Epidemic fronts in complex networks with metapopulation structure

Jason Hindes,[1,*] Sarabjeet Singh,[2] Christopher R. Myers,[1,3] and David J. Schneider[4,5]

[1]*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York*
[2]*Theoretical and Applied Mechanics, Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York*
[3]*Institute of Biotechnology, Cornell University, Ithaca, New York*
[4]*Plant-Microbe Interactions Research Unit, Robert W. Holley Center for Agriculture and Health, Agricultural Research Service,
United States Department of Agriculture, Ithaca, New York*
[5]*Department of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, New York*

Infection dynamics have been studied extensively on complex networks, yielding insight into the effects of heterogeneity in contact patterns on disease spread. Somewhat separately, metapopulations have provided a paradigm for modeling systems with spatially extended and "patchy" organization. In this paper we expand on the use of multitype networks for combining these paradigms, such that simple contagion models can include complexity in the agent interactions and multiscale structure. Using a generalization of the Miller-Volz mean-field approximation for susceptible-infected-recovered (SIR) dynamics on multitype networks, we study the special case of epidemic fronts propagating on a one-dimensional lattice of interconnected networks—representing a simple chain of coupled population centers—as a necessary first step in understanding how macroscale disease spread depends on microscale topology. Applying the formalism of front propagation into unstable states, we derive the effective transport coefficients of the linear spreading: asymptotic speed, characteristic wavelength, and diffusion coefficient for the leading edge of the pulled fronts, and analyze their dependence on the underlying graph structure. We also derive the epidemic threshold for the system and study the front profile for various network configurations.

PACS number(s): 89.75.Fb, 64.60.aq, 87.19.X−, 87.23.Cc

I. INTRODUCTION

Network theory has proven a powerful framework for studying the effects of randomness and heterogeneity on the dynamics of interacting agents with nontrivial connectivity patterns [1]. One of the most important applications of this work is to the spread of infectious diseases among human populations, where the interaction structure is highly complex, showing salient features such as power-law degree distributions, small average path lengths, and modularity [2,3]. Various models have been proposed, primarily with random graph configuration, that incorporate these complex features while remaining theoretically tractable. Within the context of disease dynamics, graph nodes are generally taken to represent individuals, and edges to represent interactions between them, through which infection can spread. Both deterministic and stochastic infection dynamics have been studied on networks as well as bond percolation for the associated branching process [4–8]. How various thermodynamic quantities of interest, such as the total proportion infected, the epidemic (percolation) threshold, and the distribution of small outbreak sizes, depend upon network topology is of great interest.

Often these approaches disregard the multiscale organization of many real systems, in which agents can be most naturally thought of as partitioned into densely connected communities with sparser coupling among neighboring communities. In some cases, it may be useful to conceptualize the topology as a network of networks, where agent-to-agent interactions and community-to-community interactions are both useful representations depending on the scale of resolution [9]. The latter has been successfully developed

in ecology, with a network of interconnected populations referred to as a "metapopulation" [10,11]. This framework is very useful in studying large-scale propagation of diseases where most infection transmission occurs in localized regions, but can be transported on larger scales by the mobility of individuals, traveling among population centers [12]. However, most metapopulation models assume that populations are fully mixed, with no inherent complexity in the connectivity between agents. Much less understood is how the multiscale structure of agent interactions affects the larger-scale propagation of infectious processes through interconnected networks [13,14].

In this paper, we expand on a possible avenue for addressing this question using a multitype generalization of random graphs with simple, metalevel topology [9,15], and a dynamical mean-field theory for the SIR infection model in multitype configuration model networks. Putting these together, we analyze the average infection dynamics and propagating front profile on a simple metapopulation composed of coupled population centers on a one-dimensional lattice and calculate the phenomenological transport properties of the system as functions of the underlying network's degree distributions. Our results are compared to stochastic simulations of the infection kinetics on various networks and found to be in good agreement in the thermodynamic limit. Broadly, we present this work as an illustration of how well-developed ideas from different areas of statistical physics and ecology can be naturally combined.

II. MULTITYPE CONFIGURATION MODEL NETWORKS

In order to incorporate relevant node attribute information into our network models, (generically applicable for such things as age, sex, ethnicity, and place of residence), we

―――――――
*jmh486@cornell.edu

use a generalization of configuration model random graphs, wherein nodes are assigned a type from an arbitrary set of M possible types and a degree to each type from an arbitrary joint distribution for degree types, $P_i(k_1, k_2, \ldots, k_M) = P_i(\vec{k})$, with degree k_j denoting the number of connections to nodes of type j [1,9,15]. Additionally, nodes of type i occupy a fraction of the total network w_i, where $\sum_i w_i = 1$. Following the configuration model prescription, we consider graphs chosen uniformly at random from the ensemble of possible graphs with the prescribed degree distributions and self-consistent edge constraint: $w_i \sum_{\vec{k}} k_j P_i(\vec{k}) = w_j \sum_{\vec{k}'} k_i' P_j(\vec{k}'), \forall (i, j)$ [1,15,16].

From this formalism, a variety of quantities can be described compactly using generating functions [1,17]. The generating function for the probability of a randomly selected node of type i to have degree $\vec{k}$, is given by

$$G_i(\vec{x}) = \sum_{\vec{k}} P_i(\vec{k}) \prod_{l=1}^{M} x_l^{k_l} : \quad (1)$$

written as a power series in $\vec{x}$, an auxiliary variable defined over the unit interval, with expansion coefficients equal to the respective probabilities. Moments of the degree distributions can be represented simply as derivatives of the corresponding generating function. For example, the average degree of a type i node to a type j node is

$$\sum_{\vec{k}} k_j P_i(\vec{k}) = \partial_{x_j} G_i(\vec{x})|_{\vec{1}} \equiv \langle k_j \rangle_i. \quad (2)$$

Since node interactions occur along edges, an important quantity in network models is the excess degree: the number of neighbors a node has which can be reached by selecting a randomly chosen edge, and not including the neighbor on the end of the selected edge. For a multitype configuration model network, the probability that a randomly chosen edge from a type i node leads to a type j node with degree $\vec{k}$ is proportional to $k_i P_j(\vec{k})$, and thus the probability for the corresponding excess degree is generated by $\partial_{x_i} G_j(\vec{x})/\partial_{x_i} G_j(\vec{x})|_{\vec{1}}$ [15], with average degree to type l nodes,

$$\langle k_l \rangle_{i-j} = \frac{\partial_{x_l} \partial_{x_i} G_j(\vec{x})|_{\vec{1}}}{\partial_{x_i} G_j(\vec{x})|_{\vec{1}}} = \frac{\langle k_l k_i \rangle_j}{\langle k_i \rangle_j} - \delta_{il}. \quad (3)$$

By construction, this framework lacks two-point correlations, in which the excess degree distributions depend on the degrees of both nodes sharing an edge [3].

