

BOSTON UNIVERSITY  
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

**DYNAMIC AND INTERACTING COMPLEX NETWORKS**

by

**MARK E. DICKISON**

B.A., University of Missouri at Rolla  
M.S., University of Missouri at Rolla

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

2012

Approved by

First Reader

---

H. Eugene Stanley, Ph.D.  
University Professor and Professor of Physics

Second Reader

---

Karl Ludwig, Ph.D.  
Professor of Physics

*To my parents;*

*I would not be here, or anywhere, without you both.*

# Acknowledgments

Firstly, I want to thank my advisor H. Eugene Stanley, for his support during my years with him. Gene has been an inspiring figure and one I'm grateful to have had as my advisor. I would also like to thank Shlomo Havlin and Lidia Braunstein, as well as all of my other collaborators, for working with me, and hope to continue to work with you all in the future. I thank Ophelia Tsui, for her time working with me and her continued good will after I stopped working with her.

It would be remiss of me not to mention my colleagues. Maksim Kitsak, I'm thankful to you for showing me the ropes when I first joined Gene's group. Alex Petersen, you were a friend outside of the office as well as the in, and set a bar for work that I could only hope to strive up towards. Joel Tenenbaum and Kevin Stokely also deserve my thanks, for their friendship and advice.

My thanks also go to out to all of Gene's family at BU: Jiayuan Luo, Elena Strelakova, Erik Lascaris, Duan Wang, Guanling Li, Jerry Paul, Bob Tomposki and too many others to mention.

Mirtha Cabello and Winna Somers have been fantastic administrative staff, friendly and helpful despite my bad habit of missing deadlines for paperwork. My thanks as well to Erich Burton, for help both in and outside the labs.

I also need to thank several people outside the university. Lauren Page, you've been a source of wisdom and comfort literally since the first day I came to Boston. I am eternally grateful. Courtney Olson, you've cheered me on and kept me sane. Martin Sachs, you and Claudine were a bright spot every week, and provided me with graphics worlds better than

I could do on my own. To everyone on Vent, even though you'll probably never read this, thanks for giving me someone to talk too while I worked.

Finally, I thank my family: my father who has been proud of me and supported me my entire life, and my mother who made sure I could never give up.

# DYNAMIC AND INTERACTING COMPLEX NETWORKS

(Order No.            )

**MARK E. DICKISON**

Boston University Graduate School of Arts and Sciences, 2012

Major Professor: H. Eugene Stanley, University Professor and Professor of Physics

## ABSTRACT

This thesis employs methods of statistical mechanics and numerical simulations to study some aspects of dynamic and interacting complex networks. The mapping of various social and physical phenomena to complex networks has been a rich field in the past few decades. Subjects as broad as petroleum engineering, scientific collaborations, and the structure of the internet have all been analyzed in a network physics context, with useful and universal results. In the first chapter we introduce basic concepts in networks, including the two types of network configurations that are studied and the statistical physics and epidemiological models that form the framework of the network research, as well as covering various previously-derived results in network theory that are used in the work in the following chapters.

In the second chapter we introduce a model for dynamic networks, where the links or the strengths of the links change over time. We solve the model by mapping dynamic networks to the problem of directed percolation, where the direction corresponds to the time evolution of the network. We show that the dynamic network undergoes a percolation phase transition at a critical concentration  $p_c$ , that decreases with the rate  $r$  at which the network links are changed. The behavior near criticality is universal and independent of  $r$ . We find that for dynamic random networks fundamental laws are changed: i) The size of the giant component at criticality scales with the network size  $N$  for all values of  $r$ , rather than as  $N^{2/3}$  in static network, ii) In the presence of a broad distribution of disorder, the

optimal path length between two nodes in a dynamic network scales as  $N^{1/2}$ , compared to  $N^{1/3}$  in a static network.

The third chapter consists of a study of the effect of quarantine on the propagation of epidemics on an adaptive network of social contacts. For this purpose, we analyze the susceptible-infected-recovered model in the presence of quarantine, where susceptible individuals protect themselves by disconnecting their links to infected neighbors with probability  $w$  and reconnecting them to other susceptible individuals chosen at random. Starting from a single infected individual, we show by an analytical approach and simulations that there is a phase transition at a critical rewiring (quarantine) threshold  $w_c$  separating a phase ( $w < w_c$ ) where the disease reaches a large fraction of the population from a phase ( $w > w_c$ ) where the disease does not spread out. We find that in our model the topology of the network strongly affects the size of the propagation and that  $w_c$  increases with the mean degree and heterogeneity of the network. We also find that  $w_c$  is reduced if we perform a preferential rewiring, in which the rewiring probability is proportional to the degree of infected nodes.

In the fourth chapter, we study epidemic processes on interconnected network systems, and find two distinct regimes. In strongly-coupled network systems, epidemics occur simultaneously across the entire system at a critical value  $\beta_c$ . In contrast, in weakly-coupled network systems, a mixed phase exists below  $\beta_c$ , where an epidemic occurs in one network but does not spread to the coupled network. We derive an expression for the network and disease parameters that allow this mixed phase and verify it numerically. Public health implications of communities comprising these two classes of network systems are also mentioned.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	What and Where are Complex Networks . . . . .	2
1.1.1	Erdős-Rényi Graphs . . . . .	4
1.1.2	Scale-Free Graphs . . . . .	5
1.2	The Configuration Model . . . . .	7
1.3	Percolation and Directed Percolation . . . . .	9
1.3.1	Percolation Threshold in Networks . . . . .	11
1.4	Monte-Carlo Methods . . . . .	15
1.5	Shortest Paths; Ordered and Disordered . . . . .	16
1.6	Epidemic Models . . . . .	18
1.7	Appendix: Directed Percolation Scaling Laws . . . . .	21
<b>2</b>	<b>Dynamic Networks and Directed Percolation</b>	<b>23</b>
2.1	Introduction . . . . .	23
2.2	Dynamic Networks . . . . .	24
<b>3</b>	<b>Quarantine Generated Phase Transition in Epidemic Spreading</b>	<b>33</b>
3.1	Introduction . . . . .	33
3.2	Analytical Approach . . . . .	34
3.3	Simulation Results . . . . .	38
3.4	Summary and Conclusions . . . . .	43



<b>4 Epidemics on Interacting Networks</b>	<b>46</b>
4.1 Introduction . . . . .	46
4.2 Model . . . . .	48
4.3 Strongly-Coupled Network Systems . . . . .	51
4.4 Weakly-Coupled Network Systems . . . . .	51
4.5 Conclusions . . . . .	57
4.6 Appendix: Interacting Square Lattices. . . . .	58
<b>5 Summary and Future Work</b>	<b>60</b>
<b>Bibliography</b>	<b>61</b>
<b>Curriculum Vitae</b>	<b>66</b>

# List of Figures

1.1	A schematic illustration of nodes, links, and degree. . . . .	3
1.2	Degree distributions and real world network examples of Erdős-Rényi (ER) and scale-free (SF) graphs. . . . .	6
1.3	A schematic illustration of the configuration model. . . . .	8
1.4	Schematic diagram of percolation and directed percolation (DP). . . . .	10
1.5	The SIR process on a network. . . . .	19
2.1	Dynamic network with optimal paths. . . . .	25
2.2	Simulation results for the survivability $P_s(t)$ and its cutoff, at criticality . . .	26
2.3	Cluster size scaling in dynamic networks. . . . .	28
2.4	Simulations showing validity of $p_c$ and its dependence on link change rate. . .	29
2.5	Optimal path scaling under strong disorder . . . . .	31
3.1	A sample $\beta$ - $w$ phase diagram. . . . .	37
3.2	Infection densities for two infection strategies . . . . .	39
3.3	Quarantine induced phase transition for various $\beta$ . . . . .	40
3.4	Quarantine induced phase transition for various $\langle k \rangle$ . . . . .	41
3.5	Quarantine induced phase transition in scale-free networks . . . . .	42
3.6	Degree dependent quarantine. . . . .	44
4.1	An interconnected network system with two networks: A and B. . . . .	47
4.2	Cluster growth in weakly- and strongly-coupled network systems. . . . .	52
4.3	Phase diagram for weakly-coupled network systems. . . . .	53

4.4	Infection densities in the mixed phase. . . . .	54
4.5	Survival probabilities in the mixed phase. . . . .	55
4.6	Survival probability gaps for different network parameters. . . . .	56
4.7	Survival probability gap universality under rescaling. . . . .	57
4.8	Approach of square lattice to complex network behavior with increasing interactions. . . . .	59

## List of Abbreviations

ER	Erdős R�nyi
SF	Scale-Free
WWW	World Wide Web
DP	Directed Percolation
GC	Giant Component
MC	Monte Carlo
MST	Minimal Spanning Tree
SIR	Susceptible-Infected-Recovered
SIS	Susceptible-Infected-Susceptible

# Chapter 1

## Introduction

Network Science is a growing, interdisciplinary, field bringing together tools from Statistical Physics and Graph Theoretical Mathematics along with the frameworks and understanding of fields such as Psychology, Sociology, Computer Science, Epidemiology, and Geology. One can ask “Why study networks”, or specifically relevant to this thesis, “Why should physicists study networks?”

Throughout much of the twentieth century, reductionism has been the driving force behind scientific advancement. Atomic and later quantum theory discarded molecular dynamics and moved towards pure particle-particle interactions. Condensed matter moved out of high temperature and bulk statistical properties into low dimensional and low temperature effects. Biology moved into sequencing genes and mapping the structures of individual proteins. While much important work remains to be done in these areas, it is clear that complete understanding will not solely be found in understanding the building blocks of the universe and of life. How these basic pieces interact is crucial to expanding our knowledge of ourselves and the world around us. It is also clear that purely random or largely regular statistical models are insufficient. While the behavior of gases is well described by a random statistical distribution, and the arrangement of atoms in a graphene layer is a regular lattice (with perhaps some defects), protein-protein interaction networks and emergent human structures such as transportation networks, social interaction networks, and communication networks all have non-trivial structural properties: they are complex.

These complex networks have unique behaviors that must be taken into account in order to understand how they act, and how they can be used to our benefit.

Aside from the fundamental need to understand the interconnections of physical systems, the social and technological environment we find ourselves in has brought networks to the forefront. The war on terror is largely about mapping and destroying the Al Qaeda terrorist network [1]. Recurrence of previously rare diseases such as whooping cough [2] demand understanding of epidemiological networks, where the small world nature of social networks exposes the vulnerability of the population to even a small number of the unvaccinated [3].

This thesis approaches three distinct, yet related topics in complex networks: 1) General properties of dynamic networks, where the network changes over time. 2) Quarantine, where the network is dynamically altered specifically to avoid infected sites. 3) Epidemic properties on interacting networks, where two or more networks are connected.

## 1.1 What and Where are Complex Networks

A definition of complex networks must start with first defining what a network is. The core features all networks share are nodes and links (Fig. 1.1). Nodes represent the fundamental units of the system in question (pores in rock, scientists in a field, routers on the Internet) and links establish which of the nodes are connected to other nodes (channels between pores in rock, scientific collaborations, or cabling between routers). All information about a network is contained in the list of nodes and the links between them, however, there are many derived properties that are useful to refer to, much as it is useful to talk about the temperature of a gas rather than the kinetic energy of each individual particle. The degree  $k$  of a node is the number of links it has, which is to say, the number of nodes it has as nearest neighbors on the network. The degree distribution of a network ( $P(k)$ ), or the probability that a random node will have a degree  $k$  is a further characteristic of all networks, and is often used to classify networks. Other quantities that are of interest are the first moment  $\langle k \rangle$  and the ratio of the second moment to the first moment  $\kappa = \langle k^2 \rangle / \langle k \rangle$  of the degree distribution.

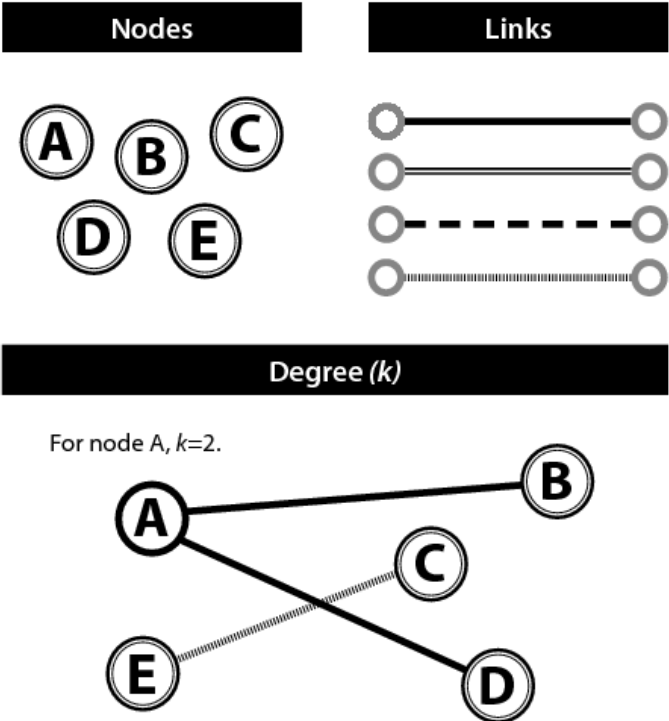


Figure 1.1: A schematic illustration of nodes, links, and degree.

So if a network is anything that can be represented by a collection of nodes and the links between them, what makes a network complex? A complex network is any network that is not uniform. A chessboard forms a network, but not a complex one. Each square is connected to exactly four other squares. Pure crystals such as diamonds or ice are again simple networks, as each atom has exactly the same links to its neighbors as every other atom. The US highway network, on the other hand, is a complex network. Different cities have different numbers of highways, and it is not possible to draw a map of US highways by looking solely at a small portion of them. Knowing where the highways are in MA does not tell you where they are located in TX. Complex networks are all around us, in our metabolisms, in our power and transport grids, and in our social interactions. A variety of models have been used to study them.

### 1.1.1 Erdős-Rényi Graphs

While network science has only recently entered the mainstream, and its roots can be traced back centuries, the modern foundations of network theory were certainly laid in the 1900s by Paul Erdős and Alfred Rényi. “On Random Graphs” [4] introduced a network formation model with  $N$  nodes, each of which was randomly linked to every other node with probability  $p$ . For  $p = 0$ , one had the empty graph, and for  $p = 1$ , the complete graph. In the range  $0 < p < 1$  lies the simplest family of non-trivial graphs. Each node has  $N - 1$  potential neighbors, and thus the average degree of each node for large  $N$  is  $\langle k \rangle = p(N - 1) \approx pN$ . When dealing with large networks, the average degree is often specified instead of the connection probability, due to  $p$  being very small in most large real-world networks, whereas  $\langle k \rangle$  is a more tractable number. The degree distribution generated by the Erdős-Rényi (ER) model follows a Poisson distribution in the limit  $N \rightarrow \infty$ .

$$P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!} \quad (1.1)$$

with  $\langle k \rangle = Np$ .

The ER model was the primary model of networks for decades. It was assumed to



represent models of acquaintance formation in small social situations, such as parties or conferences, and was used a simplified model for the evolution of communication networks. It is still used today, as it is well known, forms a standard, is analytically well tractable, and even has real world examples, such as highway or railway systems [5]. However, while examples of ER networks in the real world exist (Fig. 1.2), many if not most networks are not adequately represented by the ER model.

### 1.1.2 Scale-Free Graphs

In the 1990s, the World Wide Web (WWW) brought networks into the center of most of our lives. While scale-free (SF) networks certainly predate the Internet (scientific collaboration networks are scale free as far back as the records reach), the WWW provided an immediate and tangible example that most nodes in a network do not have a typical or average number of links. Consider the front pages of the New York Times, the Microsoft Network, the British Broadcasting Corporation, Facebook, or any one of hundreds of popular pages. Those hub pages have thousands of links or more, whereas most pages are personal Geocities pages with only a few links, or news articles that may only link back to the original site, or a few related articles. When the WWW was analyzed, it was indeed found that the degree distribution did not follow a Poisson distribution but rather, for a large range of  $k$ , a power law distribution [6].

$$P(k) \sim k^{-\lambda} \tag{1.2}$$

Many other networks were later found to be scale free, such as metabolic networks[7], the human sexual interaction network[8], and even human language[9]. However, despite being scale free over a large range of degrees  $k$ , real-world networks often have a largest degree  $k_{max}$  that is smaller on average than one would expect from a pure power law degree distribution with a matching  $\lambda$  [10, 11]. In addition, for SF networks, the second moment  $\langle k^2 \rangle \rightarrow \infty$  as  $N$  increases for fixed  $\langle k \rangle$  which, in addition to failing to match the behavior of physical networks, makes many properties impossible to calculate analytically. Thus, for both mathematical simplicity and model accuracy, it is useful to introduce an exponential

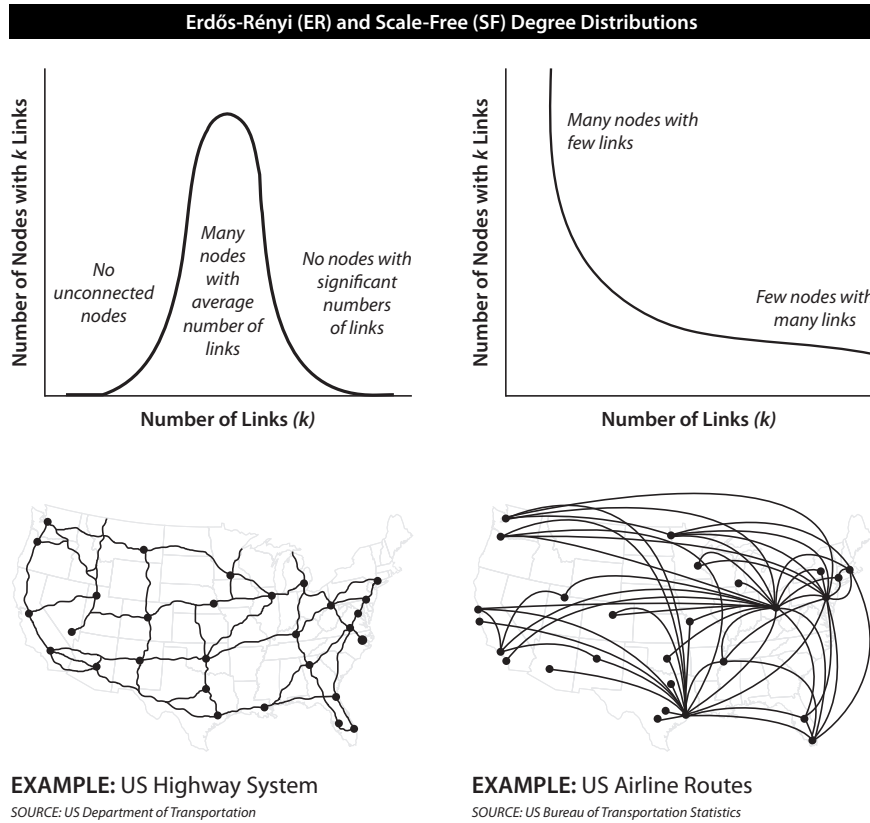


Figure 1.2: Degree distributions and real world network examples of Erdős-Rényi (ER) and scale-free (SF) graphs. The highway network is an example of an ER graph, whereas the direct flight network is SF.

cutoff for degrees above a threshold  $K$  which keeps  $\langle k^2 \rangle$  finite as  $N$  increases, and ensures that the probability of extremely high degree nodes is realistically small. Networks following such an exponential cutoff are known as finite scale-free networks, and will be used in this thesis instead of infinite SF networks. The degree distribution of finite scale-free networks is

$$P(k) = \left[ k^{-\lambda} \exp(-k/K) \right] / \left[ \text{Li}_\lambda(e^{-1/K}) \right], \quad (1.3)$$

where  $\text{Li}_\lambda(x)$  is the  $\lambda$ th polylogarithm of  $x$ . For degrees below the threshold  $K$ , the distribution is power law, and for degrees above  $K$ , the distribution is exponential.

