UNIVERSITY OF THE WITWATERSRAND

MASTER'S DISSERTATION

Threshold Functions of Colorings of Random Graphs

Student: Kim Lucas (602808)

Supervisor: Dr. Ronald J. Maartens

Co-Supervisor: Prof. Elizabeth Jonck

A Master's dissertation submitted to the Faculty of Science in fulfilment of the requirements for the degree of Master's of Science.

June 8, 2023

Declaration

I, Kim Lucas, declare that this dissertation titled "Threshold Functions of Colorings of Random Graphs" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the dissertation is based on work done by myself jointly with others, I have made clear exactly what I have contributed myself.

Signed:
Kim Lucas
on the 7th day of June 2023 at the University of the Witwatersrand, Johannesburg.

Abstract

Threshold Functions of Colorings of Random Graphs

By Elizabeth Jonck, Kim Lucas, and Ronald Maartens.

Let $F_{n,p}$ be the set of all graphs with n vertices, m edges and with probability $p = p(n)$ of an edge occurring independently. Each graph $G \in F_{n,p}$ has probability $P[G] = p^m(1-p)^{\binom{n}{2}-m}$ of occurring. This is called the Binomial Random Graph Model, denoted $\mathbb{G}(n,p)$.

Let G now be a connected graph. A *rainbow colored graph* is when every two vertices of $V(G)$ are connected by a path where each edge has a unique color.

If Q is an increasing property, then a function $t = t(n)$ is called the threshold function for Q if (i) $p << t$ so that $\lim_{n\to\infty}(P[G \text{ has property } Q]) = 0$, and (ii) $p >> t$ so that $\lim_{n\to\infty}(P[G \text{ has property } Q]) = 1$.

A function $f(n)$ is called the sharp threshold function for the property Q if there exists constants C and c such that G satisfies Q almost surely for $p \geq Cf(n)$, and G almost surely does not satisfy Q for $p \leq cf(n)$.

In this dissertation, we investigate threshold functions and sharp threshold functions of random graphs to be rainbow colored.

Acknowledgements

I would like to thank my supervisors Dr. Ronald Maartens and Prof. Elizabeth Jonck, for all the effort, teachings and guidance they have given me during my Honour's degree as well as my Master's dissertation. I would also like to thank my grandfather, Eric Lucas, for paving the way through my studies and always encouraging me to continue in the pursuit of knowledge.

Contents

List of Figures

Chapter 1

Introduction

The objective of studying random graphs is to determine if a particular graph property occurs or not in a graph. Threshold functions and sharp threshold functions are used to determine when the graph property in question occurs or not.

In Chapter 1, introductory definitions and concepts that are needed to understand random graphs are provided. Then, in Chapter 2 and Chapter 3, we investigate random graphs and the rainbow coloring of random graphs, respectively. Chapter 4 and Chapter 5 are an investigation of threshold functions and sharp threshold functions, respectively, providing techniques to determine these functions. Six proofs are given as examples, three of which are our own. Chapter 6 concludes this investigation and suggests possible future work.

Note that this dissertation is in colour.

1.1 Graph Theory

Our investigation begins with a few definitions and terminologies from Graph Theory. We refer the reader to [13] for any undefined terminologies or concepts used here regarding Graph Theory.

A *graph*, denoted G, graphically is a diagram consisting of points with line segments or curves between these points as illustrated in Figure 1.1. These points are called *vertices* (single *vertex*), for example point a, and the line segments or curves are called *edges*, for example, the edge between vertices a and b denoted by ab.

Adjacent vertices are two vertices that have an edge in common. For example, vertices d and e are adjacent in Figure 1.1. These vertices are then *neighbors* of each other. The set of all vertices that are adjacent to a particular vertex is called a *neighborhood*. For example, the neighborhood of vertex d is $\{a, c, e, f\}$ and is denoted $N(d)$. For a graph G, the finite non-empty set of all vertices in G is called the *vertex set* of G, denoted $V(G)$, and the set of all edges (possibly an empty set) in G is called the *edge set* of G, denoted $E(G)$.

