Title
Random subgraphs of a given graph

Permalink
https://escholarship.org/uc/item/5f6659fr

Author
Horn, Paul Kenneth

Publication Date
2009

Peer reviewed|Thesis/dissertation
Random Subgraphs of a Given Graph

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Paul Kenneth Horn

Committee in charge:

Professor Fan Chung Graham, Chair
Professor Samuel Buss
Professor Sanjoy Dasgupta
Professor Ronald Graham
Professor Jason Schweinsberg

2009
The dissertation of Paul Kenneth Horn is approved, and it is acceptable in quality and form for publication on microfilm:

University of California, San Diego

2009
DEDICATION

To Madeline.
We have heard much about the poetry of mathematics, but very little of it has yet been sung.

—Henry David Thoreau
# TABLE OF CONTENTS

Signature Page ........................................ iii

Dedication ............................................. iv

Epigraph ................................................ v

Table of Contents ....................................... vi

List of Figures ......................................... viii

Acknowledgements ...................................... ix

Vita and Publications ................................ xi

Abstract .............................................. xii

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>Prerequisites and Overview</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2 Notation</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1.3 Spectral Graph Theory</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1.4 Probabilistic Prerequisites</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1.5 Overview</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>The Giant Component in Percolated Graphs</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.1 Introduction</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>2.2 Preliminaries</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2.3 Several spectral lemmas</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>2.4 The range of $p$ with no giant component</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>2.5 The emergence of the giant component</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>2.6 Interesting Questions</td>
<td>34</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>The Structure of Percolated Graphs</td>
<td>37</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>37</td>
</tr>
<tr>
<td>3.2</td>
<td>Preliminaries</td>
<td>42</td>
</tr>
<tr>
<td>3.3</td>
<td>A bound on $</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>Bounding the Spectral Gap</td>
<td>52</td>
</tr>
<tr>
<td>3.5</td>
<td>Remarks and Optimality</td>
<td>59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4</th>
<th>Random Spanning Trees in General Graphs</th>
<th>65</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>65</td>
</tr>
<tr>
<td>4.2</td>
<td>Preliminaries</td>
<td>70</td>
</tr>
<tr>
<td>4.3</td>
<td>Random spanning trees generated by random walks</td>
<td>74</td>
</tr>
<tr>
<td>4.4</td>
<td>Proving a diameter Lower Bound for random spanning trees</td>
<td>76</td>
</tr>
<tr>
<td>4.5</td>
<td>Proof of Upper Bound</td>
<td>81</td>
</tr>
<tr>
<td>4.6</td>
<td>Discussion</td>
<td>86</td>
</tr>
</tbody>
</table>

| Bibliography |                                      | 91 |
LIST OF FIGURES

Figure 1.1: A host graph and a random subgraph obtained with bond per-
colation, $p = 0.1$. .................................................... 7

Figure 1.2: A host graph and a uniformly chosen spanning tree within the graph. ................................................. 11

Figure 2.1: A host graph and two random subgraphs obtained by bond percolation; the first with no giant component, and the second with a giant component. .............................................. 15

Figure 4.1: A host graph and a spanning tree within the graph. The diameter of the spanning tree is 51, while the diameter of the underlying host graph is 10. .................................................. 67

Figure 4.2: The Groundskeeper Algorithm: Run a random walk, record the first edge a node is visited through. ............... 74
ACKNOWLEDGEMENTS

First and foremost, I would like to thank Fan Chung for her immense amount of patience and support in molding me into a mathematician. Certainly I would not be where I am today without her guidance, for which I am eternally grateful. I also wish to thank my co-authors, and other mathematical collaborators notably Lincoln Lu and Jacques Verstraete for their assistance and wisdom; I have learned a great deal from all of you. Also, many thanks go to Anthony Bonato who hosted me at Fields and Lincoln (who hosted me in his own home). I thank Ron Graham for giving me my first mathematical reward check, and for exposing me to many interesting problems. My mathematical thanks would not be complete without giving a special thanks to Professor Mark Goldberg at RPI, who initiated my interest in graph theory.

My parents have also given me a tremendous amount of support from a young age to the present day; I appreciate both their exposing me to mathematics at a young age and letting me come into it at my own pace. I have made many friends here in San Diego who have all contributed in a large way to my successful work. In the math department, I have had the pleasure of writing a paper with Steve and Eric, and I also owe special thanks to many officemates over the years: Ross, Linh, Shoaib, Mary, Jake, Greg. I also must thank my roommates for putting up with me over the past 5 years; especially thanks to Spike, Sharon and Jon. Thanks also go to those who have spent time camping and hiking with me, and also I have special thanks for the mountains, canyons, and deserts of California themselves. How many times have I solved a problem while walking through Rose Canyon? What could be more inspiring than the remotest of canyons of Death Valley?

Last, but certainly not least, I want to thank Maddy whose love and support over these past few years – and especially in this final stretch – made finishing this thesis possible. I am looking forward to our next great adventure!

Chapter 2 is based on the paper ”The giant component in a random subgraph of a given graph”, Proceedings of WAW2009, Lecture Notes in Computer Science
5427, 2009, 38–49; joint with Fan Chung and Linyuan Lu and appears with kind permission of Springer Science+Business Media. The dissertation author was the primary author of this work.

Chapter 3 is partially based on the paper "The spectral gap of a random subgraph of a graph", Internet Mathematics, volume 4, 2-3, 2007, 225–244.; joint with Fan Chung, and appears with kind permission of AK Peters. The dissertation author was the primary author of this work.

Chapter 4 is partially based on the paper "Diameter of random spanning trees in a given graph", which has been submitted; joint with Fan Chung and Linyuan Lu. The dissertation author was the primary author of this work.
VITA

2003  B. S. in Mathematics *summa cum laude*, Rensselaer Polytechnic Institute

2003  B. S. in Computer Science *summa cum laude*, Rensselaer Polytechnic Institute

2003-2004  Graduate Teaching Assistant, University of Wisconsin, Madison

2004-2009  Graduate Teaching Assistant, University of California, San Diego

2006  M. A. in Mathematics, University of California, San Diego

2008  Associate Instructor, University of California, San Diego

2009  Ph. D. in Mathematics, University of California, San Diego

PUBLICATIONS

F. Chung and P. Horn, The spectral gap of a random subgraph of a graph, to appear in *Internet Mathematics*

S. Butler, P. Horn and E. Tressler, Intersecting Domino Tilings, submitted.

F. Chung, P. Horn and L. Lu, Diameter of random spanning trees in a given graph, submitted.


A. Harutyunyan, P. Horn, and J. Verstraete, Independent dominating sets in graphs of girth five, submitted.
Random Subgraphs of a Given Graph

by

Paul Kenneth Horn

Doctor of Philosophy in Mathematics

University of California San Diego, 2009

Professor Fan Chung Graham, Chair

In this thesis, we explore several problems related to understanding the relationships between a random subgraph of a host graph and the host graph itself, using the spectrum of the normalized Laplacian to understand the structure of both host graph and subgraph. In particular:

• We study the emergence of the giant component in random subgraphs of a given graph. Under some relatively mild conditions on the degree sequence and spectral gap of the Laplacian, we show that if a graph $G$ is percolated with probability $p \leq \frac{(1-\epsilon)}{\bar{d}}$, where $\bar{d}$ is the ratio of the second and first moments of the degree sequence, then the resulting subgraph $G_p$ has no giant component asymptotically almost surely, while if $p \geq \frac{(1+\epsilon)}{\bar{d}}$ then there is a unique giant component with volume a constant fraction of the volume of the entire graph. This extends earlier work of Erdős and Rényi, who considered the problem with host graph $K_n$, and Frieze, Krivelevich and Martin who considered regular pseudo-random graphs, and others.

• For a denser range of $p$, we exploit the spectral gap to show similarities between the structure of a random subgraph of a graph and the underlying host
graph. The spectral gap of the normalized Laplacian yields a great deal of structural information about a graph, including discrepancy and expansion properties. We show that the spectral gap of a random subgraph is asymptotically the same as the spectral gap of host graph so long as \( pd_{\text{min}} \gg \log^{3/2}(n) \).

- We also study another model of random subgraph, namely a uniformly randomly chosen spanning tree. We study the height of random spanning trees of general graphs; improving and generalizing earlier results of Aldous. Again, the spectral gap of the Laplacian is a key parameter in our understanding.
Chapter 1

Prerequisites and Overview

1.1 Introduction

In recent years, graph theory has emerged as an important tool in understanding large networks which occur in such disperse areas as telephone and information networks, contact and social networks, and biological networks. Often, we are not able to see the entire network we are interested in, and instead see some substructure within the network. Other times, we are interested in some process which only occurs on some of the network, for instance the spread of disease.

An important question, then, is how do properties of these observed or important subgraphs relate to their underlying host graph. One can ask, how does the structure of typical subgraphs relate to properties of host graphs? A natural way to address the question of how typical subgraphs look is to investigate properties of random subgraphs, which is the goal of this thesis. In particular, our goal is to investigate the structure of random subgraphs for the most general class of graphs we can handle.

In order to study random subgraphs of very general graphs, we still must have some control over these graphs. Our methods combine both probabilistic tools with tools from spectral graph theory, which we will use to both understand our
host graph and understand the properties that our random subgraphs possess with high probability. In this chapter, we survey some of the important tools and results which allow us to, in subsequent chapters, understand many properties of random subgraphs. We also introduce what, exactly, we mean by a random subgraph; introducing the two models we consider during this thesis.

1.2 Notation

Our graph theory terminology throughout is by and large, standard. A graph, $G$, is a collection of vertices and edges. We consider graphs which are simple (i.e. no loops and no multiple edges) and undirected. For a vertex $v$, we denote the degree of $v$ by $d_v$. The $k$-th order volume of a set $X$ of vertices is defined as

$$\text{vol}_k(X) = \sum_{v \in X} d_v^k.$$ 

We let $\text{vol}(G) = \text{vol}_1(G)$. We denote by $d$ the average degree of $G$, that is

$$d = \frac{\sum_{v \in G} d_v}{n} = \frac{\text{vol}_1(G)}{\text{vol}_0(G)}.$$ 

Another parameter which becomes important in the statement and proofs of our theorems is the second order average degree,

$$\tilde{d} = \frac{\text{vol}_2(G)}{\text{vol}_1(G)}.$$ 

For a $d$-regular graph, where each vertex has degree $d$, we have that $\text{vol}_k(X) = d^k|X|$ and $\tilde{d} = d$. In general, the Cauchy-Schwarz inequality implies that $\tilde{d} \geq d$. For irregular graphs, higher order volumes and the ratio $\tilde{d}/d$ provide some measure of the skewness of the degree distribution.

Generally speaking, if we have a graph $G$ and a random subgraph $H$, $\text{vol}(X)$ will refer to the volume of $X$ in $G$. We will use $\text{vol}_H(X)$ to denote the volume of $X$ in $H$. Throughout, unless explicitly noted, if a quantity is to be taken in $H$ it will be subscripted with $H$. Otherwise, quantities are taken to be from the underlying host graph.
1.3 Spectral Graph Theory

Spectral graph theory concerns itself with understanding the connections between eigenvalues of matrices associated with a graph $G$ and properties of the graph itself. The simplest such matrix commonly used is the adjacency matrix:

$$ [A]_{ij} = \begin{cases} 
1 & \text{if } v_i \sim v_j \\
0 & \text{otherwise.}
\end{cases} $$

Let

$$ \rho_0 \geq |\rho_1| \geq \cdots \geq |\rho_{n-1}| $$

denote the eigenvalues of the adjacency matrix in decreasing order. For a $d$-regular graph $G$, it is easy to see that $d$ is an eigenvalue of $A$ with corresponding eigenvector $1$. If $G$ is not $d$-regular then several inequalities can be obtained, for instance, $d_{\text{max}} \geq \rho_0 \geq \sqrt{d_{\text{max}}}$. Other inequalities can be proved as well; for example it is shown in Chapter 2 that (under some conditions) $\rho_0 = (1 + o(1))\bar{d}$. Overall, however, the adjacency matrix is most useful in understanding properties of regular or almost regular graphs. This can be seen, for example, in the expander mixing lemma (see eg. [KS06a])

**Expander Mixing Lemma.** Suppose $G$ is $d$-regular. For any two sets $X$ and $Y$, the number of edges between them $e(X,Y)$ satisfies

$$ \left| e(X,Y) - \frac{d|X||Y|}{n} \right| \leq |\rho_1|\sqrt{|X||Y|}. $$  \hspace{1cm} (1.1)

If $|\rho_1| = o(d)$, this suggests that $G$ looks, in some sense, like a random graph. This idea is formalized by the notion of Quasi-random graphs initiated by Chung, Graham and Wilson [CGW89]. Regular graphs on $n$ vertices with $|\rho_1| = \lambda$ are known as $(n,d,\lambda)$-graphs. Many properties of these graphs are given in [KS06a].

A disadvantage of the adjacency matrix when dealing with irregular graphs is the fact that structural information can be overwhelmed with degree information. It has been shown by Chung, Lu and Vu [CLV04] and Mihail and Papidimitrou [MP06] the eigenvalues of the adjacency matrix of random graphs obeying a power
law degree distribution (that is, the number of vertices of degree \(k\) is proportional to \(k^{-\beta}\)) follows a powerlaw themselves. Here the degree sequence dominates other useful information we could gather from the degree sequence.

Another matrix commonly studied related to graphs is the combinatorial Laplacian, defined by

\[
L = D - A
\]

where \(D = \text{diag}(d_1, \ldots, d_n)\) denotes the diagonal degree matrix. It is easy to observe that all eigenvalues of \(L\) are non-negative; and 0 always occurs as an eigenvalue with eigenvector \(1\). To see that all eigenvalues of \(L\) are non-negative, one method is to note that it is diagonally dominant and the Geršgorin disc theorem implies the result. It is also possible to prove this using the fact that \(L = BB^*\) for \(B\) the boundary operator. Here, \(B\) is a \(V \times E\) matrix such that for an edge \(e = \{v_i, v_j\}\),

\[
B(e, v) = \begin{cases} 
1 & v = v_i \\
-1 & v = v_j \\
0 & \text{otherwise}
\end{cases}
\]

Note that \(B\) gives an (arbitrary) orientation to the edges.

The combinatorial Laplacian has a number of applications, including the celebrated matrix tree theorem of Kirchhoff [Kir47]

**Matrix Tree Theorem** (Kirchhoff, 1847). The number of spanning trees of \(G\) is the absolute value of the determinant of a principal cofactor of \(L\).

In light of this, the Laplacian is occasionally known as the Kirchoff matrix of a graph.

The primary matrix representation of a graph which we will use is a normalized version of the combinatorial Laplacian. We define the (normalized) Laplacian of \(G\) to be

\[
\mathcal{L} = D^{-1/2}LD^{-1/2} = I - D^{-1/2}AD^{-1/2}.
\]
The eigenvalues of $L$ in increasing order are

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2.$$ 

We define the spectral gap of the Laplacian to be

$$\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}.$$ 

It is not difficult to show that $\sigma < 1$ if and only if $G$ is connected and not bipartite. If $\sigma$ is separated from 1, then $G$ has many nice expansion, discrepancy and mixing properties; these are discussed in more detail in [Chu97]; in Chapter 3 these are used to understand how structural properties of a host graph and random subgraph are related (cf. Corollaries 1 and 2). An advantage of $L$ is that its eigenvalues provide useful structural information even in the case that $G$ is not regular. As an example, we give the following generalized version of the expander mixing lemma.

**Expander Mixing Lemma** (for general graphs). If $G$ is a graph, with spectral gap of the Laplacian $L$, then

$$\left| e(X,Y) - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)} \right| \leq \sigma \sqrt{\frac{\text{vol}(X)\text{vol}(Y)}}.$$ 

We give a proof of this result (adapted from [Chu97]) as it illustrates a general technique which we will need to use to prove stronger results in Chapter 2.

*Proof.* We let $1_X$ denote the (column) indicative vector for the set $X$. Note that $D^{-1/2}1$ is an eigenvector of $L$ corresponding to the eigenvalue 0. Let $M = I - L$, and $\varphi_0 = \frac{D^{1/2}}{\sqrt{\text{vol}(G)}}, \varphi_1, \ldots, \varphi_{n-1}$ denote an orthonormal set of eigenvectors for $L$, so that

$$M = \sum_{i=0}^{n-1} (1 - \lambda_i) \varphi_i \varphi_i^*.$$
Then
\[
\begin{align*}
|e(X, Y) - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)}| \\
= |1_X^* A 1_Y - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)}| \\
= |1_X^* D^{1/2} M D^{1/2} 1_Y - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)}| \\
= |1_X^* D^{1/2} \left( \frac{D^{1/2} 11^* D^{1/2}}{\text{vol}(G)} + \sum_{i=1}^{n-1} (1 - \lambda_i) \varphi_i \varphi_i^* \right) D^{1/2} 1_Y - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)}| \\
\leq \|1_X^* D^{1/2}\| \cdot \left| \sum_{i=1}^{n-1} (1 - \lambda_i) \varphi_i \varphi_i^* \right| \cdot \|D^{1/2} 1_Y\| \\
= \sigma \sqrt{\frac{\text{vol}(X)\text{vol}(Y)}}{\text{vol}(G)}
\end{align*}
\]

Remark. Slightly more careful analysis in the above can lead to the slightly stronger conclusion:
\[
|e(X, Y) - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)}| \leq \sigma \sqrt{\frac{\text{vol}(X)\text{vol}(Y)\text{vol}(\bar{X})\text{vol}(\bar{Y})}{\text{vol}(G)}}.
\]

An interesting observation about $\mathcal{L}$, used extensively in Chapter 4, is the fact that eigenvalues of $\mathcal{L}$ are closely related to the rate of convergence of random walks. This occurs as $I - \mathcal{L} = D^{-1/2} A D^{-1/2}$ is similar to the random walk transition matrix $W = D^{-1} A$; in this sense $\mathcal{L}$ can be thought of as a symmetric version of $W$ which has been shifted.

### 1.4 Probabilistic Prerequisites

In addition to the spectrum, we need probabilistic tools to understand properties of random subgraphs. In this thesis, we are primarily concerned with two models of random subgraphs. The first model under consideration is that of bond
percolation. Here, given a host graph $G$, we find a random subgraph $G_p$ by independently choosing edges to be in $G_p$ with probability $p$, discarding them with probability $1 - p$. This model is well studied for certain special classes of graphs; see e.g. [Kes91]. Probabilists tend to be interested in properties of $G_p$ when $G$ is an (infinite) lattice. Here a typical question might be whether or not $G_p$ contains an infinite component. Within combinatorics, the special case where $G = K_n$ is particularly well studied. This is one of the model of random graphs introduced by Erdős and Rényi in [ER59], and has become known as the $G_{n,p}$ model (though, actually, [ER59] primarily considered a related model consisting of choosing graphs uniformly among all with exactly $m$ edges.) Erdős and Rényi’s work on random graphs was itself predated by earlier work in [Gil59] by Gilbert.

Figure 1.1: A host graph and a random subgraph obtained with bond percolation, $p = 0.1$.

In the $G_{n,p}$ model, which is our primary motivation in Chapters 2 & 3, we are interested in what properties graphs $G \in G_{n,p}$ have with high probability. In particular, we say that a property $\mathcal{A}$ holds in $G_{n,p}$ asymptotically almost surely, or a.a.s., if given a graph $G \in G_{n,p}$

$$\lim_{n \to \infty} \mathbb{P}(G \in \mathcal{A}) = 1.$$  

We use this notation throughout; of course for a (specific) general graph the error probability $\mathbb{P}(\bar{\mathcal{A}})$ is not actually zero, but our results show that this error probability decays with the size of the graph. As an illustrative example, we mention
a few properties that hold asymptotically almost surely for graphs in \( G_{n,p} \); indeed these were among the first properties observed by Erdős and Rényi. Random graph theory has become a well developed area of combinatorics, the interested reader may consult two monographs on the subject, [Bol01] and [JLR00], for many more interesting properties.

**Theorem 1** (Giant Component Threshold). Suppose \( G \in G_{n,p} \). Then

- If \( p < \frac{1-\epsilon}{n} \) the largest connected component of \( G \) has size \( \Theta(\log(n)) \) a.a.s.
- If \( p > \frac{1+\epsilon}{n} \) the largest connected component of \( G \) has size \( \Theta(n) \) a.a.s.
- If \( p = \frac{1}{n} \) the largest connected component of \( G \) has size \( \Theta(n^{2/3}) \) a.a.s.

**Theorem 2** (Connectivity Threshold). Suppose \( G \in G_{n,p} \). Then

- If \( p < (1 - \epsilon) \frac{\log n}{n} \) then \( G \) is disconnected (in fact, \( G \) has isolated vertices) a.a.s.
- If \( p > (1 + \epsilon) \frac{\log n}{n} \) then \( G \) is connected a.a.s.

In order to prove such results we often need to understand how a random variable \( X \) differs from its mean. We accomplish this by the use of several concentration inequalities. Simplest among these are Markov’s inequality and Chebyshev’s inequality.

**Proposition 1** (Markov’s Inequality). Suppose \( X \geq 0 \) is a random variable. Then

\[
P(X > k\mathbb{E}[X]) \leq \frac{1}{k}.
\]

**Proposition 2** (Chebyshev’s Inequality). Suppose \( X \) is a random variable, and \( \sigma^2 = \mathbb{E}((X - \mathbb{E}[X])^2) \). Then

\[
P(|X - \mathbb{E}[X]| \geq k\sigma) \leq \frac{1}{k^2}.
\]
For random variables which are sums of independent indicators, tighter concentration is possible via the Chernoff bounds. These can be thought of as quantitative versions of the central limit theorem. Here we give a few versions of the Chernoff bounds which we use throughout this thesis, for a more comprehensive survey of these inequalities see [CL06].

Our first Chernoff bound is the simplest and cleanest, useful when $X$ is the sum of independent indicators.

**Proposition 3** (Chernoff Bounds, v.1). Suppose $X \sim B(n,p)$ is a binomial random variable with mean $np$. Then

$$
\Pr(|X - \mathbb{E}[X]| \geq \lambda \sqrt{\mathbb{E}[X]}) \leq 2 \exp \left( -\frac{\lambda^2}{2} \right).
$$

Our next two versions, which are heavily used in Chapters 2 and 3 allow $X$ to be the sum of independent variables whose values are not necessarily $0-1$. These come in to play, for instance, when asking for concentration of the volume of a random set.