## 1.2 The Configuration Model

While ER graphs, as random graphs, can be constructed simply by picking two nodes at random and connecting them until the desired  $\langle k \rangle$  is reached, graphs that do not follow a Poisson degree distribution cannot be constructed in this fashion. The configuration model [12] is a method of constructing networks that follow an arbitrary degree distribution  $P(k)$  (See Fig. 1.3.

1. Assign each node  $n$  a degree drawn at random from the degree distribution, if the total degree of all nodes is odd, replace a node's degree with a new random degree until the total degree is even.
2. Create a list where each node appears a number of times equal to its degree.
3. While this list still has elements remaining, pick and remove two elements randomly from this list and place a link between them.

The procedure described does not proscribe self-links or repeated links. However, as will be seen from Eq. (1.5), the probability of a self link is  $k_i^2/(N\langle k \rangle)$  and the probability of a repeat link is  $k_i k_j/(N\langle k \rangle)$ . Both of these probabilities go to zero for increasing  $N$ , and the cut-off  $K$  for the finite scale free networks used here ensures that the probability of either of these events occurring is exponentially unlikely, even for finite  $N$ .

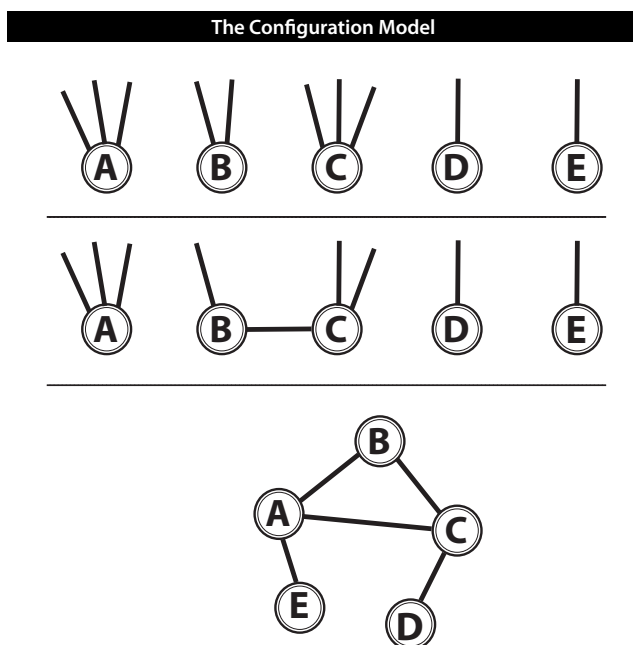


Figure 1.3: A schematic illustration of the configuration model. Nodes are first assigned a degree. In this case  $k_A = 3$ ,  $k_B = 2$ ,  $k_C = 3$ ,  $k_D = 1$ , and  $k_E = 1$ . Next, two nodes with links not yet assigned are picked at random and connected, in this case B and C. This process is repeated until all nodes have a degree equal to that assigned to them originally, producing one of the possible network configurations that satisfies the chosen degree distribution.

### 1.3 Percolation and Directed Percolation

As mentioned earlier, one possible system modeled by a network is a porous material, such as pumice, foam rubber, or a sponge, with the pores as nodes and channels between them as links. A fundamental question for such a system is: If a liquid is introduced at one edge, will that liquid be able to make it through the pores of the material and emerge from the opposite side? Investigations into questions of this nature have historically formed the field known as Percolation Theory. To model this question, Broadbent and Hammersley introduced a lattice with regular nodes where each node was connected to its nearest neighbors randomly with probability  $p$ , and disconnected with probability  $(1 - p)$ . The liquid will be able to percolate through the material if an unbroken path exists from nodes on one edge to nodes on the other. For infinite systems, this is only possible if a fraction of all nodes are connected into one large cluster, as the odds of a finite number of nodes reaching from one side to another decreases exponentially with system size. There is thus a critical  $p = p_c$  at which a first-order phase transition occurs in the infinite size limit: below  $p_c$  the largest cluster consists of a finite number of nodes, and percolation occurs with probability 0; above  $p_c$  the largest cluster occupies a fraction of the infinite nodes, and percolation occurs with probability 1. For the square lattice in two dimensions, it has been shown that  $p_c = 0.5$  [13]. For finite lattices the percolation transition is still very sharp, even for system sizes as small as  $10 \times 10$ . For systems with  $d \neq 2$ , the value of  $p_c$  varies until a critical dimension is reached, above which the system behaves according to mean field theory, and the dimensionality no longer plays a role.

There are many varieties of percolation processes; directed percolation (DP) is a particularly notable one in which the dimensions are no longer isotropic. Along one of the dimensions, the perpendicular dimension, spreading is restricted to travel in only one direction. This can be interpreted either spatially, as in the case of a liquid penetrating a porous medium where capillary action is insufficient to overcome gravitational forces, so that the fluid may only flow downwards, or temporally, as when modeling the spread of some agent

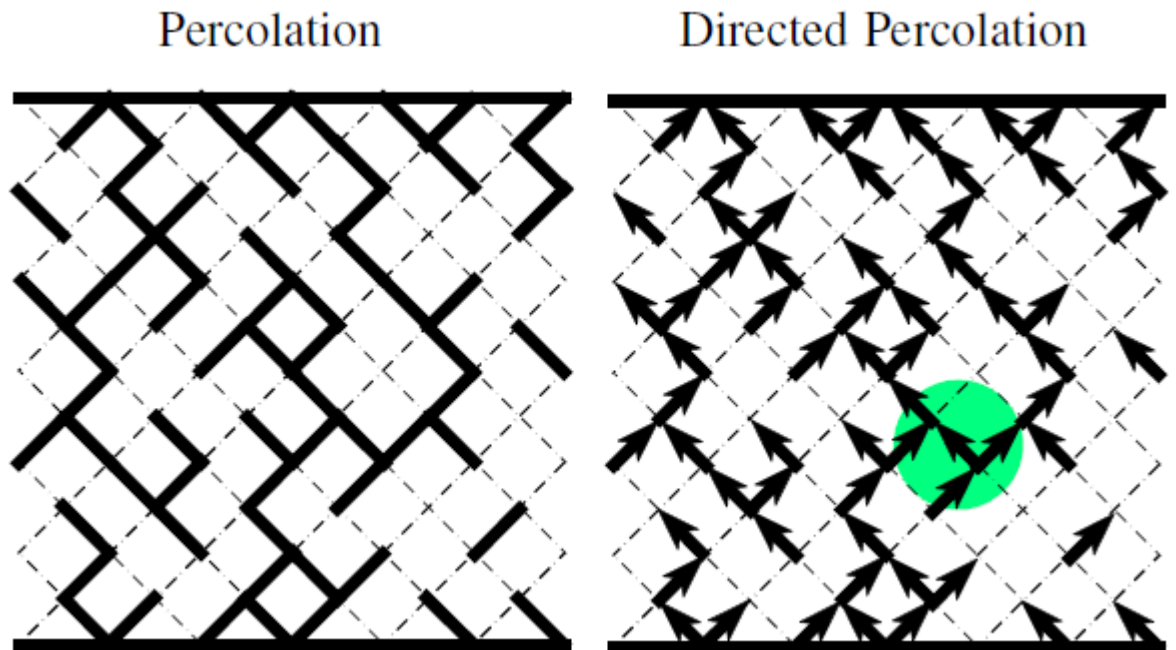


Figure 1.4: Schematic diagram of percolation and directed percolation (DP). Percolation occurs only if an unbroken line exists between the bottom and the top. In DP, only paths that advance along the direction of the arrows are allowed.

through time, where connections that existed in the past cannot be traversed, or nodes that were neighbors in the past are now separated. DP has a percolation transition similar to isotropic percolation, although the value of  $p_c$  is necessarily higher, as the directional requirement invalidates some paths that would exist otherwise. Figure (1.4) shows two lattices, one with regular percolation and the other with DP at the same  $p$ . Note that the regular percolation lattice has a path connecting the upper and lower edges, and so is above  $p_c$ , but the DP lattice is not.

Just as results that were originally found for simple lattices in graph theory came to be extended to complex networks in network theory, one can likewise extend percolation models past the lattice model into complex networks, such as ER and SF networks. In percolation theory, the condition of traversing from one side of the lattice to the other requires a large, connected cluster of sites. In complex networks, there are no longer “sides”, but if a parameter (such as the average degree, or the chance of deleting links) is varied, the point

at which a giant cluster – a cluster where the fraction of connected nodes does not decrease with increasing system size  $N$ – emerges is still referred to as the percolation threshold. If the network were to be embedded in a physical space, the giant cluster would necessarily span it. In complex networks, the percolation threshold is  $p_c = 1/(\kappa - 1)$ , the derivation of which follows.

### 1.3.1 Percolation Threshold in Networks

This method of calculating the percolation threshold follows the methods of Newman [14]. First, assume a start from a random node  $i$  with degree  $k_i$ . What is the probability that a random link leading from  $i$  leads to a node  $j$  with degree  $k_j$ ? The probability that  $i$  is not connected to  $j$  is the probability that none of the  $k_i$  links of node  $i$  are any of the  $k_j$  links of node  $j$ , out of  $\langle k \rangle N$  total links, or

$$P(i \not\leftrightarrow j | k_i, k_j) = \left(1 - \frac{k_j}{\langle k \rangle N}\right)^{k_i} \simeq 1 - \frac{k_i k_j}{\langle k \rangle N} \quad (1.4)$$

for  $N \rightarrow \infty$  with  $k_j \ll N$ . The probability that  $i$  does link to  $j$  is thus

$$P(i \leftrightarrow j | k_i, k_j) = \frac{k_i k_j}{\langle k \rangle N}. \quad (1.5)$$

The probability that an arbitrary node  $i$  of unknown degree connects to a node  $j$  with degree  $k_j$  is

$$P(i \leftrightarrow j | k_j) = \sum_i \frac{P(k_i) k_i k_j}{\langle k \rangle N} = \frac{k_j}{N}, \quad (1.6)$$

and, lastly, the probability of two arbitrary nodes being nearest neighbors is

$$P(i \leftrightarrow j) = \sum_{k_j} \frac{P(k_j) k_j}{N} = \frac{\langle k \rangle}{N}. \quad (1.7)$$

If it is known that node  $i$  is connected to a random node  $j$ , what is the probability  $P(k_j | i \leftrightarrow j)$  that node  $j$  has degree  $k_j$ ? Bayesian probability tells us

$$P(k_j | i \leftrightarrow j) P(k_j) = P(i \leftrightarrow j | k_j) P(i \leftrightarrow j). \quad (1.8)$$

Substituting in from Eq. (1.6) and Eq. (1.7) gives

$$P(k_j|i \leftrightarrow j) = \frac{P(k_j)k_j}{\langle k \rangle}. \quad (1.9)$$

Next, to calculate the point at which the largest cluster emerges, the concept of the generating function  $G_0(x)$  of a degree distribution  $p(k)$  is introduced.

$$G_0(x) = \sum_k p_k x^k. \quad (1.10)$$

The generating function is a polynomial of degree  $k_{max}$  where the coefficient of each term is the normalized probability that a random site has that degree. The generating function for a network that is composed of one-half degree one nodes and one-half degree two nodes would thus be  $.5x + .5x^2$ , as a trivial example. Some general properties of generating functions are:

1. Moments: The  $n$ th moment of a generating function is its  $n$ th derivative at 1

$$\langle k^n \rangle = \sum_k k^n p_k = G_0^{(n)}(1). \quad (1.11)$$

2. Powers: The distribution of the sum of  $m$  independent realizations of a property is equal to the  $m^{th}$  power of a generating function for that property.

$$\sum_{j_1 \dots j_m} p_{j_1} \dots p_{j_m} x^{j_1 \dots j_m} = [G_0(x)]^m. \quad (1.12)$$

Let  $G_1(x)$  be the generating function of the probability that a randomly chosen link connects to a node of degree  $k$ . From Eq. (1.9) we know that the probability of this is  $P(k)k/\langle k \rangle$ .

This gives us

$$G_1(x) = \sum_k \frac{k p_k x^k}{\langle k \rangle} = \frac{G_0'(x)}{G_0'(1)}. \quad (1.13)$$

Now there are  $k$  nearest neighbors, each with the distribution  $G_1(x)$ ; therefore, using the powers property and ignoring the possibility of self-links or repeated links, which decreases



as  $N^{-1}$ , the generating function for the second nearest neighbors  $G_2(x)$  is

$$G_2(x) = \sum_k p_k [G_1(x)]^k = G_0(G_1(x)). \quad (1.14)$$

Let  $H_1(x)$  now be the generating function for the distribution of sizes of all the groups of nodes, not just neighbors, reachable from a random link. The generating function for the probability that a link is connected to a node of size  $k$  is  $G_1(x)$ ; therefore by the powers property

$$H_1(x) = xG_1(H_1(x)). \quad (1.15)$$

A random node has a number of these groups equal to its degree, and so the generating function for the size of the whole component  $H_0(x)$  is related to the probability of that degree

$$H_0(x) = xG_0(H_1(x)). \quad (1.16)$$

Taking the first moment of the component size  $\langle s \rangle = H_0'(1)$  we find, with normalization,

$$H_0'(1) = 1 + G_0'(1)H_1'(1). \quad (1.17)$$

Where Eq. (1.15) shows

$$H_1'(1) = \frac{1}{1 - G_1'(1)}. \quad (1.18)$$

This gives

$$\langle s \rangle = 1 + \frac{G_0'(1)}{1 - G_1'(1)} \quad (1.19)$$

which becomes infinite, signaling the emergence of the percolation cluster, when

$$G_1'(1) = 1 \quad (1.20)$$

from Eq. (1.13) this is equivalent to

$$\frac{\sum_k p_k k^2}{\sum_k p_k k} = \kappa = 2. \quad (1.21)$$

Suppose that we have a system above the percolation threshold, and we want to know how many links we can remove at random before the giant component is destroyed. Starting from an initial degree distribution  $P_k$  and removing links with probability  $q$ , we arrive at a new distribution  $P_k^*$  for the altered network with

$$P_k^* = \sum_{k_0=k}^{\infty} P_{k_0} \binom{k_0}{k} q^{(k_0-k)} (1-q)^k. \quad (1.22)$$

We can derive the critical failure probability,  $q_c$ , by substituting in the first and second moments of this degree distribution into Eq. (1.21) These first and second moments are

$$\langle k^* \rangle = \langle k \rangle (1-q) \quad (1.23)$$

and

$$\langle (k^*)^2 \rangle = \langle k \rangle q (1-q) + \langle k^2 \rangle (1-q)^2. \quad (1.24)$$

Setting  $\kappa^* = 2$  and solving, we find

$$q_c = 1 - \frac{1}{\kappa - 1}. \quad (1.25)$$

This is the complex network equivalent of the  $p_c = 1/2$  found for the 2-d square lattice. Of note is that complex network theory predicts  $p_c = 1 - q_c = 1/3$  for the square lattice, which has  $\kappa = 4$ . Even ER complex networks (ones where every node has nearly the same degree) cannot be treated simply as uniform networks. The question of how much heterogeneity is necessary for complex network theory to be valid, instead of the results from uniform percolation theory, will be examined as part of Chapter 5.

## 1.4 Monte-Carlo Methods

Many real problems on complex networks, while straightforward in principle, are intractable in practice. Real-world networks are finite, and many properties which are simple to calculate for infinite systems become difficult or impossible in finite systems. As a specific example, take Eq. (1.25) from the preceding section. For an ER network,  $\kappa = \langle k \rangle + 1$ , and so for an infinite network, we know that if we start from a network with  $\langle k \rangle = 1$ , the probability of it have a giant component is 1. However, for finite networks, the same is not true. Consider a sample network with  $N = 6$ . There are 15 (  $N(N - 1)/2$  ) possible links, and 455 (  $\binom{15}{3}$  ) possible networks with  $\langle k \rangle = 1$ . It is fairly trivial to go through the possible networks by hand and show that 10 out of these 455 possible networks have only three sets of linked pairs, with no giant component, giving the probability of a giant component as  $p_{GC}(6) \approx .978$ . However, while the  $N = 6$  case is simple enough to compute by hand, as an NP-Hard problem[15], it rapidly becomes intractable. Double the size, to  $N = 12$  and there are now over ninety thousand network configurations, requiring a computer to iterate through them. Double the size of the network once more, to  $N = 24$ , and there are now over  $3.2 \times 10^{20}$  different configurations, obviously impossible to compute explicitly even for this still very small network.

A common method to resolve problems of computational complexity is the Monte Carlo (MC) technique. Originally developed to calculate neutron penetration distances at Los Alamos National Laboratory, the technique was first named in 1949[16]. The MC label applies to a wide variety of methods, but what they all have in common is random sampling of a configuration space according to certain probabilities. In this thesis, MC methods are used to solve two main classes of problems.

The first class is similar to the problem above: networks are repeatedly generated according to the configuration model (Sec. 1.2), and each network has its properties (presence of a giant component, average degree, shortest path distances, etc..) computed. These properties are then averaged to obtain an estimate for the exact value that would be obtained

by calculating over all possible network configurations.