The *size* and *order* of a graph is the cardinality of the edge set and vertex set, respectively, denoted $|E(G)|$ and $|V(G)|$. In Figure 1.1, $|E(G)| = 8$ and $|V(G)| = 8$.

An *isolated vertex* is a vertex that has no edges connecting it to any other vertex as seen in Figure 1.1 by vertex h. We define an *empty graph* as a graph with vertices but no edges. This is sometimes referred to as the *trivial graph*.

An $a - g$ *path* is a sequence of vertices and edges, starting at vertex a and ending at vertex g, without repeating a vertex. Visualize a piece of chain where each link is an edge and the point where each link connects to another link is a vertex. If the chain is closed, then a *cycle* is formed, that is, a cycle is a path that begins and ends with the same vertex with no vertices repeated along the way. A path is denoted by P_n where n is the order of the path. The number of vertices constituting the cycle is called the *order* of a cycle, or otherwise known as the *length* of a cycle. We shall denote a cycle of order n by C_n where $n \geq 3$ by the definition of a cycle. In Figure 1.1, the cycle a, b, c, d, a is of order four.

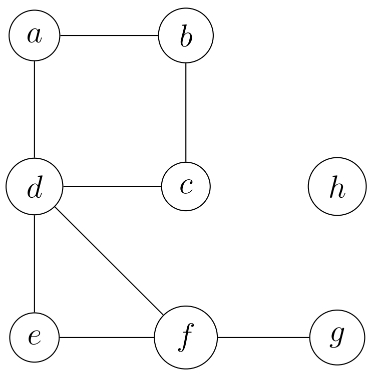

FIGURE 1.1: Simple graph G.

If every two vertices are connected by a path in a graph, then the graph is said to be *connected*. Otherwise, the graph is said to be *disconnected*. Figure 1.2 depicts a connected graph while Figure 1.1 depicts a disconnected graph. A connected graph where the removal of $k-1$ vertices, and all edges connected to these vertices, results in a graph that remains connected is called a k-*connected graph*. That is, at least k vertices need to be removed for the graph to become disconnected. In Figure 1.2, the graph is 1-connected as the removal of one of the vertices in the set $\{c, d, e, f\}$ cause the graph to be disconnected.

Let G and H be two graphs. If $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, then H is a *subgraph* of G, denoted $H \subseteq G$. If $H \subseteq G$ and there does not exist a connected graph F such that $H \subset F \subset G$, then H is said to be a *component* of G. As an example, the graph in Figure 1.2 is a component of the graph in Figure 1.1, only if the edge df is added.

A *tree* is a connected graph with no cycles, as in Figure 1.2. If all the components of a graph are trees, then the graph is referred to as a *forest*. If a connected graph has only one cycle, then the graph is called *unicyclic*.

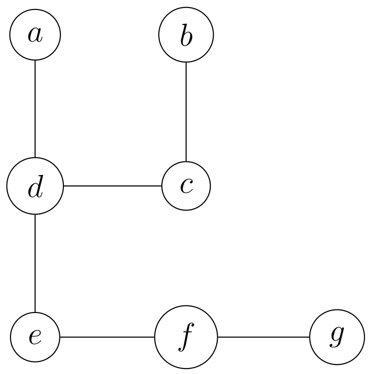

FIGURE 1.2: Subgraph of G.

Let $A \subseteq V(G)$. If every two vertices in A have no edge between them, then A is said to be an *independent set*. For example, in Figure 1.3, the set $\{a, d, e\}$ form an independent set. The *independent number of* G, denoted $\alpha(G)$, is the largest cardinality of all the independent sets of G.