**Proposition 4** (Chernoff Bounds, v.2). For $1 \leq i \leq n$, let $X_i$ be independent $0-1$ valued random variables. Let $X = \sum \alpha_i X_i$ where $0 \leq \alpha_i \leq \Delta$ for all $i$. Then

$$
\Pr(X \leq \mathbb{E}(X) - \lambda) \leq \exp \left( -\frac{\lambda^2}{2 \sum \alpha_i^2 \mathbb{E}[(\alpha_i^2)]} \right) \leq \exp \left( -\frac{\lambda^2}{2 \Delta \mathbb{E}(X)} \right).
$$

**Proposition 5** (Chernoff Bounds, v.3). For $1 \leq i \leq n$, let $X_i$ be independent random variables satisfying $|X_i| \leq M$. Let $X = \sum_{i} X_i$. Then we have, for any $\lambda > 0$,

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

In addition to versions of the Chernoff bounds, we occasionally need to use martingales in order to show concentration. These prove particularly useful when the quantity which we are asking for concentration of is not the sum of independent indicators, but is still a function of a number of other random variables such that
a small change in an given variable does not greatly affect the solution. In order to properly state the result, we give a brief definition of martingales.

A filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n$ is an increasing set of $\sigma$-fields. We say that a set of random variables $(X_i)_{i=0}^n$ is a martingale if it satisfies the following properties:

1. $X_i$ is $\mathcal{F}_i$ measurable.
2. $\mathbb{E}[X_i|\mathcal{F}_{i-1}] = X_{i-1}$.
3. $\mathbb{E}[|X_i|] < \infty$.

If, for a vector $c = (c_1, \ldots, c_n)$, a martingale $(X_i)_{i=0}^n$ satisfies

$$|X_i - X_{i-1}| \leq c_i$$

we say that $(X_i)_{i=0}^n$ is $c$-Lipschitz. For a $c$-Lipschitz martingale, the Azuma-Hoeffding inequality asserts that $X_n$ is sharply concentrated as follows:

**Proposition 6** (Azuma-Hoeffding inequality). Suppose $(X_i)_{i=0}^n$ is a $c$-Lipschitz martingale. Then

$$\mathbb{P}(|X_n - \mathbb{E}[X_n]| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2\sum_{i=1}^n c_i^2}\right).$$

Variations of Azuma’s inequality exist allowing that the martingale under consideration is almost Lipschitz. While these inequalities are not used in the present thesis (though the general idea of this argument is used in Chapter 4) versions can be found in Chung and Lu [CL06], though the ideas can be traced back to earlier important work of Kim and Vu [KV00] on polynomial concentration and even earlier work of Shamir and Spencer [SS87] on the concentration of the chromatic number, which itself was one of the first uses of martingales in combinatorics.

For any random variable $X$ and filtration $(\mathcal{F}_i)_{i=1}^n$ where $X$ is $\mathcal{F}_n$ measurable, the collection of random variables

$$X_i = \mathbb{E}[X|\mathcal{F}_i]$$
is easily checked to be a martingale (by the tower property of conditional expectation) known as an exposure or Doob martingale. Given a random variable $X$ which is determined by, not necessarily independent, random variables $X_1, \ldots, X_n$ it is natural to construct a martingale using the filtration given by $\mathcal{F}_i = \sigma(X_1, \ldots, X_i)$. Two frequently used types of martingales are the edge-exposure and vertex-exposure martingales, where $\mathcal{F}_i$ is the filtration generated by what occurs on the first $i$ edges or vertices respectively. More information on martingales is available in the book of Williams [Wil91].

Figure 1.2: A host graph and a uniformly chosen spanning tree within the graph.

As a second model of random subgraph, we consider random spanning trees within a given host graph $G$. Recall that for a connected graph $G$, a spanning tree is a loopless connected subgraph of $G$ on the same vertex set. Cayley’s theorem enumerates trees within $K_n$, stating that there precisely $n^{n-2}$ such trees. Rényi and Szekeres [RS67] determine that the height of a random tree within $K_n$ is $\Theta(\sqrt{n})$. In the early 1990’s, Aldous [Ald90] studied random spanning trees including determining (within a factor of $\log^2(n)$) the height of trees in regular graph with spectral gap $\sigma < 1 - \epsilon$. Pemantle [Pem95] studied additional properties of such trees, focusing on the degree distribution and using properties of electrical networks. Several algorithms due to Broder [Bro89] and Aldous [Ald90], as well as a later algorithm due to Wilson [Wil96] are known to generate uniformly random spanning trees. These are detailed in Chapter 4. Pemantle’s paper also provides a good survey of the subject.
1.5 Overview

The rest of this thesis is organized as follows. In Chapter 2, we study the emergence of the giant component in percolated general graphs. The results in this chapter generalize those of Erdős and Rényi and other authors to a broader class of host graphs. In Chapter 3, we examine more closely the structural properties of random subgraphs obtained by percolation. To do this, we provide a bound between the spectral gap of the random subgraph $G_p$ and the underlying host graph $G$. As the spectral gap provides a great deal of structural information about a graph, this theorem provides information about many expansion and discrepancy properties of both the subgraph and the host graph given some knowledge of the other. Finally, in Chapter 4 we study the height of random subtrees of more general graphs. In particular, we improve the lower bound of Aldous, removing a logarithmic factor, and generalize his results to graphs with a more general degree distribution.
Chapter 2

The Giant Component in Percolated Graphs

2.1 Introduction

A major theme of this thesis is to understand connections between an underlying, general, host graph and its random subgraph. In this chapter we treat the emergence of the giant component. In their first paper on random graphs, Erdős and Rényi observed that there was a sharp transition in the size of the largest component of a random subgraph of $K_n$. Later authors have shown that such a phase transition occurs for a number of graph classes. Here we give a treatment for a very general class of graphs, using spectral information about our host graph to guide our approach.

We are interested, in this chapter, in random subgraphs of $G_p$ of a graph $G$ obtained as follows: for each edge in $G$ we independently decide to retain the edge with probability $p$, and discard the edge with probability $1 - p$. A natural special case of this process is the Erdős-Rényi graph model $G_{n,p}$ which is the special case where the host graph is $K_n$. This procedure is known, in the probability literature, as bond percolation. These types of percolation problems are widely
studied [Gri99, Kes91] in theoretical physics, mainly with the host graph being a lattice, eg. $\mathbb{Z}^k$. Our study also lends a potential physical interpretation to the graph; for instance we are concerned with the case where the underlying graph is a contact graph, consisting of edges formed by pairs of people with possible contact. This situation is of special interest in the study of the spread of infectious disease and the identification of communities in social networks.

The fundamental question we ask in this chapter, is for the critical value of $p$ such that $G_p$ has a giant connected component, that is a component whose volume is a positive fraction of the total volume of the graph. For the spread of disease on contact networks, the answer to this question corresponds to the problem of finding the epidemic threshold for the disease under consideration, for instance.

For the case of $K_n$, Erdős and Rényi answered this in their seminal paper [ER59]: if $p = \frac{c}{n}$ for $c < 1$, then asymptotically almost surely $G$ contains no giant connected component and all components are of size at most $O(\log n)$, and if $c > 1$ then, indeed, there is a giant component of size $\epsilon n$. We call $\frac{1}{n}$ the giant component threshold, which we denote by $p_c$ (for critical probability.) In general, we say that $p_c$ is the giant component threshold for a class of graphs if $p < (1 - \epsilon)p_c$ implies that there is no giant component in a percolated subgraph of $G$ a.a.s., and if $p > (1 + \epsilon)p_c$ there is a giant component a.a.s.

For general host graphs, the answer has been more elusive. Results have been obtained only for very dense graphs, bounded degree and regular graphs. Bollobás, Borgs, Chayes and Riordan [BBCR07] showed that for dense graphs (where the degrees are of order $\Theta(n)$), the giant component threshold is $p_c = 1/\rho$ where $\rho$ is the largest eigenvalue of the adjacency matrix. Frieze, Krivelevich and Martin [FKM04] consider the case where the host graph is $d$-regular with adjacency eigenvalue $\lambda$ and they show that the critical probability is close to $1/d$, strengthening earlier results on hypercubes [AKS82, BKL92] and Cayley graphs [MP06]. For expander graphs with degrees bounded by $d$, Alon, Benjamini and Stacey [ABS04] proved that the percolation threshold is greater than or equal to $1/(2d)$.

It is interesting to note that in these above cases where $p_c$ is known, $p_c = \frac{1}{\rho}$.
where \( \rho \) is the largest eigenvalue of the adjacency matrix. This observation comes from the fact that the above mentioned graphs are regular and hence have largest eigenvalue equal to the degree of regularity. It is known, due to Kesten [Kes91], that for the grid \( \mathbb{Z}^2 \) that \( p_c = \frac{1}{2} \). Here, some explanation is necessary; for an infinite graph \( p_c \) is the threshold under which all components are finite a.s., but over which there is necessarily an infinite component. Kesten’s result implies that if \( p \leq \frac{1}{2} \) there is no infinite component almost surely, and if \( p > \frac{1}{2} \) there is an infinite component with probability 1. This interestingly stands in contrast to the other known cases in this problem, the largest eigenvalue of the adjacency operator of \( \mathbb{Z}^2 \) is 4. It is also interesting to note that these are notoriously difficult problems; for instance, \( p_c \) is unknown for \( \mathbb{Z}^3 \).

Figure 2.1: A host graph and two random subgraphs obtained by bond percolation; the first with no giant component, and the second with a giant component.

Here, we are interested in percolation on graphs which are not necessarily regular, and can be relatively sparse (i.e., \( o(n^2) \) edges.) Compared with the earlier
results discussed above, the main advantage of our results is the ability to handle general degree sequences.

To state our results, we recall a few definitions here. For a subset $S$ of vertices the volume of $S$, denoted by $\text{vol}(S)$, is the sum of degrees of vertices in $S$. The $k$th order volume of $S$ is the $k$th moment of the degree sequence, i.e. $\text{vol}_k(S) = \sum_{v \in S} d_v^k$. We write $\text{vol}_1(S) = \text{vol}(S)$ and $\text{vol}_k(G) = \text{vol}_k(V(G))$, where $V(G)$ is the vertex set of $G$. We denote by $\tilde{d} = \text{vol}_2(G)/\text{vol}(G)$ the second order degree of $G$, and by $\sigma$ the spectral gap of the normalized Laplacian, which we fully define in the next section. Further, recall that $f(n)$ is $O(g(n))$ if $\limsup_{n \to \infty} |f(n)|/|g(n)| < \infty$, and $f(n)$ is $o(g(n))$ if $\lim_{n \to \infty} |f(n)|/|g(n)| = 0$.

We will prove the following

**Theorem 3.** Suppose $G$ has the maximum degree $\Delta$ satisfying $\Delta = o(\frac{\bar{d}}{\sigma})$. For $p \leq 1 - \frac{\epsilon}{\bar{d}}$, a.a.s. every connected component in $G_p$ has volume at most $O(\sqrt{\text{vol}_2(G)} g(n))$, where $g(n)$ is any slowly growing function as $n \to \infty$. 

Here, an event occurring a.a.s. indicates that it occurs with probability tending to one as $n$ tends to infinity. In order to prove the emergence of giant component where $p \geq (1 + c)/\bar{d}$, we need to consider some additional conditions. Suppose there is a set $U$ satisfying

1. $\text{vol}_2(U) \geq (1 - \epsilon)\text{vol}_2(G)$.

2. $\text{vol}_3(U) \leq M\text{vol}_2(G)$

where $\epsilon$ and $M$ are constants independent of $n$. In this case, we say $G$ is $(\epsilon, M)$-admissible and $U$ is an $(\epsilon, M)$-admissible set.

We note that the admissibility measures the skewness of the degree sequence. For example, all regular graphs are $(\epsilon, 1)$-admissible for any $\epsilon$, but a graph need not be regular to be admissible. We also note that in the case that $\text{vol}_3(G) \leq M\text{vol}_2(G)$, $G$ is $(\epsilon, M)$-admissible for any $\epsilon$. 
Theorem 4. Suppose \( p \geq \frac{1+c}{d} \) for some \( c \leq \frac{1}{20} \). Suppose \( G \) satisfies \( \Delta = o\left(\frac{\tilde{d}}{\log n}\right) \) and \( \sigma = o(n^{-\kappa}) \) for some \( \kappa > 0 \), and \( G \) is \( (\frac{c\kappa}{10}, M) \)-admissible. Then a.a.s. there is a unique giant connected component in \( G_p \) with volume \( \Theta(\text{vol}(G)) \), and no other component has volume more than \( \max(2d \log n, \omega(\sigma \sqrt{\text{vol}(G)}) \).}

Here, recall that \( f(n) = \Theta(g(n)) \) if \( f(n) = O(g(n)) \) and \( g(n) = O(f(n)) \). In this case, we say that \( f \) and \( g \) are of the same order. Also, \( f(n) = \omega(g(n)) \) if \( g(n) = o(f(n)) \). For the purposes of graphs which are not regular, we note that the volume of a component is a better measure of size than the number of vertices in it. This is why, in Theorems 3 and 4 we concern ourselves with the volume of the resulting set, as opposed to its cardinality. We also observe that the volume taken in the conclusions of Theorems 3 and 4 is the volume within \( G \), the underlying host graph; that is to say that Theorem 4 asserts that, under its hypothesis, there exists a component in \( G_p \), the sum of whose degrees in \( G \) is at least \( \epsilon \text{vol}(G) \) for some \( \epsilon > 0 \). That this implies that it’s volume in \( G_p \) is at least a positive fraction of the volume of \( G_p \) with high probability follows as an application of the Chernoff bounds, Proposition 3 of Chapter 1.

We note that under the assumption that the maximum degree \( \Delta \) of \( G \) satisfying \( \Delta = o\left(\frac{\tilde{d}}{d}\right) \), it can be shown that the spectral norm of the adjacency matrix satisfies \( \|A\| = \rho = (1+o(1))\tilde{d} \); this is the content of Lemma 1 below. Under the assumption in Theorem 2, we observe that the percolation threshold of \( G \) is \( \frac{1}{d} \).

To examine when the conditions of Theorems 1 and 2 are satisfied, we note that (cf. Lemma 4 below) admissibility implies that \( \tilde{d} = \Theta(d) \), which essentially says that while there can be some vertices with degree much higher than \( d \), there cannot be too many. Chung, Lu and Vu [CLV04] show that for random graphs with a given expected degree sequence \( \sigma = O\left(\frac{1}{\sqrt{n}}\right) \), and hence for graphs with average degree \( n^\epsilon \) the spectral condition of Theorem 2 easily holds for random graphs. The results here can be viewed as a generalization of the result of Frieze, Krivelevich and Martin [FKM04] with general degree sequences and is also a strengthening of the original results of Erdős and Réyni to general host graphs.
An important note is that there is some amount of abuse of notation when stating our results. In particular, for a general graph $G$, what does it mean to say that, for instance, ‘$\Delta = o(\frac{\tilde{d}}{\sigma})$’? A more precise statement of Theorems 3 and 4 is to state them in terms of sequences of graphs. For instance, Theorem 3 may be stated as

**Theorem 5.** Suppose $\{G_n\}$ is a sequence of graphs, having maximum degree $\Delta = \Delta(n)$ satisfying $\Delta = o(\frac{\tilde{d}}{\sigma})$. For $p \leq 1 - \epsilon \frac{\tilde{d}}{d}$, a.a.s. every connected component in $G_p$ has volume at most $O(\sqrt{\text{vol}_2(G)}g(n))$, where $g(n)$ is any slowly growing function as $n \to \infty$.

This is somewhat misleading however, as there is no real continuity of the sequence of the graph needed for the result to hold.

The most correct, and most confusing, way of stating Theorem 3 is as follows:

**Theorem 6.** Fix $\epsilon, \delta$, and functions $f(n)$ and $g(n)$ such that $f(n) = o(\frac{\tilde{d}}{\sigma})$ and $g(n)$ tends to infinity as $n \to \infty$. Then there exists an $N = N(\epsilon, \delta, f, g)$ such that if $G$ is a graph on $n$ vertices satisfying $n > N$ with $\Delta \leq f(n)$ and $p = \frac{1 - \epsilon}{d}$ then with probability at least $1 - \delta$ every connected component in $G_p$ has volume at most $\frac{\sqrt{\text{vol}_2(G)}}{g(n)}$.

Equivalently this could be written, instead of in terms of fixing $\delta$, in terms of fixing $n$ and letting $\delta = \delta(\epsilon, f(n), g(n), n)$. In any event, the statements of Theorems 3 and 4 employ rather standard abuses of notation which greatly simplifies the statement of the results.

The chapter is organized as follows: In Section 2 we introduce the notation and some basic facts. In Section 3, we examine several spectral lemmas which allow us to control the expansion. In Section 4, we prove Theorem 3, and in Section 5, we complete the proof of Theorem 4.
2.2 Preliminaries

Suppose $G$ is a connected graph on vertex set $V$. Throughout this chapter, $G_p$ denotes a random subgraph of $G$ obtained by retaining each edge of $G$ independently with probability $p$.

We shall let $\Delta = \max_v d_v$ denote the maximum degree of $G$ and $\delta = \min_v d_v$ denote the minimum degree. Recall that for the normalized Laplacian

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

with eigenvalues

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$$

the spectral gap is

$$\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}.$$

We also denote by $\rho$ the largest eigenvalue of the adjacency matrix.

The following lemma measures the difference of adjacency eigenvalue and $\bar{d}$ using $\sigma$.

Lemma 1. The largest eigenvalue of the adjacency matrix of $G$, $\rho$, satisfies

$$|\rho - \bar{d}| \leq \sigma \Delta.$$

Proof. Recall that $\varphi = \frac{1}{\sqrt{\text{vol}(G)}} D^{1/2} \mathbf{1}$ is the unit eigenvector of $\mathcal{L}$ corresponding to eigenvalue 0. Letting $d = D \mathbf{1}$, We have

$$\|I - \mathcal{L} - \varphi \varphi^*\| \leq \sigma.$$

Then,

$$|\rho - \bar{d}| = \left\| A - \frac{dd^*}{\text{vol}(G)} \right\|$$

$$\leq \left\| A - \frac{dd^*}{\text{vol}(G)} \right\|$$

$$= \|D^{1/2}(I - \mathcal{L} - \varphi \varphi^*)D^{1/2}\|$$

$$\leq \|D^{1/2}\| \cdot \|I - \mathcal{L} - \varphi \varphi^*\| \cdot \|D^{1/2}\|$$

$$\leq \sigma \Delta.$$
Note that, in light of Lemma 1, our result is in line with the result of Bollobás and the other results mentioned above in the sense that the threshold given in Theorems 3 and 4 is of the form $\frac{1}{\rho}$ where $\rho$ is the largest eigenvalue of the adjacency matrix. As we observed with the case of $\mathbb{Z}^2$ this should not be true in general, and it is a fascinating question as to over how wide of a class of graphs this type of bound should hold.

For any subset of the vertices, $S$, we let $\bar{S}$ denote the complement set of $S$. The vertex boundary of $S$ in $G$, denoted by $\Gamma^G(S)$ is defined as follows:

$$
\Gamma^G(S) = \{ u \notin S \mid \exists v \in S \text{ such that } \{u, v\} \in E(G) \}.
$$

When $S$ consists of one vertex $v$, we simply write $\Gamma^G(v)$ for $\Gamma^G(\{v\})$. We also write $\Gamma(S) = \Gamma^G(S)$ if there is no confusion.

Similarly, we define $\Gamma^{G_p}(S)$ to be the set of neighbors of $S$ in our percolated subgraph $G_p$.

### 2.3 Several spectral lemmas

We begin by proving two lemmas, first relating expansion in $G$ to the spectrum of $G$, then giving a probabilistic bound on the expansion in $G_p$. Note that the strategy of proof of Lemma 2 is similar in spirit (though more complicated) than that of the Expander Mixing Lemma in the introduction.
Lemma 2. For two disjoint sets $S$ and $T$, we have

\[
\left| \sum_{v \in T} d_v |\Gamma(v) \cap S| - \frac{\text{vol}(S)\text{vol}_2(T)}{\text{vol}(G)} \right| \leq \sigma \sqrt{\text{vol}(S)\text{vol}_3(T)}.
\]

\[
\left| \sum_{v \in T} d_v |\Gamma(v) \cap S|^2 - \frac{\text{vol}(S)^2\text{vol}_3(T)}{\text{vol}(G)^2} \right| \leq \sigma^2 \text{vol}(S) \max_{v \in T} \{d_v^2\} + 2\sigma \sqrt{\text{vol}(S)^3\text{vol}_5(T)\text{vol}(G)^2}.
\]

Proof. Let $1_S$ (or $1_T$) be the indicative column vector of the set $S$ (or $T$) respectively, and $d = D1$ as before. Note

\[
\sum_{v \in T} d_v |\Gamma(v) \cap S| = 1_S^* AD1_T.
\]

\[
\text{vol}(S) = 1_S^* d.
\]

\[
\text{vol}_2(T) = d^* D1_T.
\]

Here $1_S^*$ denotes the transpose of $1_S$ as a row vector. We have

\[
\left| \sum_{v \in T} d_v |\Gamma(v) \cap S| - \frac{\text{vol}(S)\text{vol}_2(T)}{\text{vol}(G)} \right|
\]

\[
= |1_S^* AD1_T - \frac{1}{\text{vol}(G)}1_S^* dd^* D1_T|
\]

\[
= |1_S^* D^\frac{1}{2}(D^{-\frac{1}{2}}AD^{-\frac{1}{2}} - \frac{1}{\text{vol}(G)}D^\frac{1}{2}11^* D^\frac{1}{2})D^\frac{1}{2}1_T|
\]

Let $\varphi = \frac{1}{\sqrt{\text{vol}(G)}}D^{1/2}1$ denote the eigenvector of $I - \mathcal{L}$ for the eigenvalue 1. The matrix $I - \mathcal{L} - \varphi \varphi^*$, which is the projection of $I - \mathcal{L}$ to the hyperspace $\varphi^\perp$, has $L_2$-norm $\sigma$.

