Multistage Complex Contagions in Random Multiplex Networks

Yong Zhuang, *Student Member, IEEE*, and Osman Yağan, *Member, IEEE*

***Abstract*—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, coworker, family, etc.). To this end, this paper studies *multistage* complex contagions that take place over multilayer or multiplex networks. Under a linear threshold based contagion model, we first give analytic results for the expected size of *global* cascades, that is, cases where a randomly chosen node can initiate a propagation that eventually reaches a *positive* fraction of the whole population. Next, we analytically derive the probability of triggering global cascades. Then, analytic results are confirmed and supported by an extensive numerical study. In addition, we demonstrate how the dynamics of complex contagions is affected by the extra weight exerted by *hyperactive* nodes and by the structural properties of the networks. In particular, we reveal an interesting connection between the assortativity of a network and the impact of *hyperactive* nodes on the cascade size.**

***Index Terms*—Complex networks, computational modeling, communication networks, facebook, network topology, numerical models, sociology.**

I. Introduction

MODELING and analysis of dynamical processes in complex networks have 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]–[5]; disease contagion in human and animal populations [6]–[11]; cascading failures in *interdependent* infrastructures [12]–[14]; *insolvency* and *default* cascades in financial networks [15]; and the spread of computer viruses or worms on the Web [16], [17].

In this paper, 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, that is, when *multiple* sources of exposure are 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, proposed by Watts [1], considers binary-state dynamics where each node is in one of the two states, *inactive* or *active*, and is initially assigned a threshold τ in $(0, 1]$. 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 preassigned threshold τ. If $\frac{m}{d} \geq \tau$, then the node will turn *active*. Otherwise, if $\frac{m}{d} < \tau$, it stays *inactive*. Once a node is activated, it is assumed to remain active forever.

In the Watts threshold model, there are two important assumptions. The first assumption is that all active individuals exhibit the same amount of influence on their neighbors. However, individuals' standings on a given matter could vary significantly. For example, followers of a radical organization or a revolutionary movement may have varying levels of commitment to the cause, or a varying desire and ability to recruit new members. To cope with the multistate nature of individuals' activity levels, Melnik *et al.* [18] introduced a *multistage* contagion model as a generalization of the Watts threshold model. There, nodes can be inactive or in one of several active states with different levels of influence (e.g., active, hyperactive, etc).

The second assumption of the Watts model, which is also used in the multistage model by Melnik *et al.* [18], is that

Manuscript received March 13, 2019; revised March 14, 2019 and April 8, 2019; accepted May 25, 2019. Date of publication June 17, 2019; date of current version March 18, 2020. This work was supported in part by the US National Foundation by grant CCF-1813637 and in part by Army Research Office under Grant W911NF-17-1-0587. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. Recommended by Associate Editor G. Russo. *(Corresponding author: Yong Zhuang.)*

The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: yongzhua@andrew.cmu.edu; oyagan@andrew.cmu.edu).

Digital Object Identifier 10.1109/TCNS.2019.2923388

2325-5870 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

contagion is taking place over a *single* network where all edges have the same impact on spreading the influence. However, 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, etc.), and each relationship type might have a different impact on the propagation of influence in a given context. For example, video games might be more likely to spread among high-school friends rather than parents, whereas the opposite might be true for political ideas. That is, if we do not distinguish different types of relationships, dynamics of influence propagation may not be accurately captured. Hence, it is natural to consider complex contagions over multiplex networks. With this motivation, Yağan and Gligor [4] proposed and studied a threshold model in *multiplex* networks. However, they [4] still used the first assumption mentioned above in the sense that their model is not suitable for multistage contagions (where nodes can belong to a rich set of states).

In this paper, we drop both of the aforementioned assumptions and study for the first time a multistage contagion model on multiplex networks. For simplicity, we assume that the network consists of two types of links, red and blue, and individuals can be in one of three possible states, inactive, active, and hyperactive. Then, we seek to answer several important questions: *In the cases where a global spreading event is possible, could we give analytic answers to the exact probability and the condition of its taking place, as well as the final expected cascade size? Under the newly proposed model, how do hyperinfluencers and topological properties affect the cascade process?* Our contributions toward answering these questions are summarized as follows.

1) For a class of random networks generated by the *colored* configuration model (see Section II-A for precise definitions), we analytically derive the expected size of *global* cascades, that is, cases where a positive fraction of nodes (in the asymptotic limit) eventually becomes active or hyperactive when a randomly selected node is switched to the active state.
2) We are the first to give analytic results for the probability of triggering global cascades for multistage models.
3) We explore the intricate relationships between the structural properties of the underlying network and the impact of *hyperactive* nodes on the contagion dynamics. For instance, a particularly interesting scenario is when the hyperactive 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 can reach a hyperinfluential state only in one social network (e.g., Twitter) versus another (e.g., Facebook). Our main finding is the interesting connection between the assortativity (i.e., the correlation among the degree of neighboring nodes) of the network and the impact of hyperactive nodes on cascade size. For instance, when the network is highly assortative (i.e., when high-degree nodes are likely to be connected with high-degree nodes), the influence exerted by the hyperactive nodes has a much more significant impact on the cascade size as opposed to the case when the network has low assortativity (i.e., when the degrees of neighboring nodes tend to be uncorrelated). This impact is best observed when the cascade size is plotted as a function of the mean degree of the network. There, as the influence of hyperactive nodes increases, the highly assortative networks are shown to exhibit changes on not only the critical transition points (i.e., mean degree values at which expected cascade size changes from zero to a positive value, or vice versa), but also the number and *order* of transitions.
4) Considering an important property in real-world networks, clustering, we analytically derive the expected size of *global* cascades for random networks with clustering. Due to the page limit, we include all derivations and results in supplementary materials. Supplementary materials are at https://goo.gl/1nTgQS.

