

Multi-Dimensional Threshold Model With Correlation: Emergence of Global Cascades

Yurun Tian

Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, USA
yurunt@andrew.cmu.edu

Osman Yağın

Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, USA
oyagan@andrew.cmu.edu

Abstract—The study of complex contagions over networks has been receiving increasing attention across many scientific domains. Especially, linear threshold models are widely studied due to the ability to capture the mechanism that multiple sources of exposure are required for nodes in the network to take action. Most related works on influence propagation only concentrate on a single content spreading over networks. However, complex contagions usually involve multiple *correlated* contents spreading simultaneously, demonstrating significant implications for many real-life systems. In this work, we propose a *multi-dimensional threshold model* as an extension of the classical linear threshold model to incorporate multiple correlated contents spreading simultaneously over networks. We also provide analytical results that accurately predict probability of emergence of *global cascades* for correlated contents. These results reveal the interplay between the underlying network structure and contents' correlation on the spreading processes. Thus, our work advances analysis, prediction and control strategies for complex contagions over networks.

Index Terms—Influence propagation; Linear threshold model; correlated contents; 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 content 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 content (e.g., an opinion, a product, 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*. Spreading event examples include the rise of collective action to join riots, the diffusion

of beliefs, political revolutions [8], successful technologies [9], and cultural market sensations [10]. In addition, cascading failures observed in critical infrastructure networks, transportation systems and cloud computing services have sometimes been studied using complex contagion models [11], [12].

The linear threshold model has been widely used in the literature for modeling complex contagions over networks [11]. 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

$$F[m, k] \triangleq P\left[\frac{m}{k} > \tau\right] \quad (1)$$

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* content 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 [13], [14]) and extension to more than two states [15].

To the best of our knowledge, existing works on complex contagions consider only a *single* content spreading through a network. However, real-life complex contagions may often involve simultaneous spreading of *multiple correlated* contents. For example, in political events such as elections, multiple opinions spread among the population at the same time, potentially with correlations among the topics [16]. 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 [17], [18]. Modeling the simultaneous spreading of multiple correlated contents is critical to establish an understanding of these

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

phenomena and provide insights into developing spreading control and mitigation strategies.

Borodin et al. proposed a threshold model for the simultaneous spread of two contents [19]. In their model, nodes can either adopt content-1 or content-2. Although there are two contents in the spreading process, there is limited correlation between contents 1 and 2 because both contents cannot be adopted together. Zhuang et al. proposed a *vector threshold model* which allows for the simultaneous adoption of two contents [20]. Each node can be in one of the four states indicating all possible adoption scenarios of the two contents: state-0 represents adoption of neither, state-1 represents adoption of only content-1, state-2 represents adoption of only content-2, and state-3 represents adoption of both contents. State-0 is *inactive* and the others are *active*. However, each possible state's response function relies on a mutually exclusive partition of the two-dimensional proportion vector space, making it hard to extend to an arbitrary number of contents due to the increasing complexity of specifying all possible states in a high-dimensional space. In addition, Zhuang et al. did not provide an analysis of the emergence of global cascades for the influence spreading of correlated contents.

In this paper, we propose a *multi-dimensional threshold model* (MDTM), a more natural extension of Watts' linear threshold model [11], to incorporate an arbitrary number of contents with correlations. Our method explicitly quantifies the correlation among contents with normalized coefficients, and can be easily extended to incorporate non-linear correlation. We provide analytical results of probability of emergence (PE) of *global cascades* for active states, 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. Our analytical solutions disentangle the impact of multiple factors, including the degree distribution of the contact network and correlation among contents on the existence of global cascades. Extensive simulations validate our analytical results with near-perfect matches in the finite node regime. Our results provide insight into complex contagions with multiple correlated contents, helping develop efficient spreading control and mitigation strategies.