The second class of problems involves network dynamics, such as the number of infected individuals in an epidemic. In this case, not only is a random initial network configuration generated, but the network is also allowed to evolve forward in time repeatedly from the same initial conditions. As the number of time steps in the time evolution is often small compared to the size of the network, (typically scaling as  $N^\alpha$ , with  $\alpha < 1$ ), this offers a substantial speed increase, as network construction is computationally more expensive, which typically scales with  $N$ .

## 1.5 Shortest Paths; Ordered and Disordered

From previous sections we know that above the percolation transition most nodes are connected into a giant component. We can thus reach almost every node from any given node. But mere knowledge that a path exists often is not the only topic of concern. We often want to know how many links we must take to get from node to node on average. Obviously, if the nodes aren't simply connected in a ring, there will be multiple paths of varying lengths. Out of all of these possible paths, one often only cares about the one with the fewest number of links: the shortest path. There is a shortest path between all nodes on the giant component, and of these  $N_{GC}(N_{GC} - 1)/2$  shortest paths we can take various values: the maximum value to define the diameter of the network, which is the longest shortest path, or the average to calculate the average shortest path  $l$ . In a lattice,  $l$  scales as  $l \sim N^{1/d}$ . For complex networks, the average shortest path length is

$$l = \frac{-2\langle \ln k \rangle + \ln(\langle k^2 \rangle - \langle k \rangle) + \ln N - \gamma}{\ln(\kappa - 1)} + \frac{1}{2} \quad (1.26)$$

where  $\gamma = 0.5772$  is Euler's constant [20]. For ER networks, this reduces to

$$l^{ER} = \frac{\ln N - \gamma}{\ln \langle k \rangle k} + \frac{1}{2} \quad (1.27)$$

For SF networks, the exact expression does not simplify, but in the limit of large  $N$ , two general regimes emerge depending on  $\lambda$ . For  $\lambda > 3$ , the shortest path scales like an ER network, with  $l \sim \ln N$ . For  $2 < \lambda < 3$ , however  $l \sim \ln(\ln N)$ .

The distance from node to node in terms of links may not necessarily capture all of the relevant information about distances. Links are not necessarily uniform, they may also have weights attached to them. In a communications network, the latency (travel time) is different between different nodes. Transportation networks are also heterogeneous. LA is connected via direct flights to both San Francisco and Sydney, but the flight times differ greatly. Social networks can also feature links with different weights. I have many friends, but the frequency with which I see them differs greatly. If one wanted to send information or objects via direct person-to-person interactions, this heterogeneity may need to be considered. Networks with weights on the links are called weighted or disordered networks. We model this disorder by assigning different weights to the links, to represent either a cost or a time associated with traversing it. These costs can be generated by assigning each link  $i$  a random number  $r_i$  uniformly distributed between 0 and 1. The cost of the link  $\tau$  is  $\tau_i = e^{ar_i}$ . In disordered networks, the shortest path length (smallest number of links) is of less interest than the *optimal* path length,  $l_{opt}$ , which is the smallest sum of the link weights. If  $a \ll 1$ , the weak disorder case, the costs are all roughly equivalent, and the optimal path length has the same scaling as the shortest path length. If  $a \gg 1$ , the strong disorder case, then the cost distribution is very broad, and the total cost will be dominated by the single highest cost. (The crossover between weak and strong disorder is not relevant to this thesis, but has been addressed [17]). The optimal path in this case lies along the Minimal Spanning Tree (MST), which is the tree connecting clusters at the percolation threshold of the network. The optimal path between two nodes will be comprised mainly of links along the MST, and thus the average optimal path will scale with the average distance within the giant component  $l_{GC}$ , where

$$l_{GC} \sim N^{\nu_{opt}}. \quad (1.28)$$

For ER networks  $\nu_{opt} = 1/3$ , and thus under strong disorder:

$$l_{opt}^{ER} \sim N^{1/3}. \quad (1.29)$$

## 1.6 Epidemic Models

An epidemic is an occurrence of a disease in excess of normal expectancy; however, communicable disease models of all types are often referred to as epidemic models, just as the general study of disease is called epidemiology. In modeling diseases in a population, the population in question is divided into disjoint classes. In this work we consider three possible classes: The susceptible (S) class, which consists of individuals who are vulnerable to a disease, but are not currently infected; The infected (I) class, which consists of those individuals currently infected with the disease, and capable of infecting other individuals, and the recovered or removed class (R), which comprises those individuals who have been infected, and can no longer infect others or be reinfected. Various models can be constructed depending on allowable transitions between these three (or other) states, but the susceptible/infected/recovered (SIR) model [18] is the most common, and will be addressed here (See Fig. 1.5). It represents diseases to which exposure results in recovery and immunity (or death), such as HIV, influenzas such as H1N1, scarlet fever, and measles. The spread of the disease is governed by two parameters: infected individuals come into contact and infect all susceptible individuals with rate  $\beta_I$ , and infected individuals recover (or die) at rate  $\gamma$ .

For large populations with full mixing and no birth/death rate, the SIR model is governed by three coupled nonlinear differential equations:

$$\frac{ds}{dt} = -\beta_I i s, \quad \frac{di}{dt} = \beta_I i s - \gamma i, \quad \frac{dr}{dt} = \gamma i, \quad (1.30)$$

where  $s(t)$ ,  $i(t)$ , and  $r(t)$  are the relative fraction of the population in each state. To find the conditions on an epidemic (which is to say, when  $s < 1$ ), assume  $s \approx 1$ . Then  $\frac{di}{dt}$  is positive if  $\beta_I > \gamma$ , meaning that the infected population will increase until a substantial

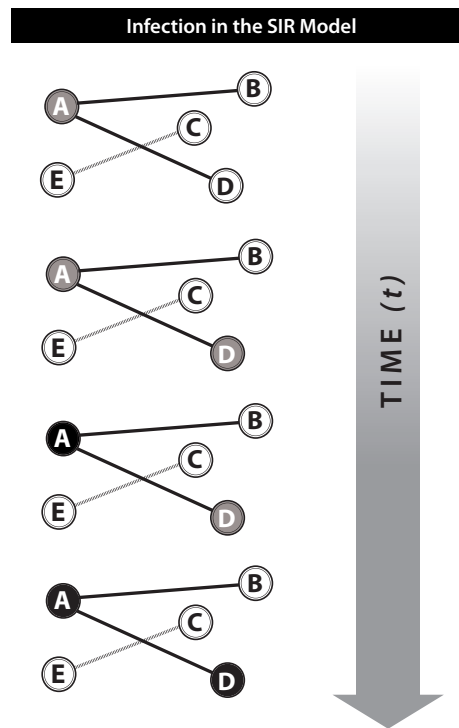


Figure 1.5: The SIR process on a network. Infected nodes (grey) infect their susceptible neighbors (white) with a certain probability at each time step. After remaining infected for the recovery time infected nodes become recovered (black). The final network state consists of only susceptible and recovered nodes.

fraction of the population has been infected, indicated by  $s < 1$ . The epidemic threshold in this model is simply  $\beta_I > \gamma$ .

The equations above makes the assumption that every individual is connected to, and can infect, every other individual. Realistically, every individual is connected only to certain other individuals. For ease of comparison to simulation, let us move to discrete time, with the average time of recovery being  $t_R$ , and let  $\beta$  be the chance of infection per time step. The transmissibility  $T$ , or the total chance that an infected individual will transmit the infection to a susceptible neighbor is then

$$T = 1 - (1 - \beta)^{t_R}. \quad (1.31)$$

Now, we vary the number of connections that each person has by assuming they are nodes on a network with a particular degree distribution  $P(k)$ . Imagine that we start with a single infected individual of degree  $k_0$ . The probability of this node infecting exactly  $k$  of its  $k_0$  edges is given by a binomial probability distribution  $\binom{k_0}{k} T^k (1 - T)^{(k_0 - k)}$ . The probability distribution of all possible infected nodes is

$$P_{k_0} \binom{k_0}{k} T^k (1 - T)^{(k_0 - k)}, \quad (1.32)$$

which is simply Eq. (1.22) with  $T = (1 - q)$ . Thus, we can immediately conclude from Eq. (1.25) that

$$T_c = \frac{1}{\kappa - 1}. \quad (1.33)$$

It can be shown [19] that if  $\beta$  or  $t_R$  are independently identically distributed random variables drawn from some distributions  $P(\beta)$  and  $P(t_R)$ , the epidemic threshold depends simply on the first moment of  $T$

$$\langle T \rangle = \int_0^\infty P(t_R) dt_R \sum_{\beta=0}^\infty P(\beta) P(t_R) (1 - \beta)^{t_R} \quad (1.34)$$

and the critical condition is  $\langle T \rangle_c = (\kappa - 1)^{-1}$ .



## 1.7 Appendix: Directed Percolation Scaling Laws

According to Janssen and Grassberger, all models that meet the following four conditions are members of the Directed Percolation (DP) universality class.

1. The existence of a continuous phase transition from a fluctuating active state into a unique absorbing state.
2. A positive one-component order parameter.
3. No long-range interactions.
4. No additional symmetries or quenched disorder.

The susceptible-infected-susceptible (SIS) model satisfies all of these properties, and so it will be used to derive the scaling exponents for DP in general. Equations (1.30) become

$$\frac{ds}{dt} = -\beta_I i s + \gamma i, \quad \frac{di}{dt} = \beta_I i s - \gamma i, \quad (1.35)$$

and, making use of the condition  $i + s = 1$ , we arrive at a single rate equation

$$\frac{di}{dt} = (\beta_I - \gamma)i - \beta_I i^2. \quad (1.36)$$

Approaching criticality ( $\beta_I = \gamma$ ) from above ( $0 < i \ll 1$ ), the stationary density of  $i$  vanishes as  $i^{stat} \sim \beta_I$ . The mean field density is thus linear in the order parameter, and the density exponent is  $\beta' = 1$ . Approaching criticality from below, we have

$$\frac{di}{dt} \approx -\delta i. \quad (1.37)$$

Which shows the density in the inactive phase decaying with time as  $i(t) \sim e^{-t} = e^{-t/\xi_{\parallel}}$ , and thus the temporal scaling exponent is  $\nu_{\parallel} = 1$ . In order to determine the spatial scaling exponent  $\nu_{\perp}$ , a term for particle diffusion must be added:

$$\frac{di(x,t)}{dt} = (\beta_I - \gamma)i(x,t) - \beta_I i^2(x,t) + D\nabla^2 i(x,t), \quad (1.38)$$

where  $D$  is the diffusion constant and corresponds to nearest-neighbor interactions in a lattice model. Eq. (1.38) should be invariant under the rescaling  $x \rightarrow \Lambda x$ ,  $t \rightarrow \Lambda^{\nu_{\perp}/\nu_{\parallel}}$  and  $i(x, t) \simeq t^{\beta/\nu_{\parallel}}$  [21], which gives us  $\nu_{\perp} = 1/2$ .

Lastly, from the generalized hyperscaling relation [22]

$$d\nu_{\perp} = 2\beta' \tag{1.39}$$

we find the upper critical dimension  $d_c = 4$ . The volume of the active cluster  $S(t)$  at criticality is thus

$$S(t) \sim (\xi_{\perp})^4 \sim t^2. \tag{1.40}$$

These relations will be used in the following section to confirm that Dynamic Networks are a member of the DP universality class.

## Chapter 2

# Dynamic Networks and Directed Percolation

### 2.1 Introduction

Network theory has answered many questions concerning static networks [3, 8, 10, 19, 23–32], but many real networks are dynamic in the sense that their links, or the strengths of their links, change with time. For example, in social networks friendships are formed and dissolved, while in communication networks, such as the Internet, the load (weight) on the links changes continually. Models for dynamic networks have been studied in the context of epidemic models in biology [33], as well as for routing and gossiping algorithms in computer science [34]. In social networks, dynamic models such as the reciprocity model and the actor-oriented model [35–38], include rate and objective functions that allow the control and optimization of the changes in the network.

In this section we focus on the general physical aspects of dynamic networks. Fundamental questions that have been extensively studied in static networks are still open for dynamic networks. Here we ask: i) Does the dynamic network undergo a percolation phase transition, above which order  $N$  of the network nodes are still connected and below which the network breaks into small clusters? ii) If so, what is the critical concentration of links for which the transition occurs, and how does it depend on the dynamics? iii) What are the properties near criticality?

## 2.2 Dynamic Networks

Consider an  $N$ -node network with  $M$  links where each link has a weight  $w$  chosen from a given distribution. In the particular example of a percolation process, we consider the percolating fluid to be a walker on our dynamic network. Our dynamic processes occur at a time scale such that the walker can traverse a single link in one unit time step. At the end of each time step, links are rewired with probability  $r$ . Without loss of generality<sup>1</sup>, we assume that a walker traversing the network cannot remain at a given node, and must advance to a new node in each time step. If a walker is unable to do so in a given time step, it is removed from the network. Now even if there is no path between nodes  $A$  and  $B$  at a specific time, a walker traversing the network may be able to pass from point  $A$  to point  $B$  because new links are continually appearing. Likewise, even if a path between  $A$  and  $B$  exists at a given time it may be disconnected before a walker is able to traverse it. Moreover, even if a path between two nodes always exists, the shortest path and optimal paths may change. Figure (Fig. 2.1) demonstrates a scenario where the path between node  $A$  and  $E$  after four steps is the optimal path, rather than the shorter path that existed between the two nodes after two steps.

In a percolation process on a dynamic network, after the links are rewired on each step, each link is set to be traversable with probability  $p$ . We now argue that percolation on a dynamic network is equivalent to the problem of DP (see Section 1.3) in infinite dimensions[21, 40]. To show this, the time evolution of the network is represented by adding another axis, which corresponds to a time axis. In this extended representation, every two successive rows along the time axis represent a layer that corresponds to the network configuration at a different time step. As a result, each node in the original network is represented by a set of nodes, one for each time step, in the extended representation. In a percolation process in this extended representation, walkers are restricted to advance only in one direction, along the time axis (fig 2.1) and therefore the percolation process is actually

---

<sup>1</sup>The critical exponents do not change regardless of the time the walker is allowed to remain at a given node, as long as that time is a finite number of time steps

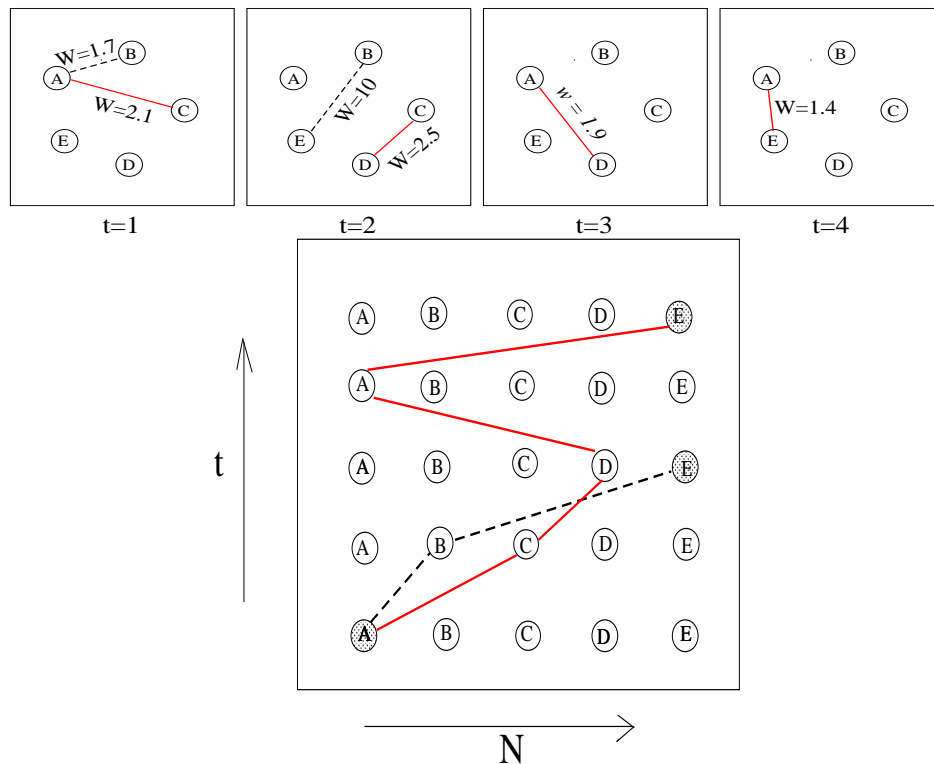


Figure 2.1: (a) A five-node dynamic network at different time steps with weights  $w = a * r_i$  (b) The  $y$  axis corresponds to the time dimension and each successive layer of two rows corresponds to a network configuration at a different time step. The solid-line path is the optimal path (sum of its weights is minimal) between node A and node E, even though a shorter path exists, shown as the dashed line. Note that not all links from (a) are presented, only the links along the paths from A to E, and the weights remain the same as in (a)

a directed percolation process, where the time  $t$  is equivalent to step  $t$ . This correspondence between dynamic networks and DP not only gives a meaning to DP in networks but, more importantly, allows us to apply the results known from the critical scaling of DP to dynamic networks.

Networks can be regarded as infinite dimensional structures, since no spatial constraints exist. Therefore, since our mapping is exact, we expect the critical properties of dynamic networks to be the same as DP in infinite dimensions. The relevant critical properties for DP are [21, 40]:  $S(t)$ , the giant component size, scales as  $S(t) \sim t^2$ ; and  $P_s(t)$ , the survivability (the probability of reaching layer  $t$  when growing a cluster), scales as  $P_s(t) \sim$

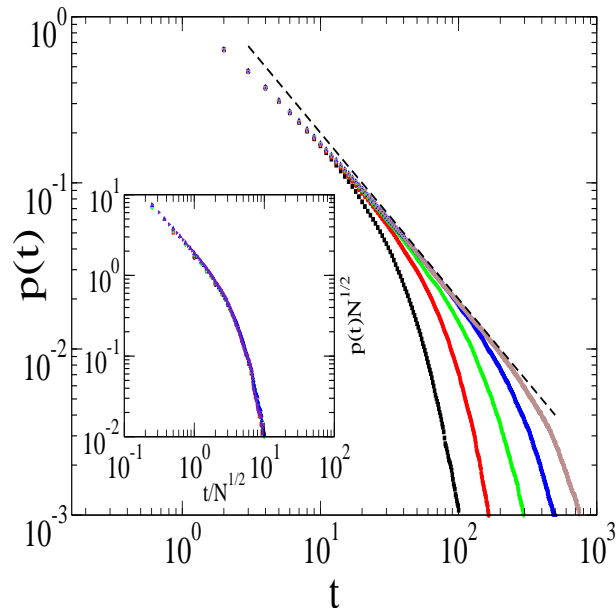


Figure 2.2: Simulation results for the survivability  $P_s(t)$  and its cutoff, at criticality, for dynamic networks of different network sizes. From left to right:  $N = 100, 400, 1600, 6400, 25600$ . The data collapse shown in the inset demonstrates that  $P_s(t)$  in dynamic networks is universal when scaled by  $N^{1/2}$ .

