Clustering determines the dynamics of complex contagions in multiplex networks

Yong Zhuang,[1] Alex Arenas,[2] and Osman Yağan[1]

[1]*Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
[2]*Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain*

We present the mathematical analysis of generalized complex contagions in a class of clustered multiplex networks. The model is intended to understand spread of influence, or any other spreading process implying a threshold dynamics, in setups of interconnected networks with significant clustering. The contagion is assumed to be general enough to account for a content-dependent linear threshold model, where each link type has a different weight (for spreading influence) that may depend on the content (e.g., product, rumor, political view) that is being spread. Using the generating functions formalism, we determine the conditions, probability, and expected size of the emergent *global* cascades. This analysis provides a generalization of previous approaches and is especially useful in problems related to spreading and percolation. The results present nontrivial dependencies between the clustering coefficient of the networks and its average degree. In particular, several phase transitions are shown to occur depending on these descriptors. Generally speaking, our findings reveal that increasing clustering decreases the probability of having global cascades and their size, however, this tendency changes with the average degree. There exists a certain average degree from which on clustering favors the probability and size of the contagion. By comparing the dynamics of complex contagions over multiplex networks and their monoplex projections, we demonstrate that ignoring link types and aggregating network layers may lead to inaccurate conclusions about contagion dynamics, particularly when the correlation of degrees between layers is high.

I. INTRODUCTION

The study of dynamical processes on real-world complex networks has been an active research area over the past decade. Some of the most widely studied problems include cascading failures in interdependent networks [1–6], *simple* contagions (e.g., disease propagation in human and animal populations [7–25], etc.), and *complex* contagions (e.g., diffusion of influence, beliefs, norms, and innovations in social networks [26–29]). Recently, the attention was shifted from single, isolated networks to multiplex and multilayer networks [17,18,29–43]. This shift is primarily driven by the observation that links in a network might be categorized according to the nature of the relationship they represent (e.g., friends, family, office mates) as well as according to the social network to which they belong (e.g., Google+ vs Facebook links), and each link type might play a different role in the dynamical process.

In this work, we focus on the analysis of complex contagion processes that take place on multiplex (or multilayer) networks. In doing so, we adopt the content-dependent linear threshold model of social contagions proposed by Yağan and Gligor [29]. Their framework is a generalization of the linear threshold model introduced by Watts [28] and is based on individuals adopting a behavior when their *perceived* proportion of *active* neighbors exceeds a certain threshold; the key to that modeling framework is that one's perceived influence depends on the *types* of the relationships they have and the *context* in which diffusion is being considered. More precisely, each individual in the network can be in one of the two states, *active* or *inactive*. Each link type i is associated with a content-dependent weight c_i in $[0,\infty]$ that encodes the relative importance of this link type in spreading the given content. Then, an inactive node with m_i active neighbors and $d_i - m_i$ inactive neighbors via type-i links turns active

only if

$$\frac{\sum_i c_i m_i}{\sum_i c_i d_i} \geqslant \tau,$$

where τ is the node's threshold drawn from a distribution $P(\tau)$. It is assumed that nodes update their state synchronously and once *active*, a node stays active forever.

Yağan and Gligor analyzed [29] the content-dependent linear threshold model in multiplex networks and derived the conditions, probability, and expected size of *global* cascades, i.e., cases where activating a randomly selected node leads to activation of a positive fraction of the population in the limit of large system sizes. However, their multiplex network model was formed by combining independent layers of networks (one for each link type), where each layer is generated by the *configuration model* [44,45]. Although a good starting point, configuration model is known to generate networks that can not accurately capture some important aspects of real-world social networks, most notably the property of *high clustering* [46,47]. Informally known as the phenomenon that "friends of our friends" are likely to be our friends, clustering has been shown to affect the dynamics of various diffusion processes [31,48–55] significantly.

With this in mind, our main contribution in this paper is to provide a thorough analysis of influence spread process in a class of *clustered* multiplex networks. In particular, we study the content-dependent linear threshold model in a multiplex network model where each link type (or network layer) is formed by the clustered random network model proposed by Newman [48] and Miller [56]. We solve for the critical threshold, probability, and expected size of global cascades and confirm our analytical results via extensive simulations. The main observation from our results is that clustering has a double-faceted impact on the probability and expected size of

global cascades. Namely, we show that clustering decreases the probability and size of cascades when average degree in the network is small, whereas after a certain value of average degree, clustering is shown to facilitate cascades.

We also compare the dynamics of complex contagions over multiplex networks and their *monoplex* projections. There has been recent interest [30] in understanding whether monoplex projection of a multiplex network (obtained by ignoring the colors of edges and aggregating the layers) can still capture the essential properties (e.g., cascade threshold and size) of a diffusion process. In the affirmative, this would eliminate the need for considering the full multiplex structure of real-world systems in tackling similar problems. We show that even in the simplest case where all link types have the same influence weight (i.e., $c_1 = c_2 = \cdots$), monoplex theory may *not* be able to capture contagion dynamics accurately, reinforcing the need for studying multiplex networks in its correct setup. We observe that the accuracy of monoplex theory in capturing cascade dynamics over multiplex networks depends tightly on the assortativity (i.e., correlation between the degrees of connected pairs) of the network. For instance, when assortativity is negligible, monoplex theory is seen to predict cascade dynamics very well, while in highly assortative cases its ability to predict contagion behavior diminishes significantly.

Finally, we proof the possibility of an unforeseen behavior in the dynamics of complex contagions in multiplex networks, i.e., that of observing more than two phase transitions in the cascade size as the mean degrees in network layers increase. It has been reported [18,28] many times that threshold models of complex contagion exhibit two phase transitions as the average degree increases; a second-order transition at *low* degrees marking the formation of a giant component of *vulnerable* nodes and a first-order transition at *high* degrees due to increased *local* stability of nodes. Here, we consider a multiplex where one layer has degree distribution Poi(λ) while the degree in the second layer follows Poi(λ/α) with probability α and is zero with probability $1 - \alpha$. In this setting, we observe that in general there exist two intervals of λ for which cascades are possible, amounting to *four* phase transitions as opposed to two; also, it is seen that only the first transition is second order while the remaining ones are first order. However, depending on the value of α, these regions may overlap (with overlap starting when α exceeds a critical value) resulting again with only two phase transitions; see Sec. VI C for details.

The paper is organized as follows. We give details of the models applied in this study and the problem to be considered in Sec. II. In Secs. III and IV we present the main results of this work, and confirm it through extensive computer simulations in Sec. V. In Sec. VI, we make a comparison between complex contagions in monoplex networks and multiplex networks, and also demonstrate the phenomenon about the number of phase transitions. Finally, Sec. VII summarizes our work and gives future directions.

II. MODEL: STRUCTURE AND DYNAMICS

A. Random graphs with clustering

Our goal is to study complex contagion processes in synthetic networks that capture some important aspects of real-world networks but otherwise are generated *randomly*. It is known [44,45] that the widely used configuration model [44] generates treelike graphs with number of cycles approaching to zero as the number of nodes gets large. However, most social networks exhibit *high clustering*, informally known as the propensity of a "friend of a friend" to be one's friend. Put differently, real-world social networks are usually not treelike and instead have considerable number of cycles, particularly of size three; i.e., *triangles*. With this in mind, Miller [57] and Newman [45] proposed a modification on the configuration model to enable generating random graphs with given degree distributions and tunable clustering.

The model proposed in [45,57] is often referred to as random networks with clustering and is based on the following algorithm. Given a joint degree distribution $\{p_{st}\}_{s,t=0}^{\infty}$ that gives the probability that a node has s single edges and t triangles, each node will be given s stubs labeled as *single* and $2t$ stubs labeled as *triangles* with probability p_{st}, for any $s,t = 0,1,\dots$. Then, stubs that are labeled as single are joined *randomly* to form single edges that are not part of a triangle, whereas *pairs* of triangle stubs from three nodes are *randomly* matched to form triangles between the three participating nodes; the total degree of a node is then distributed by $p_k = \sum_{s,t:s+2t=k} p_{st}$. As in the standard configuration model, it can be shown that the number of cycles formed by single edges goes to zero as n gets large, and so does the number of cycles of length larger than three [44].

We quantify the level of clustering using the widely recognized *global* clustering coefficient [45], defined via

$$C_{\text{global}} = \frac{3 \times (\text{number of triangles in network})}{\text{number of connected triples}}.$$

Here, "connected triples" means a single vertex connected by edges to two others. It was shown in [44] that C_{global} is *positive* in the random clustered network model, while it approaches to zero with increasing network size in the standard configuration model.

B. Multilayer and multiplex network models

In this paper, we consider a multiplex network where links are classified into different types, *or* colors; they are referred to as *edge-colored multigraphs* by some authors. For ease of exposition, we consider the case with only two colors, *red* and *blue*, but the discussion can easily be extended to arbitrary number of colors. Let $\mathbb{R}$ and $\mathbb{B}$ denote the subnetworks formed by red and blue edges, respectively. A possible motivation is that $\mathbb{R}$ models the *kinship* contact network among individuals, while the network $\mathbb{B}$ stands for the colleagueship network. Alternatively, we can think of $\mathbb{B}$ modeling the physical (e.g., face-to-face) relationships among human beings while $\mathbb{R}$ models connections through an online social network (e.g., Facebook).