III. MILLER-VOLZ MEAN-FIELD SIR IN MULTITYPE NETWORKS

In this report we consider simple dynamics for disease spread: the susceptible-infected-recovered (SIR) model, wherein each individual is assigned a disease state, $Y \in \{S, I, R\}$, and may undergo reactions to other states depending on its state and the state of its neighbors. In this model, if a node of type i is susceptible and has a single infected neighbor of type j, then it will change its state to infected with a constant probability per unit time β_{ji}. Likewise, an infected node of type i will recover with a constant probability per unit time γ_i. Since the underlying dynamics is a continuous time

Markov process, a complete analysis would describe the full probability distribution for all system trajectories. However for our purposes, it will be sufficient to focus on the behavior of extensive outbreaks (i.e., those which scale with the system size), the average dynamics of which can be derived in the limit when the number of nodes tends to infinity, by applying a generalization of the mean-field technique developed by Miller and Volz [18]. Below, we follow the basic structure of the derivations presented in Refs. [20,21].

In the thermodynamic limit, configuration model random graphs are locally treelike [16], which by construction allows them to satisfy many of the generic criteria for the applicability of mean-field theory assumptions [22]. In our case, we assume that nodes are differentiated by their degree and disease state alone and that susceptible nodes feel a uniform force of infection along every edge, related to the average number of edges connecting susceptible and infected nodes at any given time in the network: a Curie-Weiss type approximation [23]. Furthermore, from the perspective of susceptible nodes, all infection attempts along different edges can be treated as uncorrelated—a consequence of the local treelike property [5,16,20]—and thus we assume that the states of neighbors of susceptible nodes are effectively independent.

Let the probability that a node of type j has not transmitted the infection to a node of type i along a randomly chosen i-j edge, be θ_{ij}. This quantity is interpretable as the complement of the average cumulative hazard function along such edges. Given θ_{ij}, it follows that the fraction of susceptible nodes of type i at time t is

$$\begin{aligned} S_i(t) &= \sum_{\vec{k}} P_i(\vec{k}) \prod_{j=1}^{M} \theta_{ij}^{k_j}(t) \\ &= G_i[\theta_{i1}(t), \theta_{i2}(t), \ldots, \theta_{iM}(t)] \\ &\equiv G_i[\vec{\theta}_i(t)]. \end{aligned} \quad (4)$$

The fractions of infected and recovered nodes of type i follow from probability conservation, $S_i + I_i + R_i = 1$, and a constant recovery rate for infected nodes γ_i

$$\frac{dI_i}{dt} = -\frac{d\vec{\theta}_i}{dt} \cdot \vec{\nabla} G_i(\vec{x})\bigg|_{\vec{\theta}_i} - \gamma_i I_i, \quad \frac{dR_i}{dt} = \gamma_i I_i, \quad (5)$$

with the total fraction of susceptible nodes

$$S = \sum_i w_i G_i(\vec{\theta}_i) \equiv \vec{w} \cdot \vec{G}(\underline{\theta}). \quad (6)$$

The central probability and order parameter, θ_{ij}, can be subdivided into three compartments depending on the disease state of the terminal node j,

$$\theta_{ij} = \theta_{ij}^S + \theta_{ij}^I + \theta_{ij}^R, \quad (7)$$

and its dynamics determined by tracking the fluxes among these compartments. Since θ can only change when an infected node transmits the disease, the rate at which θ_{ij} changes is equal to the rate at which a corresponding neighbor infects, and therefore $d\theta_{ij} = -\beta_{ji}\theta_{ij}^I dt$. Similarly, since θ^R can only change if an infected node recovers, the rate at which θ_{ij}^R changes is equal to the rate at which a corresponding neighbor recovers, and thus $d\theta_{ij}^R = \gamma_j \theta_{ij}^I dt$. Lastly, the probability that a type j neighbor of a type i node has not transmitted and

is susceptible, θ_{ij}^S, is simply the probability that the corresponding neighbor is susceptible. Because this neighbor could not have been infected along any of its other edges and has excess degree distribution generated by $\partial_{x_i} G_j(\vec{x})/\partial_{x_i} G_j(\vec{x})|_{\vec{1}}$, it follows that $\theta_{ij}^S = \partial_{x_i} G_j(\vec{x})|_{\vec{\theta}_j}/\partial_{x_i} G_j(\vec{x})|_{\vec{1}}$. Combining the latter with the two flux relations and the initial conditions (7), $\theta_{ij}(0) = 1$ and $\theta_{ij}^R(0) = 0$, we find

$$\frac{d\theta_{ij}}{dt} = \beta_{ji}\left(\frac{\partial_{x_i} G_j(\vec{x})|_{\vec{\theta}_j}}{\partial_{x_i} G_j(\vec{x})|_{\vec{1}}} - \theta_{ij}\right) + \gamma_j(1 - \theta_{ij}). \quad (8)$$

These M^2, first-order, and coupled ODEs, $\dot{\underline{\theta}} = \underline{F}(\underline{\theta})$, define the full system's approximate mean dynamics, and form the basis of our subsequent analysis. For a more detailed derivation of the analogous results for the special case of a single type network, see [18–21].

