DISSERTATION

CHARACTERISTICS OF CERTAIN FAMILIES OF RANDOM GRAPHS

Submitted by

Christian Paul Hampson

Department of Mathematics

In partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2009

COLORADO STATE UNIVERSITY

August 17, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY CHRISTIAN PAUL HAMPSON ENTITLED "CHARACTERISTICS OF CERTAIN FAMILIES OF RANDOM GRAPHS" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Dr. Anton Betten

Dr. Daniel Rudolph

Dr. Dave Theobald

Adviser: Dr. Jeff Achter

Department Head: Dr. Simon Tavener

ABSTRACT OF DISSERTATION

CHARACTERISTICS OF CERTAIN FAMILIES OF RANDOM GRAPHS

Many random network models can be expressed as the product space of the probability space of the individual edges. In these cases, the model can be expressed using a matrix of the probabilities of each edge. I then analyze these models using their respective probability matrices. Degree distribution and the larger eigenvalues are among the attributes whose values can be bound by examining the same attributes of the probability matrix. I also bound the difference between the eigenvalues of the adjacency matrix of a member of a random graph model and the eigenvalues of the probability matrix for the model. In addition I find the neighborhood expansion properties for three separate edge-product models.

Christian Paul Hampson
Department of Mathematics
Colorado State University
Fort Collins, Colorado 80523
Fall 2009

ACKNOWLEDGEMENTS

There are many people without whom this dissertation would never have occurred and whom I thank for all they have done for me.

The first of these is Jeff Achter, my advisor. Without your guidance, I would not have been able to write this paper. Your insistence on clarity and explicit statements is greatly appreciated. You taught me much about research and writing. I am proud to be your student.

I also thank the members of my committee; Anton Betten, Dan Rudolph, and Dave Theobald; for their comments and suggestions through the process of research and writing.

To the staff and faculty in the Mathematics department: thank you for your teaching and for shepherding me through the process of graduate school. I have learned much here. To my fellow students: thank you for listening to me and providing ideas to pursue and different ways of looking at my material.

I thank the members of the Program in Interdisciplinary Math, Ecology, and Statistics for helping me to realize how others can apply the mathematics that I am doing and for providing the chances I have had to work with real world problems. In particular, gratitude is due to the members of the Webb Lab for shaping the way that I have viewed applications of networks.

Lastly, and most importantly, I owe the last six years to my wife, Malia, and to our children Jake and Caralynn. For the past six years they have stayed with me as I went back and forth across the country. They have had a husband and father who was thoroughly wrapped up in classes and research. You have stood by me the whole time and I am extremely grateful.

Funding for the author was provided by NSF-IGERT Grant DGE-#0221595, administered by the PRIMES program at Colorado State University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

This dissertation begins with an exploration of general applications of graph theory. This includes some history as well as current applications and some attributes of interest. The second chapter provides an overview of random graphs. The third chapter focuses specifically on random graph models in which edges are determined independently of each other. It is in chapter three that I extend some of what is known about the Erdös-Rényi model to other random graph models. Chapters four and five deal with specific random graph models.

1.1 Applications of Graphs

The study of graphs began with an application to the real world. In 1736, L. Euler published a paper demonstrating that it was impossible to walk through the city of Königsberg, crossing each of the city's seven bridges exactly once. Since then, there have been many studies of graph theory that pertain to applications. Outside of the mathematical world, however, graphs are usually known as networks. In this thesis, the terms network and graph are used interchangeably, with preference being given to the language in the appropriate reference document.

One of the best-known examples of a real-world network study is S. Milgram's Small World experiment [32][44]. In this experiment, a set of people were given a letter to somehow deliver to a target, about whom several items of information were provided. Information was gathered about each participant along the way. The goal of this experiment was to gather information about existing social networks. In particular, Milgram found that in the cases where the letter was successfully delivered to its target, there was an average of six links in the chain. While the social network itself was never calculated, these papers did infer information about the network.

M. Newman, in his review paper, "The Structure and Function of Complex Networks" [37], notes several applications of graph theory. In particular he notes that social networks have have been studied since the 1930s. In a social network, the vertices typically represent people and the edges represent some sort of connection. This connection can require marriage or parentage, as in the case of a family tree, or it may require only simple contact, as in a network that is used to predict the spread of an easily transmitted disease.

Since there is ambiguity regarding the nature of connection, there are issues with accuracy in reporting real-world social networks. As D. Watts noted in his book *Small Worlds*, people do not always agree on whether they are friends [47]. Collaboration and communication networks have thus been studied as surrogates, since it is possible to obtain reliable data regarding such connections as coauthorship [34][35][36], appearances in the same movie [39][49], or even telephone calls made in a single day [1]. Once we have a defined social network, we can then simulate the progress of something that can be communicated such as a disease or an idea. Perhaps it is

even possible to find characteristics of such a network that would allow us to predict the outcome of some process.

Similar to social networks are technological networks. Studies here have included power grids, roads, airline routes, and, of course, the physical structure of the Internet. Transportation networks have vertices which represent locations and edges which represent means of getting to those locations, such as roads, railroads, scheduled flights, or shipping lanes. In the case of a computer network, the vertices represent computers, routers, hubs, and switches, while the edges most often represent physical wires or wireless connections.

A third category given by Newman is information networks. These are directed networks in which an edge is placed from vertex v_i to vertex v_j if the paper represented by v_i refers to the paper represented by v_j. Examples of information networks include citation networks and the logical structure of the World Wide Web [3]. A similar network in an ecological setting can be found in the flow of water through a watershed. Further examples of applications are given in laypeople's terms in the trade books by D. Watts [48][47] and A-L. Barabási [9].

As real-world networks have been examined, many of them have been found to possess a power-law distribution for the degrees of the vertices. R. Albert and A-L. Barabási note several of these in their review paper [2]. In a graph with a power-law distribution (sometimes called a scale-free graph) with parameters α and γ, the number of vertices with degree x is approximately $e^{\alpha}x^{-\gamma}$. In particular, the Internet, the World Wide Web, the movie actor network, the collaboration network in mathematics and neuroscience, and the citation network all were found to have power-law

distributions with $2 < \gamma < 3$. This conclusion is not unanimous as S. Bansal *et al.* argued for an exponential distribution in several social networks [8].

1.2 Interesting Attributes

There are several definitions and useful facts that will be used throughout this paper. This first set of definitions is available in almost any graph theory textbook, but these formulations are taken from [17].

Definition 1.2.1. *A graph G is a set of vertices, denoted by $V(G)$, together with a set of edges, $E(G)$, such that each edge is an unordered pair of vertices. An edge between two vertices v_i and v_j is denoted by $\{v_i, v_j\}$.*

Definition 1.2.2. *A simple graph is a graph such that no vertex is adjacent to itself and there exists at most one edge between any two vertices.*

Definition 1.2.3. *For any two vertices v_i and v_j, v_i is adjacent to v_j if $\{v_i, v_j\} \in E(G)$.*

Definition 1.2.4. *The adjacency matrix of a graph or a digraph is a matrix A, in which the entry a_{ij} denotes the number of edges from vertex i to vertex j. In an undirected graph, the adjacency matrix is symmetric, so $a_{ij} = a_{ji}$. In a simple graph, each a_{ij} is either 0 or 1.*

Definition 1.2.5. *A path from vertex u to v of length t is an ordered sequence of distinct vertices $u = v_0, v_1, \ldots, v_t = v$ such that for $0 \leq i \leq t$, $\{v_i, v_{i+1}\} \in E(G)$.*

Definition 1.2.6. *For any two vertices $v_i, v_j \in V(G)$, the distance between v_i and v_j, denoted by $d(v_i, v_j)$ is the shortest path length between v_i and v_j.*

Definition 1.2.7. *A graph is connected if for any two vertices v_i and v_j there is a path from v_i to v_j.*

Definition 1.2.8. *In a connected graph G, the diameter of G is* $\text{diam}(G) = \max_{v_i,v_j \in V(G)} d(v_i, v_j)$. *If G is not connected,* $\text{diam}(G) = \max_{K_i \in K(G)} \text{diam}(K_i)$, *where $K(G)$ is the set of all connected components of G.*

Definition 1.2.9. *In a graph G, the average distance of G is the average over distances between pairs of vertices. If a graph has multiple components, the average distance of G is the average over all finite distances between pairs of vertices.*

Definition 1.2.10. *The degree of a vertex $v_j \in V(G)$, is given by* $d_j = |\{e \in E(G) \text{ such that } e = \{v_j, v_i\} \text{ for some } v_i \in V(G)\}|$.

Definition 1.2.11. *The volume of a subset of $S \subseteq V(G)$ is defined to be* $\text{vol}(S) = \sum_{v_j \in S} d_j$.

Definition 1.2.12. *The kth order volume of $S \subseteq V(G)$ is defined to be* $\text{vol}_k(S) = \sum_{v_j \in S} d_j^k$.

Definition 1.2.13. *The kth order volume of a graph G is* $\text{vol}_k(G) = \text{vol}_k(V(G))$.

Definition 1.2.14. *The average degree of a graph G is* $d(G) = \frac{\text{vol}(G)}{|V(G)|}$.

Definition 1.2.15. *The second-order average degree of a graph G is* $\tilde{d}(G) = \frac{\text{vol}_2(G)}{\text{vol}(G)}$.

A variety of network attributes have been studied over the past several years. Some of these include connectivity, the size of components, distance between vertices, degrees of vertices, clustering, degree correlation, and

spectrum. While this information is in itself interesting, when it is gathered from a real-world network we can learn how different processes may affect this network. If a disease hits a social network, these attributes will affect how the disease progresses. Likewise, if our network represents a patchy habitat, these attributes will determine where an animal can go, along with how quickly it can get there. Other processes that have been explored are removal processes, in which vertices or edges are removed from the graph to see how this changes the attributes of the graph.

Is a graph connected? If a path exists from each vertex to every other vertex, the graph is connected. A process on a disconnected graph cannot pass from one component to another, a fact recognized by the application of quarantines to prevent disease spread. Likewise, a graph is k-vertex-connected if the graph cannot be disconnected by the removal of any k vertices. The level of vertex connectivity is often used as a measure of robustness for a network. Likewise a graph is called k-edge-connected if it cannot be disconnected by the removal of any k edges. The level of edge-connectivity would likely be of great interest when examining a transportation network. If a graph is not connected, the immediate question is how large is the largest component? This information tells us the largest number of nodes that any process starting on a single component can affect.

What is the degree distribution of a graph? The answer to this question can inform us how fast a process can take place. If all vertices are of low degree, it will take several steps for a process that moves by single steps to reach all vertices. For example, if the highest degree in a graph on n vertices is 3, a process will take at least $\lceil \log_3(n) \rceil$ time steps to reach every vertex. If, on the other hand, there are a few vertices of large degree

(noted in disease modeling papers as superspreaders), then the process can drastically accelerate [8]. Another piece of information that is helpful for analysis is the distribution of the second-order degree of each vertex, i.e. the number of paths of length two which start at each vertex.

The question of transitivity or clustering addresses the likelihood that two vertices are adjacent if each of them is adjacent to a third vertex. This notion was first promoted by D. Watts and S. Strogatz [49], who observed that many real-world networks appear to have a higher level of transitivity than the random graphs which had been developed at the time. In short the transitivity, or clustering coefficient, of a graph G, $C(G)$, is given by $C(G) = \frac{3T(G)}{A(G)}$, where $T(G)$ represents the number of triangles in G and $A(G)$ represents the number of adjacent pairs of edges in G. When the degree sequence for G is given by $\{d_1, d_2, \ldots, d_n\}$, the denominator $A(G)$ is simply $\frac{1}{2}\sum_{i=1}^{n}\binom{d_i}{2}$. The clustering coefficient, however, gives a misleading characterization of a graph in which triangles are rare but cycles of other low degrees are common. One example of this type of network is the (mostly bipartite) sexual contact network. If we were to attempt to predict the spread of a sexually transmitted disease over a network assuming that the near-zero clustering coefficient implies a random structure, we would likely fail.

1.3 Graph Spectra

Broad bounds exist in the literature regarding some of the eigenvalues and their relationships to other characteristics of the graph. Each graph has five matrices which are recognized as being of particular concern. Recall that the adjacency matrix, A, is formed by letting $a_{ij} = 1$ if there is an edge

from vertex v_i to vertex v_j. The degree matrix, D, is the diagonal matrix formed by letting d_{ii} be the degree of vertex v_i. The combinatorial Laplacian is given by $L = D - A$. The transition matrix is given by $M = D^{-1}A$. The normalized Laplacian is given by

$$\mathcal{L} = D^{-1/2}LD^{-1/2} = I - D^{-1/2}AD^{-1/2}.$$

If $d_i = 0$ for some i, then the (i,i)th value of D^{-1} is defined to be 0. This, of course, occurs only when there is an isolated vertex.

For a nondirected graph A, D, L, and $\mathcal{L}$ are symmetric and real, resulting in all of the eigenvalues being real. The matrix M is similar to the matrix $D^{-1/2}AD^{-1/2}$ which is real and symmetric. Thus the eigenvalues of M are also real.

For the purposes of this paper, $\lambda_i(T)$ is the ith largest eigenvalue of a matrix T and $\mu_i(T)$ is the ith smallest eigenvalue of T. Since the eigenvalues of these matrices are real, the ordering of eigenvalues is done according to real values. Thus $\mu_1(A) \leq mu_2(A)$, while $|\mu_1(A)| \geq |\mu_2(A)|$. Since $\mu_i(\mathcal{L}) = 1 - \lambda_i(M)$, it is only necessary to determine the eigenvalues of one of these two matrices.

In some ways, the largest eigenvalue of A determines how a process can spread. From [17], the spectral radius of a graph G, $\lambda_1(A)$, is related to the number of walks of length k on G. As k becomes very large, the ratio of walks of length k to walks of length $k-1$ is $\lambda_1(A)$. If all walks of length k have equal probability, the entries in the unit principal eigenvector indicate what proportion of a walk of length k will be on each vertex for large k. Also $\sqrt{(1/n)\sum_{i=1}^{n} d_i^2} \leq \lambda_1(A) \leq d_{\max}$ [29][19]. If a connected graph G has m edges and n vertices, $\lambda_1(A) \leq \sqrt{2m - n + 1}$ [50][19].

M. Draief, A. Ganesh, and L. Massoulié found in their 2008 article that $\lambda_1(A)$ gives specific thresholds for virus spread on networks using an SIR model with each infection attempt having probability of success β:

Theorem 1.3.1. *[21, Theorem 2.1]If $\beta\lambda_1(A) < 1$, then the total number of nodes removed in an SIR model, $|Y(\infty)|$, satisfies for any $\epsilon > 0$, and for some constant $C(\epsilon)$, $\Pr[|Y(\infty)| > \sqrt{|X(0)|}n^{1/2+\epsilon}] \leq Cn^{-\epsilon}$, where $|X(0)|$ is the initial number of infectives.*

Thus $\lambda_1(A)$ determines whether the infection will spread to some fraction of the entire network or be limited to a number of nodes.

The largest eigenvalue of M is 1. The eigenvector of M corresponding to the eigenvalue 1 gives the proportion of time spent on each vertex in a random walk [30]. The second-largest eigenvalue of M gives the mixing rate of random walks on G, indicating how long it will take to reach the stable distribution [30]. Also, from [18], $\mu_2(\mathcal{L}) \geq \frac{1}{\text{diam}(G)\text{vol}(G)}$.

Another result from F. Chung relates the size of the neighborhood of a set $S \subseteq V(G)$ to the size of S and $\mu_2(\mathcal{L})$. Let $\Gamma(S) = \{v_j : \exists v_i \in S, \{v_i, v_j\} \in E(G)\}$. Let $S^C = \{v_i \in V(G) : v_i \notin S\}$. Then the following holds:

Lemma 1.3.2. *[18][Lemma 6.2] For G not a complete graph,*

$$\frac{\text{vol}(\Gamma(S))}{\text{vol}(S)} > \frac{1}{1-(1-(\lambda_2(M))^2)\frac{\text{vol}(S^C)}{\text{vol}(G)}}.$$

This lemma specifically indicates how quickly the neighborhood of a set expands.

From N. Alon comes have the following result. For two disjoint sets of vertices B and C, let $c = \frac{|C|}{n}$, let $b = \frac{|B|}{n}$, let σ be the distance between C and B. Then $b \leq \frac{1-c}{1+\mu_2(L)/d_{\max}c\sigma^2}$, or $\sigma^2 \leq \frac{(1-c-b)d_{\max}}{b\mu_2(L)c}$. In particular, if C and B are both single vertices, then $(\text{diam}(G))^2 \leq \frac{(n^2-2n)d_{\max}}{\mu_2(L)}$ [4].

Chapter 2

AN OVERVIEW OF RANDOM GRAPHS

2.1 Definitions

It is difficult to find a definition of a random graph that is simultaneously broad enough and precise enough to be useful. F. Chung and L. Lu define a random graph as being "a probability space (consisting of some family $\mathcal{F}$ of graphs) together with a probability distribution (which assigns to each member of $\mathcal{F}$ a probability of being chosen)" [17]. I find that this definition is better suited for a random graph model, with the term "random graph" applied to a selection from the random graph model. More specifically, by adapting R. Diestel's explanation [20] of the most common random graph model, we can view any random graph model on N vertices as being a probability space $(\mathcal{G}(N), \mathcal{S}, P)$ where $\mathcal{G}(N)$ is the set of all graphs G on the vertex set $\{v_1, v_2, \ldots, v_N\}$ and the probability measure P is determined by the particular model we are using. Once the probability measure is chosen, any attribute α of a graph on N vertices defines a random variable. The advantage of this formulation is that if α maps $\mathcal{G}(N)$ to $\mathbb{R}$, then we can talk about the expectation and variance of α. For the random graph models in which the number of edges between any two vertices is determined independently, the probability space for the whole model is simply the product space of probability spaces for the individual edges. For

models in which the number of edges between each pair of vertices v_i and v_j is determined by an independent Bernoulli random variable a_{ij}, we can set up a probability matrix P. To construct this matrix, let $p_{ij} = \Pr(a_{ij} = 1)$. The distribution of the random graph model is then denoted by $G(P)$. In this case, the adjacency matrix A is a matrix-valued random variable.

A *random matrix* is a matrix whose entries are random variables. Thus the matrix itself can be seen as a random variable. When analyzing edge-product models, we treat the adjacency matrix of $G \sim G(P)$ as a random matrix. Thus we can use results for random matrices to characterize edge-product random graphs.

Some models of random graphs that have been published recently do not fall within this definition, as the algorithms for their construction may involve the introduction of new vertices as well as the addition of edges. Thus if our random graph model produces graphs iteratively, we use what B. Bollobás terms a *random graph process* [15]: a Markov chain whose state space is graphs. In [15], these graphs are on a fixed set of vertices. Since models have been introduced which affect the size of the vertex set, we can expand the definition by defining $\mathcal{G} = \bigcup_{N \geq 1} \mathcal{G}(N)$ and allowing the random graph process to have state space $\mathcal{G}$.

There are few real-world examples of networks that could truly be considered random, ad-hoc wireless networks being a notable exception. If, however, we find that a large proportion of graphs with certain characteristics behave in some way, we can generalize the effects of a process on this type of graph by simulating the process on a set of random graphs that share these characteristics, or even by analyzing the model that produces such graphs. If, for example, we assume that all social networks have an

exponential degree distribution, then we can simulate the progress of a certain disease on random graphs with the proper exponential distribution and use the results to predict what will happen in real life. There are two major assumptions here. The first is that we do not need any characteristics other than degree distribution to model this disease with sufficient accuracy. The second assumption is that the real-world network behaves in the same manner as the networks which are drawn from our chosen model.