We have

\[
\sum_{v \in T} d_v |\Gamma(v) \cap S| - \frac{\text{vol}(S)\text{vol}_2(T)}{\text{vol}(G)}
\]

\[
= |1_S^* D^\frac{1}{2}(I - \mathcal{L} - \varphi \varphi^*)D^\frac{3}{2}1_T|
\]

\[
\leq \sigma \|D^\frac{1}{2}1_S\| \cdot \|D^\frac{3}{2}1_T\|
\]

\[
\leq \sigma \sqrt{\text{vol}(S)\text{vol}_3(T)}.
\]
Let $e_v$ be the column vector with $v$-th coordinate 1 and 0 elsewhere. Then $|\Gamma^G(v) \cap S| = 1_S^*Ae_v$. We have

$$\sum_{v \in T} d_v |\Gamma^G(v) \cap S|^2 = \sum_{v \in T} d_v 1_S^*Ae_v e_v^*A1_S = 1_S^*AD_TA1_S.$$ 

Here $D_T = \sum_{v \in T} d_v e_v e_v^*$ is the diagonal matrix with degree entry at vertex in $T$ and 0 elsewhere. We have

$$\left| \sum_{v \in T} d_v |\Gamma^G(v) \cap S|^2 - \frac{\text{vol}(S)^2 \text{vol}_3(T)}{\text{vol}(G)^2} \right|$$

$$= |1_S^*AD_TA1_S - \frac{1}{\text{vol}(G)^2} 1_S^*dd^*D_Tdd^*1_S|$$

$$\leq |1_S^*AD_TA1_S - \frac{1}{\text{vol}(G)} 1_S^*dd^*D_TA1_S|$$

$$+ |\frac{1}{\text{vol}(G)} 1_S^*dd^*D_TA1_S - \frac{1}{\text{vol}(G)^2} 1_S^*dd^*D_Tdd^*1_S|$$

$$= |1_SD^{\frac{1}{2}}(I - \mathcal{L} - \mathcal{A}^*\mathcal{A}) D^{\frac{1}{2}}D_TA1_S|$$

$$+ |\frac{1}{\text{vol}(G)} 1_S^*dd^*D_TD^{\frac{1}{2}}(I - \mathcal{L} - \mathcal{A}^*\mathcal{A}) D^{\frac{1}{2}}1_S|$$

$$\leq |1_SD^{\frac{1}{2}}(I - \mathcal{L} - \mathcal{A}^*\mathcal{A}) D^{\frac{1}{2}}D_TD^{\frac{1}{2}}(I - \mathcal{L} - \mathcal{A}^*\mathcal{A}) D^{\frac{1}{2}}1_S|$$

$$+ 2|\frac{1}{\text{vol}(G)} 1_S^*dd^*D_TD^{\frac{1}{2}}(I - \mathcal{L} - \mathcal{A}^*\mathcal{A}) D^{\frac{1}{2}}1_S|$$

$$\leq \sigma^2 \text{vol}(S) \max_{v \in T} \{d_v^2\} + 2\sigma \sqrt{\text{vol}(S)^3 \text{vol}_5(T)} \frac{\text{vol}(T)}{\text{vol}(G)}.$$

The worthwhile comparison to the above Lemma is the type of discrepancy bound given by the expander mixing lemma. For an irregular graph the expander mixing lemma (see Chapter 1) implies

$$\left| e(X,Y) - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)} \right| \leq \sigma \sqrt{\text{vol}(X)\text{vol}(Y)}.$$

Taking $Y = X$ in the above formula provides a bound on the number of edges leaving a set $X$, which provides us a mechanism to guarantee expansion. We can
use this to show expansion in $G_p$, but unfortunately this does not tell us enough about the set we expand into to continue the process. Lemma 2 provides tighter bounds on not just the number of edges leaving $X$ but also on the volume of the neighborhood of $X$ which provides necessary information to control expansion for a number of steps.

We now prove a lemma concerning the actual growth within $G_p$, using Lemma 2 to understand the expansion within $G$. Note that Lemma 3 provides a bound on the volume of the neighborhood of a set in the percolated graph, $G_p$, not the underlying host graph. Also note that it allows growth into a particular set of vertices $T$ (which, in our case, will be unexplored vertices so as to avoid dependency issues) instead of the entirety of the host graph.

**Lemma 3.** Suppose that two disjoint sets $S$ and $T$ satisfy

\[
\begin{align*}
\text{vol}_2(T) & \geq \frac{5p\sigma^2 \max_{v \in T} \{d_v^2\}}{2\delta} \text{vol}(G) \quad \text{(2.1)} \\
\frac{25\sigma^2 \text{vol}_3(T) \text{vol}(G)^2}{\delta^2 \text{vol}_2(T)^2} & \leq \text{vol}(S) \leq \frac{2\delta \text{vol}_2(T) \text{vol}(G)}{5p \text{vol}_3(T)} \quad \text{(2.2)} \\
\text{vol}(S) & \leq \frac{\delta^2 \text{vol}_2(T)^2}{25p^2 \sigma^2 \text{vol}_5(T)} \quad \text{(2.3)}
\end{align*}
\]

Then we have that

\[
\text{vol}(\Gamma_{G_p}^G(S) \cap T) > (1 - \delta)p \frac{\text{vol}_2(T)}{\text{vol}(G)} \text{vol}(S).
\]

with probability at least $1 - \exp \left( -\frac{\delta(1 - \delta)p \text{vol}_2(T) \text{vol}(S)}{10 \Delta \text{vol}(G)} \right)$.

**Proof.** For any $v \in T$, let $X_v$ be the indicative random variable for $v \in \Gamma_{G_p}^G(S)$. We have

\[
\mathbb{P}(X_v = 1) = 1 - (1 - p)^{|\Gamma_v \cap S|}.
\]

Let $X = |\Gamma_{G_p}^G(S) \cap T|$. Then $X$ is the sum of independent random variables $X_v$.

\[
X = \sum_{v \in T} d_v X_v.
\]
Note that
\[\mathbb{E}(X) = \sum_{v \in T} d_v \mathbb{E}(X_v)\]
\[= \sum_{v \in T} d_v (1 - (1 - p)^{\Gamma^G(v) \cap S})\]
\[\geq \sum_{v \in T} d_v (p|\Gamma^G(v) \cap S| - \frac{p^2}{2} |\Gamma^G(v) \cap S|^2)\]
\[\geq p \left( \frac{\text{vol}(S) \text{vol}_2(T)}{\text{vol}(G)} - \sigma \sqrt{\text{vol}(S) \text{vol}_3(T)} \right)\]
\[-\frac{p^2}{2} \left( \frac{\text{vol}(S)^2 \text{vol}_3(T)}{\text{vol}(G)^2} + \sigma^2 \text{vol}(S) \max_{v \in T} \{d_v^2\} + 2\sigma \sqrt{\text{vol}(S)^3 \text{vol}_5(T) / \text{vol}(G)} \right)\]
\[> (1 - \frac{4}{5} \delta) p \frac{\text{vol}_2(T)}{\text{vol}(G)} \text{vol}(S)\]
by using Lemma 2 and the assumptions on $S$ and $T$.

We apply the following Chernoff inequality, Proposition 4 in the introduction,
\[\mathbb{P}(X \leq \mathbb{E}(X) - a) \leq e^{-\frac{a^2}{2\Delta \mathbb{E}(X)}}.\]

We set $a = \alpha \mathbb{E}(X)$, with $\alpha$ chosen so that $(1 - \alpha)(1 - \frac{4}{5} \delta) = (1 - \delta)$. Then
\[\mathbb{P}(X \leq (1 - \delta) p \frac{\text{vol}_2(T)}{\text{vol}(G)} \text{vol}(S)) < \mathbb{P}(X \leq (1 - \alpha) \mathbb{E}(X))\]
\[\leq \exp \left( -\frac{\alpha^2 \mathbb{E}(X)}{2\Delta} \right)\]
\[< \exp \left( -\frac{\alpha (1 - \delta) p \text{vol}_2(T) \text{vol}(S)}{2\Delta \text{vol}(G)} \right).\]

To complete the proof, note $\alpha > \delta / 5$. \ hfill \qed

### 2.4 The range of $p$ with no giant component

In this section, we will prove Theorem 3.

**Proof of Theorem 3.** It suffices to prove the following claim:
Claim A: If $p\rho < 1$, where $\rho$ is the largest eigenvalue of the adjacency matrix, with probability at least $1 - \frac{1}{C^2(1-pp)}$, all components have volume at most $C\sqrt{\text{vol}_2(G)}$.

Proof of Claim A. Let $x$ be the probability that there is a component of $G_p$ having volume greater than $C\sqrt{\text{vol}_2(G)}$. Now we choose two random vertices with the probability of being chosen proportional to their degrees in $G$. Under the condition that there is a component with volume greater than $C\sqrt{\text{vol}_2(G)}$, the probability of each vertex in this component is at least $\frac{C\sqrt{\text{vol}_2(G)}}{\text{vol}(G)}$. Therefore, the probability that the random pair of vertices are in the same component is at least

$$x \left( \frac{C\sqrt{\text{vol}_2(G)}}{\text{vol}(G)} \right)^2 = \frac{C^2x\tilde{d}}{\text{vol}(G)}.$$

(2.4)

On the other hand, for any fixed pair of vertices $u$ and $v$ and any fixed path $P$ of length $k$ in $G$, the probability that $u$ and $v$ is connected by this path in $G_p$ is exactly $p^k$. The number of $k$-paths from $u$ to $v$ is at most $\frac{d_u}{\text{vol}(G)}A^k1_v$. Since the probabilities of $u$ and $v$ being selected are $\frac{d_u}{\text{vol}(G)}$ and $\frac{d_v}{\text{vol}(G)}$ respectively, the probability that the random pair of vertices are in the same connected component is at most

$$\sum_{u,v} \frac{d_u}{\text{vol}(G)} \frac{d_v}{\text{vol}(G)} \sum_{k=0}^{n} p^k 1_u A^k 1_v = \sum_{k=0}^{n} \frac{1}{\text{vol}(G)^2} p^k \text{d}^* A^k \text{d}.$$

We have

$$\sum_{k=0}^{n} \frac{1}{\text{vol}(G)^2} p^k \text{d}^* A^k \text{d} \leq \sum_{k=0}^{\infty} \frac{p^k \rho^k \text{vol}_2(G)}{\text{vol}(G)^2} \leq \frac{\tilde{d}}{(1-pp)\text{vol}(G)}.$$

Combining with (2.4), we have $\frac{C^2x\tilde{d}}{\text{vol}(G)} \leq \frac{\tilde{d}}{(1-pp)\text{vol}(G)}$, which implies $x \leq \frac{1}{C^2(1-pp)}$. \qed

Claim A is proved, and the theorem follows taking $C$ to be $g(n)$. \quad \Box

Note that we actually proved a slightly stronger theorem than Theorem 3. In particular, the hypothesis that $\Delta = o\left(\frac{\tilde{d}}{\rho}\right)$ is only necessary so that Lemma 1 implies that $\rho = (1 + o(1))\tilde{d}$; in general this establishes a lower bound on the giant component threshold of $\frac{1}{\rho}$.
It is also somewhat instructive to give an example of where the threshold is below \( \tilde{d} \). In particular, we construct an example here where percolating by taking \( p = \frac{1-\epsilon}{\tilde{d}} \) still results in a giant component for some small (but positive) epsilon.

**Example 1.** Consider a graph \( G \) consisting of \( K\sqrt{n} \) and \( \sqrt{n} \) disjoint 4 regular graphs on \( \sqrt{n} \) vertices, such that each vertex in the \( K\sqrt{n} \) is connected to all vertices in one of the other components. Then

\[
\text{vol}(G) = (2\sqrt{n} - 1)\sqrt{n} + 5n = 7n - \sqrt{n} \geq 6n
\]

and

\[
\text{vol}_2(G) = (2\sqrt{n} - 1)^2\sqrt{n} + 25n \leq 5n^{3/2}.
\]

for \( n \) sufficiently large. Then

\[
\tilde{d} \leq \frac{5}{6}n^{1/2}
\]

so \( \frac{1}{d} \leq \frac{6}{5}n^{-1/2} \) (for \( n \) sufficiently large.) Let \( p = 7/6n^{-1/2} \) so \( p\tilde{d} = \frac{35}{36} < 1 \). Percolate \( G \) with probability \( p \). With probability \( 1-o(1) \), (via the Chernoff bounds, Proposition 3 applies) the total volume of the remaining graph is \( \frac{49}{6}\sqrt{n} + O(\sqrt{n}\log n) \).

On the other hand; note that within the \( K\sqrt{n} \), percolation is just considering an Erdős-Renyi random graph with \( n = \sqrt{n} \) and \( p = 7/6n^{-1/2} \). Since \( np > 1 \), there is a connected component within the \( K\sqrt{n} \) of size \( \geq \delta \sqrt{n} \) with probability \( 1 - o(1) \).

But the expected volume of this component is at least \( \frac{14}{6}\delta \sqrt{n} \), and by the Chernoff bounds, Proposition 3, the total volume within this component is at least \( \epsilon \sqrt{n} \) for some \( \epsilon > 0 \) a.a.s., demonstrating that although \( p < \frac{1}{d} \) there can still be a giant component.

### 2.5 The emergence of the giant component

As an initial step, we first make an observation as to some properties of a graphs containing \((\epsilon, M)\)-admissible sets which are important to the proof of Theorem 4.

**Lemma 4.** Suppose \( G \) contains an \((\epsilon, M)\)-admissible set \( U \). Then we have

1. \( \tilde{d} \leq \frac{M}{(1-\epsilon)^2} d \).
2. For any $U' \subset U$ with $\text{vol}_2(U') > \eta \text{vol}_2(U)$, we have

$$\text{vol}(U') \geq \frac{\eta^2(1-\epsilon)\tilde{d}}{Md} \text{vol}(G).$$

Proof. Since $G$ is $(\epsilon, M)$-admissible, we have a set $U$ satisfying

(i) $\text{vol}_2(U) \geq (1-\epsilon)\text{vol}_2(G)$

(ii) $\text{vol}_3(U) \leq M\text{vol}_2(G)$.

We have

$$\tilde{d} = \frac{\text{vol}_2(G)}{\text{vol}(G)} \leq \frac{\text{vol}_2(G)}{\text{vol}(U)} \leq \frac{1}{1-\epsilon} \frac{\text{vol}_2(U)}{\text{vol}(U)} \leq \frac{1}{(1-\epsilon)^2} \frac{\text{vol}_2(U)}{\text{vol}_2(U')}$$

For any $U' \subset U$ with $\text{vol}_2(U') > \eta \text{vol}_2(U)$, we have

$$\text{vol}(U') \geq \frac{\text{vol}_2(U')^2}{\text{vol}_3(U')} \geq \frac{\eta^2 \text{vol}_2(U)^2}{\text{vol}_3(U)} \geq \frac{\eta^2(1-\epsilon)\text{vol}_2(G)}{Md} \geq \frac{\eta^2(1-\epsilon)\tilde{d}}{Md} \text{vol}(G).$$

Before proving Theorem 4, we make some comment as to the methods and motivation of the proof. In Erdős and Réyni’s original paper [ER59], the existence of the giant component was shown, roughly speaking, by a counting method. The
expected number of components of size \( k \) was computed, and the expected number of components of size \( C\log(n) \leq k \leq \epsilon n \) was shown to tend to zero. An additional argument showed that not too many vertices lay in components of size \( k \), hence proving the existence of a giant component. This is the same general method used by Frieze, Krivelevich and Martin in [FKM04]. Here, the authors use the pseudo-randomness of the graph to insure that components in some size range would necessarily expand. The regularity of a graph, enabled the authors to get a good estimate on the number of trees of a given size in the graph. As survival of a tree within a set \( X \) is necessary for \( X \) to be connected, this enabled them to show a gap in component sizes.

A more modern method of showing the existence of giant components is via branching processes. Janson, Luczak and Rucinkski’s book [JLR00] gives a good treatment of the birth of the giant component in \( G_{n,p} \) using this method. The gist of the method is that the neighborhoods of a vertex in \( G_{n,p} \) when \( p > \frac{1+\epsilon}{n} \) can be viewed as a super critical branching process, and one of the branching processes is likely to survive. The regularity of the underlying graph here plays a large role in being able to analyze the survival of these branching processes. In our case, while we have structural information on expansion from the spectrum it is not strong enough to understand the survival of a branching process started from a particular vertex in our graph.

Instead, we proceed by strapping together many branching processes. We begin with a set of vertices with our graph of fairly large volume (large enough so that we may apply Lemma 3). We repeatedly take neighborhoods of that in \( G_p \) until the volume of the resulting set is giant (and too large to apply Lemma 3). At this point we do not necessarily have a giant component, as our set need not be connected, but we can then take the largest components of our large set, and repeat from there. This decreases the number of components we are growing from. After repeating some number of times, where this is a function of \( \log(\sigma) / \log(n) \), we actually will have a set which is connected and we are able to apply Lemma 3. We can then perform a further growing step to show the existence of a giant component within \( G_p \).
Proof of Theorem 4. It suffices to assume \( p = \frac{1+c}{d} \) for some \( c < \frac{1}{25} \).

Let \( \epsilon = \frac{c}{10} \) be a small constant, and \( U \) be a \((\epsilon, M)\)-admissible set in \( G \). Define \( U' \) to be the subset of \( U \) containing all vertices with degree at least \( \sqrt{\epsilon d} \). We have

\[
\text{vol}_2(U') \geq \text{vol}_2(U) - \sum_{d_v < \sqrt{\epsilon d}} d_v^2 \\
\geq (1 - \epsilon)\text{vol}_2(G) - cnd^2 \\
\geq (1 - 2\epsilon)\text{vol}_2(G).
\]

Hence, \( U' \) is a \((2\epsilon, M)\) admissible set. We will concentrate on the neighborhood expansion within \( U' \).

Let

\[
\delta = \frac{c}{2}
\]

and

\[
C = \frac{25M}{\delta^2(1 - 4\epsilon)^2}.
\]

Take an initial set \( S_0 \subset U' \) satisfying

\[
\max(C\sigma^2\text{vol}(G), \Delta \ln n) \leq \text{vol}(S_0) \leq \max(C\sigma^2\text{vol}(G), \Delta \ln n) + \Delta.
\]

Let \( T_0 = U' \setminus S_0 \). For \( i \geq 1 \), we will recursively define

\[
S_i = \Gamma^G_{\delta}(S_{i-1}) \cap U'
\]

and

\[
T_i = U' \setminus \bigcup_{j=0}^{i} S_j
\]

until

\[
\text{vol}_2(T_i) \leq (1 - 3\epsilon)\text{vol}_2(G) \quad \text{or} \quad \text{vol}(S_i) \geq \frac{2\delta\text{vol}_2(T_i)\text{vol}(G)}{5p\text{vol}_3(T_i)}.
\]

Condition 1 in Lemma 3 is always satisfied, to verify this we observe that

\[
\frac{5p}{2\delta} \sigma^2 \max_{v \in T_i} d_v^2 \text{vol}(G) \leq \frac{5(1+c)}{2d\delta} \sigma^2 \Delta^2 \text{vol}(G) \\
= \left(\frac{\sigma \Delta}{d}\right)^2 \frac{5(1+c)}{2\delta} \text{vol}_2(G) \\
= o(\text{vol}_2(G)).
\]
Thus, assuming $n$ is sufficiently large

\[
\frac{5p}{2\delta} \sigma^2 \max_{v \in T_i} d_v^2 \mathrm{vol}(G) \leq \mathrm{vol}_2(T_i).
\]

Condition 3 in Lemma 3 is also trivial because

\[
\frac{\delta^2 \mathrm{vol}_2(T_i)^2}{25p^2 \sigma^2 \mathrm{vol}_5(T_i)} \geq \frac{\delta^2 \mathrm{vol}_2(T_i)^2}{25p^2 \sigma^2 \Delta^2 \mathrm{vol}_3(T_i)} \geq \frac{\delta^2(1 - 3\epsilon) \mathrm{vol}_2(G)}{25p^2 \sigma^2 \Delta^2 M d} \geq \left(\frac{d}{\sigma \Delta}\right)^2 \frac{\delta^2(1 - 3\epsilon)}{25(1 + c)^2 M} \mathrm{vol}(G) = \omega(\mathrm{vol}(G)).
\]

Now we verify condition 2. We have

\[
\mathrm{vol}(S_0) > C\sigma^2 \mathrm{vol}(G) = \frac{25M}{\delta^2(1 - 4\epsilon)^2} \sigma^2 \mathrm{vol}(G) \geq \frac{25\sigma^2 \mathrm{vol}_3(T_0) \mathrm{vol}(G)^2}{\delta^2 \mathrm{vol}_2(T_0)^2}.
\]

The conditions of Lemma 3 are all satisfied. Then we have that

\[
\mathrm{vol}(\Gamma^G_p(S_0) \cap T_0) \geq (1 - \delta)p \frac{\mathrm{vol}_2(T_0)}{\mathrm{vol}(G)} \mathrm{vol}(S_0).
\]

with probability at least $1 - \exp\left(-\frac{\delta(1 - \delta) p \mathrm{vol}_2(T_0) \mathrm{vol}(S_0)}{10\Delta \mathrm{vol}(G)}\right)$.

Since $(1 - \delta)p \frac{\mathrm{vol}_2(T_0)}{\mathrm{vol}(G)} \geq (1 - \delta)(1 - 3\epsilon)(1 + c) = \beta > 1$ by our assumption that $c$ is small (noting that $\epsilon$ and $\delta$ are functions of $c$), the neighborhood of $S_i$ grows exponentially, allowing condition 2 of Lemma 3 to continue to hold and us to continue the process. We stop when one of the following two events happens,

- $\mathrm{vol}(S_i) \geq \frac{2\delta \mathrm{vol}_2(T_i) \mathrm{vol}(G)}{5p \mathrm{vol}_3(T_i)}$.
- $\mathrm{vol}_2(T_i) \leq (1 - 3\epsilon) \mathrm{vol}_2(G)$.

Let us denote the time that this happens by $t$.

If the first, but not the second, case occurs we have

\[
\mathrm{vol}(S_i) \geq \frac{2\delta \mathrm{vol}_2(T_i) \mathrm{vol}(G)}{5p \mathrm{vol}_3(T_i)} \geq \frac{2\delta(1 - 3\epsilon)}{5M(1 + c)} \mathrm{vol}(G).
\]
In the second case, we have

\[
\text{vol}_2(\bigcup_{j=0}^{t} S_j) = \text{vol}_2(U') - \text{vol}_2(T_t) \\
\geq \epsilon \text{vol}_2(G) \geq \epsilon \text{vol}(U').
\]

By Lemma 4 with \(\eta = \epsilon\), we have

\[
\text{vol}(\bigcup_{j=0}^{t} S_j) \geq \frac{\epsilon^2(1-2\epsilon)\bar{d}}{Md} \text{vol}(G).
\]

On the other hand, note that since

\[
\text{vol}(S_i) \geq \beta \text{vol}(S_{i-1}),
\]

we have that

\[
\text{vol}(S_i) \leq \beta^{i-t} \text{vol}(S_t),
\]

and hence that

\[
\text{vol}(\bigcup_{j=0}^{t} S_j) \leq \sum_{j=0}^{t} \beta^{-j} \text{vol}(S_t)
\]

which implies

\[
\text{vol}(S_t) \geq \frac{\epsilon^2(1-2\epsilon)\bar{d}(\beta - 1)}{Md\beta} \text{vol}(G).
\]

In either case we have \(\text{vol}(S_t) = \Theta(\text{vol}(G))\). For the moment, we restrict ourselves to the case where \(C\sigma^2 n > \Delta \ln n\).