In line with the second motivation, we assume that network $\mathbb{B}$ is defined on the vertices $\mathcal{N} = \{1, \dots, n\}$, while $\mathbb{R}$ contains only a *subset* of the nodes in $\mathcal{N}$ to account for the fact that not every individual participates in online social networks; e.g., we assume that each vertex in $\mathcal{N}$ participates in $\mathbb{R}$ independently with probability $\alpha \in (0,1]$, meaning that the set of vertices of $\mathbb{R}$ constitutes α fraction of the whole population. We illustrate
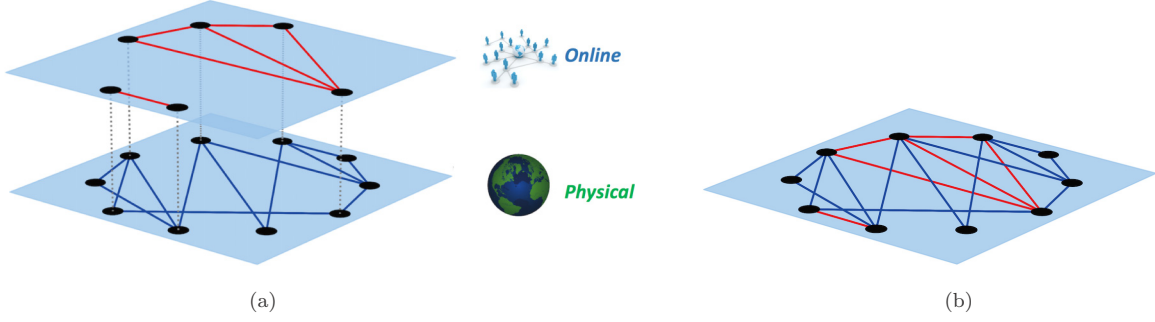
FIG. 1. Illustration of multilayer and multiplex network representations of our model. In (a), we see a multilayer network (e.g., a physical communication layer and an online social network layer) with overlapping vertex sets; vertical dashed lines represent nodes corresponding to the same individual. In (b), we see the equivalent representation of this model by a multiplex network, i.e., an edge-colored multigraph. Edges from online social networks are shown in red and edges from the physical network are shown in blue; although not shown in this particular example, it is possible for a pair of nodes to be connected by both blue and red links, rendering the resulting representation a *multigraph*.

in Fig. 1 two equivalent representations of this model, first shown as a multilayer network with overlapping vertex sets, and second as a multiplex network.

We generate both $\mathbb{R}$ and $\mathbb{B}$ from the generalized configuration model described in Sec. II A, i.e., both are random networks with clustering. In particular, we let $\{p_{st}^r, s, t = 0, 1, \dots\}$ and $\{p_{st}^b, s, t = 0, 1, \dots\}$ denote the joint distributions for single edges and triangles for $\mathbb{R}$ and $\mathbb{B}$, respectively. Then, both networks are generated independently according to the algorithm described in Sec. II A, and they are denoted, respectively, by $\mathbb{R} = \mathbb{R}(n; \alpha, p_{st}^r)$ and $\mathbb{B} = \mathbb{B}(n; p_{st}^b)$. We define the overall network $\mathbb{H}$ over which influence spreads as the *disjoint* union $\mathbb{H} = \mathbb{R} \bigsqcup \mathbb{B}$ and represent it by $\mathbb{H}(n; \alpha, p_{st}^r, p_{st}^b)$. Here, the disjoint union operation implies that we still distinguish $\mathbb{R}$ edges from $\mathbb{B}$ edges in $\mathbb{H}$, meaning that it is a *multiplex* network.

We denote the *colored* degree $\mathbf{d}$ of a node in $\mathbb{H}$ by

$$\mathbf{d} = (d_{rs}, 2n_{rt}, d_{bs}, 2n_{bt}) \qquad (1)$$

meaning that it has d_{rs} single edges and $2n_{rt}$ triangle edges in network $\mathbb{R}$, and d_{bs} single edges and $2n_{bt}$ triangle edges in network $\mathbb{B}$; here, n_{rt} and n_{bt} are defined as the number of *triangles* assigned to this node in $\mathbb{R}$ and $\mathbb{B}$, respectively. The distribution of this colored degree is denoted by $p_{\mathbf{d}}$ and can be computed directly from p_{st}^r, p_{st}^b, and α.

C. Content-dependent linear threshold model for social contagion

The classical linear threshold model by Watts [28] is based on individuals adopting a behavior when the fraction of their *active* neighbors exceed a certain threshold. Namely, an inactive node i with m_i active neighbors and $d_i - m_i$ inactive neighbors will become active only if m_i/d_i exceeds τ_i drawn from a distribution $P(\tau)$. More precisely, nodes update their states *synchronously* at discrete time steps $t = 0, 1, \dots$, and an *inactive* node will be activated at time t if the fraction of its *active* neighbors exceeds its threshold at time $t - 1$; once active, a node can not be deactivated. A major concern with this model is that it assumes all links in the network have the same importance, irrespective of the context that the spreading is being considered. However, in real-world contagion processes, each link type (e.g., co-workership versus family or physical

links versus online social network links) may play a different role in different cascade processes. For example, in the spread of a new consumer product amongst the population, a video game would be more likely to be promoted among high school classmates rather than among family members; the situation would be exactly the opposite in the case of a new cleaning product [58].

To address the aforementioned drawbacks, Yağan and Gligor [29] proposed a content-dependent linear threshold model for social contagion in multiplex networks. In this model, each link type is associated with a content-dependent parameter c_i in $[0, \infty]$ that measures the relative bias type-i links have in spreading the content. Then, an inactive node with m_i active neighbors and $d_i - m_i$ inactive neighbors through link type i will turn *active* if $\frac{\sum_i c_i m_i}{\sum_i c_i d_i} \geqslant \tau$.

In this work, we will analyze complex contagions in $\mathbb{H}$ under the content-dependent threshold model introduced in [29]. Consider a node with colored degree $\mathbf{d} = (d_{rs}, 2n_{rt}, d_{bs}, 2n_{bt})$ and *active* degree

$$\mathbf{m} = (m_{rs}, m_{rt1}, m_{rt2}, m_{bs}, m_{bt1}, m_{bt2}),$$

where m_{rs} (resp. m_{bs}) gives the number of *active* neighbors connected through *red single* edges (resp. *blue single* edges), and m_{rt1} and m_{rt2} (resp. m_{bt1} and m_{bt2}) give the number of red (resp. blue) triangles with one and two active neighbors, respectively; see Fig. 2 for demonstration of three cases counted as m_{rs}, m_{rt1}, and m_{rt2}, respectively. Next, for a given content to spread over $\mathbb{H}$, let c_r and c_b denote the weight of *red* and *blue* edges, respectively, in spreading this content. Without
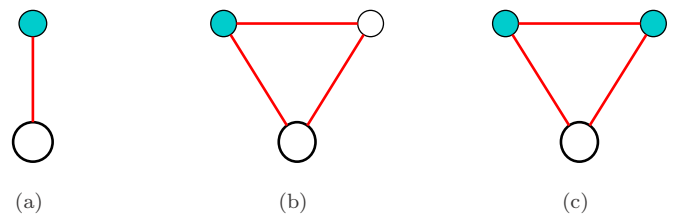

FIG. 2. Illustration of three cases that would be counted as m_{rs}, m_{rt1}, and m_{rt2}, respectively, for the number of active nodes. Nodes shown in filled (green) circles are *active* while those shown in nonfilled circles are *inactive*.

loss of generality, we set $c := \frac{c_r}{c_b}$. Then, the probability that an inactive node with degree $\mathbf{d}$ and active degree $\mathbf{m}$ turns active is given by

$$
\begin{aligned}
&F(\mathbf{m},\mathbf{d}) \\
&= \mathbb{P}\left[\frac{c(m_{rs}+m_{rt1}+2m_{rt2})+m_{bs}+m_{bt1}+2m_{bt2}}{c(d_{rs}+2n_{rt})+d_{bs}+2n_{bt}} \geqslant \tau\right].
\end{aligned}
\tag{2}
$$

Hereafter, the function $F(\mathbf{m},\mathbf{d})$ will be referred to as the *neighborhood response function* [52,59].

D. Problem

We consider the diffusion of influence over $\mathbb{H}$ that is initiated by a node selected uniformly at random. Our main goal is to derive the conditions, probability, and expected size of *global* cascades, i.e., cases where influence starts from a single individual and reaches a positive fraction of the population in the large n limit. Of particular interest will be to reveal the effect of clustering coefficient C_{global} and content parameter c on these quantities.

III. CONDITION AND PROBABILITY OF GLOBAL CASCADES

In this section, we derive the condition and probability of global cascades in clustered multiplex networks; expected size of global cascades is handled separately in Sec. IV. As mentioned in Sec. II B, we restrict our attention to networks with only two link types, labeled as *blue* and *red* edges, respectively. Distinguishing further the edges based on whether or not they are part of a triangle, we obtain four types of edges in our clustered multiplex network model, labeled as *red single* edges, *red triangle edges*, *blue single* edges, and *blue triangle edges*; these are denoted by $rs-$, $rt-$, $bs-$, and $bt-$, respectively.

To analyze the influence diffusion process, we consider a branching process [60] that starts by activating a node selected uniformly at random from among all nodes. Starting with the neighbors of the seed node, we explore and identify all nodes that are reached and activated, continuing recursively until the branching process stops. Since the contagion model considered here is *monotone*, i.e., nodes that are activated stay active forever, the branching process is guaranteed to stop, and the resulting number of nodes reached gives the cascade size.

