

Direct computation of contagion triggering probabilities for generalized and bipartite random networks

Kameron Decker Harris,^{1,*} Joshua L. Payne,^{2,†} and Peter Sheridan Dodds^{3,4,‡}

¹ *Applied Mathematics, University of Washington,
Lewis Hall #202, Box 353925, Seattle, WA 98195-3925.*

² *Institute of Evolutionary Biology and Environmental Sciences,
University of Zurich, Winterthurerstrasse 190, 8057, Zurich, Switzerland.*

³ *The University of Vermont, Burlington, VT 05401.*

⁴ *Complex Systems Center, Computational Story Lab,
the Vermont Advanced Computing Core, & Department of Mathematics & Statistics,
The University of Vermont, Burlington, VT 05401.*

(Dated: July 1, 2015)

We derive a general expression for the probability of global spreading starting from a single infected seed for contagion processes acting on generalized, correlated random networks. We employ a simple probabilistic argument that encodes the spreading mechanism in an intuitive, physical fashion. We use our approach to directly and systematically obtain triggering probabilities for contagion processes acting on a collection of random network families including bipartite random networks. We find the contagion condition, the location of the phase transition into an endemic state, from an expansion about the disease-free state.

PACS numbers: 89.75.Hc,64.60.aq,64.60.Bd,87.23.Ge

I. INTRODUCTION

Spreading is a pervasive dynamic phenomenon, ranging in form from simple physical diffusion to the complexities of socio-cultural dispersion and interaction of ideas and beliefs [1–11]. Successful spreading in systems may manifest as an expanding front, such as in the spread of disease through medieval Europe [12], or through inherent or revealed networks, such as in pandemics in the modern era of global travel [13]. Here, we focus on spreading processes operating on generalized random networks, which have proven over the last decade to be illustrative of spreading on real networks and at the same time to be analytically tractable [3, 14–24].

In contributing to the wealth of already known results for contagion on random networks, we make two main advances here. First, we obtain, in the most general terms possible, an expression for the probability of global spreading from a single seed for a broad range of contagion processes acting on generalized, correlated random networks. By global spreading we mean a non-zero fraction of nodes in an infinite network are eventually infected. Second, we use an argument that is physically motivated and direct. Existing approaches rely on a range of mathematical techniques, such as probability generating functions [14, 25, 26], which, while being entirely successful in determining spreading probabilities and higher moments of cascade size distribution, obscure the underlying physical mechanisms.

The present paper is a companion to our earlier work where we derived a general condition for the possibility (rather than probability) of global spreading for single-seed contagion processes acting on random networks [27]. We used specific results from both works in a separate investigation of exactly solvable network spreading models [28]. As we show below, our expression for the probability of spreading naturally allows us to recover our expression for the possibility of spreading, and this is a purely mathematical exercise. Our key contribution is the direct derivation of triggering probabilities via physical arguments, as illustrated in Fig. 1.

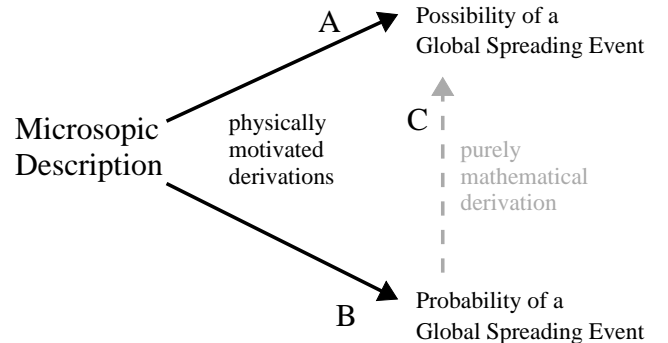


FIG. 1: Physical and mathematical explanations of two fundamental aspects of broad classes of contagion processes acting on generalized random networks. In the present paper, we provide a physical approach to determining the probability of spreading from a single seed (derivation B). We use mathematical arguments to arrive again at the binary contagion condition (derivation C), which we obtained in a previous work [27] using a direct physical explanation (derivation A).

*Electronic address: kamdh@uw.edu

†Electronic address: joshua.payne@ieu.uzh.ch

‡Electronic address: peter.dodds@uvm.edu

the broadest class of correlated random networks allowing for directed and undirected edges and arbitrary node and edge properties. In Sec. III, we define the general class of contagion processes that our treatment can encompass. In Sec. IV, we compute the probability that seeding a node of a given type generates a global spreading event. For completeness, in Sec. V, we derive the contagion condition (location of the endemic phase transition) result found in [27], and we show how non-physical expressions may arise through this mathematical route. We use our formalism for six interrelated random network families with general contagion processes acting on them in Sec. VIA. In Secs. VIB and VIC, we show how our approach readily applies to random bipartite networks, and we offer some concluding remarks in Sec. VII.

