Correlated Social Contagions with Multiple Topics: A Generalized Linear Threshold Model

Yurun Tian and Osman Yağan
Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA, 15213 USA
{yurunt, oyagan}@andrew.cmu.edu

***Abstract*—The study of influence propagation over networks has received increasing attention in many scientific domains. In particular, the linear threshold model is widely studied due to its ability to capture the mechanism by which multiple sources of exposure are required for nodes in the network to take action. Most existing research on influence propagation concentrate on a *single* topic spreading in isolation over the network. However, real-life social contagions often involve *multiple* topics spreading simultaneously in a *correlated* manner; e.g., different conspiracy theories, political opinions on taxes, immigration, gun control, etc. In this work, we propose a *multi-dimensional threshold model with correlated influences* as an extension of the classical linear threshold model to incorporate multiple correlated topics spreading simultaneously over networks. We provide analytical results that accurately predict the threshold, probability, and expected size of *global* cascades, i.e., cases where a *significant* fraction of the population gets *influenced*. Through extensive simulations, we demonstrate that our analytical results match the numerical results near-perfectly in the finite node regime. These results reveal the interplay between the underlying network structure, the correlation among spreading topics, and the heterogeneous thresholds on the *final* results of the propagation.**

***Index Terms*—Influence propagation; Linear threshold model; Correlated influence; Complex networks.**

I. Introduction

THE mathematical modeling of spreading processes over networks has drawn increasing attention in different contexts, including cascading failures [1], social contagions [2], epidemics [3], traffic jams [4], risks in banking systems [5], and networks of spiking neurons [6]. These spreading processes are typically studied through two different phenomena [7]. *Simple* contagions, also referred to as *information propagation*, are used to model cases where a single source of exposure is enough for an individual to get *infected* and start spreading the topic to their contacts (e.g., news articles, disease spreading, etc.). *Complex* contagions are used to model spreading processes where multiple sources of exposure to a *topic* (e.g., an opinion, a product, a new behavior, a neuronal spike, etc.) are needed for nodes to change their states.

This paper focuses on analysis of complex contagions, also referred to as *influence propagation* [8], [9]. Examples include the rise of collective action to join riots, the diffusion of beliefs, political revolutions [10], adoption of new technologies [11], and cultural market sensations [12]. In addition, cascade of overload failures in critical infrastructure networks, transportation systems and cloud computing services, have sometimes been studied using complex contagion models [13]–[15].

The linear threshold model has been widely used in the literature for modeling complex contagions over networks [13]. In this model, each node can be in one of the two states, *active* or *inactive*. Initially, all of the nodes are inactive, and a small number of nodes are chosen at random to become active as *seeds*. An inactive node with degree k, where m of these k neighbors are active, is activated with probability

$$F[m,k] \triangleq P\left[\frac{m}{k} \geq \tau\right] \tag{1}$$

where $F[m,k]$ is referred to as the *response function*, and τ describes the smallest fraction of active neighbors for an inactive node to turn active. The threshold τ for an individual in the population is drawn from a distribution $P(\tau)$. This model provides a framework for a *single* topic spreading over a network with *binary* influence. There has been an interest in studying threshold models with an increasing complexity of the underlying contact network (e.g., multiplex networks [16], [17]) and extension to more than two states [18].

An important gap in the literature is that most existing works on complex contagions consider only a *single* topic spreading through a network. However, real-life complex contagions may often involve simultaneous spreading of *multiple correlated* topics. For example, in political events such as elections, multiple opinions spread among the population at the same time, potentially with correlations among the topics [19]. In the context of adopting new technological devices, multiple products might gain traction in a population simultaneously (e.g., iPhone and iPad by Apple Inc.), affecting each other. Finally, in the context of cascading failures in critical infrastructures, nodes may be subject to multiple *types* of failures (e.g., affecting different functionalities such as computing and communication) and failures of different types might spread in the network simultaneously in a correlated manner [20], [21]. Modeling the simultaneous spreading of multiple correlated topics is critical to establishing an understanding of these phenomena and provide insights into developing spreading control and mitigation strategies.

This research was supported in part by the the Air Force Office of Scientific Research through grant #FA9550-22-1-0233, the National Science Foundation through grant CIF-#2225513, and the Army Research Office through grant #W911NF-22-1-0181. YT acknowledges the Knight Fellowship awarded by the IDeaS Center at Carnegie Mellon University.

Borodin et al. proposed a threshold model for the simultaneous spread of two topics [22]. In their model, nodes can adopt either topic-1 or topic-2. While the model incorporates two distinct topics in the spreading process, it lacks the flexibility to capture more complex correlations between topics. For instance, the model does not allow for the simultaneous adoption of both topics by a single node. Zhuang et al. proposed a *vector threshold model* allowing simultaneous adoption of two topics, where each node can occupy one of four states: state-0 (adoption of neither topic), state-1 (topic-1 only), state-2 (topic-2 only), and state-3 (adoption of both) [23]. While this model's visualization of correlation in the two-dimensional *received proportion vector* space is intuitive and flexible, it becomes challenging to extend to a higher number of topics due to the increased complexity of defining states in higher dimensions. Tian et al. [24] proposed a threshold model with correlation, however, the model lacks investigation on the emergence of global cascades and analysis of expected size of global cascades under varying seed sizes, limiting its broader applicability.

In this paper, we propose a *multi-dimensional threshold model with correlated influences* (MDTM), which is a natural extension of Watts' linear threshold model (LTM) [13]. Our method explicitly models the correlation among topics using the *aggregate influence* as a function of the fractions of neighbors that support different topics (see Sec. II-B). We derive results for probability of emergence (PE), cascade threshold, and expected size (ES) of *global cascades*, in which the spreading process leads to the activation of a positive fraction of the population in the limit of the number of nodes going to infinity. Extensive simulations validate our analytical results with near-perfect matches in the finite node regime.

Our results demonstrate that correlations between spreading topics significantly impact both the probability of global cascade emergence and the expected cascade size in distinct ways. Analytical solutions indicate potential dependencies between global cascades of different active states. We further show that positive topic correlations can lead to an increase in *hyper-active* nodes supporting all topics. Although the findings are presented within the context of social contagion, we also provide a case study using our model in the cascading failure context, highlighting the impact of correlation on the final *system size* when the initial *shock size* varies.

The structure of this paper is as follows. In Sec. II, we formally introduce the network model and the proposed multi-dimensional threshold model with correlated influences (MDTM). Sec. III presents our theoretical analysis, where we derive analytical results for the probability of emergence, the cascade threshold, and the expected size of global cascades. In subsequent sections, we provide both numerical and analytical results to investigate the impact of model parameters on influence propagation, with our analytical results closely matching simulation results. Sec. IV investigates the trade-off between the mean degree of the network and correlated influence on the cascade threshold within the MDTM framework. Sec. V explores the effects of network mean degree and correlated influence on the probability of emergence and expected cascade size, specifically for global cascades initiated from a single seed. In Sec. VI, we investigate the dependencies between emergence of global cascades of different active states by providing solutions for the probability of emergence of global cascades of one set of active states, conditioned on the emergence of global cascades of another set. Sec. VII examines the impact of correlation and network mean degree on global cascade size when seeding the network with a randomly selected positive fraction of active nodes. Sec. VIII extends this analysis to cascading failures, exploring the effect of initial *shock size* on the final *system size*. Finally, the conclusion section summarizes these findings.

II. MODEL

A. Network Model

Consider a population with N individuals labeled as $\mathcal{N} = \{1, \ldots, N\}$. Within the network, each node corresponds to an individual in $\mathcal{N}$, and an edge is drawn between two nodes if they are in *contact* (e.g., have friendship on social media) and can transmit topics to each other via the contact. To be able to incorporate arbitrary degree distributions, we generate the contact network $\mathbb{G}$ by the *configuration model* [25], [26]. Put differently, the topology of the network $\mathbb{G}$ is generated *randomly* from its degree distribution $\{p_k\}$, where $k = 0, 1, \ldots$. Here, $\{p_k\}$ gives the probability that an arbitrary node on network $\mathbb{G}$ has degree k. We denote the random networks that has N nodes generated via configuration model with degree distribution $\{p_k\}, k = 0, 1, ...$ as $\mathbb{G}(N, \{p_k\})$. Our analytical solutions in Sec. III are valid for *well-behaved* distributions (i.e., moments of arbitrary order being finite [27], [28]), e.g., Poisson degree distributions, power-law with exponential cutoff, etc. However, it is worth noting that if the second moment of the degree distribution is finite when N approaches infinity, the expected clustering coefficient of a typical node approaches zero. This indicates that the graph is locally tree-like [26].

B. Multi-dimensional Threshold Model with Correlated Influence

In this section, we introduce our proposed MDTM as an extension of the linear threshold model to account for multiple topics propagating through the network in a correlated manner. Consider $M \geq 1$ distinct topics, labeled as topic-1 through topic-M, respectively. Each node in the population holds a binary *state* for each topic-i, determining whether the node adopts it.

We use the matrix $\mathbf{D}$ of size $N \times M$ to represent the states of the N nodes in the population across all M topics:

$$\mathbf{D}_{N\times M}[v, i] = \mathbb{1}[\text{node } v \text{ adopts topic-}i]$$

for $i = 1, ..., M$ and $v = 1, ..., N$. Here, $\mathbb{1}[\cdot]$ is an indicator function. Matrix $\mathbf{D}$ is referred to as the *state matrix*. For a node v, the v-th row of matrix $\mathbf{D}$, denoted by $\mathbf{d}_v$, is referred to as the *state vector* for node v.

We observe that $\mathbf{d}_v \in \{0, 1\}^M$, where $|\{0, 1\}^M| = 2^M$. The state of a node is considered *active* if it adopts at least one topic, and *inactive* if it does not adopt any topic. Following the works in [1]–[4], we assume that once an inactive node

becomes active, its state remains unchanged. In the context of social contagion, this assumption underscores the *finality* of decisions as a key characteristic of scenarios of interest. For example, in political referendums, individuals make binary decisions (yes/no) on multiple issues. Once submitted, these votes are final and cannot be altered.

In an influence propagation process, each inactive node keeps receiving influence from its contacts. We classify topics' influences on nodes into different categories. In our model, the influence a node receives from contacts is measured by the *aggregate influence* from different correlated topics. For $1 \leq i, j \leq M$, the correlation from topic-j to topic-i is measured by a normalized coefficient $w_{ij} \in [-1, 1]$. Additionally, when $i = j$, $w_{ij} = 1$; if $i \neq j$, $w_{ij} \leq 1$. Formally, the aggregate influence of topic-i an inactive node receives, denoted by $f_i(\mathbf{m})$, is given by

$$f_i(\mathbf{m}) = \sum_{j=1}^{M} w_{ij} \cdot m_j \tag{2}$$

where $i, j = 1, ..., M$, and m_j represents the fraction of node v's neighbors that adopt topic-j. Specifically, we have

$$m_j = \frac{\sum_{u \in \mathcal{N}_v} \mathbf{d}_u[j]}{k_v}. \tag{3}$$

where $\mathcal{N}_v$ denotes the set of node v's neighbors, k_v denotes node v's degree, and $|\mathcal{N}_v| = k_v$. For notational convenience, let the vector $\mathbf{m} = [m_1, \ldots, m_M]$, referred to as the *proportion vector* of an inactive node, denote the fraction of neighbors of the inactive node that adopt each topic.