Each vertex in \(S_t\) is in the same component as some vertex in \(S_0\), which has size at most \(\frac{\text{vol}(S_0)}{\sqrt{\bar{d}}} \leq C'\sigma^2 n\). We now combine the \(k_1\) largest components to form a set \(W^{(1)}\) with \(\text{vol}(W^{(1)}) > C\sigma^2 \text{vol}(G)\), such that \(k_1\) is minimal. If \(k_1 \geq 2\),

\[
\text{vol}(W^{(1)}) \leq 2C\sigma^2 \text{vol}(G).
\]

Note that since the average size of a component is \(\frac{\text{vol}(S_t)}{|S_0|} \geq C_1 \frac{\text{vol}(G)}{\sigma^2 n}\),

\[
k_1 \leq C'\sigma^4 n.
\]
We grow as before: Let $W_0^{(1)} = W^{(1)}$, $Q_0^{(1)} = T_{t-1} \setminus W_0^{(1)}$. Note that the conditions for Lemma 3 are satisfied by $W_0^{(1)}$ and $T_0^{(1)}$. We run the process as before, setting $W_t^{(1)} = \Gamma(W_t^{(1)}) \cap Q_t^{(1)}$ and $Q_t^{(1)} = Q_{t-1}^{(1)} \setminus W_t^{(1)}$ stopping when either

$$\text{vol}(Q_t^{(1)}) < (1 - 4\epsilon)\text{vol}_2(G)$$

or

$$\text{vol}(W_t^{(1)}) > \frac{2\delta \text{vol}_2(Q_t^{(1)}) \text{vol}(G)}{5p \text{vol}_3(Q_t^{(1)})} \geq \frac{2\delta(1 - 4\epsilon)}{5M(1 + \epsilon)} \text{vol}(G).$$

As before, in either case $\text{vol}(W_t^{(1)}) = \Theta(\text{vol}(G))$.

Note that if $k_1 = 1$, we are now done as all vertices in $W_t^{(1)}$ lie in the same component of $G_p$.

Now we iterate. Each of the vertices in $W_t^{(1)}$ lies in one of the $k_1$ components of $W_0^{(1)}$. We combine the largest $k_2$ components to form a set $W^{(2)}$ of size at least $C\sigma^2\text{vol}(G)$. If $k_2 = 1$, then one more growth finishes us, otherwise

$$\text{vol}(W^{(2)}) < 2C\sigma^2\text{vol}(G),$$

the average size of components is at least

$$C_2 \frac{\text{vol}(G)}{\sigma^4n}$$

and hence

$$k_2 \leq C'_2 \sigma^6 n.$$

We iterate, growing $W^{(m)}$ until either

$$\text{vol}(Q_t^{(m)}) < (1 - (m + 3)\epsilon)\text{vol}_2(G) \quad \text{or} \quad \text{vol}(W_t^{(m)}) > \frac{2\delta \text{vol}_2(Q_t^{(m)}) \text{vol}(G)}{5p \text{vol}_3(Q_t^{(m)})},$$

so that $W_t^{(m)}$ has volume $\theta(\text{vol}(G))$ and then creating $W^{(m+1)}$ by combining the largest $k_{m+1}$ components to form a $W^{(m+1)}$ with volume at least $C\sigma^2 n$. Once $k_m = 1$ for some $m$ all vertices in $W^{(m)}$ are in the same component and one more growth round finishes the process, resulting in a giant component in $G$. Note that the average size of a component in $W_n^{(m)}$ has size at least $C_m \frac{\text{vol}(G)}{\sigma^2(m+1)n}$ (that is,
components must grow by a factor of at least $\frac{1}{2^m}$ each iteration) and if $k_m > 1$, we must have $k_m \leq C'_m \sigma^{2(m+1)n}$. If $m = \lceil \frac{1}{2m} \rceil - 1$, this would imply that $k_m = o(1)$ by our condition $\sigma = o(n^{-\kappa})$, so after at most $\lceil \frac{1}{2m} \rceil - 1$ rounds, we must have $k_m = 1$ and the process will halt with a giant connected component.

In the case where $\Delta \ln n > C\sigma^2 n$, we note that

$$|S_0| \leq \frac{\text{vol}(S_0)}{\sqrt{ed}} \leq C' \frac{\Delta \ln n}{\sqrt{ed}},$$

and the average volume of components in $S_t$ is at least

$$\frac{C'' \text{vol}(G)d}{\Delta \ln n} = \omega(\Delta \ln n),$$

so we can form $W^{(1)}$ by taking just one component for $n$ large enough, and the proof goes as above.

We note that throughout, if we try to expand we have that

$$\text{vol}(Q_{t}^{(m)}) > (1 - (m + 3)\varepsilon) \text{vol}_2(G)$$

$$> \left(1 - \left(\frac{1}{2\kappa} + 4\right) \varepsilon\right) \text{vol}_2(G)$$

$$> \left(1 - \frac{9c}{20}\right) \text{vol}_2(G).$$

By our choice of $c$ being sufficiently small, $(1 - (m + 3)\varepsilon)(1 - \delta)(1 + c) > 1$ at all times so throughout, noting that $\text{vol}(S_t)$ and $\text{vol}(W_i^{(m)})$ are at least $\Delta \ln n$, we are guaranteed our exponential growth by Lemma 3 with an error probability bounded by

$$\exp\left(-\frac{\delta(1 - \delta)p\text{vol}_2(T)\text{vol}(S_t)}{10\Delta \text{vol}(G)}\right) \leq \exp\left(-\frac{\delta(1 - \delta)(1 - \frac{9c}{20})\text{vol}(S_t)}{2\Delta}\right) \leq n^{-K}.$$

We run for a constant number of phases, and run for at most a logarithmic number of steps in each growth phase as the sets grow exponentially. Thus, the probability of failure is at most $C'' \log(n)n^{-K} = o(1)$ for some constant $C''$, thus completing our argument that $G_p$ contains a giant component with high probability.

Finally, we prove the uniqueness assertion. With probability $1 - C'' \log(n)n^{-K}$ there is a giant component $X$. Let $u$ be chosen at random; we estimate the probability that $u$ is in a component of volume at least $\max(2d \log n, \omega(\sigma \sqrt{\text{vol}(G)}))$. 
Let $Y$ be the component of $u$. Theorem 5.1 of [Chu97] asserts that if

$$\text{vol}(Y) \geq \max(2d \log n, \omega(\sigma \sqrt{\text{vol}(G)})),$$

then

$$e(X,Y) \geq \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)} e(X,Y) - \sigma \sqrt{\text{vol}(X)\text{vol}(Y)} \geq 1.5d \log n$$

Note that the probability that $Y$ is not connected to $X$ given that $\text{vol}(Y) = \omega(\sigma \sqrt{\text{vol}(G)})$ is $(1 - p)^{e(X,Y)} = o(n^{-1})$, so with probability $1 - o(1)$ no vertices are in such a component - proving the uniqueness of large components.

\begin{proof}
\end{proof}

### 2.6 Interesting Questions

While the above results illustrate the emergence of the giant component on a much larger class of irregular graphs than previously possible, many interesting questions remain on this problem. Perhaps one of the most obvious areas for improvement is in weakening the conditions of Theorem 3. An interesting question is

**Question 1.** Under what conditions is $\frac{1}{\rho}$, where $\rho$ is the largest eigenvalue of the adjacency matrix, the giant component threshold?

As we have seen from the above discussion as well as this result, it holds for a rather large set of graphs, however percolation on the integer grid $\mathbb{Z}^2$ indicates that this should not always be the right example. A key element in establishing the existence of the giant component is to understand expansion in the underlying host graph. It seems likely that graphs which have thresholds larger than $\frac{1}{\rho}$ will have poor expansion, and in particular will have many short loops. Since random (and pseudo-) graphs have a small number of short cycles at most vertices, most paths leaving any particular vertex are different and not repeating edges and hence force the analysis in the proof of Theorem 3 to be largely tight; these are exactly the cases where we expect $\frac{1}{\rho}$ to actually be the threshold. Still, such heuristic reasoning is far from a precise conjecture or proof; but there is clearly room for further research.
Another interesting avenue for exploration has been explored recently in several papers, mainly studying percolation on special classes of graphs. These papers have gone further, to nail down the precise critical window during which component sizes grow from $\log(n)$ vertices to a positive proportion of the graph. The first such glimpse into the birth of the giant component was given by Erdős and Rényi, who showed that when $p = \frac{1}{n}$, then $G \in G_{n,p}$ contains a unique component of size $n^{2/3}$. This behavior, where when $p = \frac{1+\epsilon}{n}$ there all components are small, when $p = \frac{1}{n}$ there is a unique component of size $n^{2/3}$ and when $p = \frac{1+\epsilon}{n}$ there is a unique giant component has become known as the double jump. Bollobás [Bol84] later proved that, in fact, that component sizes grow from $\Theta(\log n)$ to $\Theta(n)$ in a window of order $O\left(n^{-4/3+o(1)}\right)$ about the critical threshold $\frac{1}{n}$. In 1990, Luczak [Luc90] removed the $o(1)$ in the previous result demonstrating that giant component arises in a window of order $\Theta(n^{-4/3})$.

In [BCvdH+05, BCvdH+06], Borgs, et. al. find the order of this critical window for transitive graphs, and cubes. Nachmias [Nac07] looks at a similar situation to that of Frieze, Krivelevich and Martin [FKM04] and uses random walk techniques to study percolation within the critical window for quasi-random transitive graphs. Percolation within the critical window on random regular graphs is also studied by Nachmias and Peres in [NP08]. Our results differ from these in that we study percolation on graphs with a much more general degree sequence. The greater preciseness of these results, however, is quite desirable.

**Question 2.** What is the size of the scaling window for the phase condition under the hypothesis of Theorem 4.

**Question 3.** What is the size of the largest component for a graph satisfying the hypothesis of Theorem 4 when $p = \frac{1}{d}$.

It should be noted that the answer to the last question was given in the case of $G_{n,p}$ (that is, for the host graph taken to be $K_n$) by Erdős and Rényi [ER59], long before Luczak determined the precise window. Also, for the case where the host graph is the infinite grid $G = \mathbb{Z}_2$, it is known that $p_c = \frac{1}{2}$, but it is also known that
if $p = \frac{1}{2}$, then $G_p$ contains no infinite components. Other host graphs are known, however, in the probability literature where percolation at the critical value yields an infinite component.

Acknowledgement: Part of this chapter is based on the paper “The giant component in a random subgraph of a given graph”, written jointly with Fan Chung and Linyuan Lu. This paper has appeared in Proceedings of WAW2009, Lecture Notes in Computer Science 5427, 38–49, 2009. Material appears with kind permission of Springer Science+Business Media.
Chapter 3

The Structure of Percolated Graphs

3.1 Introduction

Often, when we examine a large graph, perhaps arising from some realistic setting (e.g. webgraphs, biological networks, or some information network derived from a large database), we are unable to see the entire graph. Instead, what we can observe are relatively small subgraphs of the large graph. We are interested, then, in understanding the relationship between the large host graph, and the subgraphs that we actually observe. In this chapter, we are primarily concerned with understanding relationships between the structure of a random subgraph and a percolated subgraph.

In the last chapter, we found that the spectral gap enabled us to understand the structure of a host graph to such an extent that we were able to nail down the emergence of the giant component in percolated subgraphs. Our key observation was that the spectral gap, $\sigma$ provides a great deal of structural information about the underlying host graph. Thus an understanding of the spectral gap of a random subgraph of a graph allows us to draw a number of conclusions about the structure
of the subgraph.

Indeed, this idea goes further. In this chapter we again consider a random subgraph obtained by bond percolation. We consider the relationship between the spectral gap \( \sigma \) of host graph \( G \) and \( \sigma_H \) of a random subgraph \( H \). We prove that on \(| \sigma - \sigma_H |\) that, in particular, implies that if \( p\delta \gg \log^{3/2}(n) \) (recall here this means that \( p\delta = \omega(\log^{3/2}(n)) \) that \( |\sigma - \sigma_H| = o(1) \), where \( \delta \) denotes the minimum degree. Note that such a bound actually provides information in both directions; it both reveals information about the structure of a random subgraph if the spectral gap of the underlying host graph is known, as well as information about structure of the host graph based on information about the spectrum of the random subgraph.

As mentioned in the introduction, a major motivation for this work is the fact that we often only see a part of a larger network. If the reason for this is a corruption that affects edges roughly independently and with roughly the same probability, our result gives a way to understand the structure of the host graph.

We prove the following theorem:

**Theorem 7.** Suppose \( G \) is a graph on \( n \) vertices with spectral gap \( \sigma \) and minimum degree \( \delta \). A random subgraph \( H = G_p \) selected by bond percolation with probability \( p \) has \( \sigma_H \) satisfying

\[
|\sigma_H - \sigma| = O\left( \sqrt{\frac{\log n}{p\delta}} + \frac{\log^3 n}{p\delta \log \log n} \right).
\]

An equivalent statement for the above theorem is the following: For \( p\delta \geq \frac{(\log n)^2}{(\log \log n)^3} \), we have

\[
|\sigma_H - \sigma| = O\left( \sqrt{\frac{\log n}{p\delta}} \right)
\]

and, for \( p\delta < \frac{(\log n)^2}{(\log \log n)^3} \), we have

\[
|\sigma_H - \sigma| = O\left( \frac{\log^3 n}{p\delta \log \log n} \right).
\]

Note that under the condition that \( p\delta < (1 - \epsilon) \log(n) \), the random subgraph \( H \) need not be connected. In this case, \( \sigma_H = 0 \) and we have no hope of getting a
bound. On the other hand, our results provide actual information in the case that
\( p \delta \gg \frac{(\log n)^{3/2}}{(\log \log n)^{3/2}} \). It would be interesting to understand what occurs in the gap.

The methods that we used to prove Theorem 7 are based on Wigner’s high moment methods \([\text{Wig58}]\). Such an approach has been extensively utilized in the early work on random graphs and matrices in numerous research papers including the early work by Füredi and Komlós \([\text{FK81}]\) as well as in some recent work on random sparse graphs in \([\text{CLV04}]\) and \([\text{Vu07}]\). However, these previous works belong to the special case that the host graph is taken to be the complete graph (or the full matrix). Here similar techniques are used and modified in order to deal with the spectral gap of a random subgraph of a given host graph.

It would also be of interest to understand structural properties of random subgraphs directly, without dealing with the spectrum. This would be especially useful in cases where \( p \delta \) is small enough that \( \sigma_H \) may be zero. As examples for more specific host graphs, Nachmias and Peres \([\text{NP08}]\) have studied properties related to the spectrum, in particular diameter and mixing time, in percolated regular graphs. Similarly, Ofek \([\text{Ofe07}]\) has studied expansion in the giant component of percolated pseudorandom graphs; this is the case where Frieze, Krivelevich and Martin \([\text{FKM04}]\) established the existence of giant a component. Alon, Benjamini and Stacey \([\text{ABS04}]\) also studied isoperimetric properties of percolated expanders with bounded degree. These results give hope that similar results may be provable, even with more general host graphs and provide an avenue for future research beyond the scope of this thesis.

As a brief reminder, the normalized Laplacian of \( G \) is
\[
\mathcal{L} = I - D^{-1/2} AD^{-1/2},
\]
where \( A \) denotes the adjacency matrix and \( D \) denotes the diagonal degree matrix. Let
\[
0 = \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{n-1}
\]
denote the eigenvalues of \( \mathcal{L} \). The spectral gap is
\[
\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}.
\]
Occasionally it is more natural to use the parameter

\[ \lambda = 1 - \sigma = \min\{\lambda_1, 2 - \lambda_{n-1}\} \]

in place of \( \sigma \); the two may be used interchangeably.

Throughout this chapter, we have need to consider both the volume of a set \( X \) in \( G \); which we recall is

\[ \text{vol}(X) = \sum_{v \in X} d_v \]

where \( d_v \) denotes the degree of \( v \) in \( G \) and

\[ \text{vol}_H(X) = \sum_{v \in X} d'_v \]

where \( d'_v \) is the degree of \( v \) in \( H \).

With these definitions, we can immediately deduce the following two consequences of Theorem 7 (see [Chu97, KS06b]):

**Corollary 1.** For a graph \( G \) on \( n \) vertices with spectral gap \( \sigma = 1 - \lambda \) and minimum degree \( \delta \), a subgraph \( H \) with edge-selection probability \( p \) almost surely satisfies the following properties:

1. \( H \) satisfies the expansion property as follows: For \( X \subseteq V(H) \), the number of edges in \( H \) leaving \( X \), denoted by \( \partial_H(X) \) satisfies

\[ \partial_H(X) \geq \left( \lambda - O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right) \right) \text{vol}_H(X) \]

2. \( H \) satisfies the discrepancy property as follows: For \( X, Y \subseteq V(H) \), the number of edges of \( H \) between \( X \) and \( Y \), denoted by \( e_H(X,Y) \), satisfies

\[ |e_H(X,Y) - \frac{\text{vol}_H(X)\text{vol}_H(Y)}{\text{vol}_H(G)}| \leq \left( \sigma + O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right) \right) \sqrt{\text{vol}_H(X)\text{vol}_H(Y)}. \]
For a random walk on $H$ with transition probability matrix $P_H$, the total variation distance after $t$ steps from the stationary distribution $\pi$, denoted by $\Delta_{TV}(t)$, is bounded above by

$$\Delta_{TV}(t) = \max_{A \subseteq V} \max_{y \in V} \left| \sum_{x \in A} (P_H^t(y, x) - \pi(x)) \right|$$

$$\leq e^{-c}$$

for any $c > 0$ provided $t$ satisfies

$$t \geq \frac{1}{\lambda - O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right)} \left(\log \frac{\text{vol}_H(G)}{\min_x d(x)} + c\right).$$

**Corollary 2.** If $G$ is a graph on $n$ vertices and $H$ is a random subgraph obtained by bond percolation with probability $p$ with spectral gap $\sigma_H = 1 - \lambda_H$ and minimum degree $p\delta$ then $G$ surely satisfies the following properties:

1. $G$ satisfies the expansion property as follows: For $X \subseteq V(G)$, the number of edges in $G$ leaving $X$, denoted by $\partial(X)$ satisfies

   $$\partial(X) \geq \left(\lambda_H - O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right)\right)\text{vol}(X)$$

2. $G$ satisfies the discrepancy property as follows: For $X, Y \subseteq V(G)$, the number of edges of $G$ between $X$ and $Y$, denoted by $e(X, Y)$, satisfies

   $$|e(X, Y) - \frac{\text{vol}(X)\text{vol}(Y)}{\text{vol}(G)}| \leq$$

   $$\left(\sigma_H + O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right)\right)\sqrt{\text{vol}(X)\text{vol}(Y)}.$$

3. For a random walk on $H$ with transition probability matrix $P_H$, the total variation distance after $t$ steps from the stationary distribution $\pi$, denoted by $\Delta_{TV}(t)$, is bounded above by

   $$\Delta_{TV}(t) = \max_{A \subseteq V} \max_{y \in V} \left| \sum_{x \in A} (P_H^t(y, x) - \pi(x)) \right|$$

   $$\leq e^{-c}$$
for any $c > 0$ provided $t$ satisfies

$$t \geq \frac{1}{\lambda_H - O(\sqrt{\frac{\log n}{p^2}} + \frac{(\log n)^{1/2}}{p^2(\log \log n)^{3/2}})} \left(\log \frac{\text{vol}(G)}{\min_x d(x)} + c\right).$$

The conclusions of Corollary 1 above rely on the volume of a set within $H$. Likewise the conclusions of Corollary 2 rely on the volume of a set within $G$. In applications, we may be familiar with the host graph $G$ and want to apply Corollary 1, or with $H$ and want to apply Corollary 2 without full knowledge of what the $H$ and $G$ volumes of certain sets are. However, under the conditions that Theorem 7 gives us useful information it is easy to see that

$$d'_v = (1 + o(1))pd_v$$

for all vertices a.a.s., which in particular implies that for all $X$

$$\text{vol}_H(X) = (1 + o(1))p\text{Vol}(X).$$

This is the content of Lemma 6 below, and is a simple consequence of the Chernoff bounds. This fact, however, allows us to apply Corollaries 1 and 2 and draw conclusions about the likely properties of an unknown random subgraph or host graph, based solely on properties of the other.

### 3.2 Preliminaries

Let $G = (V, E)$ be a graph. We denote by $H$ a random graph obtained from $G$ by taking each edge independently with probability $p$. That is,

$$\mathbb{P}(\{u, v\} \in E(H)) = \begin{cases} p & \text{if } \{u, v\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

Let $A$ and $A_H$ denote the adjacency matrix of $G$ and $H$, respectively. We denote the diagonal matrices whose entries consist of the degrees of the vertices in $G$ and $H$ respectively, by $D$ and $D_H$. Let $0 = \eta_0 \leq \cdots \leq \eta_{n-1}$ denote the eigenvalues
of $\mathcal{L}_H$, and let $\varphi_i$ for $i = 0, \ldots, n - 1$ denote a set of orthonormal eigenvectors associated with the $\eta_i$ (represented here as row vectors). The projection to $\varphi_i$, for each $i$, is $P_i = \varphi_i^* \varphi_i$ where $\varphi^*$ denotes the transpose of $\varphi$. Then we have

$$\mathcal{L}_H = \sum_i \eta_i P_i.$$ 

We consider, then, the matrix

$$M = I - \mathcal{L}_H - P_0 = \sum_{i \neq 0} (1 - \eta_i) P_i.$$ 

We will use the fact that for any integer $k$, we have

$$\text{Tr}(M^{2k}) = \sum_{i \neq 0} (1 - \eta_i)^{2k}.$$ 

Immediately we have the following:

**Fact 1.** For any positive integer $k$

$$\max_{i \neq 0} |1 - \eta_i| = ||M|| \leq (\text{Tr}(M^{2k}))^{1/(2k)}.$$ 

Hence we have

$$\sigma_H = 1 - \max_{i \neq 0} |1 - \eta_i| = 1 - ||M||.$$ 

Let $K$ denote the all ones matrix. We can rewrite $M$ as:

$$M = D_H^{-1/2} A_H D_H^{-1/2} - P_0 = D_H^{-1/2} A_H D_H^{-1/2} - \varphi_0^* \varphi_0 = D_H^{-1/2} A_H D_H^{-1/2} - \frac{1}{\text{vol}(H)} D_H^{1/2} K D_H^{1/2}.$$ 

Instead of directly dealing with $M$, we consider the simpler matrix

$$C = p^{-1} D^{-1/2} A_H D^{-1/2} - \frac{1}{p \text{vol}(G)} p D^{1/2} K D^{1/2} = p^{-1} D^{-1/2} A_H D^{-1/2} - \frac{1}{\text{vol}(G)} D^{1/2} K D^{1/2}. \quad (3.1)$$
where one may note that $p\text{vol}(G)$ is the expected volume of $H$.

In a way, $C$ can be thought of as an estimate for the expectation of $M$. Our plan is to first carefully consider $||C||$ in the next section, and then bound the norm of the difference between $M$ and $C$ in Section 4.

### 3.3 A bound on $||C||$

In this section we prove the following theorem.