II. GENERALIZED RANDOM NETWORKS

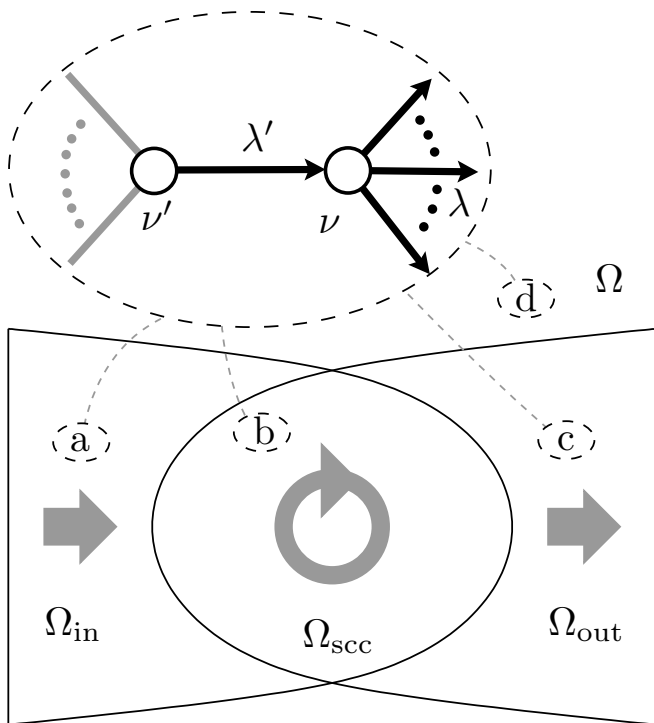


FIG. 2: Schematic showing the configuration of the potential triggering node subnetwork using the present work’s formalism for generalized random networks described in Sec. II, and the basic form of a random network with directed edges and a giant component. The ellipses labelled a–d show four possible locations of the subnetwork in the overall network Ω . Global spreading events can be successfully generated only if the subnetwork is part of the giant in-component Ω_{in} , either within or outside of the giant strongly connected component Ω_{scc} (ellipses a and b). No spreading is possible if the subnetwork is instead part of the giant out-component outside of the strongly connected component ($\Omega_{\text{out}}/\Omega_{\text{scc}}$, ellipse c) or outside of all three giant components (ellipse d).

Our theoretical treatment builds on a formalism we introduce here for representing generalized random networks, an expansion of what we used in our connected, earlier work [27]. Our theory applies to large random networks with bounded degrees (such as the configuration model), since these graphs are all locally tree-like and can be approximated by multitype branching processes. Generalized random networks may contain a combination of directed and undirected edges, so they are in general nonsimple graphs.

We depict the essential features of a random network with possibly directed edges in Fig. 2, noting that our analytic treatment will also cover more specialized random networks, such as those induced by bipartite graphs, or networks with multipartite structure (see Sec. VIB). The most basic elements of networks are nodes and edges, and here we allow the following features encoded in two types of labels:

- Node type, $\nu \in \mathcal{N}$: arbitrary node characteristics such as node age, susceptibility to a given disease or message, etc. The node type implicitly includes information about its degree, which we explain below.
- Edge type, $\lambda \in \Lambda$: arbitrary edge characteristics such as age, strength, conductance, etc. Since edges may be directed, edge type includes whether an edge is directed or not and its orientation if so. We thus use the notation $\bar{\lambda}$ to indicate the edge’s type when considered as travelling in the disallowed direction. (There is no need to distinguish λ or $\bar{\lambda}$ for undirected edges.) In other words, if there is a directed edge of type λ from node u to node v , we say there its type is $\bar{\lambda}$ when viewing v as the source and u as the target.

We take \mathcal{N} and Λ to be discrete. We denote the entire network by Ω , and the set of edge types incident to a node of type ν by Λ_ν .

We define degree as the number of edges of a certain type emanating from a node. In simple networks, we let $k(\nu, \lambda)$ denote the number of edges of type λ emanating from a node of type ν . In more general networks we let the multi-index $\vec{k}(\nu, \lambda) = [k_{\text{u}}(\nu, \lambda), k_{\text{i}}(\nu, \lambda), k_{\text{o}}(\nu, \lambda)]$ denote the number of undirected, inward, and outward edges of type λ belonging to a node of type ν .