If $w_{ij} > 0$ (resp., $w_{ij} < 0$), it means topic-j is positively (resp., negatively) correlated to topic-i, and thus contributes positively (resp., negatively) to the aggregate influence of topic-i. If $w_{ij} = w_{ji} = 0$, it indicates that topic-i and j are independent of each other. Note that it is not required that $w_{ij} = w_{ji}$ because the impact of one topic on another may not be symmetric in real life. For example, a popular movie series may help increase the sales of related posters but the posters may not be able to help promote the movie as much. We represent the correlation among all M topics as a correlation matrix $\mathbf{W}_{M \times M}$, where

$$\mathbf{W}_{M \times M}[i, j] = w_{ij}, -1 \leq w_{ij} \leq 1. \tag{4}$$

As mentioned, the diagonal elements of $\mathbf{W}$ are ones, and $\mathbf{W}$ may not be symmetric.

Next, we describe the rule for an inactive node, say, node v, to become active by receiving *aggregate influence* from its neighbors. Similar to the linear threshold model, we sample a threshold τ_i from a distribution $P(\tau_i)$, as the minimum received aggregate influence for an inactive node to adopt topic-i, for each $i = 1, ..., M$. The probability for the inactive node v to adopt topic-i is thus given by:

$$\mathbb{P}[f_i(\mathbf{m}) \geq \tau_i]. \tag{5}$$

C. Influence Propagation

We now describe the influence propagation process considered in this work. Suppose that there are M topics that spread simultaneously in a correlated manner. There are 2^M mutually exclusive states, representing all possible combinations of binary states for each of the M topics.

As mentioned earlier, a node is considered active if it adopts at least one topic, and inactive if it does not adopt any topic. For notational convenience, we say a node $v \in \mathcal{N}$ is in state-$\mathbf{s}$ if its state vector

$$\mathbf{d}_v = \mathbf{s}$$

where

$$\mathbf{s} = [s_1, s_2, \cdots, s_M],$$

with $s_i = 1$ denoting adoption of topic-i and $s_i = 0$ otherwise. We observe $\mathbf{s} \in \{0, 1\}^M$. In addition, if the node adopts at least one topic (i.e., $\exists\, i = 1, \cdots, M$ such that $s_i = 1$), we say the node is $\mathbf{s}$-active. In contrast, if $s_i = 0$ for all $i = 1, \cdots, M$, we say the node is inactive. We use the vector $\mathbf{0} = [0]^M$ to denote the inactive state.

Given these, the probability for an inactive node with proportion vector $\mathbf{m}$ to become $\mathbf{s}$-active is given by

$$F_{\mathbf{s}}(\mathbf{m}) = \mathbb{P}\Big[\cap_{i=1}^{M} \{(f_i(\mathbf{m}) \geq \tau_i)^{s_i} \cdot (f_i(\mathbf{m}) < \tau_i)^{1-s_i}\}\Big], \tag{6}$$

where $\mathbf{s} \in \{0, 1\}^M \setminus \{\mathbf{0}\}$. Eq. (6) says that for an inactive node to be $\mathbf{s}$-active, the aggregate influence for each topic i that state-$\mathbf{s}$ adopts must exceed its corresponding threshold τ_i. In other words, (6) are the *response function*s for the corresponding active states. In addition, we also reserve $F_{\mathbf{0}}$ to denote the probability that an inactive node remains inactive:

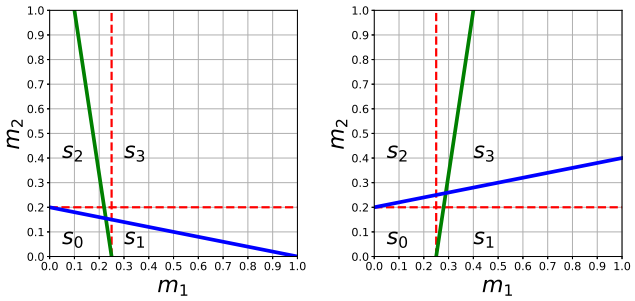
$$F_{\mathbf{0}}(\mathbf{m}) = \mathbb{P}\Big[\cap_{i=1}^{M} \{f_i(\mathbf{m}) < \tau_i\}\Big]. \tag{7}$$

In fact, (6)-(7) provide a mutually exclusive partition of the M-dimensional space spanned by the proportion vector $\mathbf{m}$ into different regions for possible states. Fig. 1 shows examples of two-dimensional regions when $M = 2$[1] The two topics are positively correlated 1(a) and negatively correlated 1(b). We can see the positive correlation increases the area for state-3 compared to the independent case (red dashed lines), resulting in a higher probability of a node adopting both topic-1 and topic-2. On the contrary, the negative correlation decreases the area for state-3, indicating the negative impact on the aggregate influence from each other.

Furthermore, note that (6) and (7) are in fact not dependent on the formation of (2). The modularity allows $f_i(\mathbf{m})$ to be an arbitrary function of the proportion vector $\mathbf{m}$, without impacting the analytical results provided in Sec. III. In other words, our results also accommodate more general modeling of aggregate influence from correlated spreading topics.

Finally, we describe the influence propagation process considered. Assume all nodes are initially inactive (i.e., not adopting any topics). A set of nodes is chosen uniformly at random and set to be active as the *seed*. Other nodes begin changing their states according to (6)–(7) synchronously

[1]Although we generally represent a state using the M-dimensional vector $\mathbf{s} \in \{0, 1\}^M$, for convenience, when $M = 2$, we follow the convention in [23], where integers 0 to 3 are used to denote states [0,0], [1,0], [0,1], and [1,1], respectively.



(a) Positively correlated (b) Negatively correlated

Fig. 1: *An example of the two-dimensional proportion vector space partition for possible states where the first dimension represents m_1, and second dimension represents m_2. The annotation s_i for $i = 0, 1, 2, 3$ represent state-0 to state-3, respectively. In (a), topic-1 and 2 are positively correlated where $w_{12} = 0.15$, $w_{21} = 0.20$; in (b) they are negatively correlated where $w_{12} = -0.15$, $w_{21} = -0.20$. In both (a) and (b), $\tau_1 = 0.25$, $\tau_2 = 0.2$. Green lines separate the space of adopting topic-1 from that of not adopting topic-1. The blue lines separate the space for adopting topic-2 from that for not adopting topic-2. Red dashed lines represent the partition when topic-1 and topic-2 are independent of each other.*

at times $t = 0, 1, \ldots$ (i.e., the influence starts propagating throughout the network). We assume that once an inactive node becomes active, it cannot change its state. Therefore, the contagion process is monotonic and will eventually stop, i.e., it will reach a steady state.

III. Analytical Results

In this section, we present our theoretical analysis, deriving expressions for the probability of emergence, cascade threshold, and expected global cascade size within the framework of the proposed MDTM.

In what follows, assume $M \geq 1$ topics for the influence propagation process. For notational convenience, let $\Omega_A = \{0,1\}^M \setminus \{\mathbf{0}\}$ denote the set that contains all active states, and $\Omega \subseteq \Omega_A$ denotes a non-empty subset of active states. A node is called Ω-active if its state is a member of Ω. See table I for a succinct description of notations for this section.

TABLE I: Table of notations in MDTM.

	Indices
N	Number of nodes in the population ($v = 1, \cdots, N$).
M	Number of topics spreading simultaneously ($i = 1, \cdots, M$).
k	Degree of a node in the population, where $k = 0, 1, \cdots$.
	General variables
s_i	$= \begin{cases} 1, & \text{adopting topic-}i, \\ 0, & \text{otherwise.} \end{cases}$
$\mathbf{s}$	A state, where $\mathbf{s} = [s_1, \cdots, s_M]$, and $\mathbf{s} \in \{0,1\}^M$.
$\mathbf{0}$	Inactive state, where $\mathbf{0} = [0]^M$.
Ω_A	The set containing all the active states, i.e., $\Omega_A = \{0,1\}^M \setminus \{\mathbf{0}\}$.
Ω	A non-empty subset of active states, i.e., $\Omega \subseteq \Omega_A$.
$\mathbf{d}_v$	State vector for node v, where $\mathbf{d}_v \in \{0,1\}^M$.
$\mathbf{D}_{N\times M}$	State matrix, where the v-th row of $\mathbf{D} = \mathbf{d}_v$.
m_i	The fraction of neighbors of an inactive node who adopt topic-i.
$\mathbf{m}$	Proportion vector, where $\mathbf{m} = [m_1, \cdots, m_M]$.
$\mathbb{G}$	Contact network.
	Model parameters
$\mathbf{W}_{M\times M}$	Correlation matrix, where $\mathbf{W}[i,j] = w_{ij} \in [-1, 1]$ for $i, j = 1, \cdots, M$.
τ_i	Threshold for topic-i.
$\boldsymbol{\tau}$	Threshold vector, where $\boldsymbol{\tau} = [\tau_1, \cdots, \tau_M]$.
$\rho_{\mathbf{s}}$	Fraction of randomly selected seed nodes that adopt state $\mathbf{s}$ in the population.
$\boldsymbol{\rho}$	Seed vector, i.e., fraction of randomly selected seed nodes for all states. $\boldsymbol{\rho} = [\rho_{\mathbf{s}}], \mathbf{s} \in \{0,1\}^M$.
$\{p_k\}$	Degree distribution of the network.
	Random variables
N	Number of nodes in the population.
$X^{\Omega}_{\mathbf{s}}(N)$	The fraction of nodes reached and turned Ω-active in the population with N nodes by a $\mathbf{s}$-active seed.
$X_{\boldsymbol{\rho}}(N)$	The fraction of nodes in the population that are *active* given the seed vector $\boldsymbol{\rho}$.
	Function names
f_i	Aggregate influence for topic-i.
$F_{\mathbf{s}}$	Response function for state $\mathbf{s}$.
$h_{\mathbf{s}}(\Omega, x)$	Probability generating function for "the *finite* number of nodes reached and activated to become Ω-active by following a randomly selected edge from an $\mathbf{s}$-active node".
$H_{\mathbf{s}}(\Omega, x)$	Probability generating function for "the *finite* number of nodes reached and activated to Ω-active by following a randomly selected state-$\mathbf{s}$ seed node".

A. Emergence of Global Cascades

We present the analytical results of the probability of emergence and cascade threshold under the proposed MDTM. We start by giving formal definitions of probability of emergence and cascade threshold for the considered influence propagation process.