II. MODEL

A. Network model

Consider a population of size n with individuals labeled as $\mathcal{N} = \{1, \dots, n\}$. Within the network, each node corresponds to an individual in \mathcal{N} , and an edge is drawn between two nodes if they have a chance to transmit contents to each other. To be able to incorporate arbitrary degree distributions, we generate the contact network \mathbb{G} by the *configuration model* [21], [22]. Put differently, the topology of the network \mathbb{G} is generated *randomly* from its degree distribution $\{p_k\}$, where $k = 0, 1, \dots$. Here, $\{p_k\}$ gives the probability that an arbitrary node on network \mathbb{G} has degree k . We denote the random networks size of n generated via configuration model with degree distribution $\{p_k\}, k = 0, 1, \dots$ as $\mathbb{G}(n, \{p_k\})$. Our analytical solutions are valid for *well-behaved* distributions

(i.e., moments of arbitrary order being finite [23], [24]), e.g., Poisson degree distributions, power-law degree distributions with exponential cutoff, etc. However, it is worth noting that if the second moment of the degree distribution is finite when n approaches infinite, the expected clustering coefficient of a typical node approaches zero. This indicates that the graph is locally treelike.

B. Threshold model with correlation

Here we introduce our proposed MDTM model as an extension of Watts' linear threshold model to incorporate an arbitrary number of contents, together with their correlation. Consider $M \geq 1$ different contents, denoted as content-1, ..., content- M , respectively. For each content- i , each node in the population makes a binary decision of whether to support it or not. We use matrix \mathbf{D} size of $N \times M$ to represent nodes' binary decisions in the population size of N for all M contents:

$$\mathbf{D}_{N \times M}[u, i] = \mathbb{1}[\text{node } u \text{ supports content-}i],$$

where $i = 1, \dots, M$, $u = 1, \dots, N$, and $\mathbb{1}[\cdot]$ is an indicator function that returns 1 if the condition inside is true and 0 otherwise. A node's state is *active* if it supports at least one content and *inactive* if it does not support any content. In this work, we assume that once an inactive node turns active, it cannot change its state.

In an influence propagation process, each inactive node keeps receiving influence from its contacts. We classify contents' influences on nodes into different categories. In our model, the influence a node receives from contacts is measured by the *accumulative influence* from different correlated contents. For $1 \leq i, j \leq M$, the correlation from content- j to content- i is measured by a normalized coefficient $w_{ij} \in [-1, 1]$. The largest coefficient value of one will be reached when $i = j$, if $i \neq j$, $w_{ij} < 1$. This means that a content most highly influences the same content, and other correlated contents only have additional influence, which, scaled by the coefficient, might increase or decrease the main influence.

Formally, the accumulative influence of content- i a node receives is given by

$$\text{Influence}_i = \sum_{j=1}^M w_{ij} \cdot m_j \quad (2)$$

where $i, j = 1, \dots, M$, and m_j represent the fraction of node v 's neighbors that support content- j . Specifically, we have

$$m_j = \frac{\sum_{u \in \mathcal{N}_v} \mathbf{D}[u, j]}{k}. \quad (3)$$

where \mathcal{N}_v denotes the set of node v 's direct neighbors, k denotes node v 's degree, and $|\mathcal{N}_v| = k$.

If $w_{ij} > 0$ (resp., $w_{ij} < 0$), it means content- j is positively (resp., negatively) correlated to content- i , and thus contributes positively (resp., negatively) to the accumulative influence of content- i . If $w_{ij} = w_{ji} = 0$, it indicates that content- i and j are independent of each other. Note that it is not required that $w_{ij} = w_{ji}$ because the impact of one content on another may

not be symmetric in real life. For example, a popular movie series may help increase the selling of related posters but the posters may not be able to help promote the movie as much. However, in the social influence propagating context, w_{ij} and w_{ji} should share the same sign to retain consistency. For notational convenience, we represent the correlation among all M contents as a correlation matrix $\mathbf{W}_{M \times M}$, where

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

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 *accumulative influence* from its neighbors. Similar to Watts' linear threshold model, we sample a threshold τ_i from a distribution $P(\tau_i)$, as the minimum received accumulative influence for an inactive node to adopt content- i , for each $i = 1, \dots, M$. The probability for the inactive node v to support content- i is thus given by:

$$\mathbb{P}[\text{Influence}_i \geq \tau_i]. \quad (5)$$

Note that even though (2) is a linear combination of the fraction of neighbors supporting each content, we can extend (2) to a non-linear function of $m_i, i = 1, \dots, M$ due to the flexibility of function design.