The steady state is given by the fixed point of (8),

$$\bar{\theta}_{ij} = (1 - T_{ji}) + T_{ji}\frac{\partial_{x_i} G_j(\vec{x})|_{\vec{\bar{\theta}}_j}}{\partial_{x_i} G_j(\vec{x})|_{\vec{1}}}, \quad (9)$$

which upon substitution into (4), gives the cumulative infection, $P = 1 - S$, at equilibrium (i.e., the final epidemic size), with $T_{ji} = \beta_{ji}/(\beta_{ji} + \gamma_j)$ the corresponding bond percolation probability, or transmissibility [15]. This can have a nontrivial solution corresponding to the existence of extensive outbreaks, if the disease-free state, $\theta_{ij} = 1 \ \forall(i,j)$, is unstable. The threshold or phase transition, which signifies the region in parameter space that separates the epidemic and nonepidemic phases, can be obtained through a stability analysis of the disease-free state, where the eigenvalue of the Jacobian for (8) with the largest real part, is real and vanishes when

$$\det(\underline{\underline{N}} - \underline{\underline{I}}) = 0, \quad \text{with} \quad N_{(i,j)(k,l)} = T_{ji}\delta_{jk}\langle k_l \rangle_{i-j} \quad (10)$$

an $M^2 \text{x} M^2$ matrix [24]. Similar results for the equilibrium properties are derivable from a multitype bond percolation approach [15].

IV. FRAMEWORK FOR MULTISCALE NETWORKS

Of interest to us are systems where type structure adds an additional scale of relevant topology, and not just demographic complexity [9,15]. For instance, we can apply the multitype network formalism to a simple model for a metapopulation by affiliating population centers with node types and coupling among populations with edges connecting their constituent nodes. In this way, a complex topology can be encoded on a microscale with a macroscale adjacency matrix, $\underline{\underline{A}}$, describing which populations are directly connected through node interactions [9]. We envisage example systems where $\underline{\underline{A}}$ describes the connectivity among urban centers, such as cities, towns, or villages, facilitated by roads or airlines. By conceptualizing the topology in this manner, we can study the phenomenology of infection propagation among population centers and describe how the propagation properties depend on the underlying connectivity patterns. A schematic is shown in Fig. 1(a) for a simple system with the pertinent structure. More broadly, we advance this approach as an avenue for combining the frameworks of network theory, metapopulations, and front propagation, which will be particularly useful if the interaction topology is coherent after some level of coarse graining.
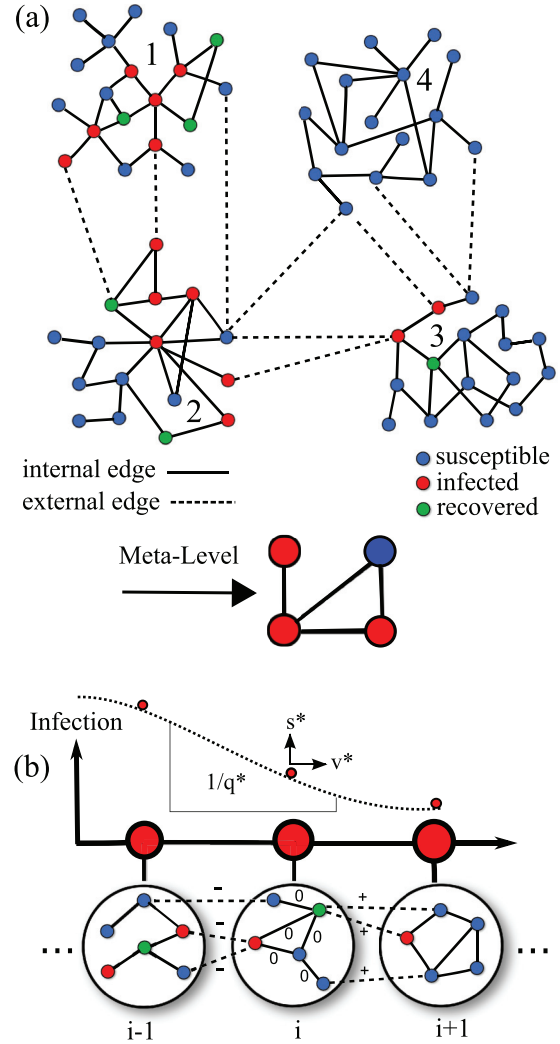

FIG. 1. (Color online) (a) A schematic of SIR dynamics on a metapopulation, where infection spreads along edges connecting nodes of various types at the finest scale (shown with integer, population labels), and the macroscale topology identifies which populations are connected through agent interactions. (b) An example of this framework, in which the macroscale topology takes the form of a one-dimensional (1D) lattice. In Sec. V, we focus on a simple case with configuration model construction, where each site has an identical degree distribution, specifying the probability of having a given number of internal (0), right (+), and left (−) external edges (shown above with labels for site i), and calculate the velocity, v^*, growth rate, s^*, and spatial decay rate, q^*, for infection fronts.

V. 1D LATTICE METAPOPULATION DYNAMICS

To illustrate this approach, we consider a special case of the above where the macroscale topology is an infinite one-dimensional (1D) lattice, $M \to \infty$, in which agents interact with other agents of the same type and agents of neighboring types, $A_{nj} = (\delta_{j,n} + \delta_{j,n+1} + \delta_{j,n-1})$ [12]. If infection is started at a single site (e.g., site 0) in a fully susceptible system, a strict directionality applies: in order for site i to be infected, sites $i - 1, i - 2, \ldots$ must be infected first. In such a case, we expect a well-defined infection front to propagate through the lattice. In keeping with the above, we focus

on an effective force of infection model among populations with static configuration networks having prescribed degree distributions, a generalization of the paradigmatic, spatial SIR model in one dimension, where the assumption of well mixed populations is relaxed to include complexity in agent interactions [11,25]. A schematic is shown in Fig. 1(b).

Since each node has three edge types, the mean equations of motion describe a three-component field, $\vec{\theta}_n(t) \equiv (\theta_{nn}, \theta_{nn+1}, \theta_{nn-1}) \equiv [\theta_n^0(t), \theta_n^+(t), \theta_n^-(t)]$, where (0), (+), and (−) denote internal, right-external, and left-external edges, at the corresponding site. For simplicity, homogeneity is assumed, with β, w, γ, and G all uniform, reducing the field equations to

$$\frac{d\theta_n^0}{d\tau} = (1 - \theta_n^0) + T\left(\frac{G_0(\vec{\theta}_n)}{G_0(\vec{1})} - 1\right)$$
$$\frac{d\theta_n^\pm}{d\tau} = (1 - \theta_n^\pm) + T\left(\frac{G_\mp(\vec{\theta}_{n\pm1})}{G_\mp(\vec{1})} - 1\right), \quad (11)$$

where the time, τ, is measured in units of $1/(\beta + \gamma)$, and the subscript in G denotes a partial derivative with respect to the corresponding variable. For edge number consistency, $G_+(\vec{1}) = G_-(\vec{1})$, but in general we allow for other asymmetries in the degree distributions.

