



















# 1

## INTRODUCTION

### 1.1. BACKGROUND AND MOTIVATION

Viral spreading in complex networks has attracted huge interest in past decades. Many propagation processes, e.g. the propagation of information, epidemics and computer virus [1–3], can be studied with the help of viral spreading models. The most widely used type of epidemiological models is the homogeneous model (or the “fully-mixed” model) [4, 5]. In this model, each entity is assumed to connect to several randomly chosen entities, and all entities have approximately the same number of connections in the same time. Network topology is ignored in fully-mixed models and every entity tends to be equally likely infected [6]. However, it has been proved in many real networks that (e.g. [7–9]), the number of connections that an individual has (this is also called the “degree” of an individual in graph theory) follows a power-law distribution and thus individuals are actually not infected with the same probability. Taking the degree heterogeneity into account, Pastor-Satorras and Vespignani [10, 11] developed analytic models for approximating epidemic spreading in networks with power-law degree distribution. Since their model only works well under strict constraints [12], the N-Intertwined Mean-Field Approximation (NIMFA) [13] is then proposed to more accurately investigate the influence of network topology on viral spreading, of which the dynamics is modelled by susceptible-infected-susceptible (SIS) process. Beyond single-layer networks, recently great in-

terest has been drawn in exploring the epidemiological performance in interconnected networks [14–16], where the different infection rates could be applied in different network layers. In classical SIS model, the infection rate and recovery rate are assumed to be homogeneous, i.e., both rates are the same for every node in a network. In recent years few papers (e.g., [17–21]) have discussed the influence of heterogeneous recovery rate and infection rate on viral spreading processes. Many network metrics, including the betweenness centrality [22], have been studied for investigating network epidemiology. In order to identify a representative set of metrics, the correlations between network metrics have been investigated during the past few years [23, 24].

All above researches have been concentrated on viral spreading in networks where all the links facilitate infections (e.g., friends, trusted or alliance), while in reality there also exist negative relations that may suppress viral spreadings (e.g., rivals, distrusted or competitors) [25–29]. Generally, networks that have both types of relationships are called signed networks, in which every link (or edge) is labelled as a positive link or a negative link, representing, e.g., cooperative or competitive relation between two entities ([30–32]). Signed networks have been commonly used to model social networks, e.g., a customer network on Epinions where customers can trust or distrust others’ reviews [33]; a participant network on news website Slashdot where users can claim others as “friends” or “foes” [34]; and a voting network on Wikipedia where each voter express their support or objection to the promotion of others as an administrator [35].

Great interest has been drawn on the study of signed networks in recent years. In 2006, based on the concept of structural balance theory [36], Antal et al. studied how relationships can evolve in signed networks to achieve a social balance [37]. Besides, Kunegis et al. revealed that the Laplacian matrix of a signed networks is always positive-semidefinite, and it only becomes positive-definite when a signed network is unbalanced (i.e., it contains cycles with an odd number of negative edges) [38]. In order to predict the sign of links in signed networks, Leskovec et al. studied principles about the determination of link signs in large social networks and developed a method which was claimed of high prediction accuracy [31].

Despite the previous constructive research in the field of signed networks, the dynamics of viral spreading in signed networks has been barely studied. Intuitively, it is believed that negative relations between individuals can suppress the overall spread of virus. For instance, in a business network, an enterprise may consider to adopt a new technique especially when its cooperators and/or

its competitors have utilised it. However, as more of its competitors have made their adoption, this company would be more reluctant to apply this new technique for, e.g., maintaining its independence from other competitors.

Inspired by the above scenario, in this thesis we propose an original Signed Susceptible-Infected-Susceptible (SSIS) viral spreading model to study the dynamics of propagation processes in signed networks. In the SSIS model, for instance, the adoption of a new technique by a company can be represented by the infection of a node, and cooperative and competitive relations between companies can be modelled by positive links and negative links in signed networks respectively. Instead of a commonly used constant infection rate, we introduce a dynamic infection rate to capture the possible influence of negative relations on overall viral spreading - The infection rate of an individual node is proportional to the fraction of negative neighbours that are susceptible. We investigate the effect of following parameters on the overall performance of viral spreading in signed networks: (1) the relative infection rate on negative links with respect to positive links; (2) the correlation between the number of positive connections and negative connections of nodes; and (3) the topologies of the positive and negative connections. We develop Monte-Carlo (MC) simulations to observe epidemic threshold and fraction of infected nodes in meta-stable state. In order to theoretically analyse the SSIS model, we also derive the Individual-Based Mean-Field Approximation (IBMFA) method to provide analytical solution of the probability that each node can be infected over time. Based on the solutions of all nodes, theoretical epidemic threshold and meta-stable fraction of infected nodes can be obtained. We show that, as a satisfying theoretical tool, IBMFA is able to provide well approximations of MC simulations. We explore the performance of the SSIS model in Erdős-Rényi and scale-free network models, under the conditions that the degrees in positive and negative networks are independent and correlated respectively. In order to verify our findings, we also investigate the performance of the SSIS model in two real-world networks. According to our best knowledge, our work is the first one that studies the dynamics of viral spreading model in signed networks.

## 1.2. THESIS OUTLINE

The rest of the thesis is organised as follows:

In Chapter 2, some background knowledge in graph theory is introduced. Besides, the concept of signed network, along with some classic network models, is given as well.

Our original model was developed based on SIS model, and thus we would like firstly introduce SIS in Chapter 3. This is a classic model that has been broadly used for decades to model dynamic processes, such as epidemic spreadings, in single-layer networks. As a theoretical tool for analysing the SIS model, the NIMFA method is also introduced in this chapter.

In Chapter 4 we propose our original SSIS model to characterise viral spreading processes in signed networks. We also introduce two methods to investigate the performance of the SSIS model: The Monte-Carlo (MC) simulation is used for observing the properties of SSIS; and the Individual-Based Mean Field Approximation (IBMFA) is used to numerically analyse the model. We conclude this chapter with the discussion on the accuracy of IBMFA method in approximating the MC simulation results.

Chapter 5 is focused on investigating the performance of SSIS in degree-independent signed networks. We would like to discover the effect of viral spreading via negative links on the epidemic threshold and meta-stable fraction of infected nodes in signed networks, mainly regarding to the relative infection rate of negative links with respect to positive links.

In Chapter 6, we study the performance of SSIS in degree-correlated signed networks. Our interest is mainly concentrated on how can correlations between degrees in positive and negative networks influence the overall viral spreading in signed networks.

In order to verify the conclusions drawn in the former two chapters, in Chapter 7 we explore the performance of SSIS in two real-world signed networks.

Finally, we summarise this thesis and give some suggestions on possible future works in Chapter 8.

# 2

## BACKGROUND KNOWLEDGE

In graph theory, a real-world complex system is usually described by a network (or a graph)  $G=(N,L)$ , where  $N$  stands for the number of nodes and  $L$  denotes the number of links in the network. Objects in a real-world system are represented by nodes in a graph, and the pairwise relations between objects are indicated by links connecting nodes. Mathematically, a graph can be represented by an adjacency matrix which is introduced in Section 2.1. Each element in the adjacency matrix indicates whether one node is connected to another or not. In Section 2.2 we would like to introduce two network metric: degree and spectral radius. These two metrics are significant in the study of viral spreadings. Aiming at modelling the complex real-world systems, several random graph models have been developed in the past decades. In Section 2.3 two classical network models - Erdős-Rényi random network and scale-free network - are presented. Besides, the concept of multi-layer network and signed network are given as well.

### 2.1. ADJACENCY MATRIX

In graph theory, an adjacency matrix  $A$  is a binary square matrix used to describe the pairwise rela-

tions between nodes. For a network with  $N$  nodes, its adjacency matrix is shown as below:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & a_{N3} & \cdots & a_{NN} \end{bmatrix} \quad (2.1)$$

The value of an element  $a_{ij}$  indicates the existence or absence of a link between node  $i$  and node  $j$ .  $a_{ij}=1$  if there is a link connecting node  $i$  and node  $j$ ; otherwise the value is 0. For finite simple graphs, all diagonal values of their adjacency matrices are zeros (i.e.  $a_{ii} = 0$  for  $i=1,2,\dots,N$ ), meaning that no self loops are allowed. An adjacency matrix is symmetric, i.e.  $a_{ij}=a_{ji}$ , if a network is undirected.

## 2.2. NETWORK METRICS

### 2.2.1. DEGREE

The degree  $d_i$  of node  $i$  represents the number of nodes that are connected to  $i$ . In other words,  $d_i$  stands for the number of neighbours that node  $i$  has. Given an adjacency matrix  $A$  of size  $N \times N$ , the degree  $d_i$  can be calculated as the summation of all elements in the  $i$ -th row of  $A$ , i.e.

$$d_i = \sum_{k=1}^N a_{ik} \quad (2.2)$$

For a graph  $G(N,L)$ , the average degree  $E[D]$  is obtained as:

$$E[D] = \frac{2L}{N} \quad (2.3)$$

The degree distribution  $\Pr[D=k]$ , which is the fraction of nodes with degree  $k$  in a network, is a statistical network feature that reflects global network property. Given the degree distribution of a network,  $E[D]$  can also be obtained follows:

$$E[D] = \sum_{k=k_{min}}^{k_{max}} k \Pr[D = k] \quad (2.4)$$

### 2.2.2. SPECTRAL RADIUS

The spectral radius of a network is defined as the largest eigenvalue  $\lambda_{max}(A)$  of its adjacency matrix  $A$ . It is a significant network metric that characterises the performance of epidemic spreading in networks. For instance, epidemic threshold is an important metric in epidemiology that determines whether a disease persists or it dies out. According to [39], the epidemic threshold in real networks can be approximated as  $\frac{1}{\lambda_{max}(A)}$ .

## 2.3. NETWORK MODELS

### 2.3.1. ERDÖS-RÉNYI RANDOM NETWORK

The Erdős-Rényi (ER) random network [40, 41] is one of the popular network models that allows many problems being analytically solved [42]. In ER networks, every two random nodes can be equally likely connected with probability  $p$ , and all links are independent from each other. Given that an ER network has  $N$  nodes, the average number of links is then determined as  $E[L] = \frac{N(N-1)}{2}p$ , where  $L_{max} = \frac{N(N-1)}{2}$  stands for the maximum number of links may exist in the network. The definition of ER networks implies that each element  $a_{ij}$  in an adjacency matrix  $A$  (except for  $a_{ii}$ ) is a Bernoulli random variable with mean  $p$ , leading to a binomial degree distribution:

$$Pr[D = k] = \binom{N-1}{k} p^k (1-p)^{N-1-k} \quad (2.5)$$

where  $Pr[D = k]$  represents the probability that the degree  $D$  of a random node equals to  $k$ . As the size of network enlarges, the degree distribution of ER networks converges to a Poisson distribution:

$$Pr[D = k] = \frac{(pN)^k}{k!} e^{-Np} \quad (2.6)$$

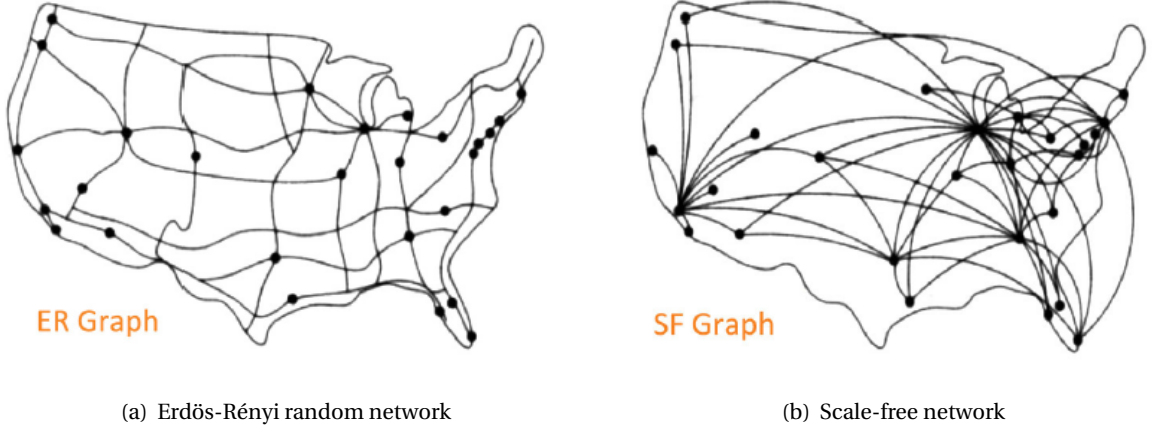


Figure 2.1: An example of network models [43]

### 2.3.2. SCALE-FREE NETWORK

Another classical network model is the scale-free (SF) network [44], of which the property of power-law degree distribution  $Pr[D = k] \propto k^{-\gamma}$  has been observed in many real-world networks [44–47]. For most real-world networks, the exponent  $\gamma$  lies between 2 and 3 [48]. In a SF network, the majority of nodes have small degrees; while only a few nodes, which are called “hubs”, have a great number of neighbours. For example, in online social networks such as Facebook, most people have a small number of friends; while a few celebrities can have connections to many other people.

Given a power-law exponent  $\gamma$ , the probability distribution function (PDF) of degree distribution in SF networks can be expressed as

$$Pr[D = k] = ck^{-\gamma} \quad (2.7)$$

where  $c$  is a constant and can be calculated by

$$c = \frac{1}{\sum_{k=k_{min}}^{k_{max}} k^{-\gamma}} \quad (2.8)$$

We use the configuration model [49] to generate SF networks in simulations. The configuration model generates a random graph using a given degree sequence. The detailed steps are summarised as follows:

1. Given the PDF of a desired SF network, the cumulative distribution function (CDF)  $F(x) = Pr[D \leq x] = \left( \sum_{k=k_{min}}^x Pr[D = k] \right)$  is calculated.



2. Based on the CDF, a degree sequence  $K = \{k_1, k_2, \dots, k_N\}$  is generated, where  $k_i$  stands for the degree of node  $i$  ( $i = 1, 2, \dots, N$ ). For each node, a uniformly distributed random number  $m \in [0, 1)$  is firstly generated, and then  $k_i$  is obtained according to the following rule.

$$k_i = \begin{cases} k_{min}, & 0 \leq m < F(m_{min}) \\ x, & x > k_{min}, \text{ and } F(x-1) \leq m < F(x) \end{cases} \quad (2.9)$$

3. For every node  $i$ ,  $k_i$  “stubs” are assigned in preparation for the further establishment of links.
4. Every time two randomly selected spare stubs are connected, and the connecting procedure continues until all stubs are connected.
5. Finally, all self-loops and duplicated links are removed and the entire generation concludes.

### 2.3.3. MULTI-LAYER NETWORK

In reality, rather than single networks, many complex systems are modelled by multi-layer interdependent networks. For example, the functioning of a computer network depends on the power supply from electrical grids; while the operation of power stations in turn relies on computer networks [50]. Multi-layer networks, also called as multiplex networks, consist of several interdependent or interconnected network layers. Commonly, connections between the same set of nodes vary in different layers of networks [51–53]. In other words, a node  $i$  may have 4 neighbours in layer 1, while in layer 2 it is possibly connected to other 3 totally different nodes. In general, signed networks, which will be introduced later, can be regarded as multi-layer networks.

### 2.3.4. SIGNED NETWORK

As defined in graph theory, a signed graph is a graph in which every link is labelled as either “positive (+)” or “negative (-)”. A positive link represents, e.g., the cooperative relationship between two companies or the friendship between two individuals in the real world; while a negative link indicates the competitive relationship between two companies or the hostility between two individuals. Generally, a signed network can be considered as a multi-layer network of two layers - one

layer consists of only positive links (positive network) and the other layer is comprised by only negative links (negative network). The positive network and the negative network share the same set of nodes, while the neighbours of each node are different in positive and negative networks. An example of signed networks is given in Fig. 2.2. Each pair of nodes are either not connected or are connected by only one type of link, hence no overlapping link exists in signed networks. In this thesis, the term “positive neighbours” denotes the neighbour nodes connecting to a node via positive links; and “negative neighbours” refers to the neighbours connecting to a node via negative links. The number of positive and negative neighbours that a node has are called “positive degree” and “negative degree”, respectively.

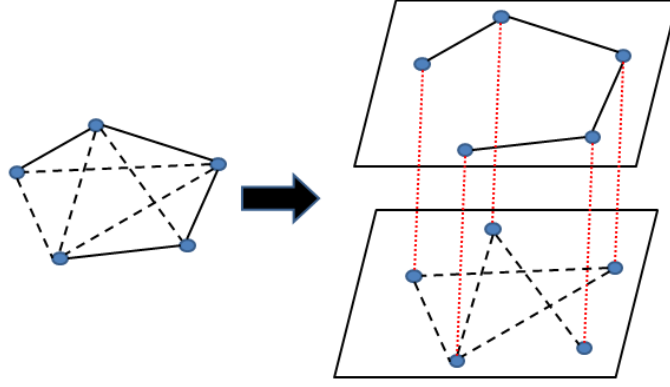


Figure 2.2: An example of signed network. Solid lines represent positive links and dash lines represent negative links. The red dotted lines indicate that the positive network shares the same node set with the negative network. As shown in the figure, one signed network (on the left) can be considered as a multi-layer network of two layers (on the right) - one layer only has positive links (positive network) and the other one only has negative links (negative network)

The positive-negative degree correlation  $p_D$  is a quantity that reflects the probable difference between degrees in positive networks and negative networks.  $p_D$  is a linear correlation coefficient and it is defined as:

$$p_D = \frac{\sum_{i=1}^N (d_i^+ - \langle d_+ \rangle)(d_i^- - \langle d_- \rangle)}{\sqrt{\sum_{i=1}^N (d_i^+ - \langle d_+ \rangle)^2} \sqrt{\sum_{i=1}^N (d_i^- - \langle d_- \rangle)^2}} \quad (2.10)$$

where  $d_i^+$  and  $d_i^-$  are positive and negative degree of node  $i$  respectively, and  $\langle d_+ \rangle$  and  $\langle d_- \rangle$  denote average positive and negative degree. If  $p_D = 1$ ,  $d_i^+ = d_i^-$  holds for every node  $i$ ; and when  $p_D = 0$  the positive and negative degree are independent of each other.

# 3

## SIS MODEL

In this chapter we briefly introduce the classical Susceptible-Infected-Susceptible (SIS) model, on which the proposal of our original viral spreading model is based. Besides, details about the N-Intertwined Mean-Field Approximation (NIMFA) method, which is used for theoretically analysing the SIS model, are given as well. NIMFA is a state-of-art mathematical tool used for deriving the probability of infection over time for each node. Based on the results of NIMFA, some common problems, including meta-stable fraction of infected nodes and the epidemic threshold, can be solved.