C. Influence propagation

Now we describe the influence propagation process in this work. For clarity, we present a case when $M = 2$ (i.e., there are two contents considered). Analysis for $M > 2$ can be achieved in a similar manner.

Suppose content-1 and content-2 are in the influence propagation process, and nodes in the population can have four mutually exclusive states in total: state-0 for supporting neither contents, state-1 for supporting only content-1, state-2 for supporting only content-2, and state-3 for supporting both contents. As mentioned, state-0 is *inactive*, and the rest are *active*. For notational convenience, we say a node is i -active if it is in an active state- i . The probability for an inactive node to become i -active, $i = 1, 2, 3$ is given by

$$F_1(\mathbf{W}, \mathbf{m}) = \mathbb{P}[\text{Influence}_1 \geq \tau_1, \text{Influence}_2 < \tau_2]; \quad (6)$$

$$F_2(\mathbf{W}, \mathbf{m}) = \mathbb{P}[\text{Influence}_2 \geq \tau_2, \text{Influence}_1 < \tau_1]; \quad (7)$$

$$F_3(\mathbf{W}, \mathbf{m}) = \mathbb{P}[\text{Influence}_1 \geq \tau_1, \text{Influence}_2 \geq \tau_2], \quad (8)$$

respectively. In other words, (6)-(8) are the *response functions* for the corresponding active states. In addition, we also reserve F_0 to denote the probability that an inactive node remains inactive:

$$F_0(\mathbf{W}, \mathbf{m}) = \mathbb{P}[\text{Influence}_1 < \tau_1, \text{Influence}_2 < \tau_2]. \quad (9)$$

Assume all nodes are initially inactive (i.e., not supporting any contents). A node is chosen uniformly at random and set to be active for all contents (i.e., 3-active when $M = 2$) as the *seed*. Other nodes start changing their states according to (6)-(8) synchronously at times $t = 0, 1, \dots$ (i.e., the influence starts propagating throughout the network). We assume once an

inactive node turns active, it cannot change its state. Therefore, the contagion process is monotone and will eventually stop, i.e., reach a steady state.

In fact, (6)-(9) provide a mutually exclusive partition of the two-dimensional space spanned by the \mathbf{m} vector into different regions for possible states. Figure 1 shows examples of space partitions when the two contents 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 supporting both content-1 and content-2. On the contrary, the negative correlation decreases the area for state-3, indicating the negative impact on the accumulative influence from each other.

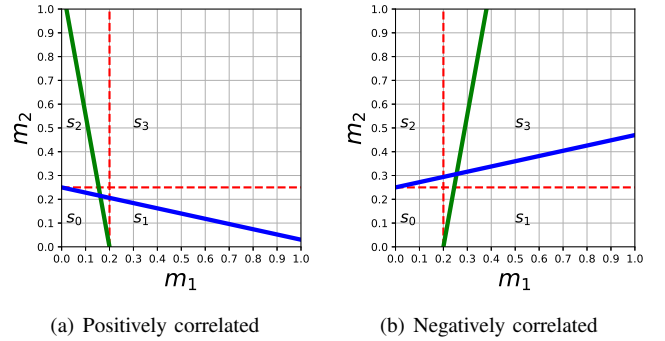


Fig. 1. An example of the two-dimensional vector space partition for possible states where the first dimension represents m_1 , and second dimension represents m_2 . In (a), content-1 and 2 are positively correlated where $w_{12} = 0.18$, $w_{21} = 0.22$; in (b) they are negatively correlated where $w_{12} = -0.18$, $w_{21} = -0.22$. In both (a) and (b), $\tau_1 = 0.2$, $\tau_2 = 0.25$. Green lines separate the space of supporting content-1 from not supporting content-1. Blue lines separate the space of supporting content-2 from not supporting content-2. Red dashed lines represent the partition when content-1 and content-2 are independent of each other. $s_i, i = 0, 1, 2, 3$ denote state-0, 1, 2 and 3, respectively.