The ‘total degree’ of a node of type ν is then $\vec{k}(\nu) = \sum_{\lambda \in \Lambda_\nu} \vec{k}(\nu, \lambda)$, and we define the effective degree, a scalar important for spreading mechanisms, as $k^{(\text{eff})}(\nu, \lambda) = k_{\text{o}}(\nu, \lambda) + k_{\text{u}}(\nu, \lambda)$. We also introduce a directedness indicator function $d(\lambda)$ which equals one if edges of type λ are directed and zero if not.

To characterize a random network with arbitrary node-edge-node correlations, we need to specify a number of interrelated probabilities, and these must further satisfy certain restrictions and detailed balance equations [15]. First, we have the node and edge distributions $\mathbf{Pr}(\nu)$ and $\mathbf{Pr}(\lambda)$. Note that we immediately have the restriction

$\mathbf{Pr}(\lambda) = \mathbf{Pr}(\bar{\lambda})$. Also, these induce the usual degree distributions via

$$\mathbf{Pr}(\vec{k}) = \sum_{\nu \in \mathcal{N}} \mathbf{Pr}(\nu) \delta_{\vec{k}, \vec{k}(\nu)}$$

where δ is the Kronecker delta.

Next we need $\mathbf{Pr}(\nu\lambda)$, defined as the probability that, in randomly choosing an edge and traversing it (in the allowed direction if directed or a random direction if undirected), we find it is of type λ and that we are travelling away from a node of type ν .

Finally, we encode correlations via the transition probability $\mathbf{Pr}(\nu|\nu'\lambda')$ which is the probability that we reach a type ν node, given that we are following a type λ' edge away from a type ν' node. This includes the usual degree-degree transition probabilities (see Sec. VI A and [28] for notation):

$$\begin{aligned} P^{(i)}(\vec{k}''|\vec{k}') &= \sum_{\nu, \nu' \in \mathcal{N}} \sum_{\lambda \in \Lambda_{\nu'}} \mathbf{Pr}(\nu|\nu'\lambda') d(\lambda') \\ &\quad \times \delta_{\vec{k}', \vec{k}(\nu')} \delta_{\vec{k}'', \vec{k}(\nu)}, \\ P^{(u)}(\vec{k}''|\vec{k}') &= \sum_{\nu, \nu' \in \mathcal{N}} \sum_{\lambda \in \Lambda_{\nu'}} \mathbf{Pr}(\nu|\nu'\lambda') [1 - d(\lambda')] \\ &\quad \times \delta_{\vec{k}', \vec{k}(\nu')} \delta_{\vec{k}'', \vec{k}(\nu)}. \end{aligned} \quad (1)$$

We are now forced to connect and constrain the probabilities $\mathbf{Pr}(\nu\lambda)$ and $\mathbf{Pr}(\nu|\nu'\lambda')$ according to a detailed balance constraint. Consider $\mathbf{Pr}(\nu'\lambda'\nu)$ defined as the probability that a randomly selected edge is of type λ' and runs from a type ν' node to a type ν node (corresponding to the subnetwork in Fig. 2). Then,

$$\mathbf{Pr}(\nu'\lambda'\nu) = \mathbf{Pr}(\nu|\nu'\lambda') \mathbf{Pr}(\nu'\lambda').$$

Now, if we traversed the edge in the disallowed direction, it would ‘‘connect’’ a type ν node to a type ν' node. Then we must also have $\mathbf{Pr}(\nu'\lambda'\nu) = \mathbf{Pr}(\nu\bar{\lambda}'\nu')$. We therefore arrive at the detailed balance condition:

$$\underbrace{\mathbf{Pr}(\nu|\nu'\lambda') \mathbf{Pr}(\nu'\lambda')}_{\mathbf{Pr}(\nu'\lambda'\nu)} = \underbrace{\mathbf{Pr}(\nu'|\nu\bar{\lambda}') \mathbf{Pr}(\nu\bar{\lambda}')}_{\mathbf{Pr}(\nu\bar{\lambda}'\nu')}. \quad (2)$$

Note that the detailed balance condition, Eq. (2), is more general for typed random networks than the detailed balance conditions in terms of the degree distributions $\mathbf{Pr}(\vec{k}, \vec{k}')$ and $\mathbf{Pr}(\vec{k})$ found by [15]. If the types of the nodes are their degrees and the edge types are $\Lambda = \{\text{undirected, incoming, outgoing}\}$, then Eq. (2) reduces to the well-known detailed balance conditions given in [15] and [28], which can all be written as

$$\underbrace{P^{(\lambda)}(k|k') \frac{k'_\lambda \mathbf{Pr}(k')}{\langle k_\lambda \rangle}}_{P^{(\lambda)}(k, k')} = \underbrace{P^{(\bar{\lambda})}(k'|k) \frac{k_\lambda \mathbf{Pr}(k)}{\langle k_\lambda \rangle}}_{P^{(\bar{\lambda})}(k', k)}. \quad (3)$$