### 3.1. INTRODUCTION OF SIS MODEL

The Susceptible-Infected-Susceptible (SIS) model [13, 54–56] is a classical continuous-time model that has been broadly used to study dynamic processes such as virus spreading in computer networks. In a SIS process, the viral state of a node  $i$  at time  $t$  can be represented by a Bernoulli random variable  $X_i(t) \in \{0,1\}$ :  $X_i(t) = 0$  represents a susceptible node (or a healthy node) and  $X_i(t) = 1$

indicates an infected node. A healthy individual can be infected by a virus and transforms into an infected node, and vice versa. At time  $t$ , node  $i$  can be infected with probability  $\nu_i(t) = \Pr[X_i(t)=1]$ , or can be healthy but susceptible to infections with probability  $1-\nu_i(t)$ . Once a node is infected, it becomes an epidemic disseminator and can infect its healthy neighbours till the node recovers. It is assumed that the curing process, where an infected node  $i$  recovers and becomes susceptible again, is a Poisson process with rate  $\delta$ . Similarly, the infection process, where a healthy node get infected by a neighbour, is also a Poisson process with rate  $\beta$  per link. The curing and infection Poisson processes are independent, and the time that a node can be infected or recovered are i.i.d. (independent identically distributed) exponential random variables. Since exponential random variables are memoryless, i.e.  $\Pr[T > s + t \mid T > t] = \Pr[T > s]$ , the SIS model has in nature Markov property: past states have no influences on the conditional probability distribution of future states of a node.

The effective infection rate is defined as  $\tau \triangleq \frac{\beta}{\delta}$ . As mentioned by many researchers (e.g. [3, 4, 10, 57, 58]), there exists a critical infection rate  $\tau_c$ , around which an epidemic phase transition can be observed: When  $\tau > \tau_c$ , a virus constantly persists and a positive fraction of infected node can be found in the steady state [59]; whereas when  $\tau \leq \tau_c$ , the epidemic rapidly dies out [27, 60]. In SIS model, any epidemic spreading process will eventually terminate with an absorbing state, where all nodes are healthy and virus can no longer spread through a network. In general, when  $\tau \leq \tau_c$  the absorbing state is achieved at least exponentially fast [39, 61]. However, when  $\tau > \tau_c$ , the absorbing state can only be reached after an extremely long time [59], and instead for the most of the time a network stays in a long-live state called meta-stable state, in which the epidemic rapidly converge - the fraction of infected nodes remains relative constant and reduces to zero extremely slowly [59].

### 3.2. N-INTERTWINED MEAN FIELD APPROXIMATION (NIMFA)

Due to the Poisson property of both curing and infection processes, the entire SIS viral spreading processes are Markov processes. The two-state Markov chain of the state transition of a node  $i$  is shown as in Fig. 3.1. It can be characterised by a infinitesimal generator  $Q_i(t)$  defined as:

$$Q_i(t) = \begin{bmatrix} -q_i(t) & q_i(t) \\ \delta & -\delta \end{bmatrix} \quad (3.1)$$

where  $\delta$  is the recovery rate of each node. The element  $q_i(t)$ , which is the rate that a susceptible node  $i$  being infected at time  $t$ , is defined as:

$$q_i(t) = \beta \sum_{j=1}^N a_{ji} 1_{\{X_j(t)=1\}} \quad (3.2)$$

which is the summation of infection rates  $\beta$  of all links connecting to infected neighbours of node  $i$ . For an undirected network, the adjacency matrix  $A$  is symmetric and hence  $a_{ji} = a_{ij}$ . The rate  $q_i$ , which couples node  $i$  to the rest of the network [13], is a random variable as mentioned in [54]. Since Markov theory requires that  $q_i$  should be a real number, the randomness in  $q_i$  needs to be removed. One method for removing the random nature is to condition all the possible states of the neighbours of node  $i$ . However, this method leads to a dramatic increase in the number of basic states in the Markov chain. For a network of  $N$  nodes, the conditioning ends up with a  $2^N$ -state Markov chain [54], which causes great demanding computational cost.

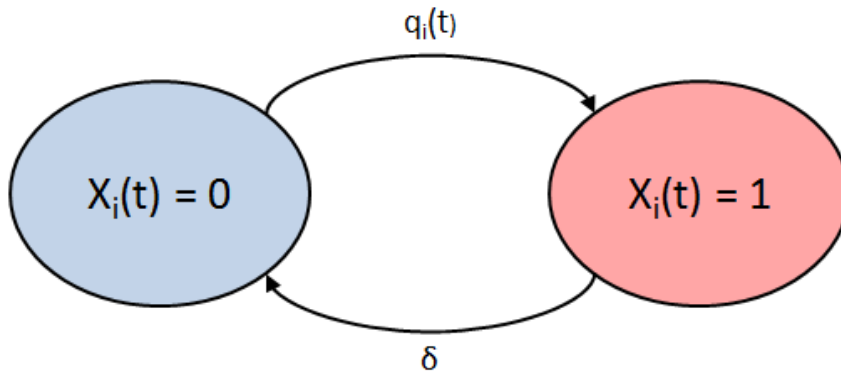


Figure 3.1: Markov chain of node state in SIS model ( $X_i(t) = 0$ : healthy  $X_i(t) = 1$ : infected)

Instead of conditioning, the mean-field approximation [62] is widely considered as an effective method for randomness removal [13]. Without considering the combination of states, in the mean-field approximation the total infection rate  $q_i$  is replaced by its average  $E[q_i]$ , which is shown below:

$$E[q_i(t)] = E \left[ \beta \sum_{j=1}^N a_{ji} 1_{\{X_j(t)=1\}} \right] \quad (3.3)$$

$E[q_i]$  is a real number and satisfies the condition for implementing Markov theory [63]. Using  $E[1_x] = \Pr[x]$ , Eq.(3.3) can be written as:

$$E[q_i(t)] = \beta \sum_{j=1}^N a_{ji} \Pr[X_j(t) = 1] = \beta \sum_{j=1}^N a_{ji} v_j(t) \quad (3.4)$$

Thus the infinitesimal generator Eq.(3.1) becomes:

$$Q_i(t) = \begin{bmatrix} -E[q_i(t)] & E[q_i(t)] \\ \delta & -\delta \end{bmatrix} \quad (3.5)$$

By mean-field approximation the computational burden is significantly reduced: one only need to obtain  $N$  random variables of  $v_i(t)$  rather than computing the  $2^N$ -state Markov process.

Since the summation of dependent indicators  $\sum_{j=1}^N 1_{\{X_j(t)=1\}}$  is replaced by the average probability of node infection  $v_i(t)$ , the mean-field approximation suffers a cost as a bit loss of accuracy. Nevertheless, the mean-field approximation is still widely regarded as an effective method for approximating the SIS viral spreading processes.

By applying (3.4), the NIMFA governing equation [63, pp. 182] can be derived as:

$$\begin{cases} \frac{dv_1(t)}{dt} = (1 - v_1(t))\beta \sum_{j=1}^N a_{1j}v_j(t) - v_1(t)\delta \\ \frac{dv_2(t)}{dt} = (1 - v_2(t))\beta \sum_{j=1}^N a_{2j}v_j(t) - v_2(t)\delta \\ \vdots \\ \frac{dv_N(t)}{dt} = (1 - v_N(t))\beta \sum_{j=1}^N a_{Nj}v_j(t) - v_N(t)\delta \end{cases} \quad (3.6)$$

The physical interpretation of the differential equation is given as follows:

1. When node  $i$  is healthy with probability  $1 - v_i(t)$ , each of its infected neighbours tries to infect this node with rate  $\beta$ ; and probability of the event that node  $i$  gets infected at time  $t$  is  $\sum_{j=1}^N a_{ij}v_j(t)$ .
2. When node  $i$  is infected with probability  $v_i(t)$ , it recovers with a rate  $\delta$ .

Eq.(3.6) can also be written in matrix form as [54]:

$$\frac{dV(t)}{dt} = \beta AV(t) - \text{diag}(v_i(t))(\beta AV(t) + \delta u) \quad (3.7)$$

where vector  $V(t) = [v_1(t) \ v_2(t) \cdots v_N(t)]^T$ , and  $u$  is an all-one vector [54]. By solving Eq.(3.6), the fraction of infected nodes in steady state, denoted as  $\rho$ , is obtained as:

$$\rho = \frac{\sum_{i=1}^N v_{i\infty}}{N} \quad (3.8)$$

where  $v_{i\infty}$  denotes the infection probability of node  $i$  in the steady state. The epidemic threshold in NIMFA is defined as  $\tau_c = \frac{1}{\lambda_1}$ , where  $\lambda_1$  is the largest eigenvalue of the adjacency matrix. According to [59], NIMFA generally gives the upper bounds of both meta-stable fraction of infected nodes and epidemic threshold in SIS model.





# 4

## SSIS MODEL AND APPROACHES

In this chapter, we propose our original Signed-SIS (SSIS) model that describes the dynamics of viral spreading in signed networks. We developed two methods to study the performance of SSIS model. The first one is the Individual-Based Mean Field Approximation (IBMFA) method, which is used to theoretically analyse the model; and the second one is the Monte-Carlo (MC) method, which is used to simulate the real-time epidemic process in SSIS.

### 4.1. THE SSIS MODEL

The development of our SSIS model was inspired by the observation of the following real-world scenario: In a business network, the relation between two companies could be cooperative (positive link), competitive (negative link) or unrelated (no connection). A company is likely to adopt a new technique to improve its own business if either its partners or its competitors have utilised this technique. However, as more of its competitors have made their adoptions, this company may become more reluctant to apply this new technique in order to, e.g., protect their independency from

other competitors.

According to the above observation, we propose the following SSIS model. Instead of a constant infection rate  $\beta$  in SIS, in SSIS we use a dynamic infection rate  $\beta_i(t)$  for each node: The infection rate of node  $i$  at time  $(t + \Delta t)$  is determined by the infection status of its negative neighbours at time  $t$ . If a susceptible node has no negative neighbours, it can be infected only at a constant rate of  $\beta$  per positive link. Otherwise, at time  $t$ , node  $i$  may be infected by a positive neighbour at a rate of:

$$\beta_i^+(t) = \beta \left( 1 - \frac{\sum_{j=1}^N b_{ij} X_j}{d_i^-} \right) \quad (4.1)$$

where  $b_{ij}$  denotes the element of the adjacency matrix  $B$  of negative networks,  $X_j$  represents the infection state of node  $j$  (0 for susceptible and 1 for infected), and  $d_i^-$  stands for the negative degree of  $i$ . Similarly, node  $i$  can also be infected via a negative link at a rate of:

$$\beta_i^-(t) = \alpha \beta_i^+(t) = \alpha \beta \left( 1 - \frac{\sum_{j=1}^N b_{ij} X_j}{d_i^-} \right) \quad (4.2)$$

where  $\alpha$  is the relative infection rate of negative links with respect to positive links (in the following text we will call it as “relative negative-link infection rate”, or “RN infection rate”).  $\alpha > 0$  represents that negative links can make contribution to the overall viral spreading; Otherwise, if  $\alpha = 0$  virus are not allowed to propagate via negative links.  $\alpha$  can also be viewed as the ratio between infection rates of negative and positive links. For instance,  $\alpha = 0.5$  indicates that the infection rate of negative links is in general half of the infection rate of positive links.

Note that  $\beta/\alpha\beta$  is the rate that a virus transmits through a positive/negative link; whereas  $\beta_i^+(t)/\beta_i^-(t)$  is the actual rate the node  $i$  can be infected by a positive/negative neighbour. According to the definition in Eq.(4.1) and Eq.(4.2),  $\beta$  and  $\alpha\beta$  are the maximum rate that one node can be infected via a positive and negative link respectively. As the viral spreading proceeds, with more negative neighbours being infected, a single susceptible node would be infected at a lower rate, and thus this node can become more resistant to infection. In other words,  $\beta_i^+(t)$  and  $\beta_i^-(t)$  are proportional to the number of uninfected negative neighbours. In the following text, without specific instructions the term “infection rate” always refers to the maximum positive infection rate  $\beta$ . In the real world, Eq.(4.1) and Eq.(4.2) can be interpreted as an “aware” mechanism: A company’s willingness of utilising a new technique can decrease as they have collected sufficient information to find out that

more competitors are using it.

## 4.2. APPROACHES FOR INVESTIGATING SSIS MODEL

### 4.2.1. INDIVIDUAL-BASED MEAN FIELD APPROXIMATION

Aiming at theoretically analysing the SSIS model, we develop the Individual-Based Mean Field Approximation (IBMFA), which is inspired by NIMFA introduced in Section 3.2 [13]. IBMFA can be used to compute the infection probability  $v_i(t)$  of node  $i$  at any time  $t$ , including the meta-stable state. Based on the infection probabilities of all nodes, we can solve the meta-stable fraction of infection  $\rho$ . The epidemic threshold  $\tau_c$  in signed networks can also be obtained with the help of IBMFA.

#### THE GOVERNING EQUATION OF IBMFA

The IBMFA governing equation for a node  $i$  in our SSIS model is given as:

$$\frac{dv_i(t)}{dt} = -v_i(t)\delta + (1-v_i(t)) \left[ \alpha\beta \left( 1 - \frac{\sum_{j=1}^N b_{ij}v_j(t)}{\sum_{j=1}^N b_{ij}} \right) \sum_{j=1}^N b_{ij}v_j(t) + \beta \left( 1 - \frac{\sum_{j=1}^N b_{ij}v_j(t)}{\sum_{j=1}^N b_{ij}} \right) \sum_{j=1}^N a_{ij}v_j(t) \right] \quad (4.3)$$

The interpretation of the equation is given as follows:

- **Term**  $(-v_i(t)\delta)$ : When node  $i$  is infected with probability  $v_i(t)$ , it can recover at a rate of  $\delta$ .
- **Term**  $(1-v_i(t)) \left[ \beta \left( 1 - \frac{\sum_{j=1}^N b_{ij}v_j(t)}{\sum_{j=1}^N b_{ij}} \right) \sum_{j=1}^N a_{ij}v_j(t) \right]$ : When node  $i$  is healthy with probability  $1-v_i(t)$ , each of its infected positive neighbours attempts to infect this node at rate  $\beta \left( 1 - \frac{\sum_{j=1}^N b_{ij}v_j(t)}{\sum_{j=1}^N b_{ij}} \right)$ , where  $b_{ij}$  is the element in the adjacency matrix of a negative network (denoted as  $B$ ).  $\sum_{j=1}^N b_{ij}$  is the negative degree of node  $i$  and  $\sum_{j=1}^N b_{ij}v_j(t)$  denotes the number of infected negative neighbours of node  $i$  at time  $t$ . The probability that node  $i$  is infected by its positive neighbours at time  $t$  is  $\sum_{j=1}^N a_{ij}v_j(t)$ .
- **Term**  $(1-v_i(t)) \left[ \alpha\beta \left( 1 - \frac{\sum_{j=1}^N b_{ij}v_j(t)}{\sum_{j=1}^N b_{ij}} \right) \sum_{j=1}^N a_{ij}v_j(t) \right]$ : When node  $i$  is healthy with probability

$1 - v_i(t)$ , each of its infected negative neighbours tries to infect this node at rate  $\alpha\beta \left(1 - \frac{\sum_{j=1}^N b_{ij} v_j(t)}{\sum_{j=1}^N b_{ij}}\right)$ , where  $\alpha$  represents the RN infection rate. The probability that node  $i$  is infected by its negative neighbours at time  $t$  is given by  $\sum_{j=1}^N b_{ij} v_j(t)$ .

According to Eq.(4.3), the more negative neighbours of a susceptible node  $i$  are infected, the less probable  $i$  can be infected. The IBMFA governing equation can also be written in the matrix form as:

$$\begin{aligned} \frac{dV(t)}{dt} = & -\delta V(t) + \text{diag}(u - V(t)) \cdot \beta(A + \alpha B) \cdot V(t) \\ & - \text{diag}(B \cdot V(t)) \cdot \text{diag}(D_-) \cdot \text{diag}(u - V(t)) \cdot \beta(A + \alpha B) \cdot V(t) \end{aligned} \quad (4.4)$$

where  $u$  is an all-one vector,  $A$  is the adjacency matrix a positive network,  $B$  is the adjacency matrix of a negative network.  $D_-$  is an  $N$ -element vector of which each element  $d_{-i}$  is the reciprocal of negative degree  $d_i^-$  of node  $i$  if  $d_i^- > 0$ , or  $d_{-i} = 0$  if  $d_i^- = 0$ .

#### THE SOLUTION OF IBMFA

The actual steady state of a Markovian process is the absorbing state, where the epidemic dies out. However, for realistic sizes of networks, when  $\tau > \tau_c$  to reach the absorbing state an unrealistically long time is required [61]. We are interested in the meta-stable state which can be reached fast and maintained for an extremely long time [59]. With  $\frac{dv_i}{dt} = 0$ , a trivial solution  $v_{i\infty} = 0$  is always obtained corresponding to the absorbing state, as well as a possible non-zero solution  $v_{i\infty} > 0$  which stands for the probability of infection in meta-stable state.

The numerical solution of  $v_{i\infty}$  could be computed by using the ode45 function in MATLAB. The ode45 function is used to solve non-stiff ordinary differential equations. It is a commonly preferred method to solve numerical problems in complex network models. Since the meta-stable state of a Markov chain does not depend on the initial condition [59, pp. 191], any value between 0 and 1 could be set as initial value for solving Eq.(4.4). In our IBMFA realisations the initial values  $v_i(0)$  were chosen to be 0.5 for all nodes. For a network of  $N$  nodes, the final solution  $V(t)$  consists of  $N$  elements, and the meta-stable fraction of infection  $\rho$  is calculated as the average of  $V(t)$ .

$$\rho = \frac{\sum_{i=1}^N v_{i\infty}}{N} \quad (4.5)$$

#### 4.2.2. MONTE-CAROL SIMULATION

We developed discrete-time Monte-Carol simulations to approximate the SSIS epidemic processes. Although in reality epidemic processes are continuous-time Markov, as pointed out in [59], with a sample rate  $\frac{1}{\Delta t}$  larger than the maximum transition rate  $\max_i q_i$  in Markov processes, the meta-stable state in discrete-time Markov process is guaranteed to be equivalent to that in continuous-time Markov process [59, Section 10.4]. In our simulations, each final result is obtained as the average over at least 200 realizations. During every realisation, we first generate a new signed network based on the given network parameters, such as number of nodes, average positive and negative degrees particularly for ER networks, and power-law exponent particularly for SF networks, etc.. We use a sample time of  $\Delta t = 0.01s$  when  $\beta < 1$  (which means that each time step in our simulation stands for 0.01s in real world), and  $\Delta t = 0.001s$  when  $\beta > 1$ . Initially, 10% of randomly selected nodes are set to be infected at time  $t = 0$ . At each time step, every healthy node  $i$  could be infected by an infected positive neighbour with a probability of  $\Delta t \cdot \beta_i^+$ , or by an infected negative neighbour with a probability of  $\Delta t \cdot \alpha \beta_i^-$ ;  $\beta_i^+$  and  $\beta_i^-$  are calculated according to Eq.(4.1) and Eq.(4.2) respectively. At the same time step, each infected node may be recovered with a probability of  $\Delta t \cdot \delta$ . We run enough time steps to achieve the precise meta-stable fraction of infection  $\rho$ . We record  $\rho$  for every 20 time steps and check the difference between adjacent records. If the difference is small than a pre-determined value for at least five consecutive times of check, we would confirm that the meta-stable state has been reached. In this thesis, all networks used in both MC and IBMFA are undirected.