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 hyperinfluencers on complex contagions under different levels of assortativity. Further details about the impact of hyperinfluencers are given in the Appendix. We conclude the paper in Section V where we also suggest several directions for future work.

II. Model Definition: Networks and Dynamics

A. Multilayer and Multiplex Network Models

A multiplex network is a network model where links are classified into different types (or colors), which can capture the different types of connections between nodes in networks. 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, whereas the other for edges in Instagram. Alternatively, one link color may be representing close friendship links, whereas the other representing "acquaintances" in a social network. In this network model, we let $\mathcal{N} = \{1, 2, \ldots, n\}$ 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 $\alpha \in (0, 1]$

$$\mathbb{P}[i \in \mathcal{N}_r] = \alpha, \ i = 1, \ldots, n. \tag{1}$$

With this assumption, by the law of large numbers, we can easily conclude that $\frac{|\mathcal{N}_r|}{n} \xrightarrow{\text{a.s.}} \alpha$, where $|\mathcal{N}_r|$ denotes the cardinality of $\mathcal{N}_r$ and $\xrightarrow{\text{a.s.}}$ indicates almost sure convergence.

This network model can be interpreted in two different ways. The first one is a multilayer network where each network layer

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

Fig. 1. Illustration of a multilayer and a multiplex network representation of our model. (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. (b) Equivalent representation of this model by a multiplex network. Edges from the online network are red, whereas edges from the physical network are blue.

is generated by the widely used configuration model [19]–[21]; this case is illustrated in Fig. 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 [19] and [20] 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 [22]. 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 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 [22] for details. Fig. 1(b) is an illustration of this multiplex network model.

B. Multistage Content-Dependent Linear Threshold Model

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 rumor, product, etc.), scalars $c_i, i = 1, \ldots, r$, represent the weight 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, $i = 1, \ldots, r$, 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 $\mathbf{m} = (m_1, \ldots, m_r)$ and $\mathbf{d} = (d_1, \ldots, d_r)$ will turn *active* with probability

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

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 $i = 1, \ldots, r$, then this model reduces to the Watts threshold model [1]. The content-dependent threshold model enables modeling 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 be only in two states, the multistage linear threshold model [18] allows nodes to be in a richer set of active states. In this paper, we assume that nodes can belong to three states: inactive, active, and hyperactive. In the following discussion, we use state-0, state-1, and state-2 to represent the inactive, active, and hyperactive state, respectively. Let τ_1 and τ_2 denote the thresholds associated with transitioning to the active and hyperactive states, respectively. The hyperactive 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 hyperactive, the probability of switching to state-i from the inactive state (i.e., state-0) is given by

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

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 multistage content-dependent linear threshold model. Assume that there are two types of links in the network, red and blue, and that nodes can be in three states, inactive, active, and hyperactive. 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

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

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 hyperactive neighbors connected through a red (resp. blue) edge, and d_r and d_b denote the number of red and blue neighbors, respectively. Assume that all nodes are initially inactive and the contagion process starts by randomly choosing a node and setting it as active. The influence might then propagate in the network according to (4) and other nodes might turn active, and so on. More precisely, we consider a discrete-time process, which is an approximation to study the continuous-time Markov process model and has been applied in many works [23], [24]. In discrete-time approaches, we discretize time into uniform time steps of length $\Delta t = 1$. Then, nodes update their states

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

synchronously at the discrete time steps $t = 0, 1, \ldots$, and an inactive node will be activated at time t according to the influence they receive at time $t-1$ and formula (4). Once active, a node cannot be deactivated. Since we are only interested in the fraction of nodes in active or hyperactive state in the *steady state* (i.e., when nodes no longer change their states), our results do not depend on the length of the discrete time steps, that is, we can use arbitrarily small time steps and get arbitrarily close to continuous-time dynamics.

III. Main Results

Since the proposed 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. A *global* cascade is said to take place if the fraction of nodes that are activated is *positive* in the limit of large network sizes. Our main goals are as follows:

1) determining the conditions (in terms of network parameters) for global cascades to be possible;
2) calculating the expected size of global cascades when they are possible;
3) 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

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

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 $n \to \infty$, and our main goal is to derive

$$\lim_{n\to\infty} \mathbb{E}\left[S | S > 0\right]$$

which gives the expected size of global cascades when they exist. For simplicity, in our analysis we omit *self-loops*, that is, the possibility of having more than one edge between two nodes. It is a simple matter to show that such self-loops occur very rarely in the construction of the configuration model and they have negligible impact on the cascade dynamics (e.g., see [25]). In fact, our experiments also confirm that the impact of this omission is negligible.