**Theorem 8.** Let $G$ be a given graph with spectral gap $\lambda$ and minimum degree $\delta$. Let $H$ be a random subgraph of $G$ with edge-selection probability $p$. Then the matrix $C$ as defined in (3.1) asymptotically almost surely satisfies

$$||C|| = \sigma + O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta \log \log n}\right).$$

**Proof.** To bound the norm of $C$, we express $C$ as a sum of two parts, a ‘random’ part and a ‘non-random’ part derived from the host graph $G$:

$$C = B + M'$$

where

$$B = p^{-1}D^{-1/2}A_H D^{-1/2} - D^{-1/2}AD^{-1/2}$$

and

$$M' = D^{-1/2}AD^{-1/2} - \frac{1}{\text{vol}(G)}D^{1/2}K D^{1/2}.$$ 

Note that $M'$ is deterministic, and is equivalent to the matrix $M$ for the graph $G$. Hence we have

$$||M'|| = \sigma.$$

It suffices to show that almost surely we have

$$||B|| = O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta \log \log n}\right).$$
In other words, we wish to prove that for any $\epsilon > 0$, there is an absolute constant $c$ so that for $n$ sufficiently large, we can bound the following probability as follows:

$$
P\left(||B|| \geq c\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right)\right) \leq \epsilon.
$$

The matrix $B$ is a random matrix, where the entries $b_{ij}$ are independent random variables defined by:

$$
b_{ij} = \begin{cases} 
\frac{1}{p\sqrt{d_id_j}} - \frac{1}{\sqrt{d_id_j}} & \text{if } v_i \sim v_j \in H \\
-\frac{p}{p\sqrt{d_id_j}} & \text{if } v_i \not\sim v_j \in H \text{ and } v_i \sim v_j \in G \\
0 & \text{otherwise.}
\end{cases}
$$

Here $d_i$ denotes the degree of $v_i$ in $G$. It follows from the definition that the expected value of $b_{ij}$ satisfies

$$
E[b_{ij}] = 0.
$$

Now consider the $(i,i)$th entry of $B^{2k}$. A typical term of such an entry is of the form

$$
b_{i_1i_2}b_{i_3i_4} \ldots b_{i_{2k-1}i_{2k}}
$$

with $i_1 = i_{2k} = i$. Assuming this term is non-zero, this corresponds to a closed walk in $G$ starting and ending at $v_i$. Taking expectations, we note that

$$
E[b_{i_1i_2}b_{i_3i_4} \ldots b_{i_{2k-1}i_{2k}}] \neq 0
$$

only if each $b_{ij}$ occurs at least twice (since $E[b_{ij}] = 0$ and all $b_{ij}$’s are independent). In other words, each edge must occur at least twice in the closed walk. We refer to a such a closed walk, which contributes to the expected trace, as a surviving walk.

To determine the expected contribution of a surviving walk to $\text{Tr}(B^{2k})$, we consider the expected value $E[b_{ij}^m]$. Note that for $m \geq 2$

$$
|E[b_{ij}^m]| \leq \frac{|(1-p)^m + (-p)^m(1-p)|}{p^m(d_id_j)^{m/2}} \leq \frac{p}{p^m(d_id_j)^{m/2}} = \frac{1}{p^{m-1}(d_id_j)^{m/2}}.
$$

The last inequality follows from the easy fact that $|(1-p)^m + (-p)^m(1-p)| \leq p$ when $p \leq 1$. 

To bound the trace of $B^{2k}$, we must get a handle on the number of surviving walks and their contribution to the trace. Consider a surviving walk of length $2k$ on vertices $v_1, \ldots, v_{l+1}$ and let us assume that the vertices are labelled by their first occurrence in the walk. Thus to get to vertex $v_i$, we must have followed an edge from one of $v_1, \ldots, v_{i-1}$. We define the exposure sequence of the walk to be a vector $(a_1, \ldots, a_l)$ such that we first travel to vertex $v_i$ from vertex $v_{a_i-1}$. Clearly, $a_i \in \{1, \ldots, i\}$. Hence there are at most $l!$ possible exposure sequences. We seek to enumerate our surviving walks by their exposure sequences.

Consider a surviving walk on vertices $v_1, \ldots, v_{l+1}$ whose exposure sequence is $e = (a_1, \ldots, a_l)$. Let us assume that the walk contains edges $e_1, \ldots, e_k$ with multiplicities $m_1, \ldots, m_k$, respectively. Then the contribution of the walk to the expected value of the trace is at most

$$\mathbb{E}[b_{e_1}^{m_1} \ldots b_{e_k}^{m_k}] = \prod_i \mathbb{E}[b_{e_i}^{m_i}] \leq \prod_{e_i = (v_{a_{i-1}}, v_{a_i})} \frac{1}{p^{m_i-1}(d_{a_{i-1}}d_{a_i})^{m_i/2}} \leq \frac{1}{(\prod_{i=1}^{l} d_{a_i}) (pd)^{2k-l}}$$

where the $d_{a_i}$ terms comes from the fact that there must exist an edge contributing $1/(pd_{a_i})$ to the product since an edge incident to $a_i$, must occur with multiplicity at least 2, while replacing all other terms with their minimum possible values.

For a set of vertices $S = \{v_1, \ldots, v_{l+1}\}$ along with an exposure sequence $e = (a_1, \ldots, a_l)$, we let $W(S, e, k)$ denote the number of surviving walks of length $2k$ on vertices $S$ with exposure sequence $e$. We can upper bound the number of surviving walks by the number of surviving walks on these vertices in a complete graph of the same size. Let $W'(k, l)$ denote the number of surviving walks of length $2k$ on the complete graph $K_{l+1}$ such that the vertices are visited in order $v_1, \ldots, v_{l+1}$. (Clearly the labeling does not affect the number of paths; just the fact that the vertices are visited in a particular order.) We note that for a given set $S$ with $|S| = l + 1$, there can be at most $l!$ exposure sequences. Furthermore, for a set $S$ with $l + 1$ vertices and an exposure sequence $e$, we have

$$W(S, e, k) \leq W'(k, l).$$
This inequality is immediate, as each surviving walk on $G|_S$ corresponds injectively to a walk on the complete graph $K_{l+1}$. Further note that this inequality holds independently of the exposure sequence $e$.

Füredi and Komlós [FK81] gave an upper bound on $W'(k,l)$. Recently this bound was improved by Vu, and it is this new bound we use. Inequality (9) in [Vu07] asserts

**Lemma 5.**

$$W'(k,l) \leq \binom{2k}{2l} 2^{k+2(k-l)+1}(l+1)^{3(k-l)}$$

We can now bound the expected value of $\text{Tr}(B^{2k})$ by both applying the above bound for $W'(k,l)$ and using the fact that we are counting surviving walks. We use the notation $u \sim v$ to denote that $u$ and $v$ are adjacent in $G$. 
\[ \mathbb{E}[\text{Tr}(B^{2k})] \leq \sum_{l=1}^{k} \sum_{e=(a_1, \ldots, a_l)} \sum_{\substack{v_1 \sim \cdots \sim v_{a_1} \sim \cdots \sim v_{a_l} \sim v_{l+1}}} \frac{1}{\left( \prod_{i=1}^{l} d_{a_i} \right) (p\delta)^{2k-l}} \]

\[ \leq \sum_{l=1}^{k} \sum_{e=(a_1, \ldots, a_l)} \sum_{\substack{v_1 \sim v_{a_1} \sim v_{a_2} \sim \cdots \sim v_{a_l} \sim v_{l+1}}} \frac{W(S, e, k)}{\left( \prod_{i=1}^{l} d_{a_i} \right) (p\delta)^{2k-l}} \]

\[ \leq \sum_{l=1}^{k} \sum_{e=(a_1, \ldots, a_l)} \sum_{\substack{v_1 \sim v_{a_1} \sim v_{a_2} \sim \cdots \sim v_{a_l} \sim v_{l+1}}} \frac{W'(k, l)}{\left( \prod_{i=1}^{l} d_{a_i} \right) (p\delta)^{2k-l}} \]

\[ = \sum_{l=1}^{k} \sum_{e=(a_1, \ldots, a_l)} \sum_{v_1} W'(k, l) \frac{1}{(p\delta)^{2k-l}} \]

\[ \leq \sum_{l=1}^{k} \sum_{e=(a_1, \ldots, a_l)} l n \left( \frac{2k}{2l} \right)^{2k+2(k-l)+1(l+1)^{3(k-l)}} \frac{1}{(p\delta)^{2k-l}} \]

\[ \leq \sum_{l=1}^{k} l! n \left( \frac{2k}{2l} \right)^{2k(l+1)^{3(k-l)}} \frac{1}{(p\delta)^{2k-l}} \]

\[ \leq \sum_{l=1}^{k} l! n \left( \frac{2k}{2l} \right)^{2k(l+1)^{3(k-l)}} \frac{1}{(p\delta)^{2k-l}} \]

\[ \leq \sum_{l=1}^{k} n 32^{k(l+1)^{3k-2l}} \frac{1}{(p\delta)^{2k-l}} \]

\[ = \sum_{l=1}^{k} n 32^{k} s_{l,k} \]

where we define

\[ s_{l,k} = \frac{(l+1)^{3k-2l}}{(p\delta)^{2k-l}}. \]

For a fixed \( \epsilon > 0 \), with \( \epsilon < 1/4 \) we now choose

\[ k = \lfloor \log n + \log(1/\epsilon) \rfloor \quad (3.2) \]

and set

\[ p\delta = \alpha \left( \frac{k}{\log k} \right)^{3/2}. \]
Note that $\alpha$ is a function of $k$ (and hence $n$). We wish to show the following:

Claim:

\[ s_{l,k} \leq \left( \frac{c}{\alpha \min\{\alpha, \frac{k^{1/2}}{\log^{3/2} k}\}} \right)^k \]  

for some absolute constant $c$.

We let

\[ f(l) := \frac{s_{l,k}}{s_{l-1,k}} = \frac{(1 + l^{-1})^{3k-2lp\delta}}{l^2} = c_0 e^{3k/lp\delta/l^2} \]  

where $c_0$ is upper and lower bounded by some absolute constants. For a given value of $p\delta$ and for the range of $0 \leq l \leq k$, the function $s_{l,k}$ either attains its maximum at $l_0$ satisfying $l_0 = k$ and $f(k) > 1$, or $l_0$ is one of the two integers closest to the solution of $f(l) = 1$. Note that for the first case, we have

\[ s_{l_0,k} = s_{k,k} \leq c \left( \frac{k + 1}{p\delta} \right)^k \leq c \left( \frac{k}{\alpha k^{3/2} \log^{-3/2} k} \right)^k = c \left( \frac{1}{\alpha k^{1/2} \log^{-3/2} k} \right)^k \]

which implies (3.3). We may assume that $l_0$ is one of the two integers closest to the solution of $f(l) = 1$. Furthermore, for $l < k/(2 \log k)$, we have, from (3.4),

\[ \frac{s_{l,k}}{s_{l-1,k}} > \frac{e^{2 \log k}}{k^2} > 1. \]

Therefore we may assume that

\[ l_0 \geq \frac{k}{2 \log k}. \]

There are two possibilities:
Case 1: \( l_0 \leq 100k / \log k \). Then

\[
\begin{align*}
sl_{l_0,k} & \leq \left( c' \frac{l_0^3}{(p\delta)^2} \right)^k \\
& \leq \left( c' \frac{k^3}{(\log k)^3} \right)^k \\
& = \left( \frac{c}{\alpha^2} \right)^k \\
\end{align*}
\]

which implies (3.3).

Case 2: \( l_0 \geq 100k / \log k \). We use the fact that \( l_0 \) is one of the two integers closest to the solution of \( f(l) = 1 \). From equation (3.4), we have:

\[
\begin{align*}
sl_{l_0,k} & \leq \left( \frac{l_0^{3-2l_0/k}}{(p\delta)^{2-l_0/k}} \right)^k \\
& \leq \left( c'' \frac{l_0^{3-2l_0/k}}{(l_0^2 e^{-3k/l_0})^{2-l_0/k}} \right)^k \\
& \leq \left( \frac{c'' e^{6k/l_0}}{l_0} \right)^k \\
\end{align*}
\]

(3.5)

One can check that for the given range of \( l_0 \) this satisfies:

\[
\begin{align*}
sl_{l_0,k} & \leq \left( \frac{c'''k}{l_0^2} \right)^k \\
& \leq \left( \frac{c'''k e^{3k/l_0}}{l_0^2} \right)^k \\
& \leq \left( \frac{c_0}{\alpha k^{1/2} \log^{-3/2} k} \right)^k \\
\end{align*}
\]

which again implies (3.3). In this section, \( c_0, c, c', c'', \ldots \) are suitably chosen integers. The proof for the claim is completed.

We are now ready to consider bounding the norm of \( B \).

\[
\begin{align*}
\mathbb{E}[\text{Tr}(B^{2k})] & \leq n^2 32^k \max_l s_{l,k} \\
& \leq 2n^2 \left( \frac{32c_0}{\alpha \min\{\alpha, k^{1/2} \log^{-3/2} k\}} \right)^k.
\end{align*}
\]

Since \( \mathbb{E}[||B||^{2k}] \leq \mathbb{E}[\text{Tr}(B^{2k})] \), we have

\[
\mathbb{E}[||B||^{2k}] \leq 2n^2 \left( \frac{c}{\min\{\alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k\}} \right)^{2k}.
\]
By the previous equation and Markov’s equality (Proposition 1 from Chapter 1), we have

\[
\P \left( \left\| B \right\| \geq 2 \frac{c}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right) = \P \left( \left\| B \right\|^{2k} \geq 2^{2k} \left( \frac{c}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right)^{2k} \right) \leq \E \left[ \left\| B \right\|^{2k} \right] \leq \frac{2n^2 \left( \frac{c}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right)^{2k}}{2^{2k} \left( \frac{c}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right)^{2k}} \leq \frac{2n^2}{2^{2k}} \leq \epsilon
\]

for the given \( \epsilon > 0 \) (noting this holds as \( \epsilon < 1/4 \)) and our choice of \( k \) in (3.2). Hence we have proved that almost surely we have

\[
\left\| C \right\| \leq \left\| M' \right\| + \left\| B \right\| = 1 - \lambda + O\left( \frac{1}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right).
\]

In a similar way, we can use the fact that \( \left\| C \right\| \geq \left\| M' \right\| - \left\| B \right\| \) to get

\[
\left\| C \right\| = 1 - \lambda \pm O\left( \frac{1}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right).
\]

It is easily verified that

\[
O\left( \frac{1}{\min \{ \alpha, \alpha^{1/2} k^{1/4} \log^{-3/4} k \}} \right) = O\left( \sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}} \right)
\]

which completes the proof of Theorem 8.

\[\square\]

**Remark.** Recall that the error term in Theorems 7 and 8 are effective when \( p\delta \gg \frac{\log^{3/2}(n)}{(\log \log n)^{3/2}} \). In Chung, Lu and Vu’s [CLV04] result for \( G(w) \), their theorem was effective in the case where the expected minimum degree is \( \gg \log^2(n) \). One may ask where the difference in the bounds came from. The main difference lies in the use of the improved path counting bound of Vu [Vu07] (Lemma 5), while the factor \( \log \log(n)^{-3/2} \) comes from more careful analysis. Using Lemma 5 in the
earlier paper of Chung, Lu and Vu immediately improves their result to work while the expected minimum degree is $\gg \log^{3/2}(n)$. As mentioned before, if the minimum expected minimum degree is less than $(1 - \epsilon) \log n$ the random subgraph need not be connected - hence no real control of the spectral gap is possible. There is still, however, a gap where some improvement is possible.

3.4 Bounding the Spectral Gap

In this section, we plan finish the proof of Theorem 7. Namely, we wish to show that for a graph $G$ with spectral gap $\sigma$ and minimum degree $\delta$, a random subgraph $H$ obtained from $G$ with edge-selection probability $p$ asymptotically almost surely has eigenvalues of the Laplacian $L_H$ of $H$ satisfying:

$$\sigma_H = \max_{i \neq 0} |1 - \eta_i| = \sigma + O\left(\sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta(\log \log n)^{3/2}}\right).$$

As a matter of notation, we let $d_i$ refer to the degree of vertex $v_i$ in $G$ and $d'_i$ refer to the degree of vertex $i$ in $H$. We also let $a'_{ij}$ refer to the $ij$th entry of $A_H$, the adjacency matrix of $H$.

To prove Theorem 1, recall that the eigenvalues of the Laplacian $L_H$ satisfy

$$\max_{i \neq 0} |1 - \eta_i| = ||M||$$

where $M = D_H^{-1/2} A_H D_H^{-1/2} - \frac{1}{\text{vol}(H)} D_H^{1/2} K D_H^{1/2}$. Now we write:

$$M = E + C + R + S$$

where we define

$$E = D_H^{-1/2} A_H D_H^{-1/2} - \frac{1}{p} D^{-1/2} A_H D^{-1/2} - \frac{p}{\text{vol}(G)} D D_H^{-1/2} K D_H^{-1/2} D$$

$$+ \frac{1}{\text{vol}(G)} D^{1/2} K D^{1/2}$$

$$R = \frac{p}{\text{vol}(G)} D D_H^{-1/2} K D_H^{-1/2} D - \frac{1}{p\text{vol}(G)} D_H^{1/2} K D_H^{1/2}$$

$$S = \left(\frac{1}{p\text{vol}(G)} - \frac{1}{\text{vol}(H)}\right) D_H^{1/2} K D_H^{1/2}$$
and $C$ is as defined in (3.1). Thus

\[ e_{ij} = \left( a'_{ij} - pd_i d_j \right) \left( \frac{1}{\sqrt{d_i d_j'}} - \frac{1}{p \sqrt{d_i d_j}} \right) \]

\[ r_{ij} = \frac{1}{p \text{vol}(G)} \frac{p^2 d_i d_j - d'_i d'_j}{\sqrt{d_i d_j'}} \]

\[ s_{ij} = \left( \frac{1}{p \text{vol}(G)} - \frac{1}{\text{vol}(H)} \right) \sqrt{d_i d_j'} \]

and

\[ c_{ij} = \frac{a'_{ij}}{p \sqrt{d_i d_j'}} - \frac{1}{\text{vol}(G)} \sqrt{d_i d_j}. \]

Clearly

\[ ||M|| \leq ||E|| + ||C|| + ||R|| + ||S||. \]

Hence, it suffices to establish the appropriate upper bounds for the norms of $E$, $C$, $R$, and $S$.

To bound these, we use the following Chernoff bounds (Proposition 5 from Chapter 1): if $X_i$ be independent random variables satisfying $|X_i| \leq M$, and $X = \sum_i X_i$. Then we have, for any $a > 0$,

\[ P(|X - \mathbb{E}[X]| > a) \leq e^{-\frac{a^2}{\mathbb{E}[X] + \text{Var}(X) + \text{Vol}(M)}}. \]

We will prove the following lemma (whose proof we delay until after the proof of Theorem 1).

**Lemma 6.** Assuming that $p\delta \gg \log n$, almost surely every vertex $v_i$ satisfies

\[ d'_i = pd_i (1 + O \left( \sqrt{\log n \over p\delta} \right)). \]

Let $X_e$, for $e \in E(G)$, be the random indicator variable which is 1 if $e \in H$ and 0 otherwise. We can write

\[ \text{vol}(H) = \sum_{e \in E(G)} 2X_e. \]
We can show that almost surely
\[ \left| \text{vol}(H) - p\text{vol}(G) \right| < 2\sqrt{p\text{vol}(H)}g(n) \] (3.8)
for any function \( g(n) \) that goes to infinity as \( n \) approaches infinity.

We also have the following lemma (whose proof will be given later).

**Lemma 7.** Suppose that \( p\delta \gg \log n \). Almost surely the vector \( \chi \) with \( \chi(i) = (d'_i - pd_i)/\sqrt{pd_i} \) satisfies
\[ ||\chi||^2 \leq (1 + o(1))n. \]

**Proof of Theorem 7.** We note that we already established a bound on \( ||C|| \) in the last section. By Theorem 8, we have that almost surely
\[ ||C|| = 1 - \lambda + O\left( \sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta \log \log n} \right) \]
a.a.s.

For convenience, we define
\[ \beta = \sqrt{\frac{\log n}{p\delta}} + \frac{(\log n)^{3/2}}{p\delta \log \log n}. \]

For \( ||R|| \), we have the following a.a.s. by using equation (3.7) and the Cauchy-
Schwartz inequality.

\[ ||R|| = \max_{||y||=1} \langle y, Ry \rangle \]
\[ \leq \max_{||y||=1} \frac{1}{p\text{vol}(G)} \sum_{i,j} y_i y_j d'_i (d'_j - pd_j) + (d'_i - pd_i)pd_j \sqrt{d'_i d'_j} \]
\[ = \frac{1}{p\text{vol}(G)} \max_{||y||=1} \left\{ \sum_i \sqrt{d'_i y_i} \sum_j \frac{(d'_j - pd_j)y_j}{\sqrt{d'_j}} + \sum_i \frac{(d'_i - pd_i)y_i}{\sqrt{d'_i}} \sum_j \frac{pd_j y_j}{\sqrt{d'_j}} \right\} \]
\[ \leq 1 \max_{||y||=1} \left\{ (\sum_i d'_i)^{1/2} ||y|| (\sum_j \frac{(d'_j - pd_j)^2}{d'_j})^{1/2} + ||y|| (\sum_j \frac{(d'_j - pd_j)^2}{d'_j})^{1/2} \right\} \]
\[ \leq (2 + o(1)) \sqrt{\frac{n}{p\text{vol}(G)}} \]
\[ \leq (1 + o(1)) \frac{2}{\sqrt{pd}} \]
\[ = o(\beta). \]

For \( ||S|| \), by using (3.8) and the Cauchy-Schwartz inequality we have:

\[ ||S|| = \max_{||y||=1} \langle y, Sy \rangle = \max_{||y||=1} \sum_{i,j} y_i y_j \left( \frac{1}{p\text{vol}(G)} - \frac{1}{\text{vol}(H)} \right) \sqrt{d'_i d'_j} \]
\[ \leq \left( \frac{1}{p\text{vol}(G)} - \frac{1}{\text{vol}(H)} \right) \max_{||y||=1} \sum_{i,j} |y_i \sqrt{d'_i}| |y_j \sqrt{d'_j}| \]
\[ \leq \frac{2\sqrt{p\text{vol}(G) \log(n)}}{p\text{vol}(H) \text{vol}(G)} \max_{||y||=1} \left( \sum_i |y_i |^2 \right)^{1/2} \]
\[ \leq \frac{2\sqrt{p\text{vol}(G) \log(n)}}{p\text{vol}(H) \text{vol}(G)} \left( \sum_i y_i^2 \right)^{1/2} \left( \sum_i d'_i \right) \]
\[ = O\left( \frac{\log n}{p\text{vol}(G)} ||y||^2 \right) \]
\[ = o\left( \frac{\log n}{\sqrt{pd}} \right) \]
\[ = o(\beta). \]
Finally, it remains to bound $||E||$. We recall
\[ e_{ij} = (a'_{ij} - \frac{p_{d_i d_j}}{\text{vol}(G)})\left(\frac{1}{\sqrt{d'_i d'_j}} - \frac{1}{p\sqrt{d_i d_j}}\right) = c_{ij} p_{\sqrt{d_i d_j} - \sqrt{d'_i d'_j}}. \]

Thus we have
\[
||E|| = \max_{||y||=1} \langle y, Ey \rangle = \max_{||y||=1} \sum_{i,j} y_i y_j c_{i,j} \frac{\sqrt{d'_i} (\sqrt{d'_j} - \sqrt{pd_j}) + (\sqrt{d'_i} - \sqrt{pd_i}) \sqrt{pd_j}}{\sqrt{d'_i d'_j}}.
\]

Let $y'_i = y_i (\sqrt{d'_i} - \sqrt{pd_i})/\sqrt{d'_i}$ and $y''_i = y_i \sqrt{pd_i}/\sqrt{d'_i}$. Then we have almost surely
\[
||E|| \leq \max_{||y||=1} \langle y, C'y' \rangle + \langle y', C'y'' \rangle \\
\leq \max_{||y||=1} ||C'|| ||y'|| + ||C'|| ||y'|| ||y''|| \\
\leq O(\beta).
\]

This last observation follows from Lemma 3, which implies
\[
||y'|| \leq ||y|| \left(1 - \frac{1}{\sqrt{1 + O(\sqrt{n/p\delta})}}\right) = O(\beta)
\]
and, $||y''|| = (1 + o(1))||y|| = O(1)$. Note that we have already observed that $||C'|| = O(1)$. Combining these results, we have
\[
\max_{i \neq 0} |1 - \eta_i| = ||M|| \\
\leq ||E|| + ||C'|| + ||R|| + ||S|| \\
\leq \sigma + O(\beta).
\]

In the other direction, the lower bound follows as $||M|| \geq ||C'|| - ||E|| - ||R|| - ||S|| = \sigma - O(\beta)$.