One difficulty regarding research into what is already known about random graphs is that much of the available information is contained in either the physics literature or in the literature specific to a particular discipline in which random networks are used. One reason for this is that with current computing power it is easier to run simulations and analyze the results than it is to prove theorems.

2.2 Published models

The paper cited by M. Newman in [37] as the first systematic study of a random graph, *Connectivity of Random Nets* [43], was published by R. Solomonoff and A. Rapoport in 1951. Solomonoff and Rapoport gave a general algorithm for generating a random net with N neurons, where each neuron has exactly a axones connecting it to other neurons. Then, generalizing to the instance where the average number of axones per neuron is a, they proceeded to analyze expected connectivity and the expected number of neurons that are at distance k from an initial neuron. Most of the analysis in that paper uses the fact that all neurons are qualitatively the same, so it suffices to examine only one. Further, they specifically take advantage of the fact that the network in the neighborhood of any

given neuron is very close to being a tree. In the same paper Solomonoff and Rapoport note the usefulness of random graphs in epidemiology and mathematical genetics.

In 1959, P. Erdös and A. Rényi [23] built on previous efforts, exploring a model of a random graph with n vertices and m edges. These edges are then distributed randomly among the pairs of vertices with equal probability. This family of graphs was denoted by $G_{n,m}$. In the probability space formulated in section 1.2, each graph with n vertices and m edges has equal probability. In this same paper, an alternative model was suggested, $G_{n,p}$, which is now called the Erdös-Rényi model. To generate a graph according to $G_{n,p}$ start with n vertices and fix $0 \leq p \leq 1$. Then let the adjacency of each vertex pair be determined by a Bernoulli random variable a_{ij} with $\Pr(a_{ij} = 1) = p$. This graph model is then denoted by $G_{n,p}$. Thus the probability measure is supported on the set of simple graphs with $|V(G)| = n$ and $0 \leq E(G) \leq \binom{n}{2}$. The number of edges in a graph $G \sim G_{n,p}$ is then given by a binomial distribution.

In 1978, E. Bender and E.R. Canfield determined the number of graphs with a specific degree distribution [13]. Building from here, M. Molloy and B. Reed designed an algorithm for generating a random graph with a specified degree distribution by assigning a certain number of half-edges to each vertex and then proceeding to pair them randomly [33]. This paper, according to M. Newman, introduced the first random graph model to allow for a degree distribution that is neither Poisson nor constant [37]. The model was then analyzed to find conditions for the formation of a giant component of a graph generated by this model.

According to R. Durrett [22], the first "small-world" graph was developed by B. Bollobás and F. Chung in 1988 [14]. While Milgram had only

inferred information about a social network, Bollobás and Chung designed a random graph model to have a small average degree and small diameter. This model is generated by creating an N-cycle on N vertices and then adding edges using a random matching between vertices. The results of that paper show that the diameter of such a graph is proportional to $\sqrt{\log N}$.

D. Watts and S. Strogatz published their "small-world" network model in 1998 to attempt to better model social networks [49]. This model starts by establishing a one-dimensional periodic lattice (or ring lattice) of N vertices, each vertex having initial degree $2k$. This lattice is formed by arranging the vertices in a circle and then creating an edge between each vertex and its k closest neighbors on each side. An example is given in Figure 2.1. Then one of the endpoints of each edge is changed with a certain probability, p. In this model, self-adjacencies and multiple edges are explicitly prohibited. The structural properties of the graphs are quantified by characteristic path length $L(p)$, or average distance between all vertex pairs, and the clustering coefficient $C(p)$, the likelihood of two vertices being adjacent given that they are both adjacent to a third vertex. While it would be expected that a high $C(p)$ is correlated directly with a high $L(p)$, Watts and Strogatz found that relatively low values of p (around 0.01) result in both a high $C(p)$ and a low $L(p)$. While this became the standard small-work network and is easily generated for simulation, it has significant disadvantages for mathematical analysis.

The difficulties of analyzing the Watts-Strogatz model led Newman and Watts to publish a slightly different model [42], in which the structure of a lattice is maintained and an edge is added between each vertex pair

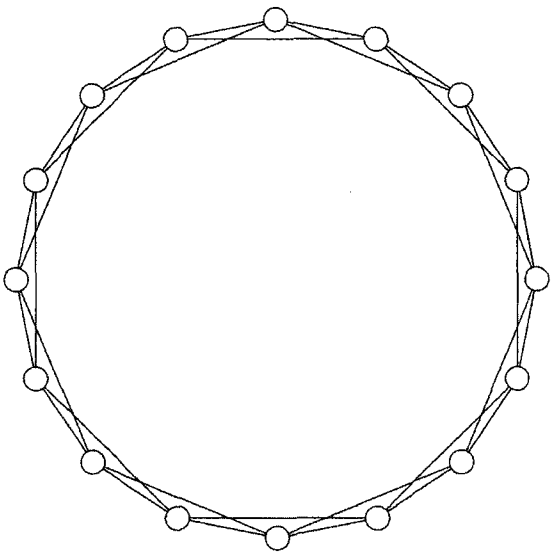

Figure 2.1: A ring lattice with 16 vertices and radius 2

with a certain probability, ϕ. In the Newman-Watts model, multiple edges between vertex pairs are allowed. This greatly eases the analysis of the model, as the probability of events relies on only 2 parameters: N and ϕ. At low values of ϕ, they found that the characteristics of graphs generated by this model are empirically very close to those of graphs generated by the Watts-Strogatz model.

Further advancement of this type of model occurred when A.D. Barbour and G. Reinert [11] analyzed small world networks using a continuous circle model. Starting from the parameters N, k, and p from the Watts-Strogatz model, we embed a network of N vertices in $\mathbb{R}^2$ as a circle of circumference N and assume that the given radius k is the distance around the circle that a process can move in one step. We then insert a set of random chords, the number of which is determined by a random variable drawn from a Poisson distribution with mean nkp. The chords are then treated as having length zero. This then provided some level of rigor to the information that had already been gathered empirically.

In 1999, Barabási and Albert [10] published a preferential attachment model in order to try to replicate the power-law distribution they and others had found in various networks that were found empirically. This model starts with a small number of vertices in place and then iteratively adds a new vertex with a certain degree. Edges are then assigned between the new vertex and the existing vertices according to a formula that makes it more likely for a vertex with a higher degree to receive a new edge than a vertex with a lower degree. This was an attempt to provide the mechanism for the power-law distribution of vertex degrees that had been observed in multiple settings, such as the World Wide Web and collaboration networks. The guiding principle for this model is "the rich get richer".

In their 2000 paper [1], W. Aiello, F. Chung, and L. Lu proposed a model for a generalized random graph, $G(\mathbf{w})$. To generate a graph using the $G(\mathbf{w})$ model, start with a sequence of expected degrees, $w_1, \ldots, w_N$, for N vertices, $v_1, \ldots, v_n$. The probability that any two vertices v_i and v_j are adjacent is then given by $w_i w_j \rho_w$, where $\rho_w = \sum_{k=1}^{N} w_k$. In order for none of the probabilities to be greater than one, it is necessary that $\max_i w_i^2 \leq \sum_{k=1}^{n} w_k$. The $G(\mathbf{w})$ model is a probability space on $\mathcal{G}(N)$, being the product space of the individual edges. This model was initially used to analyze graphs whose degree sequences obey a power-law distribution. It can, however, be applied to most degree sequences. Unlike earlier algorithms which resulted in a power-law degree distribution, this model does not try to mimic how any real-world network might be formed.

This is by no means an exhaustive list of random graph models. These models are those which pertain most directly to the present work. The models examined in this dissertation are either modifications or analogs of those presented above.

The following definitions apply only to random graphs.

Definition 2.2.1. *The expected volume of a random graph G is denoted by* $\mathrm{Vol}(G)$.

Definition 2.2.2. *The expected kth-order volume of a random graph G is denoted by* $\mathrm{Vol}_k(G)$.

Chapter 3

RESULTS FOR ALL EDGE-PRODUCT RANDOM GRAPH MODELS

In the case of an edge-product random graph model, we can obtain probability bounds for several attributes of the realizations of that model. The expected degree of any vertex v_i is the sum of the entries in the ith row of the probability matrix. The expected eigenvalues of a graph chosen according to a given probability matrix are not generally the eigenvalues of the probability matrix, though the eigenvalues of the probability matrix do provide bounds for the expected eigenvalues. This is particularly useful because it is much easier to find the spectrum of a single matrix than it is to find the spectra of several matrices. If our goal is to determine how a process will run on a certain type of network we can use the information gathered from the probability matrix to predict the probability of the process achieving a certain result. One example would be whether an SIR disease process will affect more than a fixed number of nodes.

3.1 Degree distributions of edge-product random graphs

There is a set of properties regarding degree distribution which hold true for all random graph models that can be generated from any probability matrix. Let P be a symmetric matrix with each $p_{ij} \in [0, 1]$. Let $\mathcal{G}(P)$ be

the product space of the probability spaces generated by $a_{ji} = a_{ij}$ and $a_{ij} \sim \text{Ber}(p_{ij})$. Let $G \sim \mathcal{G}(P)$. Recall then that the adjacency matrix A is a matrix-valued random variable. Most of these distribution bounds are derived from the Chernoff inequalities listed in the appendix.

Lemma 3.1.1. *(Analog of [17, Lemma 5.6]) For a graph $G \sim \mathcal{G}(P)$, let $w_i = \sum_{j=1}^{n} p_{ij}$. For all $c > 0$, with probability at least $1 - \exp(-c^2/2)$, the degree d_i of vertex v_i satisfies*

$$d_i > w_i - c\sqrt{w_i}.$$

With probability at least $1 - \exp\left(-\frac{c^2}{2(1+c/(3+\sqrt{w_i}))}\right)$, d_i satisfies

$$d_i < w_i + c\sqrt{w_i}.$$

Proof. This lemma is a direct result of the fact that w_i is the sum of independent Bernoulli random variables. Apply Theorem A-1 to $\sum_{j \neq i} a_{ij}$ with $\lambda = c\sqrt{w_i}$, giving the desired result. □

We gain from this lemma a probability distribution for the degree of each vertex.

Lemma 3.1.2. *(Analog of [17, Lemma 5.9]) For $G \sim \mathcal{G}(P)$ and for $0 < c \leq \sqrt{\text{Vol}(S)}$, with probability at least $1 - 2\exp(-3c^2/7)$, the volume of a subset $S \subseteq V(G)$ satisfies $|\text{vol}(S) - \text{Vol}(S)| < c\sqrt{\text{Vol}(S)}$.*

Proof. Let $S^C = \{v \in V(G), v \notin S\}$. We can view $\text{vol}(S)$ as a sum of independent random variables:

$$\text{vol}(S) = \sum_{v_i \in S} \sum_{v_j \in V(G)} a_{ij} = \sum_{v_i \in S, v_j \in S^C} a_{ij} + 2 \sum_{i<j, v_i, v_j \in S} a_{ij}.$$

By definition, the expected value of $\mathrm{vol}(S)$ is $\mathrm{Vol}(S)$. Let A be the adjacency matrix for G. In order to bound the volume using Theorem A-4 and Theorem A-5 we must know $E[\sum_{v_i \in S} \sum_{j=1}^{n} (a_{ij})^2]$. Using the fact that each $a_{ij} \sim \mathrm{Ber}(p_{ij})$ and so $a_{ij}^2 = a_{ij}$, we have

$$\begin{aligned} E\left[\sum_{v_i \in S} \sum_{v_j \in S} (a_{ij})^2\right] &= 2 \sum_{v_i \in S, v_j \in S, i<j} E[a_{ij}] + \sum_{v_i \in S, v_j \in S^C} E[a_{ij}] \\ &\leq 2E[\mathrm{vol}(S)]. \end{aligned}$$

Since we have assumed $c \leq \sqrt{\mathrm{Vol}(S)}$, Theorem A-4 gives us

$$\begin{aligned} \Pr[\mathrm{vol}(S) > \mathrm{Vol}(S) + c\sqrt{\mathrm{Vol}(S)}] &< \exp\left(-\frac{c^2 \mathrm{Vol}(S)}{2(2\mathrm{Vol}(S) + 2c\sqrt{\mathrm{Vol}(S)}/3)}\right) \\ &< \exp\left(-3c^2/16\right). \end{aligned}$$

Likewise, from Theorem A-5,

$$\Pr[\mathrm{vol}(S) < \mathrm{Vol}(S) - c\sqrt{\mathrm{Vol}(S)}] < \exp\left(-3c^2/16\right).$$

□

This lemma gives us bounds on the volume of a set of vertices. Further, if we let $S = V(G)$, we then have the following corollary.

Corollary 3.1.3. *(Analog of [17, Lemma 5.8]) For $G \sim \mathcal{G}(P)$ and for $0 < c \leq \sqrt{\mathrm{Vol}(G)}$, with probability at least $1 - 2\exp(-c^2/6)$, the volume of G satisfies $|\mathrm{vol}(G) - \mathrm{Vol}(G)| < c\sqrt{\mathrm{Vol}(G)}$.*

3.2 Spectra of edge-product random graphs

Some work has been done regarding the spectra of specific random graph models. This work has been either empirical in nature or related

only to models without any intrinsic structure, e.g. those for which all vertices of degree m have the same probability of being adjacent to any given vertex. I. Farkas *et al.* noted that for $G \sim G_{n,p}$, the eigenvalues of $A(G)$ follow a semicircle distribution other than in the case where n is large p is small [24]. In this same paper, they described the distribution of the eigenvalues for the Watts-Strogatz model as having several sharp maxima for low values of p, attributing this to the remnants of the initial lattice. Since most real-life networks have a structure which is defined by more than the degree sequence, it would be good to find a unified theory for all edge-product random graph models. To this end, we use the Courant-Weyl inequalities to determine the maximum difference between the ith largest eigenvalue of a probability matrix P and the ith largest eigenvalue of the adjacency matrix of a graph obtained from the model described by P.

Theorem 3.2.1. *[19, Theorem 2.1] For all symmetric matrices S and T,*

$$\lambda_{i+j+1}(S+T) \leq \lambda_{i+1}(S) + \lambda_{j+1}(T) \text{ and}$$
$$\lambda_{n-i-j}(S+T) \geq \lambda_{n-i}(S) + \lambda_{n-j}(T).$$

The consequence of Theorem 3.2.1 that will be used most heavily is

$$|\lambda_i(S) - \lambda_i(T)| \leq \lambda_1(S-T). \tag{3.2.1}$$

Thus we can bound the distance of any $\lambda_i(A)$ from $\lambda_i(P)$ by the spectral radius of $A-P$. Also, if the random variable with the greatest variance has expected value α, then $\lambda_1(A-P) \leq \lambda_1(B-\alpha(J-I))$, where $B \sim G_{n,\alpha}$ and J is the all 1s matrix. While direct application of this theorem does give us some bounds, we can do better.

The following statements from N. Alon *et al.*, Z. Füredi and J. Komlós, and V. Vu provide a means of finding probability bounds for $\lambda_1(A-P)$.

Proposition 3.2.2. *[5, Theorem 1 and concluding remark]Let M be an $n \times n$ symmetric random matrix with each m_{ij} a random variable supported on $[0,1]$. For every integer $1 \leq s \leq n$, the probability that $\lambda_s(M)$ deviates from its median by more than t is at most $4\exp(-t^2/8s^2)$.*

Since the conditions for this theorem require that for $1 \leq i \leq j \leq n$, $m_{ij} = m_{ji}$ be independent real random variables with $m_{ij} \in [0,1]$, this result cannot be applied directly to either the combinatorial or normalized Laplacian matrix of a random graph family. On the other hand, the results of this proposition do not depend on n or on the variance of the random variables. Thus when the expected eigenvalues are large, Proposition 3.2.2 is effective for providing bounds that are small relative to the size of the expected eigenvalues.

Suppose $G \sim \mathcal{G}(P)$. Let A be the adjacency matrix of G. Let $\sigma_{ij}^2 = p_{ij}(1-p_{ij})$, the variance of a_{ij}, and let $\sigma^2 = \max\{\sigma_{ij}^2\}$. Let $Q = A - P$. From Z. Füredi and J. Komlós [25], we have for $k \leq \sigma^{1/3}n^{1/6}$,

$$\Pr\left[\lambda_1(Q) > 2\sigma\sqrt{n} + c\right] \leq \sqrt{n}\left(1 - \frac{c}{2\sigma\sqrt{n}+c}\right)^k \tag{3.2.2}$$

$$\leq \sqrt{n}\exp\left(\frac{-kc}{2\sigma\sqrt{n}+c}\right) \tag{3.2.3}$$

This inequality of Z. Füredi and J. Komlós is an intermediate result on the way to their theorem which allows the probability of the second eigenvalue being less than the given bound to go to 1 as $n \to \infty$:

Theorem 3.2.3. *[25, Theorem 1] For a symmetric random matrix A, with each a_{ij} having a common expectation μ and variance σ^2, with probability tending to 1, $\max_{i\geq 2}|\lambda_i(A)| < 2\sigma\sqrt{n} + O(n^{1/3}\log n)$.*

Theorem 3.2.4. *[25, Theorem 2] Under the conditions of Theorem 3.2.3, in case* $\mu = 0$ *we have* $\max_{1 \le i \le n} |\lambda_i(A)| < 2\sigma\sqrt{n} + O(n^{1/3} \log n)$ *(with probability tending to 1).*

V. Vu then generalized inequality (3.2.2), again letting $n \to \infty$ for the theorem, to allow for $k \le \sigma^{1/3} n^{1/4}$, giving the following:

Theorem 3.2.5. *[46, Theorem 1.3] Let* σ *and* Q *be defined as for equation 3.2.2. Then there exists a constant* $c = c(\sigma)$ *such that*

$$\lambda_1(Q) \le 2\sigma\sqrt{n} + cn^{1/4} \ln n$$

holds almost surely.

In addition, we know that the largest absolute eigenvalue of a matrix is an operator norm. Thus the following holds for any two symmetric matrices S and T of equal dimension,

$$\lambda_1(S+T) \le \lambda_1(S) + \lambda_1(T).$$

If we are looking at a random graph model which uses k different probabilities, this can be used to our advantage by letting A_{α_i} be the adjacency matrix of edges that occur with probability α_i. Then, letting P_{α_i} be the probability matrices with their respective α_i entries,

$$\lambda_1(A - P) \le \sum_{i=1}^{k} \lambda_1(A_{\alpha_i} - P_{\alpha_i}). \tag{3.2.4}$$

We can then use inequality (3.2.2) to find probability bounds for each $\lambda_1(A_{\alpha_i} - P_{\alpha_i})$. This method will only be helpful with a small number of distinct probabilities. This will be applied to models with 2 distinct probabilities of adjacency in section 5.2.

In the special case of $G_{n,p}$, we get slightly better bounds on eigenvalues than the current known results. Z. Füredi and J. Komlós used a procedure very similar to finding the largest eigenvalues of $A-P$ to get their results for the distribution of the smaller eigenvalues of $G \sim G_{n,p}$ [25]. They, however, used the fact that $\lambda_2(A) \leq \lambda_1(A - pJ)$, where J is the all-ones matrix.