Let $H(x)$ denote the generating function [61] for the "*finite number of nodes that are reached and influenced*" by the branching process [60] initiated by a node selected uniformly at random. We will derive an expression for $H(x)$ using $h_{rs}(x)$, $h_{rt}(x)$, $h_{bs}(x)$, and $h_{bt}(x)$, where $h_{rs}(x)$ [resp. $h_{bs}(x)$] stands for the generating function for the "*finite* number of nodes reached by following a randomly selected *red single* (resp. *blue single*) edge"; $h_{rt}(x)$ and $h_{bt}(x)$ are defined similarly for *red triangle* and *blue triangle* edges, respectively. Then, $H(x)$ takes the form

$$
H(x) = x \sum_{\mathbf{d}} p_{\mathbf{d}} D(d_{rs},n_{rt},d_{ws},n_{wt}),
\tag{3}
$$

where

$$
D(y,z,m,\ell) := h_{rs}(x)^y h_{rt}(x)^z h_{bs}(x)^m h_{bt}(x)^\ell.
\tag{4}
$$

The validity of (3) can be seen as follows. The term x stands for the node that is selected randomly and set *active* to initiate the cascade. This node has a degree $\mathbf{d} = (d_{rs},2n_{rt},d_{bs},2n_{bt})$ with probability $p_{\mathbf{d}}$. The number of nodes reached by each of its d_{rs} (resp. d_{bs}) *red single* edges (resp. *blue single* edges) has a generating function $h_{rs}(x)$ [resp. $h_{bs}(x)$]. Considering its triangle edges in a similar manner, we see from the *powers* property of generating functions [44] that when the initial node has degree $\mathbf{d}$, the number of nodes influenced in this process has a generating function $h_{rs}(x)^{d_{rs}} h_{rt}(x)^{n_{rt}} h_{bs}(x)^{d_{bs}} h_{bt}(x)^{n_{bt}}$. Taking the expectation over the degree $\mathbf{d}$ of the initial node, we get (3).

For (3) to be useful, we shall derive expressions for the generating functions $h_{rs}(x)$, $h_{rt}(x)$, $h_{bs}(x)$, and $h_{bt}(x)$. As will become apparent soon, there are no explicit equations defining these functions. Instead, we should seek to derive *recursive* equations to define each generating function in terms of the others. These steps are taken in the next sections where we first focus on deriving $h_{rs}(x)$ and $h_{bs}(x)$ (Sec. III A) followed by derivations of $h_{rt}(x)$ and $h_{bt}(x)$ (Sec. III B).

Given that random networks with clustering are free of cycles of size larger than three, it is clear that the initial stages of the branching process will expand largely because of *vulnerable* nodes that can get activated either by one or two active neighbors. In our formulation, the multiplex nature of the network calls for defining the notion of the vulnerability with respect to link types as well [29]. Throughout, we say that a node is $\mathbb{R}$-vulnerable (resp. $\mathbb{B}$-vulnerable) if it gets activated by a *single* active connection through a *red* link (resp. *blue* link). We define $\rho_{\mathbf{d},rs}$ and $\rho_{\mathbf{d},bs}$ as the probability that a node is $\mathbb{R}$-vulnerable and $\mathbb{B}$-vulnerable, respectively. We also define $\rho_{\mathbf{d},rt}$ (resp. $\rho_{\mathbf{d},bt}$) as the probability that a node gets activated by having *two* active neighbors via *red* (resp. *blue*) edges. In other words, $\rho_{\mathbf{d},rt}$ (resp. $\rho_{\mathbf{d},bt}$) gives the probability that a node gets activated by having a red (resp. blue) triangle with both neighbors being active; see Fig. 2. More precisely, we set $\rho_{\mathbf{d},rs} = F[(1,0,0,0,0,0),\mathbf{d}]$, $\rho_{\mathbf{d},rt} = F[(0,0,1,0,0,0),\mathbf{d}]$, $\rho_{\mathbf{d},bs} = F[(0,0,0,1,0,0),\mathbf{d}]$, and $\rho_{\mathbf{d},bt} = F[(0,0,0,0,0,1),\mathbf{d}]$.

A. Influence propagation via *red single* edges

We start by deriving recursive equations for $h_{rs}(x)$ and $h_{bs}(x)$ by focusing on the number of nodes reached and influenced by following one end of a single edge in $\mathbb{R}$ and $\mathbb{B}$, respectively. In what follows, we only derive $h_{rs}(x)$ since the computation of $h_{bs}(x)$ follows in a very similar manner. In order to compute $h_{rs}(x)$, consider picking a *red* single edge uniformly at random (among all red single edges in $\mathbb{H}$) and assume that it is connected at one end to an *active* node. Then, we compute the generating function for the number of nodes influenced by following the other end of the edge, and obtain the following expression for the generating function $h_{rs}(x)$:

$$
\begin{aligned}
h_{rs}(x) = {}&x \sum_{\mathbf{d}} \frac{d_{rs} p_{\mathbf{d}}}{\langle d_{rs} \rangle} \rho_{\mathbf{d},rs} D(d_{rs}-1,n_{rt},d_{bs},n_{bt}) \\
&+ x^0 \sum_{\mathbf{d}} \frac{d_{rs} p_{\mathbf{d}}}{\langle d_{rs} \rangle} (1 - \rho_{\mathbf{d},rs}),
\end{aligned}
\tag{5}
$$

where D is as defined at (4).

We now explain each term appearing at (5) in turn. The explicit factor x stands for the initial vertex that is arrived at by following the randomly selected red single edge. The term $\frac{d_{rs} p_{\mathbf{d}}}{\langle d_{rs} \rangle}$ gives the *normalized* probability that the arrived vertex has colored degree $\mathbf{d}$. Since the arrived node is reached by a *red* link, it needs to be *red*-vulnerable to be added to the vulnerable component. If the arrived node is indeed *red*-vulnerable, which happens with probability $\rho_{\mathbf{d},rs}$, it can activate other nodes via its remaining $d_{rs} - 1$ *red* single edges, d_{bs} *blue* single edges, n_{rt} *red* triangles, and n_{bt} *blue* triangles. Because the number of vulnerable nodes reached by each of its *red* single edges and triangles (resp. *blue* single edges and triangles) are generated in turn by $h_{rs}(x)$ and $h_{rt}(x)$ [resp. $h_{bs}(x)$ and $h_{bt}(x)$], respectively, we obtain the term $h_{rs}(x)^{d_{rs}-1} h_{rt}(x)^{n_{rt}} h_{bs}(x)^{d_{bs}} h_{bt}(x)^{n_{bt}}$ by the powers property of generating functions. Averaging over all possible colored degrees $\mathbf{d}$ gives the first term in (5). The second term with the factor x^0 accounts for the possibility that the arrived node is *not* red-vulnerable and thus is not included in the cluster. An analogous expression can be obtained for $h_{bs}(x)$ via similar arguments.

B. Influence propagation via *red* triangles

We now derive $h_{rt}(x)$, i.e., the generating function for the number of nodes influenced by following a *red* triangle selected at random; similar arguments hold for $h_{bt}(x)$. We consider the situation where nodes u, v, and w form a triangle and the top vertex u is *active*, while others are not. We are interested in computing the generating function for the total number of nodes that will be influenced by nodes v and w. We will compute the generating function $h_{rt}(x)$ by conditioning on the following events:

(i) If neither of nodes v and w are $\mathbb{R}$-vulnerable, then the number of nodes influenced will be zero. Node v has degree $\mathbf{d}$ with *normalized* probability $\frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle}$, in which case it is not $\mathbb{R}$-vulnerable with probability $1 - \rho_{\mathbf{d},rs}$. Similarly, the probability that node w has degree $\mathbf{d}'$ and not $\mathbb{R}$-vulnerable is $(\frac{n'_{rt} p_{\mathbf{d}'}}{\langle n_{rt} \rangle})(1 - \rho_{\mathbf{d}',rs})$. Summing over all possible cases, we obtain the first term in (7) with x^0 (meaning that zero nodes will be influenced by following the red triangle in this case).

(ii) Consider the case where only one of v and w is influenced, leading to a term with the factor x^1 in (7). Without loss of generality, consider the case where v is activated while w is not. If node v has degree $\mathbf{d}$, then it is $\mathbb{R}$-vulnerable with probability $\rho_{\mathbf{d},rs}$, and can influence other nodes in the usual manner. Then, the event that node w, with degree $\mathbf{d}'$, will not be activated despite having two active neighbors (nodes u and v) has probability $1 - \rho_{\mathbf{d}',rt}$. By symmetry and exchangeability of nodes v and w, an equivalent term will be obtained for the case where w is activated but v is not. Summing over all possibilities we obtain the second term in (7).

(iii) Finally, we consider the case where both v and w become *active* giving rise to term with factor x^2. There are two possible scenarios:

(a) Both of v and w are *activated* by u. The probability that v is *activated* by u is $\rho_{\mathbf{d},rs}$ as already discussed during the computation of the first term. By symmetry, the probability for w is the same as for v. Multiplying the two probabilities leads to the third term in (7).

(b) Only one of v and w is made *active* immediately by u while the other is not; e.g., say v is activated but not w. However, w also gets activated by the joint influence from u and v. With $\mathbf{d}$ and $\mathbf{d}'$ denoting the degree of v and w, this happens with probability $(\rho_{\mathbf{d},rs})(\rho_{\mathbf{d}',rt} - \rho_{\mathbf{d}',rs})$. Here, the second term accounts for the fact that w gets activated *only if* it has two (or, more) active neighbors. Summing over all possibilities as before, and multiplying by two for the case where v and w are replaced, we obtain the last term in (7).

