Information and Influence Propagation in Multi-layer Networks: the Impact of Clustering and Multi-dimensional Content Spreading

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Yong Zhuang

B.Eng. Software Engineering Zhejiang University
M.Phil. Computer Science and Information Engineering National Taiwan University
Carnegie Mellon University
Pittsburgh, PA

September, 2019
Copyright © 2019 Yong Zhuang
All Rights Reserved
For Shimeng, my parents, my mother-in-law, and my father-in-law.
Thesis Committee Members

Prof. Osman Yağan (Advisor)
Department of Electrical and Computer Engineering
Carnegie Mellon University

Prof. José M. F. Moura
Department of Electrical and Computer Engineering
Carnegie Mellon University

Prof. H. Vincent Poor
School of Engineering and Applied Science
Princeton University

Prof. Junshan Zhang
School of Electrical, Computer and Energy Engineering
Arizona State University
Abstract

Spreading processes, such as the spread of the Zika virus (in the context of infectious disease propagation) or fake news (in the context of information propagation), are fundamental phenomena occurring over real-world social networks. These processes are closely associated with the nation’s stability, economy, and security, hence a large body of research has investigated their characteristics and how they interact with the underlying contact network. At a high level, a spreading process usually begins with an initially localized effect, such as when a person posts a rumor on their social media accounts, or when an infectious disease starts to spread in a given locality. A key question in this context is whether the initially localized effect continues to grow and reaches a significant fraction of the population, or dies out during the initial stages, i.e., whether an outbreak occurs or not. Indeed, if there is an outbreak, then it would be crucial to predicting how likely and widespread an outbreak would be.

Existing research in the area of mathematical modeling of spreading processes attempts to tackle the above question by means of providing mathematically-tractable models that resemble both the structure of real-world contact networks and the dynamics of real-world spreading processes. The proposed models pave the way for computing key quantities associated with the spread of the process on a given contact network, such as the probability of an outbreak, i.e., the probability that an initially localized effect would eventually reach a positive fraction of the nodes, and the expected outbreak size, i.e., the expected fraction of individuals reached/affected by the initial effect. Existing models, however, ignore several crucial aspects of real-world contact networks and spreading processes. In particular, most of the proposed models for the underlying contact network are single-layer networks, indicating that individuals participate in only one network, such as being only connected on Facebook. However, in the real world, individuals essentially participate in multiple-layer networks, such as simultaneously existing on Facebook, Twitter, etc. That is, multiple-layer networks provide a comprehensive and more realistic model for real-world social networks. In addition, most of the proposed models
assume that the underlying contact network is *tree-like* with no cycles and a vanishingly small clustering coefficient. Clustering is a propensity that two friends of one individual are more likely to know each other, and has been reported as an important topological property for spreading processes. Hence, the predictions obtained on network models without clustering are expected to be significantly inaccurate when compared to spreading processes propagating on real-world social networks that are typically clustered.

In the context of models for the dynamics of spreading processes, existing models do not consider the possible varieties of spreading contents. For example, existing models of influence propagation typically assume that an individual could be in one of *only* two states, namely, being inactive (not affected by an influence) or active (affected by an influence and is actively spreading it). In reality, individuals could have a richer set of possible states representing how strongly they would spread the influence (e.g., inactive, active, hyper-active, etc.). Besides, existing models also assume that there exists only one influence in a propagation process such as the spread of the purchase behaviors of iPhone and Apple HomePod. However, in the real-world influence propagation process, there may exist multiple influences simultaneously. Moreover, the spread of multiple influences could correlate with that of others. As a consequence, broader frameworks where an individual could be in one of many possible states or where multiple correlated influences could simultaneously exist are needed to resemble real-life spreading processes.

This dissertation focuses on two representative spreading processes, information propagation and influence propagation. Information propagation indicates a class of spreading processes which happen after only a single copy is received. In contrast, influence propagation is a class of spreading processes where social reinforcement from multiple copies plays an important role. For these two propagation processes, we addressed the aforementioned limitations of current literature by including several important characteristics of the real-world spreading processes. In particular, in terms of modeling underlying contact networks of information and influence propagation, we apply network models with multiple layers and clustering. Given
the fact that individuals could participate in multiple social networks at the same time and
the prevalence of clustering in social networks, clustered multiple-layer network models would
help us to accurately model real-world networks. Regarding modeling dynamics of real-world
spreading processes, we first consider influence propagation with multiple stages, which enables
us to model the case where individuals could have different levels of influence on her neighbors.
In addition, considering the existence of multiple correlated influences, we also propose a new
threshold model, the vector threshold model, which is the first model enabling us to study the
spread of multiple influences.

Going further, concerning each propagation process, we derive analytic results to the two
key metrics of spreading processes: the probability of an outbreak and the expected size of an
outbreak (if there exists one). The analytical solutions are confirmed via extensive simulations.
Then, with these analytical solutions, we could comprehensively study each spreading processes
by means of observing how the two metrics change as we vary the parameter to control the
level of each property. A key takeaway from this dissertation is that the assortativity (i.e.,
correlation between the degrees of connected pairs) generated by the nature of
multiple layers and multiple link types plays an important role in spreading pro-
cesses. In particular, in the study of information propagation, we showed that assortativity
has a multi-faceted impact on propagation processes. When the degree is at a low level, the
assortativity helps the information spread a larger portion of populations, while it reduces the
propagation process when the degree is at a high level. In addition, in the study of influence
propagation, we showed that the level of assortativity would not only change the expected size
of outbreak, but also the number of phase transitions.
Acknowledgements

First of all, I would like to thank my Ph.D. advisor, Dr. Osman Yağan for bringing me to CMU. Osman provided numerous help on my research and life, and also taught me a lot about doing research, life, etc. I enjoyed working with Osman a lot. He encouraged me to challenge myself instead of only working on the problems I know I can solve. It is my great honor and luck to work with Osman. Also, I appreciate Osman’s support on my last semester for job search. I would never forget all the help you offered. In addition, I am indebted to Dr. José M. F. Moura, Dr. H. Vincent Poor, and Dr. Junshan Zhang for serving on my Ph.D. thesis committee and giving me a lot of helpful and constructive suggestions. Your vision of research helped me a lot.

After I graduated from Zhejiang University, I joined the master program at National Taiwan University (NTU). At NTU, I met several people important to my career and life. First, I would thank Dr. Chih-Jen Lin. The very most important thing I learned from Dr. Lin is the attitude towards research. He always reminds us that we should be down to earth. In addition to that, he also taught me that to study in the U.S. does not only means learn knowledge there, but I should also learn their culture. Next, I would thank Yuchin Juan and Wei-Sheng Chin. It is always a great experience that working with you together. With you, we won many first places in data mining competitions, e.g., KDD Cup 2013, the CTR prediction competition hosted by Criteo, the competition hosted by Avazu, etc. We also grew from 3 idolts to 4 idolts. I am grateful to Tong Yu who changed my life. Without him, I do not know where I am now. I also need to thank many friends from 528, Ching-Pei Lee, Chieh-Yen Lin, Cheng-Hao Tsai, Bo-Wen Yuan, Chien-Chih Wang, Chia-Hua Ho, Bo-Yu Chu, Ya Chu, Xiaocong Zhou, Chi-Cheng Chiu, Mu-Chu Lee, Tzu-Ming Kuo, Wei-Cheng Chang, Po-Wei Wang.

After finishing the master program at NTU, I joined the Ph.D. program at Carnegie Mellon University (CMU). I am also very grateful to Prof. Ole Mengshoel, Prof. Joy Ying Zhang, Prof. Paul Steif, Prof. Bruno Sinopoli, Prof. Pulkit Grover, and Prof. Soummya Kar. Also,
I would thank CMU staff: Stephanie Caruso, Nathan Snizaski, Sari Smith, Lauren Schachar, Wendy Fong,, who gave me a lot help at CMU. In Silicon Valley, I met so many friends here, and I do appreciate the kindful help from Guan-Lin Chao, Bing Liu, Yuan Tian, Xiao Wang, Yingrui Zhang, Chiao-Ni (Maggie) Sun, Ming Zeng. William, Chirs Lee, David Cohen, Suyoun Kim, Benjamin Elizalde, Ervin Teng, Chih Hu Dongzhen Piao, Le Nguyen, Shijia Pan, Xinlei Chen, Jun Han.

After one and half years in Silicon Valley, I moved to Pittsburgh with my advisor and the whole research group. In Pittsburgh, I met so many great friends here. The very first friend I met here is Xiaonan Huang, and I lived with him together for 2 years. I would never forget the night we went back home together at late night. Also, I do enjoy the time with friends, Bowen Yu, Wenhuan Sun, Peisen Gu, Candice Ji, Dennis Zhao, Yangyang Qiao, Xi Wang, Yuqing Shi, Chenyu Wang, Tess Hellebrekers, Eric Markvicka, Stacie Markvicka, etc.

Next, I would also thank my comrades in the lab. Yingrui Zhang, Rashad Eletreby, Samarth Gupta, Mansi Sood, and Vaibhav. I would say thanks to all A-level folks, Senbo Fu, Jing Huang, Ke Wang, Yu Wang, Jiyuan Zhang, Guanglin Zhao, Zhipeng Zhao, Doru Thom Popovici, John Filleau, Daniele Spampinato, Marie Nguyen, Mark Blanco, Elliot Binder, and Shaun Stevens.

In addition, I would like to thank all great friends I met in Criteo, Suju Rajan, Pranjul Yadava, Avradeep Bhowmik, Suchismit Mahapatra, Zhengming Xing, Huifang Shen, Hamid Maei, Zheng Guo, Kaiyu Yang. It is a great experience for me to spend a summer with you.

Besides, I would thank the friends and the advisors during my undergraduate. I would thank Prof. Guillaume Wisniewski, Prof. François Yvon, Dr. Li Gong, and Qian Yu. Also, Wei Song, Wei Wang, Xingang Wang, Xuepeng Shi, Pengxiang Cui, Zhenxing Zhang, Nan Yang, etc.

In addition, I do enjoy the time at Universidad Rovira i Virgili in Spain, and I like to express my appreciation to Prof. Alex Arenas, Prof. Sergio Gómez, Prof. Clara Granell, Dr. Joan T. Matamalas, Prof. Albert Sole, Dr. Hongru Wu, Dr. Elisa Omodei.

I apologize to all of the friends that I have inadvertently left out who made my life memo-
rable.

I would not have been able to achieve anything without the constant encouragement and endless love of my parents. In addition, I also want to say thank you to my mother-in-law and father-in-law. They regard me as their son and give me constant support in my career and life choices. Finally, I would like to thank my wife, Shimeng. Words can scarcely express my gratitude and appreciation to her for all her support and sacrifice over the years. I would like to dedicate this dissertation as a gift for our first wedding anniversary.

The work presented in this thesis has been supported in part by National Science Foundation through grant CCF #1422165, the Army Research Office under Grant Number W911NF-17-1-0587, and in part by the Department of ECE at Carnegie Mellon University.
## Contents

### I Introduction and Related Work

1 Introduction
   1.1 Motivation and Research Problem .......................... 2
   1.2 Related Work and Our Contributions on Information Propagation .......................... 6
   1.3 Related Work and Our Contributions on Influence Propagation .......................... 9

### II Information Propagation

2 Information Propagation over Clustered Multi-layer Networks
   2.1 Introduction .................................................. 13
   2.2 Problem Formulation ........................................ 16
      2.2.1 Random Graphs with Clustering ......................... 16
      2.2.2 Multilayer Network Models with Clustering .......... 19
      2.2.3 Information Propagation Model: SIR .................. 21
      2.2.4 Problems of Interest .................................. 22
   2.3 Technical Background ....................................... 24
   2.4 Main Results ................................................. 25
      2.4.1 Information Propagation via Single Edges in Network $\mathcal{F}$ ................. 27
      2.4.2 Information Propagation via Triangles in Network $\mathcal{F}$ ..................... 28
      2.4.3 Computing the Final Epidemic Size .................... 30
      2.4.4 The Relationship between Our Analysis and Some Previous Studies .......... 34
   2.5 Numerical Results and Discussion .......................... 34
      2.5.1 Networks with Doubly Poisson Distributions .......... 34
      2.5.2 Networks with Power-law Degree Distributions .......... 36
      2.5.3 How does Clustering Affect the Threshold and Size of Information Epidemics? .... 37
      2.5.4 How does $\alpha$ Affect the Information Propagation Dynamics? ............. 41
   2.6 Chapter Summary ............................................. 46

### III Influence Propagation

3 Clustering determines the dynamics of complex contagions in multiplex networks
   3.1 Introduction .................................................. 49
3.2 Model: structure and dynamics ............................................. 52
  3.2.1 Random Graphs with Clustering .................................. 52
  3.2.2 Multi-layer and Multiplex Network Models ...................... 53
  3.2.3 Content-dependent Linear Threshold Model for Social Contagion .. 55
  3.2.4 The Problem .......................................................... 56
3.3 Condition and Probability of Global Cascades .......................... 57
  3.3.1 Influence Propagation via Red Single Edges ...................... 59
  3.3.2 Influence Propagation via Red Triangles ......................... 60
  3.3.3 Deriving the Condition for Global Cascades ...................... 61
3.4 Expected Cascade Size .................................................... 63
3.5 Numerical Results .......................................................... 67
  3.5.1 Networks with Doubly Poisson Distributions ..................... 67
  3.5.2 How does Clustering Affect the Cascade Size? ................. 69
3.6 Comparison between monoplex and Multiplex Networks .................. 71
  3.6.1 Multiplex Networks with Limited Assortativity .................. 72
  3.6.2 Multiplex Networks with Assortativity .......................... 73
  3.6.3 Two vs. Four Phase Transitions .................................. 76
3.7 Chapter Summary ............................................................ 77

4 Multi-Stage Complex Contagions in Random Multiplex Networks 79
  4.1 Introduction .............................................................. 79
  4.2 Model Definition:Networks and Dynamics ............................ 82
    4.2.1 Multi-layer and multiplex network models ...................... 82
    4.2.2 Multi-stage content-dependent linear threshold model ......... 84
  4.3 Main Results ............................................................. 86
    4.3.1 Expected cascade size and the condition to have a global cascade .. 87
    4.3.2 Probability of triggering a global cascade ...................... 92
  4.4 Numerical Results ....................................................... 96
    4.4.1 The agreement between our analysis and simulations ............ 96
    4.4.2 The impact of hyper-influencers in multiplex networks ......... 98
  4.5 Conclusion and Future Work ............................................. 102
  4.6 The impact of hyper-influencers on the global cascade boundary ........ 103
  4.7 The impact of hyper-influencers on the probability and expected size of global cascades .................................................. 105
  4.8 The Impact of the network size on the comparison between analytic results and experiments .................................................. 106

5 A Vector Threshold Model for the Simultaneous Spread of Correlated Influence 108
  5.1 Introduction .............................................................. 108
  5.2 Model and Problem Definition ......................................... 111
    5.2.1 The network model: The configuration model .................... 111
    5.2.2 The vector threshold model: A threshold model with multiple correlated contents .................................................. 112
    5.2.3 Problem Definition .................................................. 114
## Main Results

5.3 Main Results ........................................... 115
  5.3.1 Analysis of the expected size of global cascades .......... 115
  5.3.2 The condition of the existence of global cascades .......... 119

5.4 Numerical Experiments ..................................... 120
  5.4.1 The agreement between our analysis and experimental results .... 120
  5.4.2 The impact of correlations between spreading contents .......... 121

5.5 Conclusion and Future Work .................................. 123

### IV Concluding Remarks and Future Work

6 Concluding Remarks .............................................. 126

7 Future Work .................................................... 129
List of Figures

2.1 Illustration of process of Configuration Model. .................. 16
2.2 ................................................................. 18
2.3 ................................................................. 18
2.4 Illustration of a random network with clustering. In part (a), node 2 has 2 single edges and one triangle, whereas in part (b) it has zero single edges and two triangles. For the network in Figure 2.2, the global clustering coefficient is 0.2 while the local clustering coefficient is 0.3, while for the network in Figure 2.3, these coefficients are given by 0.4 and 0.7, respectively. .............. 18
2.5 Nodes in the upper circle and lower circle indicate the individuals in social network and physical network respectively. The nodes connected by a red line cross two networks mean they are the same individual existing in two networks. Green nodes in the upper circle belong to $\mathbb{F}$, while blue nodes in $\mathbb{W}$. Some of the nodes connected across the two networks by a red line indicates the fact that they represent the same individual. .......................... 19
2.6 Illustration of the branching process. The children of each individual node is identified recursively, while taking into account whether or not the information is transferred from the parent node to the child node. The initial vertex that starts the information is regarded as the 0th generation, and we are interested in deriving the limiting behavior of the total number of nodes reached and informed as the number nodes $n \to \infty$. ........................................ 25
2.7 The top vertex $u$ is infected, and information is transferred through an edge only if it is occupied (happens with probability $T_f$ in network $\mathbb{F}$). .............. 28
2.8 Simulation for doubly Poisson degree distributions. ................. 35
2.9 Simulation for Power-law degree distributions. ..................... 37
2.10 Comparison of the size of information epidemics between Non-clustered and Non-clustered networks (NN), Clustered and Non-clustered networks (NC), and Clustered and Clustered networks (CC). Plots are obtained from our analytical results. The value following the model abbreviation indicates the amount of overlapping between two networks. For example, NC-0.5 means that $\alpha = 0.5$. 39
2.11 Comparison of the epidemic boundary under several cases; the north and east of each curve specifies the region of $(T_f, T_w)$ values for which epidemics are possible, while the south and west part of each curve stands for the region where epidemics can not take place. Resulting statistics for clustering and assortativity is given in Table 2.3. ............................ 41
2.12 Illustration of how clustering affects the size of epidemics when $T_f = T_w = 0.3$. 42
2.13 Illustration of the effect of $\alpha$. SN is the abbreviation for Simplex Network where colors of the edges are ignored, while MN indicates the Multiplex Network case where only stubs of the same color are connected together. (Inset) The plots for $\alpha = 0.01$ are shown at a higher resolution near the phase transition point.

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

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

3.3 Assume that the top vertex $u$ is active in network $R$. Whether it will activate the other two nodes will depend on whether $v$ and/or $w$ are $R$-vulnerable.

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

3.5 Illustration of the effect of clustering coefficient on the expected cascade and probability of global cascades. We fix $\tau = 0.18$, $c = 0.25$, and $\alpha = 0.5$, then vary the degree parameter $\lambda$ defined in Table 3.1. We see (a) the probability to trigger a global cascade; (b) the global clustering coefficient introduced in Section 3.2.1; and (c) the expected cascade size.

3.6 We show the cascade regions in the Degree Parameter-Threshold plane when $\alpha = 0.5$, $\tau = 0.18$, and both networks follow doubly Poisson distributions as described in Table 3.1. Clustering increases as $\eta$ increases.

3.7 Comparison between monoplex networks and multiplex networks with limited assortativity. In (a) and (b), we fix the threshold $\tau = 0.15$, the content parameter $c = 1$, then vary the degree parameters in (3.13). For the networks obtained by projected theory and the networks in multiplex theory with $\alpha = 0.99$, assortativity is negligible. However, when $\alpha = 0.1$, the assortativity coefficient of the networks in the multiplex theory become significant; e.g., it can be up to 0.21.

3.8 Comparison between monoplex networks and multiplex networks with assortativity. Similar with the observation in Figure 3.7, networks in the projected theory and in the multiplex theory with $\alpha = 0.99$ have negligible assortativity coefficients. However, for the networks of multiplex theory with $\alpha = 0.1$, assortativity coefficient ranges from 0.19 to 0.79. In general, assortativity increases with increasing $\lambda_r$ and $\lambda_b$ in the multiplex theory.

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

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

4.3 Simulations for doubly Poisson degree distributions, $\alpha = 0.5$, $\tau_1 = 0.18$, and $\tau_2 = 0.32$. The weight of hyper-influencers is taken to be $\beta = 1.5$. 

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

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

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

4.7 Given $\tau_1 = 0.15$ and $\alpha = 0.5$, we vary the mean degree $\lambda$ and $\tau_2$ to plot the global cascade region for several $\beta$, 1.5, 2.0, and 3.0. 

4.8 Given the weight of extra influence $\beta = 1.5$ and $\alpha = 0.5$, we vary the degree parameter $\lambda = \lambda_b = \lambda_r$ and $\tau_2$ to plot the region where there exists a global cascade for several $\tau_1$, 0.15, 0.18, and 0.2. Both of the edges are assigned by the doubly Poisson distribution in Section 4.4.1. 

4.9 The comparison between different $\beta$ for the probability of triggering a global cascade. 

4.10 The comparison between different $\beta$ for the expected global cascade size. 

4.11 The comparison between the asymptotic and the finite-size results for different size of networks. 

5.1 An illustration of different splits on the parameter space for different correlations. 

5.2 The different region split strategies in Section 5.4.1. 

5.3 There is an agreement between our analysis and numerical results. The solid curve indicates the results obtained from our analysis, while the symbols are the results obtained from numerical results. The parameter space is split as shown in Figure 5.2(a).
5.4 There is an agreement between our analysis and numerical results. The solid curve indicates the results obtained from our analysis, while the symbols are the results obtained from numerical results. The parameter space is split as shown in Figure 5.2(b). ................................................................. 122

5.5 The different region splits for different correlations between contents for Section 5.4.2. The gray dashed lines indicate the region split strategy of the independent correlation. ................................................................. 123

5.6 The comparison between different correlations on the expected size of global cascades. These results are obtained by analytic results. ................................. 123
List of Tables

1.1 Table for related work on information propagation. .......................... 8
1.2 Table for related work on influence propagation .............................. 11

2.1 Statistics of network $H$ under the setting of Figure 2.10. GC denotes global clustering coefficient, while LC denotes the local clustering coefficient. .... 39
2.2 Parameters of the doubly Poisson distribution. In Figure 2.11 we set $\lambda_F = \lambda_W = 0.5$. We use $\lambda_F = 0.36$ and $\lambda_W = 0.5$ for Figure 2.12. ......................... 40
2.3 Statistics corresponding to the network $H$ in the setting of Figure 2.11. .... 41
2.4 Comparison of the assortativity values observed in the setting of Figure 2.13 for to the Simplex Network (SN) and the Multiplex Network (MN) case for different $\alpha$. As expected, for the simplex network case the degrees of the nodes are uncorrelated and assortativity is thus zero. The multiplex case exhibits assortative mixing, with the correlations getting more significant with decreasing $\alpha$. ................................................................. 43

3.1 Parameters of the doubly Poisson distribution. This choice ensures that the mean and variance of the total degree distribution (single plus triangle edges) in $B$ are independent of $\eta$, while its clustering varies greatly as $\eta$ varies in $(0, 4)$. In Figure 3.4 we set $\lambda = 0.5$. ................................................................. 69
Part I

Introduction and Related Work
Chapter 1

Introduction

1.1 Motivation and Research Problem

Spreading processes are fundamental phenomena occurring over networks and widely exist around us in different forms. Examples include the spread of information, rumors, influence, and diseases. These processes are closely associated with the nation’s stability, economy, and security. For example, as a spread of infectious disease, Zika virus is still a major challenge to global health [78] that was initially limited to Africa, but has spread throughout the world. In another instance, the spread of fake news could be a damage to the nation’s security, e.g., fake news has affected the 2016 U.S. presidential election [3]. Therefore, a large body of research has investigated their characteristics and how they interact with the underlying contact network, since a good understanding of spreading processes helps us prevent the unexpected spreading processes (e.g., fake news) or boost the spread of beneficial processes (e.g., the spread of influence of accepting vaccination).

At a high level, a spreading process usually begins with an initially localized effect. For example, in 2003, Severe Acute Respiratory Syndrome (SARS) as a severe epidemic first emerged in Guangdong Province, China, and then its outbreak spread to other cities in China [30]. Then, studying a spreading process boils down to understanding what the final state of a localized effect would be. In particular, a key question in this context is whether an initially localized effect continues to grow and reaches a significant fraction of the population, or dies out during initial stages, i.e., whether an outbreak occurs or not. Indeed, if there is an outbreak, then
it would be crucial to predicting the fraction of individuals who the initially localized effect reach, i.e., the outbreak size.

Existing research in the area of mathematical modeling of spreading processes attempts to tackle the above question by means of providing mathematically-tractable models that resemble both the structure of real-world contact networks and the dynamics of real-world spreading processes. The proposed models pave the way for computing key quantities associated with the spread of the process on a given contact network, the probability of an outbreak, i.e., the probability that an initially localized effect would eventually reach a positive fraction of the nodes, and the expected outbreak size, i.e., the expected fraction of individuals reached/affected by the initial effect. However, existing literature ignore several crucial aspects of real-world contact networks and spreading processes. This dissertation aims at resolving the limitations of existing literature of modeling real-world spreading processes, so that we could model them more accurately and better understand real-world spreading processes. In particular, we focus on two representative spreading processes, information propagation and influence propagation. Information propagation indicates a class of spreading processes which happen after only a single copy is received. In contrast, influence propagation is a class of spreading processes where social reinforcement from multiple copies plays an important role.

Regarding the existing literature of information propagation, most of the proposed models for underlying contact networks are single-layer networks, indicating that individuals participate in only one network, such as being only connected on Facebook. However, people essentially participate in multiple networks, because the advances in communication technologies and cyber-physical systems make network structures even more complex. For instance, in social networks, after receiving a message from a friend in face-to-face communication, a person could immediately send the message to his/her online friends via Facebook or Twitter. That is, multiple-layer networks provide a comprehensive and more realistic model for real-world social networks. In addition to the topological property of multiple layers, Watts et al. [95] reported that in many biological, technological and social networks, there is a propensity that two friends
of one individual are more likely to know each other. Researchers named it as “clustering,” and has reported it as an important topological property for spreading processes [95]. However, most of existing literature still assume that the underlying contact network is tree-like with no cycles and a vanishingly small clustering coefficient. To consider this topological property, we study the information propagation over clustered multi-layer networks. In this case, we could accurately model real-world contact networks in order to have a better understanding of real-world information propagation.

Concerning influence propagation, we first reveal the influence propagation in clustered multiplex networks. As explained in the study of information propagation, clustered multiplex networks make it possible for us to model the varieties of relationships (e.g., friends, family, office-mates, etc.) and the propensity that two friends of one individual are more likely to know each other. Through this work, we found that most existing influence propagation models assumed that there are only two possible states for a given influence - i.e., to be active or inactive for a given influence - while in the real world, the possible reaction to a given influence could be a richer set of active states. For example, followers of a radical organization or a revolutionary movement may have varying levels of commitment to the cause, or varying desire and ability to recruit new members. To consider the existence of multiple active states in real-world influence propagation, we studied influence propagation with multiple stages over multiplex networks. Besides, only considering the natures of multiple levels of states or multiple link types is still not enough to accurately capture the characteristics of influence propagation. In a real-world influence propagation process, there could exist multiple influences simultaneously, and moreover, the spread of these influences could be correlated with each other. For instance, the spread of the purchase behaviors of different products from the same company might exhibit positive correlation; e.g., an individual who already purchased an iPhone might be more easily influenced by their friends to buy the Apple HomePod. In contrast, one’s opinions on universal health care and proposed tax relief for “wealthy” individuals would be expected to have a negative correlation. However, existing literature assumed that only one influence
content exists in an influence propagation process. To drop this assumption, we proposed a new spreading model, a vector threshold model, in order to consider a propagation process in which multiple influences spread simultaneously.

To sum up, by including several important topological and dynamical properties in real-world spreading processes, this dissertation uses a modeling-based approach to comprehensively study information and influence propagation. In particular, to accurately model real-world contact networks and spreading processes, in our mathematical modeling, we consider the natures of multiple layers and clustering in modeling underlying contact networks. In modeling the dynamics of real-world spreading processes, we consider the simultaneous existence of multiple correlated influences and multiple possible reactions to a spreading influence. Then, we derive analytical solutions to the two key metrics to the spread of an initial localized effect: the probability of an outbreak and the expected size of an outbreak (if there is one). With these analytical solutions, we could comprehensively study each spreading processes by means of observing how the two metrics change as we vary the parameter to control the level of each property. A key takeaway from this dissertation is that the assortativity (i.e., correlation between the degrees of connected pairs) generated by the nature of multiple layers and multiple link types plays an important role in spreading processes. In particular, in the study of information propagation, we showed that assortativity has a multi-faceted impact on propagation processes. When the degree is at a low level, the assortativity helps the information spread a larger portion of populations, while it reduces the propagation process when the degree is at a high level. In addition, in the study of influence propagation, we showed that the level of assortativity would not only change the expected size of outbreak, but also the number of phase transitions.