Definition III.1 (Probability of Emergence of Global Cascades). Ω-active global cascades triggered by an $\mathbf{s}$-active seed refer to how in an initially inactive population, a randomly chosen $\mathbf{s}$-active node can initiate an influence propagation that eventually reaches and activates a positive fraction of Ω-active nodes in the entire population, where $\mathbf{s} \in \Omega_A$ and $\Omega \subseteq \Omega_A$. Formally, let a random variable $X^{\Omega}_{\mathbf{s}}(N)$ denote the fraction of nodes reached and turned Ω-active in the population of size N by a $\mathbf{s}$-active seed, the probability of emergence of Ω-active global cascades triggered by an $\mathbf{s}$-active seed is thus defined as

$$\text{PE}^{\Omega}_{\mathbf{s}} = \lim_{N\to\infty} \mathbb{P}[X^{\Omega}_{\mathbf{s}}(N) > 0].$$

Definition III.2 (Cascade Threshold). Cascade threshold refers to a boundary in the parameter space that separates the regions where global cascades can occur, i.e.,

$$\exists\ \mathbf{s} \in \Omega_A, \exists\ \Omega \subseteq \Omega_A,\ \lim_{N\to\infty} \mathbb{P}[X^{\Omega}_{\mathbf{s}}(N) > 0] > 0,$$

from regions where they cannot occur, i.e.,

$$\exists\ \mathbf{s} \in \Omega_A, \exists\ \Omega \subseteq \Omega_A,\ \lim_{N\to\infty} \mathbb{P}[X^{\Omega}_{\mathbf{s}}(N) > 0] = 0.$$

Consider random graphs $\mathbb{G}(N, \{p_k\})$ generated by the configuration model. To study the influence propagation over the network $\mathbb{G}$, we consider a branching process that begins by randomly selecting a node to serve as the seed, which then recursively discovers the set of nodes that are *reached* and *activated* by exploring its neighbors based on (6). We derive the *survival probability* of the aforementioned branching process through a *mean-field* analysis by employing the method of probability generating functions (PGF) [26], [29].

Let $h_{\mathbf{s}}(\Omega, x)$ denote the generating function for the *finite* number of nodes reached and activated to become Ω-active by following a randomly selected edge from an $\mathbf{s}$-active node. In other words, we have

$$h_{\mathbf{s}}(\Omega, x) = \sum_{k=0}^{\infty} r_k x^k,$$

where r_k denotes the probability that an arbitrary edge emanating from an $\mathbf{s}$-active node leads to a *finite* component of Ω-active nodes of size k. Similarly, let $H_{\mathbf{s}}(\Omega, x)$ denote the generating function for the *finite* number of nodes reached and activated to become Ω-active by following a randomly selected state-$\mathbf{s}$ seed node.

Given $\Omega \subseteq \Omega_A$, we now derive $h_{\mathbf{s}}(\Omega, x)$ for each $\mathbf{s} \in \Omega_A$, recursively. Consider an $\mathbf{s}$-active node, say node v, and an edge incident on it. Consider the node on the other end of this edge, say node u. We find that the following self-consistency equations hold:

$$h_{\mathbf{s}}(\Omega, x) = \sum_{k=1}^{\infty} \frac{p_k k}{\langle k \rangle} \Bigg[F_{\mathbf{0}}(\mathbf{m}) + x \cdot \sum_{s' \in \Omega} F_{\mathbf{s}'}(\mathbf{m}) \cdot {h_{\mathbf{s}'}}^{k-1}(\Omega, x) + \sum_{s' \in \Omega_A \setminus \Omega} F_{\mathbf{s}'}(\mathbf{m}) \cdot h_{\mathbf{s}'}^{k-1}(\Omega, x) \Bigg] \tag{8}$$

where $\mathbf{m}$ denotes the proportion vector for node u. Note that the $\mathbf{s}$-active node v is the only active neighbor of node u in a *naive* population (i.e., a population where all other nodes are inactive), therefore $\mathbf{m} = \mathbf{s}/k$ in (8).

We now explain each term appearing in (8). We first condition on the degree of node u being k, which is given by $p_k k/\langle k \rangle$. We can then determine node u's state based on (6)–(7).

If node u remains inactive (i.e., in state-$\mathbf{0}$), it will have no offspring, which explains the $F_{\mathbf{0}}$ term in (8). If node u is activated to $\mathbf{s}'$-active with probability $F_{\mathbf{s}'}(\mathbf{m})$, and if $\mathbf{s}' \in \Omega$, the number of nodes reached and activated to Ω-active will increase by one, which is captured by the multiplicative term x before the second summation term in (8).

In addition, the total size of this *branch* will also include all subsequent nodes that might be activated to Ω-active by u via its remaining $k-1$ edges, given that node u is reached and activated through an edge connecting to node v.

Furthermore, recall that the number of nodes reached and activated to Ω-active by node u via one of its $k-1$ edges is generated through $h_{\mathbf{s}'}(\Omega, x)$. By using the *powers property* of generating functions [26], and considering that node u influences its neighbors via these $k-1$ edges independently due to the locally tree-like structure, we obtain the term $h_{\mathbf{s}'}^{k-1}(\Omega, x)$.

If node u is activated to other active states that do not belong to Ω, i.e., state-$\mathbf{s}' \in \Omega_A \setminus \Omega$, we do not need to include node u itself in the collection. Instead, we must count all the subsequent nodes that may become Ω-active via the remaining $k-1$ edges of node u, which gives rise to the last term in (8). This completes the derivation of (8).

Using (8), we now derive the generating function $H_{\mathbf{s}}(\Omega, x)$. For each $\mathbf{s} \in \Omega_A$, given $\Omega \subseteq \Omega_A$, we have

$$H_{\mathbf{s}}(\Omega, x) = x^{\mathbb{1}[\mathbf{s} \in \Omega]} \sum_{k=0}^{\infty} p_k \cdot h_{\mathbf{s}}(\Omega, x)^k. \tag{9}$$

Here, the factor $x^{\mathbb{1}[\mathbf{s} \in \Omega]}$ corresponds to the initial node's state $\mathbf{s}$ being a member of Ω. The selected node has degree k with probability p_k. The number of nodes it reaches and activates to Ω-active through each of its k links is generated by $h_{\mathbf{s}}(\Omega, x)$. Similarly, by employing the powers property of generating functions and averaging over all possible degrees, we obtain (9).

With equations (8) in hand, the generating function $H_{\mathbf{s}}(\Omega, x)$ can be computed in the following manner. Given any x, we can solve for the recursive relations (8) to obtain $h_{\mathbf{s}}(\Omega, x)$ for all $\mathbf{s} \in \Omega_A$[2], which in turn will yield $H_{\mathbf{s}}(\Omega, x)$ for all $\mathbf{s} \in \Omega_A$ in light of (9).

We are interested in cases where the number of nodes reached and activated by the initial node is *infinite*, representing cases where a randomly chosen node who is $\mathbf{s}$-active triggers an Ω-active *global cascade*. There exists a trivial solution $h_{\mathbf{s}}(\Omega, 1) = 1$ to (8) (yielding $H_{\mathbf{s}}(\Omega, 1) = 1$) for each $\mathbf{s} \in \Omega_A$, when the number of nodes reached and activated is always *finite*. In other words, the underlying branching process is in the *sub-critical* regime, and *all* activated components have a finite size. However, when the branching process is in the *supercritical* regime, there is a positive probability the branching process will lead to an *infinite* component, indicating $h_{\mathbf{s}}(\Omega, 1) = 1$ for all $\mathbf{s} \in \Omega_A$ is not a stable solution. In this case, $\exists\ \mathbf{s} \in \Omega_A, h_{\mathbf{s}}(\Omega, 1) < 1$, which in turn yields $H_{\mathbf{s}}(\Omega, 1) < 1$. *Conservation of probability* property of generating functions indicates that if there is a component reached and turned Ω-active by the $\mathbf{s}$-active seed node has *infinite* size, we have $\mathrm{PE}_{\mathbf{s}}^{\Omega} = \lim_{N \to \infty} \mathbb{P}[X_{\mathbf{s}}^{\Omega}(N) > 0] = 1 - H_{\mathbf{s}}(\Omega, 1)$.

We can check the stability of the fixed point $h_{\mathbf{s}}(\Omega, 1) = 1$ for all $\mathbf{s} \in \Omega_A$ and $\Omega \subseteq \Omega_A$ by the linearization of the recursion (8) around it. This yields the Jacobian matrix $\boldsymbol{J}$ of size $2^M - 1 \times 2^M - 1$ in which

$$\boldsymbol{J}[\mathrm{idx}(\mathbf{s}), \mathrm{idx}(\mathbf{t})] = \left. \frac{\partial h_{\mathbf{s}}(\Omega, x)}{\partial h_{\mathbf{t}}(\Omega, x)} \right|_{x=1} \tag{10}$$

where $\mathbf{s}, \mathbf{t} \in \Omega_A$, for any given $\Omega \subseteq \Omega_A$. The function $\mathrm{idx}(\cdot)$ provides a one-to-one mapping from each vector in Ω_A to an integer from 1 to $2^M - 1$ for the purpose of matrix coordinate indexing, where the order of such indexing does not matter to the task at hand.

[2]If a $\mathbf{s}$-active node can never activate an inactive neighbor to Ω-active in a *naive* population, this branch will have zero Ω-active offspring. In this case, $h_{\mathbf{s}}(\Omega, x) = 1$, for all $|x| \leq 1$.

Specifically, if all eigenvalues of $\boldsymbol{J}$ are less than one in absolute value (i.e., if the spectral radius $\rho(\boldsymbol{J})$ of $\boldsymbol{J}$ satisfies $\rho(\boldsymbol{J}) \leq 1$, the solution $h_{\mathbf{s}}(\Omega, 1) = 1$ for all $\mathbf{s} \in \Omega_A$ and $\Omega \subseteq \Omega_A$ is stable. In this case, the fraction of active nodes will tend to zero as the number of nodes N goes to infinity for any active seed. In contrast, if $\rho(\boldsymbol{J}) > 1$, the trivial fixed point is not stable, which indicates that the branching process is in the *supercritical* regime. In other words, there is a positive probability that the branching process initiated by an active seed will lead to an *infinite* component of active nodes. In this case, the fraction of active nodes will be strictly greater than zero as the number of nodes N reaches infinity.

To further look into the implications of (10), expand $\boldsymbol{J}[\text{idx}(\mathbf{s}), \text{idx}(\mathbf{t})]$ to the following form

$$\sum_{k=1}^{\infty} \frac{p_k k}{\langle k \rangle}(k-1) F_{\mathbf{t}}(\frac{\mathbf{s}}{k}). \tag{11}$$

We can see that in (11), the term $k(k-1)$ increases monotonically as k increases, while $F_{\mathbf{t}}(\frac{\mathbf{s}}{k})$ could decrease monotonically. Therefore, (11) might lead to more than one solution (corresponding to phase transition points as $\langle k \rangle$ increases, i.e., the points from which the probability of emergence change from zero to a positive value or vice verse), or none at all, compared to the simple contagion models in which only one transition point appears if exists [29], [30]. The two transition points scenario has been reported in many previous studies [13], [31]. The first transition indicates that only when the connectivity of a network reach a certain value, global cascades can exist. On the other hand, the second transition point around *high* mean degree values presents when there is *too much* connectivity, causing the *stability* of nodes to increase. In other words, when nodes have a large number of neighbors, it is difficult for them to get influenced by a few active friends. In this case, these nodes are more *stable* in terms of remaining in their inactive state. Consequently, the second phase transition appears.