A *bipartite graph* is a graph whose vertices are divided into two disjoint and independent sets such that every edge connects a vertex from one set to a vertex in the other. These individual sets are called *partite sets*. In Figure 1.3, the sets $\{a, d, e\}$ and $\{b, c\}$ are partite sets, respectively. A *complete bipartite graph* is one where every vertex in the one set is connected to every vertex in the other, denoted K_{n_1,n_2} where n_1 and n_2 are the number of vertices in each set.

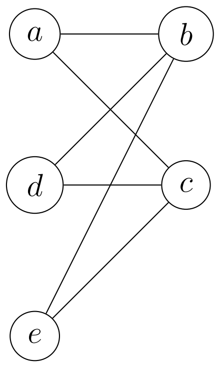

FIGURE 1.3: Bipartite graph $K_{3,2}$

A graph where every vertex has an edge from itself to every other vertex in the graph is called a *complete graph*, denoted K_n, where n is the order of the graph. Figure 1.4 illustrates the graph K_4.

If the vertex set of a graph is partitioned into $k > 0$ partite sets such that every vertex is adjacent to all other vertices in the graph except for the vertices in the same partite set, then that graph is called the *complete k-partite graph*. This graph is denoted by $K_{n_1,n_2,\cdots,n_k}$ where n_i is the order of the i^{th} partite set for $i = 1, 2, \cdots, k$.

If a graph G has two paths that have no vertices in common except for the first vertex and the last vertex, then the two paths are said to be *internally*

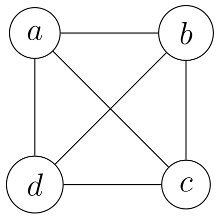

FIGURE 1.4: Complete graph K_4.

disjoint paths. In Figure 1.5, we have that the two paths a, b, d, f and a, c, e, f are two internally disjoint paths.

A *bridge* is an edge in a connected graph such that if the edge is removed, then the graph becomes disconnected. A *bridge-less graph* is a graph that contains no bridges. In Figure 1.5 the edge fg is a bridge, while ij is not.

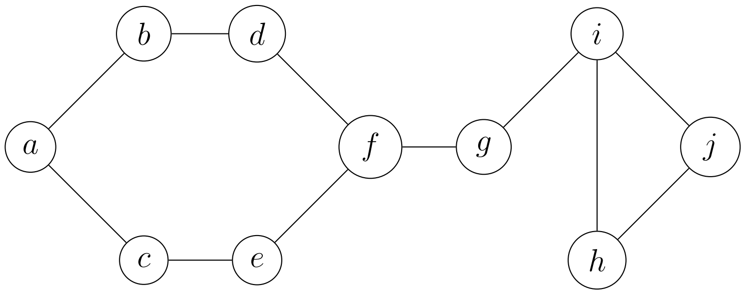

FIGURE 1.5: Internally disjoint paths.

A *geodesic* is the shortest path between two vertices, while the *distance* between two vertices is the length of the geodesic. The *diameter* of a graph G, denoted $diam(G)$, is the length of the largest geodesic in G. The graph in Figure 1.6 has diameter three which is given by the path d, a, b, c.

The *degree* of a vertex v, denoted $deg(v)$, is the number of edges incident to it. In Figure 1.6 we have $deg(a) = 4$ and $deg(f) = 1$. The *minimum degree* of a graph G, denoted $\delta(G)$, is the smallest degree of all the vertices in the graph. The *eccentricity* of a vertex is the length of the longest path between the vertex and all other vertices. The *radius* of a graph G, denoted $r(G)$, is the smallest eccentricity in the graph. In Figure 1.6 we have $r(G) = 2$ given by the path d, a, f.

The graph G is a *regular graph* if all vertices of G have the same number of neighbors. If all the vertices of G have degree k, then we say that G is a *k-regular graph*.