The rest of the dissertation is organized as follows. In the rest of this chapter, we introduce the related work on information propagation and influence propagation, respectively. In part II, we present the work on information propagation (Chapter 2). In part III, we present the work on influence propagation on clustered multiplex networks (Chapter 3), multi-stage influence
propagation (Chapter 4), and influence propagation with the simultaneous spread of correlated content (Chapter 5). In Part IV, we conclude the dissertation and give a discussion on further work.

1.2 Related Work and Our Contributions on Information Propagation

In the last decades, there are significant advances in this field of research. This section briefly introduces some of the related work on information propagation.

**Information propagation over non-clustered single layer networks:** One of the pioneering work in the field of studying information propagation can be found in [66]. This work showed that a large number of standard epidemiological models could be solved under a wide variety of networks. Through this work, it is clear that each spreading process has two important building blocks, underlying contact network models and spreading models. Underlying contact network models determine the connection information between individuals, while spreading models describe how each individual interacts with other individuals in networks. Then, Newman [66] assumed that the spreading model followed the class of susceptible/infective/removed (SIR) models, while the underlying network model is the configuration model. Given these assumptions, Newman gave the exact solutions for the probability of an outbreak, the size of an outbreak (if there is one), and a number of other quantities of interest. After this pioneering research, researchers began to branch into two directions to get a better understanding of information propagation by exploring various network models and spreading models.

**Information propagation over clustered networks:** Watts et al. [95] first reported that many biological, technological and social networks have a topological property named “clustering,” which is the propensity that two friends of one individual are more likely to know each other. However, existing research in the area of mathematical modeling of spreading
processes could not model networks with a tunable level of clustering. Then, Newman [68] and Miller [60] proposed a new network model, a generalized configuration model, which enables networks to have a tunable clustering coefficient. In the generalized configuration model, there are two types of edges, single edges and triangles. Single edges mean edges do not participate in any triangles, while triangles are only used to form triangles. Consider a vertex set \( V = 1, 2, \ldots, n \), where each vertex is independently assigned a random number of stubs according to a joint degree distribution \( \{ p_{st} \}_{s,t=0}^{\infty} \) that gives the probability that a node has \( s \) single edges and \( t \) triangles. Namely, each node will be given \( s \) stubs labeled as single and \( 2t \) stubs labeled as triangles with probability \( p_{st} \), for any \( s, t = 1, 2, \ldots \). Then, stubs that are labeled as single are randomly joined to form single edges that are not part of a triangle, whereas pairs of triangle stubs from three nodes are randomly matched to form triangles between the three participating nodes; obviously, the total degree of a node will be distributed by

\[
p_k = \sum_{s,t:s+2t=k} p_{st}.
\]

With this network model as the underlying contact network and the SIR model as the spreading model, Newman [68] and Miller [60] studied the impact of clustering on information propagation.

**Information propagation over multi-layer networks:** Instead of discussing the impact of clustering, Yağan et al. [102] studied the impact of the nature of multiple layers in underlying contact networks. They reported that the advances in communication technologies and cyber-physical systems bring even more complex topological properties in networks. For example, the smart grid where the power stations and the communication network controlling them are coupled together [21]. In social networks, this dependence is also evident. After receiving a message from a friend in face-to-face communication, a person could immediately send a message to his/her online friends via Facebook or Twitter. That is, the original information could spread over multiple networks (i.e., a face-to-face network and the Facebook network). Thus, because of the dependence between networks, they shifted the attention from single, isolated networks to multi-layer networks. By considering multi-layer network models to model underlying contact networks, Yağan et al. [102] studied the impact of multiple layers
on information propagation process, and reported an interesting finding that even if there is no outbreak in single-layer networks, an outbreak can take place in multi-layer network.

From the discussion of related work, existing literature either ignored both the natures of multiple layers and clustering or solely considered one of these two natures. Therefore, this dissertation studies the information propagation processes over clustered multi-layer networks, in order to consider both natures together. To emphasize the difference between our work and related work, we provide in Table 1.1 a summary of the literature of information propagation by distinguishing works according to the network model considered (single-layer vs. multi-layer networks, and tree-like vs. clustered networks). As can be seen from Table 1.1, this dissertation is the first work studying information propagation over multi-layer and clustered networks.

The contributions on information propagation include:

- Proposed a general framework to analyze the information propagation over non-clustered/clustered single-/multi-layer networks, and derived analytical solutions for the expected size of epidemics and the probability of the existence of epidemics.

- Showed the impact of clustering on the epidemic threshold, the final epidemic size, and the probability of the existence of epidemics in multi-layer networks. In particular, we found that increasing the level of clustering would increase the epidemic threshold and decrease the final epidemic size in the whole system.

- Reported that the nature of multiple layers could result in assortativity which exhibits a multi-faceted impact on information propagation.

<table>
<thead>
<tr>
<th>Work</th>
<th>Multi-layer</th>
<th>Clustered</th>
<th>Assortativity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newman [66]</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Newman [68]</td>
<td>X</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Miller [60]</td>
<td>X</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Yağan et al. [102]</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1.1: Table for related work on information propagation.
1.3 Related Work and Our Contributions on Influence Propagation

**Influence propagation over non-clustered monoplex networks:** As one of the pioneering work studying influence propagation, Watts [94] proposed a simple threshold model. The simple threshold model assumes that each individual has two possible states, *active* or *inactive*. Initially, the model assumes that all individuals are inactive, assigns a threshold \( \tau \) drawn from a distribution \( P(\tau) \) to each individual, and then chooses one of individuals uniformly at random and sets it as active. For each inactive individual with \( d \) neighbors and \( k \) of them are active, s/he turns active if \( \frac{k}{d} > \tau \), and then the newly activated individuals would start activating other inactive individuals. This process continues until no new individuals get activated. With this spreading model and Erdős Rényi graphs as underlying contact networks, Watts [94] derived analytical solutions to the probability of the existence of an outbreak and the expected size of an outbreak (if there exists an outbreak). One of the interesting observations is that there exist two phase transitions, while there is only one phase transition in information propagation.

**Influence propagation over non-clustered multiplex networks:** After Watts’ pioneering research, Yağan and Gligor [102] realized that in the real-world contact networks, link types might be categorized according to the nature of the relationship they represent (e.g., friends, family, office-mates, etc.). Moreover, different link types could exhibit different weights when spreading the given influence. For example, in the spread of a new consumer product amongst the population, a video game would be more likely to be promoted among high school classmates rather than among family members; the situation would be exactly the opposite in the case of a new cleaning product [90]. Then, considering the existence of multiple link types and their different weights on spreading influences, Yağan and Gligor [102] proposed the content-dependent threshold model.

**Influence propagation with multiple stages over non-clustered monoplex networks:** Instead of bringing topological features into mathematical modeling, Melnik et al. [58]
considered a dynamical property of spreading processes, multiple stages, in modeling influence propagation. This is motivated by the case where individuals could have a different reaction to a certain influence, and then they could exert different amounts of individuals on their neighbors. For example, followers of a radical organization or a revolutionary movement may have varying levels of commitment to the cause, or varying desire and ability to recruit new members. Therefore, they assume that nodes can be inactive or one of the different levels of active states (e.g., active, hyper-active, etc.), and reported that an influence outbreak could be either triggered by low-stage influencers or high-stage influencers.

In this dissertation, we addressed the limitations of current literature of influence propagation. We first reveal the impact of multiplexity and clustering on influence propagation by proposing an analytical framework by considering both properties together. Then, we consider a multi-stage influence propagation model over multiplex networks, in order to include the dynamical property of multiple active states and the topological of multiple link types. Besides, we find that all of existing work assumed that there is only one spreading influence. However, the real-world influence propagation could have multiple correlated influences spreading simultaneously. Thus, we proposed a new model for influence propagation, a vector threshold, which enables us to model the spread of correlated influence.

The contributions on influence propagation include

- Proposed a general framework to analyze the influence propagation over non-clustered/clustered single-/multi-layer networks and the influence propagation with multiple stages, and derived analytical solutions for the expected size of global cascades and the probability of the existence of global cascades.

- Showed the impact of clustering and multi-stages on the global cascades threshold, the expected cascade size, and the probability of the existence of global cascades in multiplex networks.

- Reported that the nature of multiplexity would lead to the existence of assortativity.

Then, depending on the level of assortativity, not only the expected outbreak size could
change, but also the number of phase transitions could change. In addition, we reported that in the multi-stage influence propagation, the level of assortativity could also change the impact of high-stage influencers.

- Proposed a new influence propagation model, and derived analytical solutions for the expected size of global cascades and the probability of the existence of global cascades.

To highlight the difference between our work and current literature of influence propagation, we list the difference in Table 1.2.

<table>
<thead>
<tr>
<th>Work</th>
<th>Multiple Influences</th>
<th>Multi-stage</th>
<th>Multiplex</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watts [94]</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Gleeson [40]</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>Gleeson [39]</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Yağan and Gligor [99]</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 1.2: Table for related work on influence propagation
Part II

Information Propagation
Chapter 2

Information Propagation over Clustered Multi-layer Networks

2.1 Introduction

The study of dynamical processes on real-world complex networks has been an active research area over the past decade. An interesting phenomenon that occurs in many such processes is the spreading of an initially localized effect throughout the whole (or, a very large part of the) network. These events are usually referred to as (information) cascades and can be observed in processes as diverse as adoption of cultural fads, the diffusion of belief, norms, and innovations in social networks [36, 94], disease contagion in human and animal populations [4, 63, 80, 102], failures in interdependent power systems [19, 21, 93, 101], rise of collective action to joining a riot [44], and the global spread of computer viruses or worms on the Web [8, 70].

This work focuses on an important class of dynamical process known as the information propagation or simple contagions; this is to be contrasted with complex contagions often referred to as influence propagation [99]. Although well-studied in the past across various domains, the information diffusion problem has recently taken a new form and dimension by the emergence of online social networks such as Facebook, Twitter, etc. In particular, due to the existence of multiple online social networks, information is now likely to spread among the population in an unprecedented speed and scale. Although there has been a recent surge of research on multi-layer and multiplex networks (e.g., see [54, 80, 102]), the current literature still falls short in fully quantifying this phenomenon. For instance, Yağan et al. analyzed [100, 102] the diffusion
of information in a multi-layer network, but only for the cases where all constituent layers are generated by the configuration model [71]; see also [56, 82] for works that are in the same vein. However, the configuration model produces [60, 68] networks that can not accurately capture some important aspects of real-world social networks, most notably the property of high clustering [83, 95]. Informally known as the phenomenon that “friends of our friends” are likely to be our friends, clustering has been shown to impact significantly the dynamics of various diffusion processes [48, 49, 99].

With these in mind, we study information propagation in clustered multi-layer networks. In particular, we consider a model where all constituent layers are random networks with clustering as introduced by Miller [60] and Newman [68], i.e., they are generated randomly from given distributions specifying the number of single edges and triangles for any given node; see Section 3.2.1 for details. Our modeling framework consists of a physical network where information spreads amongst people through conventional communication media (e.g., face-to-face communication, phone calls), and overlaying this network, there are online social networks offering alternative platforms for information diffusion, such as Facebook, Twitter, Google+, etc. The coupling across these networks results from nodes they have in common, i.e., individuals who participate in multiple networks simultaneously; see Section 2.2.2 for details of our multi-layer network model where the coupling level is tunable.

In this setting, we analyze the propagation of information assuming that information propagates according to the SIR epidemic model \(^1\). Namely, an individual is either susceptible (S) meaning that she has not yet received a particular information, or infectious (I) meaning that she is aware of the information and is spreading it to her contacts, or recovered (R) meaning that she is no longer spreading the information. Let \(T_{ij}\) denote the probability that an infectious individual \(i\) transmits the information to a susceptible contact \(j\). Throughout, we account for the fact that individuals’ information spreading behaviors may differ from one network to another; e.g., one may be more active in Facebook than Twitter, or vice versa. The varying

\(^1\)The analogy between the spread of diseases and information has long been recognized [45] and the SIR epidemic model is commonly used in similar studies; e.g., see [55, 89].
rate of information diffusion across different social networks is captured in our formulation by having the transmissibility $T_{ij}$ depend on the network that the link $i \sim j$ belongs to; see Section 2.2.3 for details.

Our main contributions are as follows. We solve analytically for the threshold, probability, and mean size of information epidemics, i.e., cases where information starts from a single individual and reaches a positive fraction of the population; see Section 5.2.3 for precise definitions. Our analytical approach is based on mapping the SIR propagation model to a bond percolation process and then utilizing a multi-type branching process to solve for the quantities of interest; the isomorphism between the SIR model and bond percolation has been established for certain cases in [53,59]. The analytical results are validated and extended by computer simulations.

Several interesting conclusions are drawn from these results. For example, we show that increasing the level of clustering in any one of the layers increases the epidemic threshold and decreases the final epidemic size of the whole system. Put differently, we show that i) clustering makes it more difficult for a single person to spread the information to the masses; and ii) even if the information reaches to the masses, we show that clustering decreases the total fraction of individuals informed. We also demonstrate how the overlap between the constituent networks affect the information propagation dynamics, particularly through impacting the degree-degree correlations. For instance, we show that an online social network that is small in size but large in mean connectivity is more effective (resp. less effective) in facilitating the propagation of information with low transmissibility (resp. high transmissibility) as compared to a large social network with smaller mean connectivity, with the total number of edges fixed in both cases.

Our general framework contains non-clustered multi-layer networks and single-layer clustered networks as special cases. In addition, given that information propagation problem is studied via bond percolation over a multi-layer network, our work can also be useful in the context of robustness against random attacks. Finally, although the problem is motivated here in the context of information propagation, network coupling is relevant in many simple contagion processes including diffusion of diseases [56,80]; e.g., a small community may consist
of three coupled networks corresponding to three venues people can interact at: households, hospitals, and schools [2,9].

The rest of the paper is organized as follows. In Section 2.2, we introduce the models applied in this study and the problem to be considered. In Section 2.3, we introduce the related technical background. In Section 2.4, we present and derive the main results of this work, while in Section 3.5, we confirm our analytical results via computer simulations. We conclude the paper in Section 5.5.

2.2 Problem Formulation

In this section, we give precise definitions of our system model and then describe problems that shall be studied.

2.2.1 Random Graphs with Clustering

Our modeling framework is based on random networks with clustering as introduced independently by Miller [60] and Newman [67]. This model takes its roots from the widely used configuration model [71] that generates a network randomly according to a given degree distribution. Namely, consider a vertex set \( V = 1, 2, \ldots, n \), where each vertex is independently assigned a random number of stubs according to a probability distribution \( \{p_k\}_{k=0}^{\infty} \); i.e., the degree \( d_i \) of vertex \( i \) equals \( k \) with probability \( p_k \) for any positive integer \( k \). Then, stubs are randomly paired with each other to form edges until no free stubs is left; see Figure 2.1 for an illustration of the configuration model.

![Illustration of process of Configuration Model.](image)

It is known that [67, 71] configuration model generates tree-like graphs with number of
cycles approaching to zero as the number of nodes gets large. However, most social networks exhibit high clustering, anecdotally known as the likelihood of a “friend of a friend” to be one’s friend. Put differently, real-world social networks are not tree-like and instead have considerable number of cycles, particularly of size three; i.e., triangles. With this in mind, Miller [60] and Newman [67] proposed a modification on the configuration model to enable generating random graphs with given degree distributions and tunable clustering.

The model proposed in [62, 67] is often referred to as random networks with clustering and is based on the following algorithm. Consider a joint degree distribution \( \{p_{st}\}_{s,t=0}^{\infty} \) that gives the probability that a node has \( s \) single edges and \( t \) triangles; e.g., see node 2 in Figure 2.2 that has two single edges and one triangle. Namely, each node will be given \( s \) stubs labeled as single and \( 2t \) stubs labeled as triangles with probability \( p_{st} \), for any \( s, t = 1, 2, \ldots \). Then, stubs that are labeled as single are randomly joined to form single edges that are not part of a triangle, whereas pairs of triangle stubs from three nodes are randomly matched to form triangles between the three participating nodes; of course the total degree of a node will be distributed by \( p_k = \sum_{s,t:s+2t=k} p_{st} \). As in the standard configuration model, it can be shown that the number of cycles formed by single edges goes to zero as \( n \) gets large, and so does the number of cycles of length larger than three [71].

The resulting level of clustering of the model described above can be quantified in a number of ways. Here we consider two widely used metrics known as the global clustering coefficient [67] and local clustering coefficient [10]; see [5, 40, 52] for other definitions of clustering coefficient proposed in the literature. Namely, the global clustering is defined via

\[
C_{\text{global}} = \frac{3 \times \text{(number of triangles in network)}}{\text{number of connected triples}},
\]  

(2.1)

where “connected triples” means a single vertex connected by edges to two others. On the
Figure 2.4: Illustration of a random network with clustering. In part (a), node 2 has 2 single edges and one triangle, whereas in part (b) it has zero single edges and two triangles. For the network in Figure 2.2, the global clustering coefficient is 0.2 while the local clustering coefficient is 0.3, while for the network in Figure 2.3, these coefficients are given by 0.4 and 0.7, respectively.

On the other hand, the local clustering is defined as the average

\[ C_{\text{local}} = \frac{1}{n^*} \sum_i C_i, \tag{2.2} \]

where \( C_i \) denotes the clustering coefficient for node \( i \) given by

\[ C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of connected triples centered on vertex } i}. \tag{2.3} \]

Here, \( n^* \) is the number of nodes whose \( C_i \) is well-defined in the network; i.e., number of nodes where the denominator at (2.3) is nonzero. The difference between the two definitions of clustering is illustrated in Figures 2.2 and 2.3, where networks with the same degree distribution are considered.

It was shown in [71] that both \( C_{\text{global}} \) and \( C_{\text{local}} \) are positive in the random clustered network.
Figure 2.5: Nodes in the upper circle and lower circle indicate the individuals in \textit{social network} and \textit{physical network} respectively. The nodes connected by a red line cross two networks mean they are the same individual existing in two networks. Green nodes in the upper circle belong to $F$, while blue nodes in $W$. Some of the nodes connected across the two networks by a red line indicates the fact that they represent the same individual.

model, while both quantities approach to zero with increasing network size in the standard configuration model.

2.2.2 Multilayer Network Models with Clustering

In this paper, we consider a multilayer network where each layer is generated independently and constitutes a random graph with clustering as introduced in Section 3.2.1. For brevity, we only consider two layers but most of the arguments can easily be extended higher number of layers. Namely, we let $W$ and $F$ denote the two constituent layers of networks with the possible motivation that $W$ models the \textit{physical} contact network among individuals, i.e., models face-to-face relationships, while network $F$ stands for an online social network, say Facebook. In line with this terminology, we assume that the network $W$ is defined on the vertices $\mathcal{N} = \{1, \ldots, n\}$, while $F$ contains only a \textit{subset} of the nodes in $\mathcal{N}$ to account for the fact that not every individual participates in online social networks; see Figure 2.5 for an illustration of the two-layer network model we are considering.

To specify this model further, we assume that each vertex in $\mathcal{N}$ participates in $F$ indepen-
dently with probability $\alpha \in (0, 1]$, leading by the Strong Law of Large Numbers to

$$\frac{|\mathcal{N}_F|}{n} \rightarrow_{a.s.} \alpha$$

(2.4)

where $\mathcal{N}_F$ denotes the set of vertices in network $\mathbb{F}$; here $\rightarrow_{a.s.}$ denotes convergence in almost sure sense with $n$ growing unboundedly large. In words, this implies that the fraction of nodes that belong to $\mathbb{F}$ is $\alpha$ in the large $n$ limit. The case where $|\mathcal{N}_F| = o(n)$ has been considered in [102] and it was shown that most properties pertaining to the propagation information are unaffected by the existence of the upper layer $\mathbb{F}$; i.e., when the online social network has a negligible size compared to the whole population, it does not impact the threshold or size of information epidemics.

As mentioned already, we assume that both $\mathbb{F}$ and $\mathbb{W}$ are random networks with clustering. In particular, we let $\{p_{st}^f, s, t = 0, 1, \ldots\}$ and $\{p_{st}^w, s, t = 0, 1, \ldots\}$ denote the joint distributions for single edges and triangles for $\mathbb{F}$ and $\mathbb{W}$, respectively. Then both networks are generated independently according to the algorithm described in Section 3.2.1, and they are denoted respectively by $\mathbb{F} = \mathbb{F}(n; \alpha, p_{st}^f)$ and $\mathbb{W} = \mathbb{W}(n; p_{st}^w)$. We define the multi-layer network $\mathbb{H}$ as the disjoint union $\mathbb{H} = \mathbb{F} \coprod \mathbb{W}$ and represent it by $\mathbb{H}(n; \alpha, p_{st}^f, p_{st}^w)$. Here, the disjoint union operation implies that we still distinguish $\mathbb{F}$-edges from $\mathbb{W}$-edges in network $\mathbb{H}$, and this is done to accommodate the possibly different rates (or, even rules) of information propagation across the two networks. To this end, an equivalent representation of $\mathbb{H}$ would be a multiplex network with different types (or, colors) of edges.

With these definitions in mind, let $d_{fs}$ and $d_{ws}$ to denote the random variables corresponding to the number of single edges for a vertex in $\mathbb{F}$ and $\mathbb{W}$, respectively, while $n_{ft}$ and $n_{wt}$ are defined similarly for the number of triangles assigned; i.e., the degree of a node from triangle edges in $\mathbb{F}$ is given by $d_{ft} = 2n_{ft}$ and similarly for $d_{wt}$. Then the colored degree $d$ of a vertex is given by

$$d = (d_{fs}, 2n_{ft}, d_{ws}, 2n_{wt})$$

(2.5)
meaning that the vertex has \(d_{fs}\) single edges and \(2n_{ft}\) triangle edges in network \(\mathbb{F}\), and \(d_{ws}\) single edges and \(2n_{wt}\) triangle edges in network \(\mathbb{W}\). Under the assumptions enforced here, the distribution of this colored degree is given by

\[
p_d = \left( \alpha p_f^{d_{fs},n_{ft}} + (1 - \alpha)1[d_{fs} = 0 \land n_{ft} = 0] \right) p_w^{d_{ws},n_{wt}}
\]

(2.6)

where the term \((1 - \alpha)1[d_{fs} = 0 \land n_{ft} = 0]\) accounts for the fact that if the node does not belong to \(\mathbb{F}\) (which happens with probability \(1 - \alpha\)), then its degree from single and triangle edges will both be zero.

### 2.2.3 Information Propagation Model: SIR

Consider the diffusion of a piece of information in the multi-layer network \(\mathbb{H}\) which starts from a single node. We assume that information spreads from a node to its neighbors according to the SIR epidemic model. In this context, an individual is either susceptible (S) meaning that she has not yet received a particular item of information, or infectious (I) meaning that she is aware of the information and is capable of spreading it to her contacts, or recovered (R) meaning that she is no longer spreading the information [45,55,89]. As in [66], we assume that an infectious individual \(i\) transmits the information to a susceptible contact \(j\) with probability \(T_{ij} = 1 - e^{-r_{ij} \tau_i}\). Here, \(r_{ij}\) denotes the rate of contact over the link from \(i\) to \(j\), and \(\tau_i\) is the time \(i\) keeps spreading the information; i.e., time \(i\) remains infectious.

It is expected that information propagates over the physical and social networks at different rates, which manifests from different probabilities \(T_{ij}\) across links in this case. Specifically, let \(T_{ij}^w\) stand for the probability of information transmission over a link (between and \(i\) and \(j\)) in \(\mathbb{W}\) and let \(T_{ij}^f\) denote the probability of information transmission over a link in \(\mathbb{F}\). For simplicity, we assume that \(T_{ij}^w\) and \(T_{ij}^f\) are independent for all distinct pairs \(i, j = 1, \ldots, n\). Furthermore, we assume that the random variables \(r_{ij}^w\) and \(\tau_i^w\) are independent and identically distributed (i.i.d.) with probability densities \(P_w(r)\) and \(P_w(\tau)\), respectively. We find it useful to define \(T_w\)
as the mean of $T_{ij}^w$, i.e.,

$$T_w := \langle T_{ij}^w \rangle = 1 - \int_0^\infty \int_0^\infty e^{-r\tau} P_w(r) P_w(\tau) dr d\tau.$$

We refer to $T_w$ as the transmissibility over $W$ and note that $0 \leq T_w \leq 1$. In the same manner, we assume that $r_{ij}^f$ and $\tau_i^f$ are i.i.d. with respective densities $P_f(r)$ and $P_f(\tau)$ leading to a transmissibility $T_f$ over $F$.

As shall be discussed in Section 5.2.3, under certain conditions, it can be assumed that information propagates over $W$ (resp. over $F$) as if all transmission probabilities were equal to $T_w$ (resp. to $T_f$), for the purposes of computing the threshold, probability, and expected size of epidemics.

### 2.2.4 Problems of Interest

We consider the propagation of information (or, a disease) in $H$ as explained in Section 2.2.3. The outbreak is triggered by infecting a randomly selected node and propagates in the network according to the SIR model. Given the monotonicity of the SIR process [66], a steady-state will always be reached where all nodes are either recovered or susceptible. The final size of an outbreak is defined as the number of nodes that are recovered at the steady-state, and its relative final size is its final size divided by the total size $n$ of the network. Following [53], we define a self-limited outbreak as an outbreak whose relative final size approaches zero, and an epidemic to be an outbreak whose relative final size is positive, both in the limit of large $n$. There is a critical boundary in the space of all network parameters, often defined as the epidemic threshold, or epidemic boundary, that separates the cases for which the probability of an epidemic is zero (i.e., sub-critical, or non-epidemic parameter regime) from those that lead to $\mathbb{P}[\text{epidemic}] > 0$ (i.e., super-critical, or epidemic regime), again with $n \to \infty$.

With these definitions in place, this work seeks to identify i) the epidemic boundary; ii) the relative final size of epidemics in the super-critical case; and iii) the exact probability
As we seek to study several properties of simple contagions as outlined above, a first step will be to observe that under certain conditions, the SIR propagation model is *isomorphic to a bond percolation* process [18]. More specifically, assume that each edge in $\mathbb{W}$ (resp. $\mathbb{F}$) is *occupied* – meaning that it can be used in spreading the information, disease, etc. – with probability $T_w$ (resp. $T_f$) independently from all other edges. Here, $T_w$ and $T_f$ are transmissibility parameters calculated as the mean probability of transmission between any two nodes in the corresponding networks; see Section 2.2.3. Then, the size of an outbreak started from an arbitrary node is equal to the number of individuals that can be reached from the initial node by using only the *occupied* links in $\mathbb{H}$.

This isomorphism was claimed to hold first by Newman [66] who studied the SIR model in single networks. It was later shown by several authors [53, 59] that the SIR process is isomorphic to bond percolation *only* when the infectious period distribution $P(\tau)$ is *degenerate*; i.e., when all nodes have the same recovery time $\tau_1 = \cdots = \tau_n$. When nodes have heterogeneous recovery times, the SIR process is *not* isomorphic to a bond percolation process. However, [53, 59] proved that, in the large network size limit, a bond percolation process can still be used to accurately predict the a) epidemic boundary, b) mean size of self-limited out-breaks, and c) relative final size of epidemics. With respect to our goals, it is only the probability $\mathbb{P}[\text{epidemic}]$ that can’t be obtained through analyzing the bond percolation model when the recovery times are heterogeneous; in fact, we are not aware of any technique in the literature that enables calculating $\mathbb{P}[\text{epidemic}]$ exactly in these cases. Therefore, we restrict our attention to cases where the recovery times are uniform when dealing with $\mathbb{P}[\text{epidemic}]$, while more general cases are considered for the boundary and final size of epidemics. To that end, our efforts towards analyzing information propagation (e.g., items (i)-(iii) given above) rely on mapping the SIR model to a bond percolation process.

