

Multi-Stage Complex Contagions in Random Multiplex Networks

Yong Zhuang and Osman Yağın

Department of ECE, Carnegie Mellon University, Pittsburgh, PA 15213, USA

(Dated: July 4, 2018)

Complex contagion models have been developed to understand a wide range of social phenomena such as adoption of cultural fads, the diffusion of belief, norms, and innovations in social networks, and the rise of collective action to join a riot. Most existing works focus on contagions where individuals' states are represented by *binary* variables, and propagation takes place over a single isolated network. However, characterization of an individual's standing on a given matter as a binary state might be overly simplistic as most of our opinions, feelings, and perceptions vary over more than two states. Also, most real-world contagions take place over multiple networks (e.g., Twitter and Facebook) or involve *multiplex* networks where individuals engage in different *types* of relationships (e.g., acquaintance, co-worker, family, etc.). To this end, this paper studies *multi-stage* complex contagions that take place over multi-layer or multiplex networks. Under a linear threshold based contagion model, we give analytic results for the probability and expected size of *global* cascades, i.e., cases where a randomly chosen node can initiate a propagation that eventually reaches a *positive* fraction of the whole population. Analytic results are also confirmed and supported by an extensive numerical study. In particular, we demonstrate how the dynamics of complex contagions is affected by the extra weight exerted by *hyper-active* nodes and by the structural properties of the networks involved. Among other things, we reveal an interesting connection between the assortativity of a network and the impact of *hyper-active* nodes on the cascade size.

I. INTRODUCTION

Modeling and analysis of dynamical processes in complex networks has been a very active research field in the past decade. This has led to many advances in our understanding and ability to control a wide range of physical and social phenomena. Examples include adoption of cultural fads, the diffusion of beliefs, norms, and innovations in social networks [1–7], disease contagion in human and animal populations [8–12], cascading failures in *interdependent* infrastructures [13–15], *insolvency* and *default* cascades in financial networks [16, 17], and the spread of computer viruses or worms on the Web [18, 19].

In this work, we focus on *complex* contagions, a class of dynamical processes typically used in modeling the propagation of *influence* in social networks. In particular, complex contagion models are used when *social reinforcement* plays an important role in the propagation process, i.e., when *multiple* sources of exposure is needed for an individual to adopt an activity. Examples include the spread of social movements and radical behavior, the rise of collective action to join a riot, or the decision to support one political candidate versus the other. This differs from the class of models known as *simple* contagions, where propagation often takes place after only a single copy is received; e.g., spread of diseases, viruses, etc.

Complex contagions have typically been studied in the literature using a linear threshold model. The original threshold model was proposed by Watts [1] for binary-state dynamics. In particular, each node is assumed to be in one of two states, *inactive* or *active*, and is initially given a threshold τ in $(0, 1]$. Then, a randomly chosen node is set as *active*, while all others are *inactive*. At any point in time, if an *inactive* node has d neighbors of

which m are *active*, we determine if it will be activated by checking the relationship between $\frac{m}{d}$ and the pre-assigned threshold τ . If $\frac{m}{d} \geq \tau$, then the node will turn *active*. Otherwise, if $\frac{m}{d} < \tau$, it stays *inactive*. It is also assumed that once activated, a node will remain active forever.

The Watts threshold model focuses solely on *single-stage* complex contagions, i.e., it is suitable only for binary-state dynamics. However, the characterization of an individual's standing on a given matter as a binary state might be overly simplistic, since most of our opinions, feelings, and perceptions vary over more than two states. Similarly, followers of a radical organization or a revolutionary movement may have varying levels of commitment to the cause, or might have varying desire and ability to recruit new members. Thus, it is of interest to study a model where nodes exhibit a richer set of states. To this end, Melnik et al. [20] generalized the linear threshold model and introduced a *multi-stage* contagion model. There, nodes can be inactive or can be in one of several levels of active states (e.g., active, hyper-active, etc.). The details of this model are given in Section II B.

In this work, we aim to extend the literature on multi-stage contagions from single networks to multi-layer and multiplex networks. Our intuition is that most real-world influence propagation events take place over multiple networks. For example, individuals may participate in multiple online social networks (e.g., Facebook, Twitter, etc.), and may have different levels of influence in each network. Similarly, within a single network, individuals may form different types of relationships (e.g., friendship, colleagueship, kinship, etc.), and each relationship type might have a different impact on propagation of influence in a given context. For example, video games might be more likely to spread among high-school friends rather than parents, while the opposite might be true for politi-

cal ideas. Thus, if we do not distinguish different types of relationships, dynamics of influence propagation may not be accurately captured. To address this issue, Yağın and Gligor proposed [4] a *context-dependent* linear threshold model in *multiplex* networks. In this model, each link type has its own weight on the propagation of the influence (which might change from one context to another); the detailed explanation of the model can be found in Section II B. However, the model in [4] considers only binary-state dynamics and hence is not able to model multi-stage contagions (where nodes can belong to a rich set of states).

The discussion above indicates that understanding how influence propagates in real-world networks requires incorporating two important factors: i) the multi-state nature of individuals’ activity levels; and ii) the multi-layer/multiplex nature of the networks involved in the propagation. To this end, in this work, we develop and analyze a multi-stage contagion model on multiplex networks. For simplicity, we assume that there are two types of links, *red* and *blue* in the network, and three stages, *inactive*, *active*, and *hyper-active*, representing the possible node states in the contagion dynamics. Our main results are i) calculating the probability of triggering *global cascades*, i.e., cases where a positive fraction of nodes (in the asymptotic limit) eventually becomes active or hyper-active when a randomly selected node is switched to the active state; ii) calculating the expected size of global cascades when they are possible. This is done for networks generated *randomly* from a given degree distribution, i.e., networks generated by the configuration model [21].

Our analytic results are confirmed and supported by an extensive numerical study. In particular, we demonstrate how the extra weight of hyper-active nodes in the multi-stage model and the structural properties of networks affect the contagion dynamics. For instance, a particularly interesting scenario is when the hyper-active state is manifested in only one link type. This is motivated by the case where people may be more willing to express their viewpoints to close friends instead of office-mates, or may be more influential in one social network (e.g., Twitter) versus another (e.g., Facebook). Among other things, we reveal interesting connections between the assortativity of a network and the impact of *hyper-active* nodes on cascade size. For instance, when the network is highly assortative, the influence exerted by the hyper-active nodes may change not only the critical transition points, but also the number and *order* of transitions; while the affect is much more limited in networks with low assortativity.

The rest of the paper is organized as follows. In Section II, we introduce the network and contagion models. Then, we describe the problem of interest and our main results in Section III. In Section IV, we present numerical results that demonstrate the accuracy of our analysis in the finite node regime, and discuss the impact of hyper-influencers on complex contagions. We conclude the pa-

per in Section V where we also suggest several directions for future work.

II. MODEL DEFINITION: NETWORKS AND DYNAMICS

In this section, we first present the multiplex network model considered in this work, and then describe the multi-stage complex contagion model.