According to our definition, the expected cascade size stands for the final fraction of active and hyperactive individuals in the network. Therefore, we can compute it by computing the probability that an arbitrary node is active or hyperactive at the steady state. We will compute this probability recursively using the "tree-approximation" approach [4], [18], [26], which is a mean-field treatment of the zero-temperature random-field Ising model on Bethe lattices [27]. The tree-approximation approach assumes that the network has a locally treelike structure, which is valid under the configuration model considered here [19]. In supplementary materials, we show how the results can be extended to a broader class of networks that have large clustering coefficient (i.e., a *large* number of triangles). Clustering is a common property of many real-life social networks and capture the fact that a pair of individuals are more likely to be connected with each other if they have a common friend as compared to the case where no common friend exists.

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 levels $0, 1, \ldots, \ell-1$ have finished updating. We define $q_{r,1,\ell}$ (resp. $q_{b,1,\ell}$) as the probability that a node at level ℓ that is connected to its only parent at level $\ell+1$ by a red (resp. blue) edge turns active. Similarly, we define $q_{r,2,\ell}$ (respectively, $q_{b,2,\ell}$) as the probability that a node at level ℓ that is attached to its only parent via a red (resp. blue) edge turns hyperactive. Given our assumption that nodes in the tree update their states one level at a time, these probabilities will be computed under the condition that the parent nodes at level $\ell+1$ are inactive.

In the interest of brevity, we only explain the derivation of $q_{r,1,\ell+1}$ in details. The derivations of $q_{r,2,\ell+1}$, $q_{b,1,\ell+1}$, and $q_{b,2,\ell+1}$ can be explained very similarly. 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 expression (6) for $q_{r,1,\ell+1}$ can be explained as follows. Consider an inactive node at level $\ell+1$ with colored degree $\mathbf{d} = (d_r, d_b)$ that is connected to its unique parent at level $\ell+2$ via a red edge. The probability that this node has i active children connected via red edges, s active children connected via blue edges, j hyperactive children connected via red edges, and t hyperactive children connected via blue edges, *and* that it turns active is given by

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

where $F_1[(i,j,s,t), \mathbf{d}]$ is as defined in (4), that is, it denotes the probability that an inactive node with a colored degree $\mathbf{d}$ and a group of active and hyperactive 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 in (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 treelike, 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 $\ell+2$ cannot 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] and [19] for a discussion on the

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

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)

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

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

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

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

where for $k=1,2$, we define (10), shown at the bottom of this page.

Equations (6)–(9) form a nonlinear system. Since our goal is to compute the expected size of global cascades *given that they exist*, we can initialize this dynamical system with $q_{r,1,0}, q_{r,2,0}, q_{b,1,0}, q_{b,2,0} > 0$ to obtain the fixed points, $q_{r,1,\infty}$, $q_{r,2,\infty}$, $q_{b,1,\infty}$, and $q_{b,2,\infty}$. These fixed points account for the probability of being in a corresponding state for the children of the node chosen uniformly at random. We can use them to calculate the expected size of global cascades.

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 probability that each of its d_r neighbors (via red links) is active (resp. hyperactive) 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 hyperactive with probability $q_{b,2,\infty}$, independently from each other. Then, with each possible combination of numbers of active and hyperactive neighbors, we can calculate the probability of being active or hyperactive 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], [18], and [26], this method, based on the *tree-approximation* technique, gives precise results in the asymptotic limit $n\to\infty$, when the underlying network is generated according to the configuration model. We present extensive numerical studies in Section IV that supports our results even in the finite node regime, (11) shown at the bottom of this page.

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 $n\to\infty$. 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

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

By direct inspection, we see that the recursive equations (12) have a trivial fixed point $q_1=q_2=q_3=q_4=0$, 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, when (12) has only a single solution given by this trivial fixed point, then with probability one global cascades do not take place. In general, the trivial fixed point may not be stable and there may exist *nontrivial* fixed points that yield $\mathbb{E}[S]>0$. In that case, we have $S>0$ with a positive probability in the limit $n\to\infty$ and hence global cascade may take place. To check the existence of nontrivial solutions of (6)–(9), we linearize them at $q_1=q_2=q_3=q_4=0$, which yields the Jacobian matrix $\mathbf{J}$ given as

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

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 exists a nontrivial 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.

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] and [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

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

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

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

vulnerable nodes in the network; a node is deemed *vulnerable* if it can be activated by only one *active* neighbor [1], [2], [4], [28]. 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 $n \to \infty$. 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), which 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.

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 fraction of nodes identified by the branching process described above can be analyzed using the method of probability generating functions [29] (e.g., see [1], [4], [6], [19], and [30]), 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], [30], and [31], we exclude the initially activated nodes. We have

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

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 Fig. 2. The validity of expression (14) can be seen as follows. First, we initially activate a node that 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 hyperactive 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.