Finally, we conclude that the cascade threshold, i.e., the boundary that separates the regions of parameters where $\exists\ \mathbf{s} \in \Omega_A, \exists\ \Omega \subseteq \Omega_A, \lim_{N\to\infty} \mathbb{P}[X_{\mathbf{s}}^{\Omega}(N) > 0] > 0$ from the regions where $\exists\ \mathbf{s} \in \Omega_A, \exists\ \Omega \subseteq \Omega_A, \lim_{N\to\infty} \mathbb{P}[X_{\mathbf{s}}^{\Omega}(N) > 0] = 0$ is given by $\rho(\boldsymbol{J}) = 1$. The probability of emergence of Ω-active global cascades initiated by a $\mathbf{s}$-active seed is given by $\text{PE}_{\mathbf{s}}^{\Omega} = \lim_{N\to\infty} \mathbb{P}[X_{\mathbf{s}}^{\Omega}(N) > 0] = 1 - H_{\mathbf{s}}(\Omega, 1)$.

B. Expected Size of Global Cascades

In this section, we derive the final expected global cascade size if they exist. Consider a network that has been initially seeded by activating a randomly chosen fraction $\rho_{\mathbf{s}} \in [0,1]$[3] of the N nodes to $\mathbf{s}$-active in the population for each $\mathbf{s} \in \Omega_A$. Let $\rho_{\mathbf{0}} \in [0,1]$ denote the initial fraction of nodes in state $\mathbf{0}$. We have $\sum_{\mathbf{s}\in\{0,1\}^M} \rho_{\mathbf{s}} = 1$ and $\rho_{\mathbf{0}} = 1 - \sum_{\mathbf{s}\in\Omega_A} \rho_{\mathbf{s}}$.

Let vector $\boldsymbol{\rho}$ denote the initial fraction of randomly selected seed nodes for all states $\mathbf{s} \in \{0,1\}^M$, that is, $\boldsymbol{\rho} = [\rho_{\mathbf{s}}], \mathbf{s} \in \{0,1\}^M$. Let $X_{\boldsymbol{\rho}}(N)$ be the random variable denoting the fraction of nodes in the population that are *active* given the seed $\boldsymbol{\rho}$. Our objective is to derive

$$\lim_{N\to\infty} \mathbb{E}\left[X_{\boldsymbol{\rho}}(N) \mid X_{\boldsymbol{\rho}}(N) > 0\right],$$

which we denote as ES.

Our approach is similar to that used in [31]. Since the network is locally tree-like as the network size approaches infinity [32], we can consider it as a tree structure, where there is a single inactive node at the top level (referred to as the *root*). We label the levels of the tree from $\ell = 0$ at the bottom to $\ell = \infty$ at the top. Without loss of generality, we assume that the spreading events start at the bottom of the tree and proceed toward the top. In other words, we assume that a node at level $\ell + 1$ can only be activated by its neighbors in level ℓ if it is inactive.

For each $\mathbf{s} \in \Omega_A$, let $q_{\mathbf{s},\ell}$ denote the probability that a node at level ℓ is $\mathbf{s}$-active, conditioned on its parent at level $\ell + 1$ being inactive. Therefore, the probability that the node at level ℓ is inactive is given by

$$q_{\mathbf{0},\ell} = 1 - \sum_{\mathbf{s}\in\Omega_A} q_{\mathbf{s},\ell}.$$

Our goal is to compute $q_{\mathbf{s}}$, which represents the probability that the *root* node is $\mathbf{s}$-active. Consequently, the probability that the root node is *active* is

$$\sum_{\mathbf{s}\in\Omega_A} q_{\mathbf{s}}.$$

Due to the constraint that nodes cannot change their state once they become active, $q_{\mathbf{0},\ell}$ is monotonically decreasing as $\ell \to \infty$, while $\sum_{\mathbf{s}\in\Omega_A} q_{\mathbf{s},\ell}$ is monotonically increasing as $\ell \to \infty$[4]. Moreover, considering that the root node is randomly selected, the quantity

$$\sum_{\mathbf{s}\in\Omega_A} q_{\mathbf{s}}$$

also represents the expected fraction of active nodes in the steady state of the influence propagation process.

Now, we derive $q_{\mathbf{s},\ell}$ in a recursive manner. Let $\mathbf{q}_\ell = [q_{\mathbf{s},\ell}], \mathbf{s} \in \{0,1\}^M$. For each $\mathbf{s} \in \Omega_A$, we find that

$$q_{\mathbf{s},\ell} = \rho_{\mathbf{s}} + \rho_{\mathbf{0}} \sum_k \frac{k p_k}{\langle k \rangle} \sum_{\text{sum}(\mathbf{k}_{\ell-1})=k-1} \text{Mul}(\mathbf{k}_{\ell-1}, \mathbf{q}_{\ell-1}) F_{\mathbf{s}}(\mathbf{m}) \tag{12}$$

where $\mathbf{k}_{\ell-1} = [k_{\mathbf{s},\ell-1}], \mathbf{s} \in \{0,1\}^M$, and $k_{\mathbf{s},\ell-1}$ denotes the number of nodes that are in state-$\mathbf{s}$ on level $\ell - 1$. The term $\text{Mul}(\mathbf{k}_{\ell-1}, \mathbf{q}_{\ell-1})$ denotes the multinomial coefficient of $\mathbf{k}_{\ell-1}$ following the multinomial distribution $\text{Multinomial}(k-1; \mathbf{q}_{\ell-1})$. In addition, in (12), the proportion vector $\mathbf{m} = [m_1, \cdots, m_M]$ follows

$$m_i = \sum_{\mathbf{s}\in\Omega_A} s_i \cdot k_{\mathbf{s},\ell-1}/k, \quad i = 1, \cdots, M \tag{13}$$

where $\mathbf{s} = [s_1, \cdots, s_M]$.

[3] In the case of a single $\mathbf{s}$-active seed (e.g., Section III-A), we have $\rho_{\mathbf{s}} = \lim_{N\to\infty} \frac{1}{N} = 0$.

[4] However, $q_{\mathbf{s},\ell}$ may not monotonically increase as $\ell \to \infty$, therefore $q_{\mathbf{s}}$ is not equal to the final fraction of $\mathbf{s}$-active nodes in the population in the steady state.

To see why (12) holds, let u be an inactive node at level ℓ who is connected to its unique parent at level $\ell+1$. As mentioned, $q_{\mathbf{s},\ell}$ gives the probability that u is $\mathbf{s}$-active on level ℓ. First, we condition whether node u belongs to $\mathbf{s}$-active seeds: it is $\mathbf{s}$-active w.p. (abbr. *with probability*) $\rho_{\mathbf{s}}$, and still inactive w.p. $\rho_{\mathbf{0}}$. Conditioned on u is still inactive and connects to an inactive parent on level $\ell+1$, as before, we then condition on the degree of u being k which has probability $kp_k/\langle k\rangle$. Under the assumption that nodes can only be activated by neighbors in the layers below, node u can be influenced through the $k-1$ edges in layer $\ell-1$. Conditioned on the number of neighbors at level $\ell-1$ in each of the 2^M states being $\mathbf{k}_{\ell-1}$ w.p. $\text{Mul}(\mathbf{k}_{\ell-1}, \mathbf{q}_{\ell-1})$, we can obtain the probability that node u is activated to state-$\mathbf{s}$ according to (6)-(7).

We are now able to compute $q_{\mathbf{s}}$. First, solving (12) in the limit of $\ell \to \infty$ we can compute $\mathbf{q}_\infty$ Using this, we then get

$$q_{\mathbf{s}} = \rho_{\mathbf{s}} + \rho_{\mathbf{0}} \sum_k p_k \sum_{\text{sum}(\mathbf{k})=k} \text{Mul}(\mathbf{k}, \mathbf{q}_\infty) F_{\mathbf{s}}(\mathbf{m}), \qquad (14)$$

and $q_{\mathbf{0}} = 1 - \sum_{\mathbf{s}\in\Omega_A} q_{\mathbf{s}}$. Similarly, for a randomly selected root node, we first condition on whether it is already $\mathbf{s}$-active w.p. $\rho_{\mathbf{s}}$. In the case where it is inactive w.p. $\rho_{\mathbf{0}}$, we further condition on its degree being k w.p. p_k, and the number of neighbors supporting each topic being $\mathbf{k}$ w.p. $\text{Mul}(\mathbf{k}, \mathbf{q}_\infty)$. Therefore, the root node is $\mathbf{s}$-active w.p. $F_{\mathbf{s}}(\mathbf{m})$, which concludes the derivation of (14).

With these in hand, the final fraction of active nodes in the population is thus

$$\text{ES} = \lim_{N\to\infty} \mathbb{E}[X_{\boldsymbol{\rho}}(N) | X_{\boldsymbol{\rho}}(N) > 0] = \sum_{\mathbf{s}\in\Omega_A} q_{\mathbf{s}}. \qquad (15)$$

The above concludes the analytical results for the probability of emergence (PE) of global cascades, the cascade threshold, and the expected size (ES) of global cascades if they occur. In the following sections, we will apply these analytical results to explore how various model parameters impact these key quantities and discuss their implications in practical applications, such as social contagions and cascading failures. Additionally, we will present simulation results in the finite node regime to validate our analytical findings.

IV. Cascade Window

In this section, we explore the trade-off between the mean degree of the network and the correlated influence on the cascade threshold under MDTM. For simplicity, assume $M = 2$. We use integers 0 to 3 to denote the state of adoption of neither topic, topic-1 only, topic-2 only, and both topics, respectively. Let $w_{12} = w_{21} = w$ and $\tau_1 = \tau_2 = 0.2$. The networks are generated using the configuration model with Poisson degree distribution with mean degree $\langle k\rangle$. The red dashed lines (i.e., the cascade boundary in the $(w - \langle k\rangle)$-plane that corresponds to the cascade threshold $\rho(\boldsymbol{J}) = 1$) enclose the region on the $(w - \langle k\rangle)$-plane where $\rho(\boldsymbol{J}) > 1$, i.e., where global cascades can happen. Outside the cascade window corresponds to $\rho(\boldsymbol{J}) < 1$, where there will be no global cascades.

We run 1,000 simulations at each $(w, \langle k\rangle)$ point (sampled at an interval of 0.025 along the axis of w and 0.01 along the axis of $\langle k\rangle$) on the $(w - \langle k\rangle)$-plane. In each simulation a new network with 2,000,000 nodes is generated, and three propagation processes each initiated by a randomly selected 1-active, 2-active and 3-active seed are simulated, respectively.