In networks where there are multiple types of directed or undirected edges, the detailed balance equations given

in [15, 28], which have the form of Eq. (3), are not necessarily valid. This is because not all edges or degree- k nodes are equivalent. Using Eq. (2), we can show that the symmetry of the degree distributions is conserved, $\mathbf{Pr}(k, k') = \mathbf{Pr}(k', k)$.

$$\mathbf{Pr}(k'', k') = \sum_{\nu, \nu' \in \mathcal{N}} \sum_{\lambda \in \Lambda_\nu} \mathbf{Pr}(\nu|\nu'\lambda) \delta_{k', k(\nu')} \delta_{k'', k(\nu)}$$

In considering contagion processes, we recall the well-known typical macroscopic ‘bow-tie’ form of random networks with directed edges [14, 15, 29], given that a giant component is present. As shown in Fig. 2, there are three giant components of functional importance: (1) the giant strongly connected component, Ω_{scc} , within which any pair of nodes can be connected via a path of directed and/or undirected edges, traversing the directed ones; (2) the giant in-component Ω_{in} , the set of all nodes from which paths lead to Ω_{scc} (n.b., $\Omega_{\text{scc}} \subset \Omega_{\text{in}}$); and (3) the giant out-component Ω_{out} , the set of all nodes which can be reached along directed paths starting from a node in Ω_{in} (n.b., $\Omega_{\text{scc}} \subset \Omega_{\text{out}}$). By definition, we have that $\Omega_{\text{scc}} = \Omega_{\text{in}} \cap \Omega_{\text{out}}$. Any global spreading event must begin from a seed in the giant in-component, and can at most spread to the giant out-component Ω_{out} .

III. GENERALIZED CONTAGION PROCESS

We consider contagion processes where the probability of a node’s infection may depend in any fashion on the current states of its neighbors, potentially resembling phenomena ranging from the spread of infectious diseases to socially-transmitted behaviors [20, 30–32]. Since we are interested in the probability of spreading, we can capitalize on the fact that random networks are locally pure branching structures. We therefore need to know only what the probability of infection is for a type ν node given a single neighbor of type ν' is infected, whose influence is felt along a type λ' edge. We write this probability as $B_{\nu'\lambda'\nu}$. Time is removed from this quantity, as we need to know only the probability of eventual infection. Disease spreading models with recovery [27, 32] are included, as are threshold models inspired by social contagion [20, 31].

IV. TRIGGERING PROBABILITIES

We define $Q_{\nu\lambda}$ to be the probability that seeding a type ν node generates a global spreading event along an edge of type λ . Due to the Markovian nature of random networks, this probability must satisfy a nonlinear recursion relation:

$$\begin{aligned} Q_{\nu'\lambda'} &= \sum_{\nu \in \mathcal{N}} \mathbf{Pr}(\nu|\nu'\lambda') B_{\nu'\lambda'\nu} \\ &\quad \times \left[1 - \prod_{\lambda \in \Lambda_\nu} (1 - Q_{\nu\lambda})^{k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'}} \right], \end{aligned} \quad (4)$$