A. Multi-layer and multiplex network models

We consider multiplex networks where links are classified into different types (or, colors). For convenience, in the following discussion, we focus on a multiplex network with two types of links, *red* and *blue*, but the model and results can be easily extended to an arbitrary number of link types. These two link types can be motivated by the case where one color accounts for edges in Facebook while the other for edges in Instagram. Alternatively, one link color may be representing close friendship links while the other representing “acquaintances” in a social network. In this network model, we let $\mathcal{N} = \{1, 2, \dots, n\}$ denote the vertex set, with n standing for the number of nodes. We let $\mathcal{N}_r \subset \mathcal{N}$ denote the set of vertices that have *red* edges and $\mathcal{N}_b \subset \mathcal{N}$ denote the set of vertices having *blue* edges. For simplicity, we assume $\mathcal{N}_b = \mathcal{N}$, which means all vertices in the network may have blue edges. To model the possibility that not everyone may have red links, we assume that each vertex in \mathcal{N} has red links with probability $\alpha \in (0, 1]$:

$$\mathbb{P}[i \in \mathcal{N}_r] = \alpha, \quad i = 1, \dots, n. \quad (1)$$

This network model can be interpreted in two different ways. The first one is a multi-layer network where each network layer is generated by the widely used configuration model [21–23]; this case is illustrated in Figure 1(a). In particular, we use $P(d_r)$ (resp. $P(d_b)$) as the *degree distribution* to determine the number of red (resp. blue) edges that will be assigned to each node in \mathcal{N}_r (resp. \mathcal{N}_b). Once the degree of each node is determined, we generate the networks \mathbb{R} and \mathbb{B} by selecting a graph uniformly at random from among all possible graphs that have the same degree sequence; see [21, 22] for more details. Next, we take a union of the edges in \mathbb{R} and \mathbb{B} to create a network \mathbb{H} . Equivalently, we can consider a multiplex network model generated by the *colored* configuration model [24]. Let $\mathbf{d} = (d_r, d_b)$ denote the colored degree of a node, where d_r and d_b stand for the number of red edges and blue edges incident on it. Each of the n nodes in the network is assigned a colored degree by independently drawing from the distribution $P_{\mathbf{d}}$. Then, pairs of edges of the same color are randomly chosen and connected together until none is left; see [24] for details. Figure 1(b) is an illustration of this multiplex network model.

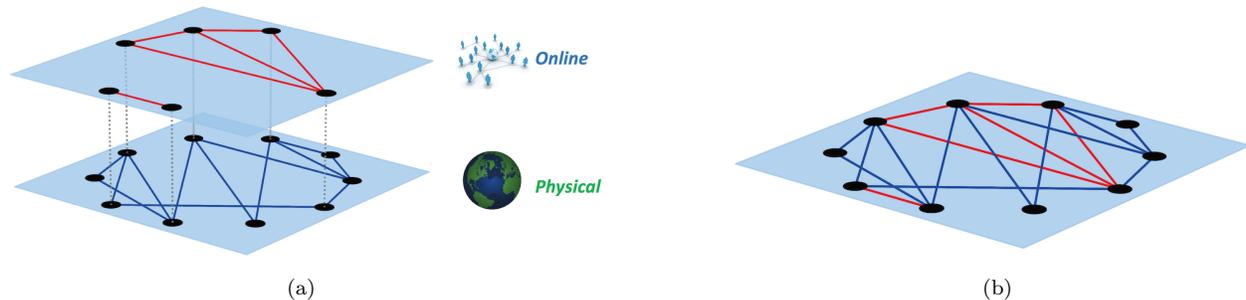


FIG. 1. Illustration of a multi-layer and a multiplex network representation of our model. In (a), we see a multi-layer network (e.g., a Physical communication layer and an online social network layer) with overlapping vertex sets; vertical dashed lines represent nodes corresponding to the same individual. In (b), we see the equivalent representation of this model by a multiplex network. Edges from Facebook are shown in red and edges from the physical network are shown in blue.

B. Multi-stage content-dependent linear threshold model for complex contagions

We first introduce the single-stage content-dependent linear threshold model [4] which is a generalization of the vanilla threshold model [1]. In the content-dependent linear threshold model, links are classified into r types. For a given content (a view, rumor, product, etc.), scalars c_i , $i = 1, \dots, r$ represent the weight (i.e., relative importance) of type- i edges on spreading this particular content. Nodes belong to either one of the two states, active or inactive, and each node is assigned a threshold τ in $(0, 1]$ drawn from a distribution $P(\tau)$. Given an inactive node with m_i active and $d_i - m_i$ inactive neighbors for each link type- i , $i = 1, \dots, r$, an inactive node will turn active if $\frac{\sum_i c_i m_i}{\sum_i c_i d_i} \geq \tau$. Namely, an inactive node with $\mathbf{m} = (m_1, \dots, m_r)$ and $\mathbf{d} = (d_1, \dots, d_r)$ will turn *active* with probability

$$F[\mathbf{m}, \mathbf{d}] \triangleq \mathbb{P} \left[\frac{\sum_{i=1}^r c_i m_i}{\sum_{i=1}^r c_i d_i} \geq \tau \right]. \quad (2)$$

Throughout, $F[\mathbf{m}, \mathbf{d}]$ is referred to as the response function. If we do not distinguish the edge types or simply set $c_i = 1$ for all $i = 1, \dots, r$, then this model reduces to the Watts' threshold model [1]. This content-dependent threshold model enables us to model the case where people's influence on others vary according to their relationship type, or the social network that they are interacting through.

Different from the single-stage threshold model where nodes can only be in two states, the multi-stage linear threshold model [20] allows nodes to be in a richer set of active states. In this work, we assume that nodes can belong to three states, inactive, active, and hyper-active. In the following discussion, we use state-0, state-1, and state-2 to represent the inactive, active, and hyper-active state, respectively. Let τ_1 and τ_2 denote the thresholds associated with transitioning to the active and hyper-active states, respectively. The hyper-active individuals are assumed to be β -times more influential than active

nodes in the propagation process (where $\beta \geq 1$). For example, an individual with d neighbors of which m_1 are active and m_2 are hyper-active, the probability of switching to state- i from the inactive state (i.e., state-0) is given by:

$$F_i[\mathbf{m}, d] \triangleq \mathbb{P} \left[\tau_i \leq \frac{m_1 + \beta m_2}{d} \leq \tau_{i+1} \right], \quad i = 0, 1, 2, \quad (3)$$

where $\mathbf{m} = (m_1, m_2)$, $\tau_0 = 0$, $\tau_3 = \infty$, and $\beta \geq 1$. Although we assume there are three states in the contagion process, our analysis can be extended to an arbitrary number of states.

Finally, we introduce the multi-stage content-dependent linear threshold model. Assume that there are two types of links, red and blue, in the network, and that nodes can be in three states, inactive, active, and hyper-active. We let c_r and c_b denote the weight of red and blue edges, respectively, and set $c = \frac{c_r}{c_b}$. With this notation, the probability of an inactive node switching to state- i is given by:

$$F_i[\mathbf{m}, \mathbf{d}] \triangleq \mathbb{P} \left[\tau_i \leq \frac{c(m_{r,1} + \beta m_{r,2}) + m_{b,1} + \beta m_{b,2}}{cd_r + d_b} \leq \tau_{i+1} \right], \quad (4)$$

where $\mathbf{m} = (m_{r,1}, m_{r,2}, m_{b,1}, m_{b,2})$, $\mathbf{d} = (d_r, d_b)$, $m_{r,1}$ and $m_{r,2}$ (resp. $m_{b,1}$ and $m_{b,2}$) denote the number of active and hyper-active neighbors connected through a red (resp. blue) edge, and d_r and d_b denote the number of red and blue neighbors respectively.

III. MAIN RESULTS

We assume that the contagion process starts by randomly choosing an initial node and setting it as *active*, while all other nodes are in the inactive state. The influence might then propagate in the network according to (4) and other nodes might turn active, and so on. Since

the contagion process is monotone (i.e., an active node can never switch back to inactive), it will eventually stop, i.e., a steady-state will be reached.

Our main goals are i) determining the conditions (in terms of network parameters) for *global cascades* to be possible, i.e., cases where influence starts from a single individual (selected uniformly at random) and eventually reaches a positive fraction of the population in the limit of large network sizes; ii) calculating the expected size of global cascades when they are possible; and iii) calculating the probability of triggering global cascades.

A. Expected cascade size and the condition to have a global cascade

We start the analysis with computing the expected size of global cascades when they occur. Consider a random variable S defined as

$$S \triangleq \frac{\# \text{ of active and hyper-active nodes at steady-state}}{n},$$

where n is the number of nodes in the network. Then, a *global cascade* is said to take place if $S > 0$ in the limit $n \rightarrow \infty$, and our main goal is to derive

$$\lim_{n \rightarrow \infty} \mathbb{E}[S \mid S > 0],$$

which gives the expected size of global cascades when they exist. Our analysis is based on the “tree-approximation” approach [4, 20, 25], which was developed to analyze the zero-temperature random-field Ising model on Bethe lattices [26]. The tree-approximation approach assumes that the network has a locally tree-like structure. Labeling the tree structure from the bottom to the top, it is assumed that the node states are updated starting from the bottom, and continuing to the top one level at a time. In other words, the nodes at level ℓ will not update their states until the nodes at level $0, 1, \dots, \ell - 1$ have finished updating. We define $q_{r,1,\ell}$ (resp. $q_{b,1,\ell}$) as the probability that the node at level ℓ is active and is connected to its only parent at level $\ell + 1$ by a red (resp. blue) edge. Similarly, we define $q_{r,2,\ell}$ (resp. $q_{b,2,\ell}$) as the probability that an inactive node at level ℓ , which is attached to its only parent via a red (resp. blue) edge, turns hyper-active. We assume that the parent nodes at level $\ell + 1$ are inactive.