### 4.3. COMPARISON BETWEEN MC AND IBMFA

The comparison between Monte-Carol simulation results and IBMFA approximations are shown in Fig. 4.1 and Fig. 4.2. Despite different epidemic thresholds and meta-stable fractions of infected nodes, the results of IBMFA approximation are consistent with those of MC simulation. In other words, IBMFA can approximately reflect the effects of different  $\alpha$  on the performance of SSIS model in both ER and SF networks. Although we have further verified that IBMFA method is generally able to give good approximation of MC simulation results (results are shown in Appendix A.1), we would

like to mainly use the precise MC simulation results to discover properties of SSIS model in the following chapters.

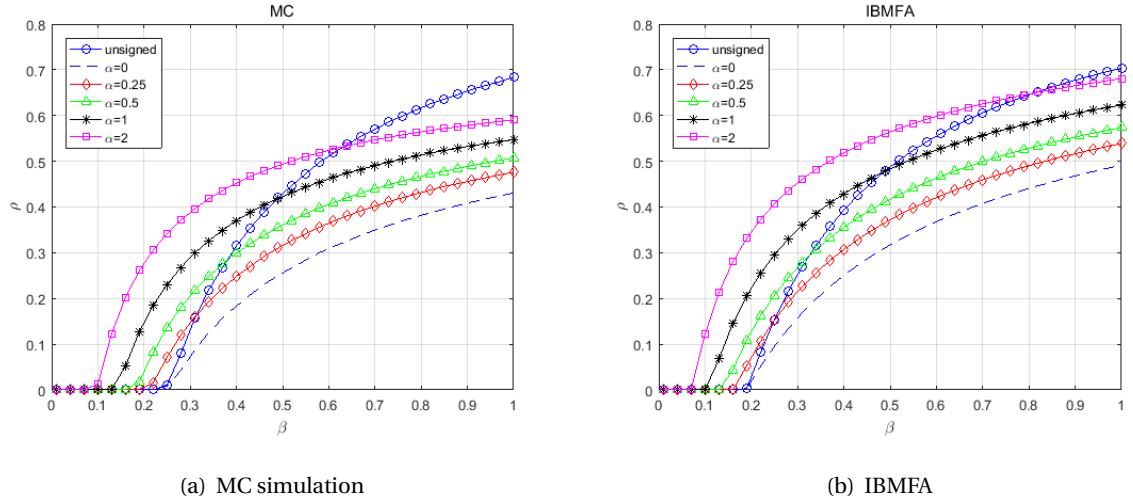


Figure 4.1: Comparison between results of Monte-Carlo simulation and IBMFA approximation. The network size is  $N = 1000$  nodes. Both positive and negative networks are ER networks with average degree  $E[D_+] = E[D_-] = 4$ . The unsigned network is obtained by removing all negative links in the negative network. The degree correlation between positive and negative networks is  $p_D \approx 0$ . RN infection rate  $\alpha$  are 0, 0.25, 0.5, 1 and 2 respectively

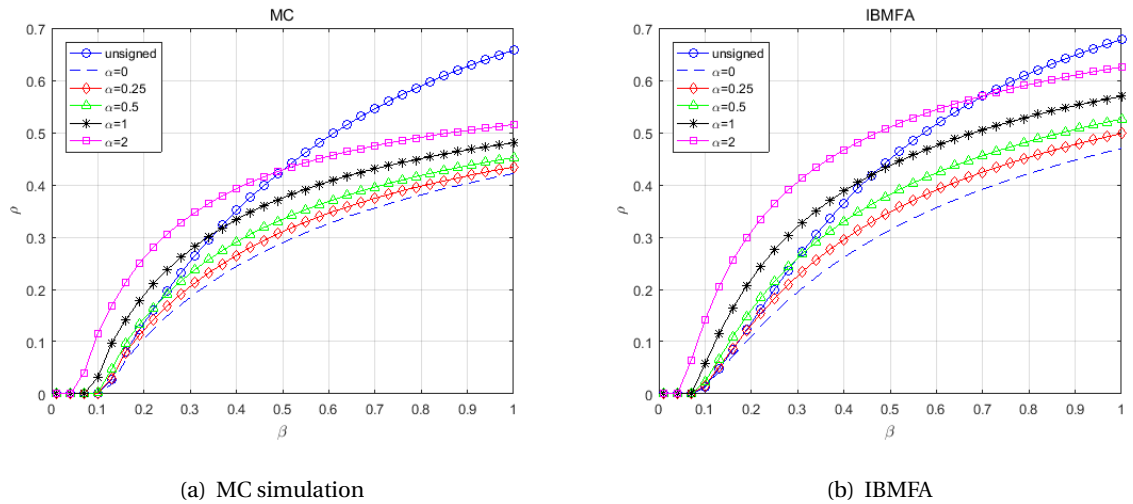


Figure 4.2: Comparison between results of Monte-Carlo simulation and IBMFA approximation. The network size is  $N = 1000$  nodes. Both positive and negative networks are SF networks with power-law exponent  $\gamma = 2.5$ . The unsigned network is obtained by removing all negative links in the negative network. The degree correlation between positive and negative networks is  $p_D \approx 0$ . RN infection rate  $\alpha$  are 0, 0.25, 0.5, 1 and 2 respectively

# 5

## SSIS MODEL IN DEGREE-INDEPENDENT SIGNED NETWORKS

In this chapter we explore the performance of the SSIS model in degree-independent signed networks. The term “degree-independent” refers to that the positive degree of a random node  $d_i^+$  is not correlated to its negative degree  $d_i^-$ , i.e.  $p_D = 0$ . We apply the SSIS model in both signed Erdős-Rényi random (ER) networks and scale-free (SF) networks. In signed ER networks, we study the epidemic performance of SSIS under two control parameters - the average negative degree  $E[D_-]$  and the RN infection rate  $\alpha$ . Our attention is mainly focused on the influence of these two parameters on the epidemic threshold  $\tau_c$  and the fraction of infected nodes in meta-stable state  $\rho$ . We also investigate the statistical relation between  $\alpha$  and a particular infection rate  $\beta^*$ , at which the meta-stable fraction of infected nodes in signed networks  $\rho_{signed}$  is equal to the one in unsigned networks  $\rho_{unsigned}$ . As to signed SF networks, the influence of  $\alpha$  on  $\rho$  is mainly studied.

## 5.1. SSIS MODEL IN SIGNED ER RANDOM NETWORKS

In this section we explore the performance of SSIS model in signed degree-independent ER random networks. We focus on the influence of two main parameters on the overall viral spreading: the average negative degree  $E[D_-]$  (in Section 5.1.1) and the RN infection rate  $\alpha$  (in Section 5.1.2). Finally in Section 5.1.3 we prove that the particular infection rate  $\beta^*$ , at which we have  $\rho_{signed} = \rho_{unsigned}$ , is approximately linearly related to  $\alpha$ .

### 5.1.1. INFLUENCE OF AVERAGE NEGATIVE DEGREE $E[D_-]$

The meta-stable fractions of infected nodes  $\rho$ , corresponding to different infection rate  $\beta$ , are plotted in Fig. 5.1 for three signed ER networks of the same average positive degree  $E[D_+]$  but different average negative degree  $E[D_-]$ . As a benchmark, we also plot  $\rho$  in the unsigned network of which the negative networks are null. It can be seen that all signed networks have smaller epidemic thresholds than the unsigned network; and the threshold decreases as  $E[D_-]$  rises up. This observation can be analytically explained as follows: The epidemic threshold can be approximated as the reciprocal of the spectral radius of a network, where spectral radius is defined as the largest eigenvalue of an adjacency matrix [1]. Since a signed network can be regarded as a two-layer interconnected network, the epidemic threshold can be approximated as  $\tau_c^{signed} = \frac{1}{\lambda_{max}(A + \alpha B)}$ , where  $A + \alpha B$  is the adjacency matrix of a interconnected network [64]. An unsigned network can also be considered as a signed network, in which the negative network is null. Thus the epidemic threshold of an unsigned network can be derived as  $\tau_c^{unsigned} = \frac{1}{\lambda_{max}(A)} = \frac{1}{\lambda_{max}(A + 0B)}$ . According to [64–66], we have the following lemma stands:

$$\lambda_{max}(A) \leq \lambda_{max}(A + \alpha B) \leq \lambda_{max}(A) + \lambda_{max}(\alpha B) \quad (5.1)$$

Therefore, if  $\alpha > 0$  it is certain to find that  $\tau_c^{signed} \leq \tau_c^{unsigned}$ . In addition, it is also observed that as  $E[D_+]$  and  $\alpha$  remain constant,  $\tau_c^{signed}$  reduces as  $E[D_-]$  increases. This is because of that, according to [60], in ER graphs  $\tau_c$  converges logarithmically fast towards zero as the number of links increases. Since we have  $E[D] = \frac{2L}{N}$ , the increase in number of negative links is equivalent to a higher  $E[D_-]$ .



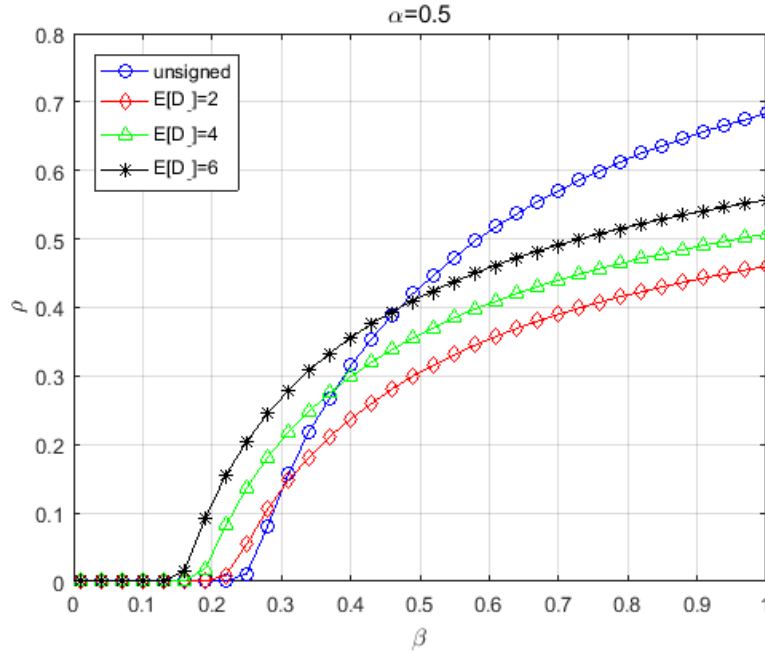


Figure 5.1: Effect of different average negative degree  $E[D_-]$  on the performance of SSIS in degree-independent signed ER networks.  $\alpha$  is a constant of 0.5, meaning that the infection rate of negative links is generally half of the rate of positive links. The average degree of positive networks is  $E[D_+]=4$  in all signed networks. The unsigned network can be considered as a signed network of which average positive degree is 4 and the negative network is null.

It is also noticed that, compared to other signed networks, at the same infection rate a signed network with a more dense negative network always has a larger  $\rho$  in meta-stable state. To our analysis, this observation can be explained as follows. The dynamic infection rate  $\beta_i^+$  and  $\beta_i^-$  of node  $i$  are both proportional to the number of susceptible negative neighbours of  $i$ . When the number of negative links increases, on average each node can have more negative neighbours (since  $E[D] = \frac{2I}{N}$ ), and thus the infection of every negative neighbours would contribute less to the decline of dynamic infection rates. As a result, the individual meta-stable probability of infection  $v_{i\infty}$  can be increased correspondingly, resulting in a greater  $\rho$ .

Although IBMFA can numerically explain the above observation, here we would like to derive a closed-form analytical expression of  $v_{i\infty}$  to prove our analysis. For simplicity, we apply the fully-mixed assumption where the network topology has been ignored [6]. In other words, all nodes have approximately the same positive (equals to  $E[D_+]$ ) and negative degree (equals to  $E[D_-]$ ) at any time  $t$ , and thus in meta-stable state (if exists) all nodes have the same positive probability of infection. Based on this assumption, Eq.(4.3) can be simplified as:

$$\frac{dv_i(t)}{dt} = (1 - v_i(t))[\alpha\beta(1 - v_i(t)) \cdot E[D_-] \cdot v_i(t) + \beta(1 - v_i(t)) \cdot E[D_+] \cdot v_i(t)] - \delta v_i(t) \quad (5.2)$$

In meta-stable state, we have  $\frac{dv_{i\infty}}{dt} = 0$  and the Eq.(5.2) can be written as:

$$\delta v_{i\infty} = (1 - v_{i\infty})^2 \cdot v_{i\infty} \cdot \beta(\alpha \cdot E[D_-] + E[D_+]) \quad (5.3)$$

Let  $\delta = 1$ , by omitting the trivial solution  $v_{i\infty} = 0$  (which represents the absorbing state) Eq.(5.3) can be further simplified as:

$$(1 - v_{i\infty})^2 = \frac{1}{\beta(\alpha \cdot E[D_-] + E[D_+])} \quad (5.4)$$

Taking into account that  $v_{i\infty} \leq 1$ , the solution of Eq.(5.3) is given as:

$$v_{i\infty} = 1 - \sqrt{\frac{1}{\beta(\alpha \cdot E[D_-] + E[D_+])}} \quad (5.5)$$

As indicated by Eq.(5.5), with all the other parameters remaining constant, in signed ER networks  $v_{i\infty}$  grows along with the increase in  $E[D_-]$ . Since  $\rho = \frac{\sum_{i=1}^N v_{i\infty}}{N}$ , we have proved that with the same numbers of positive links and the same RN infection rate  $\alpha$ , more negative links may facilitate viral spreadings in signed ER networks. This is consistent with the simulation results shown in Fig. 5.1.

Interestingly, we observe some crossing points  $(\beta^*, \rho^*)$  between the  $\beta - \rho$  curves of signed and unsigned networks in Fig. 5.1. When the infection rate  $\beta$  is smaller than  $\beta^*$ , more nodes can be infected in signed networks than in unsigned network; whereas when  $\beta > \beta^*$ ,  $\rho_{signed}$  becomes smaller than  $\rho_{unsigned}$ . We found that there is a approximate linear relation between  $\beta^*$  and  $\alpha$ , and more discussion will be given in Section 5.1.3.

### 5.1.2. INFLUENCE OF RN INFECTION RATE $\alpha$

Besides  $E[D_-]$ , we have found that the RN infection rate  $\alpha$  is another key factor that can affect the performance of SSIS model. Fig. 5.2 shows that almost all signed networks with positive  $\alpha$  have smaller epidemic thresholds than the unsigned network, and the threshold decreases as  $\alpha$  rises up. To explain this observation, given  $\alpha_1 < \alpha_2$ , we slightly modify the lemma in Eq.(5.1) into the following form:

$$\lambda_{max}(A + \alpha_1 B) \leq \lambda_{max}(A + \alpha_1 B + \alpha' B) \leq \lambda_{max}(A + \alpha_2 B) \quad (5.6)$$

where  $\alpha' = \alpha_2 - \alpha_1$ . When  $\alpha = 0$ ,  $\frac{1}{\lambda_{max}(A + \alpha B)}$  is equivalent to  $\frac{1}{\lambda_{max}(A)}$ , and under this condition signed and unsigned networks undoubtedly share the same epidemic threshold, as shown in Fig. 5.2.

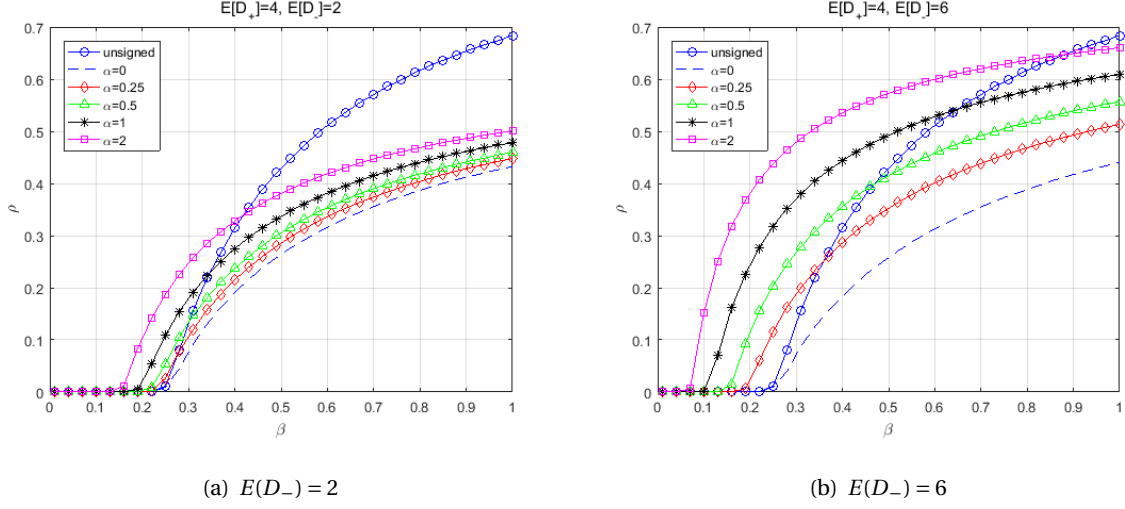


Figure 5.2: Effect of different  $\alpha$  on the performance of SSIS in degree-independent signed ER networks.  $\alpha=0, 0.25, 0.5, 1$  and  $2$  respectively. The constant average positive degree is  $E[D_+] = 4$ , and  $E[D_-]$  are set as  $2$  in (a) and  $6$  in (b). The unsigned network can be considered as a signed network of which the average positive degree is  $4$  and the negative network is null.