$t^{-1}$ . Figures 2.2(b) and 2.3(a) present simulation results confirming these scaling relations.

To learn about the size-dependent properties of dynamic networks we determine the DP properties as a function of the network size  $N$ , rather than as a function of  $t$ . In DP at criticality, the infinite dimensional relationship between  $w$ , the width in the transverse axes, and  $t$ , the length in the longitudinal axes, is  $w \sim t^{1/2}$ . The upper critical dimension  $d_c$  is the lowest dimension for which the system has the properties of an infinite dimensional system. For DP this value is  $d_c = 4 + 1$  (1 corresponds to the longitudinal axis), so the relation between the system size at the upper critical dimension and the size of a dynamic network is given by  $N \sim w^4$  (the power 4 comes from the 4 transverse dimensions of  $d_c$ ). Since  $w \sim t^{1/2}$  we conclude that:

$$t \sim N^{1/2}. \quad (2.1)$$

Therefore, for a dynamic network of size  $N$  at criticality,  $P_s(t)$  decays exponentially after

a time  $t_\times$ , with  $t_\times \sim N^{1/2}$ <sup>2</sup>. Figure 2.2(b) presents simulation results for the survivability of the giant cluster in a dynamic network at criticality. The figure shows that, for different values of  $N$  and  $t > t_\times$ ,  $P_s(t) \sim t^{-1}$ , as expected from DP in infinite dimensions. The exponential decay for  $t_\times > N^{1/2}$  can also be seen, in agreement with Eq. (2.1). The inset of Fig. 2.2(b) shows the collapse of survivability data after scaling by  $N^{1/2}$ , supporting again Eq. (2.1).

The size of the giant component  $S(N)$ , at criticality, is derived by substituting Eq. (2.1) in the DP relation  $S(t) \sim t^2$ .

$$S(N) \sim N. \quad (2.2)$$

Figure 2.3(b) presents simulation results illustrating this scaling relationship, compared to the known relationship for static networks, where  $S(N)$  is known to scale as  $S(N) \sim N^{2/3}$  [23, 39].  $P_s(t)$  for static networks is also known to decay exponentially after a time  $t_\times \sim N^{1/3}$ . The two systems clearly have different behavior and properties at criticality, and thus belong to two different universality classes.

We find that the behavior of dynamic networks at criticality is universal and independent of the rate  $r$  with which the links are changed. [inset of Fig. 2.4(a)]. However, the critical concentration,  $p_c$ , for which the phase transition occurs does depend on  $r$  [Fig. 2.4(a)].

The dependence of  $p_c$  on  $r$  can be derived as follows: For simplicity assume that—instead of rewiring each link independently at rate  $r$ —with probability  $r$ , all the links are rewired at a given step, and with probability  $1 - r$ , no links are rewired at that step. Consider a node  $i$  with degree  $k_0$  reached by traversing a link on a step followed by a rewiring. Since all links were rewired, this node has  $k_0$  links to new neighbors with probability  $[(N - 1)/N]^{k_0}$ , which is approximately unity for large  $N$ . Each of these links is thus new, and the average branching factor is  $\langle kp \rangle = p \langle k \rangle k$ .

Now suppose that instead there had been no rewiring after traversing a link. In ER

---

<sup>2</sup>Reversing the relation from DP we get  $t(w) \sim w^2$ , implying that if we limit the distance  $w$  (instead of the conventional way in DP of limiting the time  $t$ ) we will obtain a cutoff in the known critical exponents after time  $t_\times \sim w^2$ . The equivalent for networks would be a cutoff after a crossover time  $t_\times \sim N^{1/2}$  (based on Eq. (2.1)).

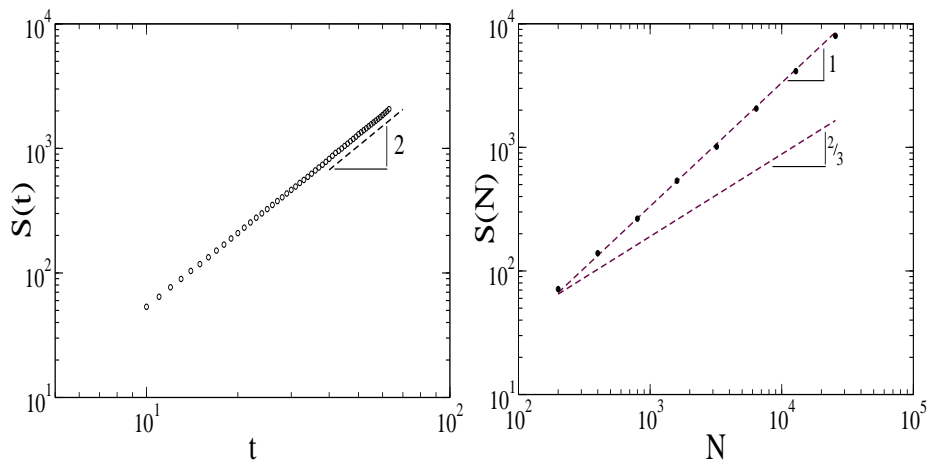


Figure 2.3: (a) The cluster size  $S(t)$  scales as  $t^2$ . (b) Simulation results (dots) supporting the relation  $S(N) \sim N$  (Eq. 3) (upper dashed line), the cluster size in dynamic networks, compared to the known  $S(N) \sim N^{2/3}$  (bottom dashed line), the cluster size in static networks.

networks, each node has on average  $p\langle k \rangle$  outgoing neighbors as before, in addition to the link through which it was reached that points to the originating (parent) node<sup>3</sup> Links to the parent node, however, also exist for all other nodes reachable from the same parent (siblings). The number of neighbors,  $x$ , of the parent (excluding node  $i$ ) is Poisson distributed with mean  $\langle k \rangle$ . Assuming node  $i$  has already reached its parent, the other  $k_0$  siblings each have a probability  $p$  to be reached and then a probability  $p$  to return. Thus, the number of neighbors,  $x$ , that can return to the parent is binomially distributed as

$$P(x) = \binom{k_0}{x} p^{2x} (1 - p^2)^{k_0 - x}. \quad (2.3)$$

When calculating the branching factor at each step, we should count the parent node only once. The contribution of the link back to the parent to the branching factor of each of the parent's siblings is therefore inversely proportional to the total number of siblings. The

<sup>3</sup>For static ER networks, the link back to the parent would not be counted, since revisiting a node will not lead to exploring new paths. For dynamic networks (where the links are rewired) even if a zero fraction of the nodes are rewired at that step, revisiting a node will eventually lead to exploring new paths and to a change in  $p_c$ . Since the formula for  $p_c$  assumes that a node can be revisited for all values of  $r$ , the limit  $p_c\langle k \rangle = 1$  is not recovered for  $r = 0$ .



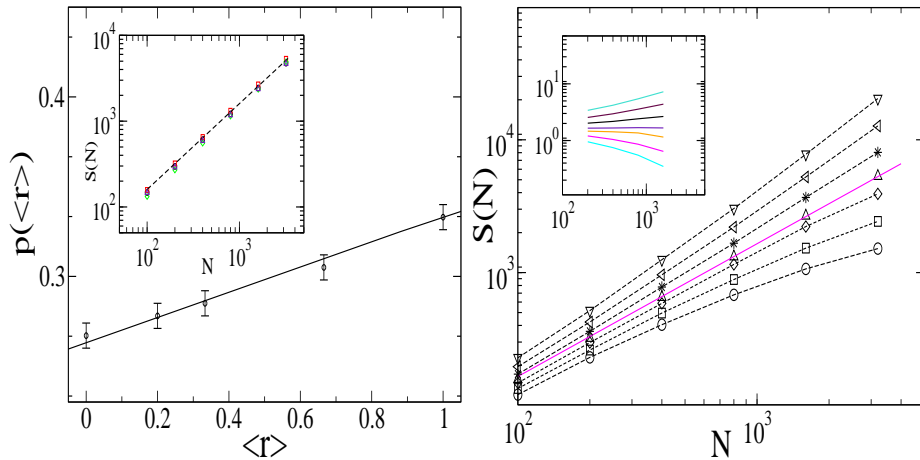


Figure 2.4: (a) Simulation fit to the formula for  $p_c$  as a function of the rate  $\langle r \rangle$  at which the links change. The value of  $p_c$  was calculated by finding the value of  $p$  for which  $S(N)$  gives a straight line. In the inset we show the value of  $S(N)$  at  $p_c$  for several different values of  $\langle r \rangle$ . In all cases the slope equals 1 as predicted by the critical exponent for dynamic networks. (b) Simulation results for  $S(N)$  in a network with  $\langle k \rangle = 3$  and  $\langle r \rangle = 2/3$  are presented for different values of  $p$  showing that the network undergoes a phase transition at  $p_c$ .

average contribution of a sibling with degree  $k_0$  is

$$\left\langle \frac{1}{x+1} \right\rangle_{k_0} = \sum_{x=0}^{k_0} \binom{x}{k_0} \frac{p^{2x}(1-p^2)^x}{x+1} = \frac{1 - (1-p^2)^{k+1}}{(k+1)p^2}. \quad (2.4)$$

Using the fact that  $x$  is Poisson distributed, the average contribution from all nodes is

$$\left\langle \frac{1}{x+1} \right\rangle = \sum_{k_0=0}^{\infty} \left\langle \frac{1}{x+1} \right\rangle_{k_0} \frac{e^{-\langle k \rangle} \langle k \rangle^{k_0}}{k_0!} = \frac{1 - e^{-p^2 \langle k \rangle}}{p^2 \langle k \rangle} \equiv f(p). \quad (2.5)$$

After  $n$  steps, we have on average  $nr$  steps with rewiring and  $n(1-r)$  steps without rewiring. Thus, the branching factor is

$$(p \langle k \rangle)^{nr} (p \langle k + f(p) \rangle)^{n(1-r)}. \quad (2.6)$$

For the process to be at criticality, this factor should be 1, leading to

$$p_c \langle k \rangle g(p_c) = 1, \quad (2.7)$$

where

$$g(p_c) = \left[ 1 + \frac{f(p_c)}{\langle k \rangle} \right]^{1-r}. \quad (2.8)$$

The agreement of our simulations with the solution of Eq. (2.7) for  $p_c$  is shown in Fig. 2.4(a).

Fig. 2.4(b) presents simulation results for  $S(N)$  at different values of  $p$  indicating the network undergoes a phase transition at some critical value of  $p$ . The inset of Fig. 2.4(a) shows that for several different values of  $r$   $S(N) \sim N$  at  $p_c$ , as expected due to the universality of the critical exponent in dynamic networks.

The correspondence to DP can also predict the general scaling of the optimal path in a dynamic network with a broad distribution of disorder. In a network where weights are assigned to links, the optimal path between any two nodes is defined as the path along which the sum of the weights is minimal. In the limit of a broad distribution of disorder, Ref. [41] has shown that, at criticality, the optimal path exists mainly along the giant cluster. Therefore for static networks the optimal path length scales with the average distance between nodes on the percolation cluster:  $\ell_{\text{opt}} \sim N^{1/3}$  (See Sec. 1.5). In our dynamic network model the average distance between nodes on the percolation cluster scales as  $\langle \ell \rangle \sim N^{1/2}$ , suggesting that in dynamic ER networks the optimal path scales as

$$\ell_{\text{opt}} \sim N^{1/2}. \quad (2.9)$$

Figure 2.5 shows simulation results for the optimal path length in a dynamic network compared to a static network. The results for dynamic networks are in full agreement with Eq. (2.9).

What makes the results in a dynamic network so different from the static case? The difference lies in the number of available configurations. While in static networks the percolation cluster is composed from paths built from  $N$  network nodes, in dynamic networks the network is represented by  $N^{3/2}$  nodes [42]. The evolution of the network over time generates many more possible configurations, enabling the percolation cluster to become much larger. Substituting  $N' = N^{3/2}$  in the percolation cluster formula for static networks,

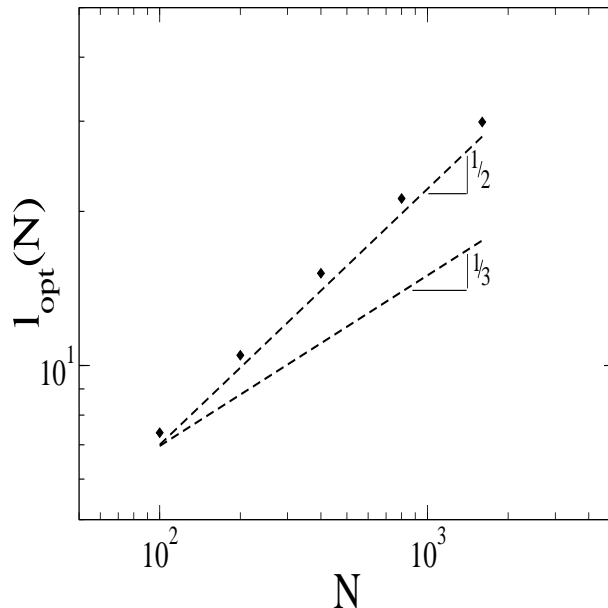


Figure 2.5: The optimal path for strong disorder scales as  $\ell(N) \sim N^{1/2}$  in dynamic networks compared to  $\ell(N) \sim N^{1/3}$  in static networks.

$S(N) \sim N^{2/3}$ , yields  $S(N') \sim S(N^{3/2}) \sim N$ , which further confirms our results for  $S(N)$  in dynamic networks.

The same explanation is also true for the optimal path, where a longer more optimal path is available due to the increased number of available configurations. For example, an optimal path reaching some node A may find it optimal not to advance to the near neighbor B at the next step, but rather to first visit C, and then only later come back to B, since at the later time node B is more optimally connected to the destination node.

Representing a dynamic network as a directed network [Fig. 2.2(a)] composed of  $N^{3/2}$  nodes allows the “same” node to be counted more than once in the percolation cluster, therefore requiring that the distinct number of nodes on the percolation cluster also scale with  $N$ . To determine the number of different nodes of the original network in a component of size  $M$  on the directed network, consider the following argument: The links between consecutive layers of the directed network are chosen randomly. Therefore, each link leads to a random node in the original network independently and with uniform distribution. The

probability to reach a new node by following a link, assuming that  $D$  nodes have already been visited, is  $1 - D/N$ . The expected number of distinct nodes  $E(D)$  reached after  $\lambda$  links have been followed from the starting node is therefore  $E(D_\lambda)/N = E(D_{\lambda-1}) + E(1 - D_{\lambda-1}/N)$ . This reduces to  $E(D_\lambda) = 1 + (1 - 1/N)E(D_{\lambda-1})$  which indicates that for large  $M$

$$\frac{E(D_M)}{N} = 1 - \left(1 - \frac{1}{N}\right)^M \approx 1 - e^{-M/N}. \quad (2.10)$$

Thus, when the size of a component in the directed network is of order  $N$  a finite fraction of the visited nodes are new and the size of the induced component on the original network is also of order  $N$ .

In summary, we introduced a model for dynamic networks which was solved by a comparison with directed percolation in  $4 + 1$  dimensions. The DP longitudinal axis is mapped to the time axis along which the dynamic network evolves. We showed that dynamic networks exhibit different properties and critical exponents near criticality. Therefore they belong to a different universality class than static networks. While in static networks  $S(N)$ , the size of the giant component at criticality, scales as  $S(N) \sim N^{2/3}$ , in dynamic networks  $S(N) \sim N$ . Even though the properties of dynamic networks are universal and independent of the rate  $r$  at which the links are changed, the critical concentration,  $p_c$ , for which the phase transition occurs, depends on  $r$ . We also showed that the optimal path in dynamic networks scales as  $\ell_{\text{opt}} \sim N^{1/2}$ , compared to  $\ell_{\text{opt}} \sim N^{1/3}$  in static networks.

## Chapter 3

# Quarantine Generated Phase Transition in Epidemic Spreading

### 3.1 Introduction

The representation of interactions in a human society as a complex network, where nodes and links play the respective roles of individuals and their contacts, has been useful for modeling, studying, and understanding many problems in epidemiology [19, 43]. Usually it is assumed that diseases evolve faster than the topological evolution of the underlying network, so that the links of the network can be regarded as static. With modern mass media, however, the presence of an epidemic can be broadcast much faster than disease propagation. This information will inevitably change the behavior of the individuals comprising this network as, for example, they try to avoid contacts with infected people. In this way a feedback loop between the state of individuals and the topology of the network is formed. Networks that exhibit such feedback are called adaptive or coevolutionary networks [44–48].

Public health services are constantly searching for new ways to try to reduce the spread of diseases. Interventions like vaccination [49] or total closure of workplaces and schools are very effective, but come with a high economic cost. As a less expensive alternative we examine here the effectiveness of a “quarantine” strategy, where healthy people are “advised” to avoid contacts with individuals that carry the disease. That is, a healthy person has a chance to suppress a contact with an infected neighbor and form a new tie with another

healthy peer (rewiring). The value of this rewiring probability could depend, for instance, on the concern that the society has about the disease and, as suggested above, in a globalized world this concern will depend on the how broadly the news is published. The degree of media attention about the disease is a parameter that could be controlled by public health services. Indeed, it is known that spontaneous quarantine in the recent H1N1 pandemic was found to have a large impact in reducing the final size of the epidemic [50].

Based on these observations, we propose two strategies for the propagation of epidemics on an adaptive network, in which individuals alter their local neighborhoods with constant quarantine probability  $w$  as described above, and systematically study the effect that the quarantine has on epidemic spreading. We find a phase transition at a critical threshold  $w_c$  above which the epidemic is stopped from spreading. We show also how the epidemic spreading in the presence of quarantine depends on the contagion and recovery parameters. More importantly, we find that the initial structure of the network plays an essential role in disease propagation at criticality, unlike in previous related models [47, 51–53] where results only depend on the average network connectivity. We also introduce a generalized form of quarantine, where the probability of rewiring is proportional to the degree of the infected nodes, producing a more efficient isolation of the nodes with high degree. This preferential rewiring is more efficient than the case with  $w$  constant.