In the interest of brevity, we only introduce the derivation of $q_{r,1,\ell+1}$, because the derivations of $q_{r,2,\ell+1}$, $q_{b,1,\ell+1}$

and $q_{b,2,\ell+1}$ can be explained in a similar way. Since $q_{r,1,\ell+1}$ cannot be expressed explicitly, we derive a recursive relation in terms of $q_{r,1,\ell}$, $q_{r,2,\ell}$, $q_{b,1,\ell}$, and $q_{b,2,\ell}$; see (6)-(9). The validity of the expression (6) for $q_{r,1,\ell+1}$ can be explained as follows. Consider an inactive node at level $\ell + 1$ that is connected to its unique parent at level $\ell + 2$ via a red edge, and that has colored degree $\mathbf{d} = (d_r, d_b)$. The probability that this node has i active children connected via red edges, s active children connected via blue edges, j hyper-active children connected via red edges, and t hyper-active children connected via blue edges, and that it turns active is given by

$$\begin{aligned} & \binom{d_r - 1}{i} \binom{d_r - 1 - i}{j} q_{r,1,\ell}^i q_{r,2,\ell}^j (1 - q_{r,1,\ell} - q_{r,2,\ell})^{d_r - 1 - i - j} \\ & \times \binom{d_b}{s} \binom{d_b - s}{t} q_{b,1,\ell}^s q_{b,2,\ell}^t (1 - q_{b,1,\ell} - q_{b,2,\ell})^{d_b - s - t} \\ & \times F_1[(i, j, s, t), \mathbf{d}], \end{aligned} \quad (5)$$

where $F_1[(i, j, s, t), \mathbf{d}]$ is as defined in (4); i.e., it denotes the probability that an inactive node with a colored degree \mathbf{d} and a group of active and hyper-active neighbors for each color represented by $\mathbf{m} = (i, j, s, t)$ switches to state-1. To simplify the notation, we use $\mathbb{F}_1[(i, j, s, t), (x, y)]$ as defined at (10), so the term given in (5) becomes equivalent to $\mathbb{F}_1[(i, j, s, t), (d_r - 1, d_b), \ell]$.

The intuition behind (5) is as follows. Since we assume that the network is tree-like, the state of each child node at level ℓ is independent from other children at the same level. Thus, we multiply together the probability of being at a specific state for each child node to get the whole expression (5) using a simple combinatorial argument. The reason behind using $d_r - 1$ rather than d_r in (5) is the fact that the node under consideration is attached to its unique parent at level $\ell + 2$ through a *red* edge, and by assumption this parent node is inactive; recall that a node at level $\ell + 2$ can not update its state until all nodes in level $\ell + 1$ finish updating. A node that is known to have at least one red edge can be seen to have colored *degree* $\mathbf{d} = (d_r, d_b)$ with probability $\frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle}$; e.g., see [4, 21] for a discussion on the *excess* degree distribution. Finally, we get the detailed expressions of $q_{r,1,\ell+1}$ (6) after taking the expectation of (5) over the degree of the node at level $\ell + 1$. We can use similar arguments to derive expressions for $q_{r,2,\ell+1}$, $q_{b,1,\ell+1}$, and $q_{b,2,\ell+1}$. The expressions of all four probabilities are shown in (6) - (9).

$$q_{r,1,\ell+1} = \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} \sum_{i=0}^{d_r-1} \sum_{j=0}^{d_r-1-i} \sum_{s=0}^{d_b} \sum_{t=0}^{d_b-s} \mathbb{F}_1 [(i, j, s, t), (d_r - 1, d_b), \ell] \quad (6)$$

$$q_{r,2,\ell+1} = \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} \sum_{i=0}^{d_r-1} \sum_{j=0}^{d_r-1-i} \sum_{s=0}^{d_b} \sum_{t=0}^{d_b-s} \mathbb{F}_2 [(i, j, s, t), (d_r - 1, d_b), \ell] \quad (7)$$

$$q_{b,1,\ell+1} = \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} \sum_{i=0}^{d_r} \sum_{j=0}^{d_r-i} \sum_{s=0}^{d_b-1} \sum_{t=0}^{d_b-1-s} \mathbb{F}_1 [(i, j, s, t), (d_r, d_b - 1), \ell] \quad (8)$$

$$q_{b,2,\ell+1} = \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} \sum_{i=0}^{d_r} \sum_{j=0}^{d_r-i} \sum_{s=0}^{d_b-1} \sum_{t=0}^{d_b-1-s} \mathbb{F}_2 [(i, j, s, t), (d_r, d_b - 1), \ell], \quad (9)$$

where for $k = 1, 2$, we define

$$\begin{aligned} \mathbb{F}_k [(i, j, s, t), (x, y), \ell] &= \binom{x}{i} \binom{x-i}{j} q_{r,1,\ell}^i q_{r,2,\ell}^j (1 - q_{r,1,\ell} - q_{r,2,\ell})^{x-i-j} \\ &\times \binom{y}{s} \binom{y-s}{t} q_{b,1,\ell}^s q_{b,2,\ell}^t (1 - q_{b,1,\ell} - q_{b,2,\ell})^{y-s-t} \times F_k [(i, j, s, t), (x, y)]. \end{aligned} \quad (10)$$

Equations (6) - (9) form a non-linear system, which can be solved recursively to obtain the steady-state values (i.e., fixed points), $q_{r,1,\infty}$, $q_{r,2,\infty}$, $q_{b,1,\infty}$, and $q_{b,2,\infty}$. Since our goal is to compute the expected size of global cascades *given that they exist*, we can initialize this dynamical system with $q_{r,1,0}, q_{r,2,0}, q_{b,1,0}, q_{b,2,0} > 0$. Because these fixed points account for the probability of being in a corresponding state for the children of the top node, we can use them to calculate the expected size of global cascades. The expected cascade size stands for the final fraction of active and hyper-active individuals in the network. This fraction is equal to the probability that the node at the top of the tree turns active or hyper-active.

We give the expected size of the cascades (given that they exist) in (11). The validity of (11) can be seen as follows: First, we randomly choose a node, whose colored degree is $\mathbf{d} = (d_r, d_b)$, with probability $p_{\mathbf{d}}$. The probabil-

ity that each of its d_r neighbors (via red links) is active (resp. hyper-active) is given by $q_{r,1,\infty}$ (resp. $q_{r,2,\infty}$). Similarly, each of the d_b neighbors (connected via blue links) of this randomly chosen node is active with probability $q_{b,1,\infty}$ and hyper-active with probability $q_{b,2,\infty}$, independently from each other. Then, with each possible combination of numbers of active and hyper-active neighbors, we can calculate the probability of being active or hyper-active for the node by the response function (4). Taking the expectation with respect to the degree \mathbf{d} yields (11). As discussed in details in [4, 20, 25], this method, based on the *tree-approximation* technique, gives precise results in the asymptotic limit $n \rightarrow \infty$, when the underlying network is *sparse* and is generated according to the configuration model. We present extensive numerical studies in Section IV that supports our results in the finite node regime.

$$\lim_{n \rightarrow \infty} \mathbb{E}[S | S > 0] = \sum_{\mathbf{d}} p_{\mathbf{d}} \sum_{i=0}^{d_r} \sum_{j=0}^{d_r-i} \sum_{s=0}^{d_b} \sum_{t=0}^{d_b-s} \{\mathbb{F}_1 [(i, j, s, t), (d_r, d_b), \infty] + \mathbb{F}_2 [(i, j, s, t), (d_r, d_b), \infty]\}. \quad (11)$$

From the recursive equations derived above, we can also obtain the conditions needed for the global cascades to be possible; i.e., conditions under which $S > 0$ with a *positive* probability in the limit $n \rightarrow \infty$. For notational convenience, we define $q_1 := q_{r,1,\infty}$, $q_2 := q_{r,2,\infty}$, $q_3 := q_{b,1,\infty}$, and $q_4 := q_{b,2,\infty}$. Then, the four recursive

equations (6) - (9) take the form

$$q_i = f_i(q_1, q_2, q_3, q_4), \quad i = 1, 2, 3, 4. \quad (12)$$

By direct inspection, we see that the recursive equations (12) have a trivial fixed point $q_{r,1,\infty} = q_{r,2,\infty} = q_{b,1,\infty} = q_{b,2,\infty} = 0$, which yields $S = 0$ almost surely; this can be seen from the fact that in that case we have $\mathbb{E}S = 0$.