It is also found that in signed ER networks, with all other parameters remaining constant, the meta-stable fraction of infected nodes  $\rho$  tends to grow along with the increase in  $\alpha$ . As implied in Eq.(5.5), as long as one node is not totally blocked from further infection (i.e.,  $\beta_i^+(t) \neq 0$  and  $\beta_i^-(t) \neq 0$ ), a larger  $\alpha$  should always result in a higher meta-stable infection possibility  $v_{i\infty}$  and thus a higher  $\rho$ . However, we have found that this conclusion only holds when the infection rate  $\beta$  is small, and more details will be given in Section 5.2. In addition, it is also inferred in Eq.(5.5) that in an extreme case, as  $\alpha \rightarrow \infty$ ,  $v_{i\infty}$  (and thus  $\rho$ ) may converge to 1. Table 5.1 and Table 5.2 demonstrate results from IBMFA computation. It can be seen that even with a small  $\beta$ ,  $v_{i\infty}$  becomes closer to 1 as  $\alpha$  increases. What's more, the denser networks are, the easier  $v_{i\infty}$  converges towards 1.

Table 5.1: IBMFA meta-stable fraction of infected nodes at large  $\alpha$  ( $\beta=0.5, \delta=1$ )

$\alpha$	50	100	200	500
$E[D_-] = 2, E[D_+] = 4$	0.7644	0.7943	0.8157	0.8349
$E[D_-] = 6, E[D_+] = 4$	0.8916	0.9185	0.9386	0.9575
$E[D_-] = 10, E[D_+] = 10$	0.9288	0.9487	0.9632	0.9764

Table 5.2: IBMFA meta-stable fraction of infected nodes at large  $\alpha$  ( $\beta=1, \delta=1$ )

$\alpha$	50	100	200	500
$E[D_-] = 2, E[D_+] = 4$	0.8157	0.8359	0.8506	0.8640
$E[D_-] = 6, E[D_+] = 4$	0.9194	0.9392	0.9540	0.9681
$E[D_-] = 10, E[D_+] = 10$	0.9491	0.9634	0.9737	0.9832

### 5.1.3. CROSSING POINTS ( $\beta^*, \rho^*$ ) BETWEEN $\beta$ - $\rho$ CURVES OF SIGNED AND UNSIGNED NETWORKS

In both Fig. 5.1 and Fig. 5.2 we have observed crossing points ( $\beta^*, \rho^*$ ) between  $\beta$ - $\rho$  curves of signed and unsigned networks, where at  $\beta = \beta^*$  the meta-stable fraction of infected nodes is the same in both signed and unsigned networks, i.e.  $\rho_{signed} = \rho_{unsigned} = \rho^*$ . When the infection rate  $\beta < \beta^*$ ,  $\rho_{signed}$  is larger than  $\rho_{unsigned}$ ; While  $\rho_{unsigned}$  becomes overwhelming when  $\beta > \beta^*$ .

In this section, we are going to investigate how does  $\beta^*$  changes as a function of  $\alpha$ . To derive this statistical relation, first we need to solve the NIMFA governing equation Eq.(3.6) to find the meta-stable probability of infection  $v_{i\infty}$  in an unsigned network. Similar to Eq.(5.2), with the additional fully-mixed assumption [6] Eq.(3.6) can be simplified as:

$$\frac{dv_{i\infty}}{dt} = (1 - v_{i\infty})\beta \cdot E[D_+] \cdot v_{i\infty} - \delta v_{i\infty} \quad (5.7)$$

With  $\frac{dv_{i\infty}}{dt} = 0$  and  $\delta=1$ , by omitting the trivial solution  $v_{i\infty}=0$  the meta-stable infection probability is given as:

$$v_{i\infty} = 1 - \frac{1}{\beta \cdot E[D_+]} \quad (5.8)$$

When  $\beta = \beta^*$ , under the fully-mixed assumption  $\rho_{signed} = \rho_{unsigned}$  is equivalent to  $v_{i\infty}^{signed} = v_{i\infty}^{unsigned}$ . Hence, by substituting Eq.(5.8) into Eq.(5.4) we have:

$$\beta^* = \frac{E[D_-]}{E[D_+]^2} \alpha + \frac{1}{E[D_+]} \quad (5.9)$$

According to Eq.(5.9),  $\beta^*$  tends to be linearly correlated with  $\alpha$ . This is consistent with results demonstrated in Fig. 5.3. Note that when  $\alpha = 0$  a signed network and an unsigned networks share the same epidemic threshold, and this is the only point where the  $\beta - \rho$  curves of these two networks “cross”. Hence, we define the point  $(\tau_c, 0)$  as the crossing point under  $\alpha = 0$  (in our case, since

$\delta = 1$  we have  $\tau_c = \beta_c$ ). Fig. 5.3 shows that in both MC and IBMFA results  $\beta^*$  is generally a linear function of  $\alpha$ , and IBMFA can better approximate the MC simulation results as the average degree of positive and/or negative networks increases. This is, in general, in line with our understanding that the mean-field approximation performs better in dense networks. In addition, it is indicated in Fig. 5.3(c) and Fig. 5.3(d) that our theoretical  $\beta^* - \alpha$  linear function Eq.(5.9) also works well in dense signed ER networks.

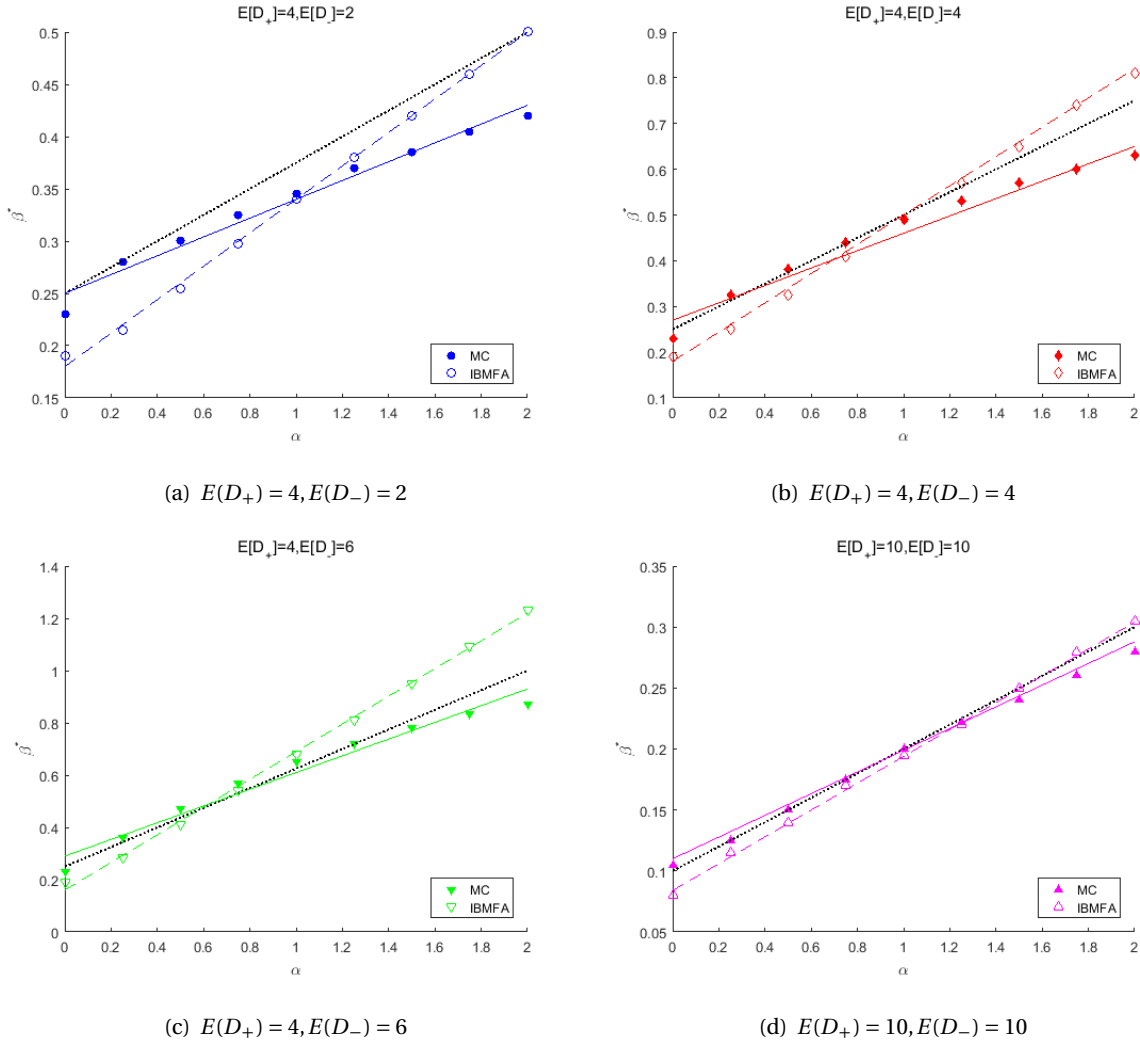


Figure 5.3: Infection rate  $\beta^*$  at crossing point corresponding to different  $\alpha$  in degree-independent signed ER networks. Solid markers show the results from MC simulations and the hollow markers represent results from IBMFA approximation. Fitting curves are provided for both MC (solid lines) and IBMFA results (dash lines). The dotted lines depict the theoretical equation  $\beta^* = \frac{E[D_-]}{E[D_+]^2} \alpha + \frac{1}{E[D_+]}$

## 5.2. SSIS MODEL IN SIGNED SF NETWORKS

In this section the performance of SSIS in degree-independent signed SF networks will be discussed, with respect to the influence of  $\alpha$  on  $\tau_c$  and  $\rho$ , and to the  $\beta^*$ - $\alpha$  relation. In our simulations, given the network size  $N = 1000$  nodes and the connectivity exponent  $\gamma = 2.5$ , the average positive and negative degrees are approximately  $E[D_+] = E[D_-] = 4$ . As a benchmark, results of unsigned networks, of which the negative networks are null, are presented as well. We set the minimum degree  $k_{min}=2$  and the maximum degree  $k_{max}=100$  when generating the networks.

As shown in Fig. 5.4(a), under  $\alpha > 0$  signed networks have smaller epidemic thresholds than unsigned network, and  $\tau_c$  declines along the increase in  $\alpha$ . Similar to the case in signed ER networks, this observation can also be explained by the lemma presented in Eq.(5.6). Compared to signed ER networks (Fig. 5.4(b)), the epidemic thresholds of signed SF networks are much smaller. The reason is that, given the similar average degree, due to the existence of hubs [60, 67],  $\tau_c$  can converge to 0 much faster than in ER networks. At small infection rate ( $\beta < 1$ ),  $\rho_{\alpha 1} < \rho_{\alpha 2}$  always holds given  $\alpha 1 < \alpha 2$ . In addition, crossing points  $(\beta^*, \rho^*)$  exist between  $\beta - \rho$  curves of signed and unsigned SF networks. Comparing Fig. 5.5(a) to Fig. 5.5(b), in signed SF networks  $\beta^*$  is still generally linearly correlated to  $\alpha$ , while the linearity between  $\beta^*$  and  $\alpha$  is a bit weaker than in signed ER networks.

We used to expect that regardless of  $\beta$ , with a larger  $\alpha$ , a higher meta-stable fraction of infected nodes should always be observed as shown in Fig. 5.6(a). However, we surprisingly find that, in degree-independent signed SF networks as shown in Fig. 5.6(b), at large  $\beta$ , a higher  $\rho$  can possibly be obtained when negative links do not participate in viral propagation ( $\alpha = 0$ ). We believe that such observation is mainly caused by the topological feature of SF networks. When  $\alpha > 0$ , hubs in negative networks would have dominantly high probability of infection at large  $\beta$ . In SSIS model, the probability of infection of a susceptible node declines as more of its negative neighbours become infected. Since hubs are neighbours of the majority of other nodes in networks, the infections of these negative hubs can in general contribute more to the decrease in infection probability of other nodes, and thus help suppress overall infection. On the other hand, when negative links are not allowed to transmit virus, the infection probability of negative hubs could be reduced, since in degree-independent signed networks negative hubs tend to have small positive degree. As a result,

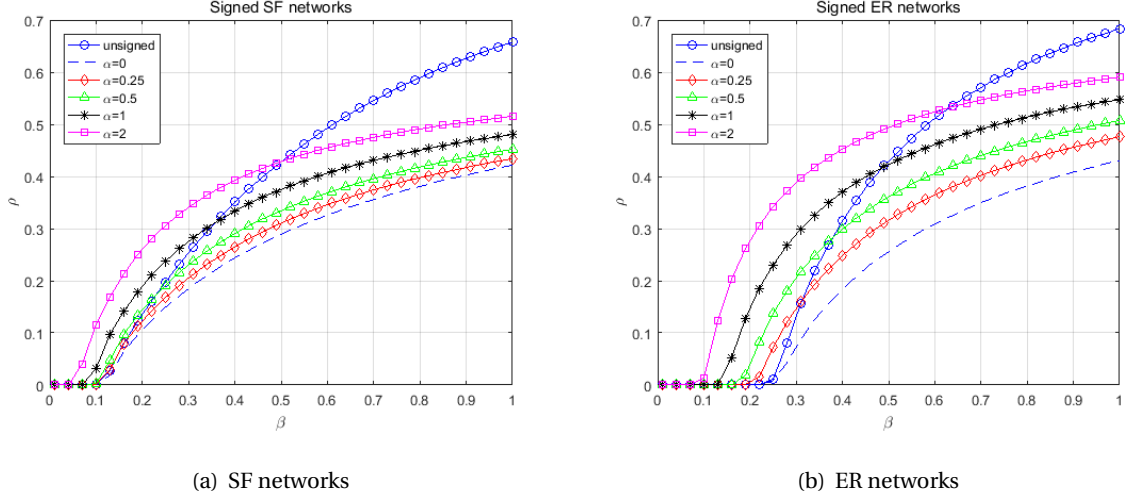


Figure 5.4: Comparison between the performance of SSIS model in signed SF networks and ER networks. Both types of networks are degree-independent.  $\alpha=0, 0.25, 0.5, 1$  and  $2$  respectively. The average degrees of ER networks are  $E[D_+] = E[D_-] = 4$ , and the power-law exponent of SF positive and negative networks is  $\gamma = 2.5$ . The unsigned network is an SF network with null negative network in (a) and an ER network with null negative network in (b).

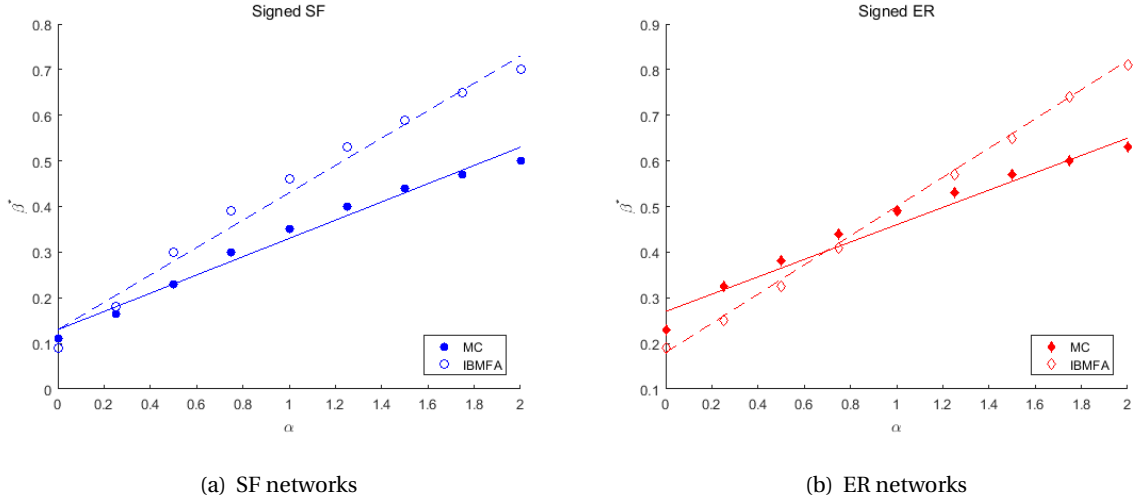


Figure 5.5: Infection rate at crossing point  $\beta^*$  corresponding to different  $\alpha$  in degree-independent signed SF networks and ER networks. Solid markers show the results from MC simulations and the hollow markers represent results from IBMFA approximation. Fitting curves are provided for both MC (solid lines) and IBMFA results (dash lines). The average degrees of positive and negative ER networks are  $E[D_+] = E[D_-] = 4$ , and the power-law exponent of SF positive and negative networks is  $\gamma = 2.5$ .

most nodes could have higher infection probabilities, leading to a possible higher meta-stable fraction of infected nodes at large  $\beta$ . As to degree-independent signed ER networks, distributions of both positive and negative degree are relatively homogeneous. Hence, regardless of  $\alpha$ , the infection probability are approximately the same for each node. In this case, a higher  $\alpha$  appears to result in

a higher meta-stable fraction of infected nodes at any  $\beta$ . Our finding can be further supported by the results of SF-ER signed networks in Fig. 5.6(c) (positive network is SF and negative network is ER) and ER-SF signed networks in Fig. 5.6(d). It can be seen that a smaller  $\alpha$  can possibly lead to a higher  $\rho$  only when negative networks are SF networks. Otherwise, when negative networks are ER networks, a higher  $\rho$  tend to be obtained with a larger  $\alpha$  at any infection rate.