## 3.2 Analytical Approach

We consider the susceptible-infected-recovered (SIR) epidemic spreading model, which is well established and accurately describes diseases such as seasonal influenza, SARS, or AIDS [49, 54]. Initially, all nodes are in the susceptible state ( $s$ ), with one node chosen at random (seed) in the infected state ( $i$ ). In the first strategy, strategy  $A$ , at each time unit, every infected node  $i$  in the network tries to transmit the disease to each susceptible neighbor  $s$  with infection probability  $\beta$ . If  $s$  does not get infected, then with rewiring probability  $w$  it disconnects its link from  $i$  and reconnects it to another randomly chosen susceptible node, different from its present neighbors. Thus, the rewiring probability  $w$  measures how fast

susceptible nodes react to the disease (the quarantine probability). Infected nodes recover ( $r$ ) after a fixed recovery time  $t_R$  since they first became infected, remaining in the recovered state forever.

To estimate  $w_c$  we start by assuming that the network has a tree structure, so that the disease spreads out from the seed and reaches a susceptible node  $s$  through only one of its neighbors  $i$ . This assumption is valid at and below criticality, as the low density of infected nodes makes collisions between infected branches (loops) unlikely. Then, if  $i$  becomes infected at time  $t_0$ , the probability that  $s$  becomes infected by  $i$  at time  $t_0 + n$ , with  $n = 1, 2, \dots, t_R$ , is  $\beta(1 - \beta)^{n-1}(1 - w)^{n-1}$ . This is the probability that  $s$  has neither become infected nor disconnected from  $i$  up to time  $t_0 + n - 1$ , times the probability that  $i$  succeeds in transmitting the disease to  $s$  at time  $t_0 + n$ . Therefore, the overall probability that  $s$  becomes infected before  $i$  recovers is given by the sum

$$\begin{aligned} T_{\beta,w}^A &\equiv \sum_{n=1}^{t_R} \beta(1 - \beta)^{n-1}(1 - w)^{n-1} \\ &= \frac{\beta \left\{ 1 - [(1 - \beta)(1 - w)]^{t_R} \right\}}{1 - (1 - \beta)(1 - w)}. \end{aligned} \quad (3.1)$$

This expression for the transmissibility  $T_{\beta,w}^A$  is equivalent to the corresponding expression in the standard SIR model  $\sum_{n=1}^{n=t_R} \beta(1 - \beta)^{n-1}$  [19], but with a non-infection probability  $(1 - \beta)(1 - w)$ , instead of  $(1 - \beta)$ . When  $w = 0$ , Eq. (3.1) reproduces the known value for the transmissibility  $T = 1 - (1 - \beta)^{t_R}$  for the SIR model on static networks [19]. In this formulation,  $w$  plays the role of a control parameter of the transmissibility: for fixed values of  $t_R$  and  $\beta$ ,  $T_{\beta,w}^A$  can be reduced by increasing  $w$ . By reducing  $T_{\beta,w}^A$  we can go from a regime in which the epidemic spreads over the population (epidemic phase) to another regime where the disease cannot spread (disease-free phase). The transition from the disease-free phase to the epidemic phase corresponds to the average number of secondary infections per infected node becoming larger than one, allowing the long-term survival of the disease and thus ensuring the epidemic spreads to a large fraction of the population. In our problem, the expected number of susceptible neighbors that a node has when it just becomes infected is

given by  $\kappa - 1$ , where  $\kappa - 1 \equiv \langle k^2 \rangle / \langle k \rangle - 1$  is called the *branching factor*, and  $\langle k \rangle$  and  $\langle k^2 \rangle$  are the first and second moments, respectively, of the degree distribution  $P(k)$ . Since  $T_{\beta,w}^A$  is the overall probability to infect a neighbor, the mean number of secondary infections per infected node is

$$N_I^A(w) = (\kappa - 1)T_{\beta,w}^A. \quad (3.2)$$

The infection will die out if each infected node does not spawn on average at least one replacement, so for a very large system, the critical point is given by the relation

$(\kappa - 1)T_{\beta,w_c}^A = 1$ , or

$$\frac{(\kappa - 1)\beta \left\{ 1 - [(1 - \beta)(1 - w_c)]^{t_R} \right\}}{1 - (1 - \beta)(1 - w_c)} = 1. \quad (3.3)$$

The transition between free-disease phase and epidemic phase is analogous to the *static link percolation* problem [30], in which each link in a network is occupied with probability  $p$  and empty with probability  $q = (1 - p)$ . When  $p$  becomes lower than a percolation threshold  $p_c$ , the giant connected component disappears. In general,  $p_c$  depends on the size of the network  $N$ [55], but in the thermodynamic limit it can be expressed as  $p_c = 1/(\kappa - 1)$  [3]. Identifying  $p$  with the transmissibility  $T_{\beta,w}^A$ , and using the relation between  $p_c$  and  $\kappa$  we find that on the epidemic/disease-free transition line:

$$T_{\beta,w_c}^A = p_c. \quad (3.4)$$

This result shows that the transition point depends on the initial topology of the network through the moments of the degree distribution, as we shall confirm via simulation.

A better strategy would be to try and avoid contact between susceptible and infected individuals before the attempt at infection, i.e., only avoid the contact when you know an individual is sick. In this second strategy, strategy B, at each time unit every susceptible node attached to an infected node disconnects its link from  $i$  with probability  $w$  and reconnects it to another randomly chosen susceptible node, different from its present neighbors.



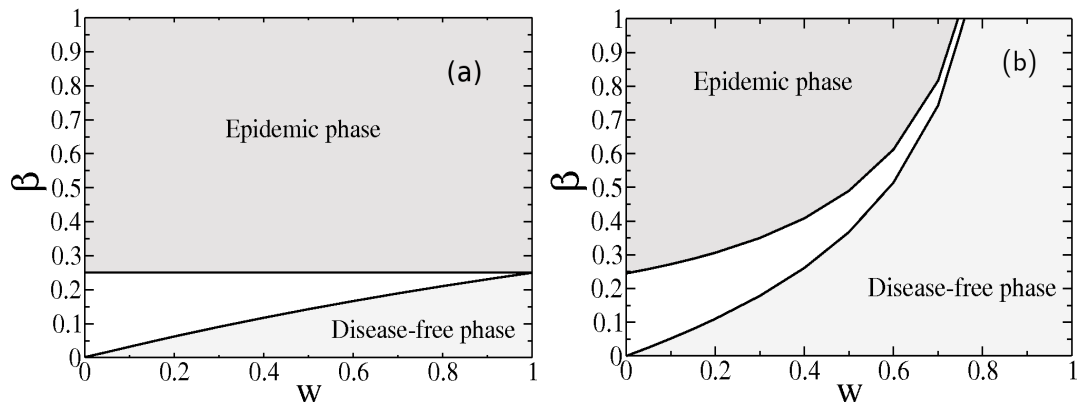


Figure 3.1:  $\beta$ - $w$  phase diagram for  $p_c = 0.25$ . The curves correspond to  $t_R = 1$  (upper) and  $t_R = \infty$  (lower). The dark gray region will always be an epidemic phase, and the light gray always a disease-free phase: (a) strategy A, (b) strategy B.

If  $s$  does not rewire its link, the infected node  $i$  tries to infect it with infection probability  $\beta$ . Notice that  $N_I^B(w) = (1 - w)N_I^A(w)$ . For this strategy Eq.(3.1) is replaced by

$$T_{\beta,w}^B \equiv \frac{\beta(1-w) \left\{ 1 - [(1-\beta)(1-w)]^{t_R} \right\}}{1 - (1-\beta)(1-w)}. \quad (3.5)$$

Comparing Eq.(3.1) and Eq.(3.5) we see that  $T_{\beta,w}^B < T_{\beta,w}^A$  for identical values of  $\beta$  and  $w$ . The new condition for the disease to die out in strategy B follows from Eq.(3.5)

$$\frac{(\kappa - 1)\beta(1-w_c) \left\{ 1 - [(1-\beta)(1-w_c)]^{t_R} \right\}}{1 - (1-\beta)(1-w_c)} = 1. \quad (3.6)$$

The two strategies represent different scenarios depending on knowledge of the state of infection of the nodes. For strategy A, susceptible nodes have no information about the state of their neighbors until they are in physical contact. By contrast, for strategy B, susceptible nodes know the state of their neighbors before they are in physical contact. We will show that this difference in the knowledge of the states of the nodes results in strategy B being more effective at stopping epidemic spreading.

Fig. 3.1 shows phase diagrams in the  $w - \beta$  plane for strategy A and strategy B, obtained

by the numerical solutions of Eq. (3.3) and Eq. (3.6) respectively, for a network with  $\kappa = 5$  ( $p_c = 0.25$ ). Two phases emerge: the epidemic phase for  $T_{\beta,w}^{A,B} > p_c$ , and the disease-free phase for  $T_{\beta,w}^{A,B} \leq p_c$ . The location of the critical line separating the two phases depends on the recovery time  $t_R$ . Fig. 3.1(a) illustrates the two limiting cases for strategy A:  $t_R = 1$  (upper curve) and  $t_R = \infty$  (lower curve), which are given by the expressions  $\beta(t_R = 1) = 1/(\kappa - 1)$  and  $\beta(t_R = \infty) = w/(\kappa + w - 2)$ . Therefore, the pairs of  $(w, \beta)$  values in the region above the curve  $\beta(t_R = 1)$  are always in the epidemic phase, and below the curve  $\beta(t_R = \infty)$  are always in the disease-free phase. A striking consequence is that if  $\beta$  is larger than the percolation threshold  $1/(\kappa - 1)$ , the propagation of an epidemic cannot be stopped, even with the largest rewiring probability  $w = 1$ . Thus, strategy A is not an efficient mechanism to control an epidemic when  $\beta$  is higher than  $p_c$ .

For strategy B, Fig. 3.1(b) illustrates the two limiting cases  $t_R = 1$  (upper curve) and  $t_R = \infty$  (lower curve), which are given by the expressions  $\beta(t_R = 1) = 1/(\kappa - 1)(1 - w)$  and  $\beta(t_R = \infty) = w/(\kappa - \kappa w + 2w - 2)$ . In contrast to strategy A above, if  $\kappa$  does not diverge, i.e., the network has finite  $p_c$ , the epidemic can always be stopped, even for  $\beta \rightarrow 1$ . Notice that the maximum value of  $w$  needed to stop a epidemic is  $w_c = (\kappa - 2)/(\kappa - 1) = 1 - p_c$ .

Our theoretical predictions illustrate a novel feature about the dynamics of adaptive networks in SIR models. While previous adaptive SIS and SIRS models have transition values that depend only on the average connectivity of the network  $\langle k \rangle$  and are independent of the heterogeneity or structural correlations of the initial topology [44], our SIR model predicts dependence on the topological structure through the higher-order moments of the degree distribution.

### 3.3 Simulation Results

Our analytical approach predicts, through Eq. (3.3) and Eq. (3.6), a dependence of  $w_c$  on the initial network. In order to test and explore this dependence, we performed extensive numerical simulations of our model starting from different topologies. We first used Erdős-Rényi (ER) networks with Poissonian degree distribution  $P(k) = e^{-\langle k \rangle} \langle k \rangle^k / k!$ . Even though

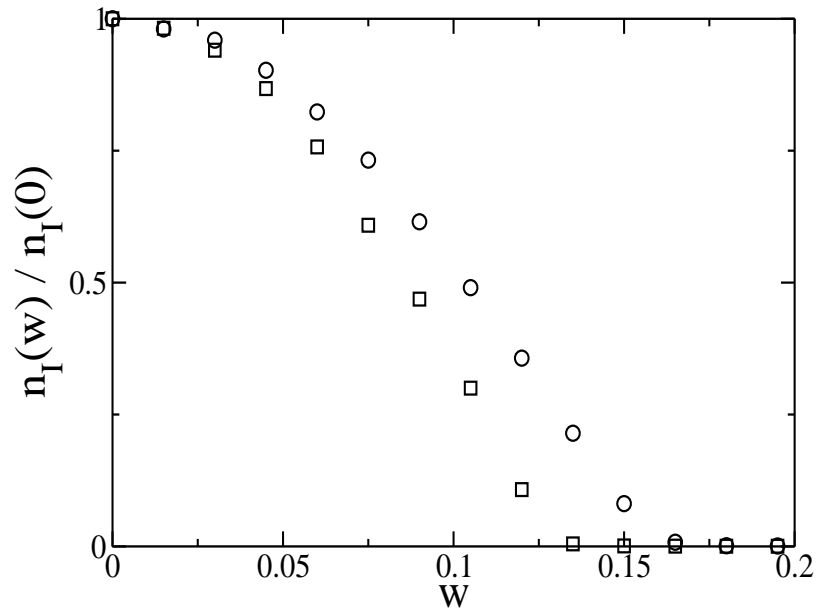


Figure 3.2: Plot of  $n_I(w)/n_I(0)$  as a function of  $w$  for strategy A ( $\circ$ ), and strategy B ( $\square$ ), for a ER network with  $\langle k \rangle_{GC} = 4.07$ ,  $N = 10^4$  and  $10^4$  realizations. As we predicted previously, strategy B is better than strategy A, shown by  $w_c^B < w_c^A$ .

these types of homogeneous networks are common in nature, many real social networks are well represented by heterogeneous networks. Thus, we also used finite scale-free (SF) networks with degree distribution  $P(k) = k^{-\lambda} \exp(-k/K) / \text{Li}_\lambda(e^{-1/K})$ , where  $K$  is the degree cutoff. This distribution represents networks with a finite threshold  $p_c$ , and appears in a variety of real-world networks [23, 56]. We only consider epidemic propagation on the largest connected cluster of the network, the giant component (GC).

In Fig. 3.2 we compare both strategies for ER networks. We plot, as a function of  $w$ , the average fraction of infected nodes  $n_I(w) = N_I(w)/N_{GC}$ , where  $N_{GC}$  is the size of the giant component, on a ER network. This fraction is normalized by the corresponding fraction on an identical fixed network— i.e., for the SIR model without quarantine,  $n_I(0)$ . We can see that  $w_c$  for strategy B is lower than  $w_c$  for strategy A, as expected. This relation will hold for any topology with the same  $\kappa$  (see Eq.(3.3) and Eq.(3.6)). Since strategy B is more effective, from here to the end of the section we will show only simulation results for strategy B. Fig. 3.3 shows,  $n_I(w)/n_I(0)$  vs  $w$  for different  $\beta$ . In this figure, we can observe

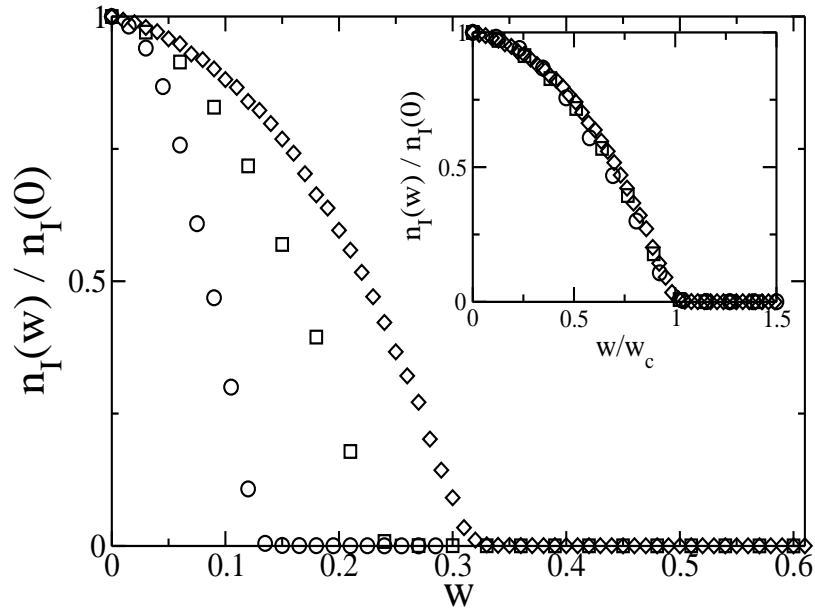


Figure 3.3: Fraction of nodes infected as a function of the applied quarantine parameter  $w$  divided by the fraction of infected nodes at  $w = 0$  for three different infection probabilities  $\beta$  in an ER network. ( $\circ$ )  $\beta = 0.05$ , ( $\square$ )  $\beta = 0.1$  and ( $\diamond$ )  $\beta = 0.15$ . Above a threshold value of  $w$ , no finite fraction of the network becomes infected. (Inset) Data with  $w$  rescaled by the appropriate  $w_c$  calculated from Eq. (3.6). The curves collapse very well, showing universal behavior and good agreement with the theory. In all simulations,  $\langle k \rangle_{GC} = 4.07$  ( $\langle k \rangle = 4$  over all nodes),  $N = 10^4$  and averages are over  $10^4$  realizations.

the strong effect of quarantine and the critical threshold  $w_c$  above which  $n_I(w)$  approaches zero. Thus, in the disease-free phase ( $w \geq w_c$ ) only a small number of individuals get infected and the disease quickly dies out. We observe, as expected, that the values of  $w_c$  increase with  $\beta$ . This behavior matches the phase diagram of Fig. 3.1 (b), where the critical line has a positive slope. Scaling the horizontal axis by the values of  $w_c$  obtained by numerically solving Eq. (3.6) collapses the curves, showing an excellent agreement between theory and simulations, as well as a scaling behavior of the form  $n_I(w) = n_I(0)f(w/w_c)$ , as shown in the inset of Fig. 3.3.