We now explain how mapping the problem to a bond percolation process paves the way to obtaining the quantities (i)-(iii) given above. Let $\tilde{\mathbb{W}}$ (resp. $\tilde{\mathbb{F}}$) be a network that contains
only the occupied edges of \( W \) (resp. \( F \)). Put differently, consider an Erdős-Rényi [15] network \( G(n; T_w) \) (resp. \( G(N_F; T_f) \)) on the nodes \( \{1, \ldots, n\} \) (resp. on the node set \( N_F \)) such that between every pair of nodes there is an edge with probability \( T_w \) (resp. \( T_f \)) independently from all other edges. Then, \( \tilde{W} = W \cap (G(n; T_w)) \) and \( \tilde{F} = F \cap (G(N_F; T_f)) \). The bond percolation network \( \tilde{H} \) that contains only the occupied edges of \( H \) is then given by \( \tilde{H} = \tilde{W} \cup \tilde{F} \). The different transmissibility properties of \( W \) and \( F \) are already incorporated into this model through distinct bond occupation probabilities \( T_w \) and \( T_f \). Thus, \( \tilde{H} \) (defined on the vertices \( \{1, \ldots, n\} \)) is a simplex network obtained by a simple union of the edges of \( \tilde{W} \) and \( \tilde{F} \).

The threshold and relative final size of epidemics can now be computed from the phase transition behavior of \( \tilde{H} \). Namely, epidemics can take place if and only if \( \tilde{H} \) has a giant component; i.e., a connected subgraph that contains a positive fraction of nodes in the large \( n \) limit. Thus, epidemic boundary is given by the phase transition threshold, i.e., the threshold for the existence of a giant component in \( \tilde{H} \). Also, a node can trigger an epidemic only if it belongs to the giant component, in which case an outbreak started from this node will reach the whole giant component. Hence, the relative size of the giant component in \( \tilde{H} \) gives both \( P[\text{epidemic}] \) as well as the relative final size of epidemics.

### 2.3 Technical Background

In what follows we introduce the technical underpinnings of our analysis. Our approach is based on exploring a branching process which starts with an arbitrary node in the network and recursively reveals all the nodes reached and informed by following its edges; see Figure 2.6. Throughout, we will be interested in various discrete random variables naturally associated with this branching process; e.g., total number of nodes reached and informed by following a randomly selected edge in \( W \) (resp. in \( F \)). Oftentimes we find it useful to characterize the probability distributions of these random variables through their generating functions [97]. This approach has been widely adopted in the literature in analyzing complex networks and has several benefits as shall soon become apparent.
We now formally define the notion of a generating function: Let $X$ be a positive-valued, discrete random variable with the distribution $\{p_k : k = 0, 1, \ldots\}$; i.e., we have $\mathbb{P}(X = k) = p_k$. Then the generating function of $X$ is given by

$$h(x) = \sum_{k=0}^{\infty} p_k x^k, \quad x \in \mathbb{R}. \quad (2.7)$$

We remark that a random variable is uniquely identified by its generating function since we have

$$p_k = h^{(k)}(x)/k!, \quad k = 0, 1, \ldots$$

where $h^{(k)}(x)$ denotes the $k$th order derivative of $h(x)$. Also, we can easily compute the moments of $X$ from the derivatives of $h(x)$ evaluated at the point $x = 1$ [71]; e.g., the first moment is given through $\mathbb{E}[X] = h'(1)$, i.e., by the first derivative of $h(x)$ evaluated at $x = 1$.

![Figure 2.6: Illustration of the branching process.](image)

The children of each individual node is identified recursively, while taking into account whether or not the information is transferred from the parent node to the child node. The initial vertex that starts the information is regarded as the $0$th generation, and we are interested in deriving the limiting behavior of the total number of nodes reached and informed as the number nodes $n \to \infty$.

### 2.4 Main Results

As described in Section 2.2.2, the clustered multilayer network in this paper consists of four kinds of edges, single edges in $\mathcal{F}$, triangle edges in $\mathcal{F}$, single edges in $\mathcal{W}$, and triangle edges in $\mathcal{W}$; these will be denoted by $fs$-, $ft$-, $ws$-, and $wt$-edges, respectively. In order to analyze the information propagation in multilayer networks, we consider a branching process that starts with informing a node selected randomly from among all nodes, $\{1, \ldots, n\}$. We then explore
all the neighbors that are reached and informed by this node, and continue recursively until
the branching process stops. The distribution of the resulting number of nodes informed will
be characterized via its generating function.

We now explain our approach based on generating functions precisely. Let $H(x)$ denote
the generating function for the “finite number of nodes that are reached and informed” by the
above branching process. We will derive an expression for $H(x)$ using four other generating
functions $h_{fs}(x)$, $h_{ft}(x)$, $h_{ws}(x)$, and $h_{wt}(x)$, where $h_{fs}(x)$ stands for the “finite number
of nodes reached and informed by following a randomly selected $fs$-edge,” and $h_{ws}(x)$ defined
similarly for the $ws$–edges. The definitions for $h_{ft}(x)$ and $h_{wt}(x)$ are a bit different in the
sense that they correspond to the “finite number of nodes reached and informed by following
a randomly selected triangle in $F$ (resp. in $W$)” for $h_{ft}(x)$ (resp. $h_{wt}(x)$). In other words, we
consider the whole triangle at once, rather than focusing on its edges separately; see see Section
2.4.2.

With these definitions in place, we now write $H(x)$ in terms of $h_{fs}(x)$, $h_{ft}(x)$, $h_{ws}(x)$, and
$h_{wt}(x)$:

$$H(x) = x \sum_{d} p_{d} h_{fs}(x)^{d_{fs}} h_{ft}(x)^{n_{ft}} h_{ws}(x)^{d_{ws}} h_{wt}(x)^{n_{wt}}, \quad (2.8)$$

where $p_{d}$ denotes the colored degree distribution given by (2.6). The validity of (2.8) can
be seen as follows. The term $x$ stands for the node that is selected randomly and given the
information to initiate the propagation. This node has a degree $d = (d_{fs}, 2n_{ft}, d_{ws}, 2n_{wt})$ with
probability $p_{d}$. The number of nodes reached and informed by each of its $d_{fs}$ (resp. $d_{ws}$) single
edges in $F$ (resp. $W$) has a generating function $h_{fs}(x)$ (resp. $h_{ws}(x)$). Similarly, the number of
nodes informed by following each of the $n_{ft}$ (resp. $n_{wt}$) triangles it participates in $F$ (resp. $W$)
has a generating function $h_{ft}(x)$ (resp. $h_{wt}(x)$). Combining, we see from the powers property
of generating functions [71] that the number of nodes reached and informed in this process when
the initial node has degree $d$ has a generating function $h_{fs}(x)^{d_{fs}} h_{ft}(x)^{n_{ft}} h_{ws}(x)^{d_{ws}} h_{wt}(x)^{n_{wt}}$.
Averaging over all possible degrees $d$ of the initial node, we get (2.8).
For (2.8) to be useful, we shall derive expressions for the generating functions \( h_{fs}(x) \), \( h_{ft}(x) \), \( h_{ws}(x) \), and \( h_{wt}(x) \). As will become apparent soon, there are no explicit equations defining these functions. Instead, we should seek for recursive equations defining each generating function in terms of others. Then, fixed points of this recursion will be explored and utilized to determine the threshold and size of information epidemics; i.e., situations where the number of people reached and informed by the original branching process is infinite. These steps are taken in the next sections where we first focus on deriving \( h_{fs}(x) \) and \( h_{ws}(x) \) (Section 2.4.1) followed by derivations of \( h_{ft}(x) \) and \( h_{wt}(x) \) (Section 2.4.2). These arguments are then combined in Section 2.4.3 to derive the epidemic threshold and final epidemic size.

### 2.4.1 Information Propagation via Single Edges in Network \( \mathbb{F} \)

We start by deriving recursive equations for \( h_{fs}(x) \) and \( h_{ws}(x) \), by focusing on the number of nodes reached and informed by following one end of a single edge in \( \mathbb{F} \) and \( \mathbb{W} \), respectively. For instance, for \( h_{fs}(x) \), we pick one of the single edges in \( \mathbb{F} \) uniformly at random and assume that it is connected at one end a node who is in the infected state. Then, we compute the generating function for the number of nodes informed by following the other end of the edge. In what follows, we only derive \( h_{fs}(x) \) since the computation of \( h_{ws}(x) \) follows in a very similar manner.

Similar to [102], we obtain the following expression for the generating function \( h_{fs}(x) \):

\[
\begin{align*}
\text{h}_{fs}(x) & = T_f x \sum_{d} \frac{d_{fs} d_{fa}}{\langle d_{fs} \rangle} h_{fs}(x)^{d_{fs}-1} h_{ft}(x)^{n_{ft}} h_{ws}(x)^{n_{ws}} h_{wt}(x)^{n_{wt}} \\
& \quad + (1 - T_f).
\end{align*}
\]

We now explain each term appearing at (2.9) in turn. First of all, it is straightforward to see that if the selected edge is not occupied, which happens with probability \( 1 - T_f \), then the number of informed nodes by following it will be zero. This leads to a term \( (1 - T_f)x^0 \) in
Figure 2.7: The top vertex $u$ is infected, and information is transferred through an edge only if it is occupied (happens with probability $T_f$ in network $\mathbb{F}$).

the generating function $h_{fs}(x)$. In words, adding the term $(1 - T_f)x^0$ to $h_{fs}(x)$ means that the probability of the underlying random variable (encoded by the generating function $h_{fs}(x)$) being zero is incremented by $1 - T_f$. On the other hand, if the selected edge is occupied, which happens with probability $T_f$, then the node at the other end of the edge will be informed. This means that the number of informed edges in this process will be one plus all the nodes that are then informed by the node at the other end of the selected edge. Adding one to a random variable is equivalent to multiplying its generating function by $x$, whence we get the term $T_fx$.

The summation term appearing at (2.9) stands for the number of nodes informed by the aforementioned end node of the randomly selected edge, and is similar in vein with the summation term used in (2.8) with two differences. First, the degree distribution of this end node is not $p_d$ since it is already known to have at least one single edge in $\mathbb{F}$. Instead, its degree distribution will be proportional to $d_{fs}p_d$, and after proper normalization we see that the end node will have degree $d = (d_{fs}, 2n_{ft}, d_{ws}, 2n_{wt})$ with probability $\frac{d_{fs}p_d}{\langle d_{fs} \rangle}$; e.g., see [71, 102] for similar arguments. Finally, if this node has degree $d$ then the number of people it informs is generated by $h_{fs}(x)^{d_{fs}-1}h_{ft}(x)^{n_{ft}}h_{ws}(x)^{d_{ws}}h_{wt}(x)^{d_{wt}}$, with the minus one term on $d_{fs}$ accounting to the fact that one of its single edges in $\mathbb{F}$ has carried the information to this node and has already been taken into account. Averaging over all possible $d$, we get (2.9).

2.4.2 Information Propagation via Triangles in Network $\mathbb{F}$

We now derive $h_{ft}(x)$, i.e., the generating function for the number of nodes informed by following a random triangle in $\mathbb{F}$; similar arguments hold for $h_{wt}(x)$. We demonstrate this situation
in Figure 2.7, where the top vertex \( u \) is \textit{infected}, and we are interested in computing the generating function for the number of nodes that will be informed \textit{by} nodes \( v \) and \( w \). Firstly, by conditioning on the state, i.e., occupied or not occupied, of the three edges forming this triangle, we compute the probabilities for neither, one, or both of \( v \) and \( w \) being informed, respectively. It is not difficult to see that

\[
\begin{align*}
\mathbb{P}[\text{none of } v \text{ and } w \text{ are informed}] &= (1 - T_f)^2 \\
\mathbb{P}[\text{one of } v \text{ and } w \text{ are informed}] &= 2T_f(1 - T_f)^2 \\
\mathbb{P}[\text{both of } v \text{ and } w \text{ are informed}] &= 2T_f^2(1 - T_f) + T_f^2.
\end{align*}
\]

We now explain why the above equations hold. Firstly, for \( v \) and \( w \) to be not informed, both of the edges \( u \sim v \) and \( u \sim w \) should be \textit{not} occupied. By independence, this occurs with probability \((1 - T_f)^2\). Secondly, we compute the probability of only one of \( v \) and \( w \) being informed, which by symmetry is given by two times the probability that \( v \) is informed but \( w \) is not. The latter happens if and only if the edge \( u \sim v \) is occupied while the edges \( u \sim w \) and \( v \sim w \) are not occupied. By independence, this has probability \( T_f(1 - T_f)^2 \). Finally, probability that both \( v \) and \( w \) are informed is given by subtracting the first two probabilities from one.

We now turn to computing the generating function \( h_{f_t}(x) \) by conditioning on the three events discussed above. As in Section 2.4.1, if neither of the nodes \( v \) and \( w \) are informed, then the number of nodes informed by this triangle will be zero, leading to an additive term \((1 - T_f)^2x^0\). Next, we derive the term corresponding to the case where only one of \( v \) or \( w \) is informed. This leads to

\[
(2T_f (1 - T_f)) x
\cdot \sum_d \frac{n_{f_t}^{P_d}}{\langle n_{f_t} \rangle} h_{f_s}(x)^{d_{f_s}} h_{f_t}(x)^{n_{f_t} - 1} h_{w_s}(x)^{d_{w_s}} h_{w_t}(x)^{n_{w_t}},
\]
where \(2T_f (1 - T_f)\) stands for the probability of the conditioning event that only one of \(v\) or \(w\) is informed, and \(x\) stands for the node that is informed. As in Section 2.4.1, the degree distribution of this informed node will not be given by \(p_d\), but instead will be proportional to the number of triangles \(n_{ft}\) assigned to it; as before this is due to the fact that the node under consideration is known to have at least one triangle in \(F\). By normalization, we see that the degree of the node will be \(d = (d_{fs}, 2n_{ft}, d_{ws}, 2n_{wt})\) with probability \(\frac{n_{ft}p_d}{\langle n_{ft} \rangle}\). The rest of the expression (2.10) follows similarly to (2.9), where a minus one term is invoked at \(n_{ft}\) in order to not double count the triangle \(u, v, w\) that is being considered. Finally, the term corresponding to the case where both \(v\) and \(w\) are informed is easily computed as the square of (2.10) as we use the powers property upon noting that \(v\) and \(w\) will inform independent sets of nodes under the enforced assumptions. Collecting, we obtain

\[
\begin{align*}
    h_{ft}(x) & = (1 - T_f)^2 + (2T_f (1 - T_f))^2 x \sum_d \left( \frac{n_{ft}p_d}{\langle n_{ft} \rangle} h_{fs}(x)^{d_{fs}} \right) \\
    & \cdot h_{ft}(x)^{n_{ft}-1} h_{ws}(x)^{d_{ws}} h_{wt}(x)^{n_{wt}} + (2T_f^2 (1 - T_f) + T_f^2) \cdot \left( x \sum_d \frac{n_{ft}p_d}{\langle n_{ft} \rangle} h_{fs}(x)^{d_{fs}} h_{ft}(x)^{n_{ft}-1} h_{ws}(x)^{d_{ws}} h_{wt}(x)^{n_{wt}} \right)^2 
\end{align*}
\]

### 2.4.3 Computing the Final Epidemic Size

We are now in a position to write the recursive equations for generating functions \(h_{fs}(x)\), \(h_{ft}(x)\), \(h_{ws}(x)\), and \(h_{wt}(x)\), whose solution will be reported into (2.8) to get the final epidemic size. Using (2.9) and (2.11) and similar expressions for \(h_{ws}(x)\) and \(h_{wt}(x)\), we obtain

\[
\begin{align*}
    h_{ft}(x) & = (1 - T_f)^2 + (2T_f (1 - T_f))^2 x \sum_d \left( \frac{n_{ft}p_d}{\langle n_{ft} \rangle} h_{fs}(x)^{d_{fs}} \right) \\
    & \cdot h_{ft}(x)^{n_{ft}-1} h_{ws}(x)^{d_{ws}} h_{wt}(x)^{n_{wt}} + (2T_f^2 (1 - T_f) + T_f^2) \cdot \left( x \sum_d \frac{n_{ft}p_d}{\langle n_{ft} \rangle} h_{fs}(x)^{d_{fs}} h_{ft}(x)^{n_{ft}-1} h_{ws}(x)^{d_{ws}} h_{wt}(x)^{n_{wt}} \right)^2 
\end{align*}
\]
The desired generating function $H(x)$ for the finite number of nodes informed in the network can now be computed in the following manner. For any $x$, we solve the recursive relations (2.12) - (2.15), i.e., find a fixed point of (2.12) - (2.15). Then reporting the resulting values of $h_{fs}(x)$, $h_{ft}(x)$, $h_{ws}(x)$, and $h_{wt}(x)$ into (2.8), we obtain $H(x)$ for this particular value of $x$. Repeating the same process for any $x$ will lead to a complete characterization of $H(x)$. However, in this work we are only interested in the cases where the number of nodes informed by the process is infinite. More precisely, we wish to derive i) the conditions for the probability of informing a positive fraction of nodes to be larger than zero in the large $n$ limit; and (ii) the exact asymptotic fraction of informed individuals when the conditions of part (i) hold. As explained in Section 5.2.3, the latter also gives the probability of triggering an epidemic starting with a random node.

In order to achieve these goals, we take advantage of the “conservation of probability” property of generating functions, i.e., the fact that $H(1) = 1$ when the number of nodes
reached and informed is always finite. If on the other hand \( H(1) < 1 \), we understand that there is a positive probability \( 1 - H(1) \) for the aforementioned branching process to lead to an infinite component of informed nodes; i.e., for the branching process to be supercritical. In this case, \( 1 - H(1) \) stands for the fraction of nodes that are in the giant component of \( \hat{\mathbb{H}} \). Recalling the discussion in Section 5.2.3, we know that information propagation will turn into an epidemic if and only if the initiator node is in this giant component. Thus, we conclude that the probability of an epidemic is given by \( 1 - H(1) \) and so is the relative final size of epidemics.

With these in mind, we now seek for a fixed point of the recursion (2.12) - (2.15) at the point \( x = 1 \). For notational convenience, we define \( h_1 := h_{fs}(1) \), \( h_2 := h_{ft}(1) \), \( h_3 := h_{ws}(1) \), and \( h_4 := h_{wt}(1) \). The recursion (2.12) - (2.15) then takes the form

\[
h_i = g_i(h_1, h_2, h_3, h_4), \quad i = 1, 2, 3, 4
\]  

(2.16)

where \( g_1, g_2, g_3, \) and \( g_4 \) are functions immediately obtainable from (2.12) - (2.15); e.g., we have

\[
g_1(h_1, h_2, h_3, h_4) = T_f \sum_d \left\{ \frac{d_{fs} p_d}{\langle d_{fs} \rangle} h_1^{d_{fs} - 1} h_2^{n_{ft}} h_3^{d_{ws}} h_4^{n_{wt}} + (1 - T_f) \right\}.
\]

With this notation, we also have

\[
H(1) = \sum_d p_d h_1^{d_{fs}} h_2^{n_{ft}} h_3^{d_{ws}} h_4^{n_{wt}}.
\]  

(2.17)

It is easy to check that the recursion (2.12) - (2.15) exhibits a trivial fixed point \( h_1 = h_2 = h_3 = h_4 = 1 \), which leads to \( H(1) = 1 \), meaning that the branching process is sub-critical and all informed components have finite size. However, the solution \( h_1 = h_2 = h_3 = h_4 = 1 \) is stable only when it is an attractor; i.e., a stable fixed point. We check the stability of this solution via linearization of (2.12) - (2.15) around \( x = 1 \), which leads to Jacobian matrix \( J \)
whose entries are given by

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

for each \( i, j = 1, 2, 3, 4 \). Namely, we have

\[ J = \begin{bmatrix}
T_f \langle \frac{d_f s - d_f t}{d_f s} \rangle & T_f \langle \frac{d_f n_f t}{d_f s} \rangle & T_f \langle \frac{d_f d_w s}{d_f s} \rangle & T_f \langle \frac{d_f n_w s}{d_f s} \rangle \\
2T_f(1 + T_f - T_f^2) \langle \frac{n_f s - n_f t}{n_f t} \rangle & 2T_f(1 + T_f - T_f^2) \langle \frac{n_f t - n_f w}{n_f t} \rangle & 2T_f(1 + T_f - T_f^2) \langle \frac{n_f d_w s}{n_f t} \rangle & 2T_f(1 + T_f - T_f^2) \langle \frac{n_f n_w s}{n_f t} \rangle \\
T_w \langle \frac{d_w s - d_w t}{d_w s} \rangle & T_w \langle \frac{d_w n_f s}{d_w s} \rangle & T_w \langle \frac{d_w d_w s}{d_w s} \rangle & T_w \langle \frac{d_w n_w s}{d_w s} \rangle \\
2T_w(1 + T_w - T_w^2) \langle \frac{n_w d_f s}{n_w s} \rangle & 2T_w(1 + T_w - T_w^2) \langle \frac{n_w d_f s}{n_w s} \rangle & 2T_w(1 + T_w - T_w^2) \langle \frac{n_w d_w s}{n_w s} \rangle & 2T_w(1 + T_w - T_w^2) \langle \frac{n_w n_w s}{n_w s} \rangle
\end{bmatrix}. \] (2.18)

Now, if the largest eigenvalue in absolute value of the Jacobian matrix \( J \), denoted by \( \sigma(J) \), is less than or equal to one, then the trivial solution mentioned above is an attractor, whence all informed components have finite size as understood from the conservation of probability; i.e., from \( H(1) = 1 \). However, if \( \sigma(J) > 1 \), then the trivial solution will not be stable and another solution with \( h_1, h_2, h_3, h_4 < 1 \) will exist. This then will lead to having \( H(1) < 1 \) meaning that information epidemics take place with probability \( 1 - H(1) > 0 \) and reach an expected fraction \( 1 - H(1) \) of the whole population, where \( H(1) \) is computed from (2.17).

Collecting, the threshold of information epidemics is given by \( \sigma(J) = 1 \), where \( \sigma(J) \) is the spectral radius of the Jacobian matrix given at (2.19). Also, the mean epidemic size (i.e., the fractional size of the giant component of the percolated network \( \bar{H} \)) can be computed by first finding the pointwise smallest solution of the recursion (2.16), and then reporting the result into (2.17) to get \( H(1) \). As discussed before, the mean size of epidemics is given by \( 1 - H(1) \).
2.4.4 The Relationship between Our Analysis and Some Previous Studies

Our results generalize some of the existing work in the literature; e.g., see [60,68,102]. First, by letting \( h_{fs}(x) = h_{ft}(x) = 1 \) in (2.12)-(2.15), we ensure that \( F \) is an empty graph, so that our system model is equivalent to the single clustered network considered in [60,68]. Similarly, if we set \( h_{ft}(x) = h_{wt}(x) = 1 \) then neither \( F \) nor \( W \) will have triangle edges, rendering our system to be equivalent to the non-clustered multi-layer network studied in [102]. A careful inspection of our results will reveal that in both special cases, our results recover the finding of [60,68,102].

2.5 Numerical Results and Discussion

This section is devoted to presenting numerical results with regard to information propagation in clustered multi-layer networks in specific settings with given degree distributions. In what follows, we first consider a simple case where both constituent networks in our model has doubly-Poisson degree distributions \( p_{st} \), while in Section 2.5.2 we consider the more realistic case where \( p_{st} \) is a power-law degree distribution with exponential cut-off. Section 2.5.3 and Section 2.5.4 are devoted to understanding the impact of clustering and of the parameter \( \alpha \) on the dynamics of information propagation, respectively.

2.5.1 Networks with Doubly Poisson Distributions

Consider the case where both \( p_{st}^{f} \) and \( p_{st}^{w} \) are doubly Poisson; i.e., the number of single edges and triangles in both networks are independent and they all follow a Poisson distribution. Namely, we set

\[
p_{st}^{f} = e^{-\mu_{f,1}} \frac{(\mu_{f,1})^s}{s!} e^{-\mu_{f,2}} \frac{(\mu_{f,2})^t}{t!}, \quad s, t = 1, 2, \ldots,
\]  

(2.20)
and

\[ p_{st}^w = e^{-\mu_{w,1}} \frac{(\mu_{w,1})^s}{s!} e^{-\mu_{w,2}} \frac{(\mu_{w,2})^t}{t!}, \quad s, t = 1, 2, \ldots, \]  

(2.21)

where \( s \) and \( t \) are the number of single edges and triangles in the corresponding networks while \( \mu_{f,1} \) and \( \mu_{f,2} \) (resp. \( \mu_{w,1} \) and \( \mu_{w,2} \)) are the mean number of them respectively in \( \mathbb{F} \) (resp. in \( \mathbb{W} \)).

Under this setting, the mean epidemic size as well as the epidemic threshold can be computed from the analytical results presented in Section 2.4.3. To check the validity of our analysis for finite-sized networks, we have also conducted an extensive numerical study. In particular, we consider \( n = 5 \times 10^5 \) nodes in the population and three different values \( \alpha = 0.1, 0.5, 0.9 \) for the size of network \( \mathbb{F} \). We let \( \mu_{f,1} = \mu_{f,2} = \lambda_{fs} = \lambda_{ft} = 0.5 \) and similarly \( \mu_{w,1} = \mu_{w,2} = \lambda_{ws} = \lambda_{wt} = 0.5 \). For various information transmissibility parameters \( T_w = T_f \) we generate 100 independent realizations of the multi-layer network \( \mathbb{H} \) and compute the size of the largest connected component (of the percolated network \( \tilde{\mathbb{H}} \)) in each case. The results are then averaged over 100 experiments to obtain the empirical size of information epidemics.

The results are depicted in Figure 2.8, where the curves stand for the theoretical results obtained from our discussion in Section 2.4.3, while the markers stand for the empirical results obtained from simulation experiments. We see that there is a perfect agreement between the analytical and experimental results confirming the validity of our results even when \( n \) is finite. We also see that as \( \alpha \) increases, the critical threshold is reduced and the epidemics size is enlarged. This is an intuitive consequence given that the network becomes denser with

![Figure 2.8: Simulation for doubly Poisson degree distributions.](image-url)
increasing $\alpha$. A more detailed discussion on the impact of the parameter $\alpha$ on the characteristics of information propagation in a multi-layer network is provided in Section 2.5.4 below.

### 2.5.2 Networks with Power-law Degree Distributions

Many real-world networks including the Internet (at the level of autonomous systems), the phone call network, the e-mail network, and the web link network are shown to exhibit power law degree distributions with exponential cut-off [26]. To gain more insight about our results for more realistic network models, we next consider the case where both $F$ and $W$ have power-law degree distribution with exponential cut-off. Namely, we have

$$p_{st}^f = \begin{cases} 0, & s = 0 \text{ or } t = 0, \\ s^{-\gamma_{f,1}} L_{1,1}(e^{-1/\Gamma_{f,1}}) L_{1,2}(e^{-1/\Gamma_{f,2}}), & s, t = 1, 2, \ldots \end{cases}$$

(2.22)

and

$$p_{st}^w = \begin{cases} 0, & s = 0 \text{ or } t = 0, \\ s^{-\gamma_{w,1}} L_{1,1}(e^{-1/\Gamma_{w,1}}) L_{1,2}(e^{-1/\Gamma_{w,2}}), & s, t = 1, 2, \ldots \end{cases}$$

(2.23)

where $L_m(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^m}$ is the $m^{th}$ polylogarithm of $z$.

In order to compute an analytical expression for the size of information epidemics we proceed similarly with the case of doubly Poisson distributions and use our results presented in Section 2.4.3.

For computer simulations, we again set $n = 2 \times 10^5$, and use $\alpha = 0.1, 0.5, 0.9$ as three sample sizes for the network $F$. The corresponding degree distributions are given by (2.22) and (2.23) with $\gamma_{f,1} = \gamma_{f,2} = \gamma_{w,1} = \gamma_{w,2} = 2.5$, and $\Gamma_{f,1} = \Gamma_{f,2} = \Gamma_{w,1} = \Gamma_{w,2} = 10$. With $T_f = T_w$ ranging from zero to one, we compute the empirical size of the information epidemics again via averaging over 100 independent experiments. The results are demonstrated in Figure 2.9 where curves are obtained analytically using our discussion in Section 2.4.3, and markers represent
2.5.3 How does Clustering Affect the Threshold and Size of Information Epidemics?

An important goal of this work is to understand how clustering affects the dynamics of information propagation in multi-layer networks. Given the complexity of the model adopted here this can be studied in several different ways; e.g. with controlling the clustering coefficient of only one of the networks $F$ or $W$, or by adjusting both networks’ clustering simultaneously. Also, the way we adjust the clustering coefficient of a given network can have a significant impact on the conclusions obtained given that such changes might also impact the degree-degree correlations (e.g., assortativity) in the network\(^2\). The situation becomes even more involved as one realizes that the choice of the parameter $\alpha$ changes the assortativity of the network as well.