A. Dispersion relation and transport coefficients

To understand the spatiotemporal dynamics (11), we first quantify how perturbations away from the unstable state propagate by linearizing the dynamics around the disease-free equilibrium, $\vec{\theta}_n(t) = \vec{1} - \vec{\epsilon}_n(t)$, and decoupling the perturbations into basis modes using the inverse discrete Fourier transform, $\vec{\epsilon}_n(t) = \frac{1}{M}\sum_{v=0}^{M-1}\vec{\epsilon}_v(t)e^{i(2\pi vn/M)}$, (IDFT). The dispersion relation can be found by substituting the IDFT into (11) and using the basis properties of orthogonality and completeness. In the limit $M \to \infty$ the site perturbations approach the integral, $\vec{\epsilon}_n(t) = \frac{1}{2\pi}\int_0^{2\pi}\vec{\epsilon}(k)e^{i[kn-\omega(k)t]}dk$.

With this prescription, we find the dispersion relation takes the form of a cubic, characteristic equation

$$\det\left(\underline{\underline{K_e}}(q) - \frac{1+s(q)}{T}\underline{\underline{I}}\right) = 0, \quad \text{with}$$

$$\underline{\underline{K_e}}(q) = \begin{bmatrix} \frac{\langle k_0^2\rangle}{\langle k_0\rangle} - 1 & \frac{\langle k_0 k_+\rangle}{\langle k_0\rangle} & \frac{\langle k_0 k_-\rangle}{\langle k_0\rangle} \\ \frac{\langle k_- k_0\rangle}{\langle k_-\rangle}e^{-q} & \frac{\langle k_- k_+\rangle}{\langle k_-\rangle}e^{-q} & \left(\frac{\langle k_-^2\rangle}{\langle k_-\rangle}-1\right)e^{-q} \\ \frac{\langle k_+ k_0\rangle}{\langle k_+\rangle}e^{q} & \left(\frac{\langle k_+^2\rangle}{\langle k_+\rangle}-1\right)e^{q} & \frac{\langle k_+ k_-\rangle}{\langle k_+\rangle}e^{q} \end{bmatrix}, \quad (12)$$

which for convenience, is written in terms of s and q, where $\omega = is$ and $k = iq$ [Fig. 1(b)]. Interestingly, this method reveals a generalization of the average excess degree matrix, $\underline{\underline{K_e}}(0)$, whose elements are found by selecting a randomly chosen edge in a particular direction, and counting the average number of reachable neighbors of a particular type, for the interconnected network system, $\underline{\underline{K_e}}(q)$, which incorporates the relative states of adjacent sites on the lattice for each mode q. We expect this operator to emerge in similar problems on interconnected networks.

Combining the above with the behavior of infection near the phase transition, where there is no exponential growth in time and each site has the same field value, $\vec{\theta}(n) = \vec{1}$: $s \to 0$ and $q \to 0$ (12), we find a simple condition for the critical transmissibility T_c

$$T_c = \frac{1}{\lambda_m^k(0)}, \quad (13)$$

where $\lambda_m^k(q)$ is the maximum eigenvalue of $\underline{\underline{K_e}}(q)$, with $\lambda_m^k(0)$ corresponding to $\underline{\underline{K_e}}(0)$. Because the addition of external edges increases the spreading capacity of the disease, the critical transmissibility in the coupled system is less than the uncoupled case, implying that transport-mediated infections from neighboring sites can sustain epidemics even when individual populations on their own cannot [13,14].

Also from the dispersion relation, we can find the asymptotic transport coefficients for rightward moving disturbances by making a standard saddle-point approximation of the perturbations' integral representation in Fourier space: expanding the integrand around its dominant contribution, k^*, in the comoving frame, $\xi = n - v^*t$,

$$e^{i[kn-\omega(k)t]} \sim e^{ik\xi}e^{it(kv^*-\omega(k^*)-\frac{d\omega}{dk}|_{k^*}(k-k^*))}e^{\frac{-it(k-k^*)^2}{2}\frac{d^2\omega}{dk^2}|_{k^*}}$$

and taking the infinite time limit while enforcing approximate constancy with no exponential growth and ξ finite, where v^* is the asymptotic speed at which perturbations to the unstable state propagate [26]. This procedure uncovers an exponential moving pulse for the leading edge of the infection profile with a diffusive correction [26]

$$1 - \theta \sim \frac{e^{-q^*\xi}e^{-\xi^2/4D^*t}}{\sqrt{D^*t}}, \quad (14)$$

where q^*, v^*, and D^* satisfy the saddle-point relations

$$v^* = \frac{ds}{dq}\bigg|_{q^*} = \frac{s(q^*)}{q^*} = T\frac{d\lambda_m^k}{dq}\bigg|_{q^*} = \frac{-1 + T\lambda_m^k(q^*)}{q^*} \quad (15)$$

$$\text{and} \quad D^* = \frac{1}{2}\frac{d^2s}{dq^2}\bigg|_{q^*} = \frac{T}{2}\frac{d^2\lambda_m^k}{dq^2}\bigg|_{q^*}, \quad (16)$$

giving a transcendental equation for q^*, which corresponds to the selection of the minimum velocity, v^*, from the front's spectrum of exponential modes, e^{st-qn}, with velocities, $v(q) = s(q)/q$ [Fig. 1(b)] [25,26]. If multiple solutions exist for the minimum, the fastest solution is selected [26].

When the average excess degree matrix is irreducible (the domain of interest to us), the dominant growth exponent for each q is real and corresponds to a positive and unique eigenvector (12) [24], and thus we expect the same selected velocity for all field components, $\vec{\theta}(n)$ [27]. Furthermore, the fields propagate in this regime with approximate proportions $\vec{1} - \epsilon(t)\vec{Q}(q^*, s^*)$, where $\vec{Q}(q^*, s^*)$ is the corresponding mode of $\underline{\underline{K_e}}(q^*)$ (12). The characteristic wavelength, $1/q^*$, is related to the asymptotic size of the front's leading edge, and diverges near the phase transition. The diffusion coefficient, D^*, gives the effective widening of the mean-field pulse in the comoving frame and is proportional to the largest finite-size correction to v^* in the limit where the number of nodes at each site tends to infinity.