This gives the following bound on the spectral gap of $H$, completing the proof of Theorem 1:
\[
|\sigma_H - \sigma| = O(\beta).
\]
It remains to prove Lemmas 6 and 7.

Proof of Lemma 6. For a vertex \(v_i \in G\), we can write \(d'_i = \sum_{v_j \sim v_i} X_j\) where \(X_i\) is the random indicator variable having value 1 if \(\{v_i, v_j\} \in E(H)\) and 0 otherwise. Then \(E[d'_i] = pd_i\) and \(\text{Var}(d'_i) = d_ip(1-p)\). By the Chernoff bounds, Proposition 5, we have

\[
P(|d'_i - pd_i| > a) \leq \exp\left(-a^2 \frac{2}{2(d_ip(1-p)) + a/3}\right).
\]

Setting \(a = 2\sqrt{\log(n)pd_i}\), we have that

\[
P(|d'_i - pd_i| > a) \leq \exp\left(-4pd_i \log(n)\right) \leq n^{-2+o(1)}
\]

Thus asymptotically almost surely, for all \(i\) we have \(|d'_i - pd_i| \leq 2\sqrt{\log(n)pd_i}\). This can be restated as, for all \(i\),

\[
|d'_i - pd_i| \leq 2pd_i \sqrt{\frac{\log n}{pd_i}} = pd_i O\left(\sqrt{\frac{\log n}{pd}}\right).
\]

\[\Box\]

The following proof of Lemma 7 is analogous to Lemma 3.3 in [CLV04].

Proof of Lemma 7. For a vertex \(v_i \in G\), let \(X_i = (d'_i - pd_i)^2\) and \(X = \sum_{i=1}^n \frac{X_i}{pd_i}\). For each \(i\), we can write

\[
X_i = \left(\sum_{v_j \sim v_i} X_{ij} - p\right)^2
\]

where \(X_{ij}\)'s are the indicator random variables of the event that \(v_i\) adjacent to \(v_j\) in \(H\) (as denoted by \(v_i \sim v_j\)). We define

\[
x_{ij} = \begin{cases} 
X_{ij} - p & \text{if } v_i \sim v_j, \\
0 & \text{otherwise}.
\end{cases}
\]
Thus, $E[x_{ij}] = 0$, and $X_i = (\sum_j x_{ij})^2$. Also,

$$
\begin{align*}
\mathbb{E}[X_i] &= \text{Var}(d'_i) = \mathbb{E}\left[\sum_{v_i \sim v_j} x_{ij}^2\right] < pd_i \\
\mathbb{E}[X_i^2] &= \mathbb{E}\left[\left(\sum_{v_i \sim v_j} x_{ij}\right)^4\right] \\
&= \sum_{i \sim j} \mathbb{E}[x_{ij}^4] + 6 \sum_{j \neq k \sim v_i, v_j \sim v_i} \mathbb{E}[x_{ij}^2 x_{ik}^2] \\
&\leq pd_i + 6p^2d_i^2.
\end{align*}
$$

If $v_i \not\sim v_j$ and $v_i \neq v_j$, then $X_i$ and $X_j$ are independent. If $v_i \sim v_j$ and $v_i \neq v_j$, we have

$$
\begin{align*}
\mathbb{E}[X_i X_j] &= \mathbb{E}\left[(d'_i - pd_i)(d'_j - pd_j)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{v_k \sim v_i} x_{ik}\right)^2 \left(\sum_{v_l \sim v_j} x_{lj}\right)^2\right] \\
&= \sum_{v_k \sim v_i, v_l \sim v_j, \{i,k\} \neq \{l,j\}} \mathbb{E}[v_{ik}^2 v_{lj}^2] + \mathbb{E}[v_{ij}^4] \\
&\leq p^2d_i d_j + p.
\end{align*}
$$

Thus,

$$
\begin{align*}
\text{Var}(X_i) &\leq pd_i + 5p^2d_i^2 \\
\text{coVar}(X_i, X_j) &\leq \begin{cases} 
0 & \text{if } v_i \not\sim v_j, \\
p & \text{otherwise}.
\end{cases}
\end{align*}
$$

Therefore

$$
\begin{align*}
\mathbb{E}[X] &= \sum_{i=1}^n \frac{1}{pd_i} \mathbb{E}[X_i] < n \\
\text{Var}(X) &= \sum_{i=1}^n \frac{1}{p^2d_i^2} \text{Var}(X_i) + 2 \sum_{i<j \leq n} \frac{1}{w_i w_j} \text{coVar}(X_i, X_j) \\
&\leq (5 + \frac{1}{pd} + \frac{1}{d}) n \\
&\leq 6n.
\end{align*}
$$
Using Chebyshev’s inequality (Proposition 2 from Chapter 1), we have, for any \( a > 0 \),
\[
\mathbb{P}(|X - E[X]| > a) \leq \frac{\text{Var}(X)}{a^2}.
\]
Setting \( a = \sqrt{ng(n)} \), with \( g(n) \gg 1 \), then almost surely we have \( X \leq (1 + o(1))n \).
From the definition, \( ||\chi||^2 = X \). Thus, almost surely
\[
||\chi||^2 \leq (1 + o(1))n
\]
as desired. \( \square \)

### 3.5 Remarks and Optimality

In the above sections, we examined the spectral relationship between a host graph \( G \) and its random subgraph with edge-selection probability \( p \). If \( G \) has \( n \) vertices with a spectral gap \( \sigma \) and minimum degree \( \delta \), then we prove that a random subgraph of \( G \) on \( n \) vertices with edge-selection probability \( p \) almost surely has a spectral gap of \( \sigma + O\left( \sqrt{\frac{\log n}{p\delta}} + \frac{\log n}{p\delta(\log \log n)^{3/2}} \right) \). The special case of having the host graph as the complete graph on \( n \) vertices and a random subgraph \( H \) chosen with edge-selection probability \( p \) is the Erdős-Rényi graph \( G_{n,p} \). Since the complete graph \( K_n \) has eigenvalue \( \lambda_1 = n/(n-1) \), our bound for \( \lambda_H \) is \( |1 - \lambda_H| = O\left( \sqrt{\frac{\log n}{n}} \right) \) which is off by a factor of \( \sqrt{\log n} \) of the best known spectral bound for \( G_{n,p} \). Therefore there is room for improvements (e.g., by a factor of \( \sqrt{\log n} \)) concerning the statements of the main theorem here.

One could also ask how good these results could possibly be. For \( k \)-regular graphs, it is known due to Alon and Boppana (see, eg. [Nil04]) that
\[
\liminf_{n \to \infty} \sigma \geq \frac{2\sqrt{k-1}}{k}.
\]  
(3.9)
This certainly suggests that for random regular graphs, \( \sigma = \Omega\left( \frac{1}{\sqrt{k}} \right) \) necessarily; Friedman [Fri08] showed that random regular graphs come close to achieving the bound in Equation (3.9).
For irregular graphs, the question is slightly more complicated. Here, the pertinent quantity seems to be the following edge harmonic average degree of $G$ by

$$\frac{n}{\hat{d}} = \sum_{\{u,v\}} \frac{2}{d_u d_v}.$$ 

Rearranging we see that

$$\hat{d} = \frac{n}{\sum_{\{u,v\}} \frac{2}{d_u d_v}}.$$ 

One can observe that this is related to the more traditional harmonic average degree, defined by

$$\frac{n}{d_H} = \sum_v \frac{1}{d_v}.$$ 

A natural question is how the edge harmonic average degree compares to other measures of average degree. This is not a completely straightforward question as the edge harmonic average degree is in some sense a structural average degree, that relies on more than just the degree sequence.

If $G$ is $k$-regular, then it is easy to observe that

$$\frac{n}{\hat{d}} = \sum_{\{u,v\}} \frac{2}{k^2} = \frac{kn}{k^2} = \frac{n}{k}.$$ 

This yields

**Proposition 7.** If $G$ is a $k$-regular graph we have that

$$\hat{d} = k.$$ 

An application of the Cauchy-Schwarz inequality yields

$$\frac{n}{\hat{d}} = \sum_{\{u,v\}} \frac{1}{d_u d_v} \leq \sqrt{\sum_{\{u,v\}} \frac{1}{d_u^2} \sum_{\{u,v\}} \frac{1}{d_v^2}} = \sum_{v \in V} \frac{1}{d_v} = \frac{n}{d_H}.$$ 

Thus we have proved
Proposition 8. Every graph $G$ satisfies

$$\hat{d} \geq d_H.$$ 

For graphs with general degree sequences, the relationship between the (usual) average degree $d$ and the edge harmonic average degree $\hat{d}$ is somewhat less clear. Here we give illustrative examples where $\hat{d} = o(d)$ and $\hat{d} = \omega(G)$.

Example 2. Consider a graph on $n$ vertices consisting of a 3-regular graph on $n/2$ vertices and a clique on the remaining $n/2$ vertices connected by a matching. Then the average degree is

$$d = \frac{2n + \frac{n^2}{4}}{n} = \frac{n}{4} + 2 \approx \frac{n}{4}. $$

On the other hand

$$\sum_{u \sim v} \frac{2}{d_u d_v} = 3 \frac{n}{4} \cdot \frac{2}{16} + \frac{n}{2} \cdot \frac{2}{4n} + \left(\frac{n/2}{2}\right) \frac{8}{n^2} \sim \frac{3}{8}n.$$ 

This implies that

$$\hat{d} \sim \frac{8}{3}.$$ 

In particular

$$\hat{d} = o(d).$$ 

Example 3. Consider the star $S_n$ on $n$ vertices. Then the average degree is

$$d = \frac{2(n - 1)}{n} \sim 2.$$ 

On the other hand

$$\sum_{u \sim v} \frac{2}{d_u d_v} = (n - 1) \cdot \frac{2}{n - 1} = 2,$$

implying that

$$\hat{d} = \frac{n}{2}.$$ 

In particular for this example

$$\hat{d} = \omega(d).$$ 

It is not difficult to prove
Theorem 9. Suppose \( \hat{d} = o(n) \), then
\[
\sigma \geq (1 + o(1)) \frac{1}{\sqrt{\hat{d}}}.
\]

Proof. Let \( M = (I - \mathcal{L}) \), then
\[
\text{Tr}(M^2) = \sum (1 - \lambda_i)^2 \leq 1 + (n - 1)\sigma^2.
\]

On the other hand,
\[
\text{Tr}(M^2) = \sum_{e \in E} \frac{2}{d_u d_v} = \frac{n}{\hat{d}}.
\]
This implies
\[
(n - 1)\sigma^2 \geq \frac{n}{\hat{d}} - 1,
\]
from whence the result follows.

The interesting part about this is it allows us to confirm the optimality of the result of Chung, Lu and Vu that for \( G(w) \), \( \sigma \leq \frac{4}{\sqrt{d}} \) (under some conditions on \( w \)) up to a constant factor. In particular, for graphs in \( G(w) \) as well as graphs generated by the configuration model of random graphs \( G(d) \) we have that
\[
\hat{d} = (1 + o(1))d,
\]
under some conditions. In particular, we have:

Theorem 10. For a graph \( G \in G(w) \) with \( w_{\text{min}} = \omega(\log(n)) \),
\[
\hat{d} = (1 + o(1))d
\]
a.a.s.

Proof. Let \( X_{uv} \) denote the indicator that \( u \sim v \in G \). Then
\[
\frac{n}{\hat{d}} = \sum_{\{u,v\}} \frac{2}{d_u d_v} X_{uv}.
\]
Note here that \( d_u, d_v \) and \( X_{u,v} \) are random variables, making this a bit difficult to control directly. Instead, consider
\[
Y = \sum_{\{u,v\}} \frac{2}{w_u w_v} X_{u,v}.
\]
Since $\mathbb{E}[X_{u,v}] = \frac{w_u w_v}{\text{Vol}(G)}$, it is easy to compute that

$$\mathbb{E}[Y] = \sum_{\{u,v\}} \frac{2w_u w_v}{w_u w_v \text{Vol}(G)} = \left( \frac{n}{2} \right) \frac{1}{\text{Vol}(G)} = d(n-1).$$

The Chernoff bounds (Proposition 3) easily imply that $Y$, as the sum of independent indicators, is tightly concentrated so $Y = (1+o(1))d(n-1)$ a.a.s. Furthermore, by our assumption that $w_{\min} = \omega(\log(n))$ it is easy to show (a variant of Lemma 6 above) that a.a.s. all degrees satisfy $d_i = (1+o(1))w_i$ and in particular this implies that $Y = (1 + o(1))\frac{n^2}{d}$. The result then follows from our computation above.

For $G \in G(w)$, Theorems 9 and 10 imply that

**Corollary 3.** For $G \in G(w)$ with $w_{\min} = \omega(\log n)$, let $d$ denote the expected average degree. Then

$$\sigma \geq (1-o(1)) \frac{1}{\sqrt{d}}.$$

Another interesting observation is that the edge harmonic average degree behaves as it should with respect to percolation. That is, if the edge harmonic average degree of $G$ is $\hat{d}$, then the edge harmonic average degree of $G$ is $(1+o(1))p^2 \hat{d}$ a.a.s. under some condition on $p\delta$.

**Theorem 11.** Suppose $G$ is a graph, and $G_p$ is a random subgraph where $p$ satisfies $p\delta = \omega(\log(n))$. Then

$$\hat{d}_{G_p} = (1 + o(1))p \hat{d}_G$$
a.a.s.

**Proof.** The proof is very similar to that of Theorem 10.

$$\frac{n}{d_{G_p}} = \sum_{\{u,v\} \in \mathcal{E}(G)} \frac{2}{d'_u d'_v} X_{uv}.$$

Where $d'_u$ and $d'_v$ denote the degrees of vertices in $G_p$, and $X_{uv}$ is the indicator that edge $\{u,v\}$ is in $G$. Note that by Lemma 6, the $d'_v = (1+o(1))d_v$, and we instead consider concentration of

$$Y = \sum_{\{u,v\}} \frac{2}{p^2 d_u d_v} X_{u,v}.$$
considering, now, degrees $d_u$ and $d_v$ in $G$.

$$\mathbb{E}[Y] = \frac{n}{pd_G}.$$  

Chernoff bounds imply that $Y = (1 + o(1)) \frac{n}{pd_G}$, and the result follows. 

These theorems suggest that the structure captured by $\tilde{d}$ is left largely intact by percolation. This also suggests that the best possible bound in Theorem 7 would be of the form

$$|\sigma_H - \sigma| = O \left( \sqrt{\frac{1}{pd}} \right),$$

in particular that it should depend on the edge harmonic average degree and not the minimum degree. This begs the following:

**Question 4.** Is it true that, under the condition that $p\delta$ is sufficiently large,

$$|\sigma_H - \sigma| = O \left( \sqrt{\frac{1}{pd}} \right)?$$

Or even

$$|\sigma_H - \sigma| = O \left( \sqrt{\frac{\log(n)}{pd}} \right),$$

as suggested in the form of the bound in Theorem 7?

**Acknowledgement:** Part of this chapter is based on the paper “The spectral gap of a random subgraph of a graph”, which appeared in Internet Mathematics, volume 4, 2-3, 2007, 225–244, and was written jointly with Fan Chung. This material appears with kind permission of AK Peters.
Chapter 4

Random Spanning Trees in General Graphs

4.1 Introduction

Many information networks or social networks have been observed to have very small diameters, as dictated by the so-called “small world phenomenon”. However, in a recent paper of Liben-Nowell and Kleinberg [LNK08], it was observed that in many social networks, trees in such often have relatively large diameter. Examples of such trees include those resulting from passing information to a small selected number of neighbors in various scenarios such as spam or gossip. (Here, while the data did not necessarily dictate that the observed structure would be a tree, people tended to only pass on the messages once, and thus the resulting observed structures were trees).

A sparse subgraph naturally has very different behavior from its host graph. It is still of interest to understand the connections between a graph and its subgraph. What invariants of the host graph can or cannot be translated to its subgraph? Under what conditions, can we predict the behavior of subgraphs? In particular, how does a random spanning tree behave; do the observations of Liben-Nowell and
Kleinberg differ from that which we would expect from a random tree, or do they indicate a more complex mechanism at work? In this chapter, we would like to address this paradox by examining the diameter of random spanning sub-tree in a given graph \( G \).

A spanning tree \( T \) of a connected graph \( G \) is a subgraph on \( V(G) \), which is isomorphic to a tree. The number of spanning trees is determined by the celebrated matrix-tree theorem of Kirchhoff [Kir47]. If \( A \) denotes the adjacency matrix and \( D \) denotes the diagonal matrix of degrees, the matrix-tree theorem states that the number of spanning tree is equal to the absolute value of the determinant of any \( n - 1 \times n - 1 \) sub-matrix of \( D - A \).

This is clearly a very different model of random subgraphs than studied in Chapters 2 and 3, and hence we expect the properties to be different. For instance, while the edge density of a random spanning tree and a percolated graph near the percolation threshold (as studied in Chapter 2) are similar, the properties of the resulting graph are very different: a spanning tree is necessarily connected and loopless. We expect neither of these properties to hold, in general, for a random subgraph generated with bond percolation with the same edge density. Still, we are interested in what properties of this type of subgraph we can capture, and we observe that many of the same parameters are important for understanding a random subtree as were important in understanding the random subgraph obtained by bond percolation.

The diameter of a subgraph is always larger than or equal to the diameter of \( G \). However, the diameter of a spanning tree could be much larger than the diameter of the graph. The case that the host graph \( G \) is the complete graph \( K_n \) is well-studied in the literature. The number of spanning trees of \( K_n \) is \( n^{n-2} \) by Cayley’s theorem. Rényi and Szekeres [RS67] showed that the diameter of a random spanning tree is of order \( \sqrt{n} \), which contrasts with the fact that the diameter of \( K_n \) is 1.

Motivated by these examples, we ask what is the true story of the diameter of random spanning trees for a general graph. Previously Aldous [Ald90] proved that in a regular graph \( G \) with spectral gap \( \sigma \), the expected diameter of a spanning tree
Figure 4.1: A host graph and a spanning tree within the graph. The diameter of the spanning tree is 51, while the diameter of the underlying host graph is 10.

$T$ of $G$, denoted by $\text{diam}(T)$ has expected value satisfying

$$\frac{c\sigma\sqrt{n}}{\log n} \leq \mathbb{E}(\text{diam}(T)) \leq \frac{c\sqrt{n \log n}}{\sqrt{\sigma}}$$

for some absolute constant $c$, where here log refers to the natural logarithm.

We partially improve Aldous’ result as follows:

**Theorem 12.** For a $d$-regular graph $G$ on $n$ vertices with spectral gap $\sigma$, a spanning tree $T$ of $G$ has expected value satisfying

$$c\sqrt{n} \leq \mathbb{E}(\text{diam}(T)) \leq c' \frac{\sqrt{n \log n}}{\sqrt{\log(1/\sigma)}}$$

for some absolute constants $c$ and $c'$ provided that $d \gg \frac{\log^2 n}{\log^2 \sigma}$.

Theorem 12 is an immediate consequence of the following result for general graphs.

**Theorem 13.** Suppose $G$ is a connected graph on $n$ vertices, with average average degree $d$, minimum degree $\delta$, and second-order average degree $\tilde{d} = \frac{\sum_v d_v^2}{\sum_u d_u}$, and $\epsilon > 0$ is fixed. Suppose the average degree satisfies

$$d \gg \frac{\log^2 n}{\log^2 \sigma}.$$
Then with probability $1 - \epsilon$, the diameter $\text{diam}(T)$ of a random spanning trees $T$ in $G$ satisfies

$$diam(T) \geq (1 - \epsilon + o(1)) \sqrt{\frac{\epsilon nd}{d}}.$$  

(4.1)

and

$$diam(T) \leq \frac{c}{\epsilon} \sqrt{\frac{nd}{\delta \log(1/\sigma)}} \log n.$$  

(4.2)

for some constant $c \leq 10$.

While the conditions look technical, they are derived from the proofs in Sections 4 and 5. It should be noted that the condition
d

$$d \gg \frac{\log^2 n}{\log^2 \sigma}$$

is really a condition on both $d$ and $\sigma$. The smaller the spectral gap $\sigma$ is, the smaller $d$ is allowed to be. We note that the average degree requirement is satisfied for any graph so long as, for instance, the average degree is $\Omega(\log^2 n)$ (a constant multiple of $\log^2 n$ for some constant), and $\sigma = o(1)$. One can observe that the result applies so long as

$$\sigma = 1 - \omega\left(\frac{\log(n)}{\sqrt{d}}\right).$$

Note that compared with the eigenvalue condition for Theorem 3 in Chapter 2, this is an extremely weak condition.