With these in mind, we consider doubly Poisson distributions in the remainder of this discussion. We first consider a scenario where one or both of the constituent networks in the system is changed from a non-clustered network to a clustered network. More precisely, we compare the following three cases

- Both networks are non-clustered (NN)

\(^2\)The assortativity coefficient is the Pearson correlation coefficient of degree between pairs of linked nodes, and the detail of the computation can be found in [65].
• Network $\mathcal{W}$ is clustered but network $\mathcal{F}$ is non-clustered (NC)

• Both networks are clustered (CC)

Here, the clustered networks are generated as discussed in Section 3.2.1 following the approach of Miller [60] and Newman [68], say with doubly Poisson degree distribution $p_{st}$ with parameter $\lambda_s$ for single edges and $\lambda_t$ for triangle edges. To ensure a fair comparison, we generate non-clustered networks with the same total degree distribution and degree-degree correlations. To this end, we generate the non-clustered networks using the multiplex (i.e., colored) version of the configuration model [86]. Namely, each node gets $\text{Poi}(\lambda_s)$ stubs of color blue and $2 \times \text{Poi}(\lambda_t)$ stubs of color red, and then stubs of the same color are randomly matched to form edges. The standard configuration model where colors are ignored would lead to the same degree distribution, but would fail in capturing the positive degree correlations inherent in the random clustered networks proposed by Miller [60] and Newman [68].

The results are depicted in Figure 2.10, where we compare the relative size of the epidemics as $T_w = T_f$ varies from zero to one, and $\alpha$ is taken to be 0.1, 0.5, or 0.9. The resulting global and local clustering coefficients, and the assortativity values of network $\mathcal{H}$ can be found in Table 2.1. For each $\alpha$ value, we see that clustering increases as we go from $\text{NN}$ to $\text{NC}$ to $\text{CC}$, while assortativity stays the same. Our main conclusion from Figure 2.10 is that for a given $\alpha$, the curve for the NN is always above that of NC, which in turn is always above that of CC. That is, the critical threshold of information epidemics increases while the final epidemic size decreases as we move from $\text{NN}$ to $\text{NC}$ to $\text{CC}$, i.e., as the clustering coefficient in the whole system increases. Therefore, we conclude that the high level of clustering not only makes it more difficult for information to reach a significant fraction of the population, but it also reduces the mean epidemic size at any level of information transmissibility.

The inhibitive effect of clustering on epidemics has been observed in the single network case as well [60], and is often attributed to the fact that the edges used for completing wedges to triangles is redundant for the purposes of information propagation; a wedge is defined as a connected triple that is not a triangle. This is particularly evident when $T_f = T_w = 1$, and
<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Non-clustering and Non-clustering</th>
<th>Non-clustering and Clustering</th>
<th>Clustering and Clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GC</td>
<td>LC</td>
<td>Ass.</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.106</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0.071</td>
</tr>
<tr>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0.044</td>
</tr>
</tbody>
</table>

Table 2.1: Statistics of network $H$ under the setting of Figure 2.10. GC denotes global clustering coefficient, while LC denotes the local clustering coefficient.

Figure 2.10: Comparison of the size of information epidemics between Non-clustered and Non-clustered networks (NN), Clustered and Non-clustered networks (NC), and Clustered and Clustered networks (CC). Plots are obtained from our analytical results. The value following the model abbreviation indicates the amount of overlapping between two networks. For example, NC-0.5 means that $\alpha = 0.5$.

In order to understand the effect of clustering better, we next consider a different setting where we control the level of clustering in network $W$ while keeping its mean total degree fixed. More precisely, we use Poisson distributions for the number of single and triangle edges in both networks with parameters given in Table 3.1. Put differently, network $F$ has a fixed clustering coefficient while with $c \in [0, 4]$ the clustering of $W$ varies between the two extremes: i) when

the size of the epidemics is equal to the giant component size in $H$. It is clear that adding an extra edge to this graph that transforms a wedge into a triangle has no effect on its giant component; in contrast it may be possible to increase the giant component size by adding this extra edge somewhere else in the network. Therefore, as long as the degree distributions and degree-degree correlations are fixed, random networks with low clustering will tend to have a larger epidemic size and a lower epidemic threshold.

In order to understand the effect of clustering better, we next consider a different setting where we control the level of clustering in network $W$ while keeping its mean total degree fixed. More precisely, we use Poisson distributions for the number of single and triangle edges in both networks with parameters given in Table 3.1. Put differently, network $F$ has a fixed clustering coefficient while with $c \in [0, 4]$ the clustering of $W$ varies between the two extremes: i) when
c = 4, \( W \) will have no single-edges and consist only of triangles resulting with a clustering coefficient close to one; and ii) with \( c = 0 \), there will be no triangles in \( W \) and hence its clustering coefficient will be close to zero. Thus, with increasing \( c \), the clustering coefficient of \( W \) increases, which in turn increases clustering in the multilayer network \( H \); see Table 2.3 for specific clustering coefficients corresponding to several \( c \) values considered. We remark that by the choice given in Table 3.1, the degree distribution (single edges plus triangle edges) of \( W \) is given by

\[
2 \text{Poi} \left( \frac{4-c}{2} \lambda \right) + 2 \text{Poi} \left( \frac{c}{2} \lambda \right).
\]

This ensures that as \( c \) varies both the mean and the variance of the degree distribution remains constant, allowing us to focus only on the effect of clustering; for instance, using \( \text{Poi}((4-c)\lambda) \) rather than \( 2 \text{Poi} \left( \frac{4-c}{2} \lambda \right) \) would change the variance of the distribution and hence the threshold for information epidemics (viz. (2.25)). As seen from Table 3.1, the assortativity of the network also remains constant with varying \( c \).

<table>
<thead>
<tr>
<th>Distribution of single-edges</th>
<th>Network ( F )</th>
<th>Network ( W )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Poi}(2\lambda_F) )</td>
<td>( 2 \text{Poi}(\frac{4-c}{2} \lambda_W) )</td>
<td></td>
</tr>
<tr>
<td>( \text{Poi}(\lambda_F) )</td>
<td>( \text{Poi}(\frac{c}{2} \lambda_W) )</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: Parameters of the doubly Poisson distribution. In Figure 2.11 we set \( \lambda_F = \lambda_W = 0.5 \). We use \( \lambda_F = 0.36 \) and \( \lambda_W = 0.5 \) for Figure 2.12.

With these in mind, we first demonstrate in Figure 2.11 the boundary of the \( T_f - T_w \) plane that identifies the threshold of information epidemics. Put differently, for each parameter pair \((c, \alpha)\), the curves in Figure 2.11 separates the region where information epidemics can take place (north and east of the curves) from the region where they can not (south-west of the curves). We see that with the same \( T_f \), clustering increases the minimum \( T_w \) that is needed for information epidemics to be possible. In other words, we see again that clustering increases the threshold of epidemics.

Next, we look at the effect of clustering on the relative final size of information epidemics for specific percolation (i.e., transmissibility) probabilities. From Figure 2.12, we see that the
size of giant component decreases as the clustering coefficient increases, again confirming that high clustering reduces the epidemic size.

Figure 2.11: Comparison of the epidemic boundary under several cases; the north and east of each curve specifies the region of \((T_f, T_w)\) values for which epidemics are possible, while the south and west part of each curve stands for the region where epidemics can not take place. Resulting statistics for clustering and assortativity is given in Table 2.3.

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(c)</th>
<th>assortativity</th>
<th>Clust. Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Global</td>
</tr>
<tr>
<td>0.1</td>
<td>0.01</td>
<td>0.010</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>0.010</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>3.99</td>
<td>0.010</td>
<td>0.185</td>
</tr>
<tr>
<td>0.9</td>
<td>0.01</td>
<td>0.009</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>0.009</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>3.99</td>
<td>0.009</td>
<td>0.126</td>
</tr>
</tbody>
</table>

Table 2.3: Statistics corresponding to the network \(H\) in the setting of Figure 2.11.

2.5.4 How does \(\alpha\) Affect the Information Propagation Dynamics?

We now shift our focus to understanding the impact of the parameter \(\alpha\), which controls the relative size of network \(F\) to network \(W\), on the information propagation dynamics. From a practical perspective, this will help understand the role of the size of an online social network, say Facebook, on propagating the information. As we shall demonstrate soon, this parameter’s impact on the overall network topology goes beyond a change in degree distribution, and thus its effect on epidemic threshold and epidemic size are highly non-trivial.

In order to focus only on the impact of \(\alpha\), we consider non-clustered networks throughout this section; however, a similar discussion would hold for clustered networks as well. Let \(W\) and
Figure 2.12: Illustration of how clustering affects the size of epidemics when $T_f = T_w = 0.3$.

If $F$ be random networks with given degree distributions, and let $H = W \coprod F$ be their disjoint union; i.e., network $H$ is a colored degree-driven random graph introduced in [86]. As before $W$ is defined on $n$ vertices, each of which belongs to vertex set of $F$ independently with probability $\alpha$. For simplicity, we assume that network $W$ has Poisson degree distribution with parameter $\lambda_{\text{red}}$, while network $F$ has degree distribution $\text{Poi}(\lambda_{\text{blue}})$. Under this setting, each of the $n$ nodes in graph $H$ will have a colored degree distribution given by

$$
\begin{align*}
    p_k^{\text{red}} &= e^{-\lambda_{\text{red}} \frac{\lambda_{\text{red}}}{k!}}, \quad k = 0, \ldots, \\
    p_k^{\text{blue}} &= \alpha e^{-\lambda_{\text{blue}} \frac{\lambda_{\text{blue}}}{k!}} + (1 - \alpha) \mathbf{1}[k = 0], \quad k = 0, \ldots.
\end{align*}
$$

(2.24)

where blue edges represent links in $F$ and red edges represent links in $W$. The multiplex network (MN) $H$ is then generated by the colored configuration model where only stubs of the same color are connected together to form an edge.

To check the impact of $\alpha$ on the size of information epidemics in a fair way, we keep the value of $\alpha \lambda_{\text{blue}}$ fixed throughout the experiments. This ensures that the mean number of blue edges in $H$ remains constant as $\alpha$ varies. Put differently, this setting allows us to compare the impact of a small but densely connected social network with a large but loosely connected one in facilitating the propagation of information. Below, we will argue why an adjustment on $\alpha$ changes not only the degree distribution but also the degree-degree correlations in the network. To make this point clearer, we also include in our comparison the simplex network (SN) case which ignores the colors of the edges and generates $H$ via the standard configuration model.
with degree distribution $p_k = p_k^{\text{red}} \oplus p_k^{\text{blue}}$, here $\oplus$ denotes the convolution operator.

The results comparing the final epidemic sizes for three specific $\alpha$ values are given in Figure 2.13. These plots are obtained via computer simulations with $n = 5 \times 10^5$, $\lambda_{\text{red}} = 1$, $\alpha\lambda_{\text{blue}} = 1$, and $T_w = T_f$ is varied from zero to one; each data point corresponds to an average over 100 independent runs. We list the resulting assortativity values for each case in Table 2.4. As expected, the simplex case that corresponds to the standard configuration model has uncorrelated degrees and thus the resulting assortativity is zero. However, we realize that aside from changing the degree distribution in the network, the relative size of $F$ also has a significant impact on the degree-degree correlations in the multiplex case. This impact, namely the positive correlations observed between the degrees of neighbors, is particularly pronounced in the case where $\alpha$ is small; e.g., assortativity is 0.96 when $\alpha = 0.01$. This can be attributed to the fact that when $\alpha$ is close to zero, a very small fraction of nodes receive a large number of blue edges (since $\alpha\lambda_{\text{blue}}$ is fixed) and these extra edges can only be used to connect with other nodes that also have extra edges; as before red edges are assigned to every node. As a result, the network $H$ exhibits a very densely connected (community-like) subgraph on the vertices that participate in $F$, and this leads to highly positive degree-degree correlations given that the nodes in $F$ have significantly larger (in the statistical sense) degrees than nodes that are not in $F$.

![Table 2.4: Comparison of the assortativity values observed in the setting of Figure 2.13 for to the Simplex Network (SN) and the Multiplex Network (MN) case for different $\alpha$. As expected, for the simplex network case the degrees of the nodes are uncorrelated and assortativity is thus zero. The multiplex case exhibits assortative mixing, with the correlations getting more significant with decreasing $\alpha$.](image-url)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>SN</th>
<th>MN</th>
<th>$\alpha$</th>
<th>SN</th>
<th>MN</th>
<th>$\alpha$</th>
<th>SN</th>
<th>MN</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.00</td>
<td>0.96</td>
<td>0.10</td>
<td>0.00</td>
<td>0.65</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

There are a number of interesting conclusions we can derive from Figure 2.13. First, by comparing the simplex and multiplex cases with each other for each $\alpha$ value (i.e., by comparing the line and the marker that are of the same color in Figure 2.13) we see that multiplex
networks have a smaller epidemic threshold as well as a smaller epidemic size as compared to the corresponding simplex network for small $\alpha$ values; on the other hand for $\alpha \simeq 1$, the differences are negligible. This observation is in line with the assortativity values seen in Table 2.4 noting the fact that assortativity is known [60,66] to reduce the critical threshold and the size of epidemics.

Second, we focus on the impact of $\alpha$ on the threshold and size of epidemics by comparing the three lines in Figure 2.13 that correspond to the multiplex case with $\alpha = 0.01$, $\alpha = 0.1$ and $\alpha = 0.99$, respectively. We observe that as $\alpha$ gets larger, the epidemic threshold increases and so does the final epidemic size. What this means is that, it will be more difficult to trigger an information epidemic when the physical network is augmented with a large online social network that is loosely connected, as compared to the case when the online social network is small but densely connected. However, when information transmissibility is high in both networks, final epidemic size is going to be larger in the case of a large but loosely connected online social network as compared to the case of a small but densely connected one. Combining, we conclude that information with low transmissibility spreads more effectively with a small but densely connected social network, whereas highly transmissible information will reach more people with the help of a large but loosely connected social network; here the basis of comparison is again the total number edges in the overlay network.

It is important to remark that the differences observed between the three lines for the
multiplex case may not be solely attributed to the changes in the assortativity levels. This is because when we adjust $\alpha$, the degree distribution of the network changes as well. For example, it is easy to see that as $\alpha$ increases the variance of the degree distribution tends to be lower, which is known [62] to increase the epidemic threshold; i.e., it has a similar impact on the epidemic threshold with reducing the assortativity. To better understand the impact of $\alpha$ on the degree distribution and hence on the information propagation dynamics, we now compare the three simplex network cases in Figure 2.13, i.e., we compare the data points shown with markers. This time as well, we see that as $\alpha$ gets larger the epidemic threshold and the final epidemic size gets larger, although the differences observed are less significant as compared to the multiplex case discussed above. Intuitively speaking, this would be expected since none of these three cases exhibit assortativity and the observed impact is only due to the change in the degree distribution. The observed impact of $\alpha$ on the degree distribution and on the epidemic threshold can in fact be quantified. First, we recall that for a single layer network, the critical threshold for epidemics is given by [62,66]

\[ T \frac{\mathbb{E}[d_i(d_i - 1)]}{\mathbb{E}[d_i]} > 1, \quad (2.25) \]

where $d_i$ is the degree of an arbitrary node $i$. With the choice of the degree distribution given in (2.24) and $T_f = T_w = T$, it is easy to see that this condition reduces to

\[ \alpha \lambda_{\text{blue}} + \lambda_{\text{red}} + \frac{\alpha \lambda_{\text{blue}}^2 - \alpha^2 \lambda_{\text{blue}}^2}{\alpha \lambda_{\text{blue}} + \lambda_{\text{red}}} > \frac{1}{T}, \]

or, equivalently to

\[ T > \frac{2}{3 + 1/\alpha} \]

with our choices of $\alpha \lambda_{\text{blue}} = \lambda_{\text{red}} = 1$. This finding quantifies how the critical threshold should increase with $\alpha$, and it is in perfect agreement with the curves for the simplex case shown in Figure 2.13 as expected.
2.6 Chapter Summary

We analyzed the propagation of information in clustered multilayer networks, where the vertex set of one network is a subset of the vertex set of the other. We solved analytically for the threshold, probability, and mean size of information epidemics, and confirmed our findings via extensive computer simulations. We showed from various angles that clustering increases the epidemic threshold and decreases the final epidemic size in multi-layer networks. We also demonstrate how the overlap between the constituent networks affects the information propagation dynamics, particularly through impacting the degree-degree correlations. For instance, we show that an online social network $F$ that is small in size but large in mean connectivity is more effective in facilitating the propagation of information as compared to a large social network with smaller mean connectivity, with the total number of edges fixed in both cases.

Our general framework contains non-clustered multi-layer networks and single-layer clustered networks as special cases. In addition, given that information propagation problem is studied via bond percolation over a multi-layer network, our work can also be useful in the context of robustness against random attacks – Assume that our system consists of two conjoint networks $F$ and $W$ and an adversary attacks edges in both networks randomly with probabilities $T_f$ and $T_w$, respectively with the aim of disconnecting the whole system. Then, the size and existence of the giant component after edge failures would be natural metrics for the robustness of this system against random attacks. To that end, we believe our results (e.g., Figure 2.11) would be useful in understanding the impact of clustering on the robustness of multi-layer networks.

There are many open problems one might consider for future work. For instance, the impact of assortativity is not fully understood on the propagation of information over multi-layer networks. Another interesting direction would be to consider networks that exhibit clustering not only through triangles, but also through larger cliques. Extending some of the
ideas presented here to the case of influence propagation (e.g., complex contagions) would also be interesting.
Part III

Influence Propagation
Chapter 3

Clustering determines the dynamics of complex contagions in multiplex networks

3.1 Introduction

The study of dynamical processes on real-world complex networks has been an active research area over the past decade. Some of the most widely studied problems include cascading failures in interdependent networks [19, 21, 77, 79, 93, 101], simple contagions (e.g., diffusion of information, disease propagation in human and animal populations [1, 4, 7, 11, 13, 24, 27, 28, 35, 56, 63, 64, 76, 80, 88, 91, 102, 104], etc.), and complex contagions (e.g., diffusion of influence, beliefs, norms, and innovations in social networks [39, 92, 94, 99]). Recently, the attention was shifted from single, isolated networks to multiplex and multi-layer networks [12, 14, 20, 23, 32–34, 43, 46, 54, 61, 80, 84, 98, 99, 102, 106]. This shift is primarily driven by the observation that links in a network might be categorized according to the nature of the relationship they represent (e.g., friends, family, office-mates) as well as according to the social network they belong (e.g., Google+ vs. Facebook links), and each link type might play a different role in the dynamical process.

In this work, we focus on the analysis of complex contagion processes that take place on multiplex (or, multi-layer) networks. In doing so, we adopt the content-dependent linear threshold model of social contagions proposed by Yağan and Gligor [99]. Their framework is a generalization of the linear threshold model introduced by Watts [94] and is based on individuals adopting a behavior when their perceived proportion of active neighbors exceed
a certain threshold; the key to that modeling framework is that one’s perceived influence depends on the *types* of the relationships they have and the *context* in which diffusion is being considered. More precisely, each individual in the network can be in one of the two states, *active* or *inactive*. Each link type-\(i\) is associated with a content-dependent weight \(c_i\) in \([0, \infty)\) that encodes the relative importance of this link type in spreading the given content. Then, an inactive node with \(m_i\) active neighbors and \(d_i - m_i\) inactive neighbors via type-\(i\) links turns active only if

\[
\frac{\sum c_i m_i}{\sum c_i d_i} \geq \tau
\]

where \(\tau\) is the node’s threshold drawn from a distribution \(P(\tau)\). Once *active*, a node stays active forever.

Yağan and Gligor analyzed [99] the content-dependent linear threshold model in multiplex networks and derived the conditions, probability, and expected size of *global cascades*, i.e., cases where activating a randomly selected node leads to activation of a positive fraction of the population in the limit of large system sizes. However, their multiplex network model was formed by combining independent layers of networks (one for each link type), where each layer is generated by the *configuration model* [67, 71]. Although a good starting point, configuration model is known to generate networks that can not accurately capture some important aspects of real-world social networks, most notably the property of *high clustering* [83, 95]. Informally known as the phenomenon that “friends of our friends” are likely to be our friends, clustering has been shown to affect the dynamics of various diffusion processes [29, 31, 42, 47–49, 68, 106] significantly.

With this in mind, our main contribution in this paper is to provide a thorough analysis of influence diffusion process (a complex contagion) in *clustered* multiplex networks. In particular, we study the content dependent linear threshold model in a multiplex network model where each link type (or, network layer) is formed by the clustered random network model proposed by Newman [68] and Miller [60]. We solve for the critical threshold, probability, and expected size of global cascades and confirm our analytical results via extensive simulations. The main
observation from our results is that clustering has a double-faceted impact on the probability and expected size of global cascades. Namely, we show that clustering decreases the probability and size of cascades when average degree in the network is small, whereas after a certain value of average degree, clustering is shown to facilitate cascades.

We also compare the dynamics of complex contagions over multiplex networks and their monoplex projections. There has been recent interest [46] in understanding whether monoplex projection of a multiplex network (obtained by ignoring the colors of edges and aggregating the layers) can still capture the essential properties (e.g., cascade threshold and size) of a diffusion process. In the affirmative, this would eliminate the need for considering the full multiplex structure of real-world systems in tackling similar problems. We show that even in the simplest case where all link types have the same influence weight (i.e., $c_1 = c_2 = \cdots$), monoplex theory may not be able to capture contagion dynamics accurately, reinforcing the need for studying multiplex networks in its correct setup. We demonstrate that the accuracy of monoplex theory in capturing cascade dynamics over multiplex networks depends tightly on the assortativity (i.e., correlation between the degrees of connected pairs) of the network. For instance, when assortativity is negligible, monoplex theory is seen to predict cascade dynamics very well, while in highly assortative cases its ability to predict contagion behavior diminishes significantly.

Finally, we proof the possibility of an unforeseen behavior in the dynamics of complex contagions in multiplex networks, i.e., that of observing more than two phase transitions in the cascade size as the mean degrees in network layers increase. It has been reported [94,102] many times that threshold models of complex contagion exhibit two phase transitions as the average degree increases; a second-order transition at low degrees marking the formation of a giant component of vulnerable nodes and a first-order transition at high degrees due to increased local stability of nodes. Here, we consider a multiplex where one layer has degree distribution $\text{Poi}(\lambda)$ while the degree in the second layer follows $\text{Poi}(\lambda/\alpha)$ with probability $\alpha$ and is zero with probability $1 - \alpha$. In this setting, we observe that in general there exist two intervals of $\lambda$ for which cascades are possible, amounting to four phase transitions as opposed to two; also,
it is seen that only the first transition is second-order while the remaining ones are first order. However, depending on the value of $\alpha$, these regions may overlap (with overlap starting when $\alpha$ exceeds a critical value) resulting again with only two phase transitions; see Section 3.6.3 for details.

The paper is organized as follows. We give details of the models applied in this study and the problem to be considered in Section 3.2. In Sections 3.3 and 3.4 we present the main results of this work, and confirm it through extensive computer simulations in Section 3.5. In Section 3.6, we make a comparison between complex contagions in monoplex networks and multiplex networks, and also demonstrate the new phenomenon about the number of phase transitions. Finally, Section 5.5 summarizes our work and gives future directions.

### 3.2 Model: structure and dynamics

#### 3.2.1 Random Graphs with Clustering

Our goal is to study complex contagion processes in synthetic networks that capture some important aspects of real-world networks but otherwise are generated randomly. It is known [67,71] that the widely used configuration model [71] generates tree-like graphs with number of cycles approaching to zero as the number of nodes gets large. However, most social networks exhibit high clustering, informally known as the propensity of a “friend of a friend” to be one’s friend. Put differently, real-world social networks are usually not tree-like and instead have considerable number of cycles, particularly of size three; i.e., triangles. With this in mind, Miller [62] and Newman [67] proposed a modification on the configuration model to enable generating random graphs with given degree distributions and tunable clustering.

The model proposed in [62,67] is often referred to as random networks with clustering and is based on the following algorithm. Given a joint degree distribution $\{p_{st}\}_{s,t=0}^{\infty}$ that gives the probability that a node has $s$ single edges and $t$ triangles, each node will be given $s$ stubs labeled as single and $2t$ stubs labeled as triangles with probability $p_{st}$, for any $s,t = 0,1,\ldots$.
Figure 3.1: Illustration of multi-layer and multiplex network representations of our model. In (a), we see a multilayer network (e.g., a Physical communication layer and a Facebook layer) with overlapping vertex sets; vertical dashed lines represent nodes corresponding to the same individual. In (b), we see the equivalent representation of this model by a multiplex network. Edges from Facebook are shown in red and edges from the physical network are shown in blue.

Then, stubs that are labeled as single are randomly joined to form single edges that are not part of a triangle, whereas pairs of triangle stubs from three nodes are randomly matched to form triangles between the three participating nodes; the total degree of a node is then distributed by $p_k = \sum_{s,t:s+2t=k} p_{st}$. As in the standard configuration model, it can be shown that the number of cycles formed by single edges goes to zero as $n$ gets large, and so does the number of cycles of length larger than three [71].

We quantify the level of clustering using the widely recognized global clustering coefficient [67], defined via

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

Here, “connected triples” means a single vertex connected by edges to two others. It was shown in [71] that $C_{\text{global}}$ is positive in the random clustered network model, while it approaches to zero with increasing network size in the standard configuration model.

### 3.2.2 Multi-layer and Multiplex Network Models

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

In line with the second motivation, we assume that network \( B \) is defined on the vertices \( \mathcal{N} = \{1, \ldots, n\} \), while \( R \) contains only a *subset* of the nodes in \( \mathcal{N} \) to account for the fact that not every individual participates in online social networks; e.g., we assume that each vertex in \( \mathcal{N} \) participates in \( R \) independently with probability \( \alpha \in (0, 1] \), meaning that the set of vertices of \( R \) constitutes \( \alpha \)-fraction of the whole population. We illustrate in Figure 3.1 two equivalent representations of this model, first shown as a multi-layer network with overlapping vertex sets, and second as a multiplex network.

We generate both \( R \) and \( B \) from the generalized configuration model described in Section 3.2.1; i.e., both are random networks with clustering. In particular, we let \( \{p_{rst}, s, t = 0, 1, \ldots\} \) and \( \{p_{bst}, s, t = 0, 1, \ldots\} \) denote the joint distributions for single edges and triangles for \( R \) and \( B \), respectively. Then both networks are generated independently according to the algorithm described in Section 3.2.1, and they are denoted respectively by \( R = R(n; \alpha, p_{rst}) \) and \( B = B(n; p_{bst}) \). We define the overall network \( \mathbb{H} \) over which influence spreads as the *disjoint* union \( \mathbb{H} = R \biguplus B \) and represent it by \( \mathbb{H}(n; \alpha, p_{rst}, p_{bst}) \). Here, the disjoint union operation implies that we still distinguish \( R \)-edges from \( B \)-edges in \( \mathbb{H} \), meaning that it is a *multiplex* network.

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

\[
d = (d_{rs}, 2n_{rt}, d_{bs}, 2n_{bt})
\] (3.1)

meaning that it has \( d_{rs} \) *single edges* and \( 2n_{rt} \) *triangle edges* in network \( R \), and \( d_{bs} \) *single edges* and \( 2n_{bt} \) *triangle edges* in network \( B \); here \( n_{rt} \) and \( n_{bt} \) are defined as the number of *triangles* assigned to this node in \( R \) and \( B \), respectively. The distribution of this colored degree is
Figure 3.2: Illustration of three cases that would be counted as $m_{rs}$, $m_{rt1}$, and $m_{rt2}$, respectively for the number of active nodes. Nodes shown in filled (green) circles are active while those shown in non-filled circles are inactive.

denoted by $p_d$ and can be computed directly from $p_{st}^r$, $p_{st}^b$, and $\alpha$.

### 3.2.3 Content-dependent Linear Threshold Model for Social Contagion

The classical linear threshold model by Watts [94] is based on individuals adopting a behavior when the fraction of their active neighbors exceed a certain threshold. Namely, an inactive node $i$ with $m_i$ active neighbors and $d_i - m_i$ inactive neighbors will become active only if $m_i/d_i$ exceeds $\tau_i$ drawn from a distribution $P(\tau)$; once active, a node can not be deactivated. A major concern with this model is that it assumes all links in the network have the same importance, irrespective of the context that the spreading is being considered. However, in real world contagion processes, each link type (e.g., co-workership vs. family or physical links vs. online social network links) may play a different role in different cascade processes. For example, in the spread of a new consumer product amongst the population, a video game would be more likely to be promoted among high school classmates rather than among family members; the situation would be exactly the opposite in the case of a new cleaning product [90].