C. Deriving the condition for global cascades

The discussion given in Secs. III A and III B leads to a set of recursive equations for $h_{rs}(x)$, $h_{rt}(x)$, $h_{bs}(x)$, and $h_{bt}(x)$. Recursions for $h_{rs}(x)$ and $h_{rt}(x)$ are given in (6) and (7), respectively; the expressions for $h_{bs}(x)$ and $h_{bt}(x)$ are very similar and omitted here for brevity. With these four recursive equations in place, it is possible to determine the generating function $H(x)$ of the *finite* number of nodes activated in the contagion process. Namely, for a given x, we shall find a *fixed point* of these recursive equations, and then use the resulting values of $h_{rs}(x)$, $h_{rt}(x)$, $h_{bs}(x)$, and $h_{bt}(x)$ in (3) to get $H(x)$:

$$h_{rs}(x) = x \sum_{\mathbf{d}} \frac{d_{rs} p_{\mathbf{d}}}{\langle d_{rs} \rangle} \rho_{\mathbf{d},rs} D(d_{rs} - 1, n_{rt}, d_{bs}, n_{bt}) + x^0 \sum_{\mathbf{d}} \frac{d_{rs} p_{\mathbf{d}}}{\langle d_{rs} \rangle}(1 - \rho_{\mathbf{d},rs}), \tag{6}$$

$$h_{rt}(x) = x^0 \sum_{\mathbf{d}} \sum_{\mathbf{d}'} \frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle}(1 - \rho_{\mathbf{d},rs}) \frac{n'_{rt} p_{\mathbf{d}'}}{\langle n'_{rt} \rangle}(1 - \rho_{\mathbf{d}',rs}) + 2x \sum_{\mathbf{d}} \sum_{\mathbf{d}'} \frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle} \rho_{\mathbf{d},rs} D(d_{rs}, n_{rt} - 1, d_{bs}, n_{bt}) \frac{n'_{rt} p_{\mathbf{d}'}}{\langle n'_{rt} \rangle}(1 - \rho_{\mathbf{d}',rt})$$

$$+ x^2 \sum_{\mathbf{d}} \sum_{\mathbf{d}'} \left[\frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle} \rho_{\mathbf{d},rs} D(d_{rs}, n_{rt} - 1, d_{bs}, n_{bt}) \right] \left[\frac{n'_{rt} p_{\mathbf{d}'}}{\langle n'_{rt} \rangle} \rho_{\mathbf{d}',rs} D(d'_{rs}, n'_{rt} - 1, d'_{bs}, n'_{bt}) \right]$$

$$+ 2x^2 \sum_{\mathbf{d}} \sum_{\mathbf{d}'} \left[\frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle} \rho_{\mathbf{d},rs} D(d_{rs}, n_{rt} - 1, d_{bs}, n_{bt}) \right] \left[\frac{n'_{rt} p_{\mathbf{d}'}}{\langle n'_{rt} \rangle}(\rho_{\mathbf{d}',rt} - \rho_{\mathbf{d}',rs}) D(d'_{rs}, n'_{rt} - 1, d'_{bs}, n'_{bt}) \right]. \tag{7}$$

By conservation of probability and the definition of generating functions, we know that $H(1) = 1$ only if final number of activated nodes is *finite* with probability one. In other words, *global* cascades that lead to a positive fraction

of influenced nodes are possible only if $H(1) < 1$. This prompts us to seek a fixed point of the recursive equations when $x = 1$. For notational convenience, we define $h_1 := h_{rs}(1)$, $h_2 := h_{rt}(1)$, $h_3 := h_{bs}(1)$, and $h_4 := h_{bt}(1)$. From (3),

this gives

$$H(1) = \sum_{\mathbf{d}} p_{\mathbf{d}} h_1^{d_{rs}} h_2^{n_{rt}} h_3^{d_{bs}} h_4^{n_{bt}}, \qquad (8)$$

while the recursions take the form

$$h_i = g_i(h_1, h_2, h_3, h_4), \quad i = 1,2,3,4. \qquad (9)$$

Here, the functions g_1, g_2 are easily obtained from (6) and (7), and similarly g_3 and g_4 can be obtained from the recursions for $h_{bs}(x)$, and $h_{bt}(x)$. To give an example, we have

$$g_3(h_1, h_2, h_3, h_4)$$
$$= \sum_{\mathbf{d}} \frac{d_{bs} p_{\mathbf{d}}}{\langle d_{bs} \rangle} \left(\rho_{\mathbf{d},bs} h_1^{d_{rs}} h_2^{n_{rt}} h_3^{d_{bs}-1} h_4^{n_{bt}} + 1 - \rho_{\mathbf{d},bs} \right).$$

It is clear that the recursions (9) have a trivial fixed point $h_1 = h_2 = h_3 = h_4 = 1$ which yields $H(1) = 1$, meaning that cascades will die out without reaching a positive fraction of the population with high probability. To check the stability of the trivial solution, we linearize the recursive equations (9) around $x = 1$, and compute the corresponding Jacobian matrix $\mathbf{J}$ via

$$\mathbf{J}(i,j) = \left. \frac{\partial g_i(h_1, h_2, h_3, h_4)}{\partial h_j} \right|_{h_1 = h_2 = h_3 = h_4 = 1} \qquad (10)$$

for each $i, j = 1,2,3,4$; the exact expression for the 4×4 matrix $\mathbf{J}$ is not given here in order to save space. Now, the trivial solution $h_1 = h_2 = h_3 = h_4 = 1$ is linearly stable if and only if the largest eigenvalue in absolute value of $\mathbf{J}$, denoted $\sigma(\mathbf{J})$, is less than one. Otherwise, if $\sigma(\mathbf{J}) > 1$, then there exists another fixed point for the recursion with $h_1, h_2, h_3, h_4 < 1$, leading to $H(1) < 1$. In that case, the probability deficit $1 - H(1) > 0$ gives the probability that the contagion process reaches infinitely many nodes, i.e., a *global* spreading event takes place. Collecting, we conclude that the condition of

global cascades is given by $\sigma(\mathbf{J}) > 1$, while the probability of global cascade equals $P_{trig} = 1 - H(1)$.

IV. EXPECTED CASCADE SIZE

Next, we are interested in computing the expected size of global cascades when they take place. Put differently, we will analyze the expected fraction of nodes that will eventually become active as we pick a node in the network uniformly at random and set it active. We follow the approach used in [29,52,62], which has been proven to be an effective way to analyze expected cascade size in networks.

First, consider the network $\mathbb{H}$ as a tree structure with an arbitrary node selected as the root. Then, label the levels of the tree from $\ell = 0$ at the bottom to $\ell = +\infty$ at the top of the tree. Similar to [29,52], we assume that nodes begin updating their states starting from the bottom of the tree and proceeding to the top. In other words, we assume that a node at level ℓ updates its state only after all nodes at the lower levels $0, 1, \ldots, \ell - 1$ finish updating. We define $q_{rs,\ell}$ as the probability that a node at level ℓ of a tree, which is connected to its unique parent by a *red* single edge, is active given that its parent at level $\ell + 1$ is inactive. Then, we consider a pair of nodes at level ℓ that together with their parent at level $\ell + 1$ form a red triangle. Given that the parent is inactive, we let $q_{rt1,\ell}$ (resp. $q_{rt2,\ell}$) denote the probability that only one (resp. both) of the two child nodes of this triangle is active.

We define $q_{bs,\ell}, q_{bt1,\ell}$, and $q_{bt2,\ell}$ for *blue* edges in the same manner.

According to our model, an *active* node is never deactivated, meaning that $q_{rs,\ell}, q_{rt1,\ell}, q_{rt2,\ell}, q_{bs,\ell}, q_{bt1,\ell}, q_{bt2,\ell}$ are all nondecreasing. Therefore, they will converge to $q_{rs,\infty}, q_{rt1,\infty}, q_{rt2,\infty}, q_{bs,\infty}, q_{bt1,\infty}, q_{bt2,\infty}$. Then, the expected cascade size (i.e., the fraction of active individuals) S is given by the probability that the arbitrary selected node at the top of the tree becomes *active*. In order to computer S, we first derive recursive relations for $q_{rs,\ell}, q_{rt2,\ell}, q_{rt2,\ell}, q_{bs,\ell}, q_{bt2,\ell}, q_{bt2,\ell}$. We have

$$q_{rs,\ell+1} = \sum_{\mathbf{d}} \frac{d_{rs} p_{\mathbf{d}}}{\langle d_{rs} \rangle} \sum_{i=0}^{d_{rs}-1} \sum_{j=0}^{n_{rt}} \sum_{x=0}^{j} \sum_{m=0}^{d_{bs}} \sum_{n=0}^{n_{bt}} \sum_{y=0}^{n} Q_\ell[(i,j,x,m,y,n),(d_{rs}-1,n_{rt},d_{bs},n_{bt})] F[(i,x,j-x,m,y,n-y),\mathbf{d}],$$

$$q_{rt1,\ell+1} = 2 \sum_{\mathbf{d},\mathbf{d}'} \frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle} \frac{n'_{rt} p_{\mathbf{d}'}}{\langle n'_{rt} \rangle} \sum_{i,i'=0}^{d_{rs},d'_{rs}} \sum_{j,j'=0}^{n_{rt}-1,n'_{rt}-1} \sum_{x,x'=0}^{j,j'} \sum_{m,m'=0}^{d_{bs},d'_{bs}} \sum_{n,n'=0}^{n_{bt},n'_{bt}} \sum_{y,y'=0}^{n,n'} \{ Q_\ell[(i,j,x,m,y,n),(d_{rs},n_{rt}-1,d_{bs},n_{bt})]$$
$$\times Q_\ell[(i',j',x',m',y',n'),(d'_{rs},n'_{rt}-1,d'_{bs},n'_{bt})] F[(i,x,j-x,m,y,n-y),\mathbf{d}]$$
$$\times (1 - F[(i',x'+1,j'-x',m',y',n'-y'),\mathbf{d}'])\},$$

$$q_{rt2,\ell+1} = \sum_{\mathbf{d},\mathbf{d}'} \frac{n_{rt} p_{\mathbf{d}}}{\langle n_{rt} \rangle} \frac{n'_{rt} p_{\mathbf{d}'}}{\langle n'_{rt} \rangle} \sum_{i,i'=0}^{d_{rs},d'_{rs}} \sum_{j,j'=0}^{n_{rt}-1,n'_{rt}-1} \sum_{x,x'=0}^{j,j'} \sum_{m,m'=0}^{d_{bs},d'_{bs}} \sum_{n,n'=0}^{n_{bt},n'_{bt}} \sum_{y,y'=0}^{n,n'} \{ Q_\ell[(i,j,x,m,y,n),(d_{rs},n_{rt}-1,d_{bs},n_{bt})]$$
$$\times Q_\ell[(i',j',x',m',y',n'),(d'_{rs},n'_{rt}-1,d'_{bs},n'_{bt})](F[(i,x,j-x,m,y,n-y),\mathbf{d}] F[(i',x'+1,j'-x',m',y',n'-y'),\mathbf{d}']$$
$$+ (F[(i,x+1,j-x,m,y,n-y),\mathbf{d}] - F[(i,x,j-x,m,y,n-y),\mathbf{d}]) F[(i',x',j'-x',m',y',n'-y'),\mathbf{d}'])\},$$

where we define

$$Q_\ell[(i,j,x,m,n,y),(d_1,d_2,d_3,d_4)] = \binom{d_1}{i} q_{rs,\ell}^i (1-q_{rs,\ell})^{d_1-i} \binom{d_2}{j} \binom{j}{x} q_{rt1,\ell}^x q_{rt2,\ell}^{j-x} (1-q_{rt1,\ell}-q_{rt2,\ell})^{d_2-j} \binom{d_3}{m} q_{bs,\ell}^m$$

$$\times (1-q_{bs,\ell})^{d_3-m} \binom{d_4}{n} \binom{n}{y} q_{bt1,\ell}^y q_{bt2,\ell}^{n-y} (1-q_{bt1,\ell}-q_{bt2,\ell})^{d_4-n}. \qquad (11)$$