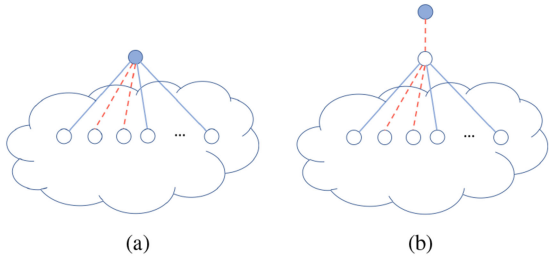


Fig. 2. Difference between (a) $G(x)$ and (b) $g_{r,1}(x)$. Red dashed lines account for red edges in our analysis, whereas other lines represent blue edges. Circles with solid fill indicate active nodes, whereas circles without fill account for inactive nodes (that can potentially be made active). Then,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

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 $\mathbf{d} = (d_r, d_b)$.

1) It turns *active*, that is, $\tau_1 \leq \frac{c}{cd_r + d_b} < \tau_2$, which happens with probability $F_1\left[(1,0,0,0),(d_r,d_b)\right]$. 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.
2) It turns *hyperactive*, that is, $\frac{c}{cd_r + d_b} \geq \tau_2$, which happens with probability $F_2\left[(1,0,0,0),(d_r,d_b)\right]$. 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.
3) It remains inactive, that is, $\frac{c}{cd_r + d_b} < \tau_1$, which happens with probability $1 - F_1\left[(1,0,0,0),(d_r,d_b)\right] - F_2\left[(1,0,0,0),(d_r,d_b)\right]$. 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 as in (15)–(18), shown at the bottom of the next page.

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*

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

forever and will not go extinct, meaning that the underlying influence propagation process will constitute a global cascade. Thus, we have

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

This approach has been introduced in [1] and used in [4] and [19] 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

$$g_i = h_i(g_1, g_2, g_3, g_4), \qquad i = 1, 2, 3, 4. \tag{20}$$

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

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

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

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

IV. Numerical Results

In this section, we present numerical results to support our analysis on the probability and expected size of global cascades. We are particularly interested in checking the accuracy of our asymptotic results when the number of nodes is finite. We will also investigate the impact of hyperinfluencers (i.e., the additional influence exerted by them) on the contagion dynamics. In particular, we aim at exploring if structural properties of networks will change the impact of hyperinfluencers. Besides, more details about the impact of hyperinfluencers are given in the Appendix.

A. Agreement Between Our Analysis and Simulations

We focus on demonstrating the accuracy of our analytic results on the expected size of 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 (respectively, p_k^b) denoting the probability that a node is assigned k red (respectively, blue) edges, we let

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

$$p_k^r = \alpha e^{-\lambda_r} \frac{(\lambda_r)^k}{k!} + (1-\alpha)\delta_{k,0}, \ \mathrm{k = 0, 1, \ldots}. \tag{24}$$

Here, λ_r (respectively, 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[1] 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 hyperactive nodes, respectively, and fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$. Then, for several values of $\lambda_r = \lambda_b$, we run 1000 independent experiments (for each parameter set), each time computing the fraction of nodes that eventually turn active or hyperactive. The results

[1]To avoid the finite size effect, we use $n = 2 \times 10^6$ as the number of nodes around the second phase transition in the simulations.

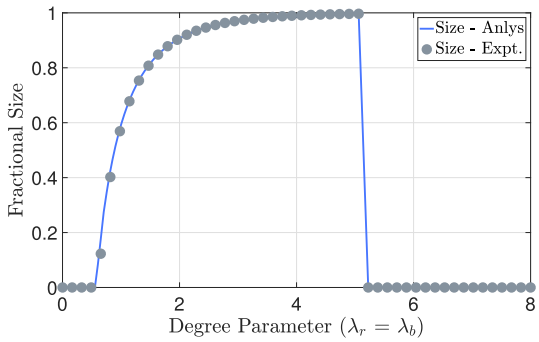
$$g_{r,1}(x) = x \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle d_r \rangle} \left[F_1\left[(1,0,0,0),(d_r,d_b)\right] g_{r,1}(x)^{d_r-1} g_{b,1}(x)^{d_b} + F_2\left[(1,0,0,0),(d_r,d_b)\right] g_{r,2}(x)^{d_r-1} g_{b,2}(x)^{d_b} \right]$$
$$+ x^0 \sum_{\mathbf{d}} \frac{d_r p_{\mathbf{d}}}{\langle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

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

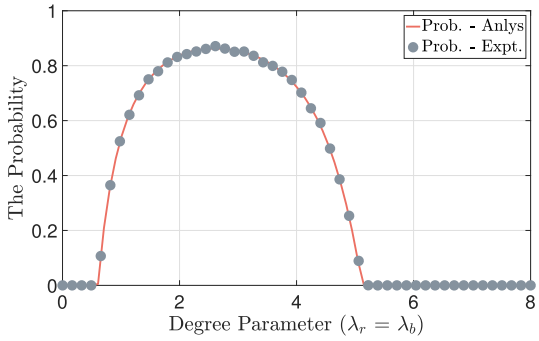
$$g_{b,1}(x) = x \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle d_b \rangle} \left[F_1\left[(0,0,1,0),(d_r,d_b)\right] g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b-1} + F_2\left[(0,0,1,0),(d_r,d_b)\right] g_{r,2}(x)^{d_r} g_{b,2}(x)^{d_b-1} \right]$$
$$+ x^0 \sum_{\mathbf{d}} \frac{d_b p_{\mathbf{d}}}{\langle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

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

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.