To address the aforementioned drawbacks, Yağan and Gligor [99] proposed a content-dependent linear threshold model for social contagion in multiplex networks. In this model, each link type is associated with a content dependent parameter $c_i$ in $[0, \infty]$ that measures the relative bias type-$i$ links have in spreading the content. Then, an inactive node with $m_i$ active neighbors and $d_i - m_i$ inactive neighbors through link type-$i$ will turn active if $\sum \frac{c_im_i}{\sum c_id_i} \geq \tau$. 

55
In this work, we will analyze complex contagions in $\mathcal{H}$ under the content-dependent threshold model introduced in [99]. Consider a node with colored degree $d = (d_{rs}, 2n_{rt}, d_{bs}, 2n_{bt})$ and active-degree

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

where $m_{rs}$ (resp. $m_{bs}$) gives the number of active neighbors connected through red single edges (resp. blue single edges), and $m_{rt1}$ and $m_{rt2}$ (resp. $m_{bt1}$ and $m_{bt2}$) give the number of red (resp. blue) triangles with one and two active neighbors, respectively; see Figure 3.2 for demonstration of three cases counted as $m_{rs}$, $m_{rt1}$, and $m_{rt2}$, respectively. Next, for a given content to spread over $\mathcal{H}$, let $c_r$ and $c_b$ denote the weight of red- and blue-edges, respectively, in spreading this content. Without loss of generality, we set $c := \frac{c_r}{c_b}$. Then, the probability that an inactive node with degree $d$ and active-degree $m$ turns active is given by

$$F(m, d) = \Pr \left[ \frac{c(m_{rs} + m_{rt1} + 2m_{rt2}) + m_{bs} + m_{bt1} + 2m_{bt2}}{c(d_{rs} + 2n_{rt}) + d_{bs} + 2n_{bt}} \geq \tau \right].$$

Hereafter, the function $F(m, d)$ will be referred to as the neighborhood response function [36, 48].

### 3.2.4 The Problem

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

In this section, we derive the condition and probability of global cascades in clustered multiplex networks; expected size of global cascades is handled separately in Section 3.4. As mentioned in Section 3.2.2, we restrict our attention to networks with only two link types, labeled as blue and red edges, respectively. Distinguishing further the edges based on whether or not they are part of a triangle, we obtain four types of edges in our clustered multiplex network model, labeled as red single edges, red triangles edges, blue single edges, and blue triangle edges; these are denoted by rs−, rt−, bs−, and bt−, respectively.

To analyze the influence diffusion process, we consider a branching process [6] that starts by activating a node selected randomly from among all nodes. Starting with the neighbors of the seed node, we explore and identify all nodes that are reached and activated, continuing recursively until the branching process stops. Let $H(x)$ denote the generating function [97] for the “finite number of nodes that are reached and influenced” by the branching process [6] initiated by a randomly selected node. We will derive an expression for $H(x)$ using $h_{rs}(x)$, $h_{rt}(x)$, $h_{bs}(x)$, and $h_{bt}(x)$, where $h_{rs}(x)$ (resp. $h_{bs}(x)$) stands for the generating function for the “finite number of nodes reached by following a randomly selected red single (resp. blue single) edge”; $h_{rt}(x)$ and $h_{bt}(x)$ are defined similarly for red triangle and blue triangle edges, respectively. Then, $H(x)$ takes the form

$$H(x) = x \sum_d p_d D(d_{rs}, n_{rt}, d_{bs}, n_{bt}),$$  \hspace{1cm} (3.3)

where

$$D(y, z, m, \ell) := h_{rs}(x)^y h_{rt}(x)^z h_{bs}(x)^m h_{bt}(x)^\ell$$  \hspace{1cm} (3.4)

The validity of (3.3) can be seen as follows. The term $x$ stands for the node that is selected
randomly and set active to initiate the cascade. This node has a degree $d = (d_{rs}, 2n_{rt}, d_{bs}, 2n_{bt})$ with probability $p_d$. The number of nodes reached by each of its $d_{rs}$ (resp. $d_{bs}$) red single edges (resp. blue single edges) has a generating function $h_{rs}(x)$ (resp. $h_{bs}(x)$). Considering its triangle edges in a similar manner, we see from the powers property of generating functions [71] that when the initial node has degree $d$ the number of nodes influenced in this process has a generating function $h_{rs}(x)^{d_{rs}} h_{rt}(x)^{n_{rt}} h_{bs}(x)^{d_{bs}} h_{bt}(x)^{n_{bt}}$. Averaging over all possible degrees $d$ of the initial node, we get (3.3).

For (3.3) to be useful, we shall derive expressions for the generating functions $h_{rs}(x)$, $h_{rt}(x)$, $h_{bs}(x)$, and $h_{bt}(x)$. As will become apparent soon, there are no explicit equations defining these functions. Instead, we should seek to derive recursive equations to define each generating function in terms of the others. These steps are taken in the next sections where we first focus on deriving $h_{rs}(x)$ and $h_{bs}(x)$ (Section 3.3.1) followed by derivations of $h_{rt}(x)$ and $h_{bt}(x)$ (Section 3.3.2).

Given that random networks with clustering are free of cycles of size larger than three, it is clear that the initial stages of the branching process will expand largely because of vulnerable nodes that can get activated either by one or two active neighbors. In our formulation, the multiplex nature of the network calls for defining the notion of the vulnerability with respect to link types as well [99]. Throughout, we say that a node is $\mathbb{R}$-vulnerable (resp. $\mathbb{B}$-vulnerable) if it gets activated by a single active connection through a red link (resp. blue link). We define $\rho_{d,rs}$ and $\rho_{d,bs}$ as the probability that a node is $\mathbb{R}$-vulnerable and $\mathbb{B}$-vulnerable, respectively. We also define $\rho_{d,rt}$ (resp. $\rho_{d,bt}$) as the probability that a node gets activated by having two active neighbors via red (resp. blue) edges. In other words, $\rho_{d,rt}$ (resp. $\rho_{d,bt}$) gives the probability that a node gets activated by having a red (resp. blue) triangle with both neighbors being active; see Figure 3.2. More precisely, we set $\rho_{d,rs} = F[(1,0,0,0,0),d]$, $\rho_{d,rt} = F[(0,0,1,0,0),d]$, $\rho_{d,bs} = F[(0,0,0,1,0),d]$, and $\rho_{d,bt} = F[(0,0,0,0,1),d]$. 

58
3.3.1 Influence Propagation via Red Single Edges

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

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

where $D$ is as defined at (3.4).

We now explain each term appearing at (3.5) in turn. The explicit factor $x$ stands for the initial vertex that is arrived at by following the randomly selected red single-edge. The term $\frac{d_{rs}p_{d}}{\langle d_{rs} \rangle}$ gives the normalized probability that the arrived vertex has colored degree $d$. Since the arrived node is reached by a red link, it needs to be red-vulnerable to be added to the vulnerable component. If the arrived node is indeed red-vulnerable, which happens with probability $\rho_{d,rs}$, it can activate other nodes via its remaining $d_{rs} - 1$ red single edges, $d_{bs}$ blue single edges, $n_{rt}$ red triangles, and $n_{bt}$ blue triangles. Because the number of vulnerable nodes reached by each of its red single edges and triangles (resp. blue single edges and triangles) are generated in turn by $h_{rs}(x)$ and $h_{rt}(x)$ (resp. $h_{bs}(x)$ and $h_{bt}(x)$) respectively, we obtain the term $h_{rs}(x)^{d_{rs}-1}h_{rt}(x)^{n_{rt}}h_{bs}(x)^{d_{bs}}h_{bt}(x)^{n_{bt}}$ by the powers property of generating functions. Averaging over all possible colored degrees $d$ gives the first term in (3.5). The second term with the factor $x^0$ accounts for the possibility that the arrived node is not red-vulnerable and
Thus is not included in the cluster. An analogous expression can be obtained for $h_{bs}(x)$ via similar arguments.

### 3.3.2 Influence Propagation via Red Triangles

We now derive $h_{rt}(x)$, i.e., the generating function for the number of nodes influenced by following a red triangle selected at random; similar arguments hold for $h_{bt}(x)$. We demonstrate this situation in Figure 3.3, where the top vertex $u$ is active, and we are interested in computing the generating function for the number of nodes that will be influenced by nodes $v$ and $w$. We will compute the generating function $h_{rt}(x)$ by conditioning on the following events:

- If neither of nodes $v$ and $w$ is $\mathbb{R}$-vulnerable, then the number of nodes influenced will be zero. Node $v$ has degree $d$ with normalized probability $\frac{n_{rt}p_d}{\langle n_{rt} \rangle}$, in which case it is not $\mathbb{R}$-vulnerable with probability $1 - \rho_{d,rs}$. Similarly, the probability that node $w$ has degree $d'$ and not $\mathbb{R}$-vulnerable is $\left(\frac{n_{rt}p_{d'}}{\langle n_{rt} \rangle}\right)(1 - \rho_{d',rs})$. Summing over all possible cases, we obtain the first term in (3.7) with $x^0$ (meaning that zero nodes will be influenced by following the red triangle in this case).

- Consider the case where only one of $v$ and $w$ is influenced, leading to a term with the factor $x^1$ in (3.7). Without loss of generality, consider the case where $v$ is activated while $w$ is not. If node $v$ has degree $d$, then it is $\mathbb{R}$-vulnerable with probability $\rho_{d,rs}$, and can influence other nodes in the usual manner. Then, the event that node $w$, with degree $d'$, will not be activated despite having two active neighbors (nodes $u$ and $v$) has probability $1 - \rho_{d',rt}$. By symmetry and exchangeability of nodes $v$ and $w$, an equivalent term will
be obtained for the case where \(w\) is activated but \(v\) is not. Summing over all possibilities we obtain the second term in (3.7).

- Finally, we consider the case where both \(v\) and \(w\) become active giving rise to term with factor \(x^2\). There are two possible scenarios:
  
  - Both of \(v\) and \(w\) are activated by \(u\). The probability that \(v\) is activated by \(u\) is \(\rho_{d,rs}\) as already discussed during the computation of the first term. By symmetry, the probability for \(w\) is the same as for \(v\). Multiplying the two probabilities leads to the third term in (3.7).
  
  - Only one of \(v\) and \(w\) is made active immediately by \(u\) while the other is not; e.g., say \(v\) is activated but not \(w\). However, \(w\) also gets activated by the joint influence from \(u\) and \(v\). With \(d\) and \(d'\) denoting the degree of \(v\) and \(w\), this happens with probability \((\rho_{d,rs})(\rho_{d',rt} - \rho_{d',rs})\). Here, the second term accounts for the fact that \(w\) gets activated only if it has two (or, more) active neighbors. Summing over all possibilities as before, and multiplying by two for the case where \(v\) and \(w\) are replaced, we obtain the last term in (3.7).

### 3.3.3 Deriving the Condition for Global Cascades

The discussion given in Section 3.3.1 and 3.3.2 leads to a set of recursive equations for \(h_{rs}(x)\), \(h_{rt}(x)\), \(h_{bs}(x)\), and \(h_{bt}(x)\). Recursions for \(h_{rs}(x)\) and \(h_{rt}(x)\) are given in (3.6) and (3.7), respectively; the expressions for \(h_{bs}(x)\) and \(h_{bt}(x)\) are very similar and omitted here for brevity. With these four recursive equations in place, it is possible to determine the characteristic function \(H(x)\) of the finite number of nodes activated in the contagion process. Namely, for a given \(x\), we shall find a fixed-point of these recursive equations, and then use the resulting values of \(h_{rs}(x)\), \(h_{rt}(x)\), \(h_{bs}(x)\), and \(h_{bt}(x)\) in (3.3) to get \(H(x)\).
\[ h_{rs}(x) = x \sum_{d} \frac{d_{rs} p_d}{\langle d_{rs} \rangle} \rho_{d,rs} D(d_{rs} - 1, n_{rt}, d_{bs}, n_{bt}) + x^0 \sum_{d} \frac{d_{rs} p_d}{\langle d_{rs} \rangle} (1 - \rho_{d,rs}) \tag{3.6} \]

\[ h_{rt}(x) = x^0 \sum_{d} \sum_{d'} \frac{n_{rt} p_d}{\langle n_{rt} \rangle} (1 - \rho_{d,rs}) \frac{n'_{rt} p_{d'}}{\langle n'_{rt} \rangle} \langle 1 - \rho_{d',rs} \rangle \tag{3.7} \]

By conservation of probability and the definition of generating functions, we know that \( H(1) = 1 \) only if final number of activated nodes is finite with probability one. In other words, global cascades that lead to a positive fraction of influenced nodes are possible only if \( H(1) < 1 \). This prompts us to seek a fixed point of the recursive equations when \( x = 1 \). For notational convenience, we define \( h_1 := h_{rs}(1), h_2 := h_{rt}(1), h_3 := h_{bs}(1), \) and \( h_4 := h_{bt}(1) \). From (3.3), this gives

\[ H(1) = \sum_{d} p_d h_1^{d_{rs}} h_2^{n_{rt}} h_3^{d_{bs}} h_4^{n_{bt}}, \tag{3.8} \]

while the recursions take the form

\[ h_i = g_i(h_1, h_2, h_3, h_4), \quad i = 1, 2, 3, 4. \tag{3.9} \]

Here the functions \( g_1, g_2 \) are easily obtained from (3.6) and (3.7), and similarly \( g_3 \) and \( g_4 \) can
be obtained from the recursions for $h_{bs}(x)$, and $h_{bt}(x)$. To give an example, we have

$$g_3(h_1, h_2, h_3, h_4) = \sum_d \frac{d_{bs}p_d}{\langle d_{bs} \rangle} \left( \rho_{d,bs} h_1^{d_{rs}} h_2^{d_{st}} h_3^{d_{bs}-1} h_4^{n_{bt}} + 1 - \rho_{d,bs} \right)$$

It is clear that the recursions (3.9) have a trivial fixed point $h_1 = h_2 = h_3 = h_4 = 1$ which yields $H(1) = 1$, meaning that cascades will die out without reaching a positive fraction of the population with high probability. However, the trivial solution is the physical solution only if it is a stable fixed point. To check its stability, we linearize the recursive equations (3.9) around $x = 1$, and compute the corresponding Jacobian matrix $J$ via

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

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

### 3.4 Expected Cascade Size

Next, we are interested in computing the expected size of global cascades when they take place. Put differently, we will analyze the expected fraction of nodes that will eventually become active as we randomly pick a node in the network and set it active. We follow the approach used in [41, 48, 99], which has been proven to be an effective way to analyze expected
cascade size in networks.

First, consider the network $H$ as a tree-structure with an arbitrary node selected as the root. Then, label the levels of the tree from $\ell = 0$ at the bottom to $\ell = +\infty$ at the top of the tree. Similar to [48, 99], we assume that nodes begin updating their states starting from the bottom of the tree and proceeding to the top. In other words, we assume that a node at level $\ell$ updates its state only after all nodes at the lower levels $0, 1, \ldots, \ell - 1$ finish updating. We define $q_{rs,\ell}$ as the probability that a node at level $\ell$ of a tree, which is connected to its unique parent by a red single edge, is active given that its parent at level $\ell + 1$ is inactive. Then, we consider a pair of nodes at level $\ell$ that together with their parent at level $\ell + 1$ form a red triangle. Given that the parent is inactive, we let $q_{rt1,\ell}$ (resp. $q_{rt2,\ell}$) denote the probability that only one (resp. both) of the two child nodes of this triangle is active. We define $q_{bs,\ell}$, $q_{bt1,\ell}$, and $q_{bt2,\ell}$ for blue edges in the same manner.

According to our model, an active node is never deactivated, meaning that $q_{rs,\ell}$, $q_{rt1,\ell}$, $q_{rt2,\ell}$, $q_{bs,\ell}$, $q_{bt1,\ell}$, $q_{bt2,\ell}$ are all non-decreasing. Therefore, they will converge to $q_{rs,\infty}$, $q_{rt1,\infty}$, $q_{rt2,\infty}$, $q_{bs,\infty}$, $q_{bt1,\infty}$, $q_{bt2,\infty}$. Then, the expected cascade size (i.e., the fraction of active individuals) $S$ is given by the probability that the arbitrary selected node at the top of the tree becomes active. In order to compute $S$, we first derive recursive relations for $q_{rs,\ell}$, $q_{rt1,\ell}$, $q_{rt2,\ell}$, $q_{bs,\ell}$, $q_{bt1,\ell}$, $q_{bt2,\ell}$. We have
\[q_{rs,\ell+1} = \sum_{d} \frac{d_{rs}p_d}{\langle d_{rs} \rangle} \sum_{n_{rt}} \sum_{j} \sum_{m=0}^{n_{bt}} \sum_{n=0}^{n} Q_{\ell} [(i, j, m, y, n), (d_{rs} - 1, n_{rt}, d_{bs}, n_{bt})] \times F [(i, x, j - x, m, y, n - y), d] \]

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

\[Q_{\ell} [(i', j', x', m', y', n'), (d'_{rs}, n'_{rt} - 1, d'_{bs}, n'_{bt})] F [(i, x, j - x, m, y, n, y - y'], d'] \]

\[Q_{\ell} [(i', j', x', m', y', d'), (d'_{rs}, n'_{rt} - 1, d'_{bs}, n'_{bt})] \]

\[(1 - F [(i', x' + 1, j' - x', m', y', n' - y'), d']) \]

\[F [(i, x + 1, j - x, m, y, n - y), d] - F [(i, x - x, m, y, n, y - y), d'] \times F [(i', x' - x', m', y', n' - y'), d'] \]

where we define

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

\[\times (1 - q_{bs,\ell})^{d_3 - m} \binom{d_4}{n} q_{bt,\ell} (1 - q_{bt,\ell})^{d_4 - n} \]

(3.11)

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

- \(i\) (resp. \(d_1 - i\)) of the \(d_1\) neighbors connected through red single edges as active (resp. inactive). Similarly, \(m\) (resp. \(d_3 - m\)) of the \(d_3\) neighbors connected through blue single edges as active (resp. inactive)
of the $d_2$ red-triangles it participates in, $x$ has one active node, $j-x$ has two active nodes, and $d_2-j$ has no active node. Similarly, of the $d_4$ blue-triangles it participates in, $y$ has one active node, $n-y$ has two active nodes, and $d_4-n$ has no active node.

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

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

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

(3.12)

Namely, we first solve for the values of $q_{rs,\infty}$, $q_{rt1,\infty}$, $q_{rt2,\infty}$, $q_{bs,\infty}$, $q_{bt1,\infty}$, $q_{bt2,\infty}$ using the recursive equations, and then substitute them into (5.5) to obtain the expected size of global cascades.
3.5 Numerical Results

3.5.1 Networks with Doubly Poisson Distributions

In our first simulation study, we use doubly Poisson distribution for the number of single edges and triangles in both networks. Namely, we set

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

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

where \(s\) and \(t\) are the number of single edges and triangles in the corresponding networks, respectively. Thus, \(\lambda_{r,1}\) and \(\lambda_{r,2}\) (resp. \(\lambda_{b,1}\) and \(\lambda_{b,2}\)) denote the mean number of single edges and triangles, respectively in \(\mathbb{R}\) (resp. in \(\mathbb{B}\)).

We consider \(n = 1 \times 10^5\) nodes in the population and \(\alpha = 0.5\) for the size of network \(\mathbb{R}\). We let \(\tau = 0.18\) and \(c = 0.25\) for the threshold and content parameters, respectively. The results are shown in Figure 3.4 where the curves stand for the theoretical results of probability \(P_{\text{trig}}\) and expected size \(S\) of cascades (obtained from our discussion in Section 3.3 and 3.4), as a function of \(\lambda_{r,1} = \lambda_{r,2} = \lambda_{b,1} = \lambda_{b,2}\). The markers stand for the empirical results for the same quantities, and are obtained by averaging over 5,000 independent experiments. We see a very good agreement between the analytical and experimental results confirming the validity of our analysis. The slight discrepancy observed in \(P_{\text{trig}}\) is due to the limited number of experiments, and can be mitigated by increasing the number of realizations.

Next, we change our experimental set-up to demonstrate the effect of content parameter on the probability and size of cascades. To that end, we fix all network parameters and observe the quantities of interest as the content parameter \(c\) varies. In particular, we set \(\lambda_{r,1} = \lambda_{r,2} = \lambda_{b,1} = \lambda_{b,2} = 0.3\) and \(\tau = 0.18\). We see that the probability and expected size of global cascades vary greatly as \(c\) changes. This can be taken as an indication that our model can capture the real-world phenomenon that over the same population certain contents can
Figure 3.4: Simulations for doubly Poisson degree distributions. In (a), we set the content parameter $c = 0.25$, the threshold as $\tau = 0.18$, and $\alpha = 0.5$, and vary the degree parameters. In (b), we fix $\tau = 0.18$, $\lambda_{r,1} = \lambda_{r,2} = \lambda_{b,1} = \lambda_{b,2} = 0.3$, and $\alpha = 0.5$ while varying content parameter $c$.

Figure 3.5: Illustration of the effect of clustering coefficient on the expected cascade and probability of global cascades. We fix $\tau = 0.18$, $c = 0.25$, and $\alpha = 0.5$, then vary the degree parameter $\lambda$ defined in Table 3.1. We see (a) the probability to trigger a global cascade; (b) the global clustering coefficient introduced in Section 3.2.1; and (c) the expected cascade size.

become widespread while others die out quickly. In the setting used here, we see that global cascades take place when the content parameter $c$ is not too small or large. The reason is that with a too small or large content parameter, the connectivity of the conjoined network is dominated by only one of the two networks. So, if neither of them has enough connectivity to trigger a global cascade by their own, then there will be no global cascades in the conjoined network. When the $c$ is neither too large nor too small (e.g., close to unity), both networks will contribute to the connectivity together and it becomes possible to trigger a global cascade. For other values of $\lambda_r$ and $\lambda_b$ a completely different situation might occur, e.g., with very small or very large $c$ promoting cascades; e.g., see [99, Fig. 2] for a few such examples.
3.5.2 How does Clustering Affect the Cascade Size?

Our next goal is to reveal the impact of clustering on the influence propagation process. In order to control the level of clustering while keeping the mean total degree fixed, we use Poisson distributions for the number of single and triangle edges in two networks with parameters given in Table 3.1. Obviously, the clustering coefficient in $\mathbb{R}$ is fixed while with $\eta \in [0, 4]$ the clustering of $\mathbb{B}$ varies between the two extremes: i) when $\eta = 4$, $\mathbb{B}$ will have no single-edges and consist only of triangles resulting with a clustering coefficient close to one; and ii) with $\eta = 0$, there will be no triangles in $\mathbb{B}$ and hence its clustering coefficient will be close to zero. Collecting, we see that the clustering coefficient of $\mathbb{B}$ and thus of $\mathbb{H}$ increases with increasing $\eta$ in this setting.

<table>
<thead>
<tr>
<th>Distribution of single-edges</th>
<th>Network $\mathbb{R}$</th>
<th>Network $\mathbb{B}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poi($2\lambda$)</td>
<td>2 Poi($\frac{\lambda^2}{2}$)</td>
<td></td>
</tr>
</tbody>
</table>

| Distribution of triangles    | Poi($\lambda$) | Poi($\frac{\lambda}{2}$) |

Table 3.1: Parameters of the doubly Poisson distribution. This choice ensures that the mean and variance of the total degree distribution (single plus triangle edges) in $\mathbb{B}$ are independent of $\eta$, while its clustering varies greatly as $\eta$ varies in $(0, 4)$. In Figure 3.4 we set $\lambda = 0.5$.

With these in mind, we first demonstrate the impact of clustering on the probability of triggering a global cascade. Figure 3.5(a) shows the probability of triggering a global cascade as a function of $\lambda$ for three different $\eta$ values. The resulting clustering coefficients are plotted in Figure 3.5(b) where we clearly see that clustering increases with increasing $\eta$. The main observation from Figure 3.5(a) is that increasing the clustering (i.e., increasing $\eta$), shifts the interval of $\lambda$ for which global cascades are possible to the right. This leads to a double-faceted conclusion that clustering decreases the probability of global cascades when average degrees are small, whereas after a certain value of average degree, clustering increases the probability of cascades.

The double-faceted impact of clustering on cascade probability can be explained as follows. It is known [94,102] that threshold models of complex contagion exhibit two phase transitions as the average degree increases, a second-order transition at low degrees that marks the formation
of a giant vulnerable component and a first-order transition at high degrees due to increased local stability of nodes; namely, due to the increased difficulty of activating high-degree nodes. Given that clustering is known to decrease the size of giant component \[106\], we expect that it will be more difficult for a clustered network to contain a giant vulnerable cluster. This is why the lower phase transition in complex contagions appear later (i.e., at larger degrees) as clustering increases. On the other hand, the cycles of size three (i.e., triangles) that are common in clustered networks can help trigger cascades when average degree is higher. For instance, in a tree-like network a single active node can only activate its vulnerable connections. However, in a triangle, an active node may first activate one of its vulnerable connections, making it possible for the third node to be activated (which now has two active neighbors) even if it is not vulnerable. This is what pushes the second phase transition to higher degrees.

Next, we explore the impact of clustering on the expected cascade size in Figure 3.4(c). Here again, we see the double-faceted impact of clustering with small average degrees favoring low clustering, while high degrees favoring high clustering in terms of having a larger cascade size.\(^1\) In fact, we see the existence of a critical average degree (around \(\lambda = 0.6\) in Figure 3.4(c)) such that when \(\lambda\) is smaller (resp. larger) than the critical value, expected cascade size decreases (resp. increases) with increasing clustering.

Finally, we consider the impact of clustering on the average degree-cascade threshold plane. For each parameter pair \((\lambda, \tau)\), the curves in Figure 3.6 separate the region where global cascade can take place (areas inside the boundaries) from the region where they cannot (areas outside the boundaries). Once again, we confirm that increasing the clustering coefficient shifts the interval where cascades are possible up (i.e., to higher degrees) for any threshold \(\tau\).

\(^1\)Hackett et al. [48] also found that clustering will decrease the expected cascade size when the average degree is small.
Figure 3.6: We show the cascade regions in the Degree Parameter-Threshold plane when \( \alpha = 0.5, \tau = 0.18, \) and both networks follow doubly Poisson distributions as described in Table 3.1. Clustering increases as \( \eta \) increases.

### 3.6 Comparison between monoplex and Multiplex Networks

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

To identify the factors affecting complex contagions, we consider two different degree distributions to generate the networks. In Section 3.6.1, we use a setting similar to previous sections, with the resulting networks having almost no degree-degree correlations; e.g., assortativity defined as the Pearson correlation coefficient between the degrees of pairs of linked nodes [65]. In Section 3.6.2 and 3.6.3, we use a different setting that leads to (tunable) assortativity for multiplex networks. In order to keep the focus on the comparison between monoplex and multiplex networks, we shall consider only non-clustered networks in the following discussion.
3.6.1 Multiplex Networks with Limited Assortativity

First we consider the limited assortativity case and use the following degree distribution to assign blue and red stubs to each node:

$$p_k^b = e^{-\lambda_b} \frac{\lambda_b^k}{k!}, \quad k = 0, \ldots,$$

$$p_k^r = \alpha e^{-\lambda_r} \frac{\lambda_r^k}{k!} + (1 - \alpha) 1[k = 0], \quad k = 0, \ldots.$$

A multiplex network is generated using the colored configuration model [86,87] where stubs that are of the same color are matched randomly. The monoplex projected theory ignores the color of the edges and matches all stubs randomly with each other. An important question is whether we lose any significant information about contagion dynamics when the monoplex projected theory is used instead of the multiplex theory developed here and in [99]. For convenience and fair comparison, we set $\lambda_b = \lambda_r$ and use $c = 1$ as the content parameter.

In Figure 3.7(a), we set $\alpha = 0.99$. We see nearly no difference between the theoretical cascade sizes obtained from monoplex and multiplex theories, and they both match the simulation results well. However, when $\alpha$ is reduced to 0.1 in Figure 3.7(b), we clearly see a difference between the two theories and only multiplex theory matches the simulation results. This shows that even in the simplest case where both link types have the same influence factor (i.e., $c = 1$), monoplex theory may be unable to capture certain properties of cascade dynamics, reinforcing the need for studying cascades using the multiplex theory.