In order to uncover the principal dependencies of the transport coefficients, we study (12)–(16) near the phase transition, where the power series expansion for the dispersion relation is a convenient representation; the latter is found by substituting $[s(q) + 1]/T = a + bq + \frac{c}{2}q^2 + \cdots$ into (12), and equating powers in q. When s^* and q^* are small in the vicinity of T_c, the expansion can be truncated at low order, giving a Fisher-Kolmogorov-like dispersion relation with the approximate scaling

$$s^* \sim \left[T\lambda_m^k(0) - 1 \right] \quad v^* \sim \left[T\lambda_m^k(0) - 1 \right]^{\frac{1}{2}} D^{*\frac{1}{2}}$$

$$q^* \sim \left[T\lambda_m^k(0) - 1 \right]^{\frac{1}{2}} D^{*-\frac{1}{2}} \quad \frac{D^*}{T\lambda_m^k(0)} \sim \delta \qquad (17)$$

$$\delta = \frac{\frac{\langle k_- k_0 \rangle \langle k_0 k_+ \rangle}{\langle k_- \rangle \langle k_0 \rangle} + \frac{\langle k_- k_+ \rangle}{\langle k_- \rangle} \left(\lambda_m^k(0) - \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle} + 1 \right)}{\left[\lambda_m^k(0) - \lambda_2^k(0) \right]\left[\lambda_m^k(0) - \lambda_3^k(0) \right]},$$

where $\lambda_2^k(0)$ and $\lambda_3^k(0)$ are the subdominant eigenvalues of $\underline{\underline{K}}_e(0)$. In this regime, we find an effective reaction-diffusion behavior with the generic dependence of the shape and speed of the propagating front's leading edge on the reproductive number $T\lambda_m^k(0)$ (a product of the spreading capacity along edges and the magnitude of topological fluctuations) and on the normalized diffusion coefficient δ: measuring the relative strength of connection between lattice sites (17). We see that the effective reaction rate is equal to the distance from the phase transition, $T\lambda_m^k(0) - 1$, and that all coefficients grow from zero with this distance, except for D^*, which varies discontinuously through T_c. Furthermore, the normalized diffusion coefficient increases from zero with $\langle k_- k_0 \rangle \langle k_0 k_+ \rangle$ and $\langle k_- k_+ \rangle$ – the correlation moments of the degree distribution which encode the propensity for transport from the $i \mp 1$ site to the $i \pm 1$ site (both of which cannot be zero, otherwise epidemics are locally confined), and with the viability of subdominant modes to support growth. In general, we find that as δ increases v^* and D^* increase, q^* decreases, and s^* remains constant, implying faster transport and greater similarity among sites, as more edge-type pairs allow for traversing the lattice, but with little change in the growth exponent.

The above demonstrates the typical trend for these models, that the front dynamics is strongly influenced by the joint degree distribution's second moments (i.e., the relevant excess degree properties are generally amplified by correlation among degree types and degree heterogeneity). For example, in analogy with the single network case, fast transport can be achieved with the presence of a few nodes with large internal and external degrees, or "transport hubs", even if the average degrees in the network are small [2,3].

B. Simple mixing example

Additional understanding of the basic form of the transport coefficients is gained by looking at a special case of the microscale degree distribution, where the generating function takes the form $G[px_0 + \frac{1-p}{2}(x_+ + x_-)]$, with total degree described by G, and a given edge connecting nodes of the same site with probability p, and nodes of left and right neighboring sites with equal probability $(1 - p)/2$, where $1 - p$ is an effective mixing parameter among populations. With this prescription, the critical transmissibility is reduced

to the inverse of the total-edge excess degree, $T_c = \frac{G'(1)}{G''(1)}$, and the normalized diffusion coefficient, to the fraction of external edges in a each direction, $\delta = (1 - p)/2$. Moreover, the dispersion relation takes the instructive form

$$s(q) = -1 + \frac{TG''(1)}{G'(1)}[p + (1 - p)\cosh(q)], \qquad (18)$$

where $s + 1$ is given by the basic reproductive number multiplied by the average relative incidence, $e^{-\Delta x q}$, on the end of a randomly selected edge, illustrating the intuitive generalization of the single network case, where different edge types are more and less likely to connect to infected nodes depending on their place in the lattice, and thus to contribute to local growth.

Likewise, from (15) and (16), we find the speed and diffusion coefficient,

$$v^* = \frac{TG''(1)}{G'(1)}(1 - p)\sinh(q^*) \qquad (19)$$

and

$$D^* = \frac{TG''(1)}{2G'(1)}(1 - p)\cosh(q^*), \qquad (20)$$

where q^* satisfies (15), and v^* is given by the basic reproductive number multiplied by the average product of relative position and incidence, $-\Delta x e^{-\Delta x q^*}$, on the end of a randomly selected edge. Figure 2 shows the transport coefficients (18)–(20), as functions of T/T_c and p, with partial scaling collapse (17) for the corresponding class of network configurations. The expected reaction-diffusion scaling can be observed near the critical point, and far away from the critical region, when $T \gg T_c$, q^*, $\frac{s^* G'(1)}{TG''(1)}$, $\frac{v^* G'(1)}{TG''(1)}$, and $\frac{D^* G'(1)}{TG''(1)}$ tend to limiting
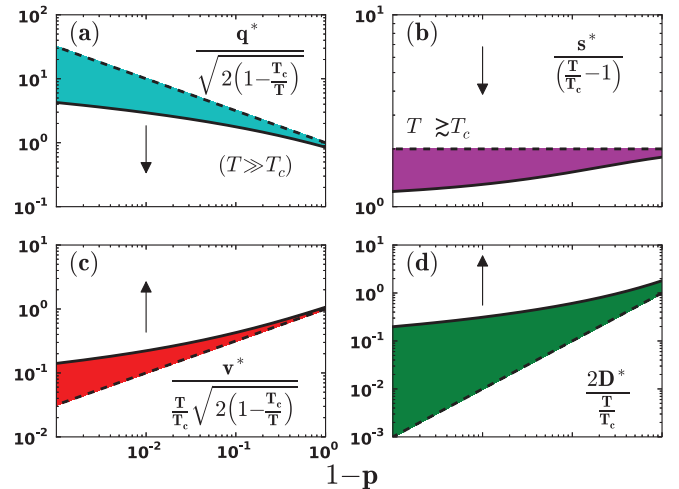

FIG. 2. (Color online) The transport coefficients for a one-dimensional lattice of configuration model networks with arbitrary total-degree distribution and inter-population mixing parameter $1 - p$, shown as functions of the latter (Sec. V B): q^* (a), s^* (b), v^* (c), and D^* (d). The colored regions mark the range of each coefficient, which are bounded by the critical-region scaling ($T \gtrsim T_c$), and the limiting behavior ($T \gg T_c$), delineated by dashed and solid curves respectively; the former are straight lines, signifying agreement with the predicted scaling (17). Each panel's arrow indicates the direction of increase in the distance from the phase transition, $T/T_c - 1$.

curves which depend only on $1 - p$, suggesting the intuitive asymptotic proportionality to the reproductive number.