In other words, if (12) has only the trivial fixed point, then global cascades are not possible to take place. In general, the trivial fixed point may not be stable and there may exist other non-trivial fixed points which can yield $\mathbb{E}S > 0$; i.e., $S > 0$ with a positive probability in the limit $n \rightarrow \infty$ and hence global cascade may take place. To check the existence of non-trivial solutions of (6) - (9), we can linearize these equations at $q_1 = q_2 = q_3 = q_4 = 0$ which yields the Jacobian matrix \mathbf{J} given as

$$\mathbf{J} = \left. \frac{\partial f_i(q_1, q_2, q_3, q_4)}{\partial q_j} \right|_{q_1=q_2=q_3=q_4=0}. \quad (13)$$

If the spectral radius, i.e., the largest eigenvalue in absolute value, of the Jacobian matrix is larger than one, then the trivial fixed point $q_1 = q_2 = q_3 = q_4 = 0$ is not stable. That is, there exist a non-trivial fixed point indicating that global cascades are possible and $S > 0$ with positive probability. Otherwise, if the spectral radius of \mathbf{J} is less than or equal to one, then there will be no global cascades.

After the analysis and discussion on the expected size and the conditions of global cascades, we focus next on the probability of triggering a global cascade. In other words, we will calculate the exact (asymptotic) probability that a node selected uniformly at random and turned active leads eventually to a global cascade; i.e., we will compute

$$\lim_{n \rightarrow \infty} \mathbb{P}[S > 0].$$

B. Probability of triggering a global cascade

We now turn our attention to computing the probability $\mathbb{P}[S > 0]$ of global cascades. As discussed in [1, 4], the possibility of a seed node to trigger a *global* cascade is closely tied to the size of (and the seed node's connectivity to) the set of *vulnerable* nodes in the network; a node is deemed *vulnerable* if it can be activated by only one *active* neighbor [1, 2, 4, 27]. The definition of vulnerable nodes and of the vulnerable component has been extended in [4] to the case of multiplex networks. There, a “vulnerable component” is defined as a set of nodes, each of which is vulnerable w.r.t. at least one of the link types, such that in the subgraph containing this set of nodes, activating any node leads to the activation of all nodes in the set. A multiplex network is said to contain a *giant* vulnerable component (GVC) if the fraction of nodes in its largest vulnerable component is positive in the limit $n \rightarrow \infty$. These definitions were then used [4] to demonstrate that an initial node can trigger a global cascade if and only if it belongs to the *extended* giant vulnerable cluster (EGVC), that contains nodes in the GVC and nodes whose activation leads to activation of a node in GVC. Put differently, the probability of a randomly selected node triggering a global cascade is equal to the fractional size of the EGVC; see [4] for details.

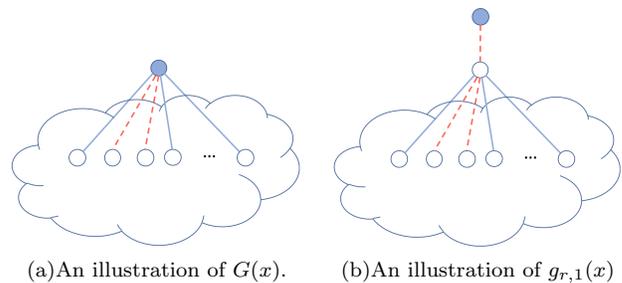


FIG. 2. The difference between $G(x)$ and $g_{r,1}(x)$. Red dashed lines account for red edges in our analysis, while other lines represent blue edges. Circles with solid fill indicate active nodes while circles without fill account for inactive nodes (that can potentially be made active). Then, $G(x)$ generates the distribution of the number of active nodes by following the initially activated node, while $g_{r,1}(x)$ generates the distribution of the number of active nodes by following a randomly chosen red edge.

Here, we use the ideas mentioned above to calculate the probability of global cascades, or equivalently the fraction of nodes that are in EGVC. This will be done through the analysis of a branching process that starts from a randomly selected and activated initial node, and keeps exploring the neighboring nodes that are activated according to nodes' response function (4). The branching process will continue by exploring the neighbors of the newly activated nodes that will also be activated, and so on. The size S of the influence cascade will then be equal to the fraction of nodes identified by this branching process.

The fraction of nodes identified by the branching process described above can be analyzed using the method of probability generating functions [28]; e.g., see [1, 4, 7, 9, 21] where this tool was demonstrated to be useful for similar purposes. The first generating function we use in our analysis is $G(x)$, and it generates the probability distribution of “the finite number of nodes reached and influenced by the above branching process”; different from [3, 7, 29], we exclude the initially activated nodes. We have

$$G(x) = \sum_{\mathbf{d}} p_{\mathbf{d}} g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b}, \quad (14)$$

where $g_{r,1}(x)$ (resp. $g_{b,1}(x)$) generates the probability distribution of “the finite number of nodes reached and influenced by following a randomly chosen *red* (resp. *blue*) edge one of whose ends is set to active.” The difference between $G(x)$ and $g_{r,1}(x)$ is illustrated in Figure 2. The validity of expression (14) can be seen as follows. First, we initially activate a node which is chosen uniformly at random. The probability that this node has a degree $\mathbf{d} = (d_r, d_b)$ is $p_{\mathbf{d}}$. In that case, the number of nodes that are reached and activated by this node will be generated (in view of the *powers property* of the generating functions) by $g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b}$. Summing over all

possible degrees \mathbf{d} of the initial node leads to (14).

For (14) to be useful, we also need to derive expressions for $g_{r,1}(x)$ and $g_{b,1}(x)$. This will be done by the help of two more generating functions. Namely, let $g_{r,2}(x)$ (resp. $g_{b,2}(x)$) generate the distribution of “the finite number of nodes reached and influenced by following a *red* (resp. *blue*) edge whose one end is connected to a hyper-active node.” The detailed expressions of the four generating functions are given in (15) - (18). Here, we only explain the derivation $g_{r,1}(x)$, as others can be explained in a similar manner.

To see why (15) holds, first note that as the randomly selected red edge whose one end is connected to an active node is followed, we will find a node with colored *degree* $\mathbf{d} = (d_r, d_b)$ with probability $\frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle}$ as already explained in the derivation of (6). There are three possible cases for this node with degree $\mathbf{d} = (d_r, d_b)$:

- It turns *active*, i.e., $\tau_1 \leq \frac{c}{cd_r + d_b} < \tau_2$, which happens with probability $F_1[(1, 0, 0, 0), (d_r, d_b)]$. Then, this newly activated node will activate

$g_{r,1}(x)^{d_r-1} g_{b,1}(x)^{d_b}$ other nodes based on the powers property of generating functions. The reason why we use $d_r - 1$ instead of d_r is because one of its d_r edges has already been considered as its connection to the *active* end.

- It turns *hyper-active*, i.e., $\frac{c}{cd_r + d_b} \geq \tau_2$, which happens with probability $F_2[(1, 0, 0, 0), (d_r, d_b)]$. Then, the number of nodes reached and influenced by this newly activated node will be generated by $g_{r,2}(x)^{d_r-1} g_{b,2}(x)^{d_b}$; this can be seen via similar arguments to the case above.
- It remains *inactive*, i.e., $\frac{c}{cd_r + d_b} < \tau_1$, which happens with probability $1 - F_1[(1, 0, 0, 0), (d_r, d_b)] - F_2[(1, 0, 0, 0), (d_r, d_b)]$. Then, there will be no newly activated nodes.