Using the fact that $\lambda_2(P) = -p$ and the results of Equation 3.2.2, however, we have the following:

Corollary 3.2.6. *For $G \sim G_{n,p}$,*

$$\Pr\left[\lambda_2(A) > 2\sigma\sqrt{n} - p\right] \leq \sqrt{n}\left(1 - \frac{1}{2\sigma\sqrt{n}}\right)^{\left\lfloor \sigma^{1/3} n^{1/6} \right\rfloor},$$

where $\sigma^2 = p(1-p)$.

Since the difference between Theorem 3.2.3 and Corollary 3.2.6 is a constant, there is no improvement on the theorem itself. The only improvement is found when trying to calculate the upper bound of $\lambda_2(A)$ for a particular model.

3.3 Constructing a probability matrix for a patchy habitat problem

From the previous results we can construct a probability matrix that will generate a random graph model with a certain set of desired attributes. The expected degree of each vertex is controlled completely by the sums of the probabilities of the individual edges. We also have some control over the eigenvalues of the adjacency matrix through the Courant-Weyl inequalities as given by Inequality (3.2.1).

If our desired network should always contain $\{v_i, v_j\}$, then we set $p_{ij} = 1$. Likewise, if the desired network should never have $\{v_i, v_j\}$, then we

set $p_{ij} = 0$. One particular application of this is modeling the boundaries of a patchy habitat. It has been noted that it is necessary to to allow for the influence of the boundaries of a study area [7]. The network models discussed by D. Urban *et al.* in their review article [45] do not address the issue of network boundaries.

Suppose there is a study area containing habitat patches for a given species that has a range larger than the study area. We might know all of the connections between the habitat patches within the study area. We may also know from which patches the species can leave or enter the study area. In order to model the boundary, we invent other patches that are outside the study area and assign probabilities of connections between the known patches and the invented patches using whatever data is available. It is important to note that these probability values need not be the same. An example of this is given in figures 3.1 through 3.4.

We can then establish the spectrum of P and the spectral radius of $A - P$. For the simplified case of one distinct probability, m "outside" nodes, and r "inside" nodes which can connect to the "outside", $A - P$ has a spectral radius at most $2\sqrt{(r+m)p(1-p)}$, which is the spectral radius of $A - P$ for $G \sim G_{m+r,p}$. Thus, any attributes of a network that are correlated with the spectrum are then bounded. If, for example, we consider the progress of an introduced species to be similar to a disease process (with low probability of "recovery" if there are not attempts to eradicate the species from the patches), we can apply the results of Theorem 1.3.1 to find out whether this species will spread through the network or if it will remain confined to a few patches.

One reason for proceeding in this manner rather than by modifying movement parameters between the patches within the study area is that

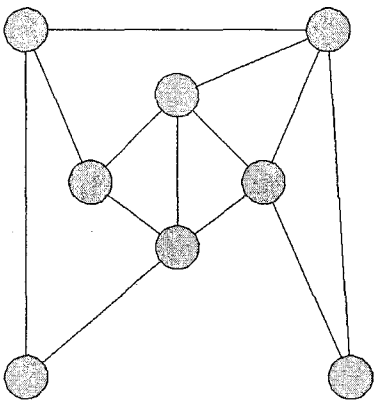

Figure 3.1: A simple set of habitat patches

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 3.2: The adjacency matrix for the graph in Figure 3.1

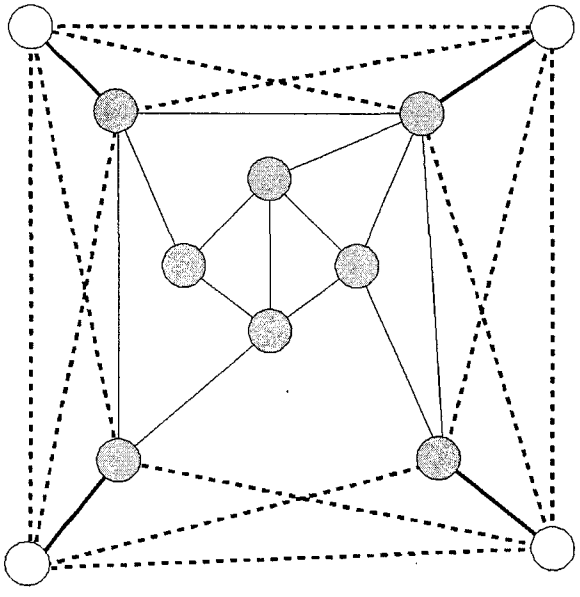

Figure 3.3: A simple set of habitat patches with added "outside" patches. The dashed lines indicate edges that exist with some probability $p \in (0, 1)$.

$$\begin{pmatrix}
0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & p & 0 & p \\
1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & p & 1 & p & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & p & 1 & p \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & p & 0 & p & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & p & 0 & p & 0 & 0 & 0 & 0 & 0 & p & p & p \\
p & 1 & p & 0 & 0 & 0 & 0 & 0 & p & 0 & p & p \\
0 & p & 1 & p & 0 & 0 & 0 & 0 & p & p & 0 & p \\
p & 0 & p & 1 & 0 & 0 & 0 & 0 & p & p & p & 0
\end{pmatrix}$$

Figure 3.4: The adjacency matrix for the graph in Figure 3.1

the use of a random graph model allows for the elimination of a species within the study area followed by a natural return from outside the study area. This would be useful when deciding whether an invasive species is likely to establish itself within the network despite efforts to remove it from each patch. We can account for any such efforts by establishing a higher "recovery" rate.

Chapter 4

THE MODIFIED CHUNG-LU MODEL

The $G(\mathbf{w})$ model as described previously allows for each vertex to be self-adjacent with probability $w_i^2\rho_w$. Since self-adjacency in real-world networks is often either much higher or much lower than this, I have modified this model slightly so that self-adjacencies are avoided. In the modified Chung-Lu model, $\widehat{G}(\mathbf{u})$ on $\mathcal{G}(n)$, we start with an n-dimensional vector $\mathbf{u}$. The probability that any two vertices v_i and v_j are adjacent is then given by

$$p_{ij} = \begin{cases} u_i u_j \rho & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases}, \text{ where}$$

$$\rho = \sum_{i=1}^{n} u_i. \tag{4.0.1}$$

Like the $G(\mathbf{w})$ model it is necessary that $u_i < \sqrt{\sum_{j=1}^{n} u_j}$ for all i. The $\widehat{G}(\mathbf{u})$ model does then avoid self-adjacency, but at the cost of simplicity of results. For example, the expected degree of vertex v_i is given by

$$E[d_i] = w_i = u_i - u_i^2\rho. \tag{4.0.2}$$

As the $\widehat{G}(\mathbf{u})$ model is based on the $G(\mathbf{w})$ model, this chapter is based largely on the work of F. Chung and L. Lu in [17]. Some of the lemmas that were developed for $G(\mathbf{w})$ have been redeveloped here for $\widehat{G}(\mathbf{u})$ with

explicit probability bounds. In this chapter we first look at bounds for the volumes of sets of vertices. We will then find probability bounds for the volume of the neighborhood of a set of vertices. We will finish our tour of this particular random graph model with a look at a special case: $G_{n,p}$.

4.1 Bounds for the volumes of sets

Several bounds exist for specific attributes of $\widehat{G}(\mathbf{u})$ relating to the volume of a set of vertices and the number of edges between sets of vertices.

We would like to know the number of edges between any two subsets, S and T of $V(G)$ where $G \sim \widehat{G}(\mathbf{u})$. If these two subsets are disjoint, then the probability that any given edge has an endpoint in S is $\frac{\text{Vol}(S)}{\text{Vol}(G)}$. The probability that the other endpoint of this edge is in T is $\frac{\text{Vol}(T)}{\text{Vol}(G)}$. So the probability that any given edge has one endpoint in S and one endpoint in T is $\frac{\text{Vol}(S)\text{Vol}(T)}{\text{Vol}^2(G)}$. Since the expected number of endpoints is $\text{Vol}(G)$, the expected number of edges between S and T is $\frac{\text{Vol}(S)\text{Vol}(T)}{\text{Vol}(G)}$.

If $S \cap T \neq \emptyset$, further definition is needed regarding the counting of edges. Let

$$e(S,T) = |\{\{v_i, v_j\} : v_i \in S, v_j \in T\}|. \tag{4.1.1}$$

This definition does count each edge with both endpoints in $S \cap T$ twice. For $G \sim G(\mathbf{w})$, we find that $e(S,T) = \frac{\text{Vol}(S)\text{Vol}(T)}{\text{Vol}(G)}$ from the fact that the proportion of endpoints of $E(G)$ in S is $\frac{\text{Vol}(S)}{\text{Vol}(G)}$, the proportion of endpoints of $E(G)$ in T is $\frac{\text{Vol}(T)}{\text{Vol}(G)}$, and the total number of endpoints of $E(G)$ is $\text{Vol}(G)$. The model $\widehat{G}(\mathbf{w})$ is not so well behaved. In order to give the general case, we must account for the fact that if vertex v_k is in $S \cap T$, then Equation 4.1.1 includes a non-zero probability of an edge from v_k to itself. An exact formula for $E[e(S,T)]$ is given by $(\sum_{v_i \in S} u_i)(\sum_{v_j \in T} u_j)\rho - (\sum_{k \in S \cap T} u_k^2 \rho)$, which

unfortunately does not appear to translate nicely into terms of expected volumes. We therefore need to introduce a new quantity.

Definition 4.1.1. *For $G \sim \widehat{G}(\boldsymbol{u})$, and $S \subseteq V(G)$, $\widehat{\mathrm{Vol}}(S) = \sum_{v_i \in S} u_i$. Likewise, $\widehat{\mathrm{Vol}}(G) = \widehat{\mathrm{Vol}}(V(G)) = \sum_{v_i \in V(G)} u_i$. We also denote $(\widehat{\mathrm{Vol}}(S))^k$ by $\widehat{\mathrm{Vol}}^k(S)$.*

Note that $\rho = \frac{1}{\widehat{\mathrm{Vol}}(G)}$. The quantity $\mathrm{Vol}(G)$ is the expected volume of G in the $G(\mathbf{u})$ random graph model. Since we have eliminated the possibility of self-adjacency in the $\widehat{G}(\mathbf{u})$ random graph model, $\widehat{\mathrm{Vol}}(G) > \mathrm{Vol}(G)$.

In order to find the probability that there is no edge between S and T when S and T are disjoint, we need to describe the neighborhood of a set S. Let S^C denote $V(G) - S$.

Definition 4.1.2. *For $S \subseteq V(G)$, the neighborhood of S, denoted by $\Gamma(S)$, consists of all vertices adjacent to some vertex in S, but not in S;*
$\Gamma(S) = \{v_j \in S^C : \exists v_i \in S \text{ such that } \{v_i, v_j\} \in E(G)\}$.

Lemma 4.1.3. *(Analog of [17, Lemma 5.12]) For $G \sim \widehat{G}(\boldsymbol{u})$ and for any two disjoint subsets S and $T \subseteq V(G)$,*

$$\Pr[\Gamma(S) \cap T = \emptyset] < \exp(-\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)\rho).$$

Proof. By definition, $\sum_{v_i \in S} u_i = \widehat{\mathrm{Vol}}(S)$, $\sum_{v_j \in T} u_j = \widehat{\mathrm{Vol}}(T)$, and $\Pr[\{v_i, v_j\} \in E(G)] = u_i u_j \rho$.
Since the presence of each edge is determined independently,

$$\begin{aligned}
\Pr[\Gamma(S) \cap T = \emptyset] &= \prod_{v_i \in S, v_j \in T} (1 - u_i u_j \rho) \\
&\leq \exp\left(- \sum_{v_i \in S, v_j \in T} u_i u_j \rho \right) \\
&= \exp\left(-\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)\rho \right).
\end{aligned}$$

□

We can also find bounds for the number of edges between any two sets.

Lemma 4.1.4. *For $G \sim \widehat{G}(\boldsymbol{u})$, and for $S, T \subseteq V(G)$,*

$$E[e(S,T)] = \frac{\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)}{\widehat{\mathrm{Vol}}(G)} - \sum_{k \in S \cap T} u_k^2 \rho.$$

Proof. By definition, $e(S,T) = \sum_{v_i \in S, v_j \in T-S} X_{ij} + 2\sum_{v_i \in S, v_j \in T \cap S} X_{ij}$ where $X_{ij} \sim \mathrm{Ber}(p_{ij})$. So

$$\begin{aligned} E[e(S,T)] &= \sum_{v_i \in S, v_j \in T, v_i \neq v_j} u_i u_j \rho \\ &= \sum_{v_i \in S, v_j \in T} u_i u_j \rho - \sum_{v_k \in S \cap T} u_k^2 \rho \\ &= \left(\sum_{v_i \in S} u_i\right)\left(\sum_{v_j \in T} u_j\right)\rho - \sum_{v_k \in S \cap T} u_k^2 \rho \\ &= \frac{\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)}{\widehat{\mathrm{Vol}}(G)} - \sum_{v_k \in S \cap T} u_k^2 \rho. \end{aligned}$$

□

With the expected value in hand, we can find bounds for the distribution.

Lemma 4.1.5. *For $G \sim \widehat{G}(\boldsymbol{u})$, let $S, T \subseteq V(G)$. Then for $0 < c \leq \sqrt{\frac{\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)}{\widehat{\mathrm{Vol}}(G)}}$, $e(S,T)$ satisfies*

$$\begin{aligned} \Pr\left[\left|e(S,T) - \left(\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)\rho - \sum_{k \in S \cap T} u_k^2 \rho\right)\right| > c\sqrt{\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}(T)\rho}\right] \\ \leq 2\exp\left(\frac{-3c^2}{8}\right). \end{aligned}$$

Proof. Let $e = e(S,T)$. Applying Theorem A-3 with $\lambda = c\sqrt{E[e(S,T)]}$, we have

$$\Pr\left[|e - E[e]| > c\sqrt{E[e]}\right] \leq \exp\left(-\frac{c^2}{2(1 + c/3\sqrt{E[e]})}\right)$$
$$\leq \exp\left(-\frac{3c^2}{8}\right).$$

Substituting the value obtained in Lemma 4.1.4 for $E[e(S,T)]$, we have our result. □

We can now apply this information to determine the expected number of edges within a set $S \subseteq V(G)$, applying Lemma 4.1.4 to the special case of $T = S$.

Lemma 4.1.6. *For $G \sim \widehat{G}(\boldsymbol{u})$ with $\widehat{\mathrm{Vol}}(G) > 0$ and $S \subseteq V(G)$,*

$$E[e(S,S)] = \frac{\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)}{\widehat{\mathrm{Vol}}(G)}.$$

Thus we can achieve probability bounds for the number of edges within a set $S \subseteq V(G)$.

Theorem 4.1.7. *For $G \sim \widehat{G}(\boldsymbol{u})$ and $S \subseteq V(G)$ and for all $0 < c \leq \sqrt{\frac{\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)}{2\widehat{\mathrm{Vol}}(G)}}$, with probability at least $1 - 2\exp(-3c/7)$, the number of edges from S to itself satisfies $\left|e(S,S) - \frac{\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)}{\widehat{\mathrm{Vol}}(G)}\right| < \frac{c}{2}\sqrt{\frac{\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)}{\widehat{\mathrm{Vol}}(G)}}$.*

Proof. By Lemma 4.1.6, $e(S,S)$ is a sum of independent Bernoulli random variables with $E[e(S,S)] = \frac{\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)}{\widehat{\mathrm{Vol}}(G)}$. Let $e = e(S,S)$. We can then

apply Theorem A-1 using $\lambda = c\sqrt{E[e]}$:

$$\begin{aligned}
\Pr\left[|e - E[e]| > c\sqrt{E[e]}\right] &\leq \exp\left(-\frac{(c\sqrt{(E[e])^2}}{2E[e] + c\sqrt{E[e]}/3}\right) \\
&= \exp\left(-\frac{3c^2 E[e]}{6E[e] + c\sqrt{E[e]}}\right) \\
&= \exp\left(-\frac{3c^2}{6 + c\sqrt{E[e]}}\right) \\
&\leq \exp\left(-\frac{3c}{7}\right).
\end{aligned}$$

□

We can also find probability bounds for the volume of a set. We first find the expected volume of a set $S \subseteq V(G)$ in relation to $\widehat{\text{Vol}}(S)$ and $|S|$. This will prepare us to be able to work on the problem of neighborhood growth.

Lemma 4.1.8. *For $G \sim \widehat{G}(\boldsymbol{u})$ and for $S \subset V(G)$,*

$$\widehat{\text{Vol}}(S) - \min\left\{\sqrt{\widehat{\text{Vol}}(G)}, |S|\right\} \leq \text{Vol}(S) \leq \widehat{\text{Vol}}(S).$$

Proof. From Lemma 3.1.1 and Equation 4.0.2,

$$\begin{aligned}
\text{Vol}(S) &= \sum_{v_i \in S} \left(u_i - u_i^2 \rho\right) \\
&= \widehat{\text{Vol}}(S) - \rho \sum_{v_i \in S} u_i^2.
\end{aligned}$$

Since each $u_i \leq \sqrt{\text{Vol}(G)}$ and $\text{Vol}_2(S) \leq \text{Vol}_2(G)$,

$$\begin{aligned}
\widehat{\text{Vol}}(S) - \rho \sum_{v_i \in S} u_i^2 &\geq \widehat{\text{Vol}}(S) - \rho \cdot \min\{(\widehat{\text{Vol}}(G))^{3/2}, |S|\widehat{\text{Vol}}(G)\} \\
&= \widehat{\text{Vol}}(S) - \min\{\sqrt{\widehat{\text{Vol}}(G)}, |S|\}.
\end{aligned}$$

Further, since $\rho \sum_{v_i \in S} u_i^2 > 0$, $\mathrm{Vol}(S) \leq \widehat{\mathrm{Vol}}(S)$, giving

$$\widehat{\mathrm{Vol}}(S) - \min\{\sqrt{\widehat{\mathrm{Vol}}(G)}, |S|\} \leq \mathrm{Vol}(S) \leq \widehat{\mathrm{Vol}}(S).$$

□

With the expected volume of a set in hand, we can then pursue probability bounds for the volume of a set.

Theorem 4.1.9. *For $G \sim \widehat{G}(\boldsymbol{u})$, $S \subseteq V(G)$, and $c \leq (\widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)})^3$, with probability at least $1 - \exp(-3c/7)$, the following hold:*

$$\widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)} - \frac{c}{\widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)}} \leq \mathrm{vol}(S)$$

$$\mathrm{vol}(S) \leq \widehat{\mathrm{Vol}}(S) + \frac{c}{\widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)}}.$$

Proof. By Lemma 4.1.8, $\widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)} \leq \mathrm{Vol}(S) \leq \widehat{\mathrm{Vol}}(S)$.