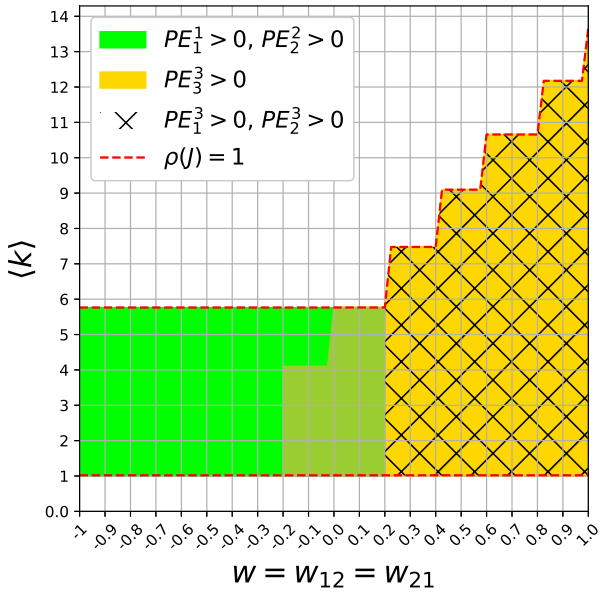


Fig. 2: *Cascade window in the $(w - \langle k\rangle)$ plane. The red dashed lines representing the cascade boundary in the $(w-\langle k\rangle)$ plane that corresponds to the cascade threshold $\rho(\boldsymbol{J}) = 1$. They enclose the region where $\rho(\boldsymbol{J}) > 1$, indicating where global cascades can occur. Conversely, the area outside the cascade window corresponds to $\rho(\boldsymbol{J}) < 1$, indicating the global cascades can not occur. The degree distribution follows Poisson degree distribution with mean degree $\langle k\rangle$. The parameters are set as $w_{12} = w_{21} = w$ and $\tau_1 = \tau_2 = 0.2$.*

The green color marks the region where a 1-active seed can trigger a 1-active global cascade, and a 2-active seed can trigger a 2-active global cascade. The hatched region marks the space where a 1-active seed can trigger a 3-active global cascade, and a 2-active seed can trigger a 3-active global cascade. We can see the simulation results match the analytical cascade boundary near perfect in a finite node regime.

In Fig. 2, we observe two phase transition points along the average degree $\langle k\rangle$, corresponding to $\rho(\boldsymbol{J}) = 1$, as implied by (1). The first transition point indicates that global cascades can only occur when the network's mean degree exceeds one. The second transition point marks the threshold in $\langle k\rangle$ beyond which global cascades no longer emerge. These results suggest a non-monotonic trend in the probability of cascade emergence as $\langle k\rangle$ increases within the cascade window, which we further explore in Sec.V.

In the yellow region, where a 3-active seed can trigger 3-active global cascades, its upper bound along $\langle k\rangle$ expands as w increases. This is due to the growing area of state-3 in the proportion vector space as w increases. This also demonstrates that positive correlation between the spreading topics mitigates the effect of *stability* of inactive nodes as the mean degree of the network increases.

In contrast, within the green region, where a 1-active (or 2-active) seed node can lead to 1-active (or 2-active) global cascades, the second transition point remains constant. This is because when only one topic is involved in influence propagation, the correlation becomes irrelevant, causing the second transition point to remain unchanged when τ is fixed. Notably, the right boundary of the green region occurs at $w = \tau = 0.2$, beyond which a 1-active (or 2-active) seed node can only trigger 3-active global cascades (hatched area). It is important to highlight that when $w \geq \tau$, the inequality $\tau/w \leq 1$ holds. Therefore, if $\tau/w \leq m_1 \leq 1$, even when $m_2 = 0$, the inactive node will become 3-active. In this case, if the population is initialized with a 1-active (or 2-active) seed, any global cascades that occur will only result in 3-active cascades An example of this phenomenon in a social contagion context could be explained through the concept of *information mutation* [33]. For instance, when a *sufficient* number of a node's neighbors like a meme, the node may modify the original meme, creating a highly positively correlated variant. This modified meme could spread simultaneously within the population.

In summary, these results demonstrate that the existence of a global cascade is a result of multiple factors, including seed state, influence correlation, threshold τ, and the interaction between these parameters. Our experiments demonstrate that the analytical cascade threshold $\rho(\boldsymbol{J}) = 1$ effectively captures the complex dynamics among these factors.

V. Impact of mean degree and correlated influence on probability of emergence and expected size of global cascades

In this section, we investigate the impact of the mean degree of the network and the correlated influence on the probability of emergence (PE) and the expected cascade size of global cascades (ES) initiated from a single seed. Specifically, we compare the effects of positively correlated, negatively correlated and independent influence on the probability of emergence and expected global cascade size when there is a single 3-active seed. The parameters are set as follows: for positive correlation, $w_{12} = w_{21} = 0.25$; for negative correlation, $w_{12} = w_{21} = -0.25$; and for independent influence, $w_{12} = w_{21} = 0.0$. In all three cases, $\tau_1 = \tau_2 = 0.25$. Fig. 3 presents the results of probability of emergence of 3-active global cascades initiated from a 3-active seed in 3(a) and expected size (ES) in 3(b), respectively. We observe that the simulation results closely match the analytical solutions near-perfect. This validates our analytical findings and confirms the usefulness of the analytical results within the finite node regime.

In Fig. 3(a), we observe that as the mean degree of the network $\langle k \rangle$ increases, both the positively correlated and independent cases exhibit two phase transition points, with the first transition point occurring at a mean degree of one. Additionally, the positively correlated curve demonstrates a larger second transition point compared to the independent case. In contrast, the negatively correlated case shows zero phase transition points as the mean degree increases. Furthermore, there is a non-monotonic trend in the probability of emergence as the mean degree increases. Specifically, it first increases and then *gradually* decreases to zero. This behavior highlights the trade-off between the network connectivity and local stability of inactive nodes as the mean degree of the network increases. Additionally, the positively correlated case shows a higher PE than the independent case when it is greater than zero. These results indicate that correlated influence affects not only the phase transition points but also the probability of emergence of global cascades within the corresponding cascade window.

In Fig. 3(b), we observe that as the mean degree of the network increases, the global cascade size for both the positively correlated and independent scenarios increases after the first transition point, matching the size of the giant component of the network when it becomes non-zero. The cascade size then sharply drops to zero at their corresponding second transition points. Notably, the positively correlated scenario exhibits a larger second transition point in $\langle k \rangle$ compared to the independent case.

Compared to the probability of emergence shown in Fig. 3(a), we find that correlation affects the probability of emergence and the expected size of the global cascades in distinct ways: correlation can increase or decrease PE, but it does not impact the value of ES within the cascade window, which equals the giant component size when global cascades occur. This suggests that in a global cascade initiated from a single active seed, the final scale of the spread is determined by the number of nodes that can be reached given the underlying network structure.

VI. Dependency between global cascades of different active state sets

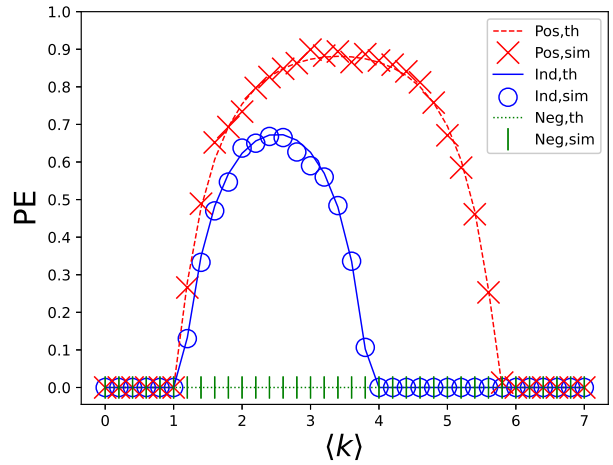
In this section, we investigate dependencies between global cascades of different active state sets. In other words, we are interested in exploring if a global cascade for a set of active states is observed, does this provide insights into the probability of emergence of global cascades for other sets of active states? To approach this question, we provide analysis towards the conditional probability of emergence of Ω_1-active global cascades triggered by a randomly selected state-$\mathbf{s}$ seed, given that a Ω_2-active global cascade is triggered by a randomly selected state-$\mathbf{s}$ seed, where $\mathbf{s} \in \Omega_A$ and $\Omega_1, \Omega_2 \subseteq \Omega_A$, denoted by $\text{PE}_{\mathbf{s}}^{\Omega_1|\Omega_2}$. Specifically, we are interested in

$$\text{PE}_{\mathbf{s}}^{\Omega_1|\Omega_2} = \lim_{N\to\infty} \mathbb{P}[X_{\mathbf{s}}^{\Omega_1}(N) > 0 | X_{\mathbf{s}}^{\Omega_2}(N) > 0]. \tag{16}$$

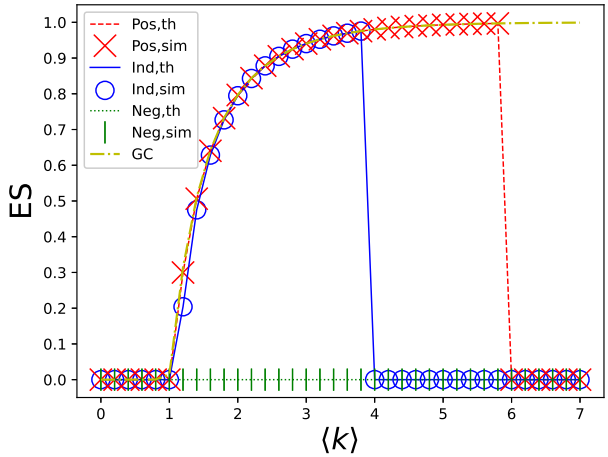
To address this problem, we apply a transformation based on Bayes' theorem and fundamental probability theory, which provides us with

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{P}[X_{\mathbf{s}}^{\Omega_1}(N) > 0 | X_{\mathbf{s}}^{\Omega_2}(N) > 0] \\ &= \lim_{N\to\infty} \frac{\mathbb{P}[X_{\mathbf{s}}^{\Omega_1}(N) > 0] + \mathbb{P}[X_{\mathbf{s}}^{\Omega_2}(N) > 0]}{\mathbb{P}[X_{\mathbf{s}}^{\Omega_2}(N) > 0]} - \\ &\qquad \frac{\mathbb{P}[X_{\mathbf{s}}^{\Omega_1}(N) > 0 \cup X_{\mathbf{s}}^{\Omega_2}(N) > 0]}{\mathbb{P}[X_{\mathbf{s}}^{\Omega_2}(N) > 0]} \end{aligned} \tag{17}$$

in which the term $X_{\mathbf{s}}^{\Omega_1}(N) > 0 \cup X_{\mathbf{s}}^{\Omega_2}(N) > 0$ represents a random event where a randomly selected $\mathbf{s}$-active seed node