Combining these three cases and summing over all possible \mathbf{d} , we get (15), where the explicit factor x accounts for the initial node that is activated. The expressions for $g_{r,2}(x)$, $g_{b,1}(x)$, and $g_{b,2}(x)$ can be derived similarly.

$$g_{r,1}(x) = x \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} \left[F_1[(1, 0, 0, 0), (d_r, d_b)] g_{r,1}(x)^{d_r-1} g_{b,1}(x)^{d_b} + F_2[(1, 0, 0, 0), (d_r, d_b)] g_{r,2}(x)^{d_r-1} g_{b,2}(x)^{d_b} \right] + x^0 \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} (1 - F_1[(1, 0, 0, 0), (d_r, d_b)] - F_2[(1, 0, 0, 0), (d_r, d_b)]) \quad (15)$$

$$g_{r,2}(x) = x \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} \left[F_1[(0, 1, 0, 0), (d_r, d_b)] g_{r,1}(x)^{d_r-1} g_{b,1}(x)^{d_b} + F_1[(0, 1, 0, 0), (d_r, d_b)] g_{r,2}(x)^{d_r-1} g_{b,2}(x)^{d_b} \right] + x^0 \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} (1 - F_1[(0, 0, 1, 0), (d_r, d_b)] - F_2[(0, 0, 1, 0), (d_r, d_b)]) \quad (16)$$

$$g_{b,1}(x) = x \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} \left[F_1[(0, 0, 1, 0), (d_r, d_b)] g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b-1} + F_2[(0, 0, 1, 0), (d_r, d_b)] g_{r,2}(x)^{d_r} g_{b,2}(x)^{d_b-1} \right] + x^0 \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} (1 - F_1[(0, 0, 1, 0), (d_r, d_b)] - F_2[(0, 0, 1, 0), (d_r, d_b)]) \quad (17)$$

$$g_{b,2}(x) = x \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} \left[F_1[(0, 0, 0, 1), (d_r, d_b)] g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b-1} + F_2[(0, 0, 0, 1), (d_r, d_b)] g_{r,2}(x)^{d_r} g_{b,2}(x)^{d_b-1} \right] + x^0 \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} (1 - F_1[(0, 0, 0, 1), (d_r, d_b)] - F_2[(0, 0, 0, 1), (d_r, d_b)]) \quad (18)$$

These recursive equations can be used to compute the probability that a global cascade is triggered in the following manner. Since $G(x)$ generates the number of *finite* nodes reached and activated by this branching process, we should have $G(1) = 1$ by the conservation of probability, *unless* there is a positive probability that the branching process leads to an *infinite* number of nodes. In other words, $1 - G(1)$ corresponds to the probability that the branching process under consideration will *survive* forever and will not go extinct, meaning that the underlying influence propagation process will constitute a global

cascade. Thus, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}[S > 0] = 1 - G(1). \quad (19)$$

This approach has been introduced in [1] and used in [4, 21] for similar calculations.

In order to calculate $G(1)$, we now solve for the fixed point of (15)-(18) at $x = 1$. Simplifying the notation as $g_1 := g_{r,1}(1)$, $g_2 := g_{r,2}(1)$, $g_3 := g_{b,1}(1)$, and $g_4 := g_{b,2}(1)$, the recursive equations (15)-(18) at $x = 1$ can be expressed as

$$g_i = h_i(g_1, g_2, g_3, g_4), \quad i = 1, 2, 3, 4. \quad (20)$$

Here, the exact form of the functions $h_1(g_1, g_2, g_3, g_4), \dots, h_4(g_1, g_2, g_3, g_4)$ will be obtained from (15)-(18). Once the fixed points of (20) are obtained, we get from (14) that

$$G(1) = \sum_{\mathbf{d}} p_{\mathbf{d}} g_1^{d_r} g_3^{d_b}. \quad (21)$$

In view of (19), we finally obtain the desired probability of global cascades as

$$\lim_{n \rightarrow \infty} \mathbb{P}[S > 0] = 1 - \sum_{\mathbf{d}} p_{\mathbf{d}} g_1^{d_r} g_3^{d_b}. \quad (22)$$

IV. NUMERICAL RESULTS

In this section, we present numerical results to support our analysis on the expected size and probability of global cascades. In particular, we are interested in checking the accuracy of our asymptotic results when the number of nodes is finite. We will also investigate via extensive simulations the impact of hyper-influencers (i.e., the additional influence exerted by them) on the contagion dynamics.

A. The agreement between our analysis and simulations

First, we focus on demonstrating the accuracy of our analytic results on the expected size of global cascades and the probability of having global cascades in the finite node regime. In our numerical simulations, we use a doubly Poisson distribution to assign the number of red and blue edges for each node. Namely, with p_k^r (resp. p_k^b) denoting the probability that a node is assigned k red (resp. blue) edges, we let

$$p_k^b = e^{-\lambda_b} \frac{(\lambda_b)^k}{k!}, \quad k = 0, 1, \dots, \quad (23)$$

$$p_k^r = \alpha e^{-\lambda_r} \frac{(\lambda_r)^k}{k!} + (1 - \alpha) \delta_{k,0}, \quad k = 0, 1, \dots \quad (24)$$

Here, λ_r (resp. λ_b) denotes the mean number of red (resp. blue) edges assigned per node, α denotes the fraction of nodes that have red edges (i.e., the relative size of the *red* network \mathbb{R}), and δ denotes the Kronecker delta. In our simulations to verify our analysis on the expected size, we use $n = 1 \times 10^6$ nodes[?] to create networks and set $\alpha = 0.5$. Besides, we use $c = 0.5$ and $\beta = 1.5$ as the content parameter and the weight of hyper-active nodes, respectively, and fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$. Then, for several values of $\lambda_r = \lambda_b$, we run 1,000 independent experiments (for each parameter set), each time computing the fraction of nodes that eventually turn active or hyper-active. The results are depicted in Figure 3 where lines represent analytical results obtained from (6) - (9), and symbols represent the average cascade size

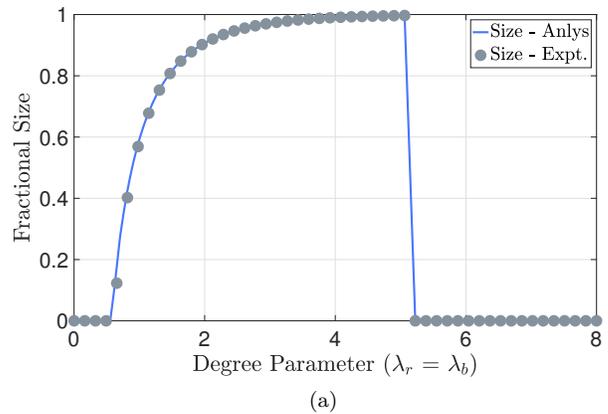


FIG. 3. Simulations for doubly Poisson degree distributions, $n = 1 \times 10^6$, $\alpha = 0.5$, $\tau_1 = 0.18$, and $\tau_2 = 0.32$. The weight of hyper-influencers is taken to be $\beta = 1.5$.

obtained in simulations (over 1,000 experiments for each data point). We see that there is a good agreement between the analytic results and the simulations.

Next, to check the correctness of our analysis on the probability, we fix all parameters except increasing the number of experiments from 1,000 to 10,000. As shown in Figure 4, we observe that our analysis on the probability (15) - (18) also match very well the simulations results. This indicates that, although asymptotic in nature, the results presented in Section III are still helpful in understanding complex contagion dynamics (e.g., the probability and expected size of global cascades) in finite networks.