To further see that this is a reasonable condition, we give a few examples of $\sigma$ for some special classes of graphs.

Ideally, we would like to let $\epsilon$ be a function of $n$ tending to zero; this can be done and we discuss the necessary requirements after the proof of Theorem 13. For simplicity we do not state the (somewhat technical) requirements here.

For random $d$-regular graphs, it is known that $\sigma$ is about $\frac{2}{\sqrt{d}}$; indeed Friedman’s proof of the Alon Conjectures [Fri08] implies that

$$\sigma \leq \frac{2\sqrt{d - 1} + \epsilon}{d}$$
for any positive $\epsilon$ as $n$ tends to infinity.

We may also consider the random graph model $G(w)$ for a given expected degree sequence $w = (w_1, w_2, \ldots, w_n)$, as introduced in [CL02]. Recall that in the $G(w)$ model, the probability $p_{ij}$ that there is an edge between $v_i$ and $v_j$ is proportional to the product $w_i w_j$. Namely,

$$p_{ij} = \frac{w_i w_j}{\sum_k w_k} = \frac{w_i w_j}{\text{Vol}(G)}.$$  

(4.3)

Note that this includes the possibility of loops at vertex $v_i$ with probability proportional to $w_i^2$.

It has been shown in [CLV04] that $G(w)$ has $\sigma = (1 + o(1)) \frac{4}{\sqrt{d}}$ provided that the minimum of weights is $\Omega(\log n)$. Theorem 13 implies the diameter of a random spanning tree is $\Omega(\sqrt{\frac{d}{3}n})$ if the average degree is $d = \Omega\left(\frac{\log n}{\log \log n}\right)^2$. The upper bound is within a multiplicative factor of $\sqrt{\frac{d}{3}} \log n$.

It has been observed that many real-world information networks satisfy the so-called power law. We say a graph satisfies a power law with exponent $\beta$ if the degree sequence of the graph satisfies the property that the number of vertices having degree $k$ is asymptotically proportional to $k^{-\beta}$. There are many models being used to capture the behavior of such power law graphs [CL06], especially for the exponent $\beta$ in the range between 2 and 3. We may use the random graph model $G(w)$ with $w$ satisfying the power law to understand typical power law graphs. In random graph model $G((w))$, the maximum degree can be as large as $\sqrt{n}$. (In other words, if the maximum degree exceeds $\sqrt{n}$, then $G(w)$ can only be used to model the subgraph with degree no larger than $\sqrt{n}$.) Also in $G(w)$ the second average degree is of order $d^{3-1} m^{3-\beta}$. Using Theorem 13, the diameter of a random spanning tree in such random power law graph is at least $cn^{(\beta-2)/4}(\log n)^{(2-\beta)/2}$ and at most $c'\sqrt{n}(\log n)^{3/2}$ for some constant $c$ and $c'$.

The remainder of the chapter is organized as follows. In section 2, we will recall some definitions and prove some useful facts on the spectrum of the Laplacian, random walks, and spanning trees. In Section 3, we describe a method of using random walks to generate a uniform spanning tree. In Section 4, we will prove
the lower bound for the diameter of a random spanning tree and give an upper bound in section 5. Finally, we conclude with some discussion as to routes to improvement.

### 4.2 Preliminaries

Recall that the normalized Laplacian of $G$ is

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

If the eigenvalues of $\mathcal{L}$ are

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2,$$

recall that

$$\sigma = \max\{1 - \lambda_1, \lambda_{n-1} - 1\}.$$

The spectral gap, $\sigma$, is such that $\sigma < 1$ if and only if $G$ is connected and not bipartite.

Let $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ be orthonormal eigenvectors of the Laplacian $\mathcal{L}$, $U = (\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$, where $\alpha_i$ is viewed as a column vector. Also we define $\Lambda = \text{diag}(\lambda_0, \ldots, \lambda_{n-1})$. We can write

$$\mathcal{L} = U \Lambda U^T.$$

For $0 \leq i \leq n - 1$, we define $\varphi_i = \alpha_i^T D \alpha_i$. Then we have

**Lemma 8.** The degree spectrum $(\varphi_0, \varphi_1, \ldots, \varphi_{n-1})$ satisfies the following properties.

1. $\varphi_0 = \tilde{d}$.

2. For $0 \leq i \leq n - 1$, $\delta \leq \varphi_i \leq \Delta$.

3. $\sum_{i=0}^{n-1} \varphi_i = \text{vol}(G)$. 
Proof. Note \( \alpha_0 = (\sqrt{\frac{d_1}{\text{vol}(G)}}, \ldots, \sqrt{\frac{d_n}{\text{vol}(G)}})^t \) since \( \mathcal{L} \alpha_0 = 0 \). We have

\[
\varphi_0 = \alpha_0^T D \alpha_0 = \sum_{i=1}^{n} \frac{\sqrt{d_i}}{\text{vol}(G)} \frac{\sqrt{d_i}}{\text{vol}(G)} = \sum_{i=1}^{n} \frac{d_i^2}{\text{vol}(G)} = \tilde{d}.
\]

We have

\[
\left| \varphi_i - \frac{\delta + \Delta}{2} \right| = \left| \alpha_i^T D \alpha_i - \frac{\delta + \Delta}{2} \right| = \left| \alpha_i^T (D - \frac{\delta + \Delta}{2} I) \alpha_i \right| \leq \| D - \frac{\delta + \Delta}{2} I \| = \frac{\Delta - \delta}{2}.
\]

Thus, we have

\[
\delta \leq \varphi_i \leq \Delta.
\]

We also have

\[
\sum_{i} \varphi_i = \text{Tr}(U^T DU) = \text{Tr}(D) = \text{vol}(G).
\]

\[
\square
\]

**Lemma 9.** For any integer \( j \geq 1 \),

\[
\text{Tr}(A(D^{-1}A)^{j-1}) \leq \tilde{d} + \sigma^j (\text{vol}(G) - \tilde{d}).
\]
Proof. We have

\[ \text{Tr}(A(D^{-1}A)^{j-1}) = \text{Tr}(D^{-1}A)^j = \text{Tr}(D^{\frac{1}{2}}AD^{-\frac{1}{2}})^j = \text{Tr}(D(I - L)^j) = \text{Tr}(DU(I - \Lambda)^jU^T) = \text{Tr}(U^TDU(I - \Lambda)^j) = \sum_{i=0}^{n-1} \varphi_i(1 - \lambda_i)^j = \tilde{d} + \sum_{i>0} \varphi_i(1 - \lambda_i)^j \leq \tilde{d} + \sum_{i>0} \varphi_i \sigma^j = \tilde{d} + (\text{vol}(G) - \tilde{d})\sigma^j. \]

\[ \square \]

A simple random walk on $G$ is a sequence of vertices $v_0, v_1, \ldots, v_k, \ldots$ with

\[ \mathbb{P}(v_k = j \mid v_{k-1} = i) = p_{ij} = \begin{cases} \frac{1}{d_i} & \text{if } ij \in E(G) \\ 0 & \text{otherwise} \end{cases} \]

for all $k \geq 1$.

The transition matrix $W$ is an $n \times n$ matrix with entries $p_{ij}$ for $1 \leq i, j \leq n$. We can write $W = D^{-1}A$.

A probability distribution over the set of vertices is a row vector $\beta (\beta^* \in \mathbb{R}^n)$ satisfying

1. The entries of $\beta$ are non-negative.

2. The $L_1$-norm $\|\beta\|_1 (= \beta \mathbf{1})$ equal to 1 where $\mathbf{1}$ denotes a column vector with all entries 1.
If $\beta$ is a probability distribution, so is $\beta W$. The stationary distribution is denoted by $\pi$ satisfying $\pi = \pi W$ and
\[
\pi = \frac{1}{\text{vol}(G)}(d_1, d_2, \ldots, d_n).
\]
The eigenvalues of $P$ are $1, 1 - \lambda_1, \ldots, 1 - \lambda_n$, since $P = D^{-\frac{1}{2}}(I - \mathcal{L})D^{\frac{1}{2}}$. In general, $W$ is not symmetric unless $G$ is regular. The following lemma concerns the mixing rate of the random walks.

**Lemma 10.** For any integer $t > 0$, any $\alpha \in \mathbb{R}^n$, and any two probability distributions $\beta$ and $\gamma$, we have
\[
\langle(\beta - \gamma)P^t, \alpha D^{-1}\rangle \leq \sigma^t \|\beta - \gamma\|_2 \|\alpha D^{-1/2}\|_2.
\]
(4.4)

In particular,
\[
\|\beta - \gamma\|_2 \leq \sigma^t \|\beta - \gamma\|_2.
\]
(4.5)

**Proof.** Let $\varphi_0 = \frac{1}{\sqrt{\text{vol}(G)}}(\sqrt{d_1}, \ldots, \sqrt{d_n}) = \text{Vol}(G)^{-\frac{1}{2}}1D^{\frac{1}{2}}$ denote the (row) eigenvector of $I - \mathcal{L}$ for the eigenvalue $1$. The matrix $(I - \mathcal{L})^t - \varphi^T\varphi$, which stands for the projection of $(I - \mathcal{L})^t$ to the hyperspace $\varphi^\perp$, has $L_2$-norm $\sigma^t$. Note that
\[
(\beta - \gamma)D^{-\frac{1}{2}}\varphi = \frac{1}{\text{Vol}(G)}(\beta - \gamma)1 = 0.
\]
We have
\[
\langle(\beta - \gamma)P^t, D^{-1}\alpha\rangle = (\beta - \gamma)D^{-\frac{1}{2}}[(I - \mathcal{L})^t - \xi\xi^T]D^{-\frac{1}{2}}\alpha
\leq \|\beta - \gamma\|_2 \|\alpha\|_2 \sigma^t \|\alpha D^{-\frac{1}{2}}\|_2.
\]
Now we choose $\alpha = [(\beta - \gamma)P^t]^T$, obtain (4.5) as desired.

The mixing rate of the random walks on $G$ measures how fast $\beta P^t$ converges to the stationary distribution $\pi$ from an initial distribution $\beta$. We can use the above lemma to show that the distribution $\beta P^t$ converges to $\pi$ rapidly if $\sigma$ is strictly less than 1.
4.3 Random spanning trees generated by random walks

The following so-called Groundskeeper algorithm gives a method of generating spanning trees: Start a random walk at a vertex, $v$. The first time a vertex is visited, we observe the edge it was visited on and add that edge to our spanning tree. Once the graph is covered, the resulting set of edges form a spanning tree. This gives a map $\Phi$ from random walks to random spanning trees. Aldous [Ald90] and Broder [Bro89] independently show that the Groundskeeper algorithm generates a uniform spanning tree:

**Theorem 14 (Groundskeeper Algorithm).** The image of $\Phi$ is uniformly distributed over all spanning trees. It is independent of the choice of initial vertex $v$.

![Groundskeeper Algorithm](image)

Figure 4.2: The Groundskeeper Algorithm: Run a random walk, record the first edge a node is visited through.

The Groundskeeper’s algorithm generates a uniform spanning tree in the cover time of the graph. In this sense, it is not the best possible; Wilson’s algorithm...
[Wil96] uses loop-erased random walks in order to generate spanning trees faster than the cover time. Wilson’s algorithm is as follows:

Two arbitrary vertices \( v_1 \) and \( v_2 \) are chosen; a loop erased random walk is started at \( v_1 \) and run until it hits \( v_2 \), the resulting path is added to the tree. Then, repeatedly, a vertex off of the path, is chosen and a loop erased random walk is started from the new vertex until it hits the path. Since all loops are erased when they are created, this algorithm results in a tree; and indeed this tree is uniformly distributed over spanning trees.

For analysis purposes however, the Groundskeeper’s algorithm is preferable for us as random walks are simpler to study than loop-erased random walks.

One method of using the Groundskeeper algorithm to bound the diameter of a random spanning tree is to measure the distance between a random walk and the starting point of a random walk after some number of steps. This will provide a lower bound on the diameter. As the walker moves about the graph, its distance from the root increases except in the case where it completes a loop. We thus are motivated to examine the situations where there are or aren’t long loops in the graph; in particular in the remainder of this section we prove a result on how long it takes for a long loop to appear in a graph.

We pick up a random initial vertex with stationary distribution \( \pi \). Then at any step \( t \), the distribution remains the same \( p_t = \pi \).

For an integer \( g \geq 3 \), consider the following \( g \)-truncated random walks. We construct a random spanning tree by collecting edges \( v_{t-1}v_t \) if \( v_t \) is first visited. We allow the backtrack step \( v_{t+1} = v_{t-i} \) for some \( i \leq g - 2 \). However, if \( v_{t+1} = v_{t-i} \) for some \( i > g - 2 \), the random walk stops.

**Lemma 11.** The probability that a \( g \)-truncated random walk stops before or at time \( t \) is at most

\[
\frac{(t - g + 3)(t - g + 2)d^2}{2nd} + (t - k)\frac{\sigma^g}{1 - \sigma}.
\]

*Proof.* When the truncated random walk stops, there exists a closed walk \( C = \)
\(v_i, v_{i+1}, \ldots, v_t, v_{i+k} = v_i\) of length \(k \geq g\) for some \(0 \leq i \leq t - k + 1\). For a fixed \(i\) and \(k\), the probability \(f(i, k)\) such a closed walk occurs is at most

\[
f(i, k) \leq \sum_{\text{closed walk}: v_i, \ldots, v_{i+k} = v_i} \frac{d_i}{\text{vol}(G)} \prod_{j=1}^{k} \frac{1}{d_{v_{i+j-1}}}
\]

\[
= \frac{1}{\text{vol}(G)} \text{Tr}(A(D^{-1}A)^{k-1})
\]

\[
\leq \frac{\bar{d}}{\text{vol}(G)} + \sigma^k (1 - \frac{\bar{d}}{\text{vol}(G)})
\]

\[
< \frac{\bar{d}}{\text{vol}(G)} + \sigma^k.
\]

By summing up for \(i \geq 0, k \geq g,\) and \(i + k \leq t + 1\), we have

\[
\sum_{i=0}^{t-g+1} \sum_{k=g}^{t-i+1} f(i, k) = \sum_{i=0}^{t-g+1} \sum_{k=g}^{t-i+1} \frac{\bar{d}}{\text{vol}(G)} + \sigma^k
\]

\[
\leq \frac{(t-g+3)(t-g+2)}{2} \frac{\bar{d}}{\text{vol}(G)} + \sum_{i=0}^{t-g+1} \sum_{k=g}^{\infty} \sigma^k
\]

\[
\leq \frac{(t-g+3)(t-g+2)}{2} \frac{\bar{d}}{\text{vol}(G)} + (t-g+2) \frac{\sigma^g}{1-\sigma}.
\]

\[\square\]

### 4.4 Proving a diameter Lower Bound for random spanning trees

In this section we will prove a diameter lower bound for spanning trees of \(G\) as stated in inequality (4.1) of Theorem 13. As alluded to above, the strategy is to run the walk for some number of steps \(t \approx \sqrt{\frac{\epsilon d}{\bar{d}}}, \) and show that with high probability no long loops (length \(\approx \log(n)\)) occur. Then, we show that the expected distance from the root is high, and apply a martingale argument to show that, with high probability the walker is far away from the root after such a number of steps.

**Proof of (4.1) in Theorem 13.** Let \(t = (1 - \epsilon)\sqrt{\frac{\epsilon d}{\bar{d}}}n\) and \(g = \lceil \frac{\log(\frac{(1-\sigma)\sqrt{\tau}}{4\sqrt{4\bar{d}}})}{\log(\sigma)} \rceil\). Note
that $g$ is chosen so that
\[ \frac{\sigma^g}{1 - \sigma} \leq \frac{\epsilon}{4t}. \]

Apply the $g$-truncated random walk. By Lemma 8, the $g$-truncated random walk will survive up to time $t$ with probability at least
\[ 1 - \frac{(t - g + 3)(t - g + 2)\tilde{d}}{2nd} - (t - g + 2)\frac{\sigma^g}{1 - \sigma} > 1 - \frac{t^2\tilde{d}}{2nd} - t\frac{\sigma^g}{1 - \sigma} \]
\[ > 1 - \frac{\epsilon}{2} - \frac{\epsilon}{4} \]
\[ \geq 1 - \frac{3\epsilon}{4}. \]

For $i = 1, \ldots, t$, we say $v_{i-1}v_i$ is a forward step if $v_i \neq v_j$ for some $j < i$; we say $v_{i-1}v_i$ is a $k$-backward step if $v_i = v_{i-k}$ for some $k \leq g - 2$.

Let $X_i = -k$ if $v_{i-1}v_i$ is a $k$-backward step and $X_i = 1$ otherwise. For all $i$, we have
\[ -(g - 2) \leq X_i \leq 1. \]

Let $Y$ be the distance of $v_0v_t$ in the random spanning tree and $X = \sum_{i=1}^t X_i$ Conditioning on that the truncated random walk survives up to time $t$, we have $Y \geq X$. Or equivalently,
\[ \mathbb{P}(Y < X) < \frac{3\epsilon}{4}. \]

Let $\mathcal{F}_i$ be the $\sigma$-algebra that $v_0, \ldots, v_i$ is revealed. For $i = 0, \ldots, t$, $\mathbb{E}(X \mid \mathcal{F}_i)$ forms a martingale. We would like to establish a Lipschitz condition for this martingale. For $1 \leq i, j \leq t$, it is enough to bound $|\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})|$. For $j < i$, $X_j$ is completely determined by the information on $v_0, v_1, \ldots, v_i$. In this case we have
\[ \mathbb{E}(X_j \mid \mathcal{F}_i) = \mathbb{E}(X_j \mid \mathcal{F}_{i-1}). \]

For $j \geq i$, $\mathbb{E}(X_j \mid \mathcal{F}_i)$ and $\mathbb{E}(X_j \mid \mathcal{F}_{i-1})$ are different because $v_i$ is exposed. For $i \leq j \leq i + 2g - 3$, we apply the trivial bound
\[ |\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})| \leq g - 1. \]

For $j \geq i + 2g - 2$, $X_j$ only depends on $v_{j-g+2}, v_{j-g+3}, \ldots, v_{j+1}$. Note that the random walk at step $i$ only depends on the current position $v_i$ and is independent
of history positions \(v_0, \ldots, v_{i-1}\). Thus \(\mathbb{E}(X_j \mid v_{j-g+2})\) is independent of \(v_i\) because \(i < j - g + 2\). We use the mixing of our random walk to show that information gained from knowing \(v_i\) is quickly lost. Let \(p\) be the distribution of \(v_i\) given \(v_{i-1}\) and \(q\) be the distribution of \(v_i\) given \(v_i\) (\(q\) is a singleton distribution). Let \(p'\) be the distribution of \(v_{j-g+2}\) given \(v_{i-1}\) and \(q'\) be the distribution of \(v_{j-g+2}\) given \(v_i\). (Note: Here \(p'\) is not \(p\) transposed.) We have

\[
\|p' - q'\|D^{-1/2} \leq \|p - q\|D^{-1/2}\|\sigma^{j-g+2-i} \leq \frac{2}{\sqrt{\delta}}\sigma^{j-g+2-i}.
\]

Therefore,

\[
|\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})| = |\sum_{u=1}^{n} (p'_u - q'_u)\mathbb{E}(X_j \mid v_{j-g+2} = u)|
\leq \|p' - q'\|_1(g - 2)
\leq \sqrt{\text{vol}(G)}\|p' - q'\|D^{-1/2}(g - 2)
\leq 2(g - 2)\sqrt{\text{vol}(G)}\sigma^{j-g+2-i}.
\]

We have

\[
|\mathbb{E}(X \mid \mathcal{F}_i) - \mathbb{E}(X \mid \mathcal{F}_{i-1})| \leq \sum_{j=1}^{t} |\mathbb{E}(X_j \mid \mathcal{F}_i) - \mathbb{E}(X_j \mid \mathcal{F}_{i-1})|
\leq 2(g - 1)^2 + \sum_{j=i+2g-2}^{j} 2(g - 2)\sqrt{\text{vol}(G)}\sigma^{j-g+2-i}
\leq 2(g - 1)^2 + 2(g - 2)\sqrt{\text{vol}(G)}\sigma^g
\leq 3g^2,
\]

noting that \(g\) has been chosen so that

\[
\frac{\sigma^g}{1 - \sigma} = \frac{\sqrt{\epsilon}\sqrt{\delta}}{4(1 - \epsilon)\sqrt{\text{vol}(G)}}
\]

is sufficiently small to make the last inequality hold.

Thus we have established that \(\mathbb{E}(X \mid \mathcal{F}_i)\). By applying the Azuma-Hoeffding inequality (Proposition 6 from Chapter 1), we have

\[
\mathbb{P}(X - \mathbb{E}(X) < -\alpha) < e^{-\frac{\alpha^2}{18g^2t}}
\]
Note that

\[ E(X) = \sum_{i=1}^{t} E(X_i) \]

\[ = \sum_{i=1}^{t} \sum_{j=1}^{n} E(X_i | v_{i-1} = j) P(V_{i-1} = j) \]

\[ \geq \sum_{i=1}^{t} \sum_{j=1}^{n} \left( (1 - \frac{g-1}{d_j}) + \sum_{k=1}^{g-2} \frac{-k}{d_j} \right) \frac{d_j}{\text{vol}(G)} \]

\[ = \sum_{i=1}^{t} \sum_{j=1}^{n} \left( 1 - \frac{g(g-1)}{2d_j} \right) \frac{d_j}{\text{vol}(G)} \]

\[ = \sum_{i=1}^{t} \left( 1 - \frac{g(g-1)n}{2\text{vol}(G)} \right) \]

\[ = (1 - \frac{g(g-1)}{2d})t. \]

By choosing \( \alpha = \sqrt{18g^4t \log \frac{4}{\epsilon}} \), we have

\[ P \left( X < (1 - \frac{g(g-1)}{2d})t - \sqrt{18g^4t \log \frac{4}{\epsilon}} \right) < \frac{\epsilon}{4}. \]

Putting all together, we have

\[ P \left( Y < (1 - \frac{g(g-1)}{2d})t - \sqrt{18g^4t \log \frac{4}{\epsilon}} \right) \]

\[ \leq P \left( X < (1 - \frac{2}{d})t - \sqrt{18g^4t \log \frac{4}{\epsilon}} \right) + P(Y < X) \]

\[ < \frac{3\epsilon}{4} + \frac{\epsilon}{4} \]

\[ = \epsilon. \]

To complete the proof, it suffices to check that our degree conditions imply

\[ (1 - \frac{g(g-1)}{2d})t - \sqrt{18g^4t \log \frac{4}{\epsilon}} = (1 - \epsilon - o(1))\sqrt{\frac{nd}{d}}. \]

In particular it suffices to check that

\[ \frac{g}{\sqrt{d}} = o(1), \]
as \( g^4 \) is clearly \( o(t) \).