C. Pulled front classification

In order to connect the transport properties of the linear equations to the full nonlinear system, we refer to the classification of fronts propagating into unstable states, which in our system is the fully susceptible metapopulation lying ahead of the infection front. In general, there are two types of deterministic fronts: pulled and pushed, with the former having fronts with asymptotic speed equal to the linear spreading speed and the latter having fronts with asymptotic speed greater than the linear spreading speed [26]. Pushed fronts occur because nonlinearities in the equations of motion tend to increase the growth of perturbations on the unstable state, resulting in nontrivial front shape dependence of the speed. However, in our system all nonlinearities are proportional to probability generating functions, $\sim G'(\theta)/G'(1)$ (11), which are monotonically increasing over the unit interval. Therefore, all nonlinearities tend to increase θ, and consequently dampen the growth of infection, a sufficient condition for pulled fronts [28], and thus we anticipate fronts in this model to be pulled; this agrees with the intuition that epidemic propagation is governed by its behavior in a fully susceptible population. In practice, the classification has importance for control strategies in systems with similar structure, implying that to mitigate the spread of infection among populations, efforts should be focused on the leading edge of the front, and not on larger outbreaks occurring farther behind.

D. Relaxation properties

In addition to quantifying the transport, the front speed can be used to extract information about the dynamics away from the unstable state. As shown above, the $\vec{\theta}$ field settles onto a solution with translational similarity, $\vec{\theta}_{n \pm x}(t) = \vec{\theta}_n(t \mp \frac{x}{v^*})$, after an initial transient period. Behind the leading edge of the front, the behavior resembles a relaxation to the stable equilibrium (9), $\vec{\theta}_n(t) \approx \vec{\bar{\theta}} + \vec{\eta}(t - \frac{n}{v^*}) \approx \vec{\bar{\theta}} + \vec{\eta}e^{-z^*(n-v^*t)}$, where the spatial rate, $|z^*|$, is the dominant eigenvalue of the nonlinear eigenvalue equation

$$\det\left(\underline{\underline{G'_e}}(\vec{\bar{\theta}}, z^*) - \frac{1 + v^* z^*}{T}\underline{\underline{I}}\right) = 0, \quad \text{with}$$

$$\underline{\underline{G'_e}}(\vec{\bar{\theta}}, z) = \begin{bmatrix} \frac{G_{00}}{G_0(\bar{1})} & \frac{G_{0+}}{G_0(\bar{1})} & \frac{G_{0-}}{G_0(\bar{1})} \\ \frac{G_{-0}}{G_-(\bar{1})}e^{-z} & \frac{G_{-+}}{G_-(\bar{1})}e^{-z} & \frac{G_{--}}{G_-(\bar{1})}e^{-z} \\ \frac{G_{+0}}{G_+(\bar{1})}e^{z} & \frac{G_{++}}{G_+(\bar{1})}e^{z} & \frac{G_{+-}}{G_+(\bar{1})}e^{z} \end{bmatrix}\Bigg|_{\vec{\bar{\theta}}}. \quad (21)$$

The latter is the analog of $\underline{\underline{K_e}}(q)$ at the stable state, which does not depend on the first two moments of the degree distribution directly, but on the generating function's properties near the equilibrium (9). In general, the two characteristic spatial rates for this system are not equal, $|z^*| \neq q^*$, and when their difference is large, it often signifies a significant separation in the time scales of growth, $1/s^*$, and relaxation $1/v^*|z^*|$. The latter provides an estimate for the amount of time a site is

infectious, with $1/|z^*|$ yielding a related estimate for the width of the propagating front (i.e., the typical spatial extent of an outbreak at a given time). In particular, when the front speed is very fast and the degree distribution's second moments are large with the first moments $\mathcal{O}(1)$, we find that $|z^*| \ll q^*$, which suggests broad front profiles. In this case, propagation and relaxation can be thought of as approximately distinct processes.

VI. COMPARISON WITH STOCHASTIC SIMULATIONS

The above predictions for the mean-field dynamics on the one-dimensional metapopulation were compared to stochastic simulations of SIR dynamics on random instances of multiscale, metapopulation networks, using Gillespie's direct method [11,29,30]. The graphs were constructed using the multitype configuration model by first generating a degree sequence from the desired degree distribution and then connecting pairs of edge "stubs", selected uniformly at random [2,15,16]. An outbreak was started by choosing one node from the first lattice site to be infected with all others susceptible. Only outbreaks which lead to epidemics with $\mathcal{O}(N)$ cumulative infection were considered for comparison with mean-field predictions. In order to ignore fluctuations in the initial transients, time was zeroed after the first 100 reactions.

We are interested in the average shape of the front that connects the fully susceptible unstable state lying ahead of the infectious wave and the fully recovered (equilibrium) state lying behind it. The average shape was computed by taking instantaneous "snapshots" of the profile for each stochastic realization, conditioned on the middle lattice site having cumulative infection equal to half the equilibrium value (9), and averaging the cumulative infection of the other sites over different realizations. In general, the snapshots did not occur at the same instant; however, aligning fronts based on a common level of infection eliminates some of the effects of diffusive wandering [Fig. 3(b)]. The measured fronts were compared to the mean-field profiles by integrating the lattice equations (11). A comparison is shown in Fig. 3 for two graphs with scale-free and Poisson degree distributions, with generating functions

$$G_{\text{S.F.}}(\vec{x}) = \text{Li}_\alpha[e^{-1/K}x_0(1 - \nu + \nu x_+)(1 - \nu + \nu x_-)]$$

and

$$G_P(\vec{x}) = \exp\{C[x_0(1 - \nu + \nu x_+)(1 - \nu + \nu x_-) - 1]\},$$

where Li_α is the polylogarithm function with exponent α [2]. The parameters for the degree distributions were chosen such that each network had the same average degree and cloning parameter, ν (i.e., given a specified internal degree distribution, each of a node's internal edges is copied to form an external edge with probability ν), but with different inherent levels of heterogeneity.