In addition, we observe from both Figure 3 and 4 that the contagion exhibits two *phase transitions*, i.e., two different $\lambda_r = \lambda_b$ values around which fractional cascade size transitions from zero to a positive value, or vice versa. These points are of great interest since they provide insights on how network connectivity affects the possibility of observing global influence spreading events. The first transition occurs around *low* values of λ , and reflects the fact that global spreading events become possible only after the network reaches a certain level of connectivity. The second phase transition occurs around *high* λ values, indicating that global cascades can not occur when nodes are locally stable; i.e., when they have a large number of friends, individuals tend to be difficult to get influenced by a few active neighbors.

After demonstrating the correctness of our analysis, we focus on exploring the impact of hyper-influencers on complex contagion dynamics in the following sections.

B. The impact of hyper-influencers on the global cascade boundary

In this section, we investigate how the parameters β, τ_1, τ_2 of the contagion model and the connectivity of

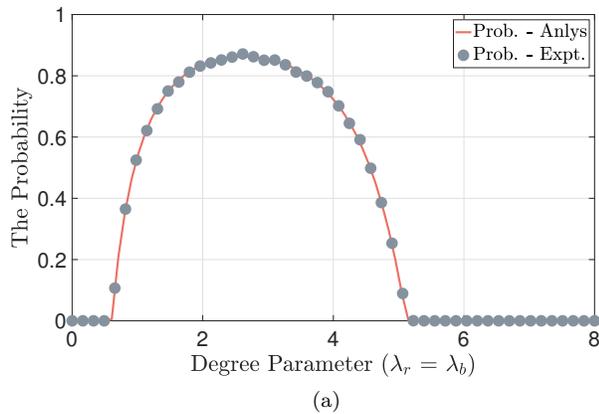


FIG. 4. Simulations for doubly Poisson degree distributions, $n = 1 \times 10^6$, $\alpha = 0.5$, $\tau_1 = 0.18$, and $\tau_2 = 0.32$. The weight of hyper-influencers is taken to be $\beta = 1.5$.

the network jointly affect the possibility of global cascades. In particular, we will determine the boundaries in the space of parameters that separate the region where global cascades are possible (i.e., $\mathbb{P}[S > 0]$) from the region where global cascades do *not* take place almost surely (i.e., $\mathbb{P}[S = 0]$). First, we will focus on the impact of the weight β of hyper-influencers on the global cascade boundary, and then move on to the discussion about the impact of the threshold τ_1 of ordinary influencers.

Figure 5 shows the global cascade boundary in the space of τ_2 and degree parameter $\lambda = \lambda_r = \lambda_b$, for several values of β . We observe that larger β values lead to a larger region of parameters τ_2, λ for which global cascades can take place; i.e., the global cascade region gets larger with increasing β . An interesting observation is that the cascade boundary is more sensitive to the changes in β values when λ is *large*; i.e., the lower parts of the boundaries seen in Figure 5 are less dependent on the choice of β as compared to the upper parts. This can be explained as follows. When λ is *small*, the existence of global cascades (and hence the cascade boundary) is mainly determined by whether the network has enough connectivity to spread the influence. However, increasing β does not change the connectivity of the network, and hence does not affect the boundary when λ is low. Differently, when λ *high*, the location of the boundary (i.e., the second phase transition points seen in Figures 3-4) is decided by the likelihood of nodes with high degree being influenced by a single active or hyper-active neighbor. Thus, the boundary is determined from a node's perceived influence, or perceived proportion of active and hyper-active neighbors, given at (4), and on how this compares with the activation thresholds τ_1 and τ_2 . From (4), we see that higher β leads to an increased perceived influence for a node that has at least one hyper-active neighbor, making it possible for the activation threshold to be exceeded at higher d_r, d_b values (equivalently at higher λ values). Thus, when λ is *high*, the boundary tends to be

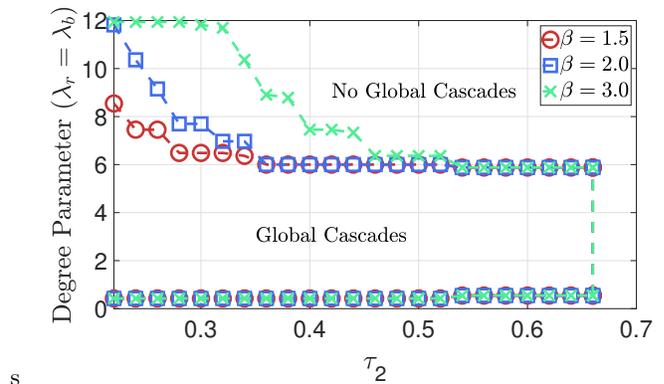


FIG. 5. Given $\tau_1 = 0.15$ and $\alpha = 0.5$, we vary the mean degree λ and τ_2 to plot the global cascade region for several β , 1.5, 2.0, and 3.0.

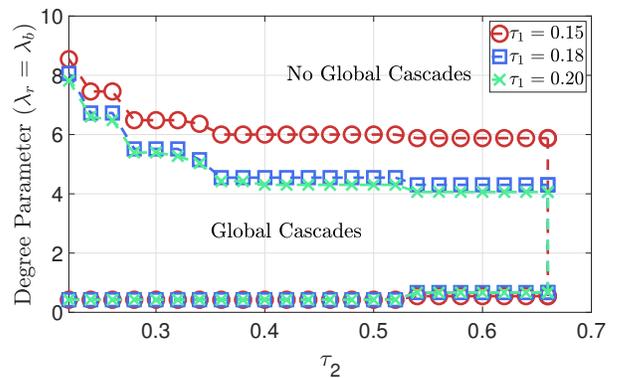


FIG. 6. Given the weight of extra influence $\beta = 1.5$ and $\alpha = 0.5$, we vary the degree parameter $\lambda = \lambda_b = \lambda_r$ and τ_2 to plot the region where there exist a global cascade for several τ_1 , 0.15, 0.18, and 0.2. Both of the edges are assigned by the doubly Poisson distribution in Section IV A.

more sensitive to the changes in β .

Next, we investigate the impact of the activation threshold τ_1 on the global cascade boundary (again considering the space of $\lambda - \tau_2$). In Figure 6, we fix $\beta = 1.5$ and plot the boundary on the $\tau_2 - \lambda$ plane that separates the regions where cascades are possible and not possible, respectively. This is done for three different values of τ_1 . We observe that the impact of τ_1 (i.e., the threshold on the perceived influence that an inactive node needs to receive in order to turn active) on the cascade boundary is opposite to that of β . That is, the higher τ_1 is, the smaller is the region where *global* cascades are possible. The reason behind this observation is as follows. From the expression of the response function (4), we see that it is decreasing with increasing τ_1 . In other words, a higher τ_1 makes it harder for nodes to become active (i.e., influenced), leading to a smaller cascade region.

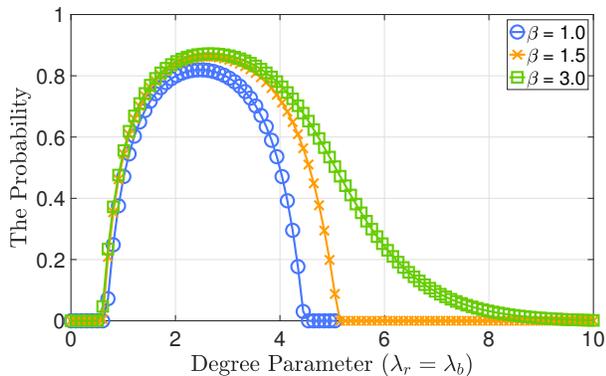


FIG. 7. The comparison between different β for the probability of triggering a *global cascade*.

C. The impact of hyper-influencers on the probability and expected size of global cascades

We start by investigating the impact of the extra influence β (that hyper-active nodes exert on their neighbors) on the probability of global cascades. From Figure 7, we observe that a larger β will increase the probability of triggering a global cascade. This observation is intuitive given that the response function (4) is increasing with respect to β . Thus, with a higher β , the perceived influence from a single active or hyper-active neighbor exceeds the threshold more easily, leading to a larger *vulnerable* component. Also, we see in Figure 7 that when the degree parameter is *large*, the cascade probability becomes more sensitive to the changes in β . This is consistent with the observations from Figure 6 and can be explained in a similar manner.