(a) Probability of Emergence



(b) Expected Global Cascade Size

Fig. 3: *Probability of emergence (PE) (a) and expected global cascade size (ES) (b) of 3-active global cascades initiated from a randomly selected node in state-3. The analysis considers positive correlation (red):* $w_{12} = w_{21} = 0.25$*; negative correlation (green):* $w_{12} = w_{21} = -0.25$*; and independent influence (blue):* $w_{12} = w_{21} = 0$*. The yellow dashed-dotted curve in (b) represents the giant component (GC) size of the network. The thresholds are set as* $\tau_1 = \tau_2 = 0.25$*, and the degree distribution follows a Poisson distribution with mean* $\langle k \rangle$*. The simulation consists of* $N = 2,000,000$ *nodes, with each data point averaged over 1,000 independent experiments, where a new network is generated via the configuration model in each experiment. The analytical results (th) show a near-perfect match with the simulation results (sim).*

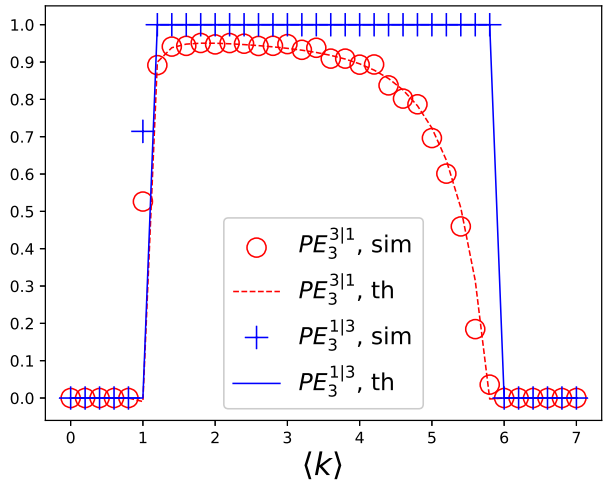


Fig. 4: *Cascade window for MDTM. The dashed lines enclose the region of the* $(w - \langle k \rangle)$*-plane in which* $\rho(\boldsymbol{J}) > 1$*. The degree distribution follows Poisson degree distribution with mean degree* $\langle k \rangle$*. The parameters are* $w_{12} = w_{21} = w$ *and* $\tau_1 = \tau_2 = 0.2$*. The analytical results (th) show a near-perfect match with the simulation results (sim).*

can lead to Ω_1 or Ω_2 active global cascades, which can be equivalently expressed as $X_{\mathbf{s}}^{\Omega_1 \cup \Omega_2}(n) > 0$. With these in hand, we have

$$\mathrm{PE}_{\mathbf{s}}^{\Omega_1|\Omega_2} = \frac{\mathrm{PE}_{\mathbf{s}}^{\Omega_1} + \mathrm{PE}_{\mathbf{s}}^{\Omega_2} - \mathrm{PE}_{\mathbf{s}}^{\Omega_1 \cup \Omega_2}}{\mathrm{PE}_{\mathbf{s}}^{\Omega_2}}, \tag{18}$$

where $\mathrm{PE}_{\mathbf{s}}^{\Omega_1}$, $\mathrm{PE}_{\mathbf{s}}^{\Omega_2}$ and $\mathrm{PE}_{\mathbf{s}}^{\Omega_1 \cup \Omega_2}$ can be computed using the method provided in Sec.III-A.

Following, we provide simulation results that validate the analytical results in (18). Consider a case where $M = 2$. Assume that the correlation parameters are as follows: $w_{12} = 0.3$ if $m_2 < 0.2$, otherwise 0; $w_{21} = 0.16$ if $m_1 < 0.12$, otherwise 0. and $\tau_1 = 0.12, \tau_2 = 0.2$. This correlation between topic-1 and 2 consists of a combination of positive and independent correlations, where the two topics are positively correlated until the fraction of neighbors adopting the other topic exceeds a certain threshold, beyond which the topics become independent, providing a more general and complex form of aggregate influence functions.

Fig. 4 presents both the analytical and simulation results for the conditional probability of the emergence of a state-1 global cascade, given the emergence of a state-3 global cascade, and the conditional probability of the emergence of a state-3 global cascade, given the emergence of a state-1 global cascade, when the randomly selected seed is in state-3. These probabilities are denoted by $\mathrm{PE}_3^{1|3}$ and $\mathrm{PE}_3^{3|1}$, respectively. The probability of the emergence of state-2 is zero under this parameter configuration and is therefore omitted from the analysis. For simplicity, the degree distributions of the networks use the Poisson distribution with mean degree $\langle k \rangle$. Each simulation data point in the figure is an average of 1,000 simulations, in which each simulation generates a new network consisting of $2,000,000$ nodes. The figure shows a good match

between the analytical and simulation results, demonstrating the validity of (18).

In Fig. 4, for $1 < \langle k \rangle < 6$, state-1 global cascades occur with probability one, given the presence of state-3 global cascades, except at $\langle k \rangle = 1.2$, where the probability is approximately 0.75. This result indicates that when $\langle k \rangle > 1.2$, the occurrence of a state-3 global cascade guarantees the simultaneous occurrence of a state-1 global cascade. In contrast, if a state-1 global cascade occurs, the probability of a state-3 global cascade is lower across the range of $\langle k \rangle$. These findings are expected, as the two topics are not negatively correlated, allowing state-1 and state-3 global cascades to coexist significantly. Additionally, the conditions for a node to adopt state-3 are stricter than those for state-1. Specifically, an inactive node must have sufficient neighbors adopting both topic-1 and topic-2 to transition to state-3, whereas it only requires enough topic-1 neighbors to transition to state-1. This stricter requirement for state-3 explains why $\mathrm{PE}_3^{1|3} > \mathrm{PE}_3^{3|1}$ for $1 < \langle k \rangle < 6$. These results demonstrate that the emergence of global cascades across different active states can exhibit dependencies at the global level, driven by the topics adopted by the states and the correlated influence between the topics.

VII. Global cascades with large seed size

In this section, we analyze the influence of correlation and the mean degree of the network on the global cascade size when a randomly selected positive fraction of nodes is seeded as active. Of particular interest, we aim to find out whether positive correlation leads to an increased number of nodes adopting both topics, as compared to independent spreading. Nodes that adopt both topics are referred to as *hyper-active* nodes (e.g., a group of consumers who purchase multiple Apple products concurrently), in contrast to nodes that adopt only one topic. To study the impact of independent and positively correlated influence with large seeds on the size of hyper-active nodes, we set parameters related to correlated influence as follows. Assume $M = 2$, with the population initially seeded with the same fraction (10%) of nodes in state-1 and state-2, respectively. Topics 1 and 2 share the same threshold, $\tau_1 = \tau_2 = 0.2$. The remaining 80% of the nodes are initially inactive. Let the vector $\boldsymbol{\rho} = [\rho_0, \rho_1, \rho_2, \rho_3]$ denote the initial fraction of nodes in each of the four states, yielding $\boldsymbol{\rho} = [0.8, 0.1, 0.1, 0]$.

We compare the following two scenarios: i) Positive correlation between the two spreading topics, where $w_{12} = w_{21} = 0.2$. ii) Independent spreading of the two topics, where $w_{12} = w_{21} = 0$. Fig. 11(a) presents the simulation results of the final sizes of states 1 through 3 for the two scenarios. Fig. 11(b) shows both the simulation and analytical results for the expected global cascade size (ES), along with the giant component size for comparison. As seen in Fig. 11(b), the simulation and analytical results of ES show a near-perfect match. Moreover, compared to the case analyzed in Sec. V, we do not observe phase transition points, as a *global cascade* has already been manually initiated by seeding a nonzero fraction of nodes.

In Fig. 5(a), the size of state-3 differs between the positively correlated and independent cases: it is higher in the former and lower in the latter. This shows a positive correlation between topics can increase the fraction of hyper-active nodes, which can be explained by the fact that inactive nodes have a higher probability of transitioning to state-3 in the positively correlated scenario compared to the independent case. Furthermore, in Fig. 5(a), for both positively correlated and independent scenarios, the final sizes of state-1 and state-2 in the steady state are identical. This result is expected since the parameters are set *symmetrically*, i.e., $\rho_1 = \rho_2$, $w_{12} = w_{21}$, and $\tau_1 = \tau_2$.

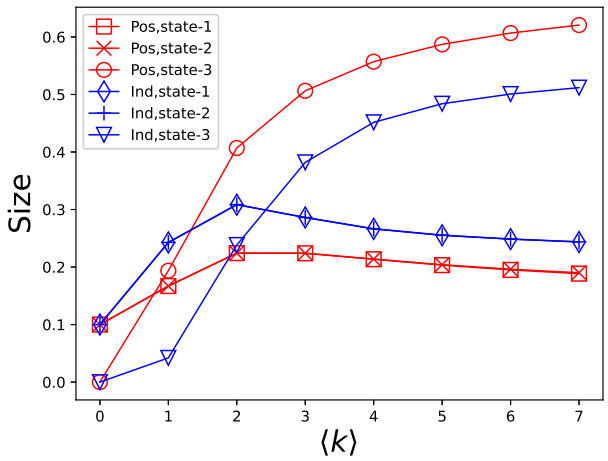
Interestingly, we do not observe a monotonic trend in the sizes of states 1 or 2 as $\langle k \rangle$ increases. Specifically, the size increases when $0 \leq \langle k \rangle \leq 2$ and decreases when $\langle k \rangle > 2$. In contrast, the size of state-3 monotonically increases as the mean degree increases. This illustrates a *competition* between different active states, with the *competency* (i.e., the probability to activate inactive nodes in its own state, measured by the response function of the state) of a specific active state varying with different values of $\langle k \rangle$. However, the overall trend of ES of all active nodes in Fig. 5(b) increases monotonically as the mean degree of the network increases, due to the underlying assumption that active nodes do not switch back to the inactive state in this model.

Moreover, as shown in Fig. 5(b), for $\langle k \rangle \leq 3$, the global cascade size exceeds that of the giant component of the network; however, it gradually converges to the size of the giant component after $\langle k \rangle > 3$. This behavior can be explained as follows. The randomly selected seed nodes are distributed among connected components, including the largest component (i.e., the giant component). Seed nodes can activate nodes within the respective components they reside in. For $1 < \langle k \rangle \leq 3$, the giant component is relatively small. As the mean degree of the network increases, the size of the giant component grows exponentially, thereby increasing the probability that the seed nodes (comprising 20% of the total nodes in this case) reside within the giant component. Consequently, the number of seed nodes positioned outside the giant component decreases, which reduces the probability of large-scale spreading initiated by these nodes in other components, increasing the probability of *significant* cascades occurring within the giant component. In Fig. 5(b), the seed nodes active all the nodes within the giant component. Additionally, we observe that both the positively correlated and independent scenarios result in the same expected global cascade size. As discussed in Section IV, both the positively correlated case (i.e., $w = 0.2, \tau = 0.2$) and the independent case (i.e., $w = 0, \tau = 0.2$) share the same cascade window along the average degree $\langle k \rangle$ when the initial seed is in state-1 or state-2. Thus, when a global cascade is initiated by each individual seed node within the cascade window, it will activate all the nodes it can reach in the network. Consider that active nodes increases monotonically as the spreading proceeds, consequently, the union of all activated nodes from these seed nodes leads to the same final cascade size in both the positively correlated and independent cases.