Since

\[
g \leq \frac{\log \left( \frac{4}{\epsilon (1 - \sigma)} \right) + \log \left( \frac{t \tilde{d}}{\delta} \right)}{\log(1/\sigma)} + 1
\]

and

\[
\log \left( \frac{t \tilde{d}}{\delta} \right) = \log((1 - \epsilon) \sqrt{\epsilon}) + \frac{1}{2} \log \left( \frac{d}{\delta} n \right)
\]

we have \( g/\sqrt{d} = o(1) \), since

\[
d \gg \frac{\log^2(n)}{\log^2(1/\sigma)}
\]

as hypothesized.

\[\square\]

The parameters \( t \) and \( g \) chosen above seem quite technical, and perhaps seem magical in the first read. Still, the rough form of \( t \), that of \( \sqrt{\frac{nd}{d}} \) can be motivated in several ways. After the proof of the upper bound, we conclude this chapter with a discussion on what the correct upper bound should be, and show how \( O(\frac{nd}{d}) \) seems a natural improvement from \( \sqrt{\frac{nd}{d}} \). It is still open, however, as to whether this improvement can be fully realized. The heuristic explained in this final section, however, explains why \( t \) is the right length of time to avoid long loops.

The proof actually yields a slightly stronger result, allowing \( \epsilon \) to vary with \( n \), though the precise necessary conditions are quite messy due to complicated relationships between \( \epsilon, \sqrt{\frac{nd}{d}} \) and \( \sigma \). Note that it is required that

\[
\log((1 - \epsilon) \sqrt{\epsilon}) = o(\sqrt{d}).
\]

This follows from our requirement that \( \frac{g}{\sqrt{d}} = o(1) \) and (4.6). Additionally, there is a requirement that

\[
\frac{g^4 \log \epsilon}{(1 - \epsilon) \sqrt{\epsilon^{\frac{nd}{d}}}} = o(1).
\]

Note that so long as \( \epsilon = \omega(\frac{1}{\log n}) \) these conditions are satisfied, however for a wide variety of graphs, \( \epsilon \) can be even smaller.
4.5 Proof of Upper Bound

For the upper bound, we follow the general strategy of Aldous in [Ald90]. In particular we provide a (relatively straightforward) generalization of theorem 15 of Aldous’ paper to give an upper bound in the general degree case.

Here, we let $X_t$ denote the position of a random walk at time $t$. We denote by $T_B$ the hitting time of a set $B$; that is

$$T_B = \min\{t : X_t \in B\}.$$ 

We denote the return time of a set $B$ to be

$$T_B^+ = \min\{t \geq 1 : X_t \in B\}.$$ 

(Note that if the random walker does not start in $B$, $T_B = T_B^+$.)

When considering the probability that our random walk has some property we use the notation $\mathbb{P}_\rho$ to denote that we condition on our random walker having initial distribution $\rho$. Likewise, $\mathbb{E}_\rho$ denotes expectation conditioning on the initial distribution. If no distribution is given, it is assumed to be starting from the stationary distribution. As a convenient abuse of notation we let $\mathbb{P}_v$, for some vertex $v$, denote the probability starting with the distribution that places weight 1 on $v$.

The first tool is the following, rather standard, mixing lemma

**Lemma 12.** For all initial distributions $\rho$ and all $B \subseteq G$, there exists an (absolute) constant $K$

$$\mathbb{P}_\rho \left( T_B > 3 \frac{\log n}{\log(1/\sigma)} \frac{\text{Vol}(G)}{\text{Vol}(B)} \right) \leq \frac{1}{2}$$

**Proof.** We begin by bounding $|P^*(\rho, B) - \pi(B)|$; where $\rho$ is an (arbitrary) initial distribution and $\pi(B) = \frac{\text{Vol}(B)}{\text{Vol}(G)}$. Write

$$\rho^* D^{-1/2} = \sum_i \alpha_i \varphi_i^*$$
where the $\varphi_i^*$ are left eigenvectors of $(I - L)$ corresponding to eigenvalues $(1 - \lambda_i)$.

Then

$$\alpha_0 = \langle \rho^* D^{-1/2}, \frac{D^{1/2}1}{\sqrt{\text{vol}(G)}} \rangle = \frac{1}{\sqrt{\text{vol}(G)}}.$$ 

Thus

$$\rho^* D^{-1/2} = \frac{1^* D^{1/2}}{\text{vol}(G)} + \sum_{i \geq 1} \alpha_i \varphi_i.$$ 

Then

$$|P^s(\rho, B) - \pi(B)| = \left| \rho^* P^s 1_B - \frac{1^* D}{\text{vol}(G)} 1_B \right|$$

$$= \left| \left( \rho^* D^{-1/2}(I - L)^s - \frac{1^* D^{1/2}}{\text{vol}(G)} \right) D^{1/2} 1_B \right|$$

$$= \left| \left( \frac{1^* D^{1/2}}{\text{vol}(G)} + \sum_{i \geq 1} (1 - \lambda_i)^s \alpha_i \varphi_i^* - \frac{1^* D^{1/2}}{\text{vol}(G)} \right) D^{1/2} 1_B \right|$$

$$\leq \sum_{i \geq 1} \sigma_s^{|\alpha_i|} |\varphi_i^* D^{1/2} 1_B|$$

$$\leq \sigma_s \sqrt{n \text{Vol}_{1/2}(B)}$$

$$\leq \sigma_s \sqrt{n |B| \text{Vol}(B)}$$

where the last step follows from an application of Cauchy-Schwarz inequality.

Let

$$s = \log \left( \frac{\sqrt{\text{vol}(B)}}{2 \sqrt{n |B| \text{Vol}(G)}} \right) / \log(\sigma)$$

so

$$\sigma^s \sqrt{n |B| \text{Vol}(B)} = \frac{\text{vol}(B)}{2\text{vol}(G)}.$$ 

Fix $t_i = is$, then

$$\mathbb{P}(T_B > x) \leq \mathbb{P}(X_{t_1} \notin B, X_{t_2} \notin B, \ldots, X_{t_{x/s}} \notin B)$$

$$= \mathbb{P}(X_{t_1} \notin B) \mathbb{P}(X_{t_2} \notin B | X_{t_1} \notin B) \ldots \mathbb{P}(X_{t_{x/s}} \notin B | X_{t_j} \notin B, \forall j < i)$$

$$\leq \left( 1 - \frac{\text{vol}(B)}{\text{vol}(G)} + \sqrt{n |B| \text{Vol}(B)} \sigma^s \right)^{x/s}$$

$$\leq \left( 1 - \frac{\text{vol}(B)}{2\text{vol}(G)} \right)^{x/s}.$$
Fix \( x = 2 \log(2) s \frac{\text{vol}(G)}{\text{vol}(B)} \) and it is easy to check that
\[
P(T_B > x) \leq \frac{1}{2}.
\]

In all, we have
\[
x = 2 \log(2) \frac{\log(\frac{2\sqrt{n|B|\text{vol}(G)}}{\sqrt{\text{vol}(B)}}) \text{vol}(G)}{\log(1/\sigma) \text{vol}(B)} \leq 3 \log n \frac{\text{vol}(G)}{\text{log}(1/\sigma) \text{vol}(B)}
\]

\[\square\]

The following result (and it’s proof) are due to Aldous [Ald90]. Let \( B = \{v_0, \ldots, v_c\} \) denote a set of vertices and let \( \mathcal{P}_B \) denote the event that the path from \( v_0 \) to the root (starting location of our random walker for generating a random spanning tree, chosen by the uniform distribution) in a uniform spanning tree starts \( v_0, v_1, \ldots, v_c \). Then

**Lemma 13.**
\[
P(T_{v_c} = \ell | \mathcal{P}_B) = \frac{\mathbb{P}_{v_c}(T_B^+ > \ell)}{\mathbb{E}_{v_c}(T_B^+)}
\]

**Proof.** For \( i < c \), we denote the event \( \mathcal{D}_i \) to be
\[
\mathcal{D}_i = \{T_{\{v_0, \ldots, v_i\}} = T_{v_i}, X_{T_{v_i}-1} = v_{i+1}\}.
\]

In words, \( \mathcal{D}_i \) is the event that \( v_i \) is hit before \( v_j \) for \( j < i \), and indeed \( v_i \) is first hit from \( v_{i-1} \), so \( \bigcap_{i<i} \mathcal{D}_i = \mathcal{P}_B \). Then:
\[
\{T_{v_c} = \ell\} \cap \mathcal{P}_B = \{T_{v_c} = \ell = T_B\}.
\]

Note that, from the Markov property, it is clear that \( P(\bigcap_{i<i} \mathcal{D}_i | T_{v_c} = T_B = \ell) \) does not depend on \( \ell \) (this is the motivation for writing \( \mathcal{P}_B \) in an obtuse way), thus:
\[
P(T_{v_c} = \ell | \mathcal{P}_B) = \alpha P(T_{v_c} = T_B = \ell)
\]
for \( \ell = 0, 1, \ldots \) and for some \( \alpha \) which (critically) does not depend on \( \ell \). We have that:

\[
P_v(T_{B} > \ell) = \sum_w P_v(X_0 = v, X_\ell = w, X_i \notin B \text{ for } 1 \leq i \leq \ell)
\]

\[
= \frac{1}{\pi(v)} \sum_w P_v(X_0 = v, X_\ell = w, X_i \notin B \text{ for } 1 \leq i \leq \ell)
\]

\[
= \frac{1}{\pi(v)} \sum_w P_v(X_0 = w, X_\ell = v, X_i \notin B \text{ for } 1 \leq i \leq \ell)
\]

\[
= \frac{1}{\pi(v)} P_v(T_\ell = v, X_i \notin B \text{ for } 1 \leq i \leq \ell)
\]

\[
= \frac{1}{\pi(v)} P_v(T_\ell = \ell)
\]

with the third to last equality following from time reversal for the stationary Markov chain. This implies:

\[
P_\pi(T_v = \ell | P_B) = \alpha \pi(v) P_v(T_B^+ > \ell).
\]

Note finally, then that

\[
1 = \sum_{\ell=0}^\infty P_\pi(T_v = \ell | P_B) = \sum_{\ell=0}^\infty \alpha \pi(v) P_v(T_B^+ > \ell) = \alpha \pi(v) E_v(T_B^+).
\]

so \( \alpha \pi(v) = E_v(T_B^+) \), implying the result.

One can observe that, actually, that while the normalizing constant is easy to compute, the exact value is unnecessary for the proof of the upper bound itself.

We now prove the upper bound, establishing (4.2) in Theorem 13; whose proof mimics that of Aldous.

**Proof of (4.2) in Theorem 13:** Let us start our random walk from the stationary distribution (unless explicitly noted, all probabilities related with the random walk which generates the spanning tree are taken to start with \( \pi \)).

We begin by fixing a path \( v_0, v_1, \ldots, v_c \) in our graph; and \( B \) be the set

\[
B = \{v_0, \ldots, v_c\}.
\]
As above, \( \mathcal{P}_B \) will denote the event that the path from \( v_0 \) to the root (that is, the starting location of our random walk, \( X_0 \)) in our uniform spanning tree starts out along the path \( v_0, \ldots, v_c \). If we let

\[
s = \left\lceil \frac{3 \text{Vol}(G)}{\log(1/\sigma) (c+1)\delta} \right\rceil \geq \frac{3 \text{vol}(G)}{\log(1/\sigma) \text{vol}(B)} \log n
\]

then, by iterating lemma 12 we have that

\[
P_{v_c}(T^+_B > js) \leq \frac{1}{2^{j-1}} P_{v_c}(T^+_B > s).
\]

We are now in the position to apply lemma 13 to both sides; note that the normalizing constant will cancel and we are left with:

\[
P(T_{v_c} = js|\mathcal{P}_B) \leq (1/2)^{j-1} P(T_{v_c} = s|\mathcal{P}_B) \leq (1/2)^{j-1} s^{-1}.
\]

where the last inequality follows from the fact that the right hand side of (13) in lemma 13 is decreasing with \( l \) and hence the left hand side must as well. This monotonicity property, and summing gives:

\[
P(js \leq T_{v_c} \leq (j+1)s|\mathcal{P}_B) \leq (1/2)^{j-1}.
\]

Further summing gives

\[
P(js \leq T_{v_c}|\mathcal{P}_B) \leq (1/2)^{j-2}.
\]

If \( \mathcal{P}_B \) occurs; then naturally we have that the distance from \( v_0 \) to the root, \( X_0 \), satisfies

\[
d(X_0, v_0) \leq d(X_0, v_c) + c \leq T_{v_c} + c.
\]

We also clearly have that if \( d(X_0, v_0) > c \), then \( \mathcal{P}_B \) occurs for some unique path \( v_0, \ldots, v_c \). Thus:

\[
P(d(X_0, v_0) > c + js) \leq (1/2)^{j-2}.
\]

Clearly \( \text{diam}(T)/2 \leq \max_v d(X_0, v) \); so

\[
P(\text{diam}(T)/2 > c + js) \leq n(1/2)^{j-2}.
\]

This gives us

\[
\mathbb{E}(\text{diam}(T)) \leq 2c + 3s \log n \leq 2c + \frac{3\text{vol}(G)}{c \log(1/\sigma) \delta} \log^2(n),
\]
with the second inequality coming from the definition of $S$. These terms are the same order of magnitude when setting $c = \sqrt{\frac{\text{vol}(G)}{\delta \log(1/\sigma)}} \log n$; giving the desired bound. To establish the bound in the form stated in (4.2), simply apply Markov’s inequality. □

Note that by minimizing

$$2c + \frac{3\text{vol}(G)}{c \log(1/\sigma) \delta} \log^2(n)$$

we actually get that

$$\mathbb{E}(\text{diam}(T)) \leq 2\sqrt{\frac{\text{vol}(G)}{\delta \log(1/\sigma)}} \log n.$$

### 4.6 Discussion

It would be of significant interest to close the gap between the upper and lower bounds of Theorem 13. In particular, it would be interesting to replace $\delta$ in the upper bound with some type of average degree. While replacing it with $\tilde{d}$ would be particularly nice, perhaps it is too much to hope for. It seems plausible, however, that a bound of the form $\sqrt{\frac{nd}{d}}$ (perhaps with some logarithmic factors) should be possible. In this section, we prove some partial results in this direction. This also ties into the choice of $t$ in the proof of the lower bound of Theorem 13.

In particular, we prove the following: First, that (regardless of where we start our random walk) in the initial part of the walk, our walker hits a set of high volume that will be visited often. We prove concentration on the volume of this set which is hit early in the random walk. Second, we note that this implies that there is a set $S$, such that the diameter of $S$ in the spanning tree is small, and that the random walk is never away from $S$ for long prior to the cover time of the graph a.a.s.

We make use of the following inequality:
Lemma 14. Suppose $x, y = o(n^{-1})$. Then

$$(1 - (x + y))^n - (1 - x)^n(1 - y)^n = o((1 - x)^n(1 - y)^n)$$

Proof. We calculate:

$$(1 - (x + y))^n - (1 - x)^n(1 - y)^n = (1 - x)^n(1 - y)^n - 1 = (1 - x)^n(1 - y)^n\left(\frac{1 - o(n^{-1})}{1 - o(n^{-1})} - 1\right) = (1 - x)^n(1 - y)^n(e^{-o(1)} - 1) = o((1 - x)^n(1 - y)^n)$$

\[\square\]

Theorem 15. Let $G$ be a graph on $n$ vertices with spectral gap $\sigma$. Then there exists a subset of vertices $S \subseteq V(G)$ such that:

1. $\text{vol}(S) = \Omega(\sqrt{\text{vol}(G)})$

2. A random walker is never away from $S$ for more than $O(\frac{1}{\lambda}\sqrt{\frac{nd}{d}\log^2(d)})$ steps a.a.s.

3. The diameter of $S$ in $T$ is at most $O(\frac{1}{\lambda}\sqrt{\frac{nd}{d}\log^2(d)})$ steps a.a.s.

Proof. First, let us handle the case where there is a vertex $v$ with degree at least $c\sqrt{\text{vol}(G)}$. In this case we can take $v$ to be the set $S$, and can skip the set building part of our idea. We can thus assume that $d_v = o(\sqrt{\text{vol}(G)})$ for all $v$ without loss of generality.
We generate a random spanning tree via the Groundskeeper algorithm; let 
\((X_t)_{t \geq 0}\) be a random walk started from the uniform distribution on \(G\). Let \(S_k\) denote the set of vertices visited by time \(\frac{1}{\lambda} k \log^2(n)\), we wish to show that \(\text{vol}(S_k)\) is large for \(k = \sqrt{\text{vol}(G)/\bar{d}}\). Let \(S'_k\) be the set of vertices visited at multiples of \(\frac{1}{\lambda} \log^2(n)\); for notational convenience let \(t_j = j \log^2(n)\); it is this set \(S'_k\) which we show has large volume a.a.s.

Note that (with our choice of \(k\) and restriction on the maximum degree)

\[
\left| \mathbb{P}(X_{t_j} = v | X_{t_{j-1}} = u) - \frac{d_v}{\text{vol}(G)} \right| \ll n^{-\log(n)}
\]

for any \(u\) and \(v\) by rapid mixing; thus

\[
\mathbb{P}(v \notin S'_k) = (1 - \frac{d_v}{\text{vol}(G)} + O(n^{-\log(n)}))^k
\]

\[
\leq e^{-k\left(\frac{d_v}{\text{vol}(G)} + O(n^{-\log(n)})\right)}
\]

\[
\leq 1 - k\frac{d_v}{\text{vol}(G)} + O(kn^{-\log(n)}) + k\frac{d^2_v}{\text{vol}(G)^2}
\]

Thus:

\[
\mathbb{P}(v \in S'_k) \geq k\frac{d_v}{\text{vol}(G)} - k^2\frac{d^2_v}{\text{vol}(G)^2} + O(kn^{-\log(n)}).
\]

The expected volume of \(S'_k\) is thus at least

\[
\mathbb{E}[\text{vol}(S'_k)] = \sum_v d_v \mathbb{P}(v \in S'_k) \geq k\bar{d} - k^2 \frac{\sum_i d^3_i}{\text{vol}(G)^2} + O(k\text{vol}(G)n^{-\log(n)}) \sim k\bar{d}.
\]

With our choice of \(k\) we get we get that

\[
\mathbb{E}[\text{vol}(S'_k)] \sim \sqrt{d\text{vol}(G)}.
\]

For the moment, let us assume that \(\text{vol}(S'_k)\) is concentrated on its mean and hence \(\text{vol}(S_k) \geq \text{vol}(S'_k) \geq C\sqrt{d\text{vol}(G)}\) for some constant \(C\), finish the proof, and then return to the problem of showing concentration.

Let \(t_j = \frac{4j}{X} \log(n)\); and consider:

\[
\mathbb{P}(X_{t_j} \in S | X_{t_{j-1}} = u) = \left(1 - \frac{\text{vol}(S_k)}{\text{vol}(G)} + O\left(\frac{1}{n^4}\right)\right).
\]
Let \( A \) denote the event that our random walk is away from \( S \) for more than \( t \ell \) steps. We can bound this probability as

\[
P(A) \leq (1 - \frac{\text{vol}(S_k)}{\text{vol}(G)} + O(\frac{1}{n^4}))^\ell
\]

\[
\leq \exp\left(-\ell\left(\frac{\text{vol}(S_k)}{\text{vol}(G)} + O(\frac{1}{n^4})\right)\right)
\]

\[
\leq (1 + tO(n^{-4}))e^{-\ell \text{vol}(S_k) / \text{vol}(G)}
\]

\[
\leq (1 + tO(n^{-4}))e^{-\ell \sqrt{\tilde{d} / \text{vol}(G)}}
\]

Setting \( \ell = 4 \sqrt{\frac{\text{vol}(G)}{\tilde{d}}} \log(n) \), we see that the probability of not returning to \( S \) for \( t \ell \) steps is \( O(n^{-4}) \). Since the expected cover time of any graph is most \( O(n^3) \), we have that a.a.s. our random walk never strays from \( S \) for more than \( t \ell = \frac{16}{\lambda} \sqrt{\text{vol}(G) / \tilde{d} \log^2(n)} \) steps. Since \( S \) is comprised of the first \( \frac{1}{\lambda} \sqrt{\text{vol}(G) / \tilde{d} \log^2(n)} \) vertices visited by the random walk, the diameter of \( S \) is bounded by \( |S| \). Combining estimates, we get what is claimed above.

To complete the proof it suffices, by Chebyshev’s inequality (Proposition 2 from Chapter 1), to show that \( \text{Var}(\text{vol}(S_k')) = o(\text{E}^2[\text{vol}(S_k)]] \). As a first step, note that

\[
P(v, u \not\in S_k') = (1 - \frac{d_u + d_v}{\text{vol}(G)} + O(n^{-\log n}))^k.
\]

Let \( X_u \) be the indicator of the event that \( u \in S_k' \), then

\[
\text{Var}(\text{vol}(S_k')) = \sum_v d_v^2 \text{E}[X_v^2] + \sum_{u \neq v} d_u d_v (\text{E}[X_u X_v] - \text{E}[X_u] \text{E}[X_v]) - \sum_v d_v^2 \text{E}[X_v^2].
\]

\[
\leq o(\sqrt{d \text{vol}(G)} \text{E}[\text{vol}(S_k')]) + \sum_{u \neq v} d_u d_v (\text{E}[X_u X_v] - \text{E}[X_u] \text{E}[X_v])
\]

\[
\leq o(\text{E}^2[\text{vol}(S_k')]) + \sum_{u \neq v} \left(d_u d_v \left(1 - \frac{d_u + d_v}{\text{vol}(G)} + O(n^{-\log n})\right)^kight.
\]

\[
- \left(1 - \frac{d_u}{\text{vol}(G)} + O(n^{-\log n})\right)^k \left(1 - \frac{d_v}{\text{vol}(G)} + O(n^{-\log n})\right)^k)
\]

\[
= o(\text{E}^2[\text{vol}(S_k')]) + \sum_{u \neq v} d_u d_v (o(\text{E}[X_u] \text{E}[X_v]) = o(\text{E}^2[\text{vol}(S_k')]);
\]

where we apply Lemma 14 to estimate the covariance term. \( \square \)
This suggests that a truly long path should be unlikely: in order for a truly long path to form, a random walker must repeatedly move out from the core, intersect the end of an already long path, and continue to build on it for a positive fraction of the time which it is away from the core. However, a proof of this has eluded us, and would be of great interest. In particular this begs the following question:

*Question 5.* Is it true that the $\delta$ in Equation (4.2), the upper bound in Theorem 13, can be replaced by the average degree $d$? By the second order average degree $\tilde{d}$?

*Acknowledgement:* Part of this chapter is based on the paper “Diameter of random spanning trees in a given graph”, written jointly with Fan Chung and Linyuan Lu. This paper has been submitted.
Bibliography