Next, we discuss the impact of hyper-influencers on the expected size of global cascades. From Figure 8, we observe that increasing β leads to an expansion of the interval of $\lambda_r = \lambda_b$ values for which expected cascade size is positive. However, over the common interval where cascade size is positive, we see that increasing β nearly does not lead to changes in the expected cascade size. The reason behind this observation is that the expected cascade size is mainly determined by the connectivity, (e.g., the mean degree) of the network, which remains invariant to changes in β . Thus, increasing β nearly does not change the expected size of global cascades. The expansion of the interval over which $S > 0$ with increasing β is explained by the response function (4) being increasing in β . In other words, a higher β makes it easier for the perceived influence to exceed the activation threshold, helping global cascades take place even at higher mean degree.

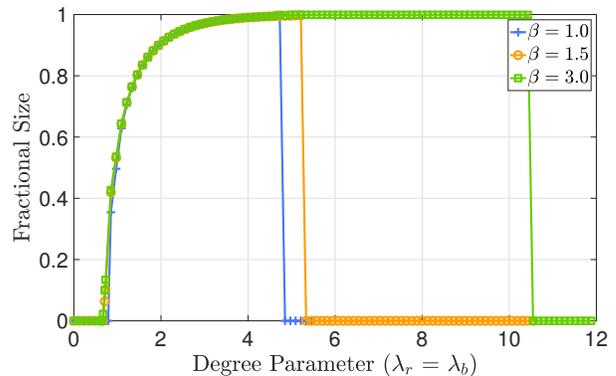


FIG. 8. The comparison between different β for the expected *global cascade size*.

D. The impact of hyper-influencers in multiplex networks

In this section, we investigate more closely how hyper-influencers affect the complex contagions. To this end, we restrict hyper-active nodes to have additional influence only through one type of edges, red or blue, rather than allowing them to exert additional influence through both types of edges. This setting is motivated by cases where people can reach a more active/influential state only in one network, or one relationship type. For example, some people may be reluctant to express their opinions freely in person (e.g., physical networks), but may be much more active on online networks (e.g., Twitter) due to anonymity. This raises an interesting question: which network or edge type would facilitate the influence propagation process most when hyper-influencers are allowed there. In what follows, we conduct several experiments to answer this question: 1) we only allow hyper-activity in red edges, i.e., hyper-active neighbors connected by blue edges will be counted as merely active when checking the response function; 2) we only allow hyper-activity in blue edges. More importantly, we conduct these experiments on a network with *low* assortativity and then a network with *high* assortativity to see if assortativity has any impact on the answer to the above question. Assortativity is defined as the Pearson correlation coefficient between the degree of nodes that are connected by a link [30]. If a network is assortative, then nodes of high degree in the network tend to attach to high degree nodes; it was noted in [30] that social networks tend to have high assortativity.

In the following experiments, we use the degree distributions (23) and (24) to assign red and blue degrees. However, to be able to control the assortativity of networks, we set $\alpha\lambda_r = \lambda_b$ rather than $\lambda_r = \lambda_b$. With this setting, when α is large, e.g., 0.99, nearly all of the nodes will have a similar number of red and blue edges, which leads to networks with limited assortativity. On the contrary, when α is *low*, e.g., 0.1, only 10% of the nodes will

have extra red edges. In addition, these nodes will have a significantly larger number of edges, since λ_r is ten times larger than λ_b . The nodes with extra red edges will tend to be connected together, which results in the network to have *high* assortativity. A more detailed discussion on this can be found in [29].

We start with the limited assortativity case, i.e., $\alpha = 0.99$. As shown in Figure 9, we observe that regardless of which network hyper-influencers are constrained to exist, there are two phase transitions as in the case of single-stage complex contagions. However, we see that the existence of hyper-influencers delays the second phase transitions to higher mean degrees. The reason behind this delay can be explained as follows. As mentioned before, the second phase transition occurs due to *high* local stability of nodes making their states hard to change by only few active neighbors. However, hyper-influencers help increase the value of the perceived influence, i.e., $\frac{c(m_{r,1} + \beta m_{r,2}) + m_{b,1} + \beta m_{b,2}}{cd_r + db}$, so that the response function could be exceeded even with few active and hyper-active neighbors, in the high mean degree region. Besides, allowing hyper-activity in blue edges leads to a larger region where global cascades take place, in comparison with the case where hyper-activity exists only in red edges. This can be explained as follows. When $\alpha = 0.99$, there are more nodes connected by blue edges in the network than red edges. That is, the impact of blue edges on impeding global cascades is more than that of red edges. Thus, allowing hyper-influence to be exerted in blue edges delays the second phase transition further.

Next, we discuss the case where $\alpha = 0.1$ that leads to a highly assortative network [29]. In Figure 10, we present numerical results for the first setting where the hyper-active state is manifested in only red edges. When $\beta = 1$, i.e., when there are no hyper-influencers in the network, four phase transitions take place. However, if we increase β from one to three, then we only observe two phase transitions. This can be explained as follows. When $\beta = 1$, multi-stage complex contagions is reduced to single-stage complex contagions, in which case four phase transitions might occur when assortativity is high [29]. As explained in [29], the first pair of phase transitions are mainly due to the red edges. When λ_b is *small*, there are too few blue edges to trigger a global cascade. However, since we have $\lambda_r = 10\lambda_b$, there are still enough red edges to have global cascades. As we increase λ_b , we observe a parameter interval where red edges are too many while blue edges are too few to have a global cascade. If we keep increasing λ_b further, global cascades start appearing again when the network has enough connectivity in blue edges to propagate the influence. However, further increasing in λ_b leads to high local stability of nodes w.r.t. both blue and red edges and global cascades become impossible again. A more detailed discussion can be found in [29].

The reason why increasing β changes the number of phase transitions is as follows. From the definition (4) of the response function, we observe that it is monotonically increasing with respect to β . Thus, when β is higher, an

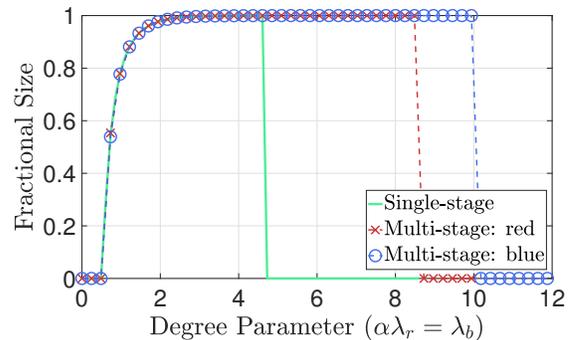


FIG. 9. Hyper-activity only appears in either red or blue edges. We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, and vary the mean degree. When $\alpha = 0.99$, the assortativity is negligible.

inactive node is easier to be activated by a hyper-active node, which makes it possible to have global cascades at higher levels of connectivity; i.e., the second phase transition tends to appear at larger λ . This leads to the second and the third phase transitions seen in Figure 10 when $\beta = 0$ disappear; i.e., the interval where we have too many red and too few blue edges disappears.

Next, we focus on the second setting where hyper-activity is only manifested in blue edges. The results are shown in Figure 11. Allowing hyper-activity in blue edges does not change the connectivity of the network, so the first and the second phase transitions caused by the connectivity w.r.t. red edges remain the same. However, the gap between the second and third transitions still exists. The gap happens between the second transition w.r.t. red and the first transition w.r.t. blue edges. A high β only shifts the second transition to the right but does not affect the first transition much. Thus, the gap disappears quickly with increasing β when we allow it in red edges, but remains when we only allow it in blue edges. Besides, compared with the case $\beta = 1$, the fourth transition is significantly delayed when $\beta = 3$. The reason behind the delay of the fourth phase transition is similar to the previous discussion: A higher β makes it easier to exceed the threshold even when the degree parameter is at a high level, so the original fourth phase transition has been extended to a larger mean degree.

From these experiments, we conclude that depending on the assortativity of the network, the impact of hyper-activity in red or blue edges on complex contagions are different: when the network is highly assortative, the additional influence exerted by the hyper-active nodes may change not only the critical transition points, but also the number and order of phase transitions, while for networks that have little or no assortativity, the additional influence mainly enlarge global cascade regions.