III. EMERGENCE OF GLOBAL CASCADES

This section presents the analytical results of the probability and threshold of emergence for global cascades under the proposed MDTM model. We start by giving formal definitions of probability of emergence and cascade threshold for the considered influence propagation process.

Definition II.1 (Probability of Emergence of Global Cascades). Global cascades refer to how in an initially inactive population, a randomly chosen node set to be active for all contents (i.e., *seed*) can initiate an influence propagation that eventually reaches a positive fraction of active nodes in the entire population. Consider M contents for the influence propagation process, and there will be 2^M different states. Specifically, for $i = 1, 2, \dots, 2^M - 1$, let a random variable $S_i(n)$ denote the fraction of nodes reached and turned i -active in the population size of n by the seed, probability of emergence of global cascades for i -active nodes (PE_i) is defined as

$$\text{PE}_i = \mathbb{P}[\lim_{n \rightarrow \infty} S_i(n) > 0].$$

Definition II.2 (Cascade Threshold). Cascade threshold refers to a boundary in the parameter space that separates the regions where global cascades can take place (i.e., $\mathbb{P}[\lim_{n \rightarrow \infty} \cup_{i=1, \dots, 2^M-1} \{S_i(n) > 0\}] > 0$) from the regions where they cannot (i.e., $\mathbb{P}[\lim_{n \rightarrow \infty} \cup_{i=1, \dots, 2^M-1} \{S_i(n) > 0\}] = 0$).

Consider random graphs $\mathbb{G}(n, \{p_k\})$ generated by the configuration model. To study the influence propagation over network \mathbb{G} , we consider a branching process that starts by randomly choosing a node to be the seed, and then recursively discovers the set of nodes that are *reached* and *activated* by exploring its neighbors based on (6)-(8). We derive the *survival probability* of the aforementioned branching process through a *mean-field* analysis by using the method of probability generating functions (PGF) [22], [25].

Here we provide an analysis for $M = 2$ for simplicity. Analysis for $M > 2$ can be conducted in a similar manner. For $i, j = 1, 2, 3$, let $h_{ij}(x)$ denote the generating function for “the *finite* number of nodes reached and activated to become j -active by following a randomly selected edge from an i -active node.” Put differently, we have $h_{ij}(x) = \sum_{s=0}^{s=\infty} v_s x^s$, where v_s denotes the “probability that an arbitrary edge emanating from an i -active node leads to a *finite* component of j -active nodes of size s .” Similarly, let $H_i(x)$ denote the generating function for “the *finite* number of nodes reached and activated to i -active by following a randomly selected state-3 seed node”.

We now derive $h_{ij}(x)$ for each $i, j = 1, 2, 3$, recursively. Consider an i -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:

$$h_{ij}(x) = \sum_{k=1}^{\infty} \frac{p_k k}{\langle k \rangle} \left[F_0 + x \cdot F_j \cdot h_{jj}^{k-1}(x) + \sum_{\ell \in \{1, 2, 3\} - j} F_\ell \cdot h_{\ell j}(x)^{k-1} \right] \quad (10)$$

for $i, j, \ell = 1, 2, 3$.

We now explain each term appearing in (10). We first condition on the degree of node u being k , which is given by $\frac{p_k k}{\langle k \rangle}$. Conditioning on node u 's degree being k , the fraction of u 's neighbors that support content- i is given by

$$m_i = \frac{\mathbb{1}[\text{node } v \text{ is in state-}i \text{ or } 3]}{k}, \quad i = 1, 2.$$

Note that we add state-3 nodes for each content because state-3 supports both contents. Given the correlation matrix \mathbf{W} , we can obtain node u 's state based on (6)-(9). If node u remains inactive (i.e., state-0), it will have no offspring which explains the F_0 term in (10).