(a)



(b)

Fig. 3. Simulations for doubly Poisson degree distributions, $\alpha = 0.5$, $\tau_1 = 0.18$, and $\tau_2 = 0.32$. The weight of hyperinfluencers is taken to be $\beta = 1.5$. (a) Expected size of global cascades. (b) Probability of the emergence of global cascades.

are depicted in Fig. 3(a), where lines represent analytic results obtained from (6) to (9), and symbols represent the average cascade size obtained in simulations (over 1000 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 1000 to 10 000. As shown in Fig. 3(b), 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 Fig. 3(a) and (b) 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 cannot occur when nodes are locally stable, that is, 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 hyperinfluencers on complex contagion dynamics in the following section.

B. Impact of Hyperinfluencers in Multiplex Networks

In this section, we investigate more closely how hyperinfluencers affect the complex contagions. We consider a case where hyperactive nodes are restricted to appear 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 hyperinfluencers are allowed there? In what follows, we conduct several experiments to answer this question: 1) we only allow hyperactivity in red edges, that is, hyperactive neighbors connected by blue edges will be counted as merely active when checking the response function; and 2) we only allow hyperactivity in blue edges. As discussed in some previous studies [30]–[32], assortativity is one of the most important structural properties on multiplex networks. Assortativity is defined as the Pearson correlation coefficient between the degree of nodes that are connected by a link [33]. If a network is assortative, then nodes of high degree in the network tend to attach to high degree nodes; it was noted in [33] that social networks tend to have high assortativity. Therefore, it is interesting to see if assortativity has any impact on the answer to the above question.

In the following experiments, we conduct these experiments on a network with *low* assortativity and then a network with *high* assortativity. We use the degree distributions (23) and (24) to assign red and blue degrees. To be able to control the assortativity of networks without changing the first moment of degree, we set $\alpha\lambda_r = \lambda_b$ rather than $\lambda_r = \lambda_b$. With this setting, when α is large, for example, 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*, for example, 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 [31].

We start with the limited assortativity case, that is, $\alpha = 0.99$. As shown in Fig. 4, we observe that regardless of which network hyperinfluencers 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 hyperinfluencers 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

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

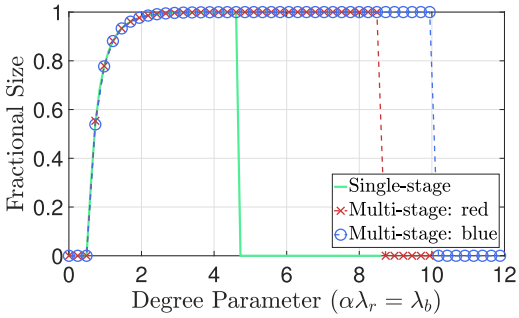


Fig. 4. Hyperactivity only appears in either red or blue edges. We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, and vary the mean degree. When $\alpha = 0.99$, the assortativity is negligible.

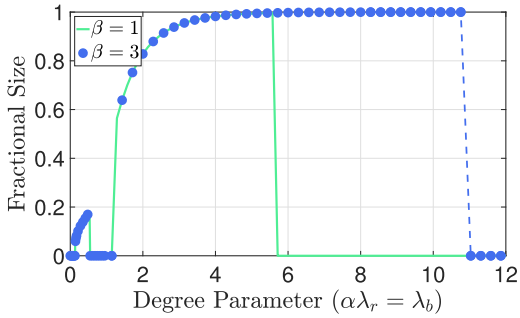


Fig. 6. Hyperactivity only appears in blue edges. We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, and vary the mean degree. When $\alpha = 0.1$, the assortativity is high (be up to 0.8).

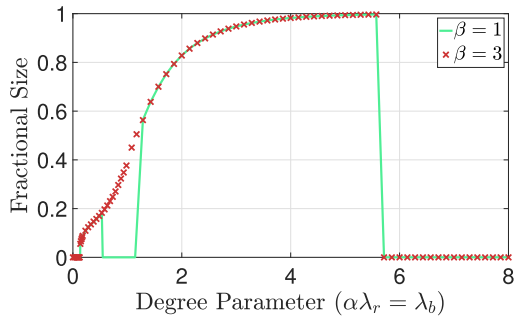


Fig. 5. Hyperactivity only appears in red edges. We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, and vary the mean degree. When $\alpha = 0.1$, the assortativity of the network is around 0.8.

of nodes making their states hard to change by only few active neighbors. However, hyperinfluencers help increase the value of the perceived influence, that is, $\frac{c(m_{r,1}+\beta m_{r,2})+m_{b,1}+\beta m_{b,2}}{cd_r+d_b}$, so that the response function could be exceeded even with few active and hyperactive neighbors, in the high mean degree region. Besides, allowing hyperactivity in blue edges leads to a larger region where global cascades take place, in comparison with the case where hyperactivity 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 hyperinfluence 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 [31]. In Fig. 5, we present numerical results for the first setting where the hyperactive state is manifested in only red edges. When $\beta = 1$, that is, when there are no hyperinfluencers in the network, four phase transitions take place. However, if we increase β from one to three, then only two phase transitions are observed. This can be explained as follows. When $\beta = 1$, multistage complex contagions are reduced to single-stage complex contagions, in which case four phase transitions might occur when assortativity is high [31]. As explained in [31], 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, whereas 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 [31].