We now explain why the two cases, $\alpha = 0.1$ and $\alpha = 0.99$, lead to different conclusions regarding the accuracy of the monoplex theory in capturing contagion dynamics over multiplex networks. One of the key differences between the two cases is the resulting assortativity. In the former case with $\alpha = 0.1$, only 10% of the nodes has red stubs, each of which shall be connected with other red stubs in the multiplex network case. Put differently, in this setting a small fraction of the population will have statistically higher degrees than the rest, and the additional links they have can only connect nodes with high degrees together. This leads
Figure 3.7: Comparison between monoplex networks and multiplex networks with limited assortativity. In (a) and (b), we fix the threshold $\tau = 0.15$, the content parameter $c = 1$, then vary the degree parameters in (3.13). For the networks obtained by projected theory and the networks in multiplex theory with $\alpha = 0.99$, assortativity is negligible. However, when $\alpha = 0.1$, the assortativity coefficient of the networks in the multiplex theory become significant; e.g., it can be up to 0.21.

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

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

3.6.2 Multiplex Networks with Assortativity

In this section, we change the setting slightly to generate multiplex networks with high assortativity. To that end, we use the degree distributions given at (3.13), but instead of setting
\[ \lambda_r = \lambda_b, \] we enforce

\[ \alpha \lambda_r = \lambda_b \tag{3.14} \]

for any \( \alpha \in (0,1) \). This setting allows us to tune assortativity without changing the mean degree in the network. In particular, assortativity will increase as \( \alpha \) decreases (by virtue of a small fraction of nodes forming a highly-connected cluster) [106]. In addition, this setting allows us to compare the contagion dynamics in multi-layer networks when the upper layer is i) small but densely connected (small \( \alpha \)) versus ii) large but loosely connected (large \( \alpha \)); see [106] for relevant results for bond percolation processes.

Using the above degree distributions, we generate monoplex and multiplex networks as in Section 3.6.1 and analyze the complex contagion process. In Figure 3.8(a), we see that when \( \alpha = 0.99 \), which leads to very limited assortativity, the difference between monoplex and multiplex networks is negligible. This is in parallel with what we observed in Section 3.6.1. However, decreasing \( \alpha \) to 0.1 leads to two interesting observations in Figure 3.8(b). First, instead of the commonly reported two phase transitions [94,102], we observe four phase transitions in the cascade size as \( \alpha \lambda_r = \lambda_b \) increases. Secondly we see a significant difference between the monoplex projected theory and multiplex theory, with multiplex theory matching the simulations perfectly. Once again, this shows that monoplex theory is unable to capture the cascade dynamics under certain settings.

The emergence of four phase transitions in Figure 3.8(b), which to the best of our knowledge was not reported before, can be explained as follows. When \( \alpha = 0.1 \), only ten percent of the nodes have red edges, but the the mean number of red edges for those nodes equals \( 10\lambda_b \) (see (3.14)). Therefore, the first couple of phase transitions taking place at very small \( \lambda_b \) values can be attributed mainly to red-edges. First, \( \lambda_b \) becomes large enough (e.g., gets to around 0.1) that the sub-network induced only on the red-edges contains a giant vulnerable cluster, giving rise to global cascades; note that at this point \( \lambda_b \) is so small that blue edges do not create enough local stability to prevent cascades from happening. However, after \( \lambda_b \) reaches a
Figure 3.8: *Comparison between monoplex networks and multiplex networks with assortativity.* Similar with the observation in Figure 3.7, networks in the projected theory and in the multiplex theory with $\alpha = 0.99$ have negligible assortativity coefficients. However, for the networks of multiplex theory with $\alpha = 0.1$, assortativity coefficient ranges from 0.19 to 0.79. In general, assortativity increases with increasing $\lambda_r$ and $\lambda_b$ in the multiplex theory.

At a certain level (around 0.65), the subgraph on red-edges, having average degree of $10\lambda_b$, reaches the second phase transition point where cascades stop due to the increased local connectivity of nodes. These first two transitions being second- and first-order, respectively, also confirms that they are primarily due to the red-edges.

As $\lambda_b$ increases further we observe an interval where there are no global cascades due to either colors of edges; nodes with red and blue-edges are highly stubborn while nodes with only blue edges are not connected enough to trigger a cascade. This interval is then followed by a region where $\lambda_b$ is large enough that the sub-graph on blue edges has a giant vulnerable cluster. However, the emergence of a second-order transition in the whole network is prevented here due to some of these nodes turning stubborn as a result of their red-edges. Eventually, however, $\lambda_b$ becomes large enough that even with occasional stubborn nodes present, a giant vulnerable cluster emerges. This point is reached much later in monoplex networks as compared to the multiplex networks. This is because in the former case stubborn nodes (with red edges) are equally likely to be connected with any other node, while in the latter case they are mostly connected with each other; thus in the latter case they are less likely to inhibit the emergence of a giant vulnerable cluster on blue edges.

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

### 3.6.3 Two vs. Four Phase Transitions

In Section 3.6.2, we have observed the possibility of having more than two phase transitions in the cascade size. As discussed there, multiple phase transitions occur mainly due to the setting (3.14) that, with small $\alpha$, ensures a small fraction of nodes having significantly higher connectivity than the rest, while also being mostly connected with each other. Since the existence of more than two transitions has not been reported in previous studies, we are interested in exploring it further. In particular, we now investigate the impact of $\alpha$ on the number of phase transitions as well as transition points. Of particular interest will be to find the critical $\alpha$ value that separates the cases where four phase transitions occur from those with only two transitions; e.g., the $\alpha$ value for which the two cascade regions overlap. For simplicity we only consider multiplex networks in this section.

Figure 3.9 shows the expected size of global cascades under (3.13)-(3.14) for three different values of $\alpha$. We see that global cascades take place over a single interval of $\alpha \lambda_r = \lambda_b$ when $\alpha$ is large (e.g., $\alpha = 0.99$) while over two disjoint intervals when $\alpha$ is small (e.g., $\alpha = 0.1$). When $\alpha$ is somewhere in between (e.g., case $\alpha = 0.166$) it is possible to have the cascade intervals partially overlap. In such cases, we only see a single interval where global cascades take place. However, an additional transition point appears, manifested by a shift of slope in cascade size,
marking possibly the overlapping point of (what would be) the two cascade intervals.

Figure 3.9 allows us to comment also on the impact of the size and density of the overlaying social network in facilitating influence propagation. With (3.14) in effect, a small $\alpha$ corresponds to a social network with few but densely connected individuals, while large $\alpha$ corresponds to a social network with many subscribers, each with few connections on average. In all cases, the total number of edges in the social network is fixed by virtue of (3.14). We see from Figure 3.9 that the comparison between the three cases leads to a multi-faceted picture as the mean number of links $\alpha \lambda_r$ varies. For instance, the large but loosely connected case of $\alpha = 0.99$ leads to the largest expected global cascade size over a certain interval, but it has the smallest cascade interval among all three. The intermediate case of $\alpha = 0.166$ seems like a stretched version (over the $x$-axis) of the case with $\alpha = 0.99$. In particular, this case leads to the largest interval where global cascades are possible, though the expected cascade size is smaller than that obtained with $\alpha = 0.99$ (and also with $\alpha = 0.1$) over certain intervals. Finally, the case of a small but densely connected overlay (i.e., with $\alpha = 0.1$), falls right under the case with $\alpha = 0.166$ for most values of $\alpha \lambda_r$, though it gives the largest size of all three in small intervals where $\alpha \lambda_r$ is very small or very large.

3.7 Chapter Summary

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

Our analysis and modeling framework subsumes some previous studies. For instance, by setting $h_{rt}(x) = h_{bt}(x) = 1$ in the recursive relations, we ensure that $\mathbb{R}$ and $\mathbb{B}$ are non-clustered random networks. So, our analysis corresponds to complex contagions in non-clustered networks, which was studied in [99]. Similarly, if we let $h_{rs}(x) = h_{rt}(x) = 1$ in the recursions and set the content parameter $c = 1$, then our analysis corresponds to complex contagions in clustered monoplex networks, which was studied in [48].

Future work may consider in more details the impact of assortativity or other topological features on the cascade dynamics. It would also be interesting to compare multiplex networks and their monoplex projections in terms of other dynamical processes; e.g., site percolation, transport processes, etc.
Chapter 4

Multi-Stage Complex Contagions in Random Multiplex Networks

4.1 Introduction

Modeling and analysis of dynamical processes in complex networks has been a very active research field in the past decade. This has led to many advances in our understanding and ability to control a wide range of physical and social phenomena. Examples include adoption of cultural fads, the diffusion of beliefs, norms, and innovations in social networks [37, 75, 94, 99, 102], disease contagion in human and animal populations [57, 66, 72, 73, 80, 81], cascading failures in interdependent infrastructures [21, 93, 101], insolvency and default cascades in financial networks [50], and the spread of computer viruses or worms on the Web [8, 70].

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

Complex contagions have typically been studied in the literature using a linear threshold model. The original threshold model, proposed by Watts [94], considers binary-state dynamics
where each node is in one of the two states, inactive or active, and is initially assigned a threshold $\tau$ in $(0, 1]$. At any point in time, if an inactive node has $d$ neighbors of which $m$ are active, we determine if it will be activated by checking the relationship between $\frac{m}{d}$ and the pre-assigned threshold $\tau$. If $\frac{m}{d} \geq \tau$, then the node will turn active. Otherwise, if $\frac{m}{d} < \tau$, it stays inactive. Once a node is activated, it is assumed to remain active forever.

In the Watts threshold model, there are two important assumptions. First assumption is that all active individuals exhibit the same amount of influence on their neighbors. However, individuals’ standings on a given matter could vary significantly. For example, followers of a radical organization or a revolutionary movement may have varying levels of commitment to the cause, or varying desire and ability to recruit new members. To cope with the multi-state nature of individuals’ activity levels, Melnik et al. [58] introduced a multi-stage contagion model as a generalization of the Watts threshold model. There, nodes can be inactive or be in one of several active states with different levels of influence; e.g., active, hyper-active, etc.

The second assumption of the Watts model, which is also used in the multi-stage model by Melnik et al. [58], is that contagion is taking place over a single network where all edges have the same impact on spreading the influence. However, most real-world influence propagation events take place over multiple networks. For example, individuals may participate in multiple online social networks (e.g., Facebook, Twitter, etc.), and may have different levels of influence in each network. Similarly, within a single network, individuals may form different types of relationships (e.g., friendship, colleagueship, etc.), and each relationship type might have a different impact on the propagation of influence in a given context. For example, video games might be more likely to spread among high-school friends rather than parents, while the opposite might be true for political ideas. That is, if we do not distinguish different types of relationships, dynamics of influence propagation may not be accurately captured. Hence, it is natural to consider complex contagions over multiplex networks. With this motivation, [99] proposed and studied a threshold model in multiplex networks. However, [99] still used the first assumption mentioned above in the sense that their model is not suitable for multi-stage...
contagions (where nodes can belong to a rich set of states).

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

- For a class of random networks generated by the *colored* configuration model (see Section 4.2.1 for precise definitions), we analytically derive the expected size of *global* cascades; i.e., cases where a positive fraction of nodes (in the asymptotic limit) eventually becomes active or hyper-active when a randomly selected node is switched to the active state.

- We are the first to give analytic results for the probability of triggering global cascades for multi-stage models.

- We explore the intricate relationships between the structural properties of the underlying network and the impact of *hyper-active* nodes on the contagion dynamics. For instance, a particularly interesting scenario is when the hyper-active state is manifested in only one link type. This is motivated by the case where people may be more willing to express their viewpoints to close friends instead of office-mates, or can reach a hyper-influential state only in one social network (e.g., Twitter) versus another (e.g., Facebook). Our main finding is the interesting connection between the assortativity (i.e., the correlation among the degree of neighboring nodes) of the network and the impact of hyper-active nodes on cascade size. For instance, when the network is highly assortative (i.e., when high-degree nodes are likely to be connected with high-degree nodes), the influence exerted by the
hyper-active nodes has a much more significant impact on the cascade size as opposed to the case when the network has low assortativity (i.e., when the degrees of neighboring nodes tend to be uncorrelated). This impact is best observed when the cascade size is plotted as a function of the mean degree of the network. There, as the influence of hyper-active nodes increases, the highly assortative networks are shown to exhibit changes on not only the critical transition points (i.e., mean degree values at which expected cascade size changes from zero to a positive value, or vice versa), but also the number and order of transitions.

- Considering an important property in real-world networks, clustering, we analytically derive the expected size of global cascades for random networks with clustering. Due to the page limit, we include all derivations and results in supplementary materials. Supplementary materials are at https://goo.gl/1nTgQS.

The rest of the paper is organized as follows. In Section 4.2, we introduce the network and contagion models. Then, we describe the problem of interest and our main results in Section 4.3. In Section 4.4, we present numerical results that demonstrate the accuracy of our analysis in the finite node regime, and discuss the impact of hyper-influencers on complex contagions under different levels of assortativity. Further details about the impact of hyper-influencers are given in Appendix. We conclude the paper in Section 5.5 where we also suggest several directions for future work.

4.2 Model Definition: Networks and Dynamics

4.2.1 Multi-layer and multiplex network models

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

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

(4.1)

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

This network model can be interpreted in two different ways. The first one is a multilayer network where each network layer is generated by the widely used configuration model [16, 62, 71]; this case is illustrated in Figure 4.1(a). In particular, we use \( P(d_r) \) (resp. \( P(d_b) \)) as the *degree distribution* to determine the number of red (resp. blue) edges that will be assigned to each node in \( \mathcal{N}_r \) (resp. \( \mathcal{N}_b \)). Once the degree of each node is determined, we generate the networks \( \mathbb{R} \) and \( \mathbb{B} \) by selecting a graph uniformly at random from among all possible graphs that have the same degree sequence; see [62, 71] for more details. Next, we take a union of the edges in \( \mathbb{R} \) and \( \mathbb{B} \) to create a network \( \mathbb{H} \). Equivalently, we can consider a multiplex network model generated by the *colored configuration model* [87]. Let \( \mathbf{d} = (d_r, d_b) \) denote the colored degree of a node, where \( d_r \) and \( d_b \) stand for the number of red edges and blue edges incident on it. Each of the \( n \) nodes in the network is assigned a colored degree by independently drawing from the distribution \( P_\mathbf{d} \). Then, pairs of edges of the same color are randomly chosen and connected together until none is left; see [87] for details. Figure 4.1(b) is an illustration of this
Figure 4.1: Illustration of a multi-layer and a multiplex network representation of our model. In (a), we see a multi-layer network (e.g., a Physical communication layer and an online social network layer) with overlapping vertex sets; vertical dashed lines represent nodes corresponding to the same individual. In (b), we see the equivalent representation of this model by a multiplex network. Edges from the online network are red while edges from the physical network are blue.

4.2.2 Multi-stage content-dependent linear threshold model

We first introduce the single-stage content-dependent linear threshold model [99] which is a generalization of the vanilla threshold model [94]. In the content-dependent linear threshold model, links are classified into $r$ types. For a given content (a rumor, product, etc.), scalars $c_i, i = 1, \ldots, r$ represent the weight of type-$i$ edges on spreading this particular content. Nodes belong to either one of the two states, active or inactive, and each node is assigned a threshold $\tau$ in $(0, 1]$ drawn from a distribution $P(\tau)$. Given an inactive node with $m_i$ active and $d_i - m_i$ inactive neighbors for each link type-$i$, $i = 1, \ldots, r$, an inactive node will turn active if $\sum_{i=1}^{r} c_i m_i \geq \tau$. Namely, an inactive node with $\mathbf{m} = (m_1, \ldots, m_r)$ and $\mathbf{d} = (d_1, \ldots, d_r)$ will turn active with probability

$$F[\mathbf{m}, \mathbf{d}] \triangleq P \left[ \sum_{i=1}^{r} \frac{c_i m_i}{c_i d_i} \geq \tau \right]. \quad (4.2)$$

Throughout, $F[\mathbf{m}, \mathbf{d}]$ is referred to as the response function. If we do not distinguish the edge types or simply set $c_i = 1$ for all $i = 1, \ldots, r$, then this model reduces to the Watts’ threshold model [94]. The content-dependent threshold model enables modeling the case where people’s influence on others vary according to their relationship type, or the social network that they
are interacting through.

Different from the single-stage threshold model where nodes can be only in two states, the multi-stage linear threshold model \[58\] allows nodes to be in a richer set of active states. In this work, we assume that nodes can belong to three states, inactive, active, and hyper-active. In the following discussion, we use state-0, state-1, and state-2 to represent the inactive, active, and hyper-active state, respectively. Let \(\tau_1\) and \(\tau_2\) denote the thresholds associated with transitioning to the active and hyper-active states, respectively. The hyper-active individuals are assumed to be \(\beta\)-times more influential than active nodes in the propagation process (where \(\beta \geq 1\)). For example, an individual with \(d\) neighbors of which \(m_1\) are active and \(m_2\) are hyper-active, the probability of switching to state-\(i\) from the inactive state (i.e., state-0) is given by:

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

(4.3)

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

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

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

(4.4)

where \(m = (m_{r,1}, m_{r,2}, m_{b,1}, m_{b,2})\), \(d = (d_r, d_b)\), \(m_{r,1}\) and \(m_{r,2}\) (resp. \(m_{b,1}\) and \(m_{b,2}\)) denote the number of active and hyper-active neighbors connected through a red (resp. blue) edge, and \(d_r\) and \(d_b\) denote the number of red and blue neighbors, respectively. Assume that all nodes
are initially inactive and the contagion process starts by randomly choosing a node and setting it as active. The influence might then propagate in the network according to (4.4) and other nodes might turn active, and so on. More precisely, we consider a discrete-time process, which is an approximation to study the continuous-time Markov process model and has been applied in many works [43,96]. In discrete-time approaches, we discretize time into uniform time-steps of length $\Delta t = 1$. Then, nodes update their states synchronously at the discrete time steps $t = 0, 1, \ldots$, and an inactive node will be activated at time $t$ according to the influence they receive at time $t - 1$ and the formula (4.4). Once active, a node can not be deactivated. Since we are only interested in the fraction of nodes in active or hyper-active state in the steady-state (i.e., when nodes no longer change their states) our results do not depend on the length of the discrete time steps; i.e., we can use arbitrarily small time steps and get arbitrarily close to continuous-time dynamics.

4.3 Main Results

Since the proposed contagion process is monotone (i.e., an active node can never switch back to inactive), it will eventually stop, i.e., a steady-state will be reached. A global cascade is said to take place if the fraction of nodes that are activated is positive in the limit of large network sizes. Our main goals are i) determining the conditions (in terms of network parameters) for global cascades to be possible; ii) calculating the expected size of global cascades when they are possible; and iii) calculating the probability of triggering global cascades.
4.3.1 Expected cascade size and the condition to have a global cascade

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

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

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

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

which gives the expected size of global cascades when they exist. For simplicity, in our analysis we omit self-loops, i.e., the possibility of having more than one edge between two nodes. It is a simple matter to show that such self-loops occur very rarely in the construction of the configuration model and they have negligible impact on the cascade dynamics; e.g., see [69]. In fact, our experiments also confirm that the impact of this omission is negligible.

According to our definition, the expected cascade size stands for the final fraction of active and hyper-active individuals in the network. Therefore, we can compute it by computing the probability that an arbitrary node is active or hyper-active at the steady-state. We will compute this probability recursively using the “tree-approximation” approach [41, 58, 99], which is a mean-field treatment of the zero-temperature random-field Ising model on Bethe lattices [85]. The tree-approximation approach assumes that the network has a locally tree-like structure, which is valid under the configuration model considered here [71]. In supplementary materials, we show how the results can be extended to a broader class of networks that have large clustering coefficient (i.e., a large number of triangles). Clustering is a common property of many real-life social networks and capture the fact that a pair of individuals are more likely
to be connected with each other if they have a common friend as compared to the case where no common friend exists.

Labeling the tree structure from the bottom to the top, it is assumed that the node states are updated starting from the bottom, and continuing to the top, one level at a time. In other words, the nodes at level \( \ell \) will not update their states until the nodes at levels 0, 1, \ldots, \( \ell - 1 \) have finished updating. We define \( q_{r,1,\ell} \) (resp. \( q_{b,1,\ell} \)) as the probability that a node at level \( \ell \) who is connected to its only parent at level \( \ell + 1 \) by a red (resp. blue) edge turns active. Similarly, we define \( q_{rt,\ell} \) (resp. \( q_{bt,\ell} \)) as the probability that a node at level \( \ell \) that is attached to its only parent via a red (resp. blue) edge turns hyper-active. Given our assumption that nodes in the tree update their states one level at a time, these probabilities will be computed under the condition that the parent nodes at level \( \ell + 1 \) are inactive.

In the interest of brevity, we only explain the derivation of \( q_{r,1,\ell+1} \) in details. The derivations of \( q_{rt,1,\ell} \), \( q_{b,1,\ell+1} \), and \( q_{bt,1,\ell} \) can be explained very similarly. Since \( q_{r,1,\ell+1} \) cannot be expressed explicitly, we derive a recursive relation in terms of \( q_{r,1,\ell} \), \( q_{rt,\ell} \), \( q_{b,1,\ell} \), and \( q_{bt,\ell} \); see (4.6) - (4.9). The validity of the expression (4.6) for \( q_{r,1,\ell+1} \) can be explained as follows. Consider an inactive node at level \( \ell + 1 \) with colored degree \( d = (d_r, d_b) \) that is connected to its unique parent at level \( \ell + 2 \) via a red edge. The probability that this node has \( i \) active children connected via red edges, \( s \) active children connected via blue edges, \( j \) hyper-active children connected via red edges, and \( t \) hyper-active children connected via blue edges, and that it turns active is given by

\[
\binom{d_r - 1}{i} \binom{d_r - 1 - i}{j} q_{r,1,\ell}^i q_{rt,\ell}^j (1 - q_{r,1,\ell} - q_{rt,\ell})^{d_r - 1 - i - j} \\
\times \binom{d_b}{s} \binom{d_b - s}{t} q_{b,1,\ell}^s q_{bt,\ell}^t (1 - q_{b,1,\ell} - q_{bt,\ell})^{d_b - s - t} \\
\times F_1 [(i, j, s, t), d],
\]

where \( F_1 [(i, j, s, t), d] \) is as defined in (4.4); i.e., it denotes the probability that an inactive
node with a colored degree $d$ and a group of active and hyper-active neighbors for each color represented by $m = (i, j, s, t)$ switches to state-1. To simplify the notation, we use $F_1[(i, j, s, t), (x, y)]$ as defined at (4.10), so the term given in (4.5) becomes equivalent to $F_1[(i, j, s, t), (d_r - 1, d_b), \ell]$.

The intuition behind (4.5) is as follows. Since we assume that the network is tree-like, the state of each child node at level $\ell$ is independent from other children at the same level. Thus, we multiply together the probability of being at a specific state for each child node to get the whole expression (4.5) using a simple combinatorial argument. The reason behind using $d_r - 1$ rather than $d_r$ in (4.5) is the fact that the node under consideration is attached to its unique parent at level $\ell + 2$ through a red edge, and by assumption this parent node is inactive; recall that a node at level $\ell + 2$ can not update its state until all nodes in level $\ell + 1$ finish updating. A node that is known to have at least one red edge can be seen to have colored degree $d = (d_r, d_b)$ with probability $\frac{d_r p_d}{\langle d_r \rangle}$, e.g., see [71,99] for a discussion on the excess degree distribution. Finally, we get the detailed expressions of $q_{r,1,\ell+1}$ (4.6) after taking the expectation of (4.5) over the degree of the node at level $\ell + 1$. We can use similar arguments to derive expressions for $q_{rt1,\ell}$, $q_{b,1,\ell+1}$, and $q_{bt1,\ell}$. The expressions of all four probabilities are shown in (4.6) - (4.9).

\begin{align}
q_{r,1,\ell+1} &= \sum_{d} \frac{d_r p_d}{\langle d_r \rangle} \sum_{i=0}^{d_r-1} \sum_{j=0}^{d_r-1-i} \sum_{s=0}^{d_b} \sum_{t=0}^{d_b-s} F_1[(i, j, s, t), (d_r - 1, d_b), \ell] \quad (4.6) \\
q_{rt1,\ell} &= \sum_{d} \frac{d_r p_d}{\langle d_r \rangle} \sum_{i=0}^{d_r-1} \sum_{j=0}^{d_r-1-i} \sum_{s=0}^{d_b} \sum_{t=0}^{d_b-s} F_2[(i, j, s, t), (d_r - 1, d_b), \ell] \quad (4.7) \\
q_{b,1,\ell+1} &= \sum_{d} \frac{d_b p_d}{\langle d_b \rangle} \sum_{i=0}^{d_r} \sum_{j=0}^{d_r-i} \sum_{s=0}^{d_b-1} \sum_{t=0}^{d_b-1-s} F_1[(i, j, s, t), (d_r, d_b - 1), \ell] \quad (4.8) \\
q_{bt1,\ell} &= \sum_{d} \frac{d_b p_d}{\langle d_b \rangle} \sum_{i=0}^{d_r} \sum_{j=0}^{d_r-i} \sum_{s=0}^{d_b-1} \sum_{t=0}^{d_b-1-s} F_2[(i, j, s, t), (d_r, d_b - 1), \ell], \quad (4.9)
\end{align}
where for $k = 1, 2$, we define

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

(4.10)

Equations (4.6) - (4.9) form a non-linear system. Since our goal is to compute the expected size of global cascades \textit{given that they exist}, we can initialize this dynamical system with $q_{r,1,0}, q_{r,2,0}, q_{b,1,0}, q_{b,2,0} > 0$ to obtain the fixed points, $q_{r,1,\infty}, q_{r,t,\infty}, q_{b,1,\infty}, q_{b,t,\ell}$. These fixed points account for the probability of being in a corresponding state for the children of the node chosen uniformly at random. We can use them to calculate the expected size of global cascades.

We give the expected size of the cascades (given that they exist) in (5.5). The validity of (5.5) can be seen as follows: First, we randomly choose a node, whose colored degree is $d = (d_r, d_b)$, with probability $p_d$. The probability that each of its $d_r$ neighbors (via red links) is active (resp. hyper-active) is given by $q_{r,1,\infty}$ (resp. $q_{r,t,\infty}$). Similarly, each of the $d_b$ neighbors (connected via blue links) of this randomly chosen node is active with probability $q_{b,1,\infty}$ and hyper-active with probability $q_{b,t,\ell}$, independently from each other. Then, with each possible combination of numbers of active and hyper-active neighbors, we can calculate the probability of being active or hyper-active for the node by the response function (4.4). Taking the expectation with respect to the degree $d$ yields (5.5). As discussed in details in [41,58,99], this method, based on the \textit{tree-approximation} technique, gives precise results in the asymptotic limit $n \to \infty$, when the underlying network is generated according to the configuration model.

We present extensive numerical studies in Section 4.4 that supports our results even in the finite node regime.
\[\lim_{n \to \infty} \mathbb{E}[S \mid S > 0] = \sum_{d} p_d \sum_{i=0}^{d_r} \sum_{j=0}^{d_i} \sum_{s=0}^{d_h} \sum_{t=0}^{d_b-s} \{ \mathbb{F}_1 [(i, j, s, t), (d_r, d_h), \infty] + \mathbb{F}_2 [(i, j, s, t), (d_r, d_h), \infty] \} . \]

(4.11)

From the recursive equations derived above, we can also obtain the conditions needed for the global cascades to be possible; i.e., conditions under which \( S > 0 \) with a positive probability in the limit \( n \to \infty \). For notational convenience, we define \( q_1 := q_{r,1,\infty} \), \( q_2 := q_{rt,\infty} \), \( q_3 := q_{b,1,\infty} \), and \( q_4 := q_{bd,\ell} \). Then, the four recursive equations (4.6) - (4.9) take the form

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

(4.12)

By direct inspection, we see that the recursive equations (5.13) have a trivial fixed point \( q_1 = q_2 = q_3 = q_4 = 0 \), which yields \( S = 0 \) almost surely; this can be seen from the fact that in that case we have \( \mathbb{E}[S] = 0 \). In other words, when (5.13) has only a single solution given by this trivial fixed point, then with probability one global cascades do not take place. In general, the trivial fixed point may not be stable and there may exist non-trivial fixed points that yield \( \mathbb{E}[S] > 0 \). In that case, we have \( S > 0 \) with a positive probability in the limit \( n \to \infty \) and hence global cascade may take place. To check the existence of non-trivial solutions of (4.6) - (4.9), we linearize them at \( q_1 = q_2 = q_3 = q_4 = 0 \) which yields the Jacobian matrix \( J \) given as

\[J = \left. \frac{\partial f_i(q_1, q_2, q_3, q_4)}{\partial q_j} \right|_{q_1 = q_2 = q_3 = q_4 = 0} .\]

(4.13)

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

### 4.3.2 Probability of triggering a global cascade

We now turn our attention to computing the probability $\mathbb{P}[S > 0]$ of global cascades. As discussed in [94,99], the possibility of a seed node to trigger a global cascade is closely tied to the size of (and the seed node’s connectivity to) the set of vulnerable nodes in the network; a node is deemed vulnerable if it can be activated by only one active neighbor [37,39,94,99]. The definition of vulnerable nodes and of the vulnerable component has been extended in [99] to the case of multiplex networks. There, a “vulnerable component” is defined as a set of nodes, each of which is vulnerable w.r.t. at least one of the link types, such that in the subgraph containing this set of nodes, activating any node leads to the activation of all nodes in the set. A multiplex network is said to contain a giant vulnerable component (GVC) if the fraction of nodes in its largest vulnerable component is positive in the limit $n \to \infty$. These definitions were then used [99] to demonstrate that an initial node can trigger a global cascade if and only if it belongs to the extended giant vulnerable cluster (EGVC), that contains nodes in the GVC and nodes whose activation leads to activation of a node in GVC. Put differently, the probability of a randomly selected node triggering a global cascade is equal to the fractional size of the EGVC; see [99] for details.