If node u is activated to j -active w.p. F_j , the number of nodes reached and activated to j -active will increase by one, and this is captured by the multiplicative term x before F_j in (10). In addition to this, the total size of this *branch* will also include all subsequent nodes that might be activated to state- j by u via its remaining $k-1$ edges, given the fact that node u is

reached and activated via an edge connecting node v . Besides, recall that the number of nodes reached and become j -active by node u via one of its $k-1$ edges is generated through $h_{jj}(x)$. By using the *powers property* of generating functions [22], and considering that node u influence its neighbors via these $k-1$ edges independently due to the locally treelike structure, we obtain the term $h_{jj}^{k-1}(x)$.

If node u is activated to other active states other than j , say state- $\ell \in \{1, 2, 3\} - j$, we do not need to count node u itself into the collection. We need to count all the subsequent nodes that may become j -active via the remaining $k-1$ edges of node u , which gives rise to the last term in (10). Above completes the derivation of (10).

Utilizing (10), we now derive generating function $H_i(x)$. For each $i = 1, 2, 3$, we have

$$H_i(x) = x^{\mathbb{1}[i=3]} \sum_{k=0}^{\infty} p_k \cdot h_{3i}(x)^k. \quad (11)$$

Here, the factor $x^{\mathbb{1}[i=3]}$ corresponds to the initial node selected arbitrarily and activated to state-3. The selected node has degree k w.p. p_k . The number of nodes it reaches and activated to i -active by each of its k links is generated through $h_{3i}(x)$. Similarly, by using the powers property of generating functions and averaging over all possible degrees, we obtain (11).

With equations (10) in hand, the generating function $H_i(x)$ can be computed in the following manner. Given any x , we can solve for the recursive relations (10) to obtain $h_{11}(x), \dots, h_{33}(x)$ ¹, which in turn will yield $H_1(x), H_2(x), H_3(x)$ in light of (11).

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 3-active triggers a *global cascade*. There exists a trivial solution $h_{ij}(1) = 1$ to (10) (yielding $H_i(1) = 1$) for each $i, j = 1, 2, 3$ 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* infected 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 $h_{ij}(1) = 1$ for all $i, j = 1, 2, 3$ is not a stable solution. In this case, there exist $p, q \in \{1, 2, 3\}$ so that $h_{pq}(1) < 1$ which in turn yields $H_q(1) < 1$. The *conservation of probability* property of generating functions indicates that $H_q(1) + \mathbb{P}[\lim_{n \rightarrow \infty} S_q(n) > 0] = 1$ if there is a component reached and become q -active by the seed node has *infinite* size. Therefore, we have $\text{PE}_q = 1 - H_q(1)$.

We can check the stability of the fixed point $h_{ij}(1) = 1$ for $i, j = 1, 2, 3$ by the linearization of the recursion (10) around it. This yields the Jacobian matrix $\mathbf{J}_{9 \times 9}$ in which

$$\mathbf{J}[a, b] = \frac{\partial h_{ij}(1)}{\partial h_{pq}(1)},$$

¹Note that, for $i, j = 1, 2, 3$, if an i -active node can never activate an inactive neighbor to j -active in a naive population, i.e., a population where all other nodes are state-0, this branch will have zero j -active offspring. In this case, $h_{ij}(x) = 1$, for all $|x| \leq 1$.

where $i, j, p, q = 1, 2, 3$; $a = \text{M2D}(i, j), b = \text{M2D}(p, q)$. M2D is a 3-base to 10-base converter, mapping the tuple ij to an integer ranging from one to nine for the ease of matrix indexing. For example, when we take the derivative of $h_{ij}(x)$ in (10) with respect to one of its inputs $h_{pq}(x)$ when $x = 1$ (denoted by $\frac{\partial h_{ij}(1)}{\partial h_{pq}(1)}$), the result corresponds to an element in the matrix \mathbf{J} with coordinates (a, b) , where $a = \text{M2D}(i, j)$ and $b = \text{M2D}(p, q)$ are integers within $1 \leq a, b \leq 9$.