Thus we can apply Theorem A-2, using $\lambda = \frac{c}{\mathrm{Vol}(S)}$, which gives us

$$\begin{aligned}
\Pr\left[|\mathrm{vol}(S) - \mathrm{Vol}(S)| > \frac{c}{\mathrm{Vol}(S)}\right] &\leq \exp\left(-\frac{(c/\mathrm{Vol}(S))^2}{2\mathrm{Vol}(S) + (c/3\mathrm{Vol}(S))}\right) \\
&= \exp\left(-\frac{3c^2\mathrm{Vol}(S)}{6(\mathrm{Vol}(S))^4 + c\mathrm{Vol}(S)}\right) \\
&= \exp\left(-\frac{3c^2}{6(\mathrm{Vol}(S))^3 + c}\right) \\
&\leq \exp\left(-\frac{3c}{7}\right).
\end{aligned}$$

Since $\mathrm{Vol}(S) \geq \widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)}$, $\frac{c}{\mathrm{Vol}(S)} \leq \frac{c}{\widehat{\mathrm{Vol}}(S) - \sqrt{\widehat{\mathrm{Vol}}(S)}}$, which completes the inequality.

□

We also can achieve alternative bounds that are based also on $|S|$. This would prove useful if $\sqrt{\widehat{\text{Vol}}(S)} > |S|$.

Theorem 4.1.10. *For $G \sim \widehat{G}(\boldsymbol{u})$, for $S \subset V(G)$ such that $|S| \leq \min\{\sqrt{\widehat{\text{Vol}}(G)}, \widehat{\text{Vol}}(S)\}$, and for $c \leq (\widehat{\text{Vol}}(S) - |S|)^3$, with probability at least $1 - \exp(-3c/7)$,*

$$\widehat{\text{Vol}}(S) - |S| - \frac{c}{\widehat{\text{Vol}}(S) - |S|} \leq \text{Vol}(S) \leq \widehat{\text{Vol}}(S) + \frac{c}{\widehat{\text{Vol}}(S) - |S|}.$$

Proof. By Lemma 4.1.8 $\widehat{\text{Vol}}(S) - \min\left\{\sqrt{\widehat{\text{Vol}}(G)}, |S|\right\} \leq \text{Vol}(S) \leq \widehat{\text{Vol}}(S)$.

Thus we can apply Theorem A-2, using $\lambda = \frac{c}{\text{Vol}(S)}$, which gives us

$$\begin{aligned}
\Pr\left[|\text{vol}(S) - \text{Vol}(S)| > \frac{c}{\text{Vol}(S)}\right] &\leq \exp\left(-\frac{(c/\text{Vol}(S))^2}{2\text{Vol}(S) + (c/3\text{Vol}(S))}\right) \\
&= \exp\left(-\frac{3c^2\text{Vol}(S)}{6(\text{Vol}(S))^4 + c\text{Vol}(S)}\right) \\
&= \exp\left(-\frac{3c^2}{6(\text{Vol}(S))^3 + c}\right) \\
&\leq \exp\left(-\frac{3c}{7}\right).
\end{aligned}$$

Since $\text{Vol}(S) \geq \widehat{\text{Vol}}(S) - |S|$, $\frac{c}{\text{Vol}(S)} \leq \frac{c}{\widehat{\text{Vol}}(S) - |S|}$. □

Note that if $\widehat{\text{Vol}}(S) \leq |S|$, then at least one vertex $v_i \in S$ has expected degree less than one. Thus, in this case, we expect at least one vertex in S to be isolated.

4.2 Neighborhood Expansion of $\widehat{G}(\mathbf{u})$ (Adaptations of [17], section 5.6 to $\widehat{G}(\mathbf{u})$)

The prevalent cause of interest in the study of graphs is a desire to know how processes work on these graphs in order to apply this knowledge to real life networks. Two problems in the field of ecology for which graphs are used are determining disease spread among a population and determining

the spread of invasive organisms across a set of habitat patches. Both of these questions are related to neighborhood expansion.

Recall the definition of the neighborhood of $S \subseteq V(G)$, Definition 4.1.2. We also consider the neighborhood of a neighborhood. The kth neighborhood of S is denoted as follows: $\Gamma^0(S) = S$ and for $k \geq 1$, $\Gamma^k(S) = \Gamma(\Gamma^{k-1}(S))$. Note that by definition, for $k \geq 1$, $\Gamma^k(S) \subseteq \Gamma^{k+2}(S)$. It is, however, entirely possible for there to exist a vertex v_i such that $v_i \in S$, but $v_i \notin \Gamma^2(S)$, by having v_i be adjacent only to other members of S. In this case, if G is connected, v_i is in $\Gamma^k(S)$ for some $k > 2$ and then also in $\Gamma^{k+2}(S)$.

For this section we will assume that we are working with a set S such that $\widehat{\text{Vol}}(S)$ is relatively small compared to $\widehat{\text{Vol}}(G)$. In particular, we assume that there exists some α such that

$$\widehat{\text{Vol}}(S)u_{\max}\rho \leq \alpha < 1. \tag{4.2.1}$$

Lemma 4.2.1. *For $G \sim \widehat{G}(\boldsymbol{u})$, for a given set $S \subset V(G)$ and a vertex $v_j \notin S$, and for α satisfying Inequality* (4.2.1),

$$\widehat{\text{Vol}}(S)u_j\rho - \alpha^2 \leq \Pr[v_j \in \Gamma(S)] \leq \widehat{\text{Vol}}(S)u_j\rho.$$

Proof. From the proof of 4.1.3,

$$\begin{aligned}\Pr[v_j \notin \Gamma(S)] &= \prod_{v_i \in S}(1 - u_i u_j \rho) \\ &= 1 - u_j\rho \sum_{v_i \in S} u_i + u_j^2\rho^2 \sum_{v_{i_1}, v_{i_2} \in S} u_{i_1}u_{i_2} - u_j^3\rho^3 \sum_{v_{i_1}, v_{i_2}, v_{i_3}} u_{i_1}u_{i_2}u_{i_3} + \dots\end{aligned}$$

Note that

$$\sum_{v_{i_1}, \dots, v_{i_k} \in S} u_{i_1}u_{i_2}\dots u_{i_k} \leq \left(\sum_{v_{i_1}, \dots, v_{i_k} \in S} u_i\right)^k = \widehat{\text{Vol}}^k(S).$$

Thus

$$u_j^k \rho^k \sum_{i_1, i_2, \ldots, i_k} u_{i_1} u_{i_2} \ldots u_{i_k} \leq (u_j \rho \widehat{\mathrm{Vol}}(S))^k. \tag{4.2.2}$$

Therefore $u_j^k \rho^k \sum_{i_1, i_2, \ldots, i_k} u_{i_1} u_{i_2} \ldots u_{i_k} \leq \alpha^k$.

Since $\alpha < 1$, the absolute values of the terms of the alternating series in equation (4.2.2) are bounded by a controlled decreasing sequence.

Thus

$$\begin{aligned} \Pr[v_j \in \Gamma(S)] &= 1 - \Pr[v_j \notin \Gamma(S)] \\ &\geq u_j \rho \widehat{\mathrm{Vol}}(S) - u_j^2 \rho^2 \left(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S) \right) \\ &\geq u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2. \end{aligned}$$

Likewise, since the absolute values of the terms of this alternating series in equation (4.2.2) are bounded by a controlled decreasing sequence,

$$\Pr[v_j \in \Gamma(S)] \leq u_j \rho \widehat{\mathrm{Vol}}(S).$$

□

The expected degree of a vertex given that it is in some neighborhood is not quite the same as the expected degree of a vertex without that condition.

Lemma 4.2.2. *For $G \sim \widehat{G}(\boldsymbol{u})$, for a given set $S \subseteq V(G)$ and a vertex $v_j \notin S$, and for α satisfying Inequality* (4.2.1), $u_j - u_j^2 \rho < E[d_j | v_j \in \Gamma(S)] < u_j - u_j^2 \rho + u_j \rho \widehat{\mathrm{Vol}}(S) \left(1 - \frac{1}{u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2}\right)$.

Proof. According to Bayes' Theorem, for any variable A and any event B, $E[A] = E[A|B] \Pr[B] + E[A|B^C] \Pr[B^C]$. Thus $E[A|B] = \frac{E[A] - E[A|B^C] \Pr[B^C]}{\Pr[B]}$.

We apply this equation using $A = d_j$ with B being the event $v_j \in \Gamma(S)$.

$$E[d_j | v_j \in \Gamma(S)] = \frac{\sum_{v_i \in V(G)} (u_i u_j \rho) - u_j^2 \rho - \Pr[v_j \notin \Gamma(S)] \left(\sum_{v_i \in S^C} (u_i u_j \rho) - u_j^2 \rho \right)}{\Pr[v_j \in \Gamma(S)]}.$$

Let $\beta_j(S) = \frac{u_j \rho}{\Pr[v_j \in \Gamma(S)]}$ and let $p = \Pr[v_j \in \Gamma(S)]$. Then

$$\begin{aligned}
&E[d_j | v_j \in \Gamma(S)] \\
&= \beta_j(S) \left(\sum_{v_i \in V(G)} u_i - u_j - (1-p) \left(\sum_{v_i \in V(G)} u_i - \sum_{v_i \in S} u_i - u_j \right) \right) \\
&= \beta_j(S) \left(\widehat{\mathrm{Vol}}(G) - u_j - (1-p)(\widehat{\mathrm{Vol}}(G) - \widehat{\mathrm{Vol}}(S) - u_j) \right) \\
&= \beta_j(S) \Big(\widehat{\mathrm{Vol}}(G) - u_j - \widehat{\mathrm{Vol}}(G) + \widehat{\mathrm{Vol}}(S) + u_j + \\
&\qquad\qquad (\widehat{\mathrm{Vol}}(G) - \widehat{\mathrm{Vol}}(S))p - u_j \Pr[v_j \in \Gamma(S)] \Big) \\
&= \beta_j(S) \left(\widehat{\mathrm{Vol}}(S)(1-p) + \widehat{\mathrm{Vol}}(G)p - u_j p \right) \\
&= \frac{u_j \rho \widehat{\mathrm{Vol}}(S)(1-p)}{p} + u_j - u_j^2 \rho \\
&= u_j - u_j^2 \rho + \left(\frac{u_j \rho \widehat{\mathrm{Vol}}(S)}{p} - u_j \rho \widehat{\mathrm{Vol}}(S) \right) \\
&\leq w_j + u_j \rho \widehat{\mathrm{Vol}}(S) \left(1 - \frac{1}{u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2} \right) \\
&= u_j - u_j^2 \rho + u_j \rho \widehat{\mathrm{Vol}}(S) \left(1 - \frac{1}{u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2} \right).
\end{aligned}$$

Since $\frac{u_j \rho \widehat{\mathrm{Vol}}(S)(1-p)}{p} \geq 0$, $E[d_j | v_j \in \Gamma(S)] \geq u_j - u_j^2 \rho = w_j$. □

If α is small, then $E[d_j | v_j \in \Gamma(S)]$ is very close to u_j.

Corollary 4.2.3. *For $G \sim \widehat{G}(\boldsymbol{u})$, for a given set $S \subseteq V(G)$, for a given vertex $v_j \in \Gamma(S)$, let α satisfy Inequality (4.2.1). Then with probability at least $1 - \exp\left(-\frac{-c^2}{2}\right)$, $d_j \geq u_j - u_j^2 \rho$ and with probability at least $1 - \exp\left(-\frac{c^2}{2(1+c/(3+\sqrt{u_j}))}\right)$, $d_j \leq u_j - u_j \rho \widehat{\mathrm{Vol}}(S) \left(\frac{1}{u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2} - 1 \right) + c\sqrt{u_j}$.*

Proof. We apply Theorems A-4 and A-5 with $\lambda = c\sqrt{u_j}$, recognizing that $u_j \geq E[d_j | v_j \in \Gamma(S)]$. □

In the case of a set consisting of a single vertex, $S = \{v_i\}$, we have

$$\begin{aligned}
\Pr[d_j = x | v_j \in \Gamma(S)] &= \Pr[d_j = x | X_{ij} = 1] \\
&= \frac{\Pr[(d_j = x) \cap (X_{ij} = 1)]}{\Pr[X_{ij} = 1]} \\
&= \frac{\sum_{r=1}^{x} \left(u_i u_j \rho \Pr\left[\sum_{v_i \in S} X_{ij} = x - 1\right]\right)}{u_i u_j \rho} \\
&= \Pr\left[\sum_{v_i \in S^C} X_{ij} = x - 1\right].
\end{aligned}$$

If u_j is small, this particular operation will almost shift the distribution up by one; $\Pr[d_j = x | v_j \in \Gamma(S)] \approx \Pr[d_j = x - 1]$.

Lemma 4.2.4. *(Analog of Lemma 5.14 in [17]) For $G \sim \widehat{G}(\boldsymbol{u})$, for a given set $S \subset V(G)$, and for α satisfying Inequality* (4.2.1)*, the expected volume of $\Gamma(S)$ satisfies*

$$\mathrm{Vol}_2(S^C)\mathrm{Vol}(S)\rho \leq \mathrm{Vol}(\Gamma(S)) \leq \rho\left(\mathrm{Vol}(S) + \alpha\right)\left(\mathrm{Vol}_2(S^C) + \mathrm{Vol}(S^C)\right).$$

Proof. From the definitions of expected volume and $\Gamma(S)$,

$$\begin{aligned}
\mathrm{Vol}(\Gamma(S)) &= \sum_{v_j \notin S} E[d_j | v_j \in \Gamma(S)] \Pr[v_j \in \Gamma(S)] \\
&\geq \sum_{v_j \notin S} \left((u_j - u_j^2\rho)(u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2) \right) \text{ by Lemma 4.2.2} \\
&\geq \sum_{v_j \notin S} \left(w_j^2 \rho \mathrm{Vol}(S) \right) \\
&= \mathrm{Vol}_2(S^C)\mathrm{Vol}(S)\rho.
\end{aligned}$$

Since $\mathrm{Vol}(S) \leq \widehat{\mathrm{Vol}}(S) \leq \mathrm{Vol}(S) + \alpha$,

$$\begin{aligned}
\mathrm{Vol}(\Gamma(S)) &= \sum_{v_j \notin S} E[d_j | v_j \in \Gamma(S)] \Pr[v_j \in \Gamma(S)] \\
&\leq \sum_{v_j \notin S} \left((u_j \rho \widehat{\mathrm{Vol}}(S)) \left(u_j - u_j^2 \rho + u_j \rho \widehat{\mathrm{Vol}}(S) \left(1 - \frac{1}{u_j \rho \widehat{\mathrm{Vol}}(S) - \alpha^2} \right) \right) \right) \\
&\leq \sum_{v_j \notin S} \left(u_j \rho \widehat{\mathrm{Vol}}(S)(u_j - u_j^2 \rho + u_j \rho \widehat{\mathrm{Vol}}(S) - 1) \right) \\
&\leq \sum_{v_j \notin S} \left((w_j + 1) \rho \widehat{\mathrm{Vol}}(S) w_j \right) \\
&\leq \sum_{v_j \notin S} \left((w_j + 1) w_j \rho (\mathrm{Vol}(S) + \alpha) \right) \\
&\leq \rho \left(\mathrm{Vol}(S) + \alpha \right) \left(\mathrm{Vol}_2(S^C) + \mathrm{Vol}(S^C) \right).
\end{aligned}$$

□

There is an immediate corollary about neighborhood growth.

Corollary 4.2.5. *For $G \sim \widehat{G}(\boldsymbol{u})$ and for a given set $S \subset V(G)$,*

$$\mathrm{Vol}_2(S^C)\rho \leq \frac{\mathrm{Vol}(\Gamma(S))}{\mathrm{Vol}(S)} \leq \left(1 + \frac{1}{u_{\max} \mathrm{Vol}(G)} \right) \left(\mathrm{Vol}_2(S^C)\rho + 1 \right).$$

Proof. This is proven by dividing the results of Lemma 4.2.4 by $\mathrm{Vol}(S)$ and recognizing that $\mathrm{Vol}(S^C)\rho < 1$. □

Proportionally, the difference between the upper and lower bounds is greatest as $\mathrm{Vol}(S)$ approaches $\frac{\mathrm{Vol}(G)}{u_{\max}}$. Also, if $\mathrm{Vol}(S^C)$ is very close to $\mathrm{Vol}(G)$, the lower bounds match the results given in Lemma 5.14 in [17].

In order to bound the volume of $\Gamma(S)$, it is necessary to determine the covariance of combinations of random variables. In general, if $X_1, \ldots, X_m$ and $Y_1, \ldots Y_n$ are random variables, then

$$\mathrm{Cov}\left(\sum_{i=1}^{m} X_i, \sum_{j=1}^{n} Y_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} \mathrm{Cov}(X_i, Y_j).$$

Also, in general, if $X_1, \ldots, X_n$ are random variables, then

$$\begin{aligned}\mathrm{Var}(\sum_{i=1}^{n} X_i) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).\end{aligned}$$

Lemma 4.2.6. *For any random variables X_1, X_2, Y_1, Y_2, Z_1, and Z_2 which are pairwise independent other than (Y_1, Y_2), (X_1, Z_1), and (X_2, Z_2),* $\mathrm{Cov}(X_1 + Y_1Z_1, X_2 + Y_2Z_2) = E[Z_1Z_2]\mathrm{Cov}(Y_1, Y_2)$

Proof. By the definition of covariance,

$$\begin{aligned}\mathrm{Cov}(X_1 + Y_1Z_1, X_2 + Y_2Z_2) &= E[(X_1 + Y_1Z_1)(X_2 + Y_2Z_2)] \\ &\quad - E[X_1 + Y_1Z_1]E[X_2 + Y_2Z_2] \\ &= E[X_1X_2 + X_1Y_2Z_2 + X_2Y_1Z_1 + Y_1Y_2Z_1Z_2] \\ &\quad - (E[X_1] + E[Y_1Z_1])(E[X_2] + E[Y_2Z_2]) \\ &= E[X_1X_2] + E[X_1Y_2Z_2] + E[X_2Y_1Z_1] + E[Y_1Y_2Z_1Z_2] \\ &\quad - (E[X_1] + E[Y_1]E[Z_1])(E[X_2] + E[Y_2]E[Z_2]) \\ &= E[X_1]E[X_2] + E[X_1]E[Y_2]E[Z_2] + E[X_2]E[Y_1]E[Z_1] \\ &\quad + E[Y_1]E[Y_2]E[Z_1]E[Z_2] - (E[X_1]E[X_2] \\ &\quad + E[X_1]E[Y_2]E[Z_2] + E[X_2](E[Y_1]E[Z_1] \\ &\quad + E[Y_1]E[Y_2]E[Z_1]E[Z_2]) \\ &= E[Y_1Y_2]E[Z_1Z_2] - E[Y_1]E[Y_2]E[Z_1Z_2] \\ &= \mathrm{Cov}(Y_1, Y_2)E[Z_1Z_2].\end{aligned}$$

□

This result will allow us to prove the following statement regarding probability bounds for $\mathrm{vol}(\Gamma(S))$:

Theorem 4.2.7. *Let $G \sim \widehat{G}(\boldsymbol{u})$. Let S be some subset of $V(G)$, such that $2 \leq \widehat{\mathrm{Vol}}(S) \leq \max\{\sqrt{2\widehat{\mathrm{Vol}}(G)}, (\widehat{\mathrm{Vol}}(G))^{3/4}\}$. Then for all $c > 1$, with probability at least $1 - \frac{1}{c^2}$,*

$$\mathrm{vol}(\Gamma(S)) > \mathrm{Vol}_2(S^C)\mathrm{Vol}(S)\rho - c\sqrt{2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + 1} \text{ and}$$

$$\mathrm{vol}(\Gamma(S)) < \rho\left(\mathrm{Vol}(S) + |S|\right)\left(\mathrm{Vol}_2(S^C) + \mathrm{Vol}(S^C)\right) + c\sqrt{2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + 1}.$$

Proof. From Lemma 4.2.4, the expected volume of $\Gamma(S)$ satisfies

$$\mathrm{Vol}_2(S^C)\mathrm{Vol}(S)\rho \leq \mathrm{Vol}(\Gamma(S)) \leq \rho\left(\mathrm{Vol}(S) + \alpha\right)\left(\mathrm{Vol}_2(S^C) + \mathrm{Vol}(S^C)\right).$$

Let $\chi_{\Gamma(S)}$ be the indicator function of $\Gamma(S)$, so that $\chi_{\Gamma(S)}(v_j)$ is one if $v_\in\Gamma(S)$, and is zero otherwise. Then we can calculate the volume of $\Gamma(S)$ by the following sum of characteristics of a disjoint union.

$$\begin{aligned}\sum_{v_j\in\Gamma(S)} d_j &= \sum_{v_j\in S^C} \chi_{\Gamma(S)}(v_j)\left(\sum_{v_i\in S} X_{ij} + \sum_{v_k\in S^C} X_{jk}\right) \\ &= \sum_{v_j\in S^C}\left(\sum_{v_i\in S} X_{ij} + \sum_{v_k\in S^C}\left(X_{jk}\chi_{\Gamma(S)}(v_j)\right)\right),\end{aligned}$$

since $\chi_{\Gamma(S)}(v_j)\sum_{v_i\in S} X_{ij} = \sum_{v_i\in S} X_{ij}$.