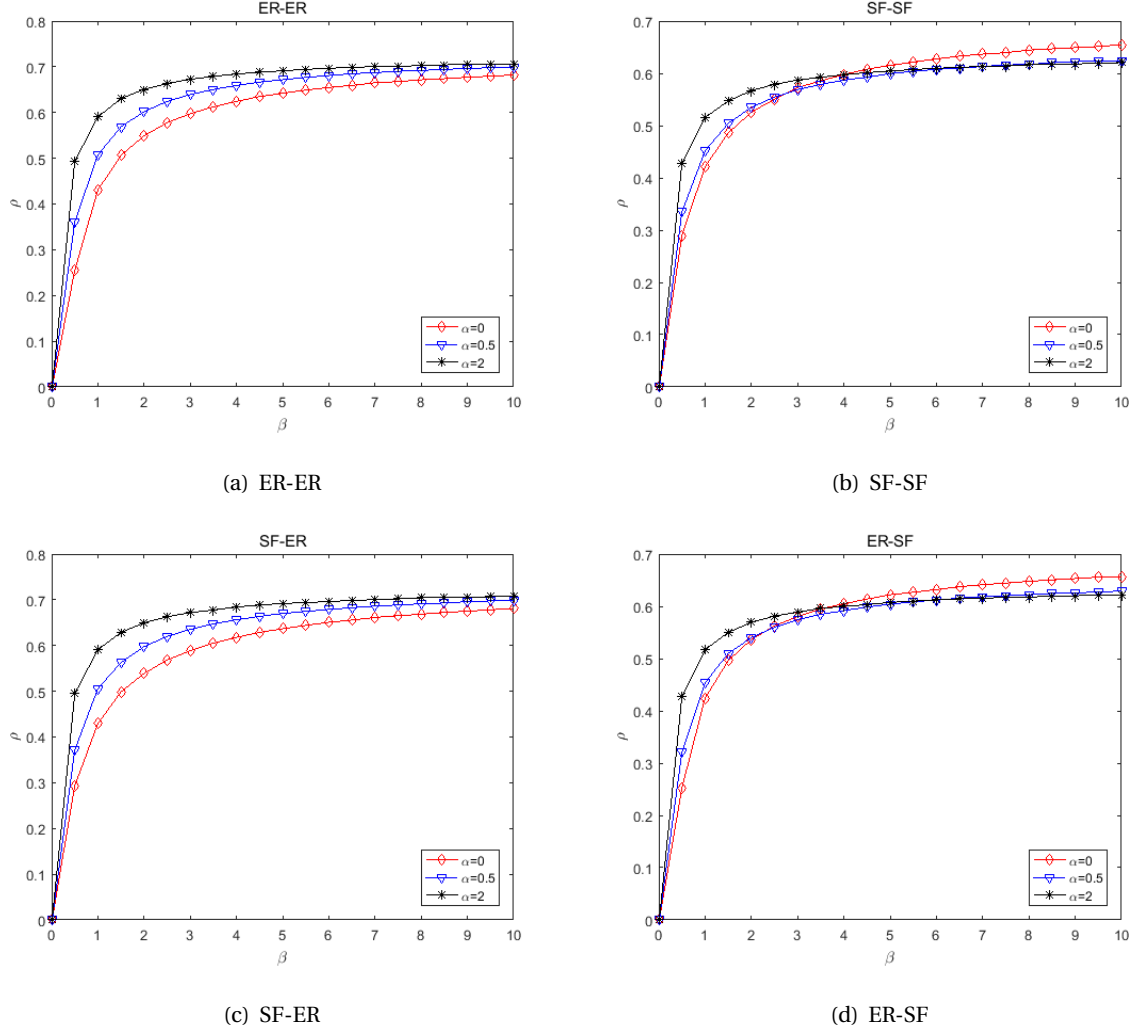


Figure 5.6: Influence of RN infection rate  $\alpha$  on the performance of SSIS in different types of degree-independent signed networks.  $\alpha=0, 0.5$ , and  $2$  respectively. The average degree of ER networks are  $E[D_+] = E[D_-] = 4$ , and the power-law exponent of SF positive and negative networks is  $\gamma = 2.5$ .

To conclude this chapter, we summarise the main discoveries in degree-independent signed networks as follows.

1. With the participant of negative links in viral spreading (i.e.  $\alpha > 0$ ), signed networks can give



smaller epidemic threshold  $\tau_c$  compared to unsigned networks. When the infection rate  $\beta$  is less than a particular value  $\beta^*$ , the meta-stable fraction of infected nodes in signed networks  $\rho_{signed}$  is larger than  $\rho_{unsigned}$ . However, when  $\beta > \beta^*$ ,  $\rho_{unsigned} < \rho_{signed}$  always holds unless  $\alpha \rightarrow \infty$ . We have proved that  $\beta^*$  is approximately linearly related to  $\alpha$ ,

2. Compare to other signed networks, when  $\alpha$  is positive and constant, a signed ER network with more negative links tends to give a lower  $\tau_c$  and a higher meta-stable fraction of infected nodes  $\rho$ .
3. In the same signed network, a higher  $\alpha$  can lead to a lower  $\tau_c$ . In signed ER network, a higher  $\alpha$  tends to result in a larger  $\rho$  at any  $\beta$ ; While in signed SF networks, compared to  $\alpha = 0$ , when  $\alpha > 0$  the overall viral spreading can be suppressed at large  $\beta$ .



# 6

## SSIS MODEL IN DEGREE-CORRELATED SIGNED NETWORKS

This chapter is focused on the performance of SSIS in degree-correlated signed networks, in which the positive degree  $d_i^+$  of node  $i$  is related to its negative degree  $d_i^-$ . We develop a method for generating degree-correlated signed networks (Section 6.1) and investigate how degree correlation  $p_D$ , along with RN infection rate  $\alpha$ , may influence epidemic threshold  $\tau_c$  and the meta-stable fraction of infected nodes  $\rho$  (Section 6.2).

### 6.1. GENERATION OF DEGREE-CORRELATED SIGNED NETWORKS

First of all, we would like to introduce our method for generating degree-correlated networks. This method was inspired by the generation of indegree-outdegree correlated networks proposed by Qu et al. [20, 68]. The details for generating degree-correlated signed networks are given as follows:

1. We firstly generate the positive network. ER network is generated by randomly selecting and

connecting nodes; As to SF network, it is generated using the method described in Section 2.3.2.

2. Next, we copy the degree sequence of the positive network for generating the negative network.
3. Given the degree correlation  $p_D$ , we randomly select  $p_D$  fraction of nodes, of which the negative degrees are kept be the same as the positive degrees. After that, the rest elements of the degree sequence are shuffled, and the negative degrees of the remaining  $1-p_D$  fraction of nodes are then randomly re-allocated.
4. We repeat step 1 to generate the negative network.

## 6.2. RESULTS AND ANALYSIS

According to the MC simulation results (shown in Appendix A.2), regardless of  $p_D$ , the influence of  $\alpha$  on  $\tau_c$  and  $\rho$  (at small infection rate  $\beta$ ) in degree-correlated networks is similar to that in degree-independent networks. Hence in this section we mainly focus on, when  $\beta$  is large, how can different  $p_D$  influence the meta-stable fraction of infected nodes with different  $\alpha$  in degree-correlated signed networks.

Based on the findings in Section 5.2, we speculate that when  $p_D$  is large, the absence of negative links in viral spreading ( $\alpha = 0$ ) might no longer lead to a larger meta-stable fraction of infected nodes in degree-correlated signed SF networks. Results in Fig. 6.1 show that, as  $p_D$  increases, the  $\beta - \rho$  curve of  $\alpha = 0$  and  $\alpha = 2$  start intersecting at larger  $\beta$ . In addition, when  $p_D = 0.8$  even no intersection between two curves can be observed in Fig. 6.1(d). This indicates that, in a less correlated signed SF network, viral spreadings tend to be more encouraged at large  $\beta$ , when nodes are not allowed to be infected by their negative neighbours (i.e.,  $\alpha = 0$ ); On the contrary, in a highly correlated signed SF network, regardless of  $\beta$ , more nodes tend to be infected when negative links participate in transmitting virus (i.e.,  $\alpha > 0$ ).

The above observations can be explained as follows. As discussed in Section 5.2, a hub node in a negative network usually has small positive degree when  $p_D$  is small. Hence, compared to  $\alpha > 0$ , hubs in negative networks are less likely to be infected when  $\alpha = 0$ . Since hubs have many neighbours, lower probabilities of infection of hubs can facilitate higher infection rates of many other nodes, leading to a higher meta-stable fraction of infected nodes at large  $\beta$ , i.e.  $\rho_{\alpha=0} > \rho_{\alpha>0}$ . On the other hand, when  $p_D$  becomes larger, hubs in the negative network are more likely to have high positive degree, and they also tend to be highly likely infected when  $\alpha = 0$ . As a result,  $\rho_{\alpha=0} < \rho_{\alpha>0}$  tend to hold at any  $\beta$ .

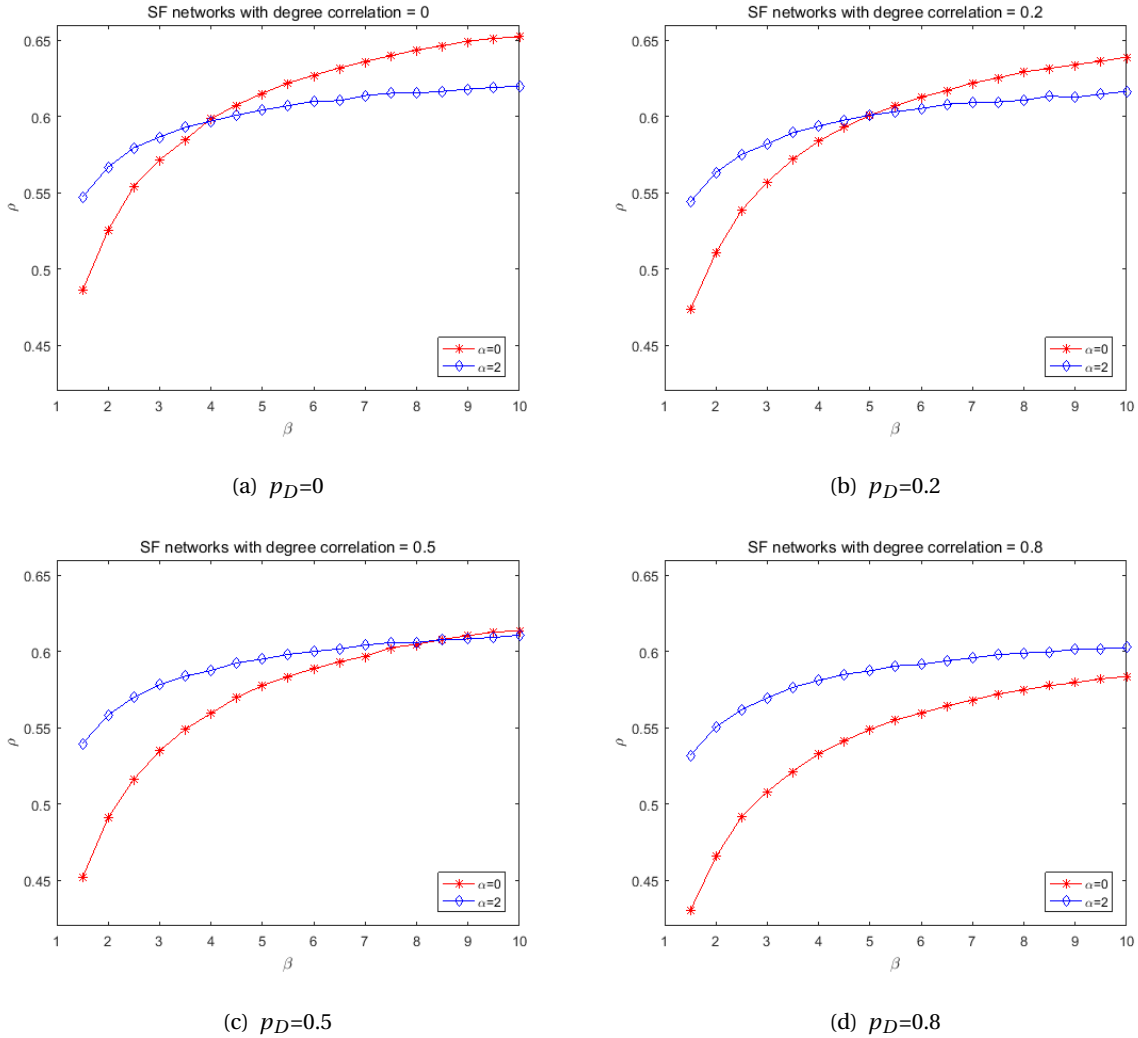


Figure 6.1: Effect of degree correlation  $p_D$  on the performance of SSIS in degree-correlated SF networks. RN infection rate  $\alpha=0$  and 2 respectively. The power-law exponent of both positive and negative networks is  $\gamma = 2.5$ .

The above interpretation of the observation in Fig. 6.1 can be further supported by results shown in Fig. 6.2. We respectively computed the infection probability of low-negative-degree ( $0 \leq d_i^- \leq 5$ ) and high-negative-degree nodes ( $d_i^- \geq 6$ ) at  $\beta=7$ . As shown in Fig. 6.2, when  $\alpha = 2$ , high-negative-degree nodes always have significantly high probability of being infected, regardless of  $p_D$ . Along with the increase in  $p_D$ , the infection probability of high-negative-degree nodes under  $\alpha = 0$  starts increasing, leading to the reduction in infection probability of nodes with small negative degree. It was found that about 90% of the nodes in our generated signed SF networks have low negative degree, and the infection probability of these nodes can, to a large extent, determine the overall  $\rho$ . Hence, in highly degree-correlated SF networks, viral spreadings tend to be relatively suppressed when negative links does not transmit virus, as less small-negative-degree can be infected under this condition.

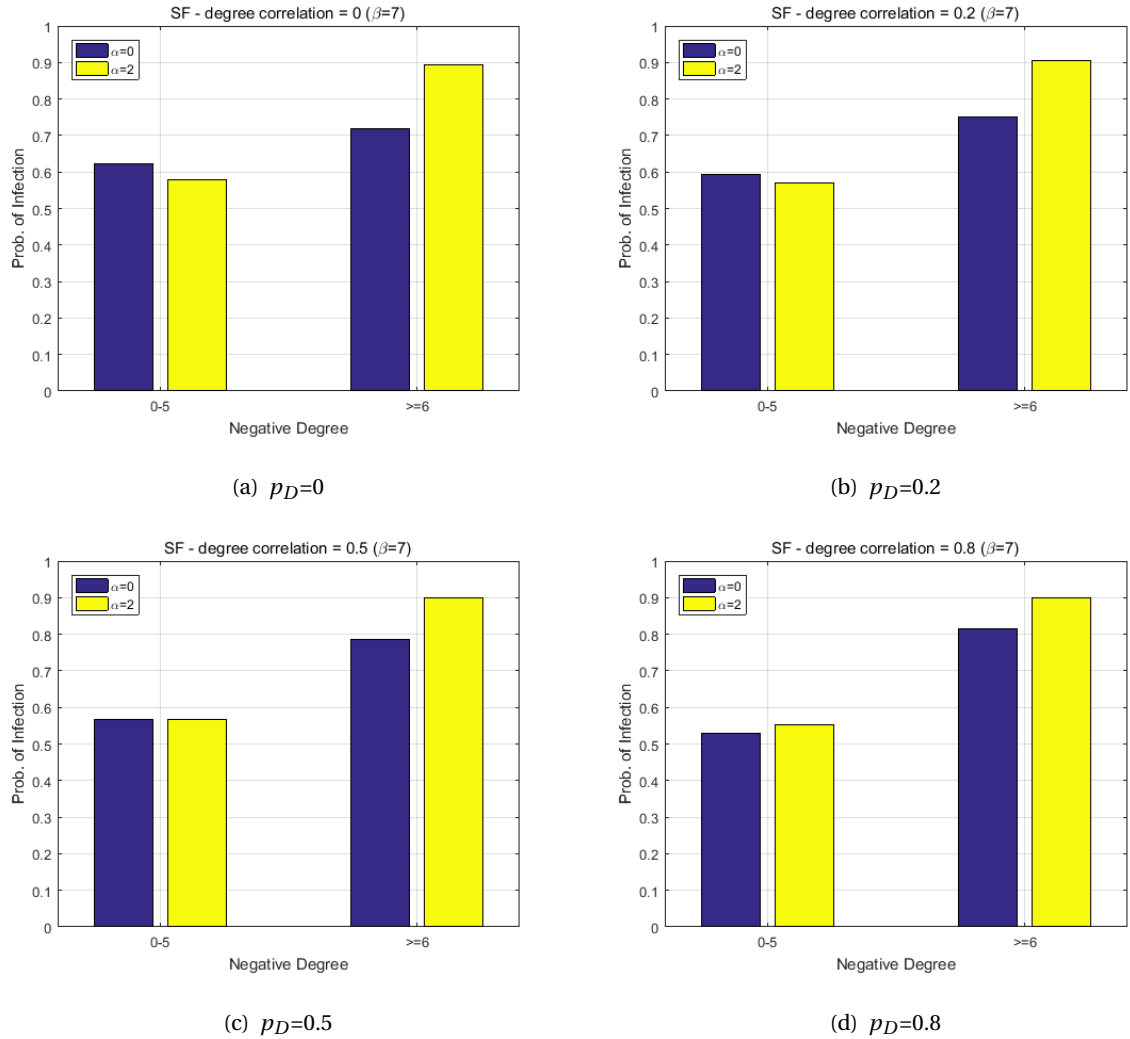


Figure 6.2: Infection probability of nodes with different degrees in degree-correlated SF networks

On the other hand, Fig. 6.3 shows that no intersection between  $\beta - \rho$  infection curves can be observed in degree-correlated signed ER networks within  $\beta \in [1.5, 10]$ , regardless of the value of  $p_D$ . About 80% of the nodes have negative degree between 0-5 in our generated signed ER networks. As shown in Fig. 6.4, although the infection probability of high-negative-degree nodes at  $\alpha = 0$  increases as  $p_D$  rises up, the infection probability of low-negative-degree nodes always maintains lower than  $\alpha > 0$  and hence  $\rho_{\alpha=0} < \rho_{\alpha>0}$  holds for all  $p_D$  presented in this figure. Taking the fact that the maximum negative degree in ER networks ( $\approx 12$ ) is much smaller than in SF networks ( $\approx 80$ ), it is not hard to understand that the infection of the so-called “high-negative-degree nodes” in ER networks appear to have very limited influence on the infection of low-negative-degree nodes.