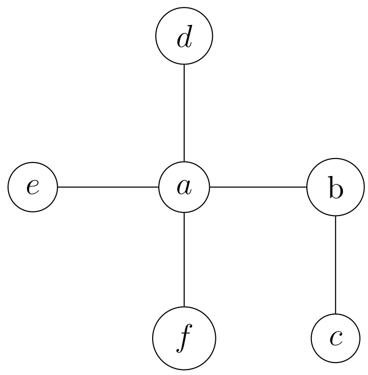

FIGURE 1.6: A graph.

Graph coloring is a way of coloring the vertices, edges, or faces (area bounded by edges) of a graph according to a list of requirements; called *vertex coloring*, *edge coloring*, and *face coloring*, respectively.

The *chromatic number* of a graph G, denoted $\chi(G)$, is the minimum number of *colors* required for the vertices of G to be colored in such a way that no two adjacent vertices have the same color.

The *compliment* of a graph G, denoted $\bar{G}$, has the same vertex set as G, however, the edges between the vertices in $\bar{G}$ are only drawn if they do not exist between the vertices in G. That is, $|V(G)| = |V(\bar{G})| = n$ and $E(G) \cup E(\bar{G}) = E(K_n)$ where $E(G) \cap E(\bar{G}) = \emptyset$.

Let G and H be two graphs. For $G \subseteq H$, a graph property Q is called *monotonically increasing* if G has property Q, then H also has property Q. Likewise, for $H \subseteq G$, a graph property Q is called *monotonically decreasing* if G has property Q, then H also has property Q. A graph has a *monotone property* Q if it is either monotonically increasing or monotonically decreasing. A graph has a *non-monotone property* Q if it is non-monotone.

Unless explicitly stated otherwise, all graphs

- have finite order,
- contain no loops (an edge that connects a vertex to itself),
- have no directed edges,
- have at most one edge between every two vertices (not a multigraph), and
- the vertices are labelled from the set $\{a, b, c, \cdots, z\}$. When the context is clear, we omit G in the notation. For example, we use V instead of $V(G)$.

1.2 Probability Theory

For any undefined terms and concepts in Probability Theory the reader is referred to [42].

A *discrete random variable* X is a measure of quantity, related to a random counting problem, and is a countable number. The *expected value* of X, denoted $E[X]$, is the average of the independent outcomes of $X = x_1 + x_2 + \ldots$. The *variance* of X, denoted $Var[X]$, is the average of the deviation from the mean squared. The variance is expressed mathematically by $Var[X] = E[X^2] - (E[X])^2$.

A *probability space* consists of three components which is denoted by (Ω, F, P). The *sample space*, denoted Ω, contains all possible outcomes (sometimes called *events*). The *event space*, denoted F, is the set of all events that occur where $F \subseteq \Omega$. The *probability function*, denoted P, is the function $P : F \to [0, 1]$ where $[0, 1]$ is the closed interval from zero to one. Note that $P[\Omega] = 1$.

The probability of an event occurring is denoted by $P[E]$. The probability of a graph G having a property Q is denoted by $P[G \in Q]$. Note that this is equivalent to counting the number of graphs with property Q out of all possible graphs in the sample space. The expected value of X is given by

$$E[X] = \sum_x xP[X = x]$$

where $P[X = x]$ depends on the graph model in use. We discuss this further in Chapter 4.

Bernoulli trials is a repeated experiment where the outcome of each experiment is only "True" or "False". Let p be the probability of a True outcome and q be the probability of a False outcome. Then, $p + q = 1$, so that $p = 1 - q$, and $q = 1 - p$. The *Boolean variable* is defined as a variable with two possible values, for example, "True" and "False".

We make use of the following well-known inequalities later on.

Theorem 1.1. ***Markov's Inequality*** *[42]*

Let X be a non-negative random variable. If $a > 0$, then

$$P[X \geq a] \leq \frac{E[X]}{a}.$$

Theorem 1.2. ***Chebyshev's Inequality*** *[42]*

Let X be a non-negative random variable. If $c > 0$, then

$$P[|X - E[X]| \geq c] \leq \frac{Var[x]}{c^2}.$$

Theorem 1.3. ***Bernoulli's Inequality*** *[42]*

Let $0 \leq x \leq 1$, and r be a positive integer. Then,

$$(1 - x)^r \geq 1 - xr.$$

1.3 Asymptotics and Combinatorics

For any undefined terms and concepts in Asymptotic Theory, the reader is referred to [41].