For convenience, we denote the following terms for each vertex $v_j \in S^C$:

$$\begin{aligned}D_j &= \sum_{v_i\in S} X_{ij} + \sum_{v_k\in S^C}\left(X_{jk}\chi_{\Gamma(S)}(v_j)\right) \\ S_j &= \sum_{v_i\in S} X_{ij} \\ S_j^C &= \sum_{v_k\in S^C} X_{jk} \\ D_j &= S_j + \chi_{\Gamma(S)}(v_j)S_j^C.\end{aligned}$$

We note that all variables S_j are mutually independent. The variables S_j^C, however, are not independent.

We then compute the covariance of these variables using Lemma 4.2.6 with $X_1 = S_j$, $X_2 = S_k$, $Y_1 = S_j^C$, $Y_2 = S_k^C$, $Z_1 = \chi_{\Gamma(S)}(v_j)$, and $Z_2 = \chi_{\Gamma(S)}(v_k)$. Thus we have the following:

$$\mathrm{Cov}(D_j, D_k) = E[\chi_{\Gamma(S)}(v_k)\chi_{\Gamma(S)}(v_j)]\mathrm{Cov}(S_j^C, S_k^C),$$

where

$$\begin{aligned}
\mathrm{Cov}(S_j^C, S_k^C) &= \left(E\left[\left(\sum_{v_m \in S^C} X_{jm} \right) \left(\sum_{v_m \in S^C} X_{km} \right) \right] \right) \\
&\quad - \left(E\left[\sum_{v_m \in S^C} X_{jm} \right] E\left[\sum_{v_m \in S^C} X_{km} \right] \right) \\
&= \left(E\left[\left(\sum_{v_m \in S^C - \{v_k\}} X_{jm} + X_{jk} \right) \left(\sum_{v_m \in S^C - \{v_j\}} X_{km} + X_{jk} \right) \right] \right) \\
&\quad - \left(E\left[\sum_{v_m \in S^C - \{v_k\}} X_{jm} + X_{jk} \right] E\left[\sum_{v_m \in S^C - \{v_j\}} X_{km} + X_{jk} \right] \right).
\end{aligned}$$

We can again use Lemma 4.2.6 to reduce this by setting $X_1 = \sum_{v_m \in S^C - \{v_k\}} X_{jm}$, $X_2 = \sum_{v_m \in S^C - \{v_j\}} X_{km}$, and $Y_1 = Y_2 = X_{jk}$:

$$\begin{aligned}
\mathrm{Cov}(S_j^C, S_k^C) &= \mathrm{Cov}(X_{jk}, X_{jk}) \\
&= \mathrm{Var}(X_{jk}).
\end{aligned}$$

Thus, $\mathrm{Cov}(D_j, D_k) = E[\chi_{\Gamma(S)}(v_k)\chi_{\Gamma(S)}(v_j)]\mathrm{Var}(X_{jk})$. Lemma 4.2.1 then tells us that $E[\chi_{\Gamma(S)}(v_j)] = \Pr[v_j \in \Gamma(S)] \leq \widehat{\mathrm{Vol}}(S)u_k\rho$. So

$$\mathrm{Cov}(D_j, D_k) \leq \widehat{\mathrm{Vol}}^2(S)u_j u_k \rho^2(u_j u_k \rho - u_j^2 u_k^2 \rho^2).$$

Then $\mathrm{Var}(D_j + D_k) \leq \mathrm{Var}(D_j) + \mathrm{Var}(D_k) + 2\widehat{\mathrm{Vol}}^2(S)\rho(p_{jk}^2 - p_{jk}^3)$.

Since S_j is a sum of independent random variables,

$$\begin{aligned}
\mathrm{Var}(S_j) &= \sum_{v_i \in S} \mathrm{Var}(X_{ij}) \\
&= \sum_{v_i \in S} u_i u_j \rho - u_i^2 u_j^2 \rho^2, \text{ and} \\
\mathrm{Var}(S_j^C) &= \sum_{v_k \in S^C} \mathrm{Var}(X_{jk}) \\
&= \sum_{v_k \in S^C} u_j u_k \rho - u_j^2 u_k^2 \rho^2.
\end{aligned}$$

We can also bound $\mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)(S_j^C)\right)$:

$$\begin{aligned}
\mathrm{Var}(\chi_{\Gamma(S)}(v_j)(S_j^C)) &= E\left[\chi_{\Gamma(S)}(v_j)^2 (S_j^C)^2\right] - \left(E\left[\chi_{\Gamma(S)}(v_j)(S_j^C)\right]\right)^2 \\
&= \left(E\left[\chi_{\Gamma(S)}(v_j)^2\right] - \left(E\left[\chi_{\Gamma(S)}(v_j)\right]\right)^2\right)\left(E\left[(S_j^C)^2\right] - \left(E\left[S_j^C\right]\right)^2\right) \\
&\quad + E\left[\left(\chi_{\Gamma(S)}(v_j)\right)^2\right]\left(E[S_j^C]\right)^2 + \left(E\left[\chi_{\Gamma(S)}(v_j)\right]\right)^2 E\left[\left(S_j^C\right)^2\right] \\
&\quad - 2E\left[\left(\chi_{\Gamma(S)}(v_j)\right)^2\right] E\left[\left(S_j^C\right)^2\right] \\
&= \mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)\right)\mathrm{Var}(S_j^C) \\
&\quad + E\left[\chi_{\Gamma(S)}(v_j)\right]\left(\left(E[S_j^C]\right)^2 + E\left[\chi_{\Gamma(S)}(v_j)\right] E[S_j^C] - 2E[(S_j^C)^2]\right) \\
&= \mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)\right)\mathrm{Var}(S_j^C) \\
&\quad + E\left[\chi_{\Gamma(S)}(v_j)\right]\left(\mathrm{Var}(S_j^C) - \left(E[(S_j^C)^2] - E[S_j^C]\right)\right).
\end{aligned}$$

If $E[S_j^C] \geq 1$, $\mathrm{Var}(S_j^C) \leq \left(E[(S_j^C)^2] - E[S_j^C]\right)$.

Thus, $\mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)(S_j^C)\right) \leq \mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)\right)\mathrm{Var}\left(S_j^C\right)$.

Also, since $\chi_{\Gamma(S)}(v_j)$ is a Bernoulli random variable,

$$\mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)\right) \leq E[\chi_{\Gamma(S)}(v_j)].$$

We can then bound $\mathrm{Var}(D_j)$ as follows:

$$\begin{aligned}
\mathrm{Var}(D_j) &= \mathrm{Var}(S_j) + \mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)(S_j^C)\right) \\
&\leq \mathrm{Var}(S_j) + \mathrm{Var}\left(\chi_{\Gamma(S)}(v_j)\right)\mathrm{Var}\left((S_j^C)\right) \\
&\leq u_j^2\rho^2\widehat{\mathrm{Vol}}^2(S) - u_j^2\rho^2\widehat{\mathrm{Vol}}_2(S) \\
&\quad + u_j\rho\widehat{\mathrm{Vol}}(S)(u_j^2\rho^2\widehat{\mathrm{Vol}}^2(S^C) - u_j^2\rho^2\widehat{\mathrm{Vol}}_2(S^C)).
\end{aligned}$$

Then

$$\begin{aligned}
\mathrm{Var}(\mathrm{vol}(\Gamma(S))) &= \sum_{v_j\in S^C}\mathrm{Var}(D_j) + \sum_{v_j\neq v_k\in S^C}\mathrm{Cov}(D_j,D_k) \\
&\leq \sum_{j\in S^C}\mathrm{Var}(D_j) + 2\widehat{\mathrm{Vol}}^2(S)\rho\left(\sum_{v_j\neq v_k\in S^C}p_{jk}^2 - \sum_{v_j\neq v_k\in S^C}p_{jk}^3\right) \\
&\leq \sum_{j\in S^C}\mathrm{Var}(D_j) + \frac{\widehat{\mathrm{Vol}}(S)^2\rho^2}{2}\left(E[e(S^C,S^C)]\right),
\end{aligned}$$

since $(p_{jk}^2 - p_{jk}^3) \leq \frac{1}{4}p_{jk}$.

Thus

$$\begin{aligned}
\mathrm{Var}(\mathrm{vol}(\Gamma(S))) &= \widehat{\mathrm{Vol}}_2(S^C)\rho^2\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S^C)\rho^2\widehat{\mathrm{Vol}}_2(S) \\
&\quad + \widehat{\mathrm{Vol}}_3(S^C)\rho^3\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}^2(S^C) \\
&\quad - \widehat{\mathrm{Vol}}_3(S^C)\rho^3\widehat{\mathrm{Vol}}(S)\widehat{\mathrm{Vol}}_2(S^C) + \frac{\widehat{\mathrm{Vol}}^2(S)\rho^2}{2}\left(E[e(S^C,S^C)]\right) \\
&= \left(\widehat{\mathrm{Vol}}_2(S^C)\rho^2 + \widehat{\mathrm{Vol}}_3(S^C)\widehat{\mathrm{Vol}}(S)\rho^3\right)\left(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)\right) \\
&\quad + \frac{\widehat{\mathrm{Vol}}^2(S)}{2\widehat{\mathrm{Vol}}^2(G)}\left(E[e(S^C,S^C)]\right) \\
&\leq 2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + \frac{\widehat{\mathrm{Vol}}^2(S)}{2\widehat{\mathrm{Vol}}^2(G)}\left(E[e(S^C,S^C)]\right),
\end{aligned}$$

since we assume that $\widehat{\mathrm{Vol}}(S) \leq (\widehat{\mathrm{Vol}}(G))^{3/4}$ and $\widehat{\mathrm{Vol}}(S) \geq 2$.

Thus

$$\begin{aligned}\mathrm{Var}(\mathrm{vol}(\Gamma(S))) &\leq 2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) \\ &\quad + \frac{\widehat{\mathrm{Vol}}^2(S)}{2\widehat{\mathrm{Vol}}^2(G)}\left(\frac{\widehat{\mathrm{Vol}}^2(S^C) - \widehat{\mathrm{Vol}}_2(S^C)}{\widehat{\mathrm{Vol}}(G)}\right) \\ &\leq 2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + \frac{\widehat{\mathrm{Vol}}^2(S)}{2(\widehat{\mathrm{Vol}}(G))}\left(\frac{\widehat{\mathrm{Vol}}^2(G) - \widehat{\mathrm{Vol}}_2(G)}{\widehat{\mathrm{Vol}}^2(G)}\right) \\ &\leq 2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + \frac{\widehat{\mathrm{Vol}}^2(S)}{2(\widehat{\mathrm{Vol}}(G))}\left(1 - \frac{\widehat{\mathrm{Vol}}_2(G)}{\widehat{\mathrm{Vol}}^2(G)}\right).\end{aligned}$$

So if $2 \leq \widehat{\mathrm{Vol}}(S) \leq \max\{\sqrt{2\widehat{\mathrm{Vol}}(G)}, (\widehat{\mathrm{Vol}}(G))^{3/4}\}$,

$$\mathrm{Var}(\mathrm{vol}(\Gamma(S))) \leq 2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + 1.$$

Therefore, by Chebyshev's Theorem, for all $c > 1$,

$$\Pr\left[|\mathrm{vol}(\Gamma(S)) - \mathrm{Vol}(\Gamma(S))| > c\sqrt{2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + 1}\right] < \frac{1}{c^2}.$$

Thus by Lemma 4.2.4, with probability at least $1 - \frac{1}{c^2}$,

$$\mathrm{vol}(\Gamma(S)) > \mathrm{Vol}_2(S^C)\mathrm{Vol}(S)\rho - c\sqrt{2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + 1)} \text{ and}$$

$$\begin{aligned}\mathrm{vol}(\Gamma(S)) &< \rho\left(\mathrm{Vol}(S) + |S|\right)\left(\mathrm{Vol}_2(S^C) + \mathrm{Vol}(S^C)\right) \\ &\quad + c\sqrt{2(\widehat{\mathrm{Vol}}^2(S) - \widehat{\mathrm{Vol}}_2(S)) + 1}.\end{aligned}$$

□

In [17], mention is made that estimates for neighborhood growth of $G(\mathbf{w})$ are only valid when $\mathrm{Vol}(S) = \mathrm{o}(\mathrm{Vol}(G)/w_{max})$. The obvious question is what happens when $\mathrm{Vol}(S)$ is comparatively large? Unfortunately, the best I can do at this time is a lower bound for the probability that a vertex is in the neighborhood of some set: $\Pr[v_j \in \Gamma(S)] \geq 1 - \exp(-\mathrm{Vol}(S)w_j\rho)$.

4.3 Special case: the Erdös-Rènyi graph model

The random graph family $\widehat{G}(n, \mathbf{u})$, where $u_i = np$ for all $i \leq n$, is precisely the Erdös-Rènyi graph model $G_{n,p}$. Thus we can obtain estimates of probability bounds for the $G_{n,p}$ using the results of the previous section. It should be noted that the number of edges incident to any set of vertices $S \subseteq V(G)$ is a binomial random variable and thus can be computed exactly. B. Bollobás in [15] provides proofs that the distributions of degrees and volumes tend toward Poisson as n becomes very large. The following results regarding the volumes of subsets, being estimates of Poisson random variables, are not the best available probability bounds, but merely illustrations of the preceding results.

Lemma 4.3.1. *(Analog of [17, Lemma 5.6]) For $G \sim G_{n,p}$, and for all $c > 0$, with probability at least $1 - \exp(-c^2/2)$, the degree d_i of vertex $v_i \in V(G)$ satisfies $d_i > (n-1)p - c\sqrt{(n-1)p}$. With probability at least $1 - \exp\left(-\frac{c^2}{2(1+c/(3+n\sqrt{p/(n-1)})}\right)$, d_i satisfies $d_i < \frac{n^2p}{n-1} + c\sqrt{(n-1)p}$.*

Lemma 4.3.2. *(Analog of [17, Lemma 5.9]) For $G \sim G_{n,p}$ and for $0 < c \leq \sqrt{|S|(n-1)p}$, with probability at least $1 - 2\exp(-3c^2/7)$, the volume of a subset S of $V(G)$ satisfies $|\mathrm{vol}(S) - |S|(n-1)p| < c\sqrt{|S|(n-1)p}$.*

Proof. Apply Lemma 3.1.2 with $\mathrm{Vol}(S) = |S|(n-1)p$. □

Corollary 4.3.3. *(Analog of [17, Lemma 5.8]) For $G \sim G_{n,p}$ and for $0 < c \leq \sqrt{n(n-1)p}$, with probability at least $1 - 2\exp(-c^2/6)$, the volume of of G satisfies $|\mathrm{vol}(G) - n(n-1)p| < c\sqrt{n(n-1)p}$.*

Proof. Apply Corollary 3.1.3 with $\mathrm{Vol}(G) = n(n-1)p$. □

The following results relate the number of edges which have an endpoint in a given set to characteristics of that set.

Lemma 4.3.4. *(Analog of [17, Lemma 5.12]) For $G \sim G_{n,p}$ and for any two disjoint subsets S and $T \subseteq V(G)$, $\Pr[\Gamma(S) \cap T \neq \emptyset] < \exp(-|S||T|p)$.*

Proof. This is a special case of Lemma 4.1.3 with $u_i u_j \rho = p$ for all $v_i \in S$ and $v_j \in T$. □

Lemma 4.3.5. *For $G \sim G_{n,p}$ and for any two subsets S and T of $V(G)$ and for $0 < c \leq \sqrt{|S||T|p}$, $e(S,T)$ satisfies*

$$\Pr\left[|e(S,T) - (|S||T|p - |S \cap T|p)| > c\sqrt{|S||T|p}\right] \leq 2e^{\frac{-3c^2}{8}}.$$

Proof. Apply Lemma 4.1.5 with $\widehat{\mathrm{Vol}}(S) = |S|np$, $\mathrm{Vol}(T) = |T|np$ and $\widehat{\mathrm{Vol}}(S \cap T) = |S \cap T|p$. □

Theorem 4.3.6. *For $G \sim G_{n,p}$ and $S \subseteq V(G)$ and for all $0 < c \leq \sqrt{\frac{n^2 p}{n-1}}$, with probability at least $1 - 2\exp(-3c/7)$,*

$$\left|e(S,S) - \frac{1}{2}|S|p(|S|-1)\right| < \frac{c}{2}\sqrt{|S|(|S|-1)p}.$$

Proof. As with Lemma 4.1.4

$$\begin{aligned} E[e(S,S)] &= \sum_{i<j, v_i \in S, v_j \in S} p \\ &= \frac{1}{2}\sum_{v_i, v_j \in S} p - \sum_{v_i \in S} p \\ &= \frac{1}{2}|S|(|S|-1)p. \end{aligned}$$

We then apply Theorem A-1 using $\lambda = c\sqrt{E[e(S,S)]}$:

$$\begin{aligned}
\Pr[|e(S,S) - E[e(S,S)]| > c\sqrt{E[e(S,S)]}] & \\
&\leq \exp\left(-\frac{c\sqrt{(E[e(S,S)])^2}}{2E[e(S,S)] + c\sqrt{E[e(S,S)]}/3}\right) \\
&= \exp\left(-\frac{3c^2 E[e(S,S)]}{6E[e(S,S)] + c\sqrt{E[e(S,S)]}}\right) \\
&= \exp\left(-\frac{3c^2}{6 + c\sqrt{E[e(S,S)]}}\right) \\
&\leq \exp\left(-\frac{3c}{7}\right).
\end{aligned}$$

□

Lemma 4.3.7. *For $G \sim G_{n,p}$, for a given set $S \subset V(G)$ and a vertex $v_j \notin S$, let $t = \min\{x, |S|\}$. Then*

$$\Pr[d_j = x | v_j \in \Gamma(S)] = \frac{\binom{|S|}{r}\binom{n-1-|S|}{x-r}p^x(1-p)^{n-1-x}}{1-(1-p)^{|S|}}.$$

Proof. We begin by using Bayes' Theorem and then substitute in the appropriate values, taking advantage of the fact that $p_{ij} = p$ for all i and j:

$$\begin{aligned}
\Pr[d_j = x | v_j \in \Gamma(S)] &= \frac{\Pr[(d_j = x) \cap (v_j \in \Gamma(S)]}{\Pr[v_j \in \Gamma(S)]} \\
&= \frac{\sum_{r=1}^{t}\binom{|S|}{r}p^r(1-p)^{t-r}\binom{n-1-|S|}{x-r}p^{x-r}(1-p)^{n-1-t+r}}{1-(1-p)^{|S|}} \\
&= \frac{\binom{|S|}{r}\binom{n-1-|S|}{x-r}p^x(1-p)^{n-1-x}}{1-(1-p)^{|S|}}.
\end{aligned}$$

□

Lemma 4.3.8. *For $G \sim G_{n,p}$ and for a given set $S \subset V(G)$,*

$$E(|\Gamma(S)|) = n - |S| - \frac{1}{(1-p)^{|S|}}.$$

Proof. The number of vertices which are not adjacent to $|S|$ is given approximately by a Poisson random variable with $\lambda = \frac{1}{(1-p)^{|S|}}$. □

Lemma 4.3.9. *(Analog of Lemma 5.14 in [17]) For $G \sim G_{n,p}$ and for $S \subset V(G)$, $\mathrm{Vol}(\Gamma(S)) = (n-1)\left(n - |S| - \frac{1}{(1-p)^{|S|}}\right)$*

Proof. This is a special case of Lemma 4.2.7 with $u_i u_j \rho = p$ for all $v_i, v_j \in V(G)$. □

Chapter 5

RESULTS FOR EDGE-PRODUCT MODELS WITH IDENTICAL DEGREE DISTRIBUTION

Two edge-product models that yield nicely to analysis by means of their probability matrices are a matrix-based small world (hereafter referred to as MSW) model and a hierarchical model.