We also explored the dependence of  $w_c$  on the connectivity of the network, by computing  $n_I(w)/n_I(0)$  vs  $w$  for different  $\langle k \rangle$  (see Fig. 3.4). The results show that  $w_c$  increases with  $\langle k \rangle$ , due to the fact that propagation is facilitated by having more neighbors, as in the

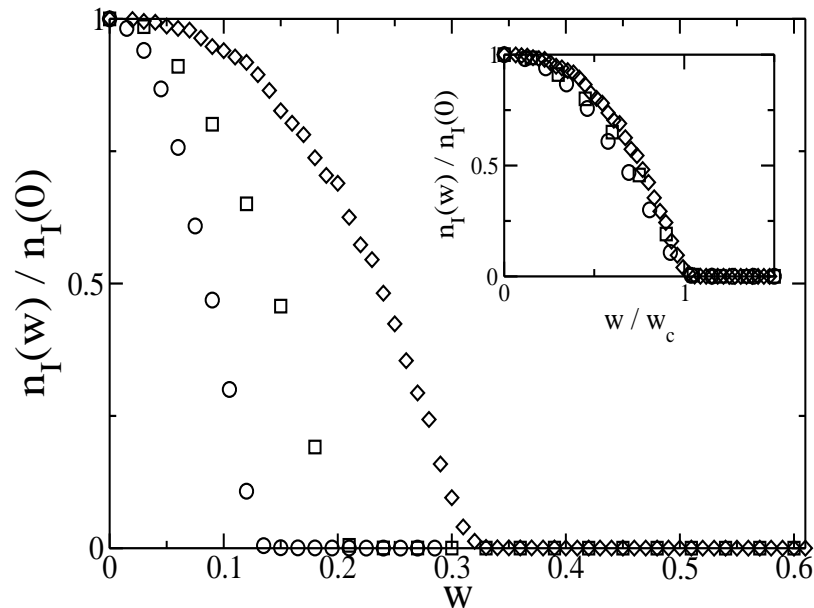


Figure 3.4: Fraction of nodes infected as a function of the applied quarantine parameter  $w$  divided by the fraction of infected nodes at  $w = 0$  for ER networks with three different  $\langle k \rangle$ :  $\langle k \rangle_{GC} = 4.07$  ( $\circ$ ),  $\langle k \rangle_{GC} = 6.015$  ( $\square$ ),  $\langle k \rangle_{GC} = 10$  ( $\diamond$ ). Again, while below  $w_c$  a finite fraction of the network become infected above  $w_c$  the epidemic vanishes. (Inset) Data rescaled by  $w_c$  calculated from Eq(3.6). In all simulations,  $\beta = 0.05$ ,  $N = 10^4$  and averages are over  $10^4$  realizations.

original SIR dynamics. The inset of Fig. 3.4 shows the collapse of all curves. Again, the good agreement between theory and simulation confirms that Eq. (3.6) is a valid expression of the transition point for the adaptive SIR model on ER networks.

Given that the second moment, and therefore  $\kappa$ , is large in heterogeneous networks, the critical value  $w_c$  turns out to be larger in heterogeneous networks than in homogeneous networks with the same mean degree  $\langle k \rangle$  and size  $N$ , as seen by considering Eq. (3.6) in the  $t_R \gg 1$  limit, where

$$w_c \simeq \frac{1}{\frac{1}{\beta(\kappa-2)} + 1}. \quad (3.7)$$

Since  $\kappa$  for heterogeneous networks is much bigger than in homogeneous networks we expect that  $w_c$  increases as the heterogeneity increases with larger  $K$ . For  $\kappa \rightarrow \infty$  we expect that  $w_c \rightarrow 1$  and that the transition will eventually disappear for large heterogeneous networks.

In Fig. 3.5, we show simulations on scale-free networks for different values of  $K$ . If we let

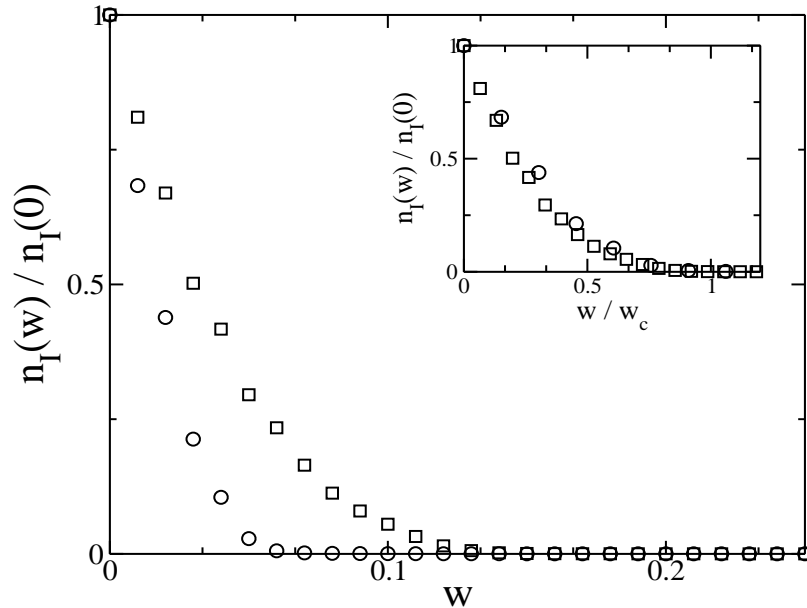


Figure 3.5: Fraction of infected nodes still infected as a function of  $w$  for  $\beta = 0.05$ ,  $t_R = 20$  on an SF network with  $K = 10$  ( $\circ$ ) and  $K = 20$  ( $\square$ ),  $\lambda = 2.1$  with  $N = 10^5$ .  $p_c \cong 0.375$  for  $K = 10$  and  $p_c \cong 0.215$  for  $K = 20$ . (Inset) Rescaled by  $w_c$ . The result shows an excellent agreement between the theory and the simulation.

$K \rightarrow \infty$  we have to replace  $p_c$  with  $p_c(N)$  in Eq. (3.4). Recent observations show that only in networks larger than  $\approx 10^9$  nodes is the approximation  $p_c \cong p_c(N)$  valid [55]. However, as we limit ourself to the case of small  $K$  this is not an issue in our simulations, and we see that in good agreement with our predictions from Eq. (3.7), as  $\kappa$  increases with  $K$ ,  $w_c$  increases. In the inset we rescale by  $w_c$  obtained from Eq. (3.6), and find good collapse (Fig. 3.5(inset)), confirming the general validity of Eq. (3.6) for heterogeneous SF networks.

In very heterogeneous networks it is well known that, due to high-degree nodes, propagation processes are very difficult to stop [30, 31]. Thus, a constant probability of rewiring is not effective for controlling epidemics in these networks, because such a strategy assumes all nodes have the same importance in an epidemic, ignoring the function of higher-degree nodes as superspreaders. It is also well known that removing the high-degree nodes of the network (targeted percolation) [30, 57, 58] is a more efficient method to stop propaga-

tion processes than random removal. This type of strategy is expected to be superior, but requires global information about the network.

To test this prediction, we propose a new strategy of type B where  $w$  depends on the degree  $k$  of the infected node, with  $w_k$  given by the general form

$$w_k \equiv w_k(\alpha) = \gamma k^\alpha, \quad (3.8)$$

where  $\gamma$  is a constant that controls the highest possible value of  $w_k$ , and it is equal to or smaller than  $k_{max}^{-\alpha}$ , where  $k_{max}$  is the largest degree of the network. For  $\alpha = 0$  and  $\gamma \in [0,1]$  we recover the results for a constant value of  $w$ , with  $w = \gamma$ . For  $\alpha > 0$  and  $\gamma = k_{max}^{-\alpha}$  the rewiring increases with the degree  $k$  of the infected node, and decreases with increasing  $\alpha$ . In the limit of  $\alpha \rightarrow \infty$  the rewiring process is equivalent to a targeted rewiring where only the links of the highest connected node(s) are rewired. To compare the cases with  $\alpha = 0$  and  $\alpha > 0$ , we use the average  $\langle w \rangle$  over the network, choosing  $\alpha$  for the targeted case such that  $\langle w \rangle$  is equal to  $w$  in the uniform case.

In Fig. 3.6 we plot  $n_I(w)/n_I(0)$  as a function of  $\langle w \rangle$  for SF networks with  $K = 10$  and  $K = 20$ . As expected,  $w_c$  for  $\alpha > 0$  is lower than for  $\alpha = 0$ . The preferential rewiring reduces the value of  $w_c$  because quarantine is more effective at isolating the superspreaders.

### 3.4 Summary and Conclusions

We have introduced and studied two strategies for the propagation of epidemics on evolving networks of social contacts. The states of the individuals are changed according to SIR dynamics, while the network evolves according to a quarantine mechanism based on local information, in which susceptible individuals replace their infected neighbors by other susceptible peers with probability  $w$ . We demonstrated by an analytical approach, and confirmed by numerical simulations, that the size of the epidemics can be largely reduced by increasing the probability of rewiring, and that the propagation can be eventually stopped by using high enough values of  $w$ . In other words, quarantine is an effective way to halt

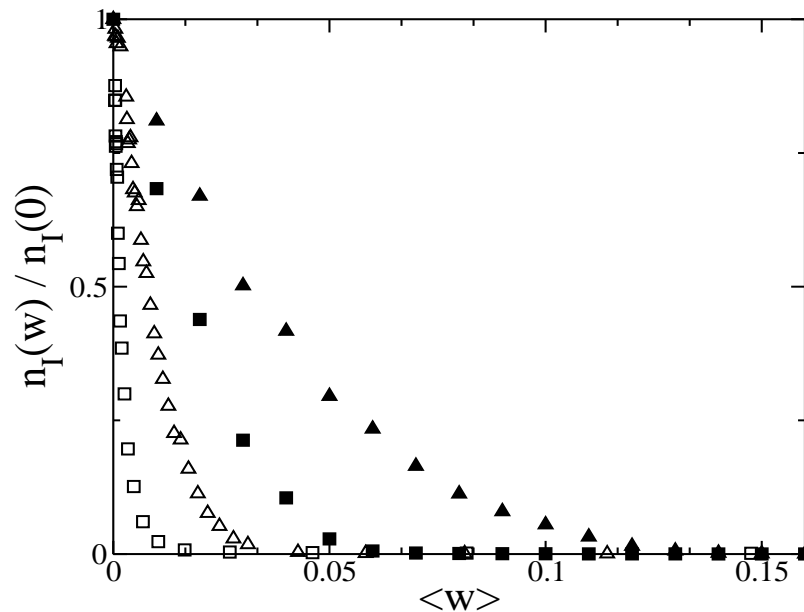


Figure 3.6: Fraction of infected nodes still infected as a function of  $\langle w \rangle$  for  $\beta = 0.05$ ,  $t_R = 20$  on an SF network,  $\lambda = 2.1$ ,  $N = 10^4$  and  $K = 10, 20$ . For general strategy B with  $\alpha = 0$ , ( $K = 10$  (■) and  $K = 20$  (▲)), and with  $\alpha > 0$ , ( $K = 10$  (□) and  $K = 20$  (△)), each  $\langle w \rangle$  represents a different  $\alpha$ .

the appearance of epidemics that would otherwise emerge in the case of a static network. For strategy A, when the infection probability is larger than a threshold, the quarantine mechanism is not effective any more and disease propagation becomes unavoidable. This is because the quarantine model only breaks contact after allowing for a chance of infection, thus for high enough infection probability the spreading is irreversible, even with a rewiring probability of unity. For strategy B, the quarantine mechanism is effective in homogeneous networks and finite heterogeneous networks, even for infection probability equal to unity. The transition disappears only for large, very heterogeneous networks. In these scale-free networks, the critical rewiring probability becomes very large and the transition eventually disappears for large enough systems; thus, only the epidemic phase is observed. This is likely due to the presence of individuals with very large connectivity that can spread the disease over a large fraction of the population, even for small infection probabilities. To confirm this, we introduced a generalized form of strategy B, where the quarantine depends



on the degree of the infected nodes. This preferential rewiring isolates the superspreaders more efficiently, reducing  $w_c$  and preserving a finite transition even for scale-free networks. Lastly, in SIR dynamics the final frozen state, where everybody is either recovered or susceptible, is reached rather quickly—in a time of order  $\ln N$  according to our simulations—and thus an SIR network does not evolve much, so the initial topology is preserved during the entire evolution of the system. In contrast, in adaptive models with SIS or SIRS dynamics, the system evolves for a very long time in the epidemic phase—times that grow exponentially larger with  $N$ —thus the network moves towards a stationary topology similar to an Erdős-Rényi network, independent of the initial topology, and rendering the initial topology irrelevant. Therefore, unlike other adaptive network models of epidemic spreading, in our model the epidemic threshold has a strong correlation with the topology of the network, which remains relatively unchanged at criticality, and the social structure of connections when epidemics begin to propagate is crucial in the state of the final outbreak. We are currently extending this work to weighted networks [41].

## Chapter 4

# Epidemics on Interacting Networks

### 4.1 Introduction

Complex network models of the interactions in human society have been used to understand many problems in epidemiology [19, 31, 43, 59–62]. These models have generally assumed that all of the nodes have interacted on a single network with a single degree distribution. Even when these degree distributions allow for large heterogeneities—as in the case of scale-free networks [57], where hubs with large numbers of connections can arise—the assumption remains that every node is part of a single network and is represented by a single underlying topology. In reality, however, societies are composed of many interconnected networks, as in Fig. 4.1, which may be communities within a larger population or separate systems entirely. A disease can spread through the network of direct personal contacts, via the water utilities network, and through travel from city to city over highway or airline networks. These interconnected network systems may be comprised of different types of nodes, which may have degrees drawn from distinct degree distributions, and may have different connectivities between them.

Interconnected network systems have been of interest to researchers in numerous different ways [63, 65–67]. Interconnected dependency networks, where failure in nodes in one network causes failures in connected nodes in the connected network, exhibit failure cascades, where the cross network dependencies result in a network much more easily frag-

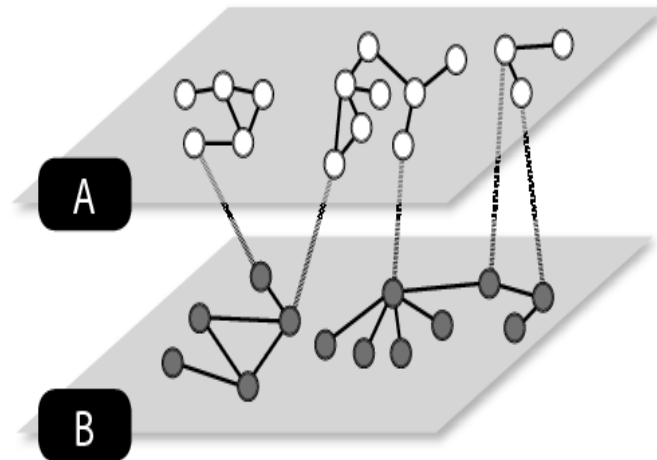


Figure 4.1: An interconnected network system with two networks: A and B. Nodes have intranetwork links within their own network, but also internetwork links connecting them to the other network.

mented than single networks of the same degree distribution [68]. Interconnected power networks, where transport capacity and failure vulnerability are competing properties, were examined and an optimal level of interconnection found [69]. Networks without dependencies, such as interconnected social networks, where populations exist at city, state, and national levels, have also been examined. In these networks, the level of movement between cities (the interconnections between them) have been shown to affect the epidemic transition on the metapopulation level [54, 70], although in this case the low-level networks were treated in a mean field fashion, classified only by rate equations and infection numbers, with no internal network features. In addition, the percolation threshold in interacting networks was found to be lower than in single networks, with a giant cluster appearing for a smaller total number of links [71].

In this work we consider a group of interconnected networks (or, alternately, interconnected communities within a single, larger network). We pose the question: Under what conditions will an epidemic spread only on the sub-networks, with minimal isolated infections on other network components, and under what conditions will it spread across the

entire interconnected network system? Depending on the parameters of the individual networks and their interconnections, connecting one network to another can have a profound or a small effect on the spread of an epidemic. Identifying the conditions in which these cases occur is vital to our understanding and management of epidemic processes.

We define two different interconnected network regimes, strongly and weakly coupled, and find the interaction strength value separating these two regimes. Our primary result is to show that in the strongly-coupled case, we find that all networks are simultaneously either disease free or part of an epidemic, while in the weakly-coupled case a new "mixed" phase can exist. In this mixed phase, the disease is epidemic on only one network, and not in other networks, despite the interconnections. The applications to public health are straightforward. If two neighboring communities comprise a strongly-coupled network system, then an outbreak in any community is cause for immediate concern in the other. Due to this, in the strongly-coupled case it becomes important to pursue a strategy of communication and joint action between public health agencies, and perhaps even intervention from a single agency with higher authority.

## 4.2 Model

In this section, we consider the case of only two interconnected networks of equal size, but it is easily possible to extend the model to an arbitrary number of networks of any size.

We form our interconnected network systems in the following way:

1. Generate two networks, A and B, with their own intranetwork ( $A \leftrightarrow A$  and  $B \leftrightarrow B$ ) degree distributions  $P_A(k)$  and  $P_B(k)$  according to the standard Molloy-Reed configuration model [12].
2. Draw a degree from the internetwork ( $A \leftrightarrow B$ ) degree distribution  $P_{AB}(k)$ , for all the nodes in both networks.
3. If the total degree assigned to nodes in network A is not equal to the total number degree assigned to nodes in network B, randomly reassign a node in B until the total

numbers in each network are equal.

4. Randomly connect nodes in network A to nodes in network B to form the interconnected network.

This method generates random, uncorrelated, interconnected network systems with specified inter- and intra-network degree distributions. While this method works for any arbitrary degree distribution,  $P_x(k)$ , we present results only for random Poissonian degree distributions.

The susceptible-infected-recovered (SIR) epidemic model is used here to study the effects of interconnected network structure on epidemic threshold. The SIR model is well established and describes diseases such as HPV, seasonal influenza, or H1N1 [49, 54]. In this model, each node has three possible states: susceptible ( $s$ ), infected ( $i$ ), or recovered ( $r$ ). Each node begins in state  $s$ , except for a single node in one network chosen in state  $i$ . Nodes in state  $i$  infect their neighbors in state  $s$  with probability  $\beta$  at each time step, changing them to  $i$ . Nodes enter state  $r$  after spending a recovery time  $t_r$  in state  $i$ .

In order to find the threshold for an epidemic, we can think of epidemic spreading as a bond percolation process [19, 72] on a network. In bond percolation, links between nodes are activated with a certain probability  $p$ . If this probability is greater than a certain critical value,  $p_c$ , then a giant cluster emerges, where the existence of a path between any two nodes is almost certain. In a disease-spreading model, nodes infect their neighbors, “activating” the links between them with a certain probability, and a disease reaches nodes through this entire network above a certain critical value,  $\beta_c$ , just as in the case for percolation.

In networks, this critical threshold for percolation if all potential links are activated is  $\kappa = 2$ , where  $\kappa$  is the expected number of nearest neighbors that a node chosen by following an arbitrary link will have, and is calculated from the ratio between the second and the first moments of the degree distribution:  $\kappa = \langle k^2 \rangle / \langle k \rangle$ . For  $\kappa \geq 2$ , a giant cluster exists, while for  $\kappa \leq 2$  only small isolated clusters exist. If some subset of bonds is activated at random with probability  $p$ , a giant cluster appears at a critical value of  $p_c = 1/(\kappa - 1)$  [3].