Network:	Edge Triggering Probability:	Node Triggering Probability, Q :
I. Undirected, Uncorrelated	$Q_{**} = \sum_{k_u} P^{(u)}(k_u *) B_{**k_u} \left[1 - (1 - Q_{**})^{k_u - 1} \right]$	$\sum_{k_u} \Pr(k_u) \left[1 - (1 - Q_{**})^{k_u} \right]$
II. Directed, Uncorrelated	$Q_{**} = \sum_{k_i, k_o} P^{(u)}(k_i, k_o *) B_{**k_i} \left[1 - (1 - Q_{**})^{k_o} \right]$	$\sum_{k_i, k_o} \Pr(k_i, k_o) \left[1 - (1 - Q_{**})^{k_o} \right]$
III. Mixed Directed and Undirected, Uncorrelated	$Q_{*u} = \sum_{\vec{k}} P^{(u)}(\vec{k} *) B_{**\vec{k}} \left[1 - (1 - Q_{*u})^{k_u - 1} (1 - Q_{*o})^{k_o} \right]$ $Q_{*o} = \sum_{\vec{k}} P^{(i)}(\vec{k} *) B_{**\vec{k}} \left[1 - (1 - Q_{*u})^{k_u} (1 - Q_{*o})^{k_o} \right]$	$\sum_{\vec{k}} \Pr(\vec{k}) \left[1 - (1 - Q_{*u})^{k_u} (1 - Q_{*o})^{k_o} \right]$
IV. Undirected, Correlated	$Q_{k'_u *} = \sum_{k_u} P^{(u)}(k_u k'_u) B_{**k_u} \left[1 - (1 - Q_{k'_u *})^{k_u - 1} \right]$	$\sum_{k_u} \Pr(k_u) \left[1 - (1 - Q_{k'_u *})^{k_u} \right]$
V. Directed, Correlated	$Q_{k'_i k'_o *} = \sum_{k_i, k_o} P^{(u)}(k_i, k_o k'_i, k'_o) B_{**k_i} \left[1 - (1 - Q_{k'_i k'_o *})^{k_o} \right]$	$\sum_{k_i, k_o} \Pr(k_i, k_o) \left[1 - (1 - Q_{k'_i k'_o *})^{k_o} \right]$
VI. Mixed Directed and Undirected, Correlated	$Q_{\vec{k}'_u} = \sum_{\vec{k}} P^{(u)}(\vec{k} \vec{k}') B_{**\vec{k}} \left[1 - (1 - Q_{\vec{k}'_u})^{k_u - 1} (1 - Q_{\vec{k}'_o})^{k_o} \right]$ $Q_{\vec{k}'_o} = \sum_{\vec{k}} P^{(i)}(\vec{k} \vec{k}') B_{**\vec{k}} \left[1 - (1 - Q_{\vec{k}'_u})^{k_u} (1 - Q_{\vec{k}'_o})^{k_o} \right]$	$\sum_{\vec{k}} \Pr(\vec{k}) \left[1 - (1 - Q_{\vec{k}'_u})^{k_u} (1 - Q_{\vec{k}'_o})^{k_o} \right]$

TABLE I: For the six classes of random networks described in Sec. VIA, the probability of triggering a global spreading events due to (1) an infected edge, and (2) an infected, randomly chosen single seed (see Eqs. 4 and 6). We indicate by the symbol * when no node or edge type is relevant.

an expression which involves the following three elements. First, we have $\Pr(\nu|\nu'\lambda')$ which is the probability of transitioning to a node of type ν . The second term $B_{\nu'\lambda'\nu}$ is the probability of successful infection. The last term contains the recursive structure. At least one of the edges leading away from the type ν node must generate a global spreading event (note that we avoid double counting the incident edge of type λ' with the indicator in the exponent). The probability this happens is the complement of the probability that none succeed, $\prod_{\lambda \in \Lambda_\nu} (1 - Q_{\nu\lambda})^{k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \lambda'}}$. Eq. (4) will rarely be analytically tractable (but see [28] for an exactly solved simple model), and will usually be solved numerically by iteration.

The probability that an infected type ν node seeds a global spreading event follows as

$$Q_\nu = 1 - \prod_{\lambda \in \Lambda_\nu} (1 - Q_{\nu\lambda})^{k^{(\text{eff})}(\nu, \lambda)}, \quad (5)$$

where again success is defined in terms of not failing. Finally, the probability that the sole infection of a randomly chosen node leads to a global spreading event is

$$Q = \sum_{\nu' \in \mathcal{N}} \Pr(\nu') Q_{\nu'} \quad (6)$$

The effects of weighted triggering schemes—where the initial node is chosen according to its degree in some

fashion—can be easily examined by replacing $\Pr(\nu')$ with the appropriate distribution.

V. CONNECTION BETWEEN TRIGGERING PROBABILITIES AND THE CONTAGION CONDITION

We show how our general expression for triggering probabilities reduces to the general cascade condition we described in [27]. The calculation involved makes an important analytic connection but is necessarily largely mathematical in nature, as represented in Fig. 2.

The cascade condition is a binary expression of possibility; when the condition is met, global spreading events initiated by single seeds are possible, and otherwise they are impossible. Starting from Eq. (4), we determine the cascade condition by examining under what circumstances the triggering probability $Q_{\nu\lambda} \rightarrow 0^+$. In this limit, the product in Eq. (4), $\prod_{\lambda \in \Lambda_\nu} (1 - Q_{\nu\lambda})^{k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \lambda'}}$ can be approximated as $1 - \sum_{\lambda \in \Lambda_\nu} (k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \lambda'}) Q_{\nu\lambda}$, to first order.