The MSW model began as an edge-product approximation to the Watts-Strogatz and Newman-Watts models. The implementation was inspired by D. Higham's 2003 paper regarding small perturbations in a transition matrix for a Markov process [28]. The primary difference between this model and Higham's is that Higham's model does not permanently establish an "edge", but rather allows for a small probability of movement from one state to a random state at each timestep.

Much work has been done on the detection of communities within networks. M. Newman noted the community structure present in social, biological, and information networks before providing an algorithm for dividing networks into communities using spectral techniques [38]. Distinct communities have been found in examination of food webs, friendship networks, and collaboration networks [26]. R. Guimera *et. al.* found distinct communities within a corporation using email messages as a proxy [27]. Thus

a random network model that incorporates such structure may be useful. The hierarchical model creates these communities using different probabilities for adjacency based on whether two vertices are in a common block.

5.1 The matrix-based small world and hierarchical models

Like the Watts-Strogatz and Newman-Watts models, the MSW model starts with a one-dimensional lattice (see Figure 2.1) with n vertices of radius r and assigns a certain probability α to the existence of edges that are contained within the lattice. The remaining vertex pairs are adjacent with probability $\beta < \alpha$. If $\alpha = 1$, this model generates a Newman-Watts small world graph without multiple edges. The goal of this model was to generate an edge-product analog of the Watts-Strogatz small world model.

If $\beta = \frac{2r(1-\alpha)}{n-2r-1}$, then the MSW model will generate a graph with a similar degree distribution to that generated by the Watts-Strogatz model with probability of rewiring $1-\alpha$. These particular parameters are a result of keeping the underlying lattice structure and mean degree. The minimum degree for any vertex in the Watts-Strogatz model with an average of $2r$ neighbors is r [12]. This is a direct result of the algorithm fixing one endpoint of each edge before rewiring. As n increases, the degree distribution for the Watts-Strogatz small world model approachs $r + \text{Binom}(\alpha, r) + \text{Poisson}((n-2r-1)(1-\alpha))$. The degree distribution for the MSW model as n increases approaches $\text{Binom}(\alpha, 2r) + \text{Poisson}((n-2r-1)(\beta))$. The degree distributions for these two models where $\beta = \frac{2r(1-\alpha)}{n-2r-1}$ differ then by $\text{Binom}(\alpha, r) - r$. The mean degree for both models is $2r$ and as r increases the degree distributions of both models approach a Poisson distribution with mean $2r$.

$$\begin{pmatrix}
0 & \alpha & \alpha & \alpha & \beta & \beta & \beta & \dots & \alpha & \alpha \\
\alpha & 0 & \alpha & \alpha & \alpha & \beta & \beta & \dots & \alpha & \alpha \\
\alpha & \alpha & 0 & \alpha & \alpha & \alpha & \beta & \dots & \beta & \alpha \\
\alpha & \alpha & \alpha & 0 & \alpha & \alpha & \alpha & \dots & \beta & \beta \\
\beta & \alpha & \alpha & \alpha & 0 & \alpha & \alpha & \dots & \beta & \beta \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\
\alpha & \alpha & \beta & \beta & \beta & \beta & \beta & \dots & 0 & \alpha \\
\alpha & \alpha & \alpha & \beta & \beta & \beta & \beta & \dots & \alpha & 0
\end{pmatrix}$$

Figure 5.1: The probability matrix for a MSW random graph model with radius 3

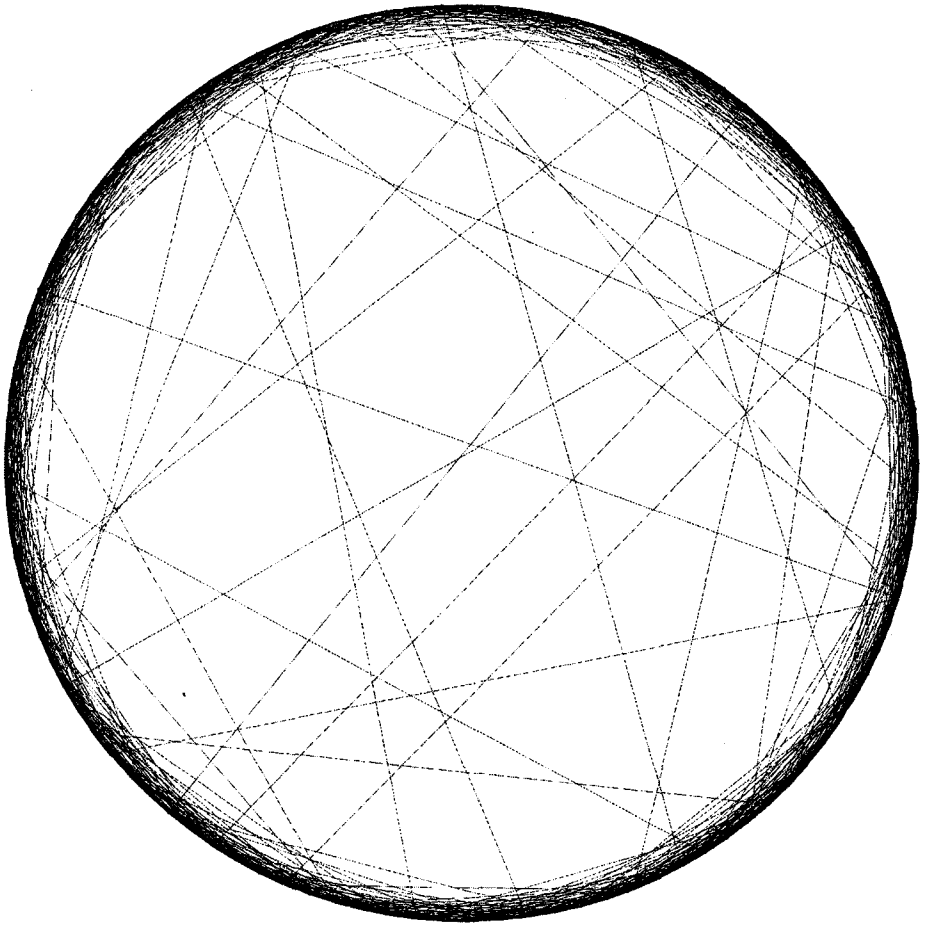

Figure 5.2: A realization of a MSW random graph model with 105 vertices, radius 10, $\alpha = 0.9$, $\beta = 0.01$.

$$\begin{pmatrix}
0 & \alpha & \alpha & \alpha & \alpha & \beta & \beta & \dots & \beta & \beta \\
\alpha & 0 & \alpha & \alpha & \alpha & \beta & \beta & \dots & \beta & \beta \\
\alpha & \alpha & 0 & \alpha & \alpha & \beta & \beta & \dots & \beta & \beta \\
\alpha & \alpha & \alpha & 0 & \alpha & \beta & \beta & \dots & \beta & \beta \\
\alpha & \alpha & \alpha & \alpha & 0 & \beta & \beta & \dots & \beta & \beta \\
\beta & \beta & \beta & \beta & \beta & 0 & \alpha & \dots & \beta & \beta \\
\beta & \beta & \beta & \beta & \beta & \alpha & 0 & \dots & \beta & \beta \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\
\beta & \beta & \beta & \beta & \beta & \beta & \beta & \dots & 0 & \alpha \\
\beta & \beta & \beta & \beta & \beta & \beta & \beta & \dots & \alpha & 0
\end{pmatrix}$$

Figure 5.3: The probability matrix for a hierarchical random graph model with cluster size 5

The hierarchical model begins with a set of clusters of vertices for which the probability of adjacency between any two vertices in the same cluster is given by α and the probability of adjacency between any two vertices in different clusters is given by $\beta < \alpha$. This model attempts to build communities within a random graph whose members are very likely to be adjacent to each other, but with weak connections between communities.

5.2 Spectra of graphs with identical degree distribution

From [19], we know that for any regular graph of degree d, $\lambda_1(A) = d$ and that the associated eigenvector is $\mathbf{1}$. Since this is true for a one-dimensional lattice with no additional edges as well as for random regular graphs, clustering or other structure within a graph appears not to have a significant effect on the spectral radius of the adjacency matrix. Recall that for any graph G, $\lambda_1(A)$ satisfies the following [29][19]:

$$\lambda_1(A) \leq \max_i d_i \tag{5.2.1}$$

$$\lambda_1(A) \geq \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i^2}. \tag{5.2.2}$$

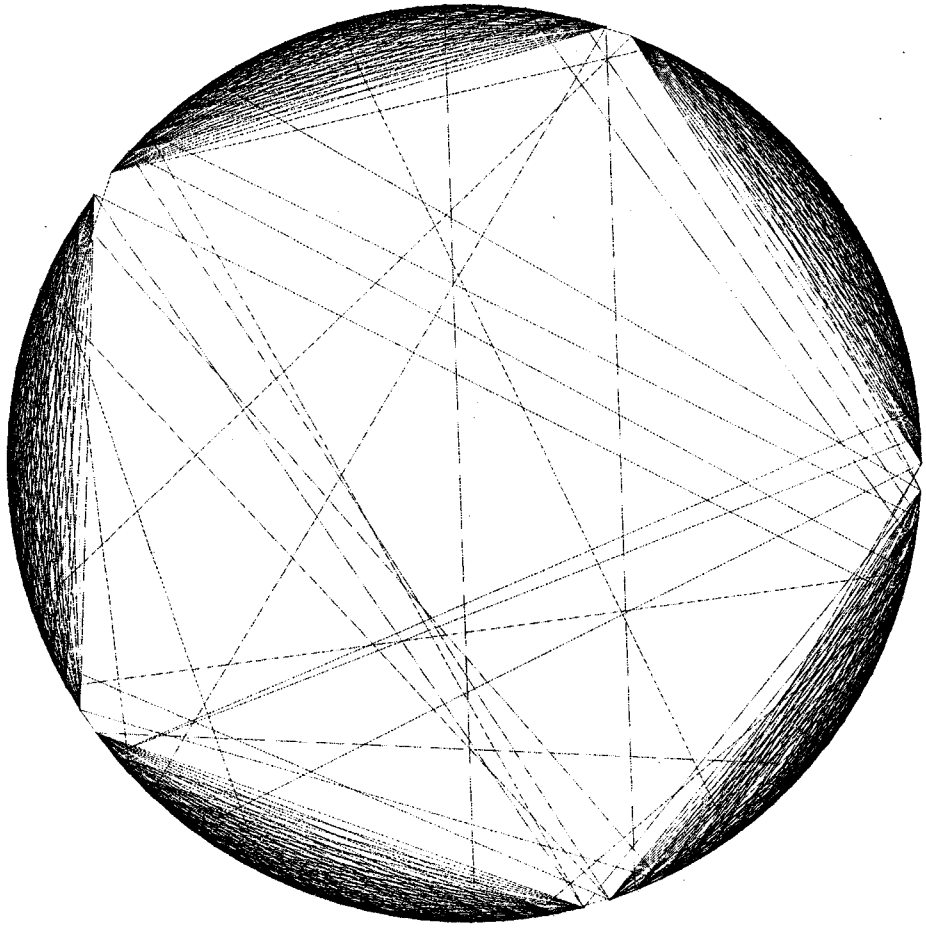

Figure 5.4: A realization of a hierarchical random graph model with cluster size 21, $\alpha = 0.9$, $\beta = 0.01$. Note the five clusters.

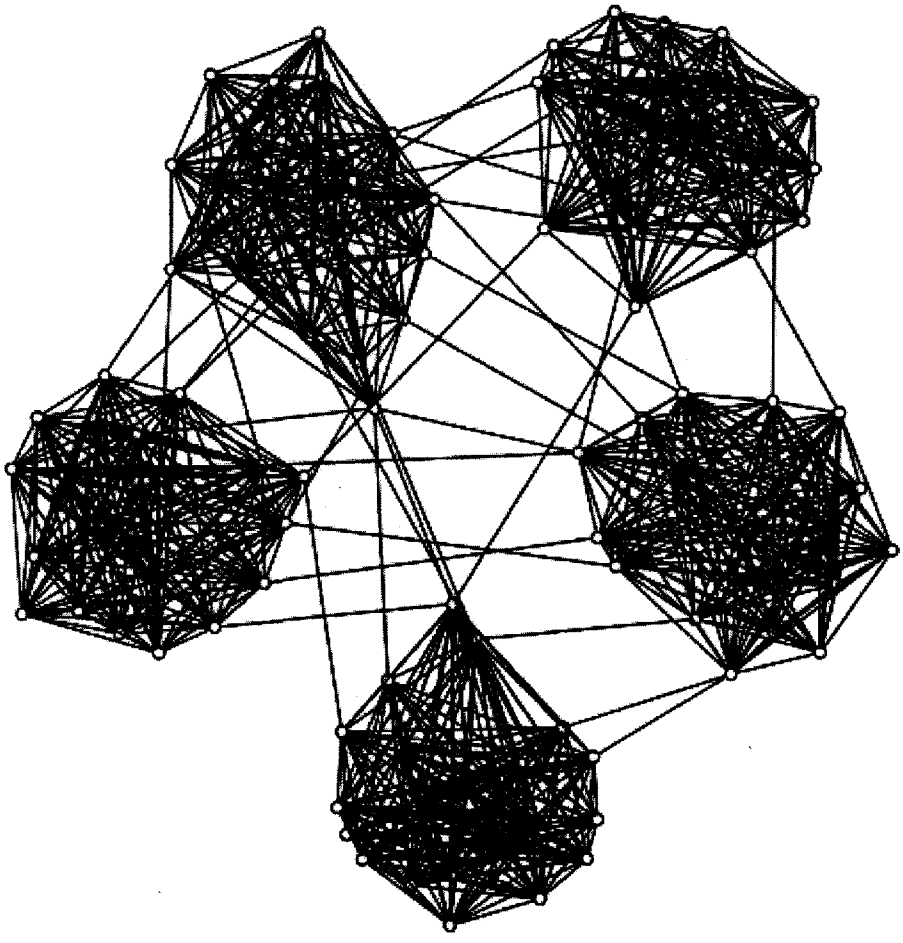

Figure 5.5: A different view of the graph in Figure 5.4. The five clusters are more clearly seen here.

Let P be the probability matrix for a family of random graphs in which each vertex has the same expected degree d. This matrix is positive semi-definite and irreducible. Since the vertices have the same expected degree, $P\mathbf{1} = d\mathbf{1}$. By the Perron-Frobenius theorem only the dominant eigenvalue has a corresponding eigenvector consisting entirely of nonnegative entries. Thus $\lambda_1(P) = d$.

Lemma 5.2.1. *Let $G(P)$ be a family of random graphs with the probability matrix P in which every vertex has the same degree distribution if $\Pr[d_j > t] = p$ for each vertex v_j, then*

$$\Pr[\max_j\{d_j\} > t] < 1 - p^n.$$

Likewise if $\Pr[d_j < t] = q$, then

$$\Pr[\min_j\{d_j\} < t] < 1 - q^n.$$

Proof. Consider two vertices, v_i and v_j. By definition, d_i and d_j are sums of independent Bernoulli random variables. There is one random variable that is both of these sums, a_{ij}. Calculating the covariance, we then find that $\mathrm{Cov}[d_i, d_j] = p_{ij}$, indicating that d_i and d_j are positively correlated. Thus for all t,

$$\Pr[d_j > t | d_i \leq t] \leq \Pr[d_j > t] \text{ and} \tag{5.2.3}$$

$$\Pr[d_j < t | d_i \geq t] \leq \Pr[d_j < t]. \tag{5.2.4}$$

□

Lemma 5.2.2. *For a family of random graphs in which every vertex has the same expected degree d,*

$$\Pr[\lambda_1(A) \leq d - t] \leq \exp(-t^2/2d).$$

Proof. From Inequality (5.2.1), $\lambda_1(A) \geq \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i^2}$. Also, since $\sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i^2} \geq \frac{1}{n}\sum_{i=1}^{n} d_i$, $\lambda_1(A) \geq d$. We can then use Chernoff's inequality, Theorem (A-2) with $\nu = d$ and find

$$\Pr[X \leq d - t] \leq \exp(-t^2/2d). \tag{5.2.5}$$

□

Lemma 5.2.3. *For a family of random graphs in which every vertex has the same degree distribution with expected degree d and variance* (σ^2),

$$\Pr[\lambda_1(A) < d + t] > (1 - \exp(-t^2/2(\sigma^2 + t/3))^n.$$

Proof. By Perron's Theorem, $||A|| \leq \max_i \sum_{j=1}^{n} a_{ij}$. Let $d_i = \sum_{j=1}^{n} a_{ij}$. Then $E[d_i] = d$ and $\mathrm{Var}[d_i] = \sum_{j=1}^{n} \mathrm{Var}[a_{ij}]$. We use Theorem A-3 to find $\Pr[d_i > d + t] \leq \exp\left(\frac{-t^2}{2(\mathrm{Var}[d_i]+t/3)}\right)$. Recall that $\lambda_1(A) \leq \max_i\{d_i\}$. By Inequality 5.2.3, $\Pr[d_i > a | d_j < a] < \Pr[d_i > a]$ for all a. Thus

$$\Pr[\lambda_1(A) > d + t] < 1 - (1 - \exp\left(-t^2/2(\mathrm{Var}[d_i] + t/3)\right))^n. \tag{5.2.6}$$

□

For a random graph model in which each vertex degree has the same distribution with $E[d_i] = d$ and $\mathrm{Var}[d_i] = \sigma^2$, let

$$\gamma_1 = 1.177\sqrt{d} \tag{5.2.7}$$

$$\gamma_2 = \sigma^2 \left(\sqrt{1/9 - \frac{2\ln(1 - (1/2)^{(1/n)})}{\sigma^2}} \right). \tag{5.2.8}$$

Note that these are the values of t if we set the right-hand sides of inequalities (5.2.5) and (5.2.6) to 1/2. Thus we are able to bound the median value of d.