If all eigenvalues of \mathbf{J} are less than one in absolute value (i.e., if the spectral radius $\rho(\mathbf{J})$ of \mathbf{J} satisfies $\rho(\mathbf{J}) \leq 1$), then the solution $h_{ij}(1) = 1$ for all $i, j = 1, 2, 3$ is stable. In this case, the fraction of active nodes will tend to zero as the number of nodes n goes to infinity. In contrast, if $\rho(\mathbf{J}) > 1$, the trivial fixed point is not stable, which indicates that the branching process is in the *supercritical* regime (i.e., there is a positive probability that the branching process will lead to an *infinite* activated component). In this case, the fraction of active nodes infected will be strictly greater than zero as the number of nodes n goes to infinity.

Finally, we conclude that the cascade threshold for an active states, i.e., the boundary that separates the parameter regions where $\mathbb{P}[\lim_{n \rightarrow \infty} \cup_{i=1, \dots, 2^M-1} \{S_i(n) > 0\}] > 0$ from those that yield $\mathbb{P}[\lim_{n \rightarrow \infty} \cup_{i=1, \dots, 2^M-1} \{S_i(n) > 0\}] = 0$ is given by $\rho(\mathbf{J}) = 1$ for $i = 1, 2, 3$. The probability of emergence of i -active global cascades (with a 3-active initiator) is thus given by $\text{PE}_i = \lim_{n \rightarrow \infty} \mathbb{P}[S_i(n) > 0] = 1 - H_i(1)$.

IV. NUMERICAL RESULTS

We next present simulation results with an eye towards validating our analytical results (which are exact in the limit of the number of nodes n goes to infinity) in the finite node regime. In doing so, we also aim to shed light on how various parameters used in the model affect the spreading process, e.g., in terms of probability of emergence of global cascades for active states. Throughout, the contact network $\mathbb{G}(n, \{p_k\})$ was generated via the configuration model with Poisson degree distribution with mean degree $\langle k \rangle$ and $n = 2,000,000$ vertices. To generate the plots, we took an average over 1,000 independent trials where, in each trial, a new contact network was generated. In what follows, we assume there are two contents, i.e., content-1 and content-2, in the influence propagation process. Nodes in the population can be in states 0 to 3 depending on if they support these two contents.

In the first experiment, we investigate the impact of network structure on the probability of emergence of global cascades for active states 1, 2 and 3. Figure 2 shows the probability of emergence when content-1 and content-2 have the positive and negative correlation given in Figure 1. First, we see the simulation results match the analytical solutions with near-perfect accuracy. This validates our analytical results and confirms the usefulness of our results in the finite node regime.

In addition, as mean degree $\langle k \rangle$ increases, there exist two *phase transition* points, i.e., the point from which the probability of emergence change from zero to a positive value or vice versa, for active states that have transition points. These two phase transition points are reported in many

works [11], [15], [26] and provide insights on the impact of network connectivity on the emergence of global cascades. 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.

It is also observed that PE_3 remains zero in Figure 2(b) while in 2(a), it has two transition points. In addition, PE_1 has lower values in Figure 2(b) than it does in 2(a). These differences show the impact of positive and negative correlation on the emergence behavior of global cascades. When content-1 and 2 are positively correlated, nodes who support content-1 are likely to support content-2 as well. In this case, state-1 nodes contribute positively to the emergence of state-3's global cascades. When negatively correlated, due to the negative impact from content-1, node that supports content-1 will not support content-2, leading to $\text{PE}_3 = 0$ in Figure 2(b).