Here, we use the ideas mentioned above to calculate the probability of global cascades, or equivalently the fraction of nodes that are in EGVC. This will be done through the analysis of a branching process that starts from a randomly selected and activated initial node, and keeps exploring the neighboring nodes that are activated according to nodes’ response function (4.4). The branching process will continue by exploring the neighbors of the newly activated nodes that will also be activated, and so on.

The fraction of nodes identified by the branching process described above can be analyzed using the method of probability generating functions [97]; e.g., see [66, 71, 94, 99, 106] where this tool was demonstrated to be useful for similar purposes. The first generating function we
use in our analysis is $G(x)$, and it generates the probability distribution of “the finite number of nodes reached and influenced by the above branching process”; different from [102,105,106], we exclude the initially activated nodes. We have

$$G(x) = \sum_{d} p_d g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b},$$

where $g_{r,1}(x)$ (resp. $g_{b,1}(x)$) generates the probability distribution of “the finite number of nodes reached and influenced by following a randomly chosen red (resp. blue) edge one of whose ends is set to active.” The difference between $G(x)$ and $g_{r,1}(x)$ is illustrated in Figure 4.2. The validity of expression (4.14) can be seen as follows. First, we initially activate a node which is chosen uniformly at random. The probability that this node has a degree $d = (d_r, d_b)$ is $p_d$. In that case, the number of nodes that are reached and activated by this node will be generated (in view of the powers property of the generating functions) by $g_{r,1}(x)^{d_r} g_{b,1}(x)^{d_b}$. Summing over all possible degrees $d$ of the initial node leads to (4.14).

For (4.14) to be useful, we also need to derive expressions for $g_{r,1}(x)$ and $g_{b,1}(x)$. This will be done by the help of two more generating functions. Namely, let $g_{r,2}(x)$ (resp. $g_{b,2}(x)$) generate the distribution of “the finite number of nodes reached and influenced by following a red (resp. blue) edge whose one end is connected to a hyper-active node.” The detailed
expressions of the four generating functions are given in (4.15) - (4.18). Here, we only explain the derivation \( g_{r,1}(x) \), as others can be explained in a similar manner.

To see why (4.15) holds, first note that as the randomly selected red edge whose one end is connected to an active node is followed, we will find a node with colored degree \( \mathbf{d} = (d_r,d_b) \) with probability \( \frac{d_r \rho_d}{\langle d_r \rangle} \) as already explained in the derivation of (4.6). There are three possible cases for this node with degree \( \mathbf{d} \):

- It turns active, i.e., \( \tau_1 \leq \frac{c}{cd_r + d_b} < \tau_2 \), which happens with probability \( F_1[(1,0,0,0),(d_r,d_b)] \). Then, this newly activated node will activate \( g_{r,1}(x)^{d_r-1}g_{b,1}(x)^{d_b} \) other nodes based on the powers property of generating functions. The reason why we use \( d_r - 1 \) instead of \( d_r \) is because one of its \( d_r \) edges has already been considered as its connection to the active end.

- It turns hyper-active, i.e., \( \frac{c}{cd_r + d_b} \geq \tau_2 \), which happens with probability \( F_2[(1,0,0,0),(d_r,d_b)] \). Then, the number of nodes reached and influenced by this newly activated node will be generated by \( g_{r,2}(x)^{d_r-1}g_{b,2}(x)^{d_b} \); this can be seen via similar arguments to the case above.

- It remains inactive, i.e., \( \frac{c}{cd_r + d_b} < \tau_1 \), which happens with probability \( 1 - F_1[(1,0,0,0),(d_r,d_b)] - F_2[(1,0,0,0),(d_r,d_b)] \). Then, there will be no newly activated nodes.

Combining these three cases and summing over all possible \( \mathbf{d} \), we get (4.15), where the explicit factor \( x \) accounts for the initial node that is activated. The expressions for \( g_{r,2}(x), g_{b,1}(x) \), and \( g_{b,2}(x) \) can be derived similarly.
Simplifying the notation as

This approach has been introduced in [94] and used in [71, 99] for similar calculations.

underlying influence propagation process will constitute a global cascade. Thus, we have

survive probability, unless and activated by this branching process, we should have

G triggered in the following manner. Since G

number of nodes. In other words, 1

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

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

This approach has been introduced in [94] and used in [71, 99] for similar calculations.

In order to calculate G(1), we now solve for the fixed point of (4.15) - (4.18) at x = 1.
Simplifying the notation as \(g_1 := g_{r,1}(1)\), \(g_2 := g_{r,2}(1)\), \(g_3 := g_{b,1}(1)\), and \(g_4 := g_{b,2}(1)\), the
recursive equations (4.15) - (4.18) at \( x = 1 \) can be expressed as

\[
g_i = h_i(g_1, g_2, g_3, g_4), \quad i = 1, 2, 3, 4.
\] (4.20)

Here, the exact form of the functions \( h_1(g_1, g_2, g_3, g_4), \ldots, h_4(g_1, g_2, g_3, g_4) \) will be obtained from (4.15) - (4.18). Once the fixed points of (4.20) are obtained, we get from (4.14) that

\[
G(1) = \sum_d p_d g_1^{d_r} g_3^{d_b}.
\] (4.21)

In view of (4.19), we finally obtain the desired probability of global cascades as

\[
\lim_{n \to \infty} \mathbb{P}[S > 0] = 1 - \sum_d p_d g_1^{d_r} g_3^{d_b}.
\] (4.22)

### 4.4 Numerical Results

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

#### 4.4.1 The agreement between our analysis and simulations

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

\begin{align}
\hat{p}_k^b &= e^{-\lambda_b} \frac{(\lambda_b)^k}{k!}, \quad k = 0, 1, \ldots, \\
\hat{p}_k^r &= \alpha e^{-\lambda_r} \frac{(\lambda_r)^k}{k!} + (1 - \alpha) \delta_{k,0}, \quad k = 0, 1, \ldots.
\end{align}

(4.23) \hspace{1cm} (4.24)

Here, $\lambda_r$ (resp. $\lambda_b$) denotes the mean number of red (resp. blue) edges assigned per node, $\alpha$ denotes the fraction of nodes that have red edges (i.e., the relative size of the red network $R$), and $\delta$ denotes the Kronecker delta. In our simulations to verify our analysis on the expected size, we use $n = 1 \times 10^6$ nodes\(^1\) to create networks and set $\alpha = 0.5$. Besides, we use $c = 0.5$ and $\beta = 1.5$ as the content parameter and the weight of hyper-active nodes, respectively, and fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$. Then, for several values of $\lambda_r = \lambda_b$, we run 1,000 independent experiments (for each parameter set), each time computing the fraction of nodes that eventually turn active or hyper-active. The results are depicted in Figure 4.3(a) where lines represent analytic results obtained from (4.6) - (4.9), and symbols represent the average cascade size obtained in simulations (over 1,000 experiments for each data point). We see that there is a good agreement between the analytic results and the simulations.

Next, to check the correctness of our analysis on the probability, we fix all parameters except increasing the number of experiments from 1,000 to 10,000. As shown in Figure 4.3(b), we observe that our analysis on the probability (4.15) - (4.18) also match very well the simulations results. This indicates that, although asymptotic in nature, the results presented in Section 4.3 are still helpful in understanding complex contagion dynamics (e.g., the probability and expected size of global cascades) in finite networks.

In addition, we observe from both Figure 4.3(a) and 4.3(b) that the contagion exhibits two phase transitions, i.e., two different $\lambda_r = \lambda_b$ values around which fractional cascade size transitions from zero to a positive value, or vice verse. These points are of great interest since they provide insights on how network connectivity affects the possibility of observing global

\(^1\)To avoid the finite size effect, we use $n = 2 \times 10^6$ as the number of nodes around the second phase transition in the simulations.
Figure 4.3: Simulations for doubly Poisson degree distributions, $\alpha = 0.5$, $\tau_1 = 0.18$, and $\tau_2 = 0.32$. The weight of hyper-influencers is taken to be $\beta = 1.5$.

influence spreading events. The first transition occurs around low values of $\lambda$, and reflects the fact that global spreading events become possible only after the network reaches a certain level of connectivity. The second phase transition occurs around high $\lambda$ values, indicating that global cascades can not occur when nodes are locally stable; i.e., when they have a large number of friends, individuals tend to be difficult to get influenced by a few active neighbors.

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

### 4.4.2 The impact of hyper-influencers in multiplex networks

In this section, we investigate more closely how hyper-influencers affect the complex contagions. We consider a case where hyper-active nodes are restricted to appear only through one type of edges, red or blue, rather than allowing them to exert additional influence through both types of edges. This setting is motivated by cases where people can reach a more active/influential state only in one network, or one relationship type. For example, some people may be reluctant to express their opinions freely in person (e.g., physical networks), but may be much more active on online networks (e.g., Twitter) due to anonymity. This raises an interesting question: which network or edge type would facilitate the influence propagation process most when hyper-influencers are allowed there. In what follows, we conduct several
experiments to answer this question: 1) we only allow hyper-activity in red edges, i.e., hyper-active neighbors connected by blue edges will be counted as merely active when checking the response function; 2) we only allow hyper-activity in blue edges. As discussed in some previous studies \[46, 105, 106\], assortativity is one of the most important structural properties on multiplex networks. Assortativity is defined as the Pearson correlation coefficient between the degree of nodes that are connected by a link \[65\]. If a network is assortative, then nodes of high degree in the network tend to attach to high degree nodes; it was noted in \[65\] that social networks tend to have high assortativity. Therefore, it is interesting to see if assortativity has any impact on the answer to the above question.

In the following experiments, we conduct these experiments on a network with low assortativity and then a network with high assortativity. We use the degree distributions (4.23) and (4.24) to assign red and blue degrees. To be able to control the assortativity of networks without changing the first moment of degree, we set \(\alpha \lambda_r = \lambda_b\) rather than \(\lambda_r = \lambda_b\). With this setting, when \(\alpha\) is large, e.g., 0.99, nearly all of the nodes will have a similar number of red and blue edges, which leads to networks with limited assortativity. On the contrary, when \(\alpha\) is low, e.g., 0.1, only 10% of the nodes will have extra red edges. In addition, these nodes will have a significantly larger number of edges, since \(\lambda_r\) is ten times larger than \(\lambda_b\). The nodes with extra red edges will tend to be connected together, which results in the network to have high assortativity. A more detailed discussion on this can be found in \[105\].

We start with the limited assortativity case, i.e., \(\alpha = 0.99\). As shown in Figure 4.4, we observe that regardless of which network hyper-influencers are constrained to exist, there are two phase transitions as in the case of single-stage complex contagions. However, we see that the existence of hyper-influencers delays the second phase transitions to higher mean degrees. The reason behind this delay can be explained as follows. As mentioned before, the second phase transition occurs due to high local stability of nodes making their states hard to change by only few active neighbors. However, hyper-influencers help increase the value of the perceived influence, i.e., \(\frac{c(m_{r,1} + \beta m_{r,2}) + m_{b,1} + \beta m_{b,2}}{cd_r + d_b}\), so that the response function could be
exceeded even with few active and hyper-active neighbors, in the high mean degree region. Besides, allowing hyper-activity in blue edges leads to a larger region where global cascades take place, in comparison with the case where hyper-activity exists only in red edges. This can be explained as follows. When \( \alpha = 0.99 \), there are more nodes connected by blue edges in the network than red edges. That is, the impact of blue edges on impeding global cascades is more than that of red edges. Thus, allowing hyper-influence to be exerted in blue edges delays the second phase transition further.

Next, we discuss the case where \( \alpha = 0.1 \) that leads to a highly assortative network [105]. In Figure 4.5, we present numerical results for the first setting where the hyper-active state is manifested in only red edges. When \( \beta = 1 \), i.e., when there are no hyper-influencers in the network, four phase transitions take place. However, if we increase \( \beta \) from one to three, then only two phase transitions are observed. This can be explained as follows. When \( \beta = 1 \), multi-stage complex contagions is reduced to single-stage complex contagions, in which case four phase transitions might occur when assortativity is high [105]. As explained in [105], the first pair of phase transitions are mainly due to the red edges. When \( \lambda_b \) is small, there are too few blue edges to trigger a global cascade. However, since we have \( \lambda_r = 10\lambda_b \), there are still enough red edges to have global cascades. As we increase \( \lambda_b \), we observe a parameter interval where red edges are too many while blue edges are too few to have a global cascade. If we keep increasing \( \lambda_b \) further, global cascades start appearing again when the network has enough connectivity in blue edges to propagate the influence. However, further increasing in \( \lambda_b \) leads to high local stability of nodes w.r.t. both blue and red edges and global cascades become impossible again. A more detailed discussion can be found in [105].

The reason why increasing \( \beta \) changes the number of phase transitions is as follows. From the definition (4.4) of the response function, we observe that it is monotonically increasing with respect to \( \beta \). Thus, when \( \beta \) is higher, an inactive node is easier to be activated by a hyper-active node, which makes it possible to have global cascades at higher levels of connectivity; i.e., the second phase transition tends to appear at larger \( \lambda \). This leads to the second and the
Figure 4.4: Hyper-activity only appears in either red or blue edges. We fix $\tau_1 = 0.18$ and $\tau_2 = 0.32$, and vary the mean degree. When $\alpha = 0.99$, the assortativity is negligible.

third phase transitions seen in Figure 4.5 disappear when $\beta = 1$; i.e., the interval where we have too many red and too few blue edges disappears.

Next, we focus on the second setting where hyper-activity is only manifested in blue edges. The results are shown in Figure 4.6. Allowing hyper-activity in blue edges does not change the connectivity of the network, so the first and the second phase transitions caused by the connectivity w.r.t. red edges remain the same. However, the gap between the second and third transitions still exists. The gap happens between the second transition w.r.t. red and the first transition w.r.t. blue edges. A high $\beta$ only shifts the second transition to the right but does not affect the first transition much. Thus, the gap disappears quickly with increasing $\beta$ when we allow it in red edges, but remains when we only allow it in blue edges. Besides, compared with the case $\beta = 1$, the fourth transition is significantly delayed when $\beta = 3$. The reason behind the delay of the fourth phase transition is similar to the previous discussion: A higher $\beta$ makes it easier to exceed the threshold even when the degree parameter is at a high level, so the original fourth phase transition has been extended to a larger mean degree.

From these experiments, we conclude that depending on the assortativity of the network, the impact of hyper-activity in red or blue edges on complex contagions are different: when the network is highly assortative, the additional influence exerted by the hyper-active nodes may change not only the critical transition points, but also the number and order of phase transitions, while for networks that have little or no assortativity, the additional influence mainly enlarge global cascade regions.
Figure 4.5: Hyper-activity only appears in red edges. We fix \( \tau_1 = 0.18 \) and \( \tau_2 = 0.32 \), and vary the mean degree. When \( \alpha = 0.1 \), the assortativity of the network is around 0.8.

4.5 Conclusion and Future Work

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

Figure 4.6: Hyper-activity only appears in blue edges. We fix \( \tau_1 = 0.18 \) and \( \tau_2 = 0.32 \), then vary the mean degree. When \( \alpha = 0.1 \), the assortativity is high (be up to 0.8).
added analysis of multi-stage complex contagions over networks with high clustering coefficient, which is an important property of real-world networks.

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

4.6 The impact of hyper-influencers on the global cascade boundary

In this section, we investigate how the parameters $\beta, \tau_1, \tau_2$ of the contagion model and the connectivity of the network jointly affect the possibility of global cascades. In particular, we will determine the boundaries in the space of parameters that separate the region where global cascades are possible (i.e., $P[S > 0]$) from the region where global cascades do not take place almost surely (i.e., $P[S = 0]$). First, we will focus on the impact of the weight $\beta$ of hyper-influencers on the global cascade boundary, and then move on to the discussion about the impact of the threshold $\tau_1$ of ordinary influencers.

Figure 4.7 shows the global cascade boundary in the space of $\tau_2$ and degree parameter $\lambda = \lambda_r = \lambda_b$, for several values of $\beta$. We observe that larger $\beta$ values lead to a larger region of parameters $\tau_2, \lambda$ for which global cascades can take place; i.e., the global cascade region gets larger with increasing $\beta$. An interesting observation is that the cascade boundary is more sensitive to the changes in $\beta$ values when $\lambda$ is large; i.e., the lower parts of the boundaries seen in Figure 4.7 are less dependent on the choice of $\beta$ as compared to the upper parts. This can be explained as follows. When $\lambda$ is small, the existence of global cascades (and hence the cascade boundary) is mainly determined by whether the network has enough connectivity to spread the influence. However, increasing $\beta$ does not change the connectivity of the network,
and hence does not affect the boundary when $\lambda$ is low. Differently, when $\lambda$ high, the location of the boundary (i.e., the second phase transition points seen in Figures 4.3(a)-4.3(b)) is decided by the likelihood of nodes with high degree being influenced by a single active or hyper-active neighbor. Thus, the boundary is determined from a node’s perceived influence, or perceived proportion of active and hyper-active neighbors, given at (4.4), and on how this compares with the activation thresholds $\tau_1$ and $\tau_2$. From (4.4), we see that higher $\beta$ leads to an increased perceived influence for a node that has at least one hyper-active neighbor, making it possible for the activation threshold to be exceeded at higher $d_r, d_h$ values (equivalently at higher $\lambda$ values). Thus, when $\lambda$ is high, the boundary tends to be more sensitive to the changes in $\beta$.

Next, we investigate the impact of the activation threshold $\tau_1$ on the global cascade boundary (again considering the space of $\lambda - \tau_2$). In Figure 4.8, we fix $\beta = 1.5$ and plot the boundary on the $\tau_2 - \lambda$ plane that separates the regions where cascades are possible and not possible, respectively. This is done for three different values of $\tau_1$. We observe that the impact of $\tau_1$ (i.e., the threshold on the perceived influence that an inactive node needs to receive in order to turn active) on the cascade boundary is opposite to that of $\beta$. That is, the higher $\tau_1$ is, the smaller is the region where global cascades are possible. The reason behind this observation is as follows. From the expression of the response function (4.4), we see that it is decreasing with increasing $\tau_1$. In other words, a higher $\tau_1$ makes it harder for nodes to become active (i.e., influenced), leading to a smaller cascade region.

![Figure 4.7](image_url)

Figure 4.7: Given $\tau_1 = 0.15$ and $\alpha = 0.5$, we vary the mean degree $\lambda$ and $\tau_2$ to plot the global cascade region for several $\beta$, 1.5, 2.0, and 3.0.
Figure 4.8: Given the weight of extra influence $\beta = 1.5$ and $\alpha = 0.5$, we vary the degree parameter $\lambda = \lambda_b = \lambda_r$ and $\tau_2$ to plot the region where there exists a global cascade for several $\tau_1$, 0.15, 0.18, and 0.2. Both of the edges are assigned by the doubly Poisson distribution in Section 4.4.1.

Figure 4.9: The comparison between different $\beta$ for the probability of triggering a *global* cascade.

### 4.7 The impact of hyper-influencers on the probability and expected size of global cascades

We start by investigating the impact of the extra influence $\beta$ (that hyper-active nodes exert on their neighbors) on the probability of global cascades. From Figure 4.9, we observe that a larger $\beta$ will increase the probability of triggering a global cascade. This observation is intuitive given that the response function (4.4) is increasing with respect to $\beta$. Thus, with a higher $\beta$, the perceived influence from a single active or hyper-active neighbor exceeds the threshold more easily, leading to a larger *vulnerable* component. Also, we see in Figure 4.9 that when the degree parameter is *large*, the cascade probability becomes more sensitive to the changes in $\beta$. This is consistent with the observations from Figure 4.8 and can be explained in a similar manner.
Next, we discuss the impact of hyper-influencers on the expected size of global cascades. From Figure 4.10, we observe that increasing $\beta$ leads to an expansion of the interval of $\lambda_r = \lambda_b$ values for which expected cascade size is positive. However, over the common interval where cascade size is positive, we see that increasing $\beta$ nearly does not lead to changes in the expected cascade size. The reason behind this observation is that the expected cascade size is mainly determined by the connectivity, (e.g., the mean degree) of the network, which remains invariant to changes in $\beta$. Thus, increasing $\beta$ nearly does not change the expected size of global cascades. The expansion of the interval over which $S > 0$ with increasing $\beta$ is explained by the response function (4.4) being increasing in $\beta$. In other words, a higher $\beta$ makes it easier for the perceived influence to exceed the activation threshold, helping global cascades take place even at higher mean degree.

4.8 The Impact of the network size on the comparison between analytic results and experiments

In this section, we explore the impact of network size on the comparison of analytic results and simulation experiments. In particular, we repeat the experiments conducted to obtain Figure 4.3(a) and compare the asymptotic and the finite-size results for $n = 2 \times 10^2, 2 \times 10^4, 2 \times 10^6$.

The results are shown in Figure 4.11, where we see that the differences between the asymptotic and the finite-size results are not significant when the mean degree is at a low level.

Figure 4.10: The comparison between different $\beta$ for the expected global cascade size.
However, when the mean degree is at a high level, simulation results are significantly different from the asymptotic results when the network size is small. In particular, when $n = 200$, the second phase transition observed at large mean degrees appears much later than suggested by the analysis. This transition also appears like a continuous one in the simulations with $n = 200$ although it is clearly a discontinuous transition in the analytic results. When $n = 20,000$ we see that simulation results match the analysis a lot better than the case with $n = 200$ although the match is not nearly as perfect as the one seen when $n = 2,000,000$. In particular, we see significant discrepancies around the second phase transition point, though it is at least suggested correctly by the simulations that the transition is discontinuous. Overall, we can conclude that our asymptotic analysis yields almost perfect predictions for networks with a million nodes or more, and relatively accurate predictions for networks with 20,000-100,000 nodes. For smaller networks with only a few hundred nodes, the asymptotic results do not yield reliable predictions.

Figure 4.11: The comparison between the asymptotic and the finite-size results for different size of networks.
Chapter 5

A Vector Threshold Model for the Simultaneous Spread of Correlated Influence

5.1 Introduction

In recent decades, mathematical modeling of propagation processes over networks have been studied in a wide range of contexts including cascading failures [101, 103], epidemics and social contagions [74, 75, 102, 105, 106], systemic risk in banking networks [51], to name a few. Contagion processes are typically modeled and studied under two different categories referred to as simple and complex contagions, respectively [25]. Simple contagion models are used for cases where a single source of exposure is enough for an individual to get infected and to start spreading the content to their contacts; e.g., news articles, disease spreading, etc. On the other hand, complex contagion models are used in cases where social reinforcement plays a key role in the spreading process. In other words, complex contagion models are used when multiple sources of exposure to a content (e.g., an opinion, a product, a political view, etc.) are needed for individuals to change their action or state. For example, an individual may not adopt a new behavior (or, change their opinion) after seeing only one friend doing so, but the situation might change if the ratio of their contacts doing it exceeds a certain level.

This paper focuses on complex contagions for which several models have been proposed in the literature. Perhaps the most widely known among them is the linear threshold model proposed by Watts in [94]. In this model, each node belongs to one of two states, inactive or active, and has a threshold $\tau$ in $(0, 1]$ which is drawn from a distribution $P(\tau)$. This threshold
indicates the required fraction of active neighbors for an inactive node to turn active. Starting from a state where all nodes are initially inactive, a small number of nodes are chosen uniformly at random and made active. Then, an inactive node with degree $d$ of which $m$ are active will get activated with probability

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

(5.1)

where $F[m, d]$ is referred to as the response function. This model allows studying the propagation of binary influence (where individuals are either active or inactive) over monoplex networks (where there exists only one link type). Under these assumptions, this model overly simplifies the dynamics for link relationships and possible states of each node. With this motivation several works have proposed and studied models that involve a richer set of node states and link types. For instance, Yağan and Gligor [99] extended this vanilla threshold model to multiplex networks in order to incorporate the fact that there may exist more than one type of edges in the network; e.g., friendship, colleagueship, family, etc. With this observation, they proposed a content-dependent multiplex threshold model. In this model, the network consists of $r$ different link types with each link type having certain influence weight that varies by the content that is being spread. For instance, their models make it possible to capture the fact that video games might be more likely to spread among high-school friends rather than parents, while the opposite might be true for political ideas. From a different perspective, Melnik et al. [58] extended the vanilla linear threshold model to multi-stage (i.e., non-binary) influence propagation. In their model, nodes can belong to a richer a set of states, i.e., inactive, active, hyper-active, etc.

The common aspect of [99] of [58] is the assumption that there is only one content being spread in network. Alternatively, Borodin et al. [17] proposed a threshold model for competitive influence. If there are two contents, $A$ and $B$, spreading over networks, then nodes satisfying the condition of supporting $A$ (resp. $B$) will turn to the state of supporting $A$ (resp. $B$). If nodes satisfy both of the conditions for supporting $A$ and $B$, then the nodes will randomly
choose a content to support. Although there are multiple contents in the propagation process, there is limited correlation between the contents arising from the specific constraint that nodes can not support A and B simultaneously. Aside from this, the spreading of one content has little impact on the spreading of the other. Additionally, even though some researchers studied the spread of correlated contents (e.g., competition, cooperation, etc.), they focused on simple contagions rather than complex contagions [22].

The discussion given above indicates that most works on complex contagions studied the spreading of either a single content (e.g., opinion, rumor, product, political view, etc.), or multiple contents spreading simultaneously but with limited correlation between each other. However, real-life influence propagation processes can have multiple contents spreading simultaneously with positive or negative correlation, and individuals might support more than one content at the same time. For example, the spread of the purchase behaviors of different products from the same company might exhibit positive correlation; e.g., an individual who already purchased an iPhone might be more easily influenced by their friends to buy the Apple HomePod. In contrast, one’s opinions on universal health care and proposed tax relief for “wealthy” individuals would be expected to have negative correlation.

With this motivation, this work aims to initiate a study on complex contagions where multiple contents spread simultaneously in a correlated manner. To this end, we propose an extension of the Watts threshold model [94] allowing correlated contents to spread at the same time, referred to as the vector threshold model; see Section 5.2.2 for description of the model. Under this model, we derive analytic results for the expected size of global cascades, i.e., cases where a randomly chosen node can initiate a propagation that eventually reaches a positive fraction of the whole population. We also derive conditions on the contagion parameters under which global cascades take place with positive probability. Our analytic results are supported by a numerical study through which we demonstrate how the correlations between spreading contents affect the expected size of global cascades. In particular, we find that when the mean degree is low, correlations have limited impact on the size of global cascades. In contrast,
we show that different levels and types of correlations could result in global cascades with significantly different size when the mean degree is high.

The rest of the paper is organized as follows. In Section 5.2, we introduce the network model, the proposed vector threshold model, and the problem of interests. In Section 5.3, we give the derivations of the expected size of global cascades and the condition for the existence of global cascades. Then, in Section 5.4, we use numerical experiments to support our analysis and discuss the impact of correlations among contents on the expected size of global cascades. In Section 5.5, we conclude our work and provide a brief discussion on future directions.