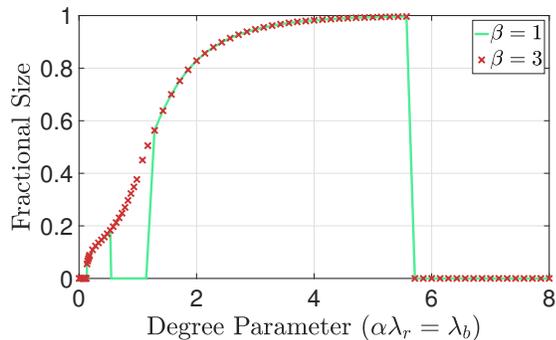


FIG. 10. *Hyper-activity only appears in red edges.* We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, and vary the mean degree. When $\alpha = 0.1$, the assortativity of the network is around 0.8.

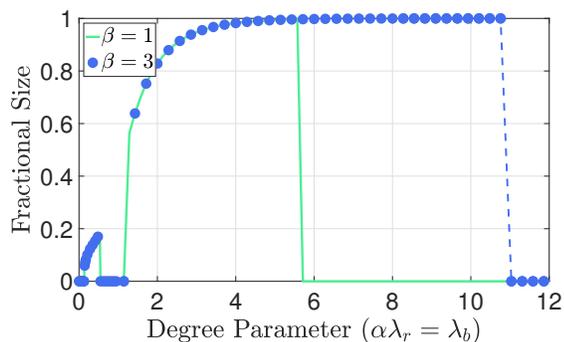


FIG. 11. *Hyper-activity only appears in blue edges.* We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, then vary the mean degree. When $\alpha = 0.1$, the assortativity is high (be up to 0.8).

V. CONCLUSION AND FUTURE WORK

In this work, we study the propagation of influence in multiplex networks under a *multi-stage* complex contagion model. We derive recursive relations characterizing the dynamics of influence propagation, and compute the probability of triggering *global cascades* and the expected size of *global cascades*, i.e., cases where a single individual can initiate a propagation that eventually influences a positive fraction of the population. The analytic results are also confirmed and supported by an numerical study. In particular, we demonstrate how the additional influence exerted by the hyper-active nodes can enlarge the network parameter region where global cascades take place. An interesting finding is that depending on the assortativity of the network, the existence of hyper-influencers affect the expected size of global cascades differently. For instance, when the network is highly assortative, the additional influence exerted by the hyper-active nodes may change not only the critical transition points, but also the number and order of phase transitions; while the affect is much more limited in networks with low assortativity.

There are many interesting directions to pursue for future work. First, it might be interesting to extend this work to more general network models than the configuration model used here. For instance, it would be interesting to consider networks that have high clustering. It would also be interesting to study multi-stage complex contagions using non-linear threshold models. Finally, it would be interesting to consider the case where it is possible for a node to transition back to the inactive state after being activated, e.g., due to the *negative* influence received by several *hyper-inactive* neighbors.

-
- [1] D. J. Watts, “A simple model of global cascades on random networks,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 5766–5771, 2002.
 - [2] P. S. Dodds and J. L. Payne, “Analysis of a threshold model of social contagion on degree-correlated networks,” *Physical Review E*, vol. 79, no. 6, p. 066115, 2009.
 - [3] O. Yağan, D. Qian, J. Zhang, and D. Cochran, “Conjoining speeds up information diffusion in overlaying social-physical networks,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1038–1048, 2013.
 - [4] O. Yağan and V. Gligor, “Analysis of complex contagions in random multiplex networks,” *Physical Review E*, vol. 86, no. 3, p. 036103, 2012.
 - [5] W. Duan, Z. Chen, Z. Liu, and W. Jin, “Efficient target strategies for contagion in scale-free networks,” *Physical Review E*, vol. 72, pp. 026 133–, 2005.
 - [6] D. Qian, O. Yağan, L. Yang, and J. Zhang, “Diffusion of real-time information in social-physical networks,” in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 2072–2077.
 - [7] Y. Zhuang and O. Yağan, “Information propagation in clustered multilayer networks,” *IEEE Transactions on Network Science and Engineering*, vol. 3, no. 4, pp. 211–224, 2016.
 - [8] R. M. Anderson and R. M. May, *Infectious Diseases of Humans*. Oxford (UK): Oxford University Press, 1991.
 - [9] M. E. Newman, “Spread of epidemic disease on networks,” *Physical Review E*, vol. 66, no. 1, p. 016128, 2002.
 - [10] P. E. Paré, J. Liu, C. L. Beck, A. Nedić, and T. Başar, “Multi-competitive viruses over static and time-varying networks,” in *American Control Conference (ACC)*, May 2017, pp. 1685–1690.
 - [11] J. Liu, P. E. Paré, A. Nedić, and T. Başar, “On a continuous-time multi-group bi-virus model with human awareness,” in *IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 4124–4129.
 - [12] P. E. Paré, C. L. Beck, and A. Nedić, “Stability analysis and control of virus spread over time-varying networks,” in *IEEE 54th Annual Conference on Decision and Control (CDC)*, 2015, pp. 3554–3559.
 - [13] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, “Catastrophic cascade of failures in interdependent networks,” *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

- [14] A. Vespignani, “Complex networks: The fragility of interdependency,” *Nature*, vol. 464, pp. 984–985, 2010.
- [15] O. Yağan, D. Qian, J. Zhang, and D. Cochran, “Optimal Allocation of Interconnecting Links in Cyber-Physical Systems: Interdependence, Cascading Failures and Robustness,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012,
- [16] T. R. Hurd and J. P. Gleeson, “A framework for analyzing contagion in banking networks,” *Available at SSRN 1945748*, 2011.
- [17] T. R. Hurd, D. Cellai, S. Melnik, and Q. H. Shao, “Double cascade model of financial crises,” *International Journal of Theoretical and Applied Finance*, vol. 19, no. 05, p. 1650041, 2016.
- [18] M. E. J. Newman, S. Forrest, and J. Balthrop, “Email networks and the spread of computer viruses,” *Physical Review E*, vol. 66, no. 3, 2002.
- [19] J. Balthrop, S. Forrest, M. E. J. Newman, and M. W. Williamson, “Technological networks and the spread of computer viruses,” *Science*, vol. 304, pp. 527–529, 2004.
- [20] S. Melnik, J. A. Ward, J. P. Gleeson, and M. A. Porter, “Multi-stage complex contagions,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 23, no. 1, p. 013124, 2013.
- [21] M. E. Newman, S. H. Strogatz, and D. J. Watts, “Random graphs with arbitrary degree distributions and their applications,” *Physical Review E*, vol. 64, no. 2, p. 026118, 2001.
- [22] M. Molloy and B. Reed, “A critical point for random graphs with a given degree sequence,” *Random structures & algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995.
- [23] B. Bollobás, W. Fulton, A. Katok *et al.*, “Cambridge studies in advanced mathematics,” in *Random graphs*. Cambridge University Press New York, 2001.
- [24] B. Söderberg, “Random graphs with hidden color,” *Physical Review E*, vol. 68, no. 1, p. 015102, 2003.
- [25] J. P. Gleeson and D. J. Cahalane, “Seed size strongly affects cascades on random networks,” *Physical Review E*, vol. 75, no. 5, p. 056103, 2007.
- [26] J. P. Sethna, K. Dahmen, S. Kartha, J. A. Krumhansl, B. W. Roberts, and J. D. Shore, “Hysteresis and hierarchies: Dynamics of disorder-driven first-order phase transformations,” *Physical Review Letters*, vol. 70, no. 21, p. 3347, 1993.
- [27] J. P. Gleeson, “Cascades on correlated and modular random networks,” *Physical Review E*, vol. 77, no. 4, p. 046117, 2008.
- [28] H. S. Wilf, *Generating functionology*. Elsevier, 2013.
- [29] Y. Zhuang, A. Arenas, and O. Yağan, “Clustering determines the dynamics of complex contagions in multiplex networks,” *Physical Review E*, vol. 95, p. 012312, Jan 2017.
- [30] M. E. Newman, “Assortative mixing in networks,” *Physical Review Letters*, vol. 89, no. 20, p. 208701, 2002.