff in the cluster size to split epidemic from outbreaks |
| $p_s(t)$ | At time $t$, the fraction of infected node with degree $k$ |
| $s_k(t)$ | At time $t$, the fraction of susceptible node with degree $k$ |
| $\Theta(t)$ | Probability that a susceptible node connects to an infected neighbor |
| $p_i(t)$ | Probability of node $i$ is infected at time $t$ |
| $s_i(t)$ | Probability of node $i$ is in the susceptible state at time $t$ |
| $r(t)$ | Fraction of removed node at time $t$ |
| $\beta_c$ | Epidemic threshold |
| $p_c$ | Critical edge occupied probability |
| $\beta_{\text{MF}}$ | Epidemic threshold predicted by the MF approach |
| $\beta_{\text{HMF}}$ | Epidemic threshold predicted by the HMF approach |
| $\beta_{\text{QMF}}$ | Epidemic threshold predicted by the QMF approach |
| $\beta_{\text{DMP}}$ | Epidemic threshold predicted by the DMP approach |
| $\beta_{\text{QMF}}$ | Epidemic threshold predicted by the simplified QMF approach |
| $\beta_{\text{LP}}$ | Epidemic threshold predicted by the LP approach |
| $\beta_{\text{EBC}}$ | Epidemic threshold predicted by the EBC approach |
| $C$ | Jacobian matrix of $C_{kk'} = \beta k k' P(k')/\langle k \rangle - \delta_{k,k'}$ |
| $\delta_{k,k'}$ | Dirac delta function |
| $\Lambda_A$ | Largest eigenvalue of adjacency matrix |
| $\Lambda_B$ | Largest eigenvalue of non-backtracking matrix |
| $\tilde{f}(\Lambda_A)$ | Eigenvector of $\Lambda_A$ |
| $f_i(\Lambda_A)$ | The $i$-th element of the eigenvector $\tilde{f}(\Lambda_A)$ of $\Lambda_A$ |
| $f_{\infty}(p)$ | Probability that a random chosen link connects to the GCC |
| $\rho(t)$ | Probability that node $i$ connects to the GCC |
| $\theta_{j \rightarrow i}(t)$ | Probability that node $j$ is infected by neighbors at time $t$ in the absence of neighbor $i$ |
| $z_{j \rightarrow i}(t)$ | Probability that node $j$ is infected by neighbors in the absence of neighbor $i$ |
| $\psi_{x_i,x_j}(t)$ | Probability that nodes $i$ and $j$ are in the $x_i$ and $x_j$ state, respectively |
| $\Phi(t)$ | Probability that a randomly selected edge has not transmitted the infection to a neighbor by time $t$ |
| $\xi(t)$ | Probability that a neighbor of a node in the susceptible state at time $t$
[1] G. Ódor, Physical Review E 88, 032109 (2013).
[2] T. Voigt, Journal of Physics A: Mathematical and General 39, R143 (2006).
[3] G. Buono, F. Vazquez, P. Macri, and L. Braunstein, Physical Review E 88, 022813 (2013).
[4] J. D. Noh and H. Park, Physical Review E 79, 056115 (2009).
[5] T. Tomé and M. J. de Oliveira, Physical Review E 72, 026130 (2005).
[6] R. Pastor-Satorras and A. Vespignani, Physical Review E 65, 035108 (2002).
[7] Y. Gang, Z. Tao, W. Jie, F. Zhong-Qian, and W. Bing-Hong, Chinese Physics Letters 22, 510 (2005).
[8] Z. Yang and T. Zhou, Physical Review E 85, 056106 (2012).
[9] L. Wang and X. Li, Chinese Science Bulletin 59, 3511 (2014).
[10] C. Kamp, M. Moslonka-Lefebvre, and S. Alizon, PLoS Comput Biol 9, e1003352 (2013).
[11] Y. Sun, C. Liu, C.-X. Zhang, and Z.-K. Zhang, Physics Letters A 378, 635 (2014).
[12] Y. Moreno, J. B. Gómez, and A. F. Pacheco, Physical Review E 68, 035103 (2003).
[13] R.-S. Wang and R. Albert, Physical Review E 87, 012810 (2013).
[14] Y. Min, X. Jin, Y. Ge, and J. Chang, PLoS one 8, e57100 (2013).
[15] M. E. Newman, Physical review letters 103, 058701 (2009).
[16] Z. Liu and B. Hu, EPL (Europhysics Letters) 72, 315 (2005).
[17] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Journal of theoretical biology 235, 275 (2005).
[18] E. H. Xu, W. Wang, C. Xu, M. Tang, Y. Do, and P. Hui, Physical Review E 92, 022812 (2015).
[19] M. E. Newman, Physical review e 64, 016131 (2001).
[20] M. E. Newman, Physical review e 64, 016132 (2001).
[21] M. E. Newman, Physical review e 64, 025102 (2001).
[22] A. Clauset, C. R. Shalizi, and M. E. Newman, SIAM review 51, 661 (2009).
[23] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science 298, 824 (2002).
[24] A.-X. Cui, Z.-K. Zhang, M. Tang, P. M. Hui, and Y. Fu, PLoS one 7, e50702 (2012).
[25] M. E. Newman, Proceedings of the national academy of sciences 103, 8577 (2006).
[26] A. Clauset, M. E. Newman, and C. Moore, Physical review e 70, 066111 (2004).
[27] C. Song, S. Havlin, and H. A. Makse, Nature 433, 392 (2005).