5.2 Model and Problem Definition

5.2.1 The network model: The configuration model

The degree sequence or the degree distribution is one of the most important parameters of networks. To be able to incorporate arbitrary degree distributions, we use the configuration model to construct networks. The configuration model is a widely used reference model for real-world networks, and is also analytically tractable [62,69,71]. Let \( N = \{1, 2, \ldots, n\} \) denote the vertex set where \( n \) is the number of vertices. For each node in \( N \), its degree is assigned by a prescribed degree distribution \( P(d) \) where \( d \) is the random variable for the degree. After the assignment, each node has the same number of half-edges as its degree. A half-edge accounts for an edge one end of which is connected to the previously assigned vertex while the other end of which is free to connect any other half-edges. Then, we randomly choose two half-edges to form an edge connecting two vertices until no half-edges are left. This model requires that the sum of the degrees is even. With the above process, the configuration model could generate networks by selecting a graph uniformly at random among all possible graphs that have the same degree distribution; see [62,71] for more details. In addition, self-loops or multi-edges could be omitted in the limit of large network size [69], which simplifies the analysis in Section 5.3.1.
5.2.2 The vector threshold model: A threshold model with multiple correlated contents

As mentioned in Section 5.1, there are many works that model the influence propagation in networks. However, few of them consider the simultaneous spread of multiple correlated contents. Therefore, we introduce a general threshold model, which is called the vector threshold model. Without loss of generality, we assume that there are two correlated contents or opinions which are denoted by Content-1 and Content-2, respectively, and that nodes could have two choices for each of the contents, support or not support. Then, with two contents spreading in a network, there exist four possible states for each node: 0) not supporting both of the contents; 1) only supporting Content-1; 2) only supporting Content-2; and 3) supporting both of the contents. For notational convenience, we use state-i to indicate the i-th state where $i = 0, 1, 2, 3$, and name state-0 as inactive while the other states as active. In addition, once a node gets activated, it will not change its state. Although we assume there are two correlated contents, the arguments could be extended to an arbitrary number of correlated contents.

In an influence propagation process, each inactive node keeps receiving influence from its contacts. Since its contacts could belong to different states, we can classify the influence by the state of its neighbors into different categories, and represent the received influence in a vector. In our model, the influence from contacts is measured by the “perceived” proportion [44], i.e., the fraction of contacts in specific states. Then, there are two types of influence, one from supporting Content-1 while the other from supporting Content-2. Because state-0 neighbors do not exert any influence and state-3 neighbors support both of the contents, the perceived proportion from state-i is defined as

$$\text{Proportion}_i = \frac{\# \text{ of neighbors in state-i and state-3}}{\# \text{ of neighbors}},$$

$i = 1 \text{ or } 2$.  

(5.2)
Therefore, each inactive node in the vector threshold model could receive a vector of perceived proportions, \( \mathbf{v} = [\text{Proportion}_1, \text{Proportion}_2] \). Based on the received vector of an inactive node, we could determine which state the inactive node will turn to by a response function (5.3), i.e., the probability of changing from the inactive state to the other active states. Under the vector threshold model, given an inactive node with degree \( d \) of which \( m_1, m_2, \) and \( m_3 \) neighbors for state-1, 2, and 3, respectively, the probability it turns to state-\( i \) is given by

\[
F_i[\mathbf{m}, d] \triangleq P[(\text{Proportion}_1, \text{Proportion}_2) \in \text{Space}_i],
\]

where \( \mathbf{m} = (m_1, m_2, m_3) \), \( \text{Proportion}_i \) accounts for the perceived proportion of state-\( i \) neighbors (5.2), and \( \text{Space}_i \) means the parameter space of being in state-\( i \). The parameter space is formed by the perceived proportions, i.e., \([\text{Proportion}_1, \text{Proportion}_2]\), and is illustrated in Figure 5.1. By the flexibility of the design of the parameter space, this model offers a general solution for different correlations among contents.

In general, there are three widely recognized correlations among contents: independent, positive, and negative correlation. We could split the parameter space accordingly to achieve these correlations. First, the independent correlation means the spread of contents is independent from each other. For example, the spread of newly released video games is usually independent from the spread of new norms. This correlation could be achieved by dividing the parameter space as Figure 5.1(a). In Figure 5.1(a), the threshold required by supporting Content-2 (resp. Content-1) is not affected by the proportion of state-1 contacts, i.e., \( \text{Proportion}_1 \) (resp. \( \text{Proportion}_2 \)). Second, the positive correlation could be interpreted as a spreading content would help the spread of other contents. For instance, the spread of opinions on same-sex marriage and gun control could be a positive correlation. If we support same-sex marriage, we may more easily accept gun control. A positive correlation between the spread of contents is illustrated in Figure 5.1(b). In this figure, the more perceived proportion of contacts with one content an individual receives, the lower threshold it requires to turn to the other states. Third, the negative correlation means the spread of one content would impede
the spread of other contents. For example, the spread of opinions on universal healthcare and
tax-relief for the “rich” could be a negative correlation. Because the more we support universal
healthcare, the less we will support tax-relief for the “rich”. This correlation is illustrated in
Figure 5.1(c). As we can observe from the figure, the more perceived proportion of state-1
neighbors an individual has, the higher threshold it requires to turn to state-2.

Moreover, we can have an even more complex split on the parameter space to model a
non-trivial correlation between the spread of contents, like non-linear boundaries for splitting
the parameter space. In this work, we only focus on a linear relationship between the perceived
proportions and the boundaries of the states.

5.2.3 Problem Definition

In this work, we consider an influence propagation process under the proposed vector threshold
model. In particular, assume that all nodes are initially inactive, i.e., state-0, a node is chosen
uniformly at random and is set as state-3. Then, other nodes start changing their states
according to (5.3) synchronously at times \( t = 0, 1, \ldots \), i.e., the influence starts propagating
over networks. Since the contagion process is monotone (i.e., an active node can never switch
back to inactive or other states), it will eventually stop, i.e., a steady-state will be reached.

For analyzing a spread of influence, there is a major metric of interest. The metric is the
expected size of global cascades; see Definition 2 for more details. In other words, for given
contents (e.g., opinions, rumors, products, etc.), we aim at calculating the fraction of people
eventually adopting these contents, in the cases where a global spreading event is possible. We

Figure 5.1: An illustration of different splits on the parameter space for different correlations.
can regard the expected size of global cascades as a measure of the extent of the propagation process. With this measure, we could answer or predict how widely an influence would reach, and then determine if we need to take further actions to control an influence. In addition, this metric could help us derive the condition of the existence of global cascades, which could help us control the propagation of an influence.

Definition 1 (Global Cascades). Global cascades mean a randomly chosen node can initiate a propagation that eventually reaches a positive fraction of active nodes among the whole population. In particular, we define a random variable $S$ for the fraction of active nodes as

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

where $n$ is the number of nodes in the network. In this case, a global cascade means $S > 0$ when $n$ approaches infinity. In our discussion, active nodes indicate nodes in state-1, state-2, and state-3.

Definition 2 (The Expected Size of Global Cascades). Given global cascades are possible, the expected global cascades size is defined as

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

where the random variable $S$ is defined in Definition 1.

5.3 Main Results

5.3.1 Analysis of the expected size of global cascades

In this section, we aim to derive the expected size of global cascades, i.e., the final fraction of active nodes in a network; see Definition 2 for more details. If a global cascade exists, we could obtain the final fraction of active nodes by calculating the probability of that a randomly
chosen node is active (i.e., in state-1, state-2, or state-3), which is expressed as (5.5).

\[
\lim_{n \to \infty} \mathbb{E}[S \mid S > 0] = P[\text{a randomly chosen node is active}]
\]

\[
= \sum_d P_d \sum_{m_1} \sum_{m_2} \sum_{m_3} \frac{d}{m_1} \left( \frac{d - m_1}{m_2} \right) \times \left( \frac{d - m_1 - m_2}{m_3} \right) q_1^{m_1} q_2^{m_2} q_3^{m_3} (1 - q_1 - q_2 - q_3)^{d - m_1 - m_2 - m_3} \times \{ F_1[m, d] + F_2[m, d] + F_3[m, d] \}, \tag{5.5}
\]

where \( q_1 \) (resp. \( q_2 \) and \( q_3 \)) indicates the probability of that the neighbors of a chosen node is in state-1 (resp. state-2 and state-3), \( P_d \) is the probability of having degree \( d \) for a randomly chosen node, and \( m_1 \) (resp. \( m_2 \) and \( m_3 \)) in \( m = [m_1, m_2, m_3] \) is the number of neighbors in state-1 (resp. \( m_2 \) and \( m_3 \)). We explain the validity of (5.5) as follows. First, for a randomly chosen node, the probability of having \( d \) contacts is \( P_d \). Given the node with \( d \) contacts, by a combinatorial argument, the probability of having \( m = (m_1, m_2, m_3) \) neighbors is

\[
\left( \frac{d}{m_1} \right) \left( \frac{d - m_1}{m_2} \right) \left( \frac{d - m_1 - m_2}{m_3} \right) \times q_1^{m_1} q_2^{m_2} q_3^{m_3} (1 - q_1 - q_2 - q_3)^{d - m_1 - m_2 - m_3}. \tag{5.6}
\]

If a node has \( m = (m_1, m_2, m_3) \) active neighbors among \( d \) contacts, the probability of turning active is

\[
F_1[m, d] + F_2[m, d] + F_3[m, d], \tag{5.7}
\]

where \( F_i[m, d] \) is the response function (5.3) of turning to state-\( i \) from state-0, \( i = 1, 2, 3 \). Therefore, (5.5) is obtained by first iterating all possible \( m \) and then taking the expectation of the multiplication of (5.6) by (5.7) with respect to \( d \).

From the expression of (5.5), it is clear that we need the help of \( q_1 \), \( q_2 \), and \( q_3 \) to calculate the expected size of global cascades. We could obtain \( q_1 \), \( q_2 \), and \( q_3 \) by the tree-approximation...
The tree-approximation approach was developed to get a mean-field solution to the Ising model. In this approach, we assume that networks have a tree-structure. That is, each node has only one parent node and several children nodes. Then, we label each layer of the tree from the bottom to the top, 0, 1, ..., \( \ell \). For each node at layer \( \ell \), there is only one parent node at layer \( \ell + 1 \) and several children at layer \( \ell - 1 \). The number of children follows the excess degree distribution mentioned; see [71] for more details. In addition, the states of nodes at layer \( \ell \) will not update their states until all nodes at layer \( i, i = 0, 1, \ldots, \ell - 1 \) finish updating. In this case, parent nodes at \( \ell + 1 \) are always inactive before the updates of the states of nodes at layer \( \ell \). With this tree-assumption, we could derive the probability of being in each state for the contacts of the initially activated nodes in the network. First, we define \( q_{i,\ell} \) as the probability that a randomly chosen node at layer \( \ell \) turns to state \( i \), \( i = 0, 1, 2, 3 \). Obviously, the probability that the chosen node at layer \( \ell \) is in state-0 is \( 1 - q_{1,\ell} - q_{2,\ell} - q_{3,\ell} \).

With these probabilities, we can recursively express the probabilities of being in any states for nodes at layer \( \ell + 1 \).

In the following, we give a detailed derivation of \( q_{1,\ell} \) (5.9), since we can derive the other probabilities in the similar way. We can see the validity of (5.9) as follows. For nodes at layer \( \ell \), the probability of having \( d \) contacts is \( \frac{\alpha P_d}{(d)} \), since we already know that there is one parent at layer \( \ell + 1 \); see excess degree in [71]. Among these \( d \) contacts, there are \( d - 1 \) children and 1 parent node. With these \( d - 1 \) children, the probability of having \( \mathbf{m} = (m_1, m_2, m_3) \) active neighbors is

\[
\frac{(d - 1)}{m_1} \frac{(d - 1 - m_1)}{m_2} \frac{(d - 1 - m_1 - m_2)}{m_3} q_{1,\ell-1}^{m_1} q_{2,\ell-1}^{m_2} q_{3,\ell-1}^{m_3} \times (1 - q_{1,\ell-1} - q_{2,\ell-1} - q_{3,\ell-1})^{d-1-m_1-m_2-m_3}.
\] (5.8)

Again, with \( \mathbf{m} = (m_1, m_2, m_3) \) active neighbors in different states, the probability of turning to state-1 for the node is given by \( F_1[\mathbf{m}, d] \). If we iterate all possible \( \mathbf{m} \) and take an expectation on the product of (5.8) and \( F_1[\mathbf{m}, d] \) with respect to the degree \( d \), we will get (5.9). Then,
using similar arguments, we could get the expressions for $q_2, \ell$ (5.10) and $q_3, \ell$ (5.11).

\[
q_1, \ell = \sum_d \frac{dP_d}{\langle d \rangle} \sum_{m_1} \sum_{m_2} \sum_{m_3} \left( d - 1 - m_1 \right) \left( d - 1 - m_1 - m_2 \right) \mathbf{F}_1[(m_1, m_2, m_3), d - 1, \ell - 1]
\]

(5.9)

\[
q_2, \ell = \sum_d \frac{dP_d}{\langle d \rangle} \sum_{m_1} \sum_{m_2} \sum_{m_3} \left( d - 1 - m_1 \right) \left( d - 1 - m_1 - m_2 \right) \mathbf{F}_2[(m_1, m_2, m_3), d - 1, \ell - 1]
\]

(5.10)

\[
q_3, \ell = \sum_d \frac{dP_d}{\langle d \rangle} \sum_{m_1} \sum_{m_2} \sum_{m_3} \left( d - 1 - m_1 \right) \left( d - 1 - m_1 - m_2 \right) \mathbf{F}_3[(m_1, m_2, m_3), d - 1, \ell - 1]
\]

(5.11)

where

\[
\mathbf{F}_k[(m_1, m_2, m_3), d - 1, \ell - 1] =
\left( d - 1 \right) \left( d - 1 - m_1 \right) \left( d - 1 - m_1 - m_2 \right)
\times q_1, \ell - 1_1 q_2, \ell - 1_2 q_3, \ell - 1_3
\times (1 - q_1, \ell - 1 - q_2, \ell - 1 - q_3, \ell - 1)^{d - m_1 - m_2 - m_3}
\times \mathbf{F}_k[(m_1, m_2, m_3), d].
\]

(5.12)

With the above derivations, we obtain a non-linear system described by the equations (5.9) - (5.11). We could recursively solve the non-linear system, and then get $q_1, \infty$, $q_2, \infty$, and $q_3, \infty$. These quantities obtained from the tree-approximation technique correspond to $q_1$, $q_2$, and $q_3$ introduced at the beginning of this section. Next, we could replace these probabilities into (5.5) to get the expected size of global cascades. As discussed in previous works [41], the tree-approximation technique generates accurate results in the asymptotic limit $n \to \infty$. We will confirm the correctness of our analysis even when the number of nodes is finite in Section

118
The condition of the existence of global cascades

In this section, we aim to find the condition of the existence of global cascades. As mentioned in Section 5.3.1, the proposed vector threshold model could be described by a non-linear system given by (5.9) - (5.11). In this case, the behaviors of this non-linear system imply the state of the dynamics.

As described in Section 5.3.1, we have the relationship $q_1 := q_{1,\infty}$, $q_2 := q_{2,\infty}$, and $q_3 := q_{3,\infty}$. Then, the recursive equations (5.9) - (5.11) have the form

$$q_i = f_i(q_1, q_2, q_3), \quad i = 1, 2, 3.$$  \hspace{1cm} (5.13)

From the equations (5.9) - (5.11), we can easily find that there exists a trivial fixed point $q_1 = q_2 = q_3 = 0$. If we replace this fixed point into (5.5), then we could obtain

$$P[\text{a randomly chosen node is active}] = 0,$$  \hspace{1cm} (5.14)

which means there is no global cascade from Definition 1. However, if there exists a non-trivial fixed point which leads to a positive value for (5.5), then it indicates the existence of global cascades. To check the existence of non-trivial fixed points, we could use the Jacobian matrix $\mathbb{J}$ (5.15) obtained by linearizing equations at $q_1 = q_2 = q_3 = 0$.

$$\mathbb{J} = \frac{\partial f_i(q_1, q_2, q_3)}{\partial q_j}|_{q_1=q_2=q_3=0}. \hspace{1cm} (5.15)$$

If the special radius, i.e., the largest eigenvalue in absolute value, of $\mathbb{J}$ is larger than one, then it means the non-linear system is not stable at $q_1 = q_2 = q_3 = 0$. In this case, there exist non-trivial fixed points, which indicates there would exist global cascades. Otherwise, we could conclude that there do not exist global cascades.
5.4 Numerical Experiments

5.4.1 The agreement between our analysis and experimental results

In this section, we aim to demonstrate the correctness of our analysis when the number of nodes is finite. We use Poisson distribution to assign the degree for each node. Namely, the probability that a node has degree \( d \) is given by:

\[
P(d) = e^{-\lambda} \frac{\lambda^d}{d!}, d = 0, 1, \ldots,
\]  

(5.16)

where \( \lambda \) denotes the mean number of edges assigned for each node. We use \( n = 2 \times 10^6 \) as the number of nodes, and two different split strategies on the parameter space which are shown in Figure 5.2(a) and 5.2(b). We can interpret these two figures as a spread of uncorrelated contents and correlated contents, respectively, which has been discussed in Section 5.2.2. Then, for each degree parameter, we conduct 1,000 experiments, take the average of the cascade size for experiments with global cascades,\(^1\) and depict them in Figure 5.3 and 5.4. From both of Figure 5.3 and 5.4, we can observe that there is a good agreement between our analysis and experimental results. These observations confirm the correctness of our analysis of the expected size of global cascades.

In addition, we observe from Figure 5.3 and 5.4 that there exist two phase transitions, i.e., the place where the expected size of global cascades change from 0 to a positive fraction or vice versa. These two phase transitions are reported in many works [94, 105, 107, 108] and provide insights on the impact of network connectivity on the expected size of global cascades. In short, the first transition indicates that global cascades exist only when the connectivity of a network achieve a certain value. In contrast, the second transition around high \( \lambda \) value appears because too much connectivity will increase the stability of nodes, i.e., when they have a large number of friends, individuals tend to be difficult to get influenced by a few active neighbors.

\(^1\)In a single experiment, if the fraction of active nodes is larger than 0.5%, then we regard the experiment as one with a global cascade.
In this case, the second phase transition appears.

5.4.2 The impact of correlations between spreading contents

As we have introduced in Section 5.2.2, the proposed threshold model offers a general solution for different correlations among contents. In this section, we aim at exploring how different correlations between contents affect the expected size of global cascades.

We make a comparison between three different correlations: independent, positive, and negative correlation, which are illustrated by the region split strategies in Figure 5.5. To make a fair comparison, we keep the thresholds right on the $x$- and $y$-axis the same. With this setting, when a node only receives Proportion$_1$ (resp. Proportion$_2$), i.e., Proportion$_2$ = 0 (resp. Proportion$_1$ = 0), the probability of being in state-1 (resp. state-2) are equal in these correlations.

What we can observe from Figure 5.6 are 1) there still exist two phase transitions; and 2) when the mean degree of nodes is at a low level, these correlations produce global cascades with similar size, while we see significant differences between these correlations when the degree of
Figure 5.4: There is an agreement between our analysis and numerical results. The solid curve indicates the results obtained from our analysis, while the symbols are the results obtained from numerical results. The parameter space is split as shown in Figure 5.2(b).

nodes is at a high level. This can be explained as follows. For the first observation, it happens with the same reason mentioned in Section 5.4.1. Limited connectivity results in the first phase transition, while high stability given by high degree leads to the second phase transition. Since different splits on the parameter space do not change the connectivity of networks, which is the reason behind the appearance of two phase transitions, two phase transitions still exist. We can explain the reason behind the second observation as follows. According to (5.3) and Figure 5.5, we find that the positive correlation makes nodes more easily be activated while it does not change the connectivity. When the mean degree is at low level, limited connectivity is the reason for determining the position of the first phase transition and the expected size of global cascades. Even a correlation makes nodes more easily be activated, the position and the size would not change much. In contrast, when the mean is at high level, there exists enough connectivity. However, high connectivity will increase the local stability of nodes; i.e., according to (5.3) and Figure 5.5, if a node has a large amount of friends, then the node is hard to be activated. The increased local stability would prevent cascades from happening. As observed in Figure 5.5(b), a positive correlation makes inactive nodes easier been activated. In this case, a positive correlation helps nodes to overcome the local stability caused by high degree, which makes global cascades possible again.

From these experiments, we can conclude that depending on the mean degree of nodes, the impact of different correlations between spreading contents on the expected size of global cascades will be significantly different. That is, when the mean degree is at a high level,
the expected size of global cascades is more sensitive to the correlations between spreading contents.

5.5 Conclusion and Future Work

In this work, we propose a vector threshold model that analyzes a simultaneous spread of correlated influence. We derive recursive relations characterizing the dynamics of influence propagation to compute the expected size of global cascades, i.e., cases where a single individual can initiate a propagation that eventually influences a positive fraction of the population. Also, we find what the conditions are in order for the existence of global cascades. Then, we use a numerical study to confirm and support our analytic results. In addition, we report an interesting observation that different correlations could lead to a significant difference on the expected size of global cascades only when the mean degree is at a high level.

There are many possible directions for future work. First, this work only derives the
expected size of global cascades. Giving analytic results for the probability of having global cascades may be a possible direction. Second, this work can be extended to more general network models than the configuration model used here. For instance, it would be interesting to consider networks that have high clustering coefficient. Finally, it would also be interesting to study vector threshold models using non-linear threshold models.
Part IV

Concluding Remarks and Future Work
Chapter 6

Concluding Remarks

In this dissertation, we used a modeling-based approach to study two representative propagation processes, information propagation and influence propagation. In particular, we consider several real-world topological properties (i.e., multiple layers and clustering) and dynamical properties (i.e., multiple stages and simultaneous existence of correlated influences) in mathematically-tractable models, in order to comprehensively study real-world spreading processes. Then, regarding each process, we derived the analytical solutions to two key metrics for each process: the probability of the existence outbreak and the expected size of an outbreak (if an outbreak exists). With the analytical solutions to these two key metrics, we gave a comprehensive study on the impact of several dynamical and topological properties on information and influence propagation, by means of observing how the two metrics change as varying the level of each property. In this chapter, we briefly conclude this dissertation.

In Chapter 2, we worked on information propagation over multi-layer networks with clustering. Including both multiple layers and clustering together in modeling underlying contact networks, we studied the impact of these two topological properties on information propagation. In this work, we found that an increase in the level of clustering decreases the probability of the existence of the outbreak and its expected size. In addition to these observations, we reported that the nature of multiple layers leads to assortativity, which has a multi-faceted impact on information propagation processes. When the degree is at a low level, the assortativity helps the information spread a larger portion of populations, while it reduces the propagation process when the degree is at a high level.
In Chapter 3, we worked on influence propagation over networks with multiple link types and clustering. We still included both multiple layers and clustering together in modeling underlying contact networks, but we shifted the focus from information propagation to influence propagation. Different from the observations in information propagation, we found that clustering exhibited a multi-faceted impact on the propagation process. When the degree is at low level, the impact of clustering on influence propagation is the same as on information propagation that high clustering decreases both of the probability of the existence of the outbreak and its expected size. However, when the degree is at high level, increasing clustering would increase the probability of the expected size. Moreover, we found that the assortativity resulting from the nature of multiple link types plays an important role on influence propagation. Depending on the level of assortativity, the linear threshold models could exhibit four phase transitions instead of two phase transitions, which is one of the first work reporting it.

In Chapter 4, we further studied the influence propagation with multiple levels of states for a given influence over multiplex networks. In this work, the most interesting finding appeared when we consider a case where the hyper-active state is manifested in only one link type. We found that depending on the level of assortativity, the influence exerted by the hyper-active nodes has a much more significant impact on the cascade size as opposed to the case when the network has low assortativity (i.e., when the degrees of neighboring nodes tend to be uncorrelated). Besides, as the influence of hyper-active nodes increases, the highly assortative networks were shown to exhibit changes on not only the critical transition points (i.e., mean degree values at which expected cascade size changes from zero to a positive value, or vice versa), but also the number and order of transitions.

In Chapter 5, we proposed a new threshold model, a vector threshold model, which considered the simultaneous spread of correlated influences. In the newly proposed model, we allowed the existence of correlations between multiple spreading influences, independent, positive, and negative correlations. In addition, we derived the analytical solutions to the probability of global cascades and the expected global cascades (if there exist). This is the first threshold
model supporting the existence of multiple correlated influences.

With these contributions, we comprehensively studied the impact of topological properties (i.e., clustering, multiple layers, multiplexity) and dynamical properties (i.e., multiple stages, and the simultaneous spread of multiple influences) on spreading processes. We reported that the impact of each topological property or dynamical property is usually multi-faceted. More importantly, we found that when we studied the spreading processes, we should consider the collective impacts of properties together. A key takeaway from this dissertation is that the assortativity (i.e., correlation between the degrees of connected pairs) generated by the nature of multiple layers and multiple link types plays an important role in spreading processes. In particular, in the study of information propagation, we showed that assortativity has a multi-faceted impact on propagation processes. When the degree is at a low level, the assortativity helps the information spread a larger portion of populations, while it reduces the propagation process when the degree is at a high level. In addition, in the study of influence propagation, we showed that the level of assortativity would not only change the expected size of an outbreak, but also the number of phase transitions.
Chapter 7

Future Work

Although this dissertation gave a comprehensive study on information propagation and influence propagation, there are still many interesting directions to explore in future work. In this chapter, we discuss some potential future directions.

First, given the different natures of information propagation and influence propagation, this dissertation discussed these two processes separately. However, in reality, there may exist spreading processes that behave like a mixture of information propagation and influence propagation. For example, in a given network, some individuals receive information from or get influenced by a single copy of the resource. In contrast, other individuals need to receive multiple copies to confirm the information, i.e., they behave more like individuals in influence propagation. In this case, proposing a new spreading process by combining the characteristics of information propagation and influence propagation is worth to explore.

Second, in terms of discussing the impact of clustering on spreading processes, we only use one type of motifs, triangles, to generate tunable clustering. In the real-world contact network, there exist more complicated motifs, e.g., the number of cycles of length larger than three, or even a combination of different motifs. In this case, it will be interesting to examine the impact of larger cycles to see if the clustering generated by other motifs behaves differently.

Third, regarding influence propagation, we consider the case where there is two or even more number of possible states for individuals’ reaction to a given influence. However, we assumed that once the state of an individual is determined or activated, its state would not change or be deactivated. In the real-world propagation processes, the states of individuals could be
dynamic, i.e., the states of individuals could be mutable between different ones. Allowing the dynamic changes of states would make our models closer to the spreading processes in the real world.

At last, this dissertation applied a modeling-based approach to study spreading processes. Although a modeling-based approach enables us to easily study the impact of model parameters, there are still some limitations on modeling-based approaches. For example, to consider or add each topological property into our analysis, we need to use handmade ways to add it. When we study the impact of clustering, we directly add triangles to make network models have tunable clustering. However, adding triangles for clustering into our analysis already makes the recursive relations in our analysis hard to solve. If we want to more complicated structures, it would make the relations even more complex. More importantly, triangles are not the only motifs generating clustering. In this case, it is not possible for us to exhaust all possible motifs to generate clustered networks. Besides, there are various topological properties in the real-world networks, so it is infeasible to exhaust all possible topological properties and add them into our analysis. With this issue in mind, we need to explore other possibilities to resolve this issue. Nowadays, learning approaches have been proven to be powerful tools, so we could also use learning approaches to study spreading processes. Therefore, another potential direction is to generate networks by learning approaches. For example, a generative adversarial network (GAN) has successfully been applied for image generation, videos generation, or text generation. With these successful applications, we are motivated to use GAN to generated networks of which properties are even closer to the real-world networks.

With this dissertation, we hope it could provide more insights on understanding spreading processes. More importantly, we hope it could inspire researchers to have more ideas on understanding spreading processes, and even make the world better.
Bibliography


[77] F. Radicchi and A. Arenas. Abrupt transition in the structural formation of intercon-

defects—reviewing the evidence for causality. *New England Journal of Medicine*,

2014.

for spreading processes over multilayer complex networks. *IEEE/ACM Transactions on

[81] F. Sahneh, A. Vajdi, J. Melander, and C. Scoglio. Contact adaption during epidemics:
A multilayer network formulation approach. *IEEE Transactions on Network Science and

[82] A. Saumell-Mendiola, M. Á. Serrano, and M. Boguñá. Epidemic spreading on intercon-
nected networks. 02 2012.


[84] M. Serrano, L. Buzna, and M. Boguñá. Escaping the avalanche collapse in self-similar