Neglecting higher order terms, Eq. (4) reduces to

$$Q_{\nu'\lambda'} \simeq \sum_{\nu \in \mathcal{N}} \mathbf{Pr}(\nu|\nu'\lambda') B_{\nu|\lambda'\nu'} \times \sum_{\lambda \in \Lambda_\nu} \left(k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'} \right) Q_{\nu\lambda}. \quad (7)$$

We introduce the notation from [27], where $\alpha = (\nu, \lambda)$ and $\alpha' = (\nu', \lambda')$, as well as $k_{\alpha'\alpha} = k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'}$ as the number of type λ edges leaving from nodes of type ν , with the exclusion of the incident type λ' edge arriving from a type ν' node. We also let $P_{\alpha'\alpha} = \mathbf{Pr}(\nu|\nu'\lambda')$, $B_{\alpha'\alpha} = B_{\nu'\lambda'\nu}$. Note that the outgoing edge of type λ' does not affect the contagion mechanism and is left as arbitrary in α' . Then the above equation becomes

$$Q_{\alpha'} \simeq \sum_{\alpha} P_{\alpha'\alpha} \bullet k_{\alpha'\alpha} \bullet B_{\alpha'\alpha} Q_{\alpha} = \sum_{\alpha} R_{\alpha'\alpha} Q_{\alpha}, \quad (8)$$

where we have identified the gain matrix R we obtained and described in [27]. Contagion is possible only when the largest eigenvalue of R exceeds unity, and we have connected the triggering probability to the cascade condition.

VI. APPLICATIONS

A. Triggering probabilities for six random network families

In Tab. I, we list the forms of $Q_{\nu'\lambda'}$ and Q for six specific families of random networks which we describe below. The last of these network families is the most general and contains the other five as special cases. Nodes potentially have three kinds of unweighted edges incident to them: undirected, in-directed, and out-directed, and we use the vector representation $\vec{k} = (k_u, k_i, k_o)$ to define node classes [15, 27]. The specific transition probabilities, $P^{(i)}(\vec{k}|\vec{k}')$, $P^{(o)}(\vec{k}|\vec{k}')$, and $P^{(u)}(\vec{k}|\vec{k}')$, give the probabilities of an edge leading from a degree \vec{k}' node to a degree \vec{k} node being oriented as undirected, incoming, or outgoing (see Refs. [27] and [28] for more details). For uncorrelated networks, we use the notation $P^{(i)}(\vec{k}|\ast)$, etc. Similarly for the triggering probabilities, where the node or edge type is irrelevant we also use \ast (e.g., $Q_{\ast\ast}$ instead of $Q_{\nu\lambda}$ for undirected, uncorrelated, unweighted networks). For simplicity, we assume infection is due only to properties of the node potentially being infected, which for these networks means the node's degree.

B. Random bipartite networks

We now show how the theory of contagion in bipartite networks [14] is a special case of the general model. Consider a bipartite network $G = (V, E)$ with the nodes partitioned into disjoint sets $A^{(1)}$ and $A^{(2)}$, such that

$V = A^{(1)} \cup A^{(2)}$ and all edges $uv \in E$ satisfy $u \in A^{(1)}$ and $v \in A^{(2)}$ or $u \in A^{(2)}$ and $v \in A^{(1)}$. Again, we consider general node types ν , but now they are also associated with either one of the sets $A^{(1)}$ or $A^{(2)}$.

Due to the bipartite structure, the triggering probability Eq. (4) separates into two coupled equations

$$Q_{\nu'\lambda'}^{(1)} = \sum_{\nu} \mathbf{Pr}^{(1)}(\nu|\nu'\lambda') B_{\nu'\lambda'\nu}^{(1)} \times \left[1 - \prod_{\lambda \in \Lambda_\nu} \left(1 - Q_{\nu\lambda}^{(2)} \right)^{k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'}} \right], \quad (9)$$

$$Q_{\nu'\lambda'}^{(2)} = \sum_{\nu} \mathbf{Pr}^{(2)}(\nu|\nu'\lambda') B_{\nu'\lambda'\nu}^{(2)} \times \left[1 - \prod_{\lambda \in \Lambda_\nu} \left(1 - Q_{\nu\lambda}^{(1)} \right)^{k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'}} \right],$$

where the superscripts denote the triggering probabilities starting in $A^{(1)}$ and $A^{(2)}$, respectively.