We see in Fig. 3 that the epidemic front is broader in the scale-free network than in the Poisson. This difference comes from the much larger front speed of the former, which had average excess degrees an order of magnitude larger than the latter, (12) and (15), and the relatively similar relaxation times (21) for the two classes of networks (implying that the time scale over which a site is infectious in each network
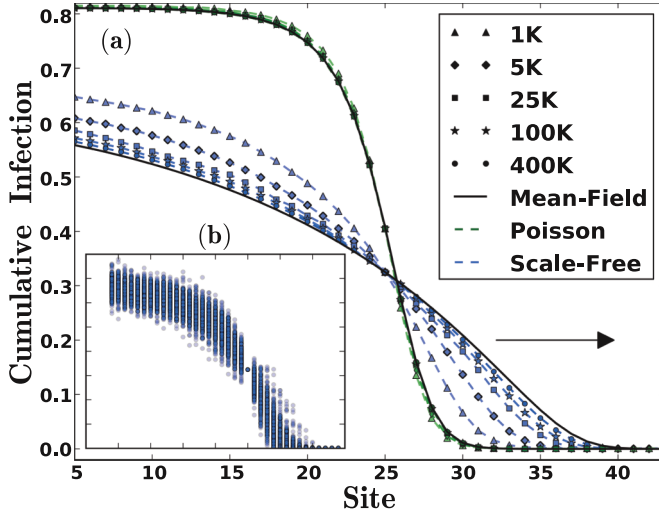
FIG. 3. (Color online) (a) A comparison between the average cumulative infection profile for stochastic simulations of SIR dynamics on a one-dimensional lattice of scale-free (blue/lower) and Poisson (green/upper) networks and mean-field predictions. Various site sizes are shown with different symbols and color shades, varying from light to dark for 10^3 to 4×10^5 respectively. Front shapes for increasingly large sizes are found to converge to the mean-field front. The parameters for the graphs were chosen to be: $K = 100$, $\alpha = 2$ and $\nu = 0.3$ for the scale-free, and $C = 2.901\,57$ and $\nu = 0.3$ for the Poisson (Sec. VI). A lattice of 50 sites was used, which was large enough to ensure uniformity with the above parameters and reaction rates $\beta = \gamma = 1$. The arrow indicates the propagation direction. (b) An ensemble of stochastic fronts conditioned on the middle lattice site having cumulative infection equal to half the equilibrium value (9). Because the infection fronts for different stochastic realizations are offset from one another in time, we align all fronts based on a common level of infection, and then average over realizations (Sec. VI).

is roughly the same). In the more homogeneous Poisson networks, the front is more narrow and propagates through the lattice on the same time scales as the local infection dynamics; whereas in the scale-free case, the leading edge of the front propagates quickly through the lattice, followed by a slower relaxation to the stable equilibrium state behind the front. This comparison shows that assumptions of homogeneity can drastically underestimate the speed and extent of fronts in systems with heterogenous interactions.

Additionally, the front speed v was numerically estimated from the average time $\langle \tau_{\text{prog}} \rangle$ required for the leading edge of the front to move forward by one lattice site (where the leading edge was defined as that site where the incidence first reached a set $\mathcal{O}(1)$ level) and averaging over such levels; i.e., $1/v = \langle \tau_{\text{prog}} \rangle$, once the initial spatial variation had decayed. Figure 4 shows the convergence of the measured speed from simulations to the mean-field prediction for each graph as a function of the steady state, cumulative infected population size at every lattice site, $\bar{N} = \bar{P} N$ (Sec. III), with total size N. The lines represent fits to the expected scaling of the largest finite-size correction for pulled fronts, $v^* - v_{\bar{N}} \sim \frac{D^* q^* \pi^2}{\ln^2(\bar{N})}$, obtained from a general $1/\bar{N}$ cutoff in the mean-field equations [26,28,31]; the coefficients are found to be $\mathcal{O}(1)$ of the expected scaling. In general, higher-order corrections in $\bar{N}$ must be calculated from
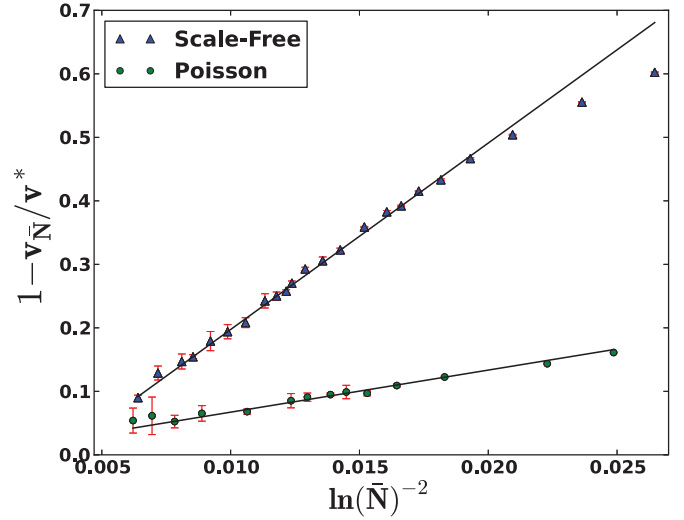

FIG. 4. (Color online) (a) Convergence of the average velocities, $v_{\bar{N}}$, to the mean-field predictions, v^*, for scale-free and Poisson networks (Fig. 3) as functions of the cumulative number of infected nodes at each site, shown with fits to the expected pulled front scaling, $v^* - v_{\bar{N}} \sim \frac{D^* q^* \pi^2}{\ln^2(\bar{N})}$ (Sec. VI).

an analysis of the full, stochastic system [28]. The very slow convergence in $\bar{N}$ comes from the transport dependence on the linearized equations where infinitesimal infection levels apply and sensitivity to stochastic effects is highest. This can be seen in the fairly large finite-size corrections to the velocity, particularly for the scale-free network, leading to a more narrow conditionally averaged front relative to the mean-field, with fewer sites initiated at a given time (Fig. 3).