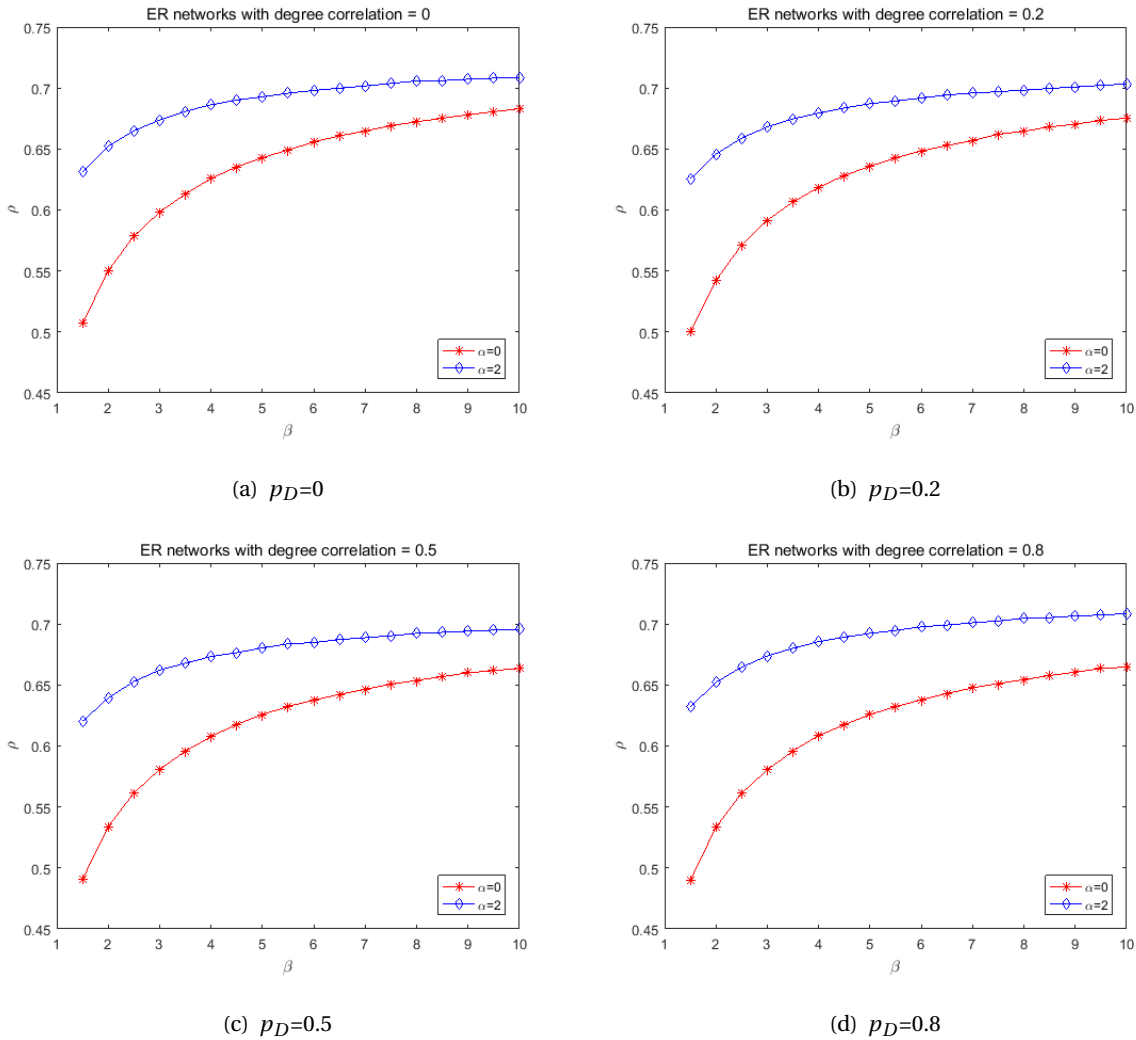


Figure 6.3: Infection of nodes in degree-correlated ER networks. RN infection rate  $\alpha=0$  and 2 respectively. The average degree of positive and negative networks are  $E[D_+] = E[D_-] = 4$ .

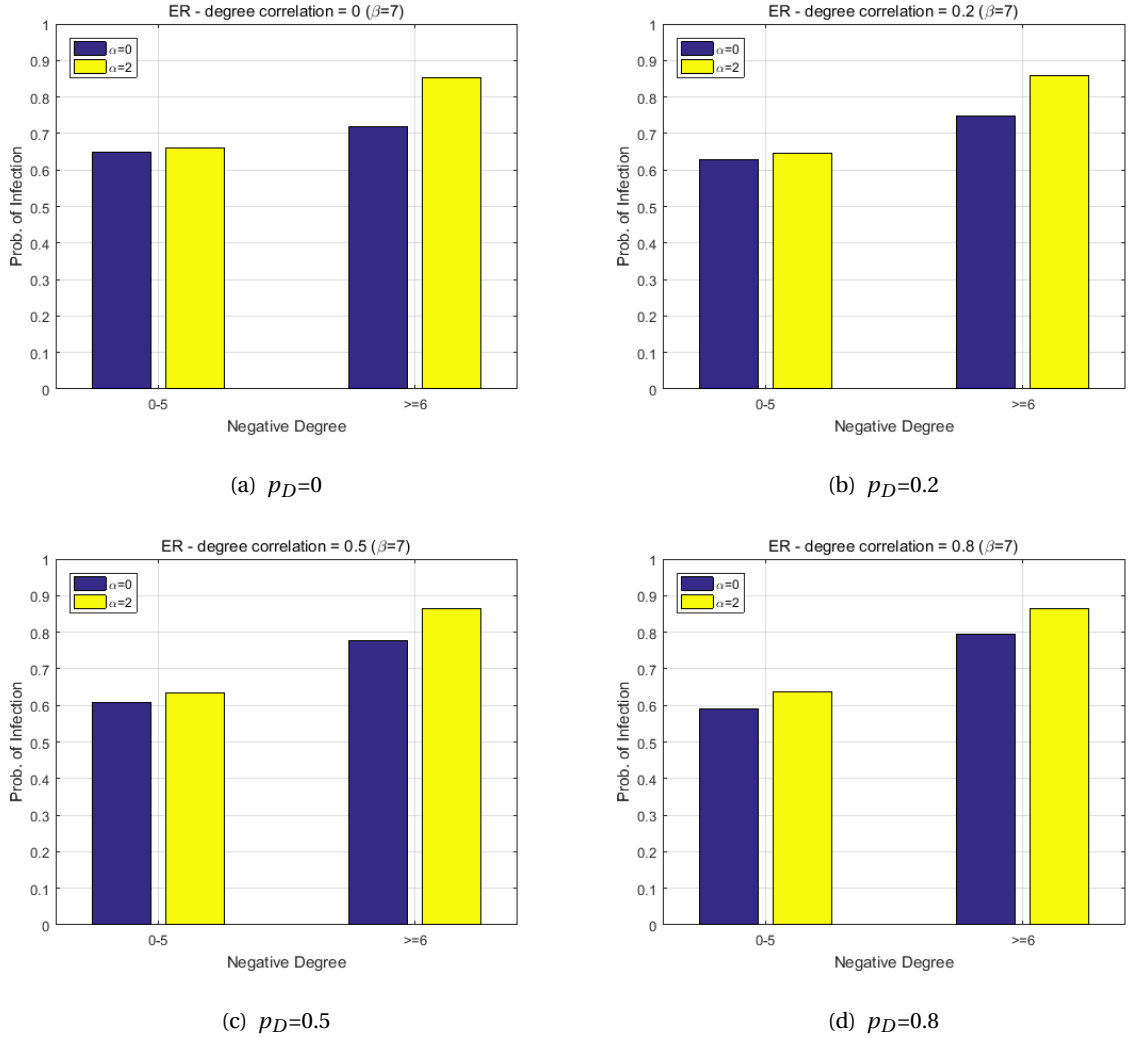


Figure 6.4: Infection probability of nodes with different degrees in degree-correlated ER networks

In a nutshell, at high infection rate  $\beta$ , degree correlation is a main factor that can determine whether the participant of negative links in virus propagation may encourage or suppress the overall spreading in degree-correlated signed SF networks. In networks of small  $p_D$ , the overall viral spreading can be encouraged at high  $\beta$  if negative links are not allowed to transmit virus (i.e.  $\alpha = 0$ ); While in highly degree-correlated networks,  $\alpha = 0$  tends to always result in relative suppression of viral spreadings, comparing to the case when virus can propagate via negative links. As to degree-correlated signed ER networks, due to the relative homogenous degree distribution, a higher  $\alpha$  tends to result in a larger  $\rho$  at any  $\beta$ , regardless of  $p_D$ .



# 7

## SSIS MODEL IN REAL-WORLD SIGNED NETWORKS

In this chapter we aim at verifying the previously discovered properties of SSIS model in two real-world signed networks. We are interested in the influence of RN infection rate  $\alpha$  and the degree correlation  $p_D$  on viral spreading in real-world networks. The simulation result indicates that the performance of our SSIS model in real-world networks completely agree with conclusions that drawn in Chapter 5 and 6.

### 7.1. DESCRIPTION OF SELECTED REAL-WORLD NETWORKS

The first considered real-world network is the voting network of Wikipedia (Wiki network) [35], where a signed link represents one user's support (positive) or objection (negative) to another user being promoted to an administrator. The original Wiki networks is directed, and we take the following measures to remove link directions and simplify the network. First, we check the number of directed links between every two nodes: If there is only one directed link between two nodes, then

we simply remove the direction and keep its original sign; If there are two directed link between two nodes, then we further check the signs of these two links. One of these two link would be kept and changed to be undirected if the signs are the same; otherwise both links would be deleted. Moreover, there also exist several neutral links in the Wiki network and they are all discarded.

The second real-world network is a product relationship network based on Amazon dataset (Amazon network) [69]. Relationships between products are modelled and predicted by the SCEPTRE (Substitute and Complementary Edges between Products from Topics in REviews) system, based on reviews and descriptions of products. A positive link denotes the complementary relation between two products, i.e., these might be viewed/purchased together; and a negative link stands for the substitutable relation between two products, which are interchangeable - such as one jacket for another. In [69] several product relationship networks based on different categories of products were provided, and we select the one that based on baby product database.

The basic topological features of the two networks are summarised in Table 7.1. Both networks have approximately hundreds of thousand links, and thousands of nodes. In Wiki network roughly 80% of links are positive and in Amazon network the positive-negative proportion is about 50%-50%. The average degrees of all networks are around 20, except the negative network of Wiki, of which the average degree is 10. The logarithmic degree distributions of all networks are given in Fig. 7.1. It can be seen that the degree distributions of all networks approximately follow power law.

Table 7.1: Topological features of real networks

Network	Nodes	Links	“+”Links	“-”Links	$E[D_-]$	$E[D_+]$	$p_D$
Wiki	7066	100390	78479(78.2%)	21911(21.8%)	25	10	0.63
Amazon	4217	77242	37769(48.9%)	39473(51.1%)	19	19	0.39

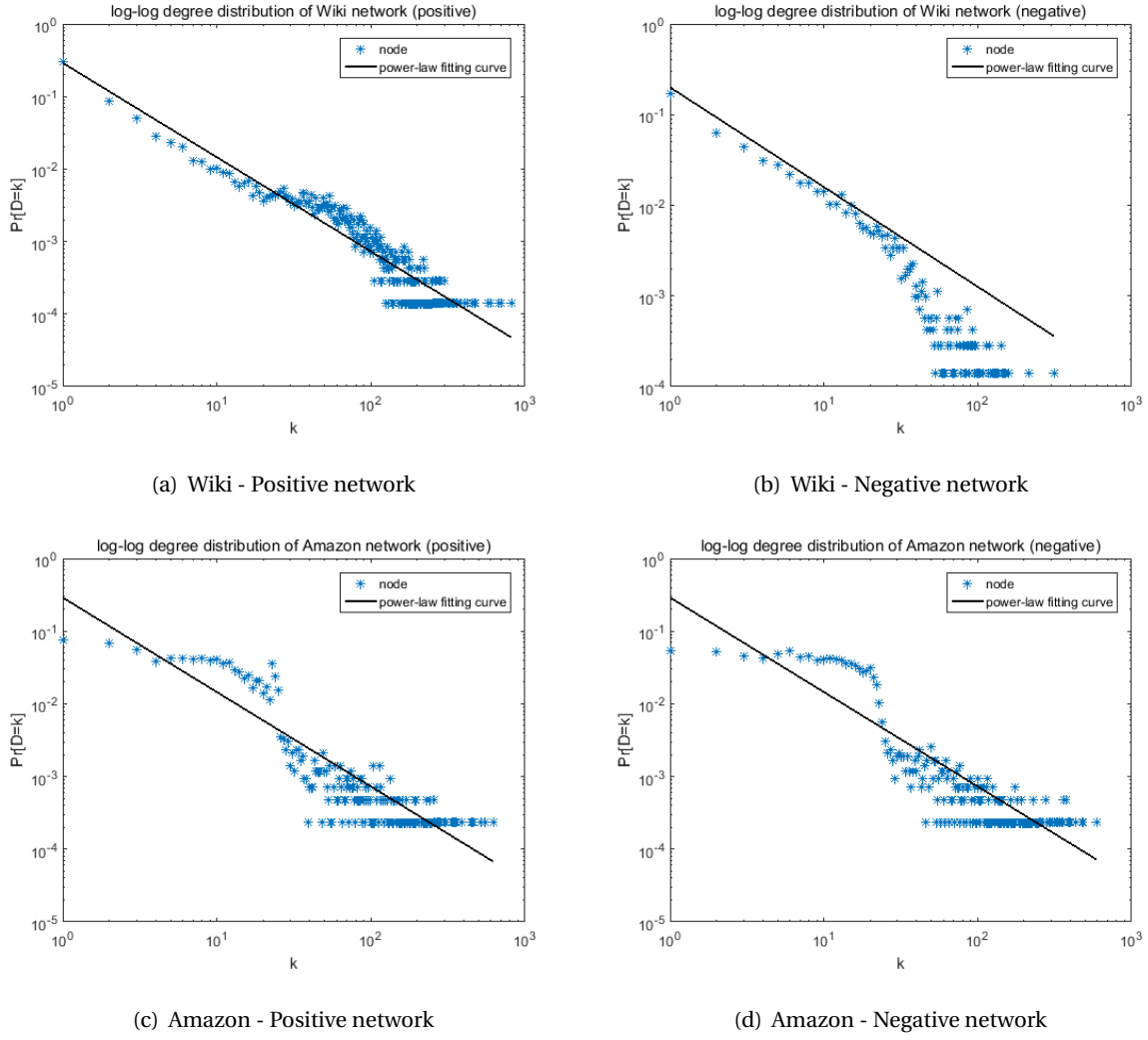


Figure 7.1: Degree distribution of signed real-world networks.

## 7.2. PERFORMANCE OF SSIS MODEL IN REAL-WORLD NETWORKS

For real-world networks, we are interested in the following four questions: (1) Does a higher RN infection rate  $\alpha$  lead to a lower epidemic threshold  $\tau_c$ ? (2) Does a higher  $\alpha$  tend to lead to a larger meta-stable fraction of infected nodes  $\rho$  at any infection rate? (3) What is the influence of degree correlation  $p_D$  on  $\rho$  with different  $\alpha$ ? (4) Do signed-unsigned crossing points  $(\beta^*, \rho^*)$  still exist for SSIS real-world networks?

First, we focus on the influence of  $\alpha$  on viral spreadings and try to verify the existence of signed-unsigned crossing points. As discussed in Chapter 5 and 6, in SSIS model the epidemic threshold

of signed SF networks is reduced as  $\alpha$  increases. For networks with power-law degree distributions, there always exist non-zero epidemic thresholds as long as the sizes of networks are finite [11]. In addition, signed-unsigned crossing points  $(\beta^*, \rho^*)$  can be observed in both degree-independent and degree-correlated signed SF networks. Hence, it is expected that such influence of  $\alpha$  on epidemic threshold may be observed, and crossing points between  $\beta - \rho$  infection curves of signed and unsigned networks could also be detected in real-world signed networks. As shown in Fig. 7.2, results of both MC and IBMFA methods confirm the existence of crossing points in both Wiki and Amazon networks. The unsigned networks are obtained by removing all negative links in each of the real-world negative networks. It is also observed that in Amazon network  $\tau_c$  obviously reduces along with the increase in  $\alpha$ ; While in Wiki network, precise epidemic thresholds of different  $\alpha$  can hardly be seen, possibly due to the high dense of links in this network.

Next, we are going to explore the influence of  $p_D$  on the epidemiological performance of SSIS model. It has been observed in Chapter 6 that: (i) In SF networks of small degree correlation, at small infection rate  $\beta$ , we can observe a higher  $\rho$  under a larger  $\alpha$ , i.e.  $\rho_{\alpha_1} < \rho_{\alpha_2}$  where  $\alpha_1 < \alpha_2$ ; However, as  $\beta$  increases,  $\rho_{\alpha_1}$  starts taking the overwhelming position. (ii) In highly degree-correlated SF networks, it appears that at any  $\beta$  we have  $\rho_{\alpha_1} < \rho_{\alpha_2}$  for  $\alpha_1 < \alpha_2$ . Since both Wiki and Amazon networks have power-law degree distribution, we expect that the viral spreading performance in Amazon network ( $p_D = 0.39$ ) follows observation (i); while observation (ii) holds in Wiki network ( $p_D = 0.63$ ). Results shown in Fig. 7.3 are completely consistent with our expectation.

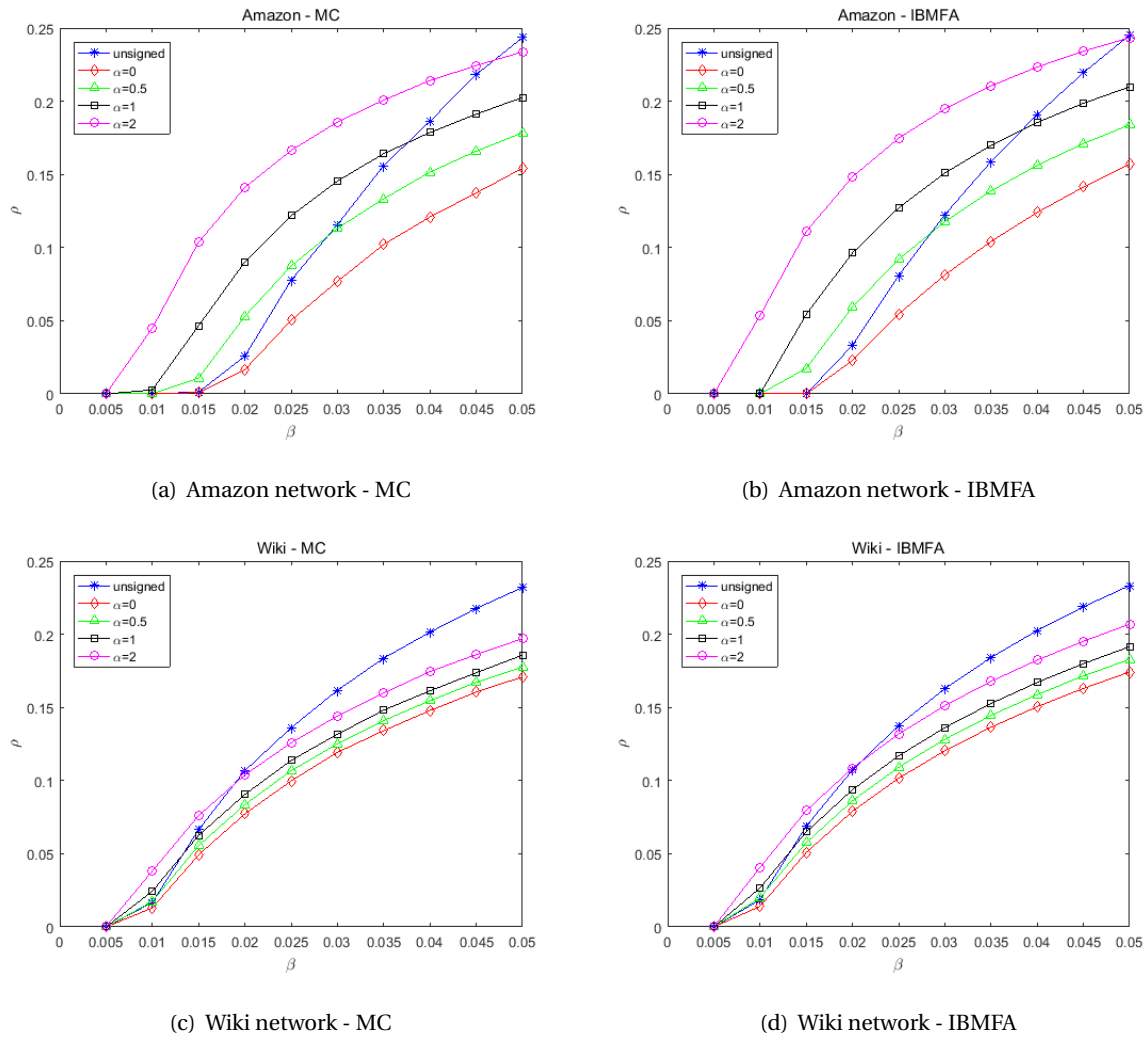


Figure 7.2: Influence of  $\alpha$  on the performance of SSIS in real-world networks.  $\alpha=0, 0.5, 1$  and  $2$  respectively. The unsigned networks are obtained by removing all negative links in each of the real-world negative networks.