The contagion condition arises again by linearizing Eq. (9) about $Q^{(1)} = Q^{(2)} = 0$. This gives the linear system of equations

$$Q_{\nu'\lambda'}^{(1)} = \sum_{\nu} \sum_{\lambda \in \Lambda_\nu} \mathbf{Pr}^{(1)}(\nu|\nu'\lambda') B_{\nu'\lambda'\nu}^{(1)} \times \left(k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'} \right) Q_{\nu\lambda}^{(2)}, \quad (10)$$

$$Q_{\nu'\lambda'}^{(2)} = \sum_{\nu} \sum_{\lambda \in \Lambda_\nu} \mathbf{Pr}^{(2)}(\nu|\nu'\lambda') B_{\nu'\lambda'\nu}^{(2)} \times \left(k^{(\text{eff})}(\nu, \lambda) - \delta_{\lambda, \bar{\lambda}'} \right) Q_{\nu\lambda}^{(1)}. \quad (11)$$

These equations are of the form

$$\begin{bmatrix} Q^{(1)} \\ Q^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & R_{12} \\ R_{21} & 0 \end{bmatrix} \begin{bmatrix} Q^{(1)} \\ Q^{(2)} \end{bmatrix} = R \begin{bmatrix} Q^{(1)} \\ Q^{(2)} \end{bmatrix}, \quad (12)$$

where the entries of R_{12} and R_{21} are shown in Eqs. (10) and (11). The structure of the gain matrix R , of course, reflects the bipartiteness of G . Spreading will occur when the spectral radius $\rho(R) > 1$ [27]. The eigenvalues of R are the solutions λ to

$$\det(\lambda^2 I - R_{12} R_{21}) = 0,$$

since the diagonal matrix λI and R_{21} commute [33]. The eigenvalues of R are thus the square roots of the eigenvalues of $R_{12} R_{21}$, meaning we can also express the contagion condition as $\rho(R_{12} R_{21}) > 1$.

There is a physical explanation for the contagion condition. Assume the contagion starts with one active node in $A^{(1)}$. It then must pass to $A^{(2)}$ before returning to $A^{(1)}$. The gain going from $A^{(1)}$ to $A^{(2)}$ is R_{12} , and the gain is R_{21} going from $A^{(2)}$ to $A^{(1)}$. If the expected number of active nodes after these two passes exceeds unity, the contagion will spread. Note that the spectra of $R_{12} R_{21}$ and $R_{21} R_{12}$ are equal, so that we could also consider starting the contagion in $A^{(2)}$.

C. Uncorrelated, undirected bipartite networks

We now confirm that the general theory gives the previously known results for uncorrelated, undirected bipartite networks. These networks are fully specified by the degree distributions $P^{(1)}(k)$ and $P^{(2)}(k)$ for nodes in sets $A^{(1)}$ and $A^{(2)}$, respectively. We set the infection probability $B^{(1)} = B^{(2)} = 1$ for all nodes, so that we are solving for the existence of a giant component. The edge probabilities are

$$P^{(1)}(k|*) = \frac{kP^{(1)}(k)}{\sum_k kP^{(1)}(k)} \quad (13)$$

$$P^{(2)}(k|*) = \frac{kP^{(2)}(k)}{\sum_k kP^{(2)}(k)} \quad (14)$$

where $P^{(1)}(k|*)$ is the probability of reaching a degree k node in $A^{(1)}$ from a random node in $A^{(2)}$, and $P^{(2)}(k'|*)$ is likewise the probability of reaching a degree k' node in $A^{(2)}$ from a random node in $A^{(1)}$.

Pick a random node $u \in A^{(1)}$ and imagine that the contagion arrives at u via one of its incoming edges. Then there are an expected $\sum_k (k-1)P^{(1)}(k|*) = R_{12}$ edges leftover, each leading to an unexplored node in $A^{(2)}$. Follow one of these to $v \in A^{(2)}$, then the expected excess edges coming from v is $\sum_{k'} (k'-1)P^{(2)}(k'|*) = R_{21}$. Multiplying these two sums together gives the expected number of new nodes reached in $A^{(1)}$ after two passes, so the contagion condition is

$$R_{12}R_{21} = \sum_{k,k'} (k-1)P^{(1)}(k|*)(k'-1)P^{(2)}(k'|*) > 1. \quad (15)$$

Substituting (13) and (14) for the conditional probabilities, taking the normalization factors to the right hand side, and simplifying, we arrive at

$$\sum_{k,k'} kk'(kk' - k - k')P^{(1)}(k)P^{(2)}(k') > 0 \quad (16)$$

which is the condition found by Newman, Strogatz, and Watts [14] using generating functions. While Eqs. 15 and 16 are equivalent, the former preserves the physics of the spreading process.

VII. CONCLUDING REMARKS

We have shown that the probability of a single infected node generating a global spreading event can be derived in a straightforward way for spreading processes on a very general class of correlated random networks. Our approach brings a physical intuition to the problem, and while more sophisticated mathematical analyses arrive at the same results, and are certainly useful for more detailed investigations, they are burdened with some degree of inscrutability.