Finally, the average epidemic profile and transmissibility threshold, (9) and (13), were compared to simulations. Figure 5 plots those comparisons for a system with left-right
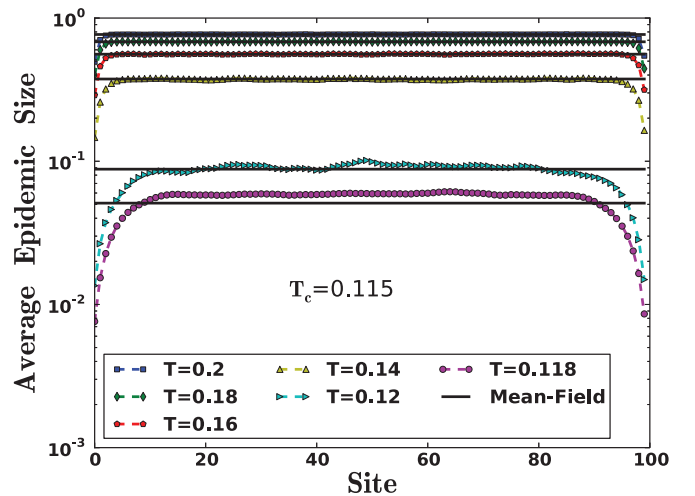

FIG. 5. (Color online) The average epidemic fraction at each site for a network with asymmetric generating function (Sec. VI) and varying transmissibilities. The point where the epidemic vanishes agrees with the prediction, $T_c = 0.115$ (13). Sites were occupied by 20 000 nodes on a lattice of 100 sites.

asymmetric generating function

$$G_{\text{Asym}}(\vec{x}) = \tfrac{1}{3}\big(2x_0^2 x_+^2 x_-^4 + x_0^4 x_+^6 x_-^2\big),$$

on a lattice of 100 sites, with 20 000 nodes on each site. Both the epidemic size for various transmissibilities and the threshold were found to be in good agreement with mean-field predictions, though the finite-size effects became increasingly important as the critical region was approached, leading to significantly smaller outbreaks near the edges of the lattice.

VII. CONCLUSION

In this paper we developed a framework for studying infection dynamics in multiscale metapopulations, combining multitype networks and a mean-field theory, to explore how macroscale disease propagation depends on microscale interaction structure. As a necessary first step in this direction, we applied the approach to a simple metapopulation model for a chain of coupled populations, and derived the transport properties for infection, including their scaling with the disease transmissibility and the statistical properties of the underlying network. We also found a threshold for the viability of epidemics, and calculated the relaxation properties of the propagating front. These were compared for different network models, with heterogeneous networks having considerably higher speeds and broader fronts than their homogeneous counterparts,

illustrating the importance of including complexity in the fine-scale topology in order to accurately capture transport phenomenology.

Various extensions of the work presented, both in terms of analyses carried out and systems studied, could be considered. We have addressed here only the average dynamics of the one-dimensional, homogeneous system, without any description of finite-size fluctuations, or consideration of the dynamics in higher-dimensional generalizations. Greater complexity could be introduced through the spatiotemporal dependence of network parameters, and/or more general network configurations [32]. An interesting extension of the model discussed here would include dynamic contacts between nodes and explicit mobility, instead of the assumed time-scale separation between topology and the overlying process [12,33,34]. However, the basic formalism presented here can enable one to study such factors and build more realistic models for infectious processes in multiscale problems.

[1] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[2] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[3] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).

[4] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).

[5] E. Kenah and J. M. Robins, Phys. Rev. E **76**, 036113 (2007).

[6] P.-A. Noël, B. Davoudi, R. C. Brunham, L. J. Dubé, and B. Pourbohloul, Phys. Rev. E **79**, 026101 (2009).

[7] P.-A. Noël, A. Allard, L. Hébert-Dufresne, V. Marceau, and L. J. Dubé, Phys. Rev. E **85**, 031118 (2012).

[8] E. Volz, J. Math. Biol. **56**, 293 (2008).

[9] A. Vazquez, Phys. Rev. E **74**, 066114 (2006).

[10] I. Hanski, Nature (London) **396**, 41 (1998).

[11] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, 2007).

[12] V. Belik, T. Geisel, and D. Brockmann, Phys. Rev. X **1**, 011001 (2011).

[13] M. Dickison, S. Havlin, and H. E. Stanley, Phys. Rev. E **85**, 066109 (2012).

[14] A. Saumell-Mendiola, M. A. Serrano, and M. Boguñá, Phys. Rev. E **86**, 026106 (2012).

[15] A. Allard, P.-A. Noël, L. J. Dubé, and B. Pourbohloul, Phys. Rev. E **79**, 036113 (2009).

[16] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).

[17] H. S. Wilf, *Generatingfunctionology*, 2nd ed. (Academic Press, London, 1994).

[18] We recently became aware of a similar treatment of SIR dynamics in multitype networks [19].

[19] J. C. Miller and E. M. Volz, arXiv:1106.6344 [q-bio.PE].

[20] J. C. Miller, J. Math. Biol. **62**, 349 (2011).

[21] J. C. Miller, A. C Slim, and E. M. Volz, J. R. Soc. Interface **9**, 890 (2011).

[22] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, Phys. Rev. E **85**, 026106 (2012).

[23] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Westview Press, Boulder, 1992).

[24] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, Philadelphia, 2001).

[25] I. Sazonov, M. Kelbert, and M. B. Gravenor, Math. Model. Nat. Phenom. **3**, 28 (2008).

[26] W. van Saarloos, Phys. Rep. **386**, 29 (2003).

[27] S. Theodorakis and E. Leontidis, Phys. Rev. E **65**, 026122 (2002).

[28] D. Panja, Phys. Rep. **393**, 87 (2004).

[29] D. T. Gillespie and L. R. Petzold, J. Chem. Phys. **119**, 8229 (2003).

[30] M. A. Gibson and J. Bruck, J. Phys. Chem. A **104**, 1876 (2000).

[31] E. Brunet and B. Derrida, Phys. Rev. E **56**, 2597 (1997).

[32] B. Karrer and M. E. J. Newman, Phys. Rev. E **82**, 066118 (2010).

[33] F. Simini, M. C. González, A. Maritan, and A. L. Barabási, Nature (London) **484**, 96 (2012).

[34] N. Perra, A. Baronchelli, D. Mocanu, B. Goncalves, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **109**, 238701 (2012).