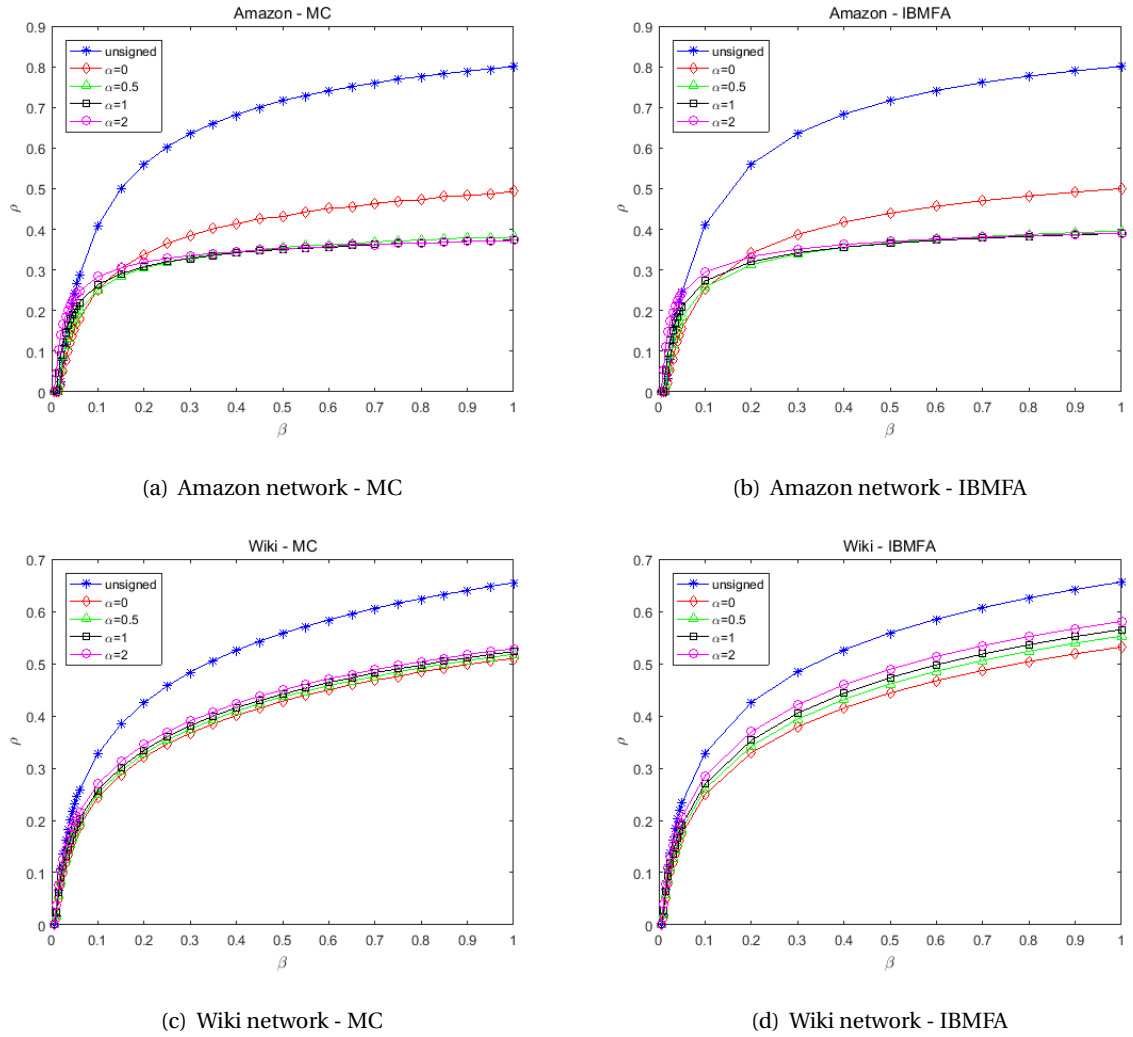


Figure 7.3: Influence of  $p_D$  on the performance of SSIS in real-world networks.  $\alpha=0, 0.5, 1$  and  $2$  respectively. The unsigned networks are obtained by removing all negative links in each of the real-world negative networks.

# 8

## SUMMARY AND FUTURE WORKS

### 8.1. SUMMARY

In this thesis, we propose a SSIS model to capture the influence of negative links on viral propagation in signed networks. In SSIS model, the dynamic infection rate of each node is determined by the infection status of its negative neighbours. We suppose that both positive and negative neighbours would participate in transmitting virus, while the infection rate via positive/negative links of a node reduces as more of its negative neighbours have been infected.

Our focus is to understand how can negative links influence the epidemiological performance in signed networks, with respect to epidemic threshold  $\tau_c$  and meta-stable fraction of infected nodes  $\rho$ . We have found that three main factors can substantially contribute to different epidemic performance of SSIS: the relative infection rate of negative links with respect to positive links ( $\alpha$ ), the degree correlation between positive and negative networks ( $p_D$ ) and degree distributions of the positive and negative links respectively (signed ER network vs. signed SF network). We develop Monte-Carlo simulations to observe the performance of SSIS in different scenarios. In addition, we

also derive the IBMFA method to theoretically analyse the SSIS model, based on the solutions of meta-stable infection probability of nodes. We show that IBMFA can in general accurately approximate the MC simulation results.

According to our simulation results, we draw the following main conclusions.

1. In signed ER networks, a larger  $\alpha$  tends to result in a greater meta-stable fraction of infected nodes at any infection rate  $\beta$ ; While in degree-independent signed SF networks, at large  $\beta$ , viral propagation via negative links may in turn lead to a smaller  $\rho$ . The main reason is given as follows. When the degree correlation  $p_D$  is small in signed SF networks, when  $\alpha > 0$  the hub nodes are usually highly likely to be infected at large  $\beta$ , leading to low probabilities of infection of the rest nodes (which are the majority of entire nodes) in general. However, when negative links are not allowed to transmit virus ( $\alpha = 0$ ), since negative hubs usually have small positive degrees in degree-independent signed SF networks, they are less likely being infected and thus more of the rest nodes can be infected in meta-stable state. As  $p_D$  rises up, negative hubs become more likely to have bigger positive degree, and in general they can be infected with high probability with any  $\alpha$ . As a result, higher  $\rho$  might be achieved only with higher  $\alpha$  at any  $\beta$ . Since the degree distribution is relatively homogeneous in signed ER network, regardless of  $\beta$ , a higher meta-stable fraction of infected nodes can hardly be obtained when  $\alpha = 0$ .
2. Compared to networks with only positive relations, the viral propagation via negative links can lead to a lower epidemic threshold in signed networks. When  $\beta$  is smaller than a particular rate  $\beta^*$ , viral propagation via negative links can facilitate a higher  $\rho$  in signed networks than in unsigned networks (i.e.  $\rho_{signed} > \rho_{unsigned}$  when  $\beta < \beta^*$ ). However,  $\rho_{unsigned}$  becomes superior as  $\beta > \beta^*$ . As a result, crossing point  $(\beta^*, \rho^*)$  can be observed between  $\beta$ - $\rho$  curves of signed and unsigned networks, where  $\rho_{signed} = \rho_{unsigned} = \rho^*$ . We have found that  $\beta^*$  is approximately linearly related to  $\alpha$ .

With regard to the real-world scenario on the technique adoptions of companies mentioned in Chapter 1, these above conclusions indicate that: (1) Even when every company adopts a new technique at a higher rate in response to the utilisations of its competitors (reflected by a higher  $\alpha$ ), it would be not always possible to have a larger percentage of companies adopting a new technique; (2) Compared to networks with only cooperative relations (unsigned networks), the competitive relations between companies may sometimes result in a higher percentage of adoption among com-



panies, especially when a company accepts a new technique at a small rate.

## 8.2. FUTURE WORKS

Taking the possible development of SSIS model into account, some suggestions on future works are given as follows:

- It has been found that IBMFA does not give a perfect approximation of SF signed networks at high  $\beta$  (See Fig. A.5). Therefore, an advanced IBMFA method which is particularly designed for SF networks may be required.
- The dynamic infection rate  $\beta^+(t)$  and  $\beta^-(t)$  we have proposed in Eq.(4.1) and Eq.(4.2) may be optimised in future works in order to better characterise the influence of negative links on the overall viral spreadings in real world.
- The proposed viral propagation model might be extended to the field of SIR model, where the immunisation processes are introduced to remove infected nodes permanently from the virus (disease).



# A

## APPENDIX

### A.1. COMPARISON BETWEEN MC AND IBMFA RESULTS IN DEGREE-INDEPENDENT SIGNED NETWORKS

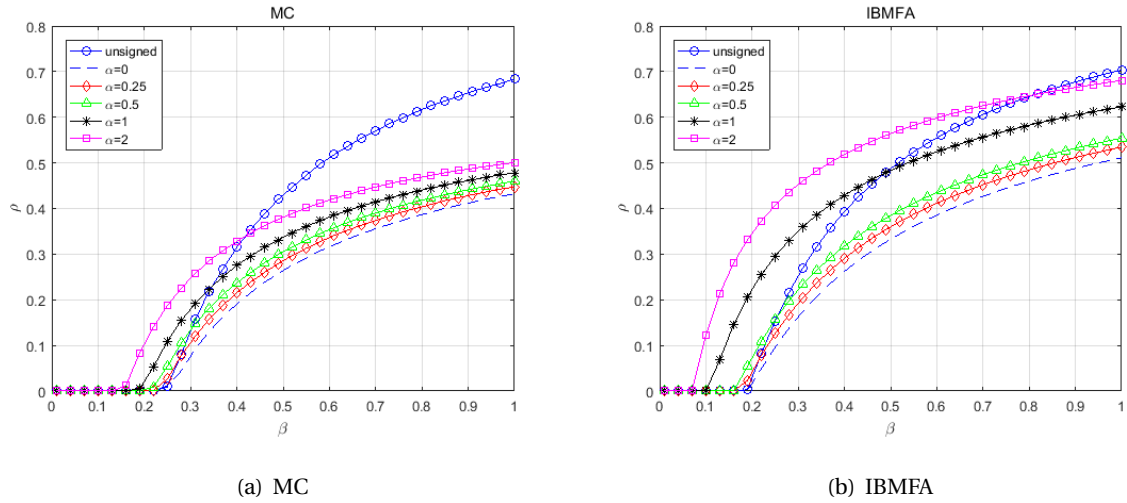


Figure A.1: Influence of  $\alpha$  on performance of SSIS model in signed degree-independent ER networks. Network size  $N=1000$  nodes. Average positive degree is  $E[D_+] = 4$ , and average negative degree is  $E[D_-] = 2$ . The unsigned network is obtained by removing all negative links in negative networks.  $\alpha=0, 0.25, 0.5, 1$  and  $2$  respectively.

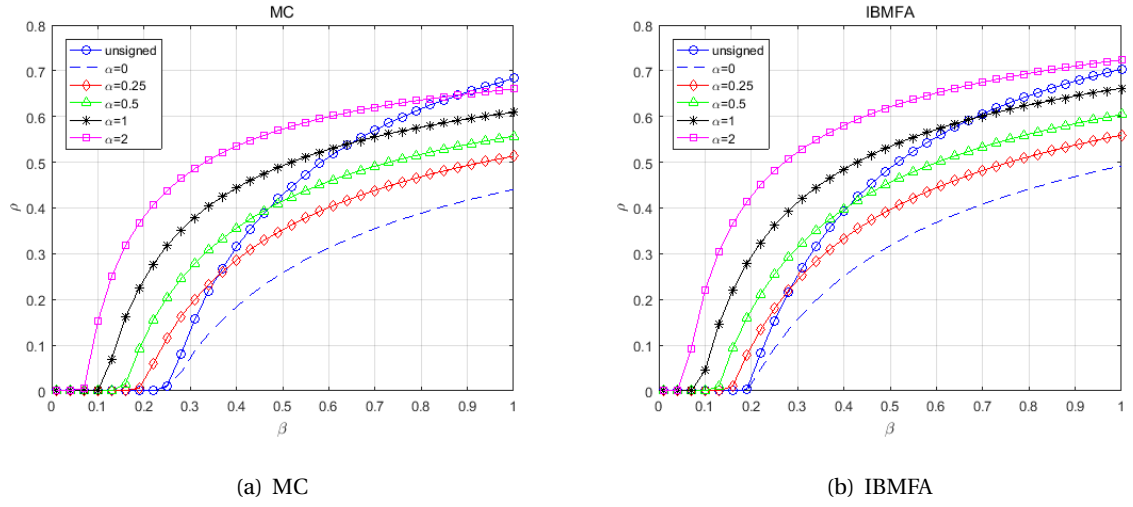


Figure A.2: Influence of  $\alpha$  on performance of SSIS model in signed degree-independent ER networks. Network size  $N=1000$  nodes. Average positive degree is  $E[D_+] = 4$ , and average negative degree is  $E[D_-] = 6$ . The unsigned network is obtained by removing all negative links in negative networks.  $\alpha=0, 0.25, 0.5, 1$  and  $2$  respectively

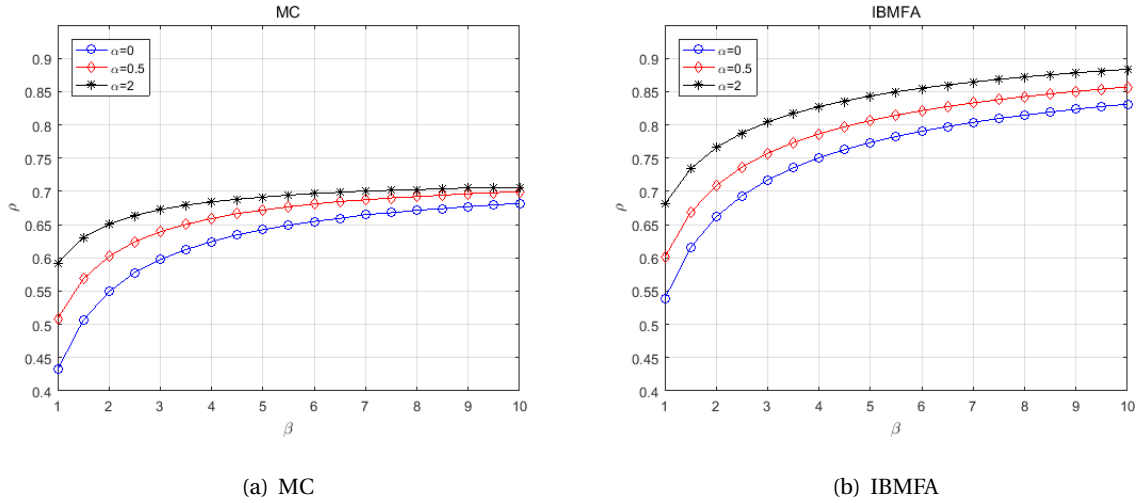


Figure A.3: Influence of  $\alpha$  on performance of SSIS model in signed degree-independent ER networks. Network size  $N=1000$  nodes. Average positive degree is  $E[D_+] = 4$ , and average negative degree is  $E[D_-] = 4$ .  $\alpha=0, 0.5$  and  $2$  respectively.

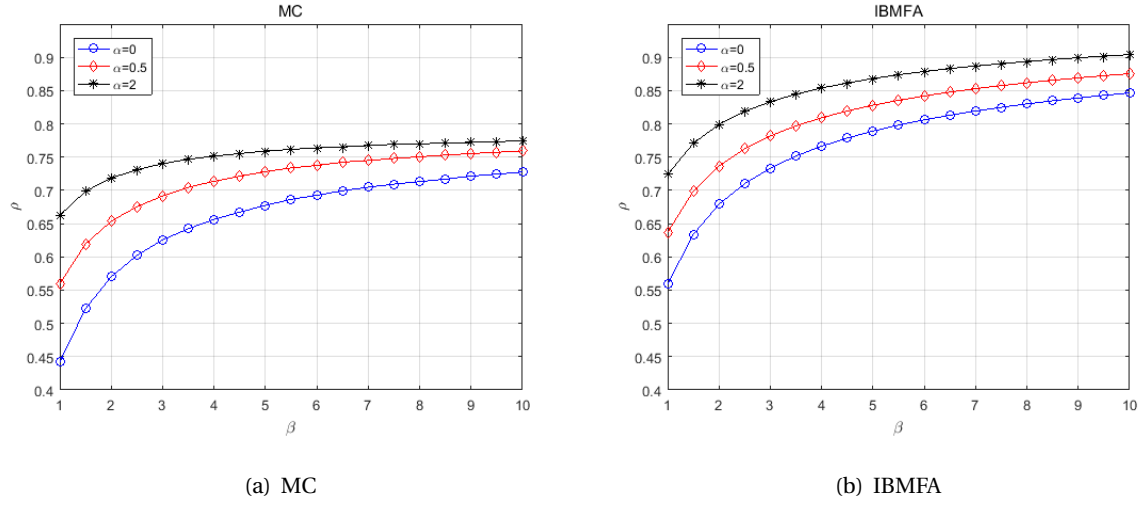


Figure A.4: Influence of  $\alpha$  on performance of SSIS model in signed degree-independent ER networks. Network size  $N=1000$  nodes. Average positive degree is  $E[D_+] = 4$ , and average negative degree is  $E[D_-] = 6$ .  $\alpha=0, 0.5$  and  $2$  respectively.