The SIR model likewise has an epidemic phase transition at a critical  $\beta = \beta_c$  below which the disease remains confined to the local neighborhood of the initial infection, and above which the disease spreads throughout the network. This transition from the disease-free phase to the epidemic phase depends on the average number of secondary infections per infected node becoming larger than one. This allows the long-term survival of the disease, as the infection density will grow over time on average, and thus ensure that the epidemic spreads to a large fraction of the population. In our problem, the expected number of susceptible neighbors that a node has when it just becomes infected is given by  $\kappa - 1$ , since the total expected number of neighbors is  $\kappa$ , and one of them must be excluded as the infected parent from which the current node descended. The transmissibility  $T_\beta = 1 - (1 - \beta)^{t_r}$  is the probability to infect a neighbor before recovery. The mean number of secondary infections per infected node is thus  $N_I = (\kappa - 1)T_\beta$ . The infection will die out if each infected node does not spawn on average at least one replacement so, for a very large network, the critical point is given by the relation  $(\kappa - 1)T_\beta = 1$ . The simple network model exhibits only a single transition at  $\beta_c$  given by [19]

$$\beta_c(\kappa) = 1 - [1 - (\kappa - 1)^{-1}]^{1/t_r}. \quad (4.1)$$

In the interconnected network model, the behavior is more complicated, as the disease can potentially be in epidemic in different combinations of the networks. The disease can either be in the epidemic phase in both networks, in the disease-free phase in both networks, or active in one network while the other remains disease free, called here the mixed phase. The boundaries of these phases are controlled by  $\kappa_A$ ,  $\kappa_B$ , and  $\kappa_T$ , where  $\kappa_A$  and  $\kappa_B$  are calculated over the individual A and B networks, disregarding internetwork connections, and  $\kappa_T$  is calculated over the entire coupled network system, including intra- and inter-network links.

### 4.3 Strongly-Coupled Network Systems

We consider an interconnected network system to be strongly-coupled if  $\kappa_T$  is larger than  $\kappa_A$ , and  $\kappa_B$ . In strongly-coupled network systems, we expect any epidemic to emerge simultaneously on networks A and B. Using Eq. 4.1 for each of the three  $\kappa$ , it can be shown that for the strongly-coupled case,  $\beta_c(\kappa_T)$ , the critical value of  $\beta$  for the disease to emerge on the giant component formed by the entire interconnected network is smaller than both  $\beta_c(\kappa_A)$  and  $\beta_c(\kappa_B)$ , the critical values of  $\beta$  for epidemics to spread on networks A or B ignoring internetwork links. As such, any pathogen virulent enough to spread in network A or B alone will have already caused an epidemic occurring across the interconnected network system. For this case, the disease spreads across the interconnected network system as a single network, with the internetwork connections bringing an epidemic into existence before any intranetwork connections can do so independently; the mixed phase will not be seen. To support this, we plot the ratio of the largest connected infected cluster formed solely from nodes connected with intranetwork links, compared to the size of the largest connected cluster formed by nodes connected with all links, in Fig. 4.2. For a strongly-coupled network system, the relative size of the largest connected infected component contained entirely in a single network decreases initially, showing that the epidemic is occurring across the interconnected network system, not locally in one of the networks. Thus in the strongly-coupled case, epidemic spreading is enhanced due to internetwork connections.

### 4.4 Weakly-Coupled Network Systems

If  $\kappa_A$  or  $\kappa_B$  is larger than  $\kappa_T$ , we define the interconnected network system to be weakly coupled. In interconnected network systems of this sort,  $\beta_c(\kappa_A)$  or  $\beta_c(\kappa_B)$  respectively will be smaller than  $\beta_c(\kappa_T)$ . Without loss of generality, we define network B as the more intranetworked network. We thus have  $\beta_c(\kappa_B) > \beta_c(\kappa_T) > \beta_c(\kappa_A)$ . In the weakly-coupled case, we expect to see a mixed phase, with the boundaries dependent on the values of  $\beta$  and  $\langle k_{AB} \rangle$ . A mixed phase indicates that the addition of the interconnections between the

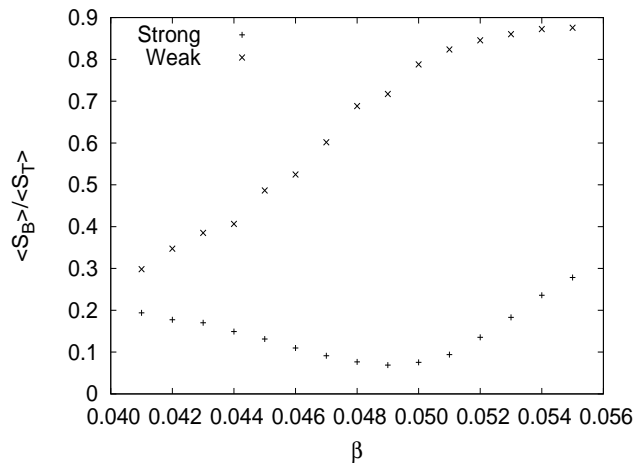


Figure 4.2: Epidemics in strongly-coupled network systems spread across all networks, remaining confined to one network in weakly-coupled network systems. We plot  $S_B$ , the size of the largest connected cluster solely in network B, divided by  $S_T$ , the size of the infected cluster across the interconnected network system, for both strongly- and weakly-coupled network systems. The B only cluster decreases in relative size until criticality ( $\beta_c(\kappa_B) = .048$ ), showing that the epidemic spreads throughout both networks rather than remaining confined in B in the strongly-coupled system. By contrast, in the weakly-coupled system, the relative size of the B only cluster grows until  $\beta = \beta_c(\kappa_T) = .054$ , showing that growth is localized in the more strongly coupled network. For the strongly-coupled network,  $\langle k_A \rangle = 1.5$ ,  $\langle k_B \rangle = 2.5$ , and  $\langle k_{AB} \rangle = 2.5$ . For the weakly-coupled network,  $\langle k_A \rangle = 1.5$ ,  $\langle k_B \rangle = 4.55$ , and  $\langle k_{AB} \rangle = 0.3$ .  $N_A = N_B = 10^4$ ,  $t_r = 5$ .

two networks is only affecting epidemic spreading on the network with weaker intranetwork connections, with the epidemic on the network with stronger intranetwork connections unchanged by the internetwork links.

Epidemic spreading is a non-competitive process. Adding more links to a network can only increase the spread of an epidemic, never decrease it, as the chance of a node infecting its neighbors is constant regardless of degree. Thus, a disease with  $\beta$  above the individual epidemic threshold of network B ( $\beta_c(\kappa_B)$ ) will enter the epidemic phase on that network, regardless of the other network and the values of  $\kappa_T$  and  $\kappa_A$ . If  $\beta$  is below  $\beta_c(\kappa_T)$ , however, the disease cannot spread to more than isolated small clusters of network A. The disease is in its mixed phase, where an epidemic is occurring on one local network, but not throughout the entire interconnected network system. The weakly-coupled case in Fig. 4.2 shows this,



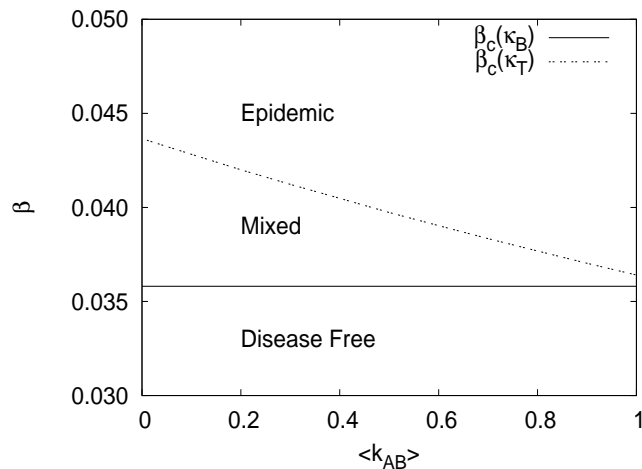


Figure 4.3: Three distinct phases exist in weakly-coupled networks. Sample phase diagram for two networks, with  $\langle k_A \rangle = 1.5$  and  $\langle k_B \rangle = 6.0$  as a function of infection strength  $\beta$  and internetwork degree  $\langle k_{AB} \rangle$ . Below  $\beta_c(\kappa_B)$  no epidemic occurs. For  $\beta_c(\kappa_B) < \beta < \beta_c(\kappa_T)$ , there exists a mixed phase, where a finite fraction of network B becomes infected, but network A has only small infected clusters. Above  $\beta_c(\kappa_T)$ , an epidemic occurs across the entire network.  $N_A = N_B = 10^4$ ,  $t_r = 5$

with the largest connected cluster contained entirely in B becoming larger compared to the size of the giant component with increasing  $\beta$  until  $\beta_c(\kappa_T)$  is reached.

If  $\beta$  is increased to above  $\beta_c(\kappa_T)$ , network B becomes capable of spreading the disease to network A, which now enters the epidemic phase, even for  $\beta < \beta_c(\kappa_A)$ . This matches our own strongly-coupled case, as well as the work done by Leicht and D'Souza [69] where a giant cluster forms consisting of nodes in both networks, even when the less intracoupled network is below its own percolation threshold. We plot a phase diagram describing these three regions for a sample network in Fig. 4.3, showing the disease-free phase, the mixed phase, and the epidemic phase. The existence of this mixed phase is important in the real-world context of interacting networks, as the communities or systems that comprise the components are likely to be governed by different bodies. If two cities, for example, together form a weakly-coupled network system, the more highly connected city can more safely disregard the links to, and response of, the less highly connected city, as the spread of the epidemic will depend on local parameters only.

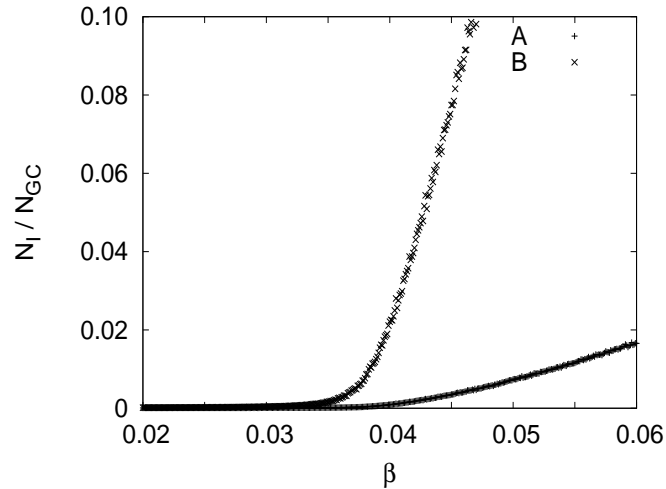


Figure 4.4: In the mixed phase, the two networks have separate transition values. Ratio of total number of infected  $N_I$  in each network to the size of the giant cluster  $N_{GC}$  for two weakly coupled networks with  $\langle k_A \rangle = 1.5$ ,  $\langle k_B \rangle = 6.0$  and  $\langle k_{AB} \rangle = .1$ . The respective epidemic thresholds calculated from Eq. (4.1) are  $\beta_c(\kappa_B) \approx 0.035$ ,  $\beta_c(\kappa_T) \approx 0.0425$ . The infection can be seen to become active in network B well before it does in network A.  $\beta$  for  $\kappa_T$ .  $N_A = N_B = 10^4$ ,  $t_r = 5$

We performed Monte-Carlo simulations to verify this result. First, Fig. 4.4, shows infection densities at different  $\beta$ , corresponding to a vertical sweep across the phase diagram seen in Fig. 4.3 at  $\langle k_{AB} \rangle = 0.1$ . The epidemic spreading first occurs at  $\beta_c(\kappa_B)$ , where the disease enters the epidemic phase and spreads through network B, while the infection density in network A remains negligible. This mixed region, in agreement with our predictions, occurs in the region  $\beta_c(\kappa_B) < \beta < \beta_c(\kappa_T)$ . In this regime, network A plays no role in the spreading of the infection on network B. Above  $\beta_c(\kappa_T)$ , we see that the infection density in network A begins to rise, showing that the entire interconnected network system is now in the epidemic phase, as predicted.

In addition to infection density, the survival probability is often used to identify the critical threshold in epidemics. At criticality, the probability of an infection started from a single infected site remaining active at a later time  $t$  is expected to scale as  $P(t) \sim t^{-1}$ [21]. Fig.4.5, shows the survival probabilities of the networks comprising an interconnected with  $\beta = \beta_c(\kappa_B)$  for both the strongly- and weakly-coupled cases. In both cases, Network B

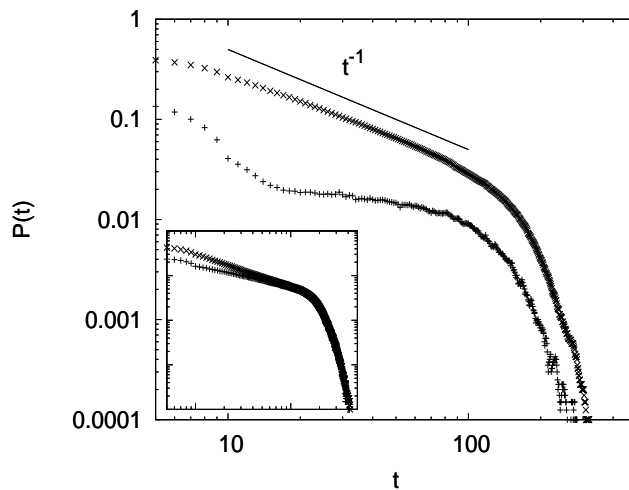


Figure 4.5: Epidemics exhibit critical survival only in one network for weakly-coupled networks. Infection survival probabilities  $P(t)$  on the individual networks with internetwork connectivity  $\langle k_{AB} \rangle = .1$  at  $\beta = \beta_c(\kappa_B)$ . The survival probability in network A (lower curve, with +), the less connected network, is a small fraction of that in network B (upper curve with x), the more connected network. The survival probability in network B falls off as  $t^{-1}$ , as expected of a system at criticality. Network A does not show a smooth decrease towards 0 typical of a network much below criticality. The inset shows the survival probabilities for each network with  $\langle k_{AB} \rangle = 1$ , in the strongly-coupled regime, showing the convergence of the two survival probabilities. For both networks we have  $\langle k_A \rangle = 1.5$ ,  $\langle k_B \rangle = 6.0$ ,  $N_A = N_B = 10^4$ , and  $t_r = 5$ ,

exhibits the expected  $t^{-1}$  fall-off in survivability with time that is expected of a system at criticality, but in the weakly-coupled case, the survival probability in network A does not fall off as expected, due to infrequent and non-epidemic instances of infections from network B. The slope of the survival probability for network A thus cannot be used directly to confirm when it enters epidemic, but the relative difference in survival probabilities can be used instead, as, if both networks are participating in an epidemic, the disease should be active in both networks at each time step. We thus introduce the survival probability gap,  $\Delta P(t) = P_B(t) - P_A(t)$ , to examine the approach to criticality in Network A.

In Fig. 4.6 we plot this survival gap at different  $\beta$ , equivalent to horizontal slices across the phase diagram seen in Fig. 4.3. We see that when  $\langle k_{AB} \rangle$  or  $\beta$  is increased and moves outside the expected mixed phase region, the difference in survival probabilities between

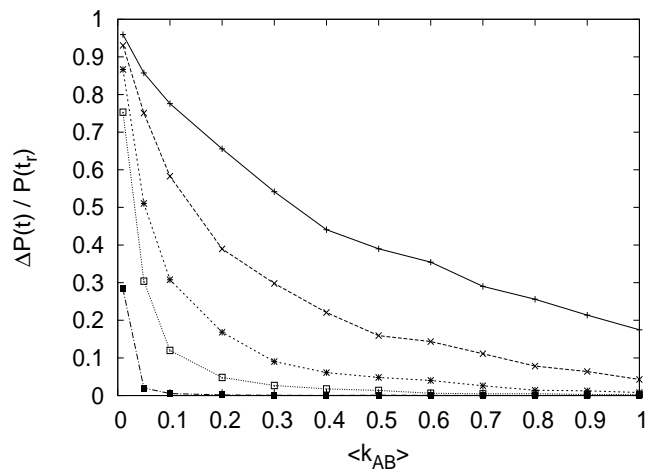


Figure 4.6: Survival probability gap shows mixed phase boundaries. Fractional size of the minimum survival probability gap,  $\Delta P(t)/P(t_r)$ , (minimum distance between the two curves in Fig. 4.5 after the time  $t_r$  has passed) between two interacting ER networks with  $\langle k_A \rangle = 1.5$  and  $\langle k_B \rangle = 6.0$  at various infection strengths. From top to bottom  $\beta = \beta_c(\kappa_B) = .0358, .038, .04, .042, \beta_c(\kappa_T) = .044$ . The gap shrinks with increasing interaction and with increasing  $\beta$ , matching Fig. 4.3. At  $\beta_c(\kappa_T)$ , the survival probability is the same in both networks for all but  $\langle k_{ab} \rangle = .01$ , where it remains distinct due to finite size effects.  $N_A = N_B = 10^4$ ,  $t_r = 5$ .

the two networks vanishes. This confirms the assertion that there can only be a gap in survival probability when one network is in the epidemic phase and the other is not, i.e. in the mixed phase. Thus a non-zero survival probability gap can serve as a good predictor for the presence of the mixed phase.

Lastly, we addressed the question of universality under different values of inter- and intra-network degree, finding that along the disease-free/mixed phase transition line, the behavior of networks with different  $\kappa$  is universal under appropriate scaling. Fig.4.7 shows three different networks, all with  $\langle k_B \rangle > \langle k_A \rangle$  and  $\beta = \beta_c(\kappa_B)$ . Rescaling the survival probabilities by  $P(t_r)$  and plotting  $\kappa_T/\kappa_B$  instead of  $\langle k_{AB} \rangle$  directly, the curves collapse, showing identical approaches toward survival probability homogeneity, and thus identical mixed phase disappearance. For small  $\delta_\kappa = \kappa_T - \kappa_B$ , the gap grows identically, implying survival probabilities in networks can be used as a measure of network connectivities near criticality, as the latter may be difficult to obtain for social and biological networks.