If an event happens with probability tending to one, then that event happens *almost surely*, abbreviated a.s. Also, if the event happens with probability tending to zero, then that event happens almost surely not. The concept of almost surely is also sometimes called *asymptotically almost surely* (a.a.s), or with high probability (whp).

Let $f(n)$ and $g(n)$ be two functions which are not equal to each other. Define the *small o notation* such that $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. Further, define the *big O notation* such that $f(n) = O(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) < \infty$ (otherwise known as an upper bound). Lastly, the *omega notation* is given by $f(n) = \Omega(g(n))$, if $\lim_{n\to\infty} f(n)/g(n) > 0$ (otherwise known as a lower bound).

The functions $f(n)$ and $g(n)$ are called *asymptotically equivalent*, abbreviated a.e, if $\lim_{n\to\infty} f(n)/g(n) \to 1$, denoted $f(n) \sim g(n)$. If there exists constants $x > 0$ and $y > 0$ such that $xg(n) < f(n) < yg(n)$, then we write $f(n) \approx g(n)$.

Theorem 1.4. *[41]*

The Gamma function, denoted $\Gamma(n)$*, is defined in the positive complex plane by*

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt,$$

where

$$\Gamma(n-1) = n!.$$

Many asymptotic approximations of the Gamma function exist, such as that of the famous mathematician Ramanujan. In this dissertation, the *Sterling's approximation for the Gamma function* is used as it is one of the sharpest approximations.

Theorem 1.5. ***Sterling's approximation*** *[41]*

The factorial function is asymptotically equivalent to

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

1.4 Calculus

For any undefined terms and concepts in Calculus, the reader is referred to [41]. In this dissertation, all functions containing "ln" are with base e, known as the *natural logarithm*, unless explicitly stated otherwise.

For the next theorem, indeterminate refers to an unknown form, for example, 0/0.

Theorem 1.6. ***L'Hopital's Rule*** *[23]*

Let $f(x)$ and $g(x)$ be two differentiable functions except possibly at a point c. If the limit of $f(x)/g(x)$ is indeterminate and $g'(x) \neq 0$, then

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}.$$

Theorem 1.7. *[41]*

Let $f(x)$ be infinitely differentiable and let a be a real or complex number. Then, the Taylor Series of $f(x)$ at the point a is given by

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!}.$$

In the case where $a = 0$, the Taylor Series is then called a Maclaurin Series. The Maclaurin Series expansion of the natural logarithm is given by

$$ln(1-x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}.$$

The next Theorem is also known as the *Sandwich Theorem.*

Theorem 1.8. ***The Squeeze Theorem*** *[41]*

Let $f(x)$, $g(x)$, and $h(x)$ be functions. If $g(x) \leq f(x) \leq h(x)$ and $\lim_{x\to\infty} g(x) = L = \lim_{x\to\infty} h(x)$, then

$$\lim_{x\to\infty} f(x) = L.$$

In the next chapter, we introduce random graphs.

Chapter 2

Random Graphs

The story of random graphs begins as mentioned below with Erdös and Rényi in 1959 with their seminal paper titled *On Random Graphs I* [39]. The main purpose of Random Graph Theory is to transform problems in Graph Theory to problems in Statistics, determine results, and then convert the results back to results in Graph Theory. Our aim is to determine the probability that a random graph has a specific graph property, for example, having diameter of length two. We investigate asymptotic distributions when the number of vertices in a graph tends to infinity. In this chapter we define a random graph for various random graph models, as well as investigate the asymptotic equivalence of random graphs.

2.1 What is a Random Graph?

Informally, *random graph models* are defined as a graph obtained through a random process.