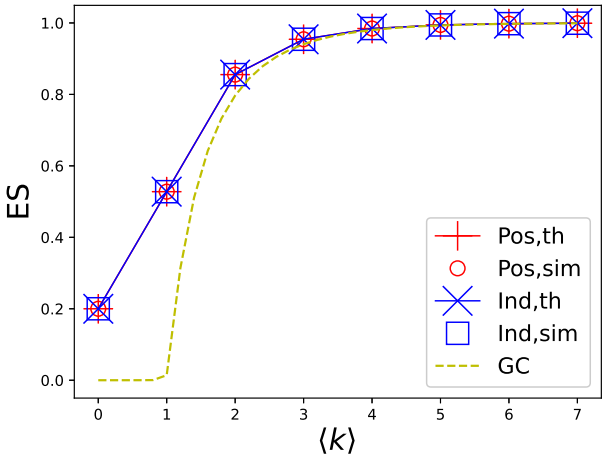
Fig. 6 shows the expected size of global cascades in a different scenario. The population is initially seeded with an equal proportion of nodes supporting state-1, state-2, and state-3. Specifically, the parameters are set as $\boldsymbol{\rho} = [0.85, 0.05, 0.05, 0.05]$, $w_{12} = 0$, $w_{21} = 0.2$, and $\tau_1 = 1$,

$\tau_2 = 0.4$. This configuration indicates that the spread of topic-1 is particularly challenging due to its high threshold, and topic-2 does not facilitate the spread of topic-1 (i.e., $w_{12} = 0$). However, topic-1 positively contributes to the spread of topic-2 (i.e., $w_{21} > 0$). As shown in Fig. 6, despite seeding the same fraction for states 1, 2, and 3, state-2 achieves the largest final size. This outcome is expected as topic-2 has a lower threshold, and the spread of topic-1 aids the spread of topic-2. Additionally, the global cascade size of all active nodes is smaller than the size of the giant component in the network with $\langle k \rangle > 1$. This result demonstrates that despite seeding a positive fraction of nodes leading to some fraction of active nodes, the process does not generate enough active nodes to trigger significant cascades capable of occupying the entire giant component as compared to the case in Fig. 5(b), due to the high thresholds of the topics and the limited positive correlation between them in this parameter configuration.

Furthermore, a non-monotonic trend is observed for all three states as the mean degree increases; the cascade size first increases and then decreases with increasing mean degree. However, the peak cascade size occurs at different values of the mean degree for each state: both state-1 and state-3 peak around $\langle k \rangle = 1$, while state-2 peaks around $\langle k \rangle = 2$. This behavior can be attributed to the trade-off between increasing connectivity and stability of inactive nodes as the mean degree of the network increases. When the number of active neighbors is fixed, achieving a high proportion of active nodes becomes easier when the total number of neighbors is smaller. As a result, topics with higher thresholds tend to reach their peak cascade size earlier along $\langle k \rangle$. This finding is interesting and somewhat counterintuitive: when a topic is harder to adopt with a high threshold, spreading it in a community with lower connectivity may result in a larger cascade of influence.



(a) Size of state-1, 2 and 3



(b) Expected Global Cascade Size

Fig. 5: *Simulation results of size of state 1, 2 and 3 (a), analytical (th) and simulation (sim) results of global cascades size (ES) (b) in the steady state when the population is initially seeded with* $\boldsymbol{\rho} = [0.8, 0.1, 0.1, 0.0]$. *Blue color denotes the results for the independent case, where* $w_{12} = w_{21} = 0$, *and the red color denotes the results for the positively correlated case where* $w_{12} = w_{21} = 0.2$. *For both cases,* $\tau_1 = \tau_2 = 0.2$. *In (b), the yellow dashed line shows the size of the giant component of the network (labeled as GC). Degree distribution follows Poisson degree distribution with mean degree* $\langle k \rangle$.

VIII. Implications on Cascading Failures: Impact of Initial Shock Size on the Final System Size

As previously discussed, when influence propagation begins with a single seed and a global cascade occurs, it typically spans the entire giant component of the network. However, we observe that when a global cascade is initiated from a set of active nodes, it does not necessarily result in a cascade large enough to occupy a giant component. To investigate this matter further, our next objective is to determine how the initial seed size affects the global cascade size under different types of topic correlations. In this section, we investigate how the initial seed size influences the final cascade size.

We believe that our MDTM in general and this investigation in particular can also shed light on the robustness of networks under cascading failures. Namely, it can help reveal how the initial *shock* size affect the final system size, i.e., the fraction of nodes that *survive* in the steady state of the failure propagation process [13], [31], [34], [35]. Consider a scenario where there are $M \geq 1$ functionalities that can cause node failure. Each node in the population can be in one of 2^M states, where state-0 denotes that the node is unaffected by the failure of any of the M functionalities in the cascading failure process, referred to as the *survival* state. In contrast, the remaining $2^M - 1$ states represent *failure* states, where each state specifies the corresponding underlying functionalities responsible for the failure.

The M functionalities can exhibit positive, negative, or independent correlations. For example, a positive correlation might occur if the failure of one functionality overloads other functionalities within a node, increasing the likelihood of failure due to these interconnected functionalities [36], [37]. Conversely, a negative correlation may model a scenario where the failure of one functionality frees up resources, making other functionalities more robust [38], [39].

Our objective is to examine how the correlation between

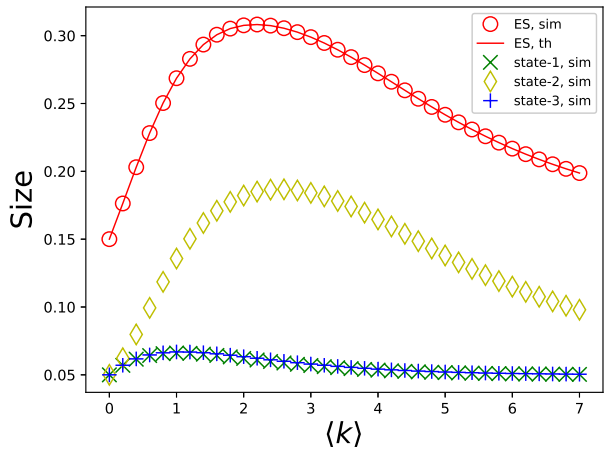


Fig. 6: *Global cascades size (ES) with large seed size* $\boldsymbol{\rho} = [0.85, 0.05, 0.05, 0.05]$, *and* $w_{12} = 0$ *and* $w_{21} = 0.2$, *and* $\tau_1 = 1$, $\tau_2 = 0.4$. *Degree distribution follows Poisson degree distribution with mean degree* $\langle k \rangle$. *The analytical results (th) show a near-perfect match with the simulation results (sim).*

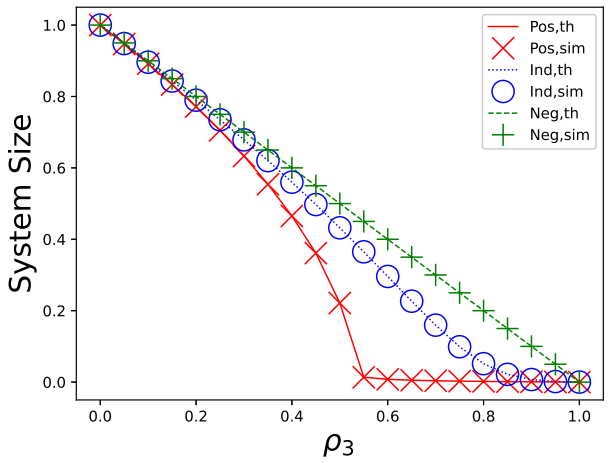


Fig. 7: *Final system size as a function of attack size (i.e., fraction of seed nodes in state-3, denoted by* ρ_3*) when the functionalities that can lead to node failure contain positive correlation (red), negative correlation (green) and independent of each other (blue). The degree distribution follows Poisson distribution with mean degree 5. In this experiment, there are zero state-1 or 2 seed nodes (i.e.,* $\rho_1 = \rho_2 = 0$*). The analytical results (th) show a near-perfect match with the simulation results (sim).*

functionalities that can cause node failures affects the *system size*, defined as the relative size of surviving nodes in state-0 in the steady state. In the following analysis, we assume each node has $M = 2$ underlying such functionalities. Each data point in the simulations represents the average of 1,000 independent experiments, where each experiment generates a new network consisting of 2,000,000 nodes.

Fig. 7 illustrates the final system size as a function of the initial shock size under three different correlation scenarios: (i) $w_{12} = 1$ if $m_2 > 0.5$, 0 otherwise; $w_{21} = 1$ if $m_1 > 0.5$ 0 otherwise (ii) $w_{12} = w_{21} = 0$ (iii) $w_{12} = -1$ if $m_2 > 0.5$, 0 otherwise; and $w_{21} = -1$ if $m_1 > 0.5$, 0 otherwise. For all cases, the thresholds are set to $\tau_1 = \tau_2 = 0.8$. The correlated influence functions represent a combination of positive and independent correlations. The network degree distribution follows a Poisson distribution with a mean degree of 5. Initially, the functionalities are independent, but as the fraction of failing neighbors in the other functionality exceeds a certain threshold, the correlations become active. This setup demonstrates a scenario where the system remains robust until the point where failure in one functionality significantly degrades the robustness of the other. Additionally, the initial shock is assumed to consist of a set of state-3 seed nodes, which represent nodes where both functionalities have failed (e.g., due to adversarial attacks). The results shown in Fig. 7 indicate that as the initial shock size increases, the system size decreases for all correlation settings. However, positive correlation leads to complete system failure with the smallest initial shock size, compared to the independent setting. In contrast, negative correlation enhances the system's resilience, mitigating cascading failure and maintain robustness against initial shocks.

Fig. 8 illustrates the final system size when the network follows a power-law distribution with an exponential cut-off (e.g., air transportation network [40], internet topology [41]) with an eye towards validating the analytical results. The power-law exponent is set to 2.5, with a cut-off value of 10, resulting in a mean degree of 1.42. A near-perfect match is observed between the simulation results and the analytical predictions. In this experiment, we compare three different correlation scenarios: (i) $w_{12} = w_{21} = 0.9$, (ii): $w_{12} = w_{21} = 0$, (iii): $w_{12} = w_{21} = -0.9$. The thresholds for all three scenarios are set as $\tau_1 = \tau_2 = 0.8$. Additionally, we randomly initialize 10% of the nodes in state-1 and state-2 (i.e., $\rho_1 = \rho_2 = 0.1$). We then vary the fraction of state-3 seed nodes ρ_3 from 0.1 to 0.8. As shown in Fig. 8, the impact of correlation on the final system size follows a similar trend across the different scenarios: positive correlation decreases the network's robustness to failures, while negative correlation increases robustness compared to the independent case.