In words, $Q_\ell[(i,j,x,m,n,y),(d_1,d_2,d_3,d_4)]$ gives the probability that a node at level ℓ with colored degree $(d_1,2d_2,d_3,2d_4)$ has

(i) i (resp. $d_1 - i$) of the d_1 neighbors connected through red single edges as active (resp. inactive). Similarly, m (resp. $d_3 - m$) of the d_3 neighbors connected through blue single edges as active (resp. inactive);

(ii) of the d_2 red triangles it participates in, x has one active node, $j - x$ has two active nodes, and $d_2 - j$ has no active node. Similarly, of the d_4 blue triangles it participates in, y has one active node, $n - y$ has two active nodes, and $d_4 - n$ has no active node

Hence, multiplying $Q_\ell[(i,j,x,m,n,y),(d_1,d_2,d_3,d_4)]$ with $F[(i,x,j - x,m,y,n - y),\mathbf{d}]$ and summing over all possibilities for $\mathbf{d}$ and i,j,x,m,n,y gives the probability that the node under consideration turns active. This confirms the first expression above. Second and third terms consider simultaneously a pair of nodes that are part of a red triangle (where the top, i.e., parent, vertex is inactive). Therefore, we first condition on the degrees of these two nodes being $\mathbf{d}$ and $\mathbf{d}'$, respectively, and consider all possibilities concerning the states (active versus inactive) of these neighbors. Then, for $q_{rt1,\ell+1}$, we realize by symmetry that the desired expression is two times the probability that the node with degree $\mathbf{d}$ turns active, and despite having one extra active neighbor, the node with degree $\mathbf{d}'$ does not turn active. The fact that the first node turns active is incorporated in the expression $(1 - F[(i',x' + 1,j' - x',m',y',n' - y'),\mathbf{d}'])$ by the term $x' + 1$. For $q_{rt2,\ell+1}$, we proceed similarly and realize that for both nodes to turn active there are two possibilities. The node with degree $\mathbf{d}$ either turns active regardless of the state of the node with degree $\mathbf{d}'$ {in which case the node with degree $\mathbf{d}'$ will turn active with probability $F[(i',x' + 1,j' - x',m',y',n' - y'),\mathbf{d}']$}, or it turns active only after the node with degree $\mathbf{d}'$ does.

With the above recursion in place, we compute the final cascade size via

$$S = \sum_{\mathbf{d}} p_{\mathbf{d}} \sum_{i=0}^{d_{rs}} \sum_{j=0}^{n_{rt}} \sum_{x=0}^{j} \sum_{m=0}^{d_{bs}} \sum_{n=0}^{n_{bt}} \sum_{y=0}^{n} Q_\infty[(i,j,x,m,n,y),$$

$$(d_{rs},n_{rt},d_{bs},n_{bt})] F[(i,x,j - x,m,y,n - y),\mathbf{d}]. \quad (12)$$

Namely, we first solve for the values of $q_{rs,\infty}$, $q_{rt1,\infty}$, $q_{rt2,\infty}$, $q_{bs,\infty}$, $q_{bt1,\infty}$, $q_{bt2,\infty}$ using the recursive equations, and then substitute them into (12) to obtain the expected size of global cascades.

V. NUMERICAL RESULTS

A. Networks with doubly Poisson distributions

In our first simulation study, we use doubly Poisson distribution for the number of *single edges* and *triangles* in both networks. Namely, we set
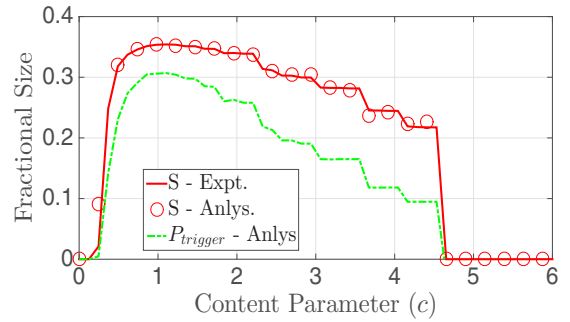
$$p_{st}^r = e^{-\lambda_{r,1}} \frac{(\lambda_{r,1})^s}{s!} e^{-\lambda_{r,2}} \frac{(\lambda_{r,2})^t}{t!}, \quad s,t = 0,1,\ldots,$$

$$p_{st}^b = e^{-\lambda_{b,1}} \frac{(\lambda_{b,1})^s}{s!} e^{-\lambda_{b,2}} \frac{(\lambda_{b,2})^t}{t!}, \quad s,t = 0,1,\ldots,$$

where s and t are the number of *single edges* and *triangles* in the corresponding networks, respectively. Thus, $\lambda_{r,1}$ and $\lambda_{r,2}$
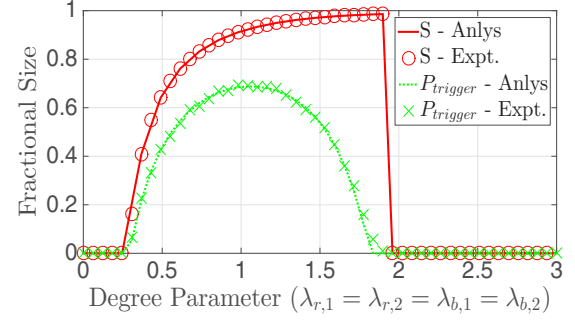
(a)

(b)

FIG. 3. Simulations for doubly Poisson degree distributions. In (a), we set the content parameter $c = 0.25$, the threshold as $\tau = 0.18$, and $\alpha = 0.5$, and vary the degree parameters. In (b), we fix $\tau = 0.18$, $\lambda_{r,1} = \lambda_{r,2} = \lambda_{b,1} = \lambda_{b,2} = 0.3$, and $\alpha = 0.5$ while varying content parameter c.

(resp. $\lambda_{b,1}$ and $\lambda_{b,2}$) denote the mean number of single edges and triangles, respectively, in $\mathbb{R}$ (resp. in $\mathbb{B}$).

We consider $n = 1 \times 10^5$ nodes in the population and $\alpha = 0.5$ for the size of network $\mathbb{R}$. We let $\tau = 0.18$ and $c = 0.25$ for the threshold and content parameters, respectively. The results are shown in Fig. 3 where the curves stand for the theoretical results of probability P_{trig} and expected size S of cascades (obtained from our discussion in Secs. III and IV), as a function of $\lambda_{r,1} = \lambda_{r,2} = \lambda_{b,1} = \lambda_{b,2}$. The markers stand for the empirical results for the same quantities, and are obtained by averaging over 5000 independent experiments. We see a very good agreement between the analytical and experimental results confirming the validity of our analysis. The slight discrepancy observed in P_{trig} is due to the limited number of experiments, and can be mitigated by increasing the number of realizations.

Next, we change our experimental setup to demonstrate the effect of content parameter on the probability and size of cascades. To that end, we fix all network parameters and observe the quantities of interest as the content parameter c varies. In particular, we set $\lambda_{r,1} = \lambda_{r,2} = \lambda_{b,1} = \lambda_{b,2} = 0.3$ and $\tau = 0.18$. We see that the probability and expected size of global cascades vary greatly as c changes. This can be taken as an indication that our model can capture the real-world phenomenon that over the same population certain contents can become widespread while others die out quickly. In the setting used here, we see that global cascades take place when the content parameter c is not too small or large. The

TABLE I. Parameters of the doubly Poisson distribution. This choice ensures that the mean and variance of the total degree distribution (single plus triangle edges) in $\mathbb{B}$ are independent of η, while its clustering varies greatly as η varies in $(0,4)$. In Fig. 3, we set $\lambda = 0.5$.

	Network $\mathbb{R}$	Network $\mathbb{B}$
Distribution of *single* edges	Poi(2λ)	$2\,\text{Poi}\left(\frac{4-\eta}{2}\lambda\right)$
Distribution of *triangles*	Poi(λ)	Poi$\left(\frac{\eta}{2}\lambda\right)$

reason is that with a too small or large content parameter, the connectivity of the conjoined network is dominated by only one of the two networks. So, if neither of them has enough connectivity to trigger a global cascade by their own, then there will be no global cascades in the conjoined network. When the c is neither too large nor too small (e.g., close to unity), both networks will contribute to the connectivity together and it becomes possible to trigger a global cascade. For other values of λ_r and λ_b, a completely different situation might occur, e.g., with very small or very large c promoting cascades; e.g., see [29, Fig. 2] for a few such examples.