Acknowledgments

We appreciate discussions with Braden Brinkman. KDH was supported by VT-NASA EPSCoR and a Boeing fellowship; JLP was supported by NIH grant # K25-CA134286; PSD was supported by NSF CAREER Award # 0846668.

-
- [1] P. J. Richerson and R. Boyd, *Not by Genes Alone* (University of Chicago Press, Chicago, IL, 2005).
 - [2] A. Chmiel, J. Sienkiewicz, M. Thelwall, G. Paltoglou, K. Buckley, A. Kappas, and J. A. Hoyst, *PLoS ONE* **6**, e22207 (2011).
 - [3] D. M. Romero, B. Meeder, and J. Kleinberg, in *Proceedings of World Wide Web Conference* (2011).
 - [4] P. Rozin and E. Royzman, *Personality and Social Psychology Review* **5**, 296 (2001).
 - [5] J. Leskovec, L. Backstrom, and J. Kleinberg, in *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (2009), pp. 497–506.
 - [6] J. Berger and G. Le Mens, *Proc. Natl. Acad. Sci.* **106**, 8146 (2009).
 - [7] A. V. Banerjee, *Quart. J. Econ.* **107**, 797 (1992).
 - [8] S. G. Barsade, *Administrative Science Quarterly* **47**, 644 (2002).
 - [9] S. Bikhchandani, D. Hirshleifer, and I. Welch, *J. Polit. Econ.* **100**, 992 (1992).
 - [10] E. Rogers, *The Diffusion of Innovations* (Free Press, New York, 1995), Fifth ed.
 - [11] P. Siczka, D. Sornette, and J. A. Holyst, *Eur. Phys. J. B* **82**, 257 (2011).
 - [12] A. D. Cliff, P. Haggett, J. K. Ord, and G. R. Versey, *Spatial diffusion: an historical geography of epidemics in an island community* (Cambridge University Press, Cambridge, UK, 1981).
 - [13] V. Colizza, A. Barrat, M. Barthelmeij, A.-J. Valleron, and A. Vespignani, *PLoS Med.* **4**, e13 (2007).
 - [14] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. E* **64**, 026118 (2001).
 - [15] M. Boguñá and M. Ángeles Serrano, *Phys. Rev. E* **72**, 016106 (2005).
 - [16] L. A. Meyers, M. Newman, and B. Pourbohloul, *J. Theor. Biol.* **240**, 400 (2006).
 - [17] J. P. Gleeson and D. J. Cahalane, *Phys. Rev. E* **75**, 056103 (2007).
 - [18] J. P. Gleeson, *Phys. Rev. E* **77**, 046117 (2008).
 - [19] J. P. Gleeson, S. Melnik, and A. Hackett, *Phys. Rev. E* **81**, 066114 (2010).
 - [20] D. J. Watts, *Proc. Natl. Acad. Sci.* **99**, 5766 (2002).
 - [21] A. Hackett, S. Melnik, and J. P. Gleeson, *Phys. Rev. E* **83**, 056107 (2011).
 - [22] Y. Ikeda, T. Hasegawa, and K. Nemoto, *Journal of Physics: Conference Series* **221**, 012005 (2010).
 - [23] P. Munz, I. Hudea, J. Imad, and R. J. Smith?, in *Infectious Disease Modelling Research Progress*, edited by J. M. Tchuente and C. Chiyaka (Nova Science Publish-

- ers, Inc., 2009), pp. 133–150.
- [24] D. J. Watts and P. S. Dodds, in *The Oxford Handbook of Analytical Sociology*, edited by P. Hedström and P. Bearman (Oxford University Press, Oxford, UK, 2009), chap. 20, pp. 475–497.
- [25] H. S. Wilf, *Generatingfunctionology* (A K Peters, Natick, MA, 2006), 3rd ed.
- [26] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
- [27] P. S. Dodds, K. D. Harris, and J. L. Payne, *Phys. Rev. E* **83**, 056122 (2011).
- [28] J. L. Payne, K. D. Harris, and P. S. Dodds, *Phys. Rev. E* **84**, 016110 (2011).
- [29] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, *Comput. Netw.* **33**, 309 (2000).
- [30] T. C. Schelling, *J. Conflict Resolut.* **17**, 381 (1973).
- [31] M. Granovetter, *Am. J. Sociol.* **83**, 1420 (1978).
- [32] J. D. Murray, *Mathematical Biology* (Springer, New York, 2002), Third ed.
- [33] J. R. Silvester, *The Mathematical Gazette* pp. 460–467 (2000).