Then we can use equation (3.2.2) to give possibly tighter bounds for the largest eigenvalue of A. Let $\gamma_1 = 1.177\sqrt{d}$. Using the values of γ_1 and γ_2. Since $\Pr[|\lambda - \lambda_{\text{med}}| > t] < 4e^{-t^2/8}$,
$\Pr[\lambda \notin (d - \gamma_1 - 3.59, d + \gamma_2 + 3.59)] < .05$. It should be noted that the variances of the random variables here only affect the value of γ_i.

Looking at the possibility of an epidemic on this network, we can then pair this result with Theorem 1.3.1 and find that if the probability of success for each attempted infection β satisfies $\beta < \frac{1}{d+\gamma_2+3.59}$, then the probability that an epidemic will occur is less than 2.5%.

We now look at the smaller eigenvalues of A. We start by finding $\lambda_i(P - A)$ in order to be able to use Theorem 3.2.1. In the case where each row of P includes $s - 1$ entries having the value α and $n - s - 1$ entries having the value β, one quantity tends to arise frequently. Let

$$\phi = 2\sqrt{n(s-1)\alpha(1-\alpha)} + 2\sqrt{n\beta(1-\beta)} \tag{5.2.9}$$

Lemma 5.2.4. *Let P be a symmetric probability matrix with $s - 1$ entries in each row having the value α and $n - s - 1$ entries in each row having the value β. Let ϕ be defined as in equation 5.2.9. For $G \sim G(P)$,*

$$\begin{aligned}\Pr[\lambda_1(A - P) > \phi + 2c] \\ < \max\left\{\exp\left(-\frac{c^4}{2(n(s-1)\alpha(1-\alpha) + c^2/3)}\right),\right. \\ \left.\exp\left(-\frac{c}{2n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right)\right\}.\end{aligned}$$

Proof. Let A_α be the adjacency matrix of edges which exist with probability α. Let P_α be the probability matrix of edges which exist with probability α. We then find an upper bound for $\lambda_1(A_\alpha - P_\alpha)$ as follows:

$$\lambda_1(A_\alpha - P_\alpha \leq \sqrt{\text{Tr}(A_\alpha - P_\alpha)^2}).$$

Algebra tells us that $E[\mathrm{Tr}(A_\alpha - P_\alpha)^2] = n(s-1)\alpha(1-\alpha)$. We can then find an upper bound for $\mathrm{Tr}(A_\alpha - P_\alpha)^2$ using Theorem A-1, giving us

$$\Pr[\mathrm{Tr}(A_\alpha - P_\alpha)^2 - n(s-1)\alpha(1-\alpha) > c] \leq \exp\left(-\frac{c^2}{2(n(s-1)\alpha(1-\alpha) + c/3)}\right).$$

Since $(E[\lambda_1(A_\alpha - P_\alpha])^2 \leq E[(\lambda_1(A_\alpha - P_\alpha)^2]$ and $a^2 + b^2 < (a+b)^2$ for all $a, b > 0$, we then have

$$\begin{aligned}\Pr[\lambda_1(A_\alpha - P_\alpha) > \sqrt{n(s-1)\alpha(1-\alpha)} + c] \\ < \exp\left(-\frac{c^4}{2(n(s-1)\alpha(1-\alpha) + c^2/3)}\right).\end{aligned}$$

Defining A_β and P_β in the same way as A_α and P_α,

$$\Pr[\lambda_1(A_\beta - P_\beta) > 2\sqrt{\beta(1-\beta)} + c] < \sqrt{n}\exp\left(\frac{c}{2n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right).$$

Then, since

$$\Pr[A + B - E[A+B] > 2t] < \max\left\{\Pr[A - E[A] > t], \Pr[B - E[B] > t]\right\},$$

we have our result. □

In most cases, including any time $n^{1/3} > c^3$, $\left(\frac{c^4}{2(n(s-1)\alpha(1-\alpha)+c^2/3)}\right) > \left(\frac{c}{2n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right)$. We will thus use the estimate

$$\Pr[\lambda_1(A - P) > \phi + 2c] < \exp\left(-\frac{c}{2n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right). \qquad (5.2.10)$$

Lemma 5.2.4, together with Theorem 3.2.1 then gives bounds for $\lambda_i(A)$.

We now look for way to bound $\lambda_2(M) = \lambda_2(D^{-1}A)$. We rely on the following theorem from A.R. Amir-Moéz to show that $\lambda_n(D^{-1})\lambda_i(A) \leq \lambda_i(D^{-1}A) \leq \lambda_1(D^{-1})\lambda_i(A)$.

Theorem 5.2.5. *[6, Special Case of Theorems 3.9 and 3.10] Let S and T be any linear transformations on $\mathbb{R}^n$. Then the following hold:*

1. $\lambda_i(ST) \geq \lambda_n(S)\lambda_i(T)$.

2. $\lambda_i(ST) \leq \lambda_1(S)\lambda_i(T)$.

The next step is determining bounds for $\lambda_1(D^{-1})$ and $\lambda_n(D^{-1})$.

Lemma 5.2.6. *For a family of random graphs in which all vertices have the same degree distribution with $E[d_i] = w$, let $H = wI - D$. Then*

$$\lambda_i(D^{-1}) = \frac{1}{w}\frac{1}{1-\left(\frac{\lambda_i(H)}{w}\right)}.$$

Proof. As defined in the statement, $D = wI - H$ so $D = wI(I-(wI)^{-1}H)$. Recall that the eigenvalues are ordered according to their real value. Thus, since $d_i \geq 0$ for all i, $\lambda_i(D^{-1}) = \frac{1}{\mu_i(D)}$. Thus

$$\begin{aligned}
\lambda_i(D^{-1}) &= \mu_i(wI(I-(wI)^{-1}H))^{-1} \\
&= \frac{1}{w}\mu_i\left(I-\frac{1}{w}H\right)^{-1} \\
&= \frac{1}{w}\left(1+\mu_i\left(-\frac{1}{w}H\right)^{-1}\right) \\
&= \frac{1}{w}\left(1-\lambda_i\left(\frac{1}{w}H\right)^{-1}\right) \\
&= \frac{1}{w}\left(1-\left(\frac{1}{w}\lambda_i(H)\right)^{-1}\right) \\
&= \frac{1}{w}\frac{1}{1-\left(\frac{\lambda_i(H)}{w}\right)}.
\end{aligned}$$

□

We can now bound the largest and smallest eigenvalues of D^{-1}.

Lemma 5.2.7. *For a family of random graphs in which all vertices have the same degree distribution with $E[d_i] = w$,*

$$\Pr[\lambda_1(D^{-1}) \leq \frac{1}{w} + \frac{2t}{w^2}] \geq 1 - n\exp\left(-t^2/2(w+t/3)\right)$$

Proof. Let $s = \max_i\{w - d_i\}$. Then $\lambda_1(E) = s$ and $s = w - \min_i\{d_i\}$. So $1 + \frac{s}{w} \leq \ _1((wI)^{-1}E) \leq 1 + \frac{2s}{w}$. So $\frac{1}{w} + \frac{s}{w^2} \leq \ _1(D^{-1}) \leq \frac{1}{w} + \frac{2s}{w^2}$. Applying Theorem A-1 we find

$$\Pr[d_i > t] < 1 - \exp\left(-t^2/2(w + t/3)\right).$$

We then observe that the bounds provided by independent variables are wider than the bounds of positively correlated variables, so

$$\Pr[d_{\max} > t] < \left(1 - \exp\left(-t^2/2(w + t/3)\right)\right)^n.$$

□

Lemma 5.2.8. *For a family of random graphs in which all vertices have the same degree distribution with* $E[d_i] = w$,

$$\Pr[\ _n(D^{-1}) \geq \frac{1}{w} - \frac{t}{w^2} + \frac{t^2}{w^3}] \geq 1 \quad n\exp(-t^2/2(w + t/3)).$$

Proof. We proceed exactly as in the proof of Lemma 5.2.7 to find $\ _n(D^{-1}) = \frac{1}{w}\sum_{j=0}^{\infty} w^{-j}\ _n(E)^j$. Then let $s = \max_i\{d_i - w\}$. Then $\ _n(E) = -s$ and $s = \max_i\{d_i\} - w$. So $1 - \frac{s}{w} \leq \ _n(W^{-1}E) \leq 1 \quad \frac{s}{w} + \frac{s^2}{w^2}$.
Thus $\frac{1}{w} - \frac{s}{w^2} \leq \ _n(D^{-1}) \leq \frac{1}{w} - \frac{s}{w^2} + \frac{s^2}{w^3}$. The remainder follows exactly as for Lemma 5.2.7.

□

This approach can be used to find the largest and smallest eigenvalues of D^{-1} when the vertices do not all have the same expected degree. It is, however, necessary in this case to multiply the different probabilities that each vertex has degree greater than t to determine $\Pr[\min_i\{d_i\} > t]$.

We are now able to find probability bounds for $\ _i(M)$:

Theorem 5.2.9. *For a family of random graphs in which each vertex is adjacent to $s-1$ vertices with probability α and to $n-s-1$ vertices with probability β and $E[d_i] = w$, let ϕ be defined as in equation (5.2.9). Then for $c < w$, $\lambda_i(M)$ satisfies the following:*

$$\Pr\left[\lambda_i(M) - \frac{\lambda_i(P)}{w} < \frac{\phi + 3c}{w}\right] \leq n\exp(-c^2/2(w+c/3)) + \exp\left(-\frac{c}{2n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right)$$

Proof. By definition, $\lambda_i(M) = \lambda_i(D^{-1}A)$. From Theorem 5.2.5, $\lambda_i(M) \leq \lambda_1(D^{-1})\lambda_i(A)$.

Using Inequality 5.2.10,

$$\Pr[\lambda_i(A) - \lambda_i(P) > \phi + c] < \exp\left(-\frac{c}{4n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right).$$

Using Lemma 5.2.7,

$$\Pr[\lambda_1(D^{-1}) \geq \frac{1}{w} + \frac{2c}{w^2}] < n\exp\left(-c^2/2(w+c/3)\right)$$

Finally, since $\frac{c}{w} < 1$, $\frac{c^2}{w^2} < \frac{c}{w}$.

□

Using this result with Lemma 1.3.2, we have with probability greater than $(1-\exp(-c^2/2(w+c/3)))^n + \exp\left(-\frac{c}{2n^{1/3}\beta^{2/3}(1-\beta)^{2/3}}\right)$,

$$\frac{\mathrm{vol}(\Gamma(S))}{\mathrm{vol}(S)} > \frac{1}{1-(1-\left(\frac{\lambda_i(P)+\phi+3c}{w}\right)^2)\frac{\mathrm{vol}(S^C)}{\mathrm{vol}(G)}}.$$

where $\phi = 2\sqrt{n(s-1)\alpha(1-\alpha)} + 2\sqrt{n\beta(1-\beta)}$.

5.3 A direct computation of neighborhood growth in the MSM and hierarchical models

In general it is not as easy to determine the rate of neighborhood growth for graphs that are distributed according to an MSM or hierarchical model as it is for graphs that are distributed according to a modified Chung-Lu model. The primary difficulty stems from the importance of the individual vertices in relationship to each other. Since a graph distributed according to a modified Chung-Lu model is locally tree-like, we can find the likelihood that any two members of a set $S \subseteq V(G)$ are adjacent to one vertex in S^C. For the MSW and hierarchical models we must consider whether members of a set $S \subseteq V(G)$ are more likely to be adjacent to a single vertex in S^C. In this discussion of neighborhood growth, it is assumed that n is very large compared to the other quantities. Thus we can get a general idea of how neighborhood expansion proceeds before it is limited by the number of vertices.

For $S \subseteq V(G)$, let $\Gamma^i(S)$ be defined as in section 4.2. Let $\Gamma_k(S) = \cup_{i=0}^{k}\Gamma^i(S)$. Let $\gamma_k(S) = |\Gamma_k(S)|$. Let $\Delta_k(S) = \Gamma_k(S) - \Gamma_{k-1}(S)$. Let $\delta_k(S) = |\Delta_k(S)|$. To investigate the growth of $\Gamma_k(S)$ it is necessary to define an induced subgraph.

Definition 5.3.1. *Let $S \subseteq V(G)$. The subgraph of G induced by S, denoted by $G[S]$, is the subgraph of G such that $V(G[S]) = S$ and $E(G[S]) = \{\{v_i, v_j\} \in G | v_i, v_j \in S\}$.*

For a ring lattice L and for nonempty $S \subseteq V(L)$ let $t_L(S)$ be the number of components of $L[S]$. Let $t_L^k(S) = t_L(\Gamma_k(S))$.

As noted in [40], for $G \sim G_{n,p}$ and for small k,

$$\begin{aligned}\gamma_k(S) &\approx |S| \sum_{i=0}^{k} ((n-1)p)^i \\ &\approx |S|((n-1)p)^k.\end{aligned}$$

Lemma 5.3.2. *On a ring lattice L with radius r and $k < \frac{n-|S|}{2rt_L(S)}$, then the following are true:*

1. $2rk + |S| \leq \gamma_k(S) \leq (2rk)t_L(S) + |S|$.

2. *If the connected components of $L[S]$ are distributed uniformly over L,* $\Pr[t_L^k(S) = m] = \sum_{j=2rm}^{2rt_L^{k-1}} \frac{(n-\gamma_{k-1}(S)-j-1)_{(n-\gamma_{k-1}(S)-j-m)}\left(t_L^{k-1}(S)\right)_{(m-1)}}{(n-\gamma_{k-1}(S)-1)_{\left(n-\gamma_{k-1}(S)-t_L^{k-1}(S)\right)}}$.

Proof. The proof of (1) is by induction on k.

As defined, $\Gamma_0(S) = S$, so equality holds. Assume that Lemma 5.3.2 holds for k and that $k+1 < \frac{n-|S|}{2rt(S)}$. Then each component of $L[\Gamma_k(S)]$ expands by at most $2r$, so $\gamma_k + 1(S) \leq \gamma_k(S) + 2rt_L^k(S)$. Expansion on the lattice requires that $1 \leq t_L^{k+1}(S) \leq t_L^k(S)$. Since $k+1 < \frac{n-|S|}{2rt_L(S)}$ and $\gamma_k(S) \leq 2rkt_L(S)$, there are at least $2rt_L(S)$ vertices in $L - L[\Gamma_k(S)]$. Again, because of the the structure of the lattice, at least $2r$ of these vertices are in $\Gamma_{k+1}(S)$. Thus $\gamma_{k+1}(S) - \gamma_k(S) \geq 2r$.

For (2) we must determine the expected rate at which the connected components of $L[\Gamma_k(S)]$ grow together. We will do this by counting the possible ways in which the gaps between the connected components of $L[\Gamma_k(S)]$ are arranged. The number of ways in which $t_L^k(S)$ of these gaps can be arranged is equivalent to the number of ordered partitions of the integer $|V(G) - \Gamma_k(S)|$ into $t_L^k(S)$ parts. So there are $\binom{n-\gamma_{k-1}(S)-1}{t_L^{k-1}(S)-1}$ ways in which these gaps can be arranged. Assuming that each arrangement is equally

likely when a set S is randomly selected, the probability that $t_L^k(S) = m$ can be determined given the values of $t_L^{k-1}(S)$ and $\gamma_{k-1}(S)$.

$$\begin{aligned}\Pr[t_L^k(S) = t_L^{k-1}(S)] &= \frac{\binom{n-\gamma_{k-1}(S)-2rt_L^{k-1}(S)-1}{t_L^{k-1}(S)-1}}{\binom{n-\gamma_{k-1}(S)-1}{t_L^{k-1}(S)-1}} \\ &= \frac{\left(n-\gamma_{k-1}(S)-2rt_L^{k-1}(S)-1\right)!\left(n-\gamma_{k-1}(S)-t_L^{k-1}(S)-1\right)!}{\left(n-\gamma_{k-1}(S)-2rt_L^{k-1}(S)-t_L^{k-1}(S)-1\right)!\left(n-\gamma_{k-1}(S)-1\right)!} \\ &= \frac{\left(n-\gamma_{k-1}(S)-2rt_L^{k-1}(S)-1\right)_{\left(n-\gamma_{k-1}(S)-2rt_L^{k-1}(S)-t_L^{k-1}(S)-1\right)}}{\left(n-\gamma_{k-1}(S)-1\right)_{\left(n-\gamma_{k-1}(S)-t_L^{k-1}(S)-1\right)}},\end{aligned}$$

where $(x)_{(y)} = \frac{x!}{y!}$.

If $t_L^{k-1}(S)$ is small compared to $n-\gamma_{k-1}(S)$, then $t_L^k(S) = t_L^{k-1}(S)$ with high probability.

If $m < t_L^{k-1}(S)$, then we need to consider the size of $\delta_k(S)$. Since we have m partitions, we know that $\delta_k(S) \geq 2rm$. In addition, at least $t_L^{k-1}(S) - m$ vertices must be in $\Delta_k(S)$, since there are $t_L^{k-1}(S) - m$ sections of $V(G) - \Gamma_{k-1}(S)$ that are now in $\Gamma_k(S)$. Thus, there are

$$\sum_{j=2rm}^{2rt_L^{k-1}(S)} \binom{n-\gamma_{k-1}(S)-j-1}{m-1} = \sum_{j=2rm}^{2rt_L^{k-1}} \frac{\left(n-\gamma_{k-1}(S)-j-1\right)_{(n-\gamma_{k-1}(S)-j-m)}}{(m-1)!}$$

ordered partitions of $V(G) - \Gamma_k(G)$ into m parts.

Therefore

$$\begin{aligned}\Pr[t_L^k(S) = m] &= \sum_{j=2rm}^{2rt_L^{k-1}(S)} \frac{\binom{n-\gamma_{k-1}(S)-j-1}{m-1}}{\binom{n-\gamma_{k-1}(S)-1}{t_L^{k-1}(S)-1}} \\ &= \sum_{j=2rm}^{2rt_L^{k-1}} \frac{\left(n-\gamma_{k-1}(S)-j-1\right)_{(n-\gamma_{k-1}(S)-j-m)}\left(t_L^{k-1}(S)\right)_{(m-1)}}{\left(n-\gamma_{k-1}(S)-1\right)_{\left(n-\gamma_{k-1}(S)-t_L^{k-1}(S)\right)}}.\end{aligned}$$

□

Observe that when $t_L^k(S)$ is small in relation to $n - \gamma_k(S)$,

$E[t_L^k(S)] \approx t_L^{k-1}(S)$

From Lemma 5.3.2 we see that neighborhoods on a ring lattice expand at most linearly. The neighborhoods on $G \sim G_{n,p}$, however, expand exponentially. Any small perturbation of a ring lattice into an MSM model, however, changes the rate of neighborhood growth dramatically.

Proposition 5.3.3. *When G is distributed according to an MSW model, $\Gamma_k(S)$ expands exponentially for small k.*

Proof. Since some edges are present in G with probability α and others are present with probability β, it is necessary to consider the two mechanisms of growth separately. Recall that $t_L^k(S)$ is the number of connected components of $L[\Gamma_k(S)]$.