B. How does clustering affect the cascade size?

Our next goal is to reveal the impact of clustering on the influence propagation process in the models considered here. To do so, we should be able to vary the level of clustering while keeping the first and second moments of the total degree distribution fixed, as they are known to affect the contagion behavior significantly [29]. In order to satisfy this constraint, we use Poisson distributions for the number of single edges and triangle in two networks with parameters given in Table I.

With the setting given in Table I, the clustering coefficient in $\mathbb{R}$ will be fixed for any $\eta \in [0,4]$, while the clustering of $\mathbb{B}$ varies between the two extremes: (i) when $\eta = 4$, $\mathbb{B}$ will have no *single* edges and consist only of *triangles* resulting with a clustering coefficient close to one; and (ii) with $\eta = 0$, there will be no triangles in $\mathbb{B}$ and hence its clustering coefficient will be close to zero. Collecting, we see that the clustering coefficient of $\mathbb{B}$ and thus of $\mathbb{H}$ increases with increasing η in this setting.

With these in mind, we first demonstrate the impact of clustering on the probability of triggering a global cascade.

Figure 4(a) shows the probability of triggering a global cascade as a function of λ for three different η values. The resulting clustering coefficients are plotted in Fig. 4(b) where we clearly see that clustering increases with increasing η. The main observation from Fig. 4(a) is that increasing the clustering (i.e., increasing η) shifts the interval of λ for which global cascades are possible to the *right*. This leads to a double-faceted conclusion that clustering decreases the probability of global cascades when average degrees are small, whereas after a certain value of average degree, clustering increases the probability of cascades.

The double-faceted impact of clustering on cascade probability can be explained as follows. It is known [18,28] that threshold models of complex contagion exhibit two phase transitions as the average degree increases, a second-order transition at *low* degrees that marks the formation of a giant vulnerable component and a first-order transition at *high* degrees due to increased *local* stability of nodes, namely, due to the increased difficulty of activating high-degree nodes. Given that clustering is known to decrease the size of giant component [31], we expect that it will be more difficult for a clustered network to contain a giant vulnerable cluster. This is why the lower phase transition in complex contagions appear later (i.e., at larger degrees) as clustering increases. On the other hand, the cycles of size three (i.e., triangles) that are common in clustered networks can help trigger cascades when average degree is higher. For instance, in a treelike network a single active node can only activate its vulnerable connections. However, in a triangle, an active node may first activate one of its vulnerable connections, making it possible for the third node to be activated (which now has two active neighbors) even if it is *not* vulnerable. This is what pushes the second phase transition to higher degrees.

Next, we explore the impact of clustering (in the setting considered here) on the expected cascade size in Fig. 3(c). Here again, we see the double-faceted impact of clustering with small average degrees favoring low clustering, while high degrees favoring high clustering in terms of having a larger cascade size. In fact, we see the existence of a critical average degree [around $\lambda = 0.6$ in Fig. 3(c)] such that when λ is smaller (resp. larger) than the critical value, expected cascade size decreases (resp. increases) with increasing clustering. This extends the observation that Hackett *et al.* [52] made for single-layer networks to multilayer networks.
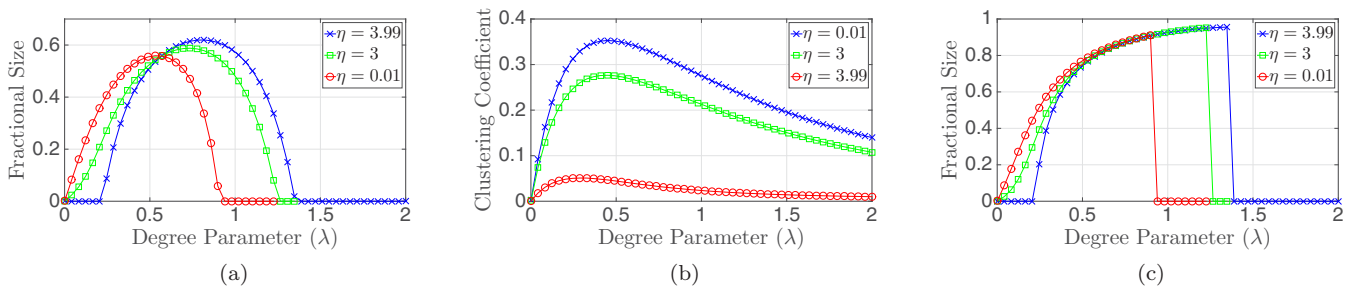

(a)

(b)

(c)

FIG. 4. Illustration of the effect of clustering coefficient on the expected cascade and probability of global cascades. We fix $\tau = 0.18$, $c = 0.25$, and $\alpha = 0.5$, then vary the degree parameter λ defined in Table I. We see (a) the probability to trigger a global cascade; (b) the global clustering coefficient described in Sec. II A; and (c) the expected cascade size.
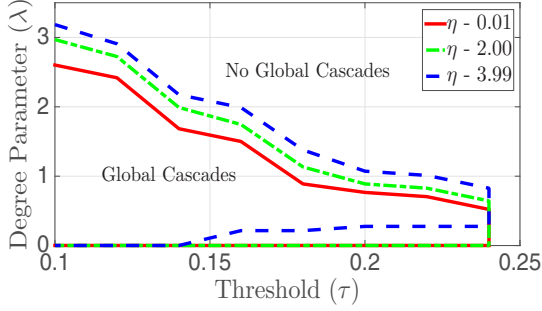
FIG. 5. We show the *cascade regions* in the degree parameter threshold plane when $\alpha = 0.5$, $\tau = 0.18$, and both networks follow doubly Poisson distributions as described in Table I. Clustering increases as η increases.

Finally, we consider the impact of clustering on the average degree-cascade threshold plane. For each parameter pair (λ, τ), the curves in Fig. 5 separate the region where global cascade can take place (areas inside the boundaries) from the region where they cannot (areas outside the boundaries). Once again, we confirm that increasing the clustering coefficient shifts the interval where cascades are possible *up* (i.e., to higher degrees) for any threshold τ.

VI. COMPARISON BETWEEN MONOPLEX AND MULTIPLEX NETWORKS

In what follows, we will compare the dynamics of complex contagions over a monoplex network with that over a multiplex network. Of particular interest will be to find out whether the projection of a multiplex network into a monoplex network leads to any significant differences in the dynamics that would warrant the separate analyses of multiplex networks as conducted here.

To identify the factors affecting complex contagions, we consider two different degree distributions to generate the networks. In Sec. VI A, we use a setting similar to previous sections, with the resulting networks having almost no degree-degree correlations, e.g., assortativity defined as the Pearson correlation coefficient between the degrees of pairs of linked nodes [63]. In Secs. VI B and VI C, we use a different setting that leads to (tunable) assortativity for multiplex networks. In order to keep the focus on the comparison between monoplex and multiplex networks, we shall consider only nonclustered networks in the following discussion.
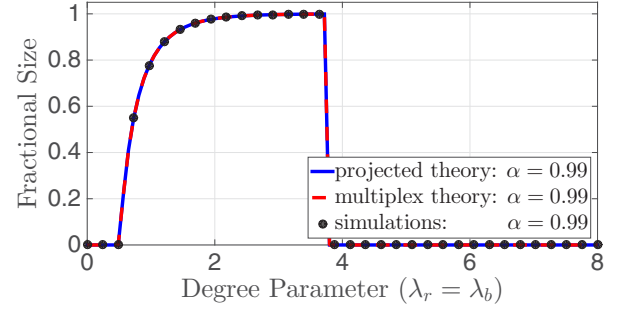
A. Multiplex networks with limited assortativity

First we consider the limited assortativity case and use the following degree distribution to assign *blue* and *red* stubs to each node:
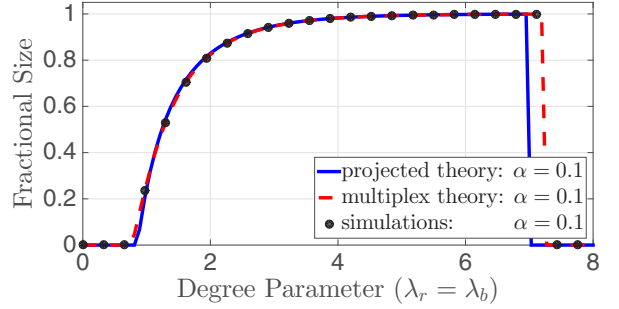
$$p_k^b = e^{-\lambda_b} \frac{\lambda_b^k}{k!}, \quad k = 0, \dots$$

$$p_k^r = \alpha e^{-\lambda_r} \frac{\lambda_r^k}{k!} + (1 - \alpha)\delta_{k,0}, \quad k = 0, \dots \quad (13)$$

where δ denotes the Kronecker delta. In other words, each node receives $\text{Poi}(\lambda_b)$ blue edges, and an α fraction of nodes receive an additional $\text{Poi}(\lambda_r)$ edges of color red. A multiplex network

(a)

(b)

FIG. 6. Comparison between monoplex networks and multiplex networks with limited assortativity. In (a) and (b), we fix the threshold $\tau = 0.15$, the content parameter $c = 1$, then vary the degree parameters in (13). For the networks obtained by projected theory and the networks in multiplex theory with $\alpha = 0.99$, assortativity is negligible. However, when $\alpha = 0.1$, the assortativity coefficient of the networks in the multiplex theory become significant, e.g., it can be up to 0.21.

is generated using the *colored* configuration model [64,65] where stubs that are of the *same* color are matched randomly. The monoplex projected theory *ignores* the color of the edges and matches all stubs randomly with each other. An important question is whether we lose any significant information about contagion dynamics when the monoplex projected theory is used instead of the multiplex theory developed here and in [29]. For convenience, we set $\lambda_b = \lambda_r$ and use $c = 1$ as the content parameter.