The reason why increasing β changes the number of phase transitions is as follows. From definition (4) of the response function, we observe that it is monotonically increasing with respect to β. Thus, when β is higher, an inactive node is easier to be activated by a hyperactive node, which makes it possible to have global cascades at higher levels of connectivity, that is, the second phase transition tends to appear at larger λ. This leads to the second and the third phase transitions seen in Fig. 5 disappear when $\beta = 1$, that is, the interval where we have too many red and too few blue edges disappear.

Next, we focus on the second setting where hyperactivity is only manifested in blue edges. The results are shown in Fig. 6. Allowing hyperactivity 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 hyperactivity in red or blue edges on complex contagions is different: When the network is highly assortative, the additional influence exerted by

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

the hyperactive nodes may change not only the critical transition points, but also the number and order of phase transitions, whereas for networks that have little or no assortativity, the additional influence mainly enlarges global cascade regions.

V. Conclusion and Future Work

In this paper, we study the propagation of influence in multiplex networks under a *multistage* complex contagion model. We derive recursive relations characterizing the dynamics of influence propagation to compute the probability and expected size of *global* cascades, that is, 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 a numerical study. An interesting finding is that depending on the assortativity of the network, the existence of hyperinfluencers affects the expected size of global cascades differently. For instance, when the network is highly assortative, the additional influence exerted by the hyperactive nodes may change not only the critical transition points, but also the number and order of phase transitions, whereas the effect is much more limited in networks with low assortativity. In addition, we relaxed the assumption that the network topology has a treelike structure and added analysis of multistage complex contagions over networks with high clustering coefficient, which is an important property of real-world networks.

There are many interesting directions for future work. For example, it would also be interesting to study the case where it is possible for a node to transition back to the inactive state after being activated, for example, due to the *negative* influence received by several *hyperinactive* neighbors. It would also be of interest to study multistage complex contagions using nonlinear threshold models, or correlated propagation of multiple opinions over the same population.

Appendix A
Impact of Hyperinfluencers 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 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 hyperinfluencers on the global cascade boundary, and then move on to the discussion about the impact of the threshold τ_1 of ordinary influencers.

Fig. 7 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, that is, 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*, that is, the lower parts of the boundaries seen in Fig. 7 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 Fig. 3(a) and (b)] is decided by the likelihood of nodes with high degree being influenced by a single active or hyperactive neighbor. Thus, the boundary is determined from a node's perceived influence, or perceived proportion of active and hyperactive 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 hyperactive 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 more sensitive to the changes in β.

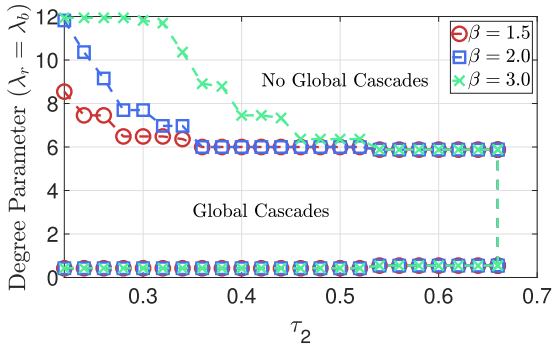


Fig. 7. Given $\tau_1 = 0.15$ and $\alpha = 0.5$, we vary the mean degree λ and τ_2 to plot the global cascade region for several β, 1.5, 2.0, and 3.0.

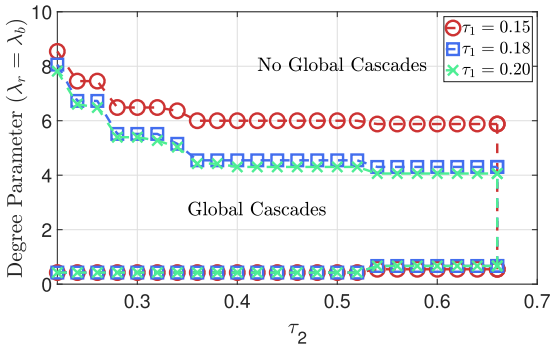


Fig. 8. Given the weight of extra influence $\beta = 1.5$ and $\alpha = 0.5$, we vary the degree parameter $\lambda = \lambda_b = \lambda_r$ and τ_2 to plot the region where there exists a global cascade for several τ_1, 0.15, 0.18, and 0.2. Both of the edges are assigned by the doubly Poisson distribution in Section IV-A.

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 Fig. 8, 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

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

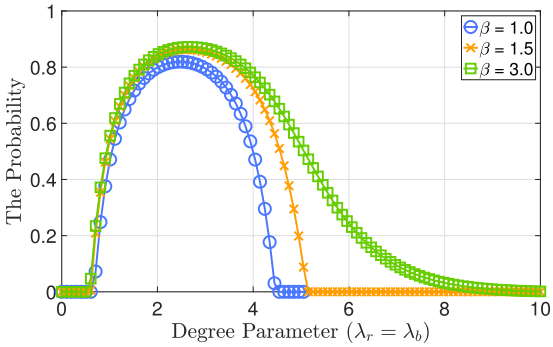


Fig. 9. Comparison between different β for the probability of triggering a *global* cascade.