[28] S. Havlin, D. ben Avraham, et al., New Journal of Physics 9, 175 (2007).
[29] F. Altarelli, A. Braunstein, L. Doll’Asta, J. R. Wakeling, and R. Zecchina, Physical Review X 4, 021024 (2014).
[30] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, The European Physical Journal B-Conondensed Matter and Complex Systems 26, 521 (2002).
[31] J. C. Miller, A. C. Slim, and E. M. Volz, Journal of the Royal Society Interface 9, 890 (2012).
[32] M. Shrestha, S. V. Scarpino, and C. Moore, Physical Review E 92, 022821 (2015).
[33] W. Wang, M. Tang, H. Yang, Y. Do, Y.-C. Lai, and G. Lee, Scientific reports 4, 5097 (2014).
[34] X. Chu, J. Guan, Z. Zhang, and S. Zhou, Journal of Statistical Mechanics: Theory and Experiment 2009, P07043 (2009).
[35] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. 90, 028701 (2003).
[36] A. Saumell-Mendiola, M. Á. Serrano, and M. Boguñá, Physical Review E 86, 026106 (2012).
[37] Q.-H. Liu, W. Wang, M. Tang, and H.-F. Zhang, Scientific reports 6, 25617 (2016).
[38] M. A. Muñoz, R. Juhász, C. Castellano, and G. Ódor, Physical Review Letters 105, 128701 (2010).
[39] R. Durrett, Proceedings of the National Academy of Sciences 107, 4491 (2010).
[40] S. Chatterjee, R. Durrett, et al., The Annals of Probability 37, 2332 (2009).
[41] O. Givan, N. Schwartz, A. Cygeling, and L. Stone, Journal of theoretical biology 288, 21 (2011).
[42] C. Li, R. van de Bovenkamp, and P. V. Mieghem, Physical review e 86, 026116 (2012).
[43] M. Boguñá, C. Castellano, and R. Pastor-Satorras, Physical Review E 79, 036110 (2009).
[44] A. Arenas, J. Borge-Holthoefer, S. Meloni, Y. Moreno, et al., EPL (Europhysics Letters) 89, 38009 (2010).
[45] P. Van Mieghem (Springer, 2011), vol. 93, pp. 147–169.
[46] C. Li, H. Wang, and P. Van Mieghem, Physical Review E 88, 062802 (2013).
[47] F. Chung, L. Lu, and V. Vu, Proceedings of the National Academy of Sciences 100, 6313 (2003).
[48] F. Radicchi and C. Castellano, Physical Review E 93, 030302 (2016).
[49] B. Karrer and M. E. Newman, Physical Review E 82, 016101 (2010).
[50] T. Martin, X. Zhang, and M. Newman, Physical Review E 90, 052808 (2014).
[51] B. Karrer, M. E. Newman, and L. Zdeborová, Physical review letters 113, 208702 (2014).
[52] K. E. Hamilton and L. P. Pryadko, Physical review letters 113, 208701 (2014).
[53] F. Radicchi and C. Castellano, Nature communications 6 (2015).
[54] F. Radicchi, Nature Physics 11, 597 (2015).
[55] D. Cellai, S. N. Dorogovtsev, and G. Bianconi, arXiv preprint arXiv:1604.05175 (2016).
[56] S.-W. Son, G. Bizhani, C. Christensen, P. Grassberger, and M. Paczuski, EPL (Europhysics Letters) 97, 16006 (2012).
[57] B. Karrer and M. E. Newman, Physical Review E 82, 016101 (2010).
[58] W. Wang, Q.-H. Liu, L.-F. Zhong, M. Tang, H. Gao, and H. E. Stanley, Scientific reports 6 (2016).
[59] D. Rand, Advanced ecological theory: principles and applications 100 (1999).
[60] K. T. Eames and M. J. Keeling, Proceedings of the National Academy of Sciences 99, 13330 (2002).
[61] A. S. Mata and S. C. Ferreira, EPL (Europhysics Letters) 103, 48003 (2013).
[62] A. Szabo-Solticzky, L. Berthouze, I. Z. Kiss, and P. L. Simon, Journal of mathematical biology 72, 1153 (2016).
[63] T. Gross and H. Sayama, Adaptive networks (Springer, 2009).
[64] T. Gross, C. J. D. Lima, and B. Blasius, Physical review letters 96, 208701 (2006).
[65] G. A. Böhme and T. Gross, Physical Review E 83, 035101 (2011).
[66] I. Z. Kiss, G. Röst, and Z. Vizi, Physical review letters 115, 078701 (2015).
[67] P. Rattana, K. B. Blyuss, K. T. Eames, and I. Z. Kiss, Bulletin of mathematical biology 75, 466 (2013).
[68] W. O. Kermack and A. G. McKendrick, Proc. Roy. Soc. Lond. A 115, 700 (1927).
[69] N. T. Bailey et al., The mathematical theory of infectious diseases and its applications (Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE, 1975).
[181] J. P. Gleeson, D. Cellai, J.-P. Onnela, M. A. Porter, and F. R. Tsochas, Proc. Nat. Acad. Sci. **111**, 10411 (2014).

[182] W. Wang, Q.-H. Liu, S.-M. Cai, M. Tang, L. A. Braunstein, and H. E. Stanley, Scientific reports **6**, 29259 (2016).

[183] C. Granell, S. Gómez, and A. Arenas, Physical review letters **111**, 128701 (2013).