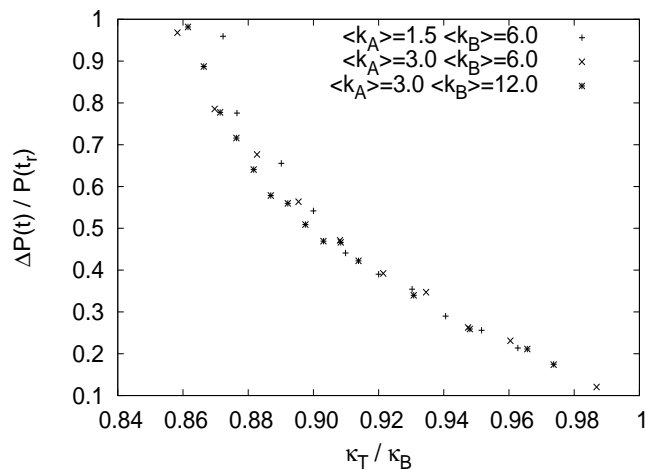


Figure 4.7: Survival probability gap scaling is universal when rescaled by  $\kappa_B$  for different networks. The data collapses onto a single curve, showing this universal behavior. The ratio of  $\kappa_T$  to  $\kappa_B$  determines  $\delta P(t)$  along the disease-free / mixed phase border. For the network systems with  $\langle k_B \rangle = 6.0, \beta = .0358$ , for  $\langle k_B \rangle = 12.0, \beta = .01725$ . Both are  $\beta_c(\kappa_B)$  of the respective systems.  $N_A = N_B = 10^4, t_r = 5$ .

## 4.5 Conclusions

In summary, we introduced two classes for interconnected network systems, strongly coupled and weakly coupled, and studied the behavior of epidemics on them. In strongly-coupled network systems, epidemics occur always across the entire interacting network system, with the presence of interconnections enhancing epidemic spreading. In weakly-coupled network systems, a mixed phase exists where epidemics do not always occur across the full interconnected network system, and interconnections affect only epidemic spreading across the less intracoupled network. We demonstrated the boundaries and behavior of the mixed phase numerically as well as analytically. Proper analysis of which groups of communities comprise strongly- or weakly-coupled systems could inform public policy and highlight the necessity of cooperation between different governing bodies.

## 4.6 Appendix: Interacting Square Lattices.

It is interesting to consider the special case of two interacting square lattices. Identical networks are always strongly-coupled, so no mixed phase exists. The square lattice still exhibits an epidemic phase transition, however. The critical value for the percolation transition in the square lattice is  $p_c = 0.5$  (see Sec. 1.3). From  $p_c = T$  we can derive the critical virulence in a square grid as  $\beta_c^g = 1 - \sqrt[5]{.5}$ , which is  $\beta_c^g \approx .129$  for  $t_r = 5$ . The square grid has  $\kappa = 4$  and so from Eq. 4.1 we would expect  $\beta_c(\kappa = 4) = 1 - \sqrt[5]{1 - 1/3}$ , which is  $\beta_c(4) \approx .078$  for  $t_r = 5$ . The ratio of these two values is  $\beta_c^g/\beta_c(\kappa) \approx 1.6$ . We conducted simulations of two interacting square lattices, in which any node in lattice A could connect to any node in Lattice B. Fig. 4.8 shows the ratio of  $\beta_c$  determined from the simulations to the value predicted by Eq. 4.1 at different interaction strengths. We can see that at very low interaction strengths ( $k_{ab} < .001$ ) we see the behavior of a pure grid; the interconnections are too rare to impact epidemic spreading. As the interaction strength is increased, the system behaves more and more like a random network. Above  $k_{ab} = 1$ , the epidemic threshold measured agrees well with Eq. 4.1, regardless of network size.

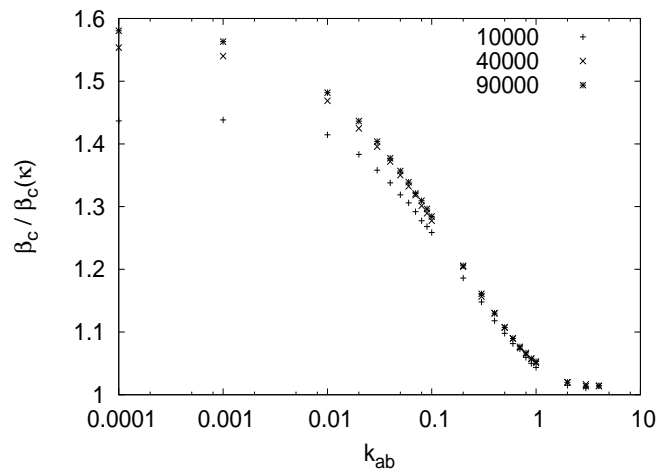


Figure 4.8: Interacting square lattices behave as ER for strong enough interactions. We plot the ratio of  $\beta_c$  measured from the critical point in simulations to  $\beta_c$  calculated from Eq. 4.1 as a function of  $\langle k_{AB} \rangle$  for three different network sizes. The critical value smoothly approaches the heterogeneous network value over a range of  $k_a b$  from  $\approx .01 - 1$  regardless of system size.  $t_r = 5$ .

# Chapter 5

## Summary and Future Work

We end with a brief overview of topics for future research building on the work done in this thesis.

1. **Growth of Interacting Networks:** In Section 4.2 we introduced a method based on the configuration model for generating interacting network systems given a fixed number of nodes and links. It will be interesting to pose the question of how to use a network growth model (such as the Barabási-Albert model) to generate these networks following stochastic attachment methods. The question of how the network parameters will depend on the attachment methods is also of interest.
2. **Quarantine on Interacting Networks:** Combining the topics of Chapter 3 and Chapter 4, it would be of interest to look at the effects of internetwork vs intranetwork quarantine efforts. The final result in Section 3.3 showed that quarantine preferentially targeted at high-degree nodes was more effective than the same average level of quarantine applied uniformly, due to the high-degree nodes functioning as super-spreaders. It is likely that quarantine targeted at internetwork links will be more effective at stopping epidemics in strongly-coupled network systems than a uniform quarantine, and that it would be more effective at preventing the spread of an epidemic to the coupled network, expanding the mixed phase, in weakly-coupled network systems.



# Bibliography

- [1] National Research Council, *Network Science*, (National Academies Press, Washington, 2005.)
- [2] B. Roehr, *BMJ* 2010; 341:c4627.
- [3] R. Cohen et al., *Physical Review Letters* **85**, 4626 (2000).
- [4] P. Erdős and A. Rényi, *Publicationes Mathematicae Debrecen*, **6**, 290 (1959).
- [5] B. R. Gadjiev and T. B. Progulova, arXiv:0805.3706.
- [6] A.-L. Barabási and R. Albert, *Science* **286**(5439), 509 (1999).
- [7] H. Jeong et al., *Nature* **407**, 651 (2000).
- [8] F. Liljeros et al., *Nature* **411**, 907 (2001).
- [9] R. F. i Cancho and R. V. Solè *Proceedings of the Royal Society B* **268**, 2261 (2001).
- [10] L. A. N. Amaral et al., *Proceedings of the National Academy of Sciences USA* **97**, 11149 (2000).
- [11] D. J. Watts and S. H. Srogatz, *Nature (London)* **393**, 440 (1998).
- [12] M. Molloy and B. Reed, *Random Structures and Algorithms* **6** 161 (1995); *Combinatorics, Probability and Computing* **7**, 295 (1998).
- [13] H. Kesten, *Percolation Theory for Mathematics*, (Birkhäuser, Boston, 1982).

- [14] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Physical Review E* **64**, 026118 (2001).
- [15] M. O. Ball, *Networks* **10**, 153 (1980).
- [16] N. Metropolis and S. Ulam, *Journal of the American Statistical Association* **44**(247), 335 (1949).
- [17] L. A. Braunstein, et. al., *International Journal of Bifurcation and Chaos* **17**(7), 2215 (2007).
- [18] H. W. Hethcote, *Mathematical Biosciences* **28**, 335 (1976).
- [19] M. E. J. Newman, *Physical Review E* **66**, 016128 (2002).
- [20] A. Fronczak, P. Fronczak and J. A. Holyst, *American Institute of Physics Conference Proceedings* **776**, 52 (2005).
- [21] H. Hinrichsen, *Advances in Physics* **49**, 815 (2000).
- [22] J. F. F. Mendes, et al., *Journal of Physics A* **27**, 3019 (1994).
- [23] R. Albert and A.-L. Barabási, *Reviews of Modern Physics* **74**, 47 (2002).
- [24] R. Albert et al., *Nature* **406**, 378 (2000).
- [25] K. I. Goh et al., *Physical Review Letters* **87**, 278701 (2001).
- [26] D.S. Calaway et al., *Physical Review Letters* **85**, 4626 (2000).
- [27] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, 2004).
- [28] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
- [29] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge, 1994).

- [30] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function*, (Cambridge Univ. Press, 2010.)
- [31] R. Pastor-Satorras and A. Vespignani Physical Review Letters **86**, 3200 (2001).
- [32] A. Barrat et al., Proceedings of the National Academy of Sciences USA **101**, 3747 (2004).
- [33] E. Volz and L. A. Meyers, Proceedings of the Royal Society B **274**, 2925 (2007); Journal of the Royal Society Interface **6**, 233 (2009).
- [34] D. Kempe, J. Kleinberg, and A. Kumar, Proceedings of the thirty-second annual ACM symposium on Theory of computing (2000); Proceedings of the thirty-third annual ACM symposium on Theory of computing (2001).
- [35] J. S. Coleman, *Introduction to Mathematical Sociology* (Free Press, New York, 1964).
- [36] S. Wasserman, Journal of the American Statistical Association **75**, 280 (1980).
- [37] T. A. B. Snijders, Journal of Mathematical Sociology **21**, 149 (1996).
- [38] T. A. B. Snijders, in *Sociological Methodology*, edited by M. E. Sobel and M. P. Becker (Basil Blackwell, Boston, 2001), pp 361-395.
- [39] P. Erdős and A. Rényi, Publicationes Mathematicae Debrecen, **6**, 290 (1959); Publications of the Mathematical Institute of the Hungarian Academy of Sciences **5**, 1760 (1960).
- [40] W. Kinzel, in *Percolation Structures and Processes*, edited by G. Deutscher, R. Zallen, and J. Adler (Hilger, Bristol, 1983).
- [41] L. A. Braunstein et al., Physical Review Letters **91**, 168701 (2003).
- [42] The directed network representation of a dynamic network as shown in Fig. 2.2(a) includes  $tN$  nodes where  $t_x \sim N^{1/2}$  according to Eq. (2.2).

- [43] A. Vespignani and G. Caldarelli, *Large Scale Structure and Dynamics of Complex Networks*, (World Scientific, Singapore, 2007).
- [44] T. Gross and H. Sayama, *Adaptive Networks: Theory, Models and Applications*, (Springer, New York, 2009).
- [45] I. B. Schwartz and L. B. Shaw, *Physics* **3**, 17 (2010).
- [46] T. Gross and B. Blasius, *Journal of the Royal Society Interface* **5**, 259 (2008).
- [47] T. Gross, Carlos. J. Dommar D’Lima, and B. Blasius, *Physical Review Letters* **96**, 208701 (2006).
- [48] F. Vazquez, V. Eguíluz and M. San Miguel, *Physical Review Letters* **100**, 108702 (2008).
- [49] R.M. Anderson and R.M. May, *Infectious Disease of Humans*, (Oxford University Press, Oxford, 1992).
- [50] K. Eastwood, D. N. Durrheim, M. Butler, and A. Jon, *Emerging Infectious Diseases* **16**(8), 1211 (2010).
- [51] L. B. Shaw and I. B. Schwartz, *Physical Review E* **77**, 066101 (2008).
- [52] D. H. Zanette and S. Risau-Gusman, *Journal of Biological Physics* **34**, 135 (2008).
- [53] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen, *Proceedings of the National Academy of Sciences USA* **106**, 6872 (2009).
- [54] V. Colizza, A. Barrat, M. Barthlemy, and A. Vespignani, *Proceedings of the National Academy of Sciences USA* **103** 2015 (2006).
- [55] Z. Wu, C. Lagorio, L. A. Braunstein, R. Cohen, S. Havlin, and H. E. Stanley., *Physical Review E* **75**, 066110 (2007).
- [56] W. Aiello, F. Chung, and L. Lu, *Experimental Mathematics* **10**(1), 53 (2001).

- [57] R. Albert, H. Jeong, and A.-L. Barabási, *Nature* **401**, 130 (1999).
- [58] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Physical Review Letters* **86**, 3682 (2001).
- [59] M. Kuperman and G. Abramson, *Physical Review Letters* **86**, 2909 (2001).
- [60] R. Cohen, S. Havlin, and D. ben-Avraham, *Physical Review Letters* **91**, 247901 (2003).
- [61] R. Parshani, S. Carmi, and S. Havlin, *Physical Review Letters* **104**, 258701 (2010).
- [62] C. Lagorio et al., *Physical Review E* **83**, 026102 (2011).
- [63] M. Kurant and P. Thiran, *Physical Review Letters* **96**, 138701 (2006).
- [64] B. Bollobàs, S. Janson, and O. Riordan, *Random Structures and Algorithms* **31**, 3-122 (2007).
- [65] A. Allard, P-A Noël, L. J. Dubè and B. Pourbohloul, *Physical Review E* **79**, (3) 036113 (2009).
- [66] S. N. Dorogovtsev, J. F. F. Mendes, A. N. Samukhin, and A. Y. Zyuzin, *Physical Review E* **78**, 056106 (2008).
- [67] E. Barreto, B. Hunt, E. Ott, and P. So, *Physical Review E* **77**, 036107 (2008).
- [68] S. Buldyrev et al., *Nature* **464**, 7291 (2010).
- [69] C. D. Brummitt, R. M. D'Souza, E. A. Leicht, arXiv:1106.4499.
- [70] V. Colizza and A. Vespignani, *Physical Review Letters* **99**, 148701 (2007).
- [71] E. A. Leicht, R. M. D'Souza, arXiv:0907.0894.
- [72] R. Cohen, S. Havlin, *Complex Networks: Structure, Robustness and Function*, (Cambridge Univ. Press, 2010.)

# Curriculum Vitae

## Mark E. Dickison

Boston University, Physics Department  
Center For Polymer Studies  
590 Commonwealth Avenue  
Boston, Massachusetts 02215 USA

Telephone: 617-353-3891  
Facsimile: 617-353-3783  
E-mail: dickison@bu.edu

### EDUCATION

- 2012 (Expected in May), Ph.D. Physics, Boston University, Boston, USA  
Advisor: H. Eugene Stanley  
Dissertation: *Dynamic and Interacting Complex Networks*
- 2005, M.Sc. Physics, University of Missouri at Rolla, Rolla, USA  
Advisor: Thomas Vojta  
Thesis: *Monte-Carlo Simulations of Disordered Non-Equilibrium Phase Transitions*

### EMPLOYMENT AND TEACHING EXPERIENCE

- 2008–present, Boston University, *Research Assistant*
  - Mapped complex network concepts onto problems from statistical mechanics.
  - Used network methods to predict outcomes of epidemics taking into consideration media coverage.
  - Contributed to research on development and prevention of extremist beliefs in opinion modeling.
  - Studied implicit and explicit social networks generated via original data collected from internet sources.
- 2009, Chinese Culture Connection, *Outreach Instructor*

- Prepared engaging lessons to present and popularize science concepts to the general public.
- Coordinated parallel activities with other instructors.
- 2005–2007, Boston University, *Teaching Assistant*
  - Led discussion sections and labs for undergraduate physics courses.
  - Developed presentation and instructional skills.
  - Acquired familiarity with Atomic Force Microscopy techniques under Prof. Ophelia Tsui.
- 2004–2005, University of Missouri at Rolla, *Research Assistant*
  - Built and maintained a large scale parallel computing environment.
  - Designed and programmed disorder incorporating models for simulation.

## AWARDS

- NSF International Travel Award to participate in the IUPAP Triennial Conference on Statistical Physics, Cairns, Australia, July 2010

## SKILLS

- Familiar with a range of operating systems: Windows, Unix/Linux
- Experience with large scale parallel computing environments
- Proficient with Fortran, C/C++, shell scripting, Python, L<sup>A</sup>T<sub>E</sub>X.
- Well versed in technical writing.
- Experienced with modeling real-world systems via statistical methods.
- Capable at extracting useful information from divergent data sources, such as web records.

## RESEARCH PAPERS

1. M. Dickison and H. E. Stanley, “Are your friends who you think they are? Implicit vs. Explicit Social Networks” (In Progress).
2. B. Podobnik, M. Dickison, D. Horvatic, and H. E. Stanley, “Preferential Attachment in the Interaction between Dynamically Generated Interdependent Networks”, Submitted to Phys. Rev. Lett. (2012).

3. M. Dickison, S. Havlin, and H. E. Stanley, “Epidemics on Interconnected Networks”, arXiv:1201.6339 (2012).
4. C. Lagorio, M. Dickison, F. Vazquez, L. A. Braunstein, P. A. Macri, M. V. Migueles, S. Havlin, and H. E. Stanley, “Quarantine-generated phase transition in epidemic spreading”, *Phys. Rev. E* **83**, 026102 (2011).
5. R. Parshani, M. Dickison, R. Cohen, H. E. Stanley and S. Havlin, “Dynamic networks and directed percolation”, *EPL* **90**,38004 (2010).
6. M. Dickison, T. Vojta, “Monte Carlo simulations of the smeared phase transition in a contact process with extended defects ”, *J. Phys. A: Math. Gen.* **38**, 1199 (2005).
7. T. Vojta, M. Dickison, “Critical behavior and Griffiths effects in the disordered contact process”, *Phys. Rev. E* **72**, 036126 (2005).

#### CONFERENCES AND PRESENTATIONS

1. American Physical Society (APS 2011) 2011 Annual Meeting, Dallas TX, USA. “Quarantine Generated Phase Transition in Epidemic Spreading” 03/2011.
2. HES70: Horizons in Emergence and Scaling, Boston MA, USA. “Quarantine Generated Phase Transition in Epidemic Spreading” 03/2011
3. International Conference on Statistical Physics of the International Union for Pure and Applied Physics (IUPAP) 2010, Cairns, Australia, “Quarantine on Dynamic Networks” 07/2010
4. The International School and Conference on Network Science (NetSci 2009) 2009, Venice, Italy. “Dynamic networks and directed percolation” 07/2009
5. American Physical Society (APS 2005) 2005 Annual Meeting, Los Angeles CA, USA. “Rare region effects at a non-equilibrium phase transition” 03/2005