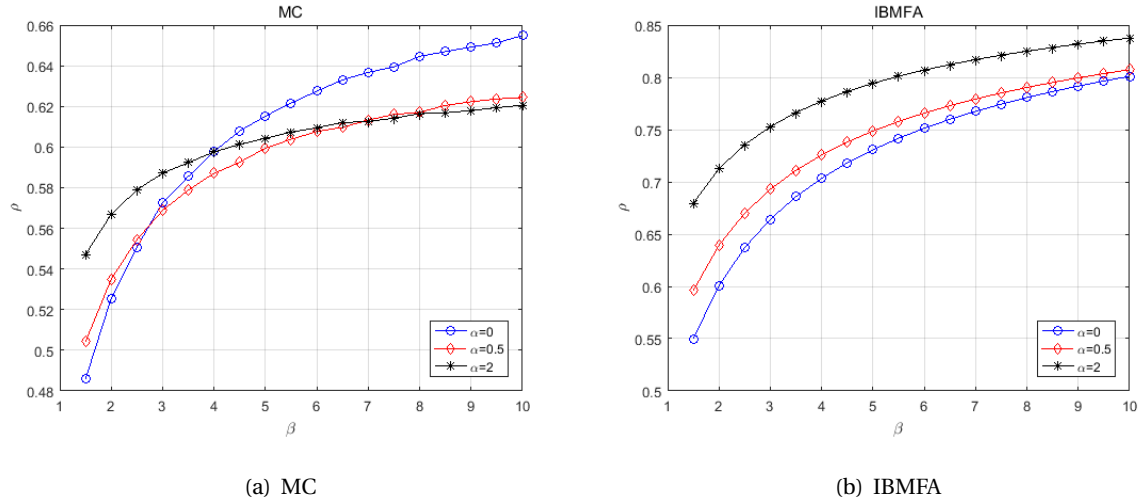


Figure A.5: Influence of  $\alpha$  on performance of SSIS model in signed degree-independent SF networks. Network size  $N=1000$  nodes. The power law exponent in both positive and negative networks is  $\gamma = 2.5$ .  $\alpha=0, 0.5$  and  $2$  respectively.

## A.2. SOME MC RESULTS IN DEGREE-CORRELATION NETWORKS

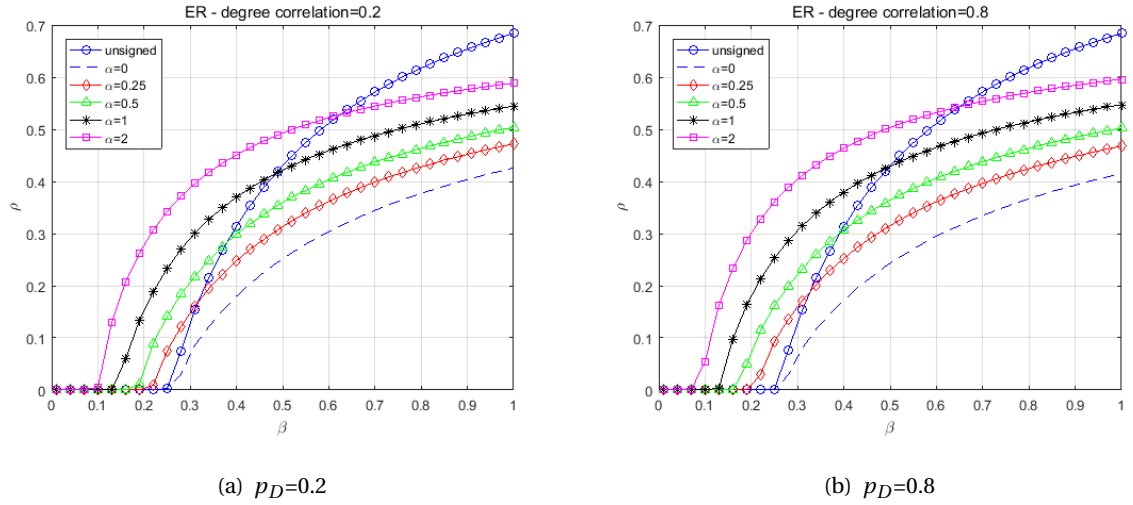


Figure A.6: Influence of  $\alpha$  on performance of SIS model in signed degree-correlated ER networks. Network size  $N=1000$  nodes. Average positive degree is  $E[D_+] = 4$ , and average negative degree is  $E[D_-] = 4$ . The unsigned network is obtained by removing all negative links in negative networks.  $\alpha=0, 0.25, 0.5, 1$  and  $2$  respectively

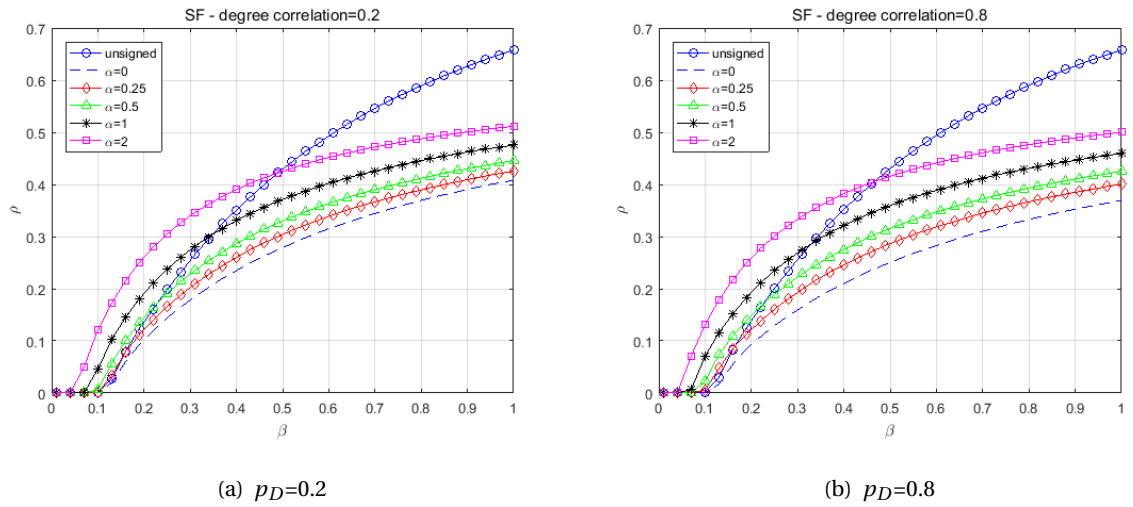


Figure A.7: Influence of  $\alpha$  on performance of SIS model in signed degree-correlated SF networks. Network size  $N=1000$  nodes. The power law exponent in both positive and negative networks is  $\gamma = 2.5$ .  $\alpha=0, 0.25, 0.5, 1$  and  $2$  respectively.

# ACKNOWLEDGEMENT

My master thesis is done in the Multimedia Computing Group, Faculty of Electrical Engineering, Mathematics, and Computer Science (EEMCS), Delft University of Technology.

First of all, I would like to give my heartfelt gratitude to my daily supervisor Dr. Huijuan Wang. As a very supportive supervisor, she did not only help me to tackle research bottlenecks, but also spent much time and effort on giving many constructive comments on my thesis work. She has great comprehensive knowledge in this field, and without her guidance this thesis would never be a success.

I would like to thank my responsible professor Prof. Dr. Alan Hanjalic for joining my thesis defense committee and taking time to read my thesis.

I would like to thank Dr. Marco Zúñiga from the Embedded Software Group in TU Delft for his willingness to be the committee member of my thesis defense.

I would like to thank Mr. Chuyi Chen, my alumni in TU Delft, for his warm-hearted and selfless help with some problems related to my thesis.

I would like to thank Mr. Bo Qu from Multimedia Computing Group for his constructive advices on some questions.

I would like to thank Mr. Hao Liu, my classmate in high school, for his suggestions in some of programming works.

Finally, I would like to express my great appreciation to my family for their wholehearted support to my study and my life, to all my friends for their encouragement, and to my girlfriend for her encouragement, support and love.

*Boning LI*

*Delft, August 2016*



## BIBLIOGRAPHY

- [1] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, *Epidemic spreading in real networks: An eigenvalue viewpoint*, in *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on* (IEEE, 2003) pp. 25–34.
- [2] A. Khelil, C. Becker, J. Tian, and K. Rothermel, *An epidemic model for information diffusion in manets*, in *Proceedings of the 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems* (ACM, 2002) pp. 54–60.
- [3] J. O. Kephart and S. R. White, *Directed-graph epidemiological models of computer viruses*, in *Research in Security and Privacy, 1991. Proceedings., 1991 IEEE Computer Society Symposium on* (IEEE, 1991) pp. 343–359.
- [4] N. T. J. Bailey, *The mathematical theory of infectious diseases and its applications* (Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975).
- [5] A. M’Kendrick, *Applications of mathematics to medical problems*, *Proceedings of the Edinburgh Mathematical Society* **44**, 98 (1925).
- [6] M. E. Newman, *Spread of epidemic disease on networks*, *Physical review E* **66**, 016128 (2002).
- [7] P. Domingos and M. Richardson, *Mining the network value of customers*, in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM, 2001) pp. 57–66.
- [8] M. Faloutsos, P. Faloutsos, and C. Faloutsos, *On power-law relationships of the internet topology*, in *ACM SIGCOMM computer communication review*, Vol. 29 (ACM, 1999) pp. 251–262.
- [9] M. Ripeanu, I. Foster, and A. Iamnitchi, *Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design*, arXiv preprint cs/0209028 (2002).
- [10] R. Pastor-Satorras and A. Vespignani, *Epidemic spreading in scale-free networks*, *Physical review letters* **86**, 3200 (2001).

- [11] R. Pastor-Satorras and A. Vespignani, *Epidemics and immunization in scale-free networks*, arXiv preprint cond-mat/0205260 (2002).
- [12] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, *Trawling the web for emerging cyber-communities*, Computer networks **31**, 1481 (1999).
- [13] P. Van Mieghem, *The  $n$ -intertwined sis epidemic network model*, Computing **93**, 147 (2011).
- [14] C. Li, H. Wang, and P. Van Mieghem, *Epidemic threshold in directed networks*, Physical Review E **88**, 062802 (2013).
- [15] D. Li, P. Qin, H. Wang, C. Liu, and Y. Jiang, *Epidemics on interconnected lattices*, EPL (Europhysics Letters) **105**, 68004 (2014).
- [16] M. Liu, D. Li, P. Qin, C. Liu, H. Wang, and F. Wang, *Epidemics in interconnected small-world networks*, PloS one **10**, e0120701 (2015).
- [17] Z. Yang and T. Zhou, *Epidemic spreading in weighted networks: an edge-based mean-field solution*, Physical Review E **85**, 056106 (2012).
- [18] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. Pappas, *Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks*, in *52nd IEEE Conference on Decision and Control* (IEEE, 2013) pp. 7486–7491.
- [19] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas, *Optimal resource allocation for network protection against spreading processes*, IEEE Transactions on Control of Network Systems **1**, 99 (2014).
- [20] B. Qu, A. Hanjalic, and H. Wang, *Heterogeneous recovery rates against sis epidemics in directed networks*, arXiv preprint arXiv:1408.6959 (2014).
- [21] B. Qu and H. Wang, *Sis epidemic spreading with heterogeneous infection rates*, arXiv preprint arXiv:1506.07293 (2015).
- [22] H. Wang, J. M. Hernandez, and P. Van Mieghem, *Betweenness centrality in a weighted network*, Physical Review E **77**, 046105 (2008).
- [23] C. Li, H. Wang, W. De Haan, C. Stam, and P. Van Mieghem, *The correlation of metrics in complex networks with applications in functional brain networks*, Journal of Statistical Mechanics: Theory and Experiment **2011**, P11018 (2011).

- [24] C. Li, Q. Li, P. Van Mieghem, H. E. Stanley, and H. Wang, *Correlation between centrality metrics and their application to the opinion model*, The European Physical Journal B **88**, 1 (2015).
- [25] D. Guo, S. Trajanovski, R. van de Bovenkamp, H. Wang, and P. Van Mieghem, *Epidemic threshold and topological structure of susceptible-infectious-susceptible epidemics in adaptive networks*, Physical Review E **88**, 042802 (2013).
- [26] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, *Theory of rumour spreading in complex social networks*, Physica A: Statistical Mechanics and its Applications **374**, 457 (2007).
- [27] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks* (Cambridge University Press, 2008).
- [28] C. Castellano, S. Fortunato, and V. Loreto, *Statistical physics of social dynamics*, Reviews of modern physics **81**, 591 (2009).
- [29] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Epidemic processes in complex networks*, Reviews of modern physics **87**, 925 (2015).
- [30] J. Leskovec, D. Huttenlocher, and J. Kleinberg, *Signed networks in social media*, in *Proceedings of the SIGCHI conference on human factors in computing systems* (ACM, 2010) pp. 1361–1370.
- [31] J. Leskovec, D. Huttenlocher, and J. Kleinberg, *Predicting positive and negative links in online social networks*, in *Proceedings of the 19th international conference on World wide web* (ACM, 2010) pp. 641–650.
- [32] J. Tang, Y. Chang, C. Aggarwal, and H. Liu, *A survey of signed network mining in social media*, arXiv preprint arXiv:1511.07569 (2015).
- [33] P. Massa and P. Avesani, *Controversial users demand local trust metrics: An experimental study on epinions. com community*, in *Proceedings of the National Conference on artificial Intelligence*, Vol. 20 (Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005) p. 121.
- [34] C. A. Lampe, E. Johnston, and P. Resnick, *Follow the reader: filtering comments on slashdot*, in *Proceedings of the SIGCHI conference on Human factors in computing systems* (ACM, 2007) pp. 1253–1262.
- [35] M. Burke and R. Kraut, *Mopping up: modeling wikipedia promotion decisions*, in *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (ACM, 2008) pp. 27–36.

- [36] F. Heider, *Attitudes and cognitive organization*, The Journal of psychology **21**, 107 (1946).
- [37] T. Antal, P. L. Krapivsky, and S. Redner, *Social balance on networks: The dynamics of friendship and enmity*, Physica D: Nonlinear Phenomena **224**, 130 (2006).
- [38] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, *Spectral analysis of signed graphs for clustering, prediction and visualization*. in *SDM*, Vol. 10 (SIAM, 2010) pp. 559–559.
- [39] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, *Epidemic thresholds in real networks*, ACM Transactions on Information and System Security (TISSEC) **10**, 1 (2008).
- [40] P. Erdős and A. Rényi, *On random graphs*, Publicationes Mathematicae Debrecen **6**, 290 (1959).
- [41] P. Erdős and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci **5**, 17 (1960).
- [42] B. Bollobás, *Random graphs* (Springer, 1998).
- [43] A.-L. Barabási and J. Frangos, *Linked: the new science of networks science of networks* (Basic Books, 2014).
- [44] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286**, 509 (1999).
- [45] D. J. Watts and S. H. Strogatz, *Collective dynamics of small-world networks*, nature **393**, 440 (1998).
- [46] R. Cohen and S. Havlin, *Scale-free networks are ultrasmall*, Physical review letters **90**, 058701 (2003).
- [47] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, *Generation of uncorrelated random scale-free networks*, Physical Review E **71**, 027103 (2005).
- [48] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman, *Search in power-law networks*, Physical review E **64**, 046135 (2001).
- [49] E. A. Bender and E. R. Canfield, *The asymptotic number of labeled graphs with given degree sequences*, Journal of Combinatorial Theory, Series A **24**, 296 (1978).
- [50] E. Quill, *When networks network: once studied solo, systems display surprising behavior when they interact*, Science News **182**, 18 (2012).

- [51] G. Baxter, S. Dorogovtsev, A. Goltsev, and J. Mendes, *Avalanche collapse of interdependent networks*, Physical review letters **109**, 248701 (2012).
- [52] M. Szell, R. Lambiotte, and S. Thurner, *Multirelational organization of large-scale social networks in an online world*, Proceedings of the National Academy of Sciences **107**, 13636 (2010).
- [53] G. Bianconi, *Statistical mechanics of multiplex networks: Entropy and overlap*, Physical Review E **87**, 062806 (2013).
- [54] P. Van Mieghem, J. Omic, and R. Kooij, *Virus spread in networks*, Networking, IEEE/ACM Transactions on **17**, 1 (2009).
- [55] E. Cator and P. Van Mieghem, *Susceptible-infected-susceptible epidemics on the complete graph and the star graph: Exact analysis*, Physical Review E **87**, 012811 (2013).
- [56] C. Li, R. van de Bovenkamp, and P. Van Mieghem, *Susceptible-infected-susceptible model: A comparison of  $n$ -intertwined and heterogeneous mean-field approximations*, Physical Review E **86**, 026116 (2012).
- [57] D. J. Daley, J. Gani, and J. M. Gani, *Epidemic modelling: an introduction*, Vol. 15 (Cambridge University Press, 2001).
- [58] J. Marro and R. Dickman, *Nonequilibrium phase transitions in lattice models* (Cambridge University Press, 2005).
- [59] P. Van Mieghem, *Performance analysis of complex networks and systems* (Cambridge University Press, 2014).
- [60] C. Castellano and R. Pastor-Satorras, *Thresholds for epidemic spreading in networks*, Physical review letters **105**, 218701 (2010).
- [61] A. Ganesh, L. Massoulié, and D. Towsley, *The effect of network topology on the spread of epidemics*, in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, Vol. 2 (IEEE, 2005) pp. 1455–1466.
- [62] T. Hotta, *Mean-field approximation*, in *Nanyscale Phase Separation and Colossal Magnetoresistance* (Springer, 2003) pp. 157–167.
- [63] P. Van Mieghem, *Performance analysis of communications networks and systems* (Cambridge University Press, 2006).

- [64] H. Wang, Q. Li, G. D'Agostino, S. Havlin, H. E. Stanley, and P. Van Mieghem, *Effect of the interconnected network structure on the epidemic threshold*, Physical Review E **88**, 022801 (2013).
- [65] C. Paige, *Eigenvalues of perturbed hermitian matrices*, Linear Algebra and its Applications **8**, 1 (1974).
- [66] I. C. Ipsen and B. Nadler, *Refined perturbation bounds for eigenvalues of hermitian and non-hermitian matrices*, SIAM Journal on Matrix Analysis and Applications **31**, 40 (2009).
- [67] M. Boguñá, C. Castellano, and R. Pastor-Satorras, *Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks*, Physical review letters **111**, 068701 (2013).
- [68] B. Qu, Q. Li, S. Havlin, H. E. Stanley, and H. Wang, *Nonconsensus opinion model on directed networks*, Physical Review E **90**, 052811 (2014).
- [69] J. McAuley, R. Pandey, and J. Leskovec, *Inferring networks of substitutable and complementary products*, in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2015) pp. 785–794.