In Fig. 6(a), we set $\alpha = 0.99$. We see nearly no difference between the theoretical cascade sizes obtained from monoplex and multiplex theories, and they both match the simulation results well. However, when α is reduced to 0.1 in Fig. 6(b), we clearly see a difference between the two theories and only multiplex theory matches the simulation results. This shows that even in the simplest case where both link types have the same influence factor (i.e., $c = 1$), monoplex theory may be unable to capture certain properties of cascade dynamics, reinforcing the need for studying cascades using the multiplex theory.

We now explain why the two cases, $\alpha = 0.1$ and 0.99, lead to different conclusions regarding the accuracy of the monoplex theory in capturing contagion dynamics over multiplex networks. One of the key differences between the two cases is the resulting assortativity. When $\alpha = 0.1$, only 10% of the nodes have red stubs, each of which can only be connected with other red stubs in the multiplex network case. Put differently,

in this setting a small fraction of the population will have statistically *higher* degrees than the rest, and the additional links they have can only connect nodes with high degrees together. This leads to a positive correlation (i.e., assortativity) between the degrees of pairs of connected nodes. However, in the monoplex projection, the additional edges can be used to connect any two nodes, resulting with very little to no assortativity in the network. Obviously, when α is close to one, almost every node will have the additional edges and the above phenomenon will not be observed. Our simulation results confirm this intuition as we see that assortativity is negligible ($\sim 10^{-4}$) in both monoplex and multiplex cases when $\alpha = 0.99$, while with $\alpha = 0.1$, assortativity varies (as $\lambda_r = \lambda_b$ increases) from 0.05 to 0.2 in the multiplex case while still being negligible in the monoplex case.

The impact of assortativity on the comparison between monoplex and multiplex theories is investigated further in the forthcoming discussion.
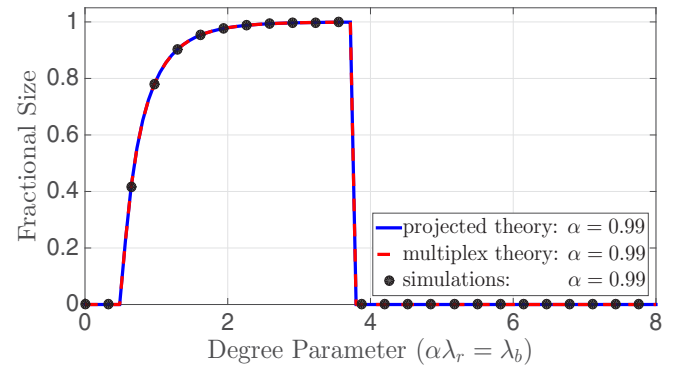
B. Multiplex networks with assortativity

In this section, we change the setting slightly to generate multiplex networks with high assortativity. To that end, we use the degree distributions given at (13), but instead of setting $\lambda_r = \lambda_b$, we enforce
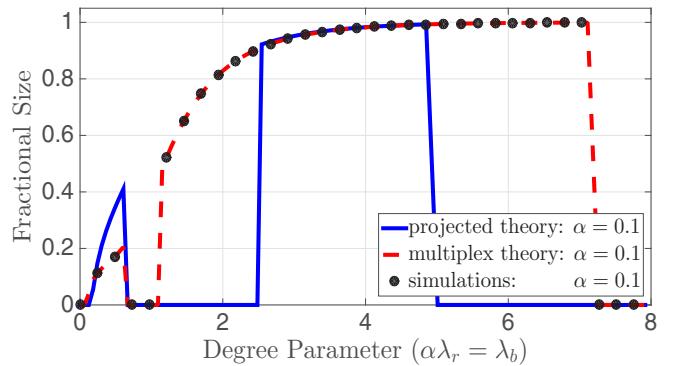
$$\alpha \lambda_r = \lambda_b \tag{14}$$

for any $\alpha \in (0,1)$. This setting allows us to tune assortativity without changing the mean degree in the network. In particular, assortativity will increase as α decreases (by virtue of a small fraction of nodes forming a highly connected cluster) [31]. In addition, this setting allows us to compare the contagion dynamics in multilayer networks when the upper layer is (i) small but densely connected (small α) versus (ii) large but loosely connected (large α); see [31] for relevant results for bond percolation processes.

Using the above degree distributions, we generate monoplex and multiplex networks as in Sec. VI A and analyze the complex contagion process. In Fig. 7(a), we see that when $\alpha = 0.99$, which leads to very limited assortativity, the difference between monoplex and multiplex networks is negligible. This is in parallel with what we observed in Sec. VI A. However, decreasing α to 0.1 leads to two interesting observations in Fig. 7(b). First, instead of the commonly reported two phase transitions [18,28], we observe four phase transitions in the cascade size as $\alpha \lambda_r = \lambda_b$ increases. Second, we see a significant difference between the monoplex projected theory and multiplex theory, with only the multiplex theory matching the simulations well. Once again, this shows that monoplex theory is unable to capture the cascade dynamics under certain settings.

The emergence of four phase transitions in Fig. 7(b), which to the best of our knowledge was *not* reported before, can be explained as follows.[1] When $\alpha = 0.1$, only 10% of the nodes

---

[1]From a practical point of view, observing four phase transitions rather than two signals a more chaotic system behavior (in terms of contagion dynamics) with respect to changes in the degree parameter λ_b. In turn, this would make the *prediction* of cascade region more

(a)

(b)

FIG. 7. Comparison between monoplex networks and multiplex networks with assortativity. Similar with the observation in Fig. 6, networks in the projected theory and in the multiplex theory with $\alpha = 0.99$ have negligible assortativity coefficients. However, for the networks of multiplex theory with $\alpha = 0.1$, assortativity coefficient ranges from 0.19 to 0.79. In general, assortativity increases with increasing λ_r and λ_b in the multiplex theory.

have *red* edges, but the the mean number of red edges for those nodes equals $10\lambda_b$ [see (14)]. Therefore, the first couple of phase transitions taking place at very small λ_b values can be attributed mainly to red edges. First, λ_b becomes large enough (e.g., gets to around 0.1) that the subnetwork induced only on the red edges contains a giant vulnerable cluster, giving rise to global cascades; note that at this point λ_b is so small that blue edges do not create enough local stability to prevent cascades from happening. However, after λ_b reaches a certain level (around 0.65), the subgraph on red edges, having average degree of $10\lambda_b$, reaches the second phase transition point where cascades stop due to the increased local connectivity of nodes. These first two transitions being second and first order, respectively, also confirms that they are primarily due to the red edges.

As λ_b increases further, we observe an interval where there are no global cascades due to either colors of edges; nodes with red *and* blue edges are highly stubborn while nodes with only blue edges are not connected enough to trigger a cascade. This interval is then followed by a region where λ_b is large enough

---

difficult in the cases where system parameters are not known exactly but estimated, e.g., social network applications.
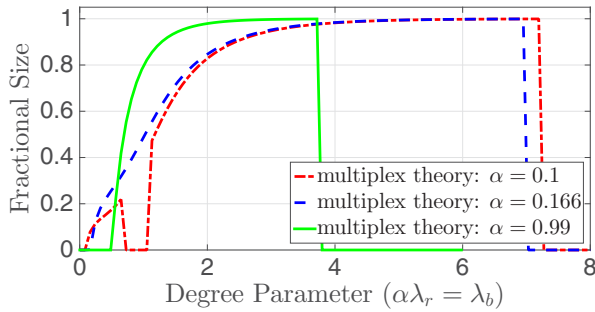
FIG. 8. Demonstration of multiple phase transitions.

that the subgraph on blue edges has a giant vulnerable cluster. However, the emergence of a second-order transition in the whole network is prevented here due to some of these nodes turning stubborn as a result of their red edges. Eventually, however, λ_b becomes large enough that even with occasional stubborn nodes present, a giant vulnerable cluster emerges. This point is reached much later in monoplex networks as compared to the multiplex networks. This is because in the former case stubborn nodes (with red edges) are equally likely to be connected with any other node, while in the latter case they are mostly connected with each other; thus, in the latter case they are less likely to inhibit the emergence of a giant vulnerable cluster on blue edges.

Finally, the system goes through a fourth transition when λ_b becomes large enough that even nodes with only blue edges become highly connected and hence stubborn. We see that this final transition point is reached much later in multiplex networks than monoplex networks, meaning that cascades take place over a broader range of λ_b values in the former case. Again, this can be attributed to the high assortativity seen in multiplex networks that leads to *extremely* stubborn nodes (that have both blue and red edges) being isolated from those that are *mildly* stubborn (that have only blue edges). On the other hand, in monoplex networks, every node is able to connect with the extremely stubborn nodes, and thus the critical value of λ_b at which cascades become impossible due to high local stability is reached much earlier than that in multiplex networks.

C. Two vs four phase transitions

In Sec. VI B, we have observed the possibility of having more than two phase transitions in the cascade size. As discussed there, multiple phase transitions occur mainly due to the setting (14) that, with small α, ensures a small fraction of nodes having significantly higher connectivity than the rest, while also being mostly connected with each other. Since the existence of more than two transitions has not been reported in previous studies, we are interested in exploring it further. In particular, we now investigate the impact of α on the number of phase transitions as well as transition points. Of particular interest will be to find the critical α value that separates the cases where four phase transitions occur from those with only two transitions, e.g., the α value for which the two cascade regions overlap. For simplicity, we only consider multiplex networks in this section.

Figure 8 shows the expected size of global cascades under (13) and (14) for three different values of α. We see that

global cascades take place over a single interval of $\alpha\lambda_r = \lambda_b$ when α is *large* (e.g., $\alpha = 0.99$) while over two disjoint intervals when α is *small* (e.g., $\alpha = 0.1$). When α is somewhere in-between (e.g., case $\alpha = 0.166$) it is possible to have the cascade intervals partially overlap. In such cases, we only see a single interval where global cascades take place. However, an additional transition point appears, manifested by a shift of slope in cascade size, marking possibly the overlapping point of (what would be) the two cascade intervals.