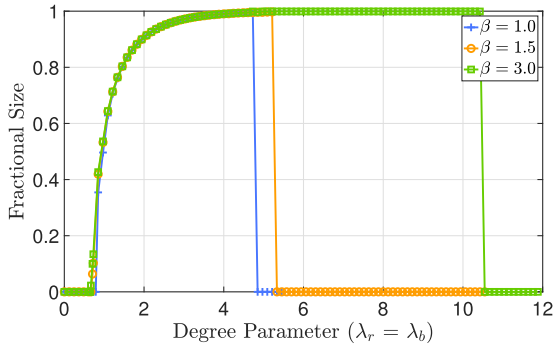


Fig. 10. Comparison between different β for the expected *global* cascade size.

(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.

Appendix B
Impact of Hyperinfluencers on the Probability and Expected Size of Global Cascades

We start by investigating the impact of the extra influence β (that hyperactive nodes exert on their neighbors) on the probability of global cascades. From Fig. 9, 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 hyperactive neighbor exceeds the threshold more easily, leading to a larger *vulnerable* component. Also, we see in Fig. 9 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 Fig. 8 and can be explained in a similar manner.

Next, we discuss the impact of hyperinfluencers on the expected size of global cascades. From Fig. 10, 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.

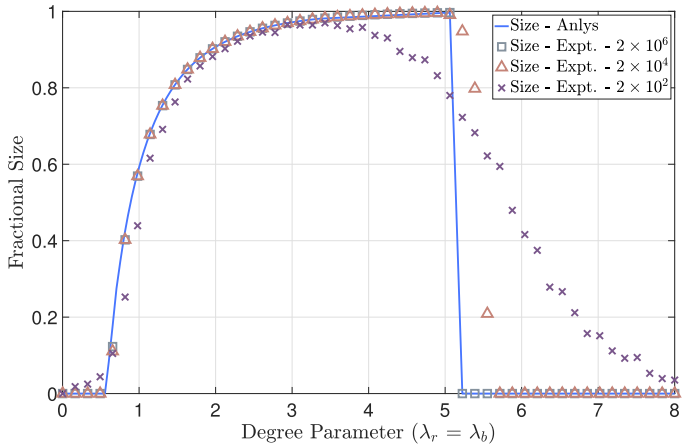


Fig. 11. Comparison between the asymptotic and the finite-size results for different size of networks.

Appendix C
Impact of the Network Size on the Comparison Between Analytic Results and Experiments

In this section, we explore the impact of network size on the comparison of analytic results and simulation experiments. In particular, we repeat the experiments conducted to obtain Fig. 3(a) and compare the asymptotic and the finite-size results for $n = 2 \times 10^2, 2 \times 10^4, 2 \times 10^6$.

The results are shown in Fig. 11, where we see that the differences between the asymptotic and the finite-size results are not significant when the mean degree is at a low level. However, when the mean degree is at a high level, simulation results are significantly different from the asymptotic results when the network size is *small*. In particular, when $n = 200$, the second phase transition observed at large mean degrees appears much later than suggested by the analysis. This transition also appears like a continuous one in the simulations with $n = 200$ although it is clearly a discontinuous transition in the analytic results. When $n = 20\,000$, we see that simulation results match the analysis a lot better than the case with $n = 200$ although the match is not nearly as perfect as the one seen when $n = 2\,000\,000$. In particular, we see significant discrepancies around the second phase transition point, though it is at least suggested correctly by

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.

the simulations that the transition is discontinuous. Overall, we can conclude that our asymptotic analysis yields almost perfect predictions for networks with a million nodes or more, and relatively accurate predictions for networks with 20 000–100 000 nodes. For smaller networks with only a few hundred nodes, the asymptotic results do not yield reliable predictions.

REFERENCES

[1] D. J. Watts, "A simple model of global cascades on random networks," *Proc. Nat. Acad. Sci. USA*, 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," *Phys. Rev. E*, vol. 79, no. 6, 2009, Art. no. 066115.

[3] O. Yağan, D. Qian, J. Zhang, and D. Cochran, "Conjoining speeds up information diffusion in overlaying social-physical networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1038–1048, Jun. 2013.

[4] O. Yağan and V. Gligor, "Analysis of complex contagions in random multiplex networks," *Phys. Rev. E*, vol. 86, no. 3, 2012, Art. no. 036103.

[5] D. Qian, O. Yağan, L. Yang, and J. Zhang, "Diffusion of real-time information in social-physical networks," in *Proc. IEEE Global Commun. Conf.*, Dec. 2012, pp. 2072–2077.

[6] M. E. Newman, "Spread of epidemic disease on networks," *Phys. Rev. E*, vol. 66, no. 1, 2002, Art. no. 016128.

[7] P. E. Paré, J. Liu, C. L. Beck, A. Nedić, and T. Başar, "Multi-competitive viruses over static and time-varying networks," in *Proc. Amer. Control Conf.*, May 2017, pp. 1685–1690.

[8] J. Liu, P. E. Paré, A. Nedić, and T. Başar, "On a continuous-time multigroup bi-virus model with human awareness," in *Proc. IEEE 56th Annu. Conf. Decis. Control*, Dec. 2017, pp. 4124–4129.