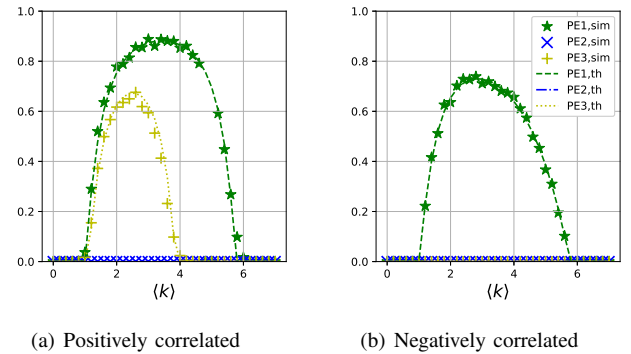


Fig. 2. Probability of Emergence of global cascades with varying mean degree $\langle k \rangle$ for state-1, 2, and 3, i.e., PE_1 , PE_2 and PE_3 , respectively. Content-1 and 2 are positively correlated in (a), where $w_{12} = 0.18$, $w_{21} = 0.22$, and negatively correlated in (b), where $w_{12} = -0.18$, $w_{21} = -0.22$. In both (a) and (b), $\tau_1 = 0.2$, $\tau_2 = 0.25$. Simulation results (marked as *sim* in the legend) show near-perfect agreement with our theoretical results (marked as *th* in the legend).

In Figure 3, we give an example of the \mathbf{m} vector space partition using a mixed correlation between content-1 and 2. The space is split into four sub-areas by four line segments. The four segments intersect in the center point with coordinates $[0.12, 0.2]$. The coordinates of the intercept on the x-axis are $[0.18, 0.0]$, followed by $[0, 0.22]$, $[1, 0.87]$ and $[1, 0.2]$ for the rest of the end points clockwise. This space partition exemplifies mixed correlation between contents, i.e., having both positive and negative correlation. For example, before m_2 reaches 0.2, content-1 and content-2 are positively correlated, while after m_2 exceeds 0.2, they are negatively correlated. Influence calculation and response functions are adapted accordingly to reflect this mixed correlation. With this complicated correlation, our analytical results still accurately

predict the probability of emergence of global cascades for each active state within the limit node regime, even around the two phase transition points, demonstrating the validity of the analysis for the proposed model.

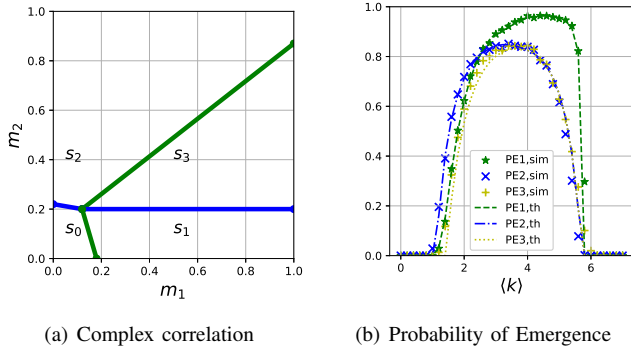


Fig. 3. Space partition of mixed correlation (a), and the corresponding probability of emergence for active states (b).

V. CONCLUSION

In this paper, we propose the *multi-dimensional threshold model* (MDTM) for the simultaneous spread of multiple correlated contents over complex networks. We provide analytical solution for the probability of emergence of the considered influence propagation process. Numerical results by extensive simulations validated the analytical solution with near-perfect match. The analytical results disentangle the interplay between multiple factors on the emergence behavior of global cascades, including the underlying network structure and the correlation among contents. 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] T. Kuran, "Sparks and prairie fires: A theory of unanticipated political revolution," *Public Choice*, vol. 61, pp. 41–74, Apr. 1989.
- [9] 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.

- [10] 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.
- [11] 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.
- [12] 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.
- [13] O. Yağan and V. Gligor, "Analysis of complex contagions in random multiplex networks," *Phys. Rev. E*, vol. 86, p. 036103, Sep 2012.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] Y. Xu and W. Wang, "Characterizing the spread of correlated failures in large wireless networks," in *2010 Proceedings IEEE INFOCOM*, pp. 1–9, 2010.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] M. Newman, "Epidemics on networks," in *Networks*, 2018.
- [26] Y. Zhuang, A. Arenas, and O. Yağan, "Clustering determines the dynamics of complex contagions in multiplex networks," *Physical Review E*, vol. 95, p. 012312, Jan. 2017.