In conclusion, the results for both Poisson degree distribution and power-law distribution with an exponential cutoff demonstrate that a positive correlation between failure factors reduces the final system size for a given attack size. In other words, fewer nodes need to be attacked to reach the same final system size in the positively correlated scenario, compared to the independent and negatively correlated cases.

CONCLUSION

In this paper, we propose the *multi-dimensional threshold model* (MDTM) for the simultaneous spread of multiple correlated topics over complex networks. We provide analytical solutions for the probability of emergence, the cascade threshold, and the expected size of global cascades when there is a randomly selected seed node. We also provide analytical results on the expected global cascade size when the seed is a positive fraction of active nodes in the population. The numerical results from extensive simulations validated the analytical

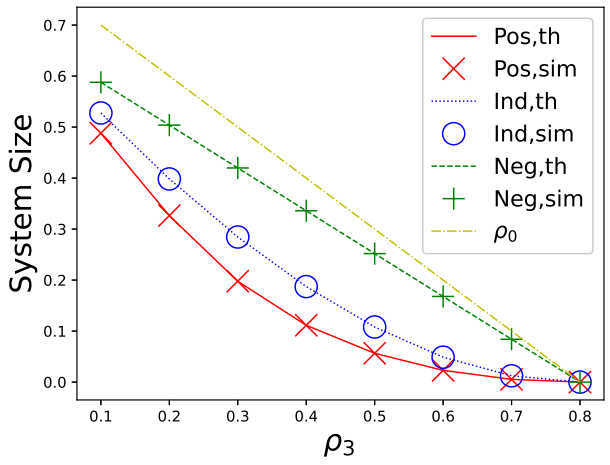


Fig. 8: *Final system size as a function of attack size (i.e., fraction of seed nodes in state-3, denoted by ρ_3) when the functionalities that can lead to node failure exhibit positive correlation (red), negative correlation (green) and independent of each other (blue). The degree distribution follows power-law with exponential cutoff where the power exponent equals 2.5, and the cutoff equals 10. The mean degree is 1.42. In this experiment, there are 10% state-1 seed nodes and 10% state-2 seed nodes (i.e., $\rho_1 = \rho_2 = 0.1$, and $\rho_0 = 1 - \rho_1 - \rho_2 - \rho_3$). The analytical results (th) show a near-perfect match with the simulation results (sim).*

solution with a near perfect match. The analytical results reveal the interplay between multiple factors in the emergence behavior of global cascades, including the underlying network structure, thresholds of topics and the correlation among the spreading topics.

Utilizing the above results, it is shown that the correlation between spreading topics affects the probability of emergence and expected global cascade size, each in a distinct manner. Specifically, with a single seed, a positive (negative, resp.) correlation can increase (decrease, resp.) the probability of emergence. However, the expected size of a global cascade is not affected by the correlation and will occupy the entire giant component if one exists within its cascade window. However, if there is a *large* seed (i.e., the seed contains a positive fraction of randomly selected active nodes in the population), the seed size impacts the final expected global cascade size.

Besides, we also provide analytical solutions for the conditional probability of emergence of a global cascade of a set of active states if a global cascade of a different set of active states exists; results show that there might be dependency between global cascades of different active states. In addition, we also see that there is *competition* between different active states during the spreading process, and the *competency* of a specific active state also varies when the mean degree of the network is different. Moreover, we see that a positive correlation between topics can create more *hyper-active* nodes (i.e., nodes that support all topics) compared to the case where they spread independently.

In summary, our results provide insights into understanding complex contagions with multiple correlated contents and help develop control and mitigation strategies for influence propagation processes.

REFERENCES

[1] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, vol. 69, p. 045104, Apr 2004.
[2] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, "Theory of rumour spreading in complex social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 374, no. 1, pp. 457–470, 2007.
[3] Y. Tian, A. Sridhar, C. W. Wu, S. A. Levin, K. M. Carley, H. V. Poor, and O. Yağan, "Role of masks in mitigating viral spread on networks," *Phys. Rev. E*, vol. 108, p. 014306, Jul 2023.
[4] Y. Liang, Z. Jiang, and Y. Zheng, "Inferring traffic cascading patterns," in *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '17, (New York, NY, USA), Association for Computing Machinery, 2017.
[5] T. R. Hurd, J. P. Gleeson, and S. Melnik, "A framework for analyzing contagion in assortative banking networks," *PLOS ONE*, vol. 12, pp. 1–20, Feb. 2017. Publisher: Public Library of Science.
[6] J. M. Beggs and D. Plenz, "Neuronal Avalanches in Neocortical Circuits," *Journal of Neuroscience*, vol. 23, pp. 11167–11177, Dec. 2003. Publisher: Society for Neuroscience Section: Behavioral/Systems/Cognitive.
[7] D. Centola and M. Macy, "Complex Contagions and the Weakness of Long Ties," *American Journal of Sociology*, vol. 113, pp. 702–734, Nov. 2007. Publisher: The University of Chicago Press.
[8] D. Centola, *How Behavior Spreads: The Science of Complex Contagions*. Princeton, NJ: Princeton University Press, 2018.
[9] O. Yağan and V. Gligor, "Analysis of complex contagions in random multiplex networks," *Physical Review E*, vol. 86, Sept. 2012.
[10] T. Kuran, "Sparks and prairie fires: A theory of unanticipated political revolution," *Public Choice*, vol. 61, pp. 41–74, Apr. 1989.
[11] V. Venkatesh and M. G. Morris, "Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior," *MIS Quarterly*, vol. 24, p. 115, Mar. 2000.
[12] M. J. Salganik, P. S. Dodds, and D. J. Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," *Science*, vol. 311, pp. 854–856, Feb. 2006.
[13] D. J. Watts, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences*, vol. 99, pp. 5766–5771, Apr. 2002. Publisher: Proceedings of the National Academy of Sciences.
[14] H. Wang, H. Shen, and Z. Li, "Approaches for resilience against cascading failures in cloud datacenters," in *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, pp. 706–717, 2018.
[15] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman, "Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 17, no. 2, 2007.
[16] O. Yağan and V. Gligor, "Analysis of complex contagions in random multiplex networks," *Phys. Rev. E*, vol. 86, p. 036103, Sep 2012.
[17] Y. Desmond Zhong, V. Srivastava, and N. E. Leonard, "On the linear threshold model for diffusion of innovations in multiplex social networks," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 2593–2598, 2017.
[18] 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, p. 013124, Feb. 2013.
[19] M. Mäs and A. Flache, "Differentiation without Distancing. Explaining Bi-Polarization of Opinions without Negative Influence," *PLOS ONE*, vol. 8, p. e74516, Nov. 2013. Publisher: Public Library of Science.
[20] Y. Xu and W. Wang, "Characterizing the spread of correlated failures in large wireless networks," in *2010 Proceedings IEEE INFOCOM*, pp. 1–9, 2010.
[21] Q. Zhang, D. Li, R. Kang, E. Zio, and P. Zhang, "Reliability analysis of interdependent networks using percolation theory," in *2013 International Conference on Signal-Image Technology Internet-Based Systems*, pp. 626–629, 2013.
[22] A. Borodin, Y. Filmus, and J. Oren, "Threshold Models for Competitive Influence in Social Networks," in *Internet and Network Economics* (A. Saberi, ed.), Lecture Notes in Computer Science, (Berlin, Heidelberg), pp. 539–550, Springer, 2010.

[23] Y. Zhuang and O. Yağan, "A Vector Threshold Model for the Simultaneous Spread of Correlated Influence," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, pp. 1–7, May 2019. ISSN: 1938-1883.
[24] Y. Tian and O. Yağan, "Multi-dimensional threshold model with correlation: Emergence of global cascades," in *ICC 2024 - IEEE International Conference on Communications*, pp. 533–538, 2024.
[25] 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.
[26] M. E. J. 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.
[27] H. Alexander and T. Day, "Risk factors for the evolutionary emergence of pathogens," *Journal of The Royal Society Interface*, vol. 7, no. 51, pp. 1455–1474, 2010.
[28] A. Allard, P.-A. Noël, L. J. Dubé, and B. Pourbohloul, "Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics," *Phys. Rev. E*, vol. 79, p. 036113, 2009.
[29] M. Newman, "Epidemics on networks," in *Networks*, 2018.
[30] 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, pp. 1038–1048, June 2013. arXiv: 1112.4002.
[31] J. P. Gleeson and D. J. Cahalane, "Seed size strongly affects cascades on random networks," *Phys. Rev. E*, vol. 75, p. 056103, May 2007.
[32] B. Söderberg, "Properties of random graphs with hidden color," *Phys. Rev. E,*, vol. 68, no. 2 Pt 2, 2003.
[33] M. Sood, A. Sridhar, R. Eletreby, C. W. Wu, S. A. Levin, O. Yağan, and H. V. Poor, "Spreading processes with mutations over multilayer networks," *Proceedings of the National Academy of Sciences*, vol. 120, no. 24, p. e2302245120, 2023.
[34] Y. Zhang and O. Yağan, "Optimizing the robustness of electrical power systems against cascading failures," *Scientific Reports*, vol. 6, p. 27625, June 2016. Publisher: Nature Publishing Group.
[35] L. Chen, D. Yue, C. Dou, Z. Cheng, and J. Chen, "Robustness of cyber-physical power systems in cascading failure: Survival of interdependent clusters," *International Journal of Electrical Power Energy Systems*, vol. 114, p. 105374, 2020.
[36] L. D. Valdez, L. Shekhtman, C. E. La Rocca, X. Zhang, S. V. Buldyrev, P. A. Trunfio, L. A. Braunstein, and S. Havlin, "Cascading failures in complex networks," *Journal of Complex Networks*, vol. 8, p. cnaa013, Apr. 2020.
[37] H. Wang, H. Shen, and Z. Li, "Approaches for resilience against cascading failures in cloud datacenters," in *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, pp. 706–717, 2018.
[38] T. Dong, F. Xue, H. Tang, and C. Xiao, "Deep reinforcement learning for fault-tolerant workflow scheduling in cloud environment," *Applied Intelligence*, vol. 53, no. 9, pp. 9916–9932, 2023.
[39] S. S. Gill and R. Buyya, "Failure management for reliable cloud computing: a taxonomy, model, and future directions," *Computing in Science & Engineering*, vol. 22, no. 3, pp. 52–63, 2018.
[40] R. Guimera, S. Mossa, A. Turtschi, and L. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 22, pp. 7794–7799, 2005.
[41] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM Computer Communication Review*, vol. 29, pp. 251–262, ACM, 1999.

IX. Biography Section

Yurun Tian received the M.S. in Electrical and Computer Engineering (ECE) from Carnegie Mellon University (CMU) in 2021. She is currently working towards the Ph.D. degree in ECE, CMU, Pittsburgh, PA, USA.

Osman Yağan is a Research Professor of ECE at CMU. He received his Ph.D. degree in ECE from the University of Maryland at College Park, MD in 2011. He is a recipient of a CIT Dean's Early Career Fellowship, an IBM Academic Award, and best paper award in ICC 2021, IPSN 2022, and ASONAM 2023.