First, consider growth along the edges that are present with probability α. Let $\Delta_\alpha^k(S)$ be $\Delta_k(S) \cap \{v_i \text{ such that } \exists v_j \in L[\Gamma_{k-1}(S)] \text{ where } p_{ij} = \alpha\}$. An upper bound for this growth is given in the proof of Lemma 5.3.2: $E[|\Delta_\alpha^k(S)|] \leq 2rt_L^{k-1}(S)$. If vertex v_i is a potential member of $\Delta_\alpha^k(S)$, then $\Pr[\{v_i, v_j\} \in E(G) \text{ for some } v_j \in \Gamma_{k-1}(S)] \geq \alpha$. Thus a lower bound for this growth is given by $|\Delta_\alpha^k(S)| \geq 2r\alpha E[t_L^{k-1}(S)]$.

Now consider the growth along edges that are present with probability β. For each of the vertices v_i in $\Delta_{k-1}(S)$, we can find the edges which cause $t_L^k(S) > t_L^{k-1}(S)$. Any edges from v_i to $v_j \in \Gamma_{k-1}(S)$ will neither increase $t_L^k(S)$ nor expand $\Gamma_k(S)$. Likewise, any edges from v_i to $v_j \in \Delta_\alpha^k(S)$ will neither increase $t_L^k(S)$ nor expand $\Gamma_k(S)$. Finally, the vertices v_j which are adjacent in L to any vertices in $L[\Gamma_{k-1}(S)]$ will expand $\Gamma_k(S)$ but will not increase $t_L^k(S)$. Thus there are $n - 2(r+1)t_L^{k-1}(S) - \gamma_{k-1}(S)$ vertices which

we must consider. Each of these vertices has probability $(1-\beta)^{\delta_{k-1}(S)}$ of not being in $\Gamma_k(S)$.

So the expected increase by means of edges which are present with probability β is $(n-2(r+1)t_L^{k-1}(S)-\gamma_{k-1}(S))\left(1-(1-\beta)^{\delta_{k-1}(S)}\right)$. Further, these vertices also increase $t_L^k(S)$. So, taking this information together with Lemma 5.3.2,

$$E[t_L^k(S)] \approx (n-2(r+1)t_L^{k-1}(S)-\gamma_{k-1}(S))\delta_{k-1}(S)\beta + \sum_{m=1}^{t_L^{k-1}(S)} m \sum_{j=2rm}^{2rt_L^{k-1}(S)} \frac{(n-\gamma_{k-1}(S)-j-1)_{(n-\gamma_{k-1}(S)-j-m)}\left(t_L^{k-1}(S)\right)_{(m-1)}}{(n-\gamma_{k-1}(S)-1)_{\left(n-\gamma_{k-1}(S)-t_L^{k-1}(S)\right)}}.$$

As in the proof of Lemma 5.3.2, if $t_L^{k-1}(S)$ and $\gamma_k(S)$ are both small relative to n, $E[t_L^k(S)] > t_L^{k-1}(S)$. Thus $E\left[\frac{\delta_k(S)}{\delta_{k-1}(S)}\right] > 1$. Therefore $\Gamma_k(S)$ expands exponentially.

□

Since the MSW model was designed to approximate the Watts-Strogatz small world model, this can help to explain why the characteristic path length decreases rapidly when the probability of rewiring is small as noted in [49].

Proposition 5.3.4. *For a graph G that is distributed according to a hierarchical model, neighborhood growth occurs exponentially.*

Proof. Let G be a graph distributed according to a hierarchical model with w clusters $C_1 \ldots C_w$, each containing s vertices. For G, neighborhood expansion is slightly slower than for $H \sim G_{n,p}$. Within each cluster vertices are added to $\Gamma_k(S)$ in the same way as for $G_{s,\alpha}$.

Let the set $T^k = \{C_i | \exists v_j \in C_i \cap \Gamma_{k-1}(S)\}$, indicating the clusters in which

vertices are added to $\Gamma_k(S)$ with probability at least α. At step k, each vertex in cluster C_j has probability $1-(1-\beta)^{|\Delta_{k-1}(S)|} \approx 1-\exp(-\beta|\Delta_{k-1}(S)|)$ of being in $\Gamma_k(S)$. Thus $\Pr[C_i \in T^k(S)] \approx 1-\exp(-s\beta|\Delta_{k-1}(S)|)$.

So $E[|T^k|] \approx (w-|T^k|)\,(1-\exp(-s\beta|\Delta_{k-1}(S)|))$. Thus the expansion between clusters also proceeds at a rate greater than for $H \sim G_{w,1-\exp(-s\beta)}$.

□

The hierarchical model also allows for a low characteristic distance paired with a high clustering coefficient. A hierarchical graph model with 1050 vertices arranged in 50 clusters of 21 vertices each with $\alpha = 1$ and $\beta = 0.0003$ has a mean characteristic distance of 5.32 and a mean clustering coefficient of 0.967.

Lemma 1.3.2 relates $\mathrm{vol}(\Gamma(S))$ to $\mathrm{vol}(S)$, $\mathrm{vol}(G)$, and $\lambda_2(M)$. Empirically, $\lambda_2(M)$ for the MSW and Hierarchical models is larger than for a ring lattice.

5.4 An example

As an example, consider the following families of random graphs on 1050 vertices: a MSW model with $r = 10$, $\alpha = 0.9$, and $\beta = 0.005$; a hierarchical model with 50 blocks of 21 vertices, $\alpha = 0.9$, and $\beta = 0.005$; and a model in which each vertex is adjacent to 20 randomly selected vertices with probability $\alpha = 0.9$, and to the other 1029 vertices with probability $\beta = 0.005$. Thus each for each vertex $E[d_i] = 23.145$ and $\mathrm{Var}[d_i] = 6.919275$. Using Theorem A-2, $\Pr[d_{\min} < 11.358] < 0.05$. Using Theorem A-3 $\Pr[d_{\max} > 38.633] < 0.05$. The differences between these models are noted in Table 5.1.

	Hierarchical	MSW	Random
$E[d_i]$	23.145	23.145	23.145
$\mathrm{Var}(d_i)$	6.919	6.919	6.919
$\lambda_1(P)$	23.145	23.145	23.145
$\lambda_2(P)$	17.883	17.895	7.717
$\mathrm{mean}(\lambda_1(A))$	23.493	23.476	23.456
$\mathrm{mean}(\lambda_2(A))$	18.756	19.418	9.367
$\mathrm{mean}(\lambda_1(A-P))$	5.556	5.555	5.546
Characteristic distance	3.0	2.8	2.6
Maximum distance	4	4	3

Table 5.1: Comparison of eigenvalues for the Hierarchial, MSW and Random families

Chapter 6

FUTURE DIRECTIONS

In this dissertation I demonstrated the use of the Courant-Weyl inequalities to find probability bounds for graphs that are distributed according to models in which all of the vertices share a common degree distribution. Such bounds can, however, be found for any edge-product model. Further work remains to be done concerning probability bounds for the eigenvalues of $A - P$ where P is an arbitrary probability matrix.

Both the MSW and hierarchical model can be further generalized. One generalization of the MSW model is the use of more than two values for p_{ij}, if intermediate probabilities of adjacency are desired. The construction of the hierarchical model does not require all clusters to be of the same size. It would also be possible to establish multiple levels of adjacency probability. Both of these aspects remain to be explored.

Bibliography

[1] William Aiello, Fan Chung, and Linyuan Lu, *A random graph model for massive graphs*, Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing (New York), ACM, 2000, pp. 171–180 (electronic). MR MR2114530

[2] Réka Albert and Albert-László Barabási, *Statistical mechanics of complex networks*, Rev. Modern Phys. **74** (2002), no. 1, 47–97. MR MR1895096 (2003d:82055)

[3] Réka Albert, Hawoong Jeong, and Albert-László Barabási, *Diameter of the world-wide web*, Nature **401** (1999), 130–131.

[4] N. Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96, Theory of computing (Singer Island, Fla., 1984). MR MR875835 (88e:05077)

[5] Noga Alon, Michael Krivelevich, and Van H. Vu, *On the concentration of eigenvalues of random symmetric matrices*, Israel J. Math. **131** (2002), 259–267. MR MR1942311 (2003i:15020)

[6] Ali R. Amir-Moéz, *Extreme properties of eigenvalues of a hermitian transformation and singular values of the sum and product of linear transformations*, Duke Math. J. **23** (1956), 463–476. MR MR0079564 (18,105j)

[7] T. Bailey and A. Gatrell, *Interactive spatial data analysis*, Longman Scientific and Technical, Essex, 1995.

[8] S Bansal, BT Grenfell, and LA Meyers, *When individual behaviour matters: homogeneous and network models in epidemiology*, Journal of the Royal Society interface **4** (2007), 879–891.

[9] Albert-László Barabási, *Linked*, Perseus, Cambridge, MA, 2003, The New Science of Networks.

[10] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512. MR MR2091634

[11] A. D. Barbour and Gesine Reinert, *Small worlds*, Random Structures Algorithms **19** (2001), no. 1, 54–74. MR MR1848027 (2002f:90006)

[12] A. Barrat and M. Weigt, *On the properties of small-world network models*, Eur. Phys. J. B **3** (2000), 547–560.

[13] Edward A. Bender and E. Rodney Canfield, *The asymptotic number of labeled graphs with given degree sequences*, J. Combinatorial Theory Ser. A **24** (1978), no. 3, 296–307. MR MR0505796 (58 #21793)

[14] B. Bollobás and F. R. K. Chung, *The diameter of a cycle plus a random matching*, SIAM J. Discrete Math. **1** (1988), no. 3, 328–333. MR MR955649 (89j:05062)

[15] Béla Bollobás, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001. MR MR1864966 (2002j:05132)

[16] Herman Chernoff, *A note on an inequality involving the normal distribution*, Annals of Probability **9** (1981), 533–535.

[17] Fan Chung and Linyuan Lu, *Complex graphs and networks (cbms regional conference series in mathematics)*, American Mathematical Society, Boston, MA, USA, 2006.

[18] Fan R. K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, vol. 92, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997. MR MR1421568 (97k:58183)

[19] Dragoš M. Cvetković, Michael Doob, and Horst Sachs, *Spectra of graphs*, second ed., VEB Deutscher Verlag der Wissenschaften, Berlin, 1982, Theory and application. MR MR690768 (84a:05046)

[20] Reinhard Diestel, *Graph theory*, third ed., Graduate Texts in Mathematics, vol. 173, Springer-Verlag, Berlin, 2005. MR MR2159259 (2006e:05001)

[21] M. Draief, A. Ganesh, and L. Massoulié, *Thresholds for virus spread on networks*, Ann. Appl. Probab. **18** (2008), no. 2, 359–378. MR MR2398760

[22] Rick Durrett, *Random graph dynamics*, Cambridge University Press, Cambridge, U.K., 2006.

[23] P. Erdös and A. Rényi, *On random graphs*, Publ. Math. Debrecen **6** (1959), 290–297.

[24] I. Farkas, I. Derényi, AL Barabási, and T Vicsek, *Spectra of "real-world" graphs: Beyond the semicircle law*, PHYSICAL REVIEW E **64** (2001), no. 2, Part 2 (English).

[25] Z. Füredi and J. Komlós, *The eigenvalues of random symmetric matrices*, Combinatorica **1** (1981), no. 3, 233–241. MR MR637828 (83e:15010)

[26] M. Girvan and M. E. J. Newman, *Community structure in social and biological networks*, Proc. Natl. Acad. Sci. USA **99** (2002), no. 12, 7821–7826 (electronic). MR MR1908073

[27] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, *Self-similar community structure in a network of human interactions*, Phys. Rev. E **68** (2003), no. 6, 065103.

[28] Desmond J. Higham, *A matrix perturbation view of the small world phenomenon*, SIAM J. Matrix Anal. Appl. **25** (2003), no. 2, 429–444 (electronic). MR MR2047427 (2005f:60158)

[29] M. Hofmeister, *Spectral radius and degree sequence*, Math. Nachr. **139** (1988), 37–44. MR MR978106 (90a:05134)

[30] L. Lovász, *Random walks on graphs: a survey*, Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud., vol. 2, János Bolyai Math. Soc., Budapest, 1996, pp. 353–397. MR MR1395866 (97a:60097)

[31] C McDiarmid, *Concentration*, Probabilitistic methods for algorithmic discrete mathematics, vol. 16, Springer, Berlin, 1998, pp. 195–248.

[32] S Milgram, *Small world problem*, Psychology Today **1** (1967), 61–67.

[33] Michael Molloy and Bruce Reed, *A critical point for random graphs with a given degree sequence*, Random Structures and Algorithms **6** (1995), 161–179.

[34] M. E. J. Newman, *Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality*, Phys. Rev. E **64** (2001), no. 1, 016132.

[35] ———, *Scientific collaboration networks. i. network construction and fundamental results*, Phys. Rev. E **64** (2001), no. 1, 016131.

[36] ———, *The structure of scientific collaboration networks*, Proceedings of the National Academy of Science USA **98** (2001), 404–409.

[37] ———, *The structure and function of complex networks*, SIAM Rev. **45** (2003), no. 2, 167–256 (electronic). MR MR2010377 (2005a:05206)

[38] ———, *Finding community structure in networks using the eigenvectors of matrices*, Phys. Rev. E (3) **74** (2006), no. 3, 036104, 19. MR MR2282139 (2007j:82115)

[39] M. E. J. Newman, C. Moore, and D. J. Watts, *Mean-field solution of the small-world network model*, Phys. Rev. Lett. **84** (2000), no. 14, 3201–3204.

[40] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Random graphs with arbitrary degree distributions and their applications*, Phys. Rev. E **64** (2001), no. 2, 026118.

[41] Mark Newman, Albert-Lázló Barabaási, and Duncan J. Watts, *The structure and dynamics of networks*, Princeton University Press, Princeton, New Jersey, 2006.

[42] M.E.J. Newman and D.J. Watts, *Scaling and percolation in the small-world network model*, Physical Review E **60** (1999), 7332–7442.

[43] Ray Solomonoff and Anatol Rapoport, *Connectivity of random nets*, Bulletin of Mathematical Biophysics **13** (1951), 107–117.

[44] J Travers and S Milgram, *Experimental study of small world problem*, Sociometry **32** (1969), 425–443.

[45] Dean L. Urban, Emily S. Minor, Eric A. Treml, and Robert S. Schick, *Graph models of habitat mosaics*, ECOLOGY LETTERS **12** (2009), no. 3, 260–273.

[46] Van H. Vu, *Spectral norm of random matrices*, Combinatorica **27** (2007), no. 6, 721–736. MR MR2384414

[47] Duncan J. Watts, *Small worlds*, Princeton Studies in Complexity, Princeton University Press, Princeton, NJ, 1999, The dynamics of networks between order and randomness. MR MR1716136 (2001a:91064)

[48] ———, *Six degrees*, W. W. Norton & Co. Inc., New York, 2003, The science of a connected age. MR MR2041642 (2004k:91169)

[49] Duncan J. Watts and Steven H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature **393** (1995), 440–442.

[50] Hong Yuan, *A bound on the spectral radius of graphs*, Linear Algebra Appl. **108** (1988), 135–139. MR MR959701 (89f:05123)

Appendix A

CHERNOFF INEQUALITIES

The following Chernoff inequalities are used to prove probabilities concerning the characteristics of edge-product random graphs. The formulations are taken from [17].

Theorem A-1. *Theorem 2.4 in [17], from [16]*

Let $X_1, \ldots, X_n$ *be independent random variables with*
$\begin{cases} \Pr[X_i = 1] = p_i \\ \Pr[X_i = 0] = 1 - p_i. \end{cases}$
If $X = \sum_{i=1}^{n} X_i$, *with expected value* $E[X] = \sum_{i=1}^{n} p_i$, *then*

(Lower tail) $\Pr[X \leq E[X] - \lambda] \leq \exp\left(-\lambda^2/2E[X]\right)$,

(Upper tail) $\Pr[X \geq E[X] + \lambda] \leq \exp\left(-\frac{\lambda^2}{2(E[X]+\lambda/3)}\right)$.

Theorem A-2. *Theorem 2.6 in [17], from [31]*

Let $X_1, \ldots, X_n$ *be independent random variables satisfying*

$X_i \leq E[X_i] + M$, *for* $1 \leq i \leq n$. *For the sum* $X = \sum_{i=1}^{n} X_i$ *with expected value*

$E[X] = \sum_{i=1}^{n} E[X_i]$ *and variance* $\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$,

$\Pr[X \geq E[X] + \nu] \leq \exp\left(-\frac{\nu^2}{2(\mathrm{Var}(X)+M\nu/3)}\right)$.

Theorem A-3. *Theorem 2.7 in [17], from [31]*

Let $X_1, \ldots, X_n$ *be independent random variables satisfying*

$X_i \leq E[X_i] + M$, *for* $1 \leq i \leq n$. *For the sum* $X = \sum_{i=1}^{n} X_i$ *with expected value*

$E[X] = \sum_{i=1}^{n} E[X_i]$ *and variance* $\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$,

$\Pr[X \leq E[X] - \nu] \leq \exp\left(-\frac{\nu^2}{2\sum_{i=1}^{n} E[X_i^2]}\right)$.

Theorem A-4. *Theorem 2.8 in [17]*

Suppose $X_1, \ldots, X_n$ *are independent random variables satisfying* $X_i \leq M$, *for* $1 \leq i \leq n$. *Let* $X = \sum_{i=1}^{n} X_i$ *and let* $\|X\| = \sqrt{\sum_{i=1}^{n} E[X_i^2]}$. *Then*

$\Pr[X \geq E[X] + \nu] \leq \exp\left(-\frac{\nu^2}{2(\|X\|^2 + M\nu/3)}\right)$.

Theorem A-5. *Theorem 2.9 in [17]*

Suppose $X_1, \ldots, X_n$ *are independent random variables satisfying* $X_i \leq M$, *for* $1 \leq i \leq n$. *Let* $X = \sum_{i=1}^{n} X_i$ *and let* $\|\|X\|\| = \sqrt{\sum_{i=1}^{n} E[X_i^2]}$. *Then*

$\Pr[X \leq E[X] - \nu] \leq \exp\left(-\frac{\nu^2}{2(\|\|X\|\|^2 + M\nu/3)}\right)$.

Lemma A-6. *Let* $a, b \in \mathbb{R}^+$ *such that* $a \geq b$. *Let* $n \in \mathbb{N}$, $n \geq \frac{a}{b}$. *Let* $(a_1, a_2, \ldots, a_n)$ *be a sequence of numbers such that* $0 \leq a_i \leq b$ *for all* i *and* $\sum_{i=1}^{n} a_i = a$. *Then* $\sum_{i=1}^{n} a_i^k \leq ab^{k-1}$ *for all* $k \in \mathbb{N}$

Proof. The proof is by induction on k:

For our base case, $k = 1$, $\sum_{i=1}^{n} a_i = a = ab^0 = ab^{k-1}$.

We then assume that $\sum_{i=i}^{n} (a_i)^k \leq ab^{k-1}$ for some $k \geq 1$.

Then $b\sum_{i=1}^{n} (a_i)^k \leq ab^k$. Since each $a_i \leq b$, $\sum_{i=1}^{n} a_i^{k+1} \leq ab^k$. Therefore $\sum_{i=1}^{n} a_i^k \leq ab^{k-1}$ for all $k \in \mathbb{N}$. □