Figure 8 allows us to comment also on the impact of the size and density of the online social network layer in facilitating influence propagation. With (14) in effect, a small α corresponds to a social network with few but densely connected individuals, while large α corresponds to a social network with many subscribers, each with few connections on average. In all cases, the total number of edges in the social network is fixed by virtue of (14). We see from Fig. 8 that the comparison between the three cases leads to a multifaceted picture as the mean number of links $\alpha\lambda_r$ varies. For instance, the large but loosely connected case of $\alpha = 0.99$ leads to the largest expected global cascade size over a certain interval, but it has the smallest *cascade interval* among all three. The intermediate case of $\alpha = 0.166$ seems like a stretched version (over the x axis) of the case with $\alpha = 0.99$. In particular, this case leads to the largest interval where global cascades are possible, although the expected cascade size is smaller than that obtained with $\alpha = 0.99$ (and also with $\alpha = 0.1$) over certain intervals. Finally, the case of a small but densely connected extra layer (i.e., with $\alpha = 0.1$), falls right under the case with $\alpha = 0.166$ for most values of $\alpha\lambda_r$, although it gives the largest size of all three in small intervals where $\alpha\lambda_r$ is very small or very large.

VII. CONCLUSION AND FUTURE WORK

We studied the diffusion of influence in a class of clustered multiplex networks. We solved analytically for the condition, probability, and expected size of *global* cascades, and confirmed our results via extensive computer simulations. One of our key findings is to show how clustering affects the probability and expected size of global cascades. We also compared several interesting properties of complex contagions on a multiplex network and its monoplex projection. We demonstrate that ignoring link types and aggregating network layers may lead to inaccurate conclusions about contagion dynamics, particularly when assortativity is high. Finally, we show that linear threshold models do not necessarily exhibit two phase transitions as previously reported. Depending on assortativity, we show that both in monoplex and multiplex cases (with two link types) it is possible to observe four phase transitions.

Our analysis and modeling framework subsumes some previous studies. For instance, by setting $h_{rt}(x) = h_{bt}(x) = 1$ in the recursive relations, we ensure that $\mathbb{R}$ and $\mathbb{B}$ are nonclustered random networks. So, our analysis corresponds to complex contagions in nonclustered networks, which was studied in [29]. Similarly, if we let $h_{rs}(x) = h_{rt}(x) = 1$ in the recursions and set the content parameter $c = 1$, then our analysis corresponds to *complex* contagions in *clustered monoplex* networks, which was studied in [52].

Future work may consider in more details the impact of assortativity or other topological features on the cascade

dynamics. It would also be interesting to compare multiplex networks and their monoplex projections in terms of other dynamical processes, e.g., site percolation, transport processes, etc. Another interesting direction would be to consider networks that exhibit clustering not only through triangles, but also through larger cliques [66]. Finally, it would be interesting to study multiplex threshold dynamics that implement *nonlinear* rules as well, e.g., see [32,67].

[1] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature (London) **464**, 1025 (2010).

[2] A. Vespignani, Nature (London) **464**, 984 (2010).

[3] O. Yağan, D. Qian, J. Zhang, and D. Cochran, IEEE Trans. Parallel Distrib. Syst. **23**, 1708 (2012).

[4] C. D. Brummitt, R. M. D'Souza, and E. Leicht, Proc. Natl. Acad. Sci. USA **109**, E680 (2012).

[5] F. Radicchi and A. Arenas, Nat. Phys. **9**, 717 (2013).

[6] S. D. Reis, Y. Hu, A. Babino, J. S. Andrade Jr, S. Canals, M. Sigman, and H. A. Makse, Nat. Phys. **10**, 762 (2014).

[7] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).

[8] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).

[9] R. Cohen, D. Ben-Avraham, and S. Havlin, Phys. Rev. E **66**, 036113 (2002).

[10] E. Leicht and R. M. D'Souza, arXiv:0907.0894.

[11] J. Nagler, A. Levina, and M. Timme, Nat. Phys. **7**, 265 (2011).

[12] J. D. Murray, *Mathematical Biology*, 3rd ed. (Springer, New York, 2002).

[13] S.-W. Son, P. Grassberger, and M. Paczuski, Phys. Rev. Lett. **107**, 195702 (2011).

[14] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. **109**, 248701 (2012).

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

[16] R. M. Anderson and R. M. May, *Infectious Diseases of Humans* (Oxford University Press, Oxford, UK, 1991).

[17] F. Sahneh, C. Scoglio, and P. Van Mieghem, IEEE/ACM Trans. Networking **21**, 1609 (2013).

[18] O. Yağan, D. Qian, J. Zhang, and D. Cochran, IEEE J. Sel. Areas Commun. **31**, 1038 (2013).

[19] L. D. Valdez, P. A. Macri, H. E. Stanley, and L. A. Braunstein, Phys. Rev. E **88**, 050803 (2013).

[20] K. Zhao and G. Bianconi, J. Stat. Mech.: Theory Exper. (2013) P05005.

[21] D. Cellai, E. López, J. Zhou, J. P. Gleeson, and G. Bianconi, Phys. Rev. E **88**, 052811 (2013).

[22] N. Azimi-Tafreshi, J. Gómez-Gardeñes, and S. N. Dorogovtsev, Phys. Rev. E **90**, 032816 (2014).

[23] G. Bianconi and S. N. Dorogovtsev, Phys. Rev. E **89**, 062814 (2014).

[24] F. Radicchi, Nat. Phys. **11**, 597 (2015).

[25] A. A. Saberi, Phys. Rep. **578**, 1 (2015).

[26] T. W. Valente, Social Networks **18**, 69 (1996).

[27] J. P. Gleeson, Phys. Rev. E **77**, 046117 (2008).

[28] D. J. Watts, Proc. Natl. Acad. Sci. USA **99**, 5766 (2002).

[29] O. Yağan and V. Gligor, Phys. Rev. E **86**, 036103 (2012).

[30] A. Hackett, D. Cellai, S. Gómez, A. Arenas, and J. P. Gleeson, Phys. Rev. X **6**, 021002 (2016).

[31] Y. Zhuang and O. Yağan, IEEE Trans. Network Sci. Eng. **3**, 211 (2016).

[32] C. D. Brummitt, K.-M. Lee, and K.-I. Goh, Phys. Rev. E **85**, 045102 (2012).

[33] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, J. Complex Networks **2**, 203 (2014).

[34] M. De Domenico, A. Solé-Ribalta, S. Gómez, and A. Arenas, Proc. Natl. Acad. Sci. USA **111**, 8351 (2014).

[35] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, Phys. Rev. X **3**, 041022 (2013).

[36] B. Min, S. D. Yi, K.-M. Lee, and K.-I. Goh, Phys. Rev. E **89**, 042811 (2014).

[37] M. Á. Serrano, L. Buzna, and M. Boguñá, New J. Phys. **17**, 053033 (2015).

[38] C. Granell, S. Gómez, and A. Arenas, Phys. Rev. Lett. **111**, 128701 (2013).

[39] G. J. Baxter, S. N. Dorogovtsev, J. F. F. Mendes, and D. Cellai, Phys. Rev. E **89**, 042801 (2014).

[40] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, Nat. Phys. **12**, 901 (2016).

[41] H. Wu, A. Arenas, and S. Gómez, arXiv:1606.01688.

[42] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, Phys. Rep. **544**, 1 (2014).

[43] D. Cellai and G. Bianconi, Phys. Rev. E **93**, 032302 (2016).

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

[45] M. E. Newman, SIAM Rev. **45**, 167 (2003).

[46] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[47] M. A. Serrano and M. Boguñá, Phys. Rev. E **74**, 056114 (2006).

[48] M. E. J. Newman, Phys. Rev. Lett. **103**, 058701 (2009).

[49] J. P. Gleeson, S. Melnik, and A. Hackett, Phys. Rev. E **81**, 066114 (2010).

[50] J. Badham and R. Stocker, Theor. Pop. Biol. **77**, 71 (2010).

[51] A. Hackett, J. P. Gleeson, and S. Melnik, Int. J. Complex Syst. Sci. **1**, 25 (2011).

[52] A. Hackett, S. Melnik, and J. P. Gleeson, Phys. Rev. E **83**, 056107 (2011).

[53] X. Huang, S. Shao, H. Wang, S. V. Buldyrev, H. E. Stanley, and S. Havlin, Europhys. Lett. **101**, 18002 (2013).

[54] P. Colomer-de-Simón and M. Boguñá, Phys. Rev. X **4**, 041020 (2014).

[55] E. Cozzo, M. Kivelä, M. De Domenico, A. Solé-Ribalta, A. Arenas, S. Gómez, M. A. Porter, and Y. Moreno, New J. Phys. **17**, 073029 (2015).

[56] J. C. Miller, Phys. Rev. E **80**, 020901 (2009).

[57] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[58] S. Tang, J. Yuan, X. Mao, X.-Y. Li, W. Chen, and G. Dai, in *Proceedings of IEEE INFOCOM 2011* (IEEE, Piscataway, NJ, 2011), pp. 2291–2299.

[59] P. S. Dodds and D. J. Watts, Phys. Rev. Lett. **92**, 218701 (2004).

[60] K. B. Athreya and P. E. Ney, *Branching Processes*, Vol. 196 (Springer, Berlin, 2012).

[61] H. S. Wilf, *Generating Functionology* (Elsevier, Amsterdam, 2013).

[62] J. P. Gleeson and D. J. Cahalane, Phys. Rev. E **75**, 056103 (2007).

[63] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).

[64] B. Söderberg, Phys. Rev. E **68**, 015102 (2003).

[65] B. Söderberg, Phys. Rev. E **66**, 066121 (2002).

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

[67] K.-M. Lee, C. D. Brummitt, and K.-I. Goh, Phys. Rev. E **90**, 062816 (2014).