[9] F. D. Sahneh, C. Scoglio, and P. Van Mieghem, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1609–1620, Oct. 2013.

[10] P. E. Paré, C. L. Beck, and A. Nedić, "Stability analysis and control of virus spread over time-varying networks," in *Proc. IEEE 54th Annu. Conf. Decis. Control*, 2015, pp. 3554–3559.

[11] F. D. Sahneh, A. Vajdi, J. Melander, and C. M. Scoglio, "Contact adaption during epidemics: A multilayer network formulation approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 6, no. 1, pp. 16–30, Jan.–Mar. 2019.

[12] 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.

[13] A. Vespignani, "Complex networks: The fragility of interdependency," *Nature*, vol. 464, pp. 984–985, 2010.

[14] O. Yağan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures and robustness," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1708–1720, Sep. 2012.

[15] T. R. Hurd, D. Cellai, S. Melnik, and Q. H. Shao, "Double cascade model of financial crises," *Int. J. Theor. Appl. Finance*, vol. 19, no. 5, 2016, Art. no. 1650041.

[16] M. E. J. Newman, S. Forrest, and J. Balthrop, "Email networks and the spread of computer viruses," *Phys. Rev. E*, vol. 66, no. 3, 2002, Art. no. 035101.

[17] 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.

[18] S. Melnik, J. A. Ward, J. P. Gleeson, and M. A. Porter, "Multi-stage complex contagions," *Chaos, Interdisciplinary J. Nonlinear Sci.*, vol. 23, no. 1, 2013, Art. no. 013124.

[19] M. E. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Phys. Rev. E*, vol. 64, no. 2, 2001, Art. no. 026118.

[20] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random Struct. Algorithms*, vol. 6, no. 2/3, pp. 161–180, 1995.

[21] B. Bollobás *et al.*, "Cambridge studies in advanced mathematics," in *Random Graphs*. New York, NY, USA: Cambridge Univ. Press, 2001.

[22] B. Söderberg, "Random graphs with hidden color," *Phys. Rev. E*, vol. 68, no. 1, 2003, Art. no. 015102.

[23] C. Granell, S. Gómez, and A. Arenas, "Dynamical interplay between awareness and epidemic spreading in multiplex networks," *Phys. Rev. Lett.*, vol. 111, no. 12, 2013, Art. no. 128701.

[24] X. Wei, N. C. Valler, B. A. Prakash, I. Neamtiu, M. Faloutsos, and C. Faloutsos, "Competing memes propagation on networks: A network science perspective," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 6, pp. 1049–1060, Jun. 2013.

[25] M. Newman, *Networks: An Introduction*. London, U.K.: Oxford Univ. Press, 2018.

[26] J. P. Gleeson and D. J. Cahalane, "Seed size strongly affects cascades on random networks," *Phys. Rev. E*, vol. 75, no. 5, 2007, Art. no. 056103.

[27] 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," *Phys. Rev. Lett.*, vol. 70, no. 21, pp. 3347–3350, 1993.

[28] J. P. Gleeson, "Cascades on correlated and modular random networks," *Phys. Rev. E*, vol. 77, no. 4, 2008, Art. no. 046117.

[29] H. S. Wilf, *Generating Functionology*. New York, NY, USA: Elsevier, 2013.

[30] Y. Zhuang and O. Yağan, "Information propagation in clustered multilayer networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 3, no. 4, pp. 211–224, Oct.–Dec. 2016.

[31] Y. Zhuang, A. Arenas, and O. Yağan, "Clustering determines the dynamics of complex contagions in multiplex networks," *Phys. Rev. E*, vol. 95, Jan. 2017, Art. no. 012312.

[32] A. Hackett, D. Cellai, S. Gómez, A. Arenas, and J. P. Gleeson, "Bond percolation on multiplex networks," *Phys. Rev. X*, vol. 6, no. 2, 2016, Art. no. 021002.

[33] M. E. Newman, "Assortative mixing in networks," *Phys. Rev. Lett.*, vol. 89, no. 20, 2002, Art. no. 208701.

Yong Zhuang (S'12) received the B.S. degree in computer science from Zhejiang University, Hangzhou, China, in 2012, and the M.S. degree in computer science from National Taiwan University, Taipei City, Taiwan, in 2014. He is currently working toward the Ph.D. degree in network science in computer engineering at the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.

His research interests include network science, random graphs, and social and information networks.

Osman Yağan (S'07–M'12) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2007, and the Ph.D. degree in electrical and computer engineering from the University of Maryland–College Park, College Park, MD, USA, in 2011.

He was a Visiting Postdoctoral Scholar with Arizona State University during Fall 2011 and then a Postdoctoral Research Fellow with the Cyber Security Laboratory, Carnegie Mellon University, until 2013. In 2013, he joined the faculty of the Department of Electrical and Computer Engineering, Carnegie Mellon University, as an Assistant Research Professor. His research interests include wireless communications, security, random graphs, social and information networks, and cyber-physical systems.

Authorized licensed use limited to: Carnegie Mellon Libraries. Downloaded on October 02,2020 at 17:05:30 UTC from IEEE Xplore. Restrictions apply.