Let $F_{n,M}$ be the set of all graphs that contain n labelled vertices and M edges with $0 \leq M \leq \binom{n}{2}$. Each graph $G \in F_{n,M}$ has the equal probability

$$P[G] = \frac{1}{\binom{\binom{n}{2}}{M}}$$

of occurring, where P is a probability distribution function on $F_{n,M}$. We refer to this model as the *Uniform model* or the *Erdös-Rényi model* which was first defined in [39]. A graph G which is from the Erdös-Rényi model is denoted by $G \in \mathbb{G}(n, M)$.

Around the same time, another random graph model was established by Gilbert in [19]. Let $F_{n,p}$ be the set of all graphs with n labelled vertices, m edges, and a probability $p = p(n)$ of an edge occurring independently where $0 \leq p \leq 1$. Each graph $G \in F_{n,p}$ has probability

$$P[G] = p^m(1-p)^{\binom{n}{2}-m}$$

of occurring where m is the number of edges occurring in G. This model is known as the *Binomial Model* or sometimes the *Bernoulli Model*. A graph G which is part of the Binomial graph model is denoted by $G \in \mathbb{G}(n, p)$.

As M is an integer and $0 \leq p \leq 1$, there is no ambiguity in using the notation $\mathbb{G}(n, M)$ and $\mathbb{G}(n, p)$ except for $\mathbb{G}(n, 1)$. In this case we either have a graph with n vertices and only one edge, or the complete graph of order n. Should the need arise to use $\mathbb{G}(n, 1)$, we clearly indicate the meaning there of.

In the Binomial model, each edge has probability $p = p(n)$ of occurring in the graph independently of each other, while in the Uniform model, each edge occurs uniformly from the set of all possible edges. Note that the Binomial model has two parameters, being the number of vertices and the probability of an edge occurring. This implies that the graphs in the set $F_{n,p}$ have different number of edges and as such the calculations in the Binomial model tend to be simpler as it is independent of the number of edges.

Random graphs are thought of as random processes. The process begins with an empty graph G with n vertices. Iteratively, an edge with probability p is placed between two randomly chosen vertices of G, ending after $\binom{n}{2}$ iterations. The aim is to investigate and study the properties present in these graphs.

The Uniform Model and the Binomial Model have been studied extensively over the years. For an in depth understanding of random graphs, the reader is referred to the series of papers published by Erdös and Rényi between 1959 and 1968; see [35], [39], [36], [37], [38], and [40].

2.2 Other Random Graph Models

The Uniform model and Binomial model are not the only random graph models being studied. Another more popular random graph model is the random tree graph model. There are two types of trees involved here which are called *labelled trees* and *recursive trees*, respectively. An in depth analysis of these two random tree models is available in [1].

Let F'_n be the set of all labelled trees with n vertices. There are n^{n-2} such trees. A *random labelled tree* is a tree chosen at random from F'_n where each tree $G \in F'_n$ has Uniform probability of being selected.

Let $G \in F''_n$ be an empty graph with n vertices labelled from the set $\{1, 2, \ldots, n\}$. The process of adding edges to F''_n begins with vertex 1, called the *root* of G. Each vertex $2, 3, \ldots, n$ in order, randomly places edges between itself and any labelled vertices that came before it. This is called a *recursive tree.*

A *random recursive tree* is a tree chosen at random from the set of all $(n-1)!$ recursive trees with n vertices and Uniform distribution.

These two are not the only random tree models. Nowadays, random forests has become a popular area of study due to its applications in machine learning classification problems.

We now define the k-out random graph model. More information of the k-out random graph model as well as the underlying digraphs that the model is built on is found in [44].

Let D_n be a graph with n labelled vertices from the set $\{1, 2, \ldots, n\}$ and let $k \in \{1, 2, \ldots, n-1\}$. Each vertex $x \in V(D_n)$ randomly connects with k vertices creating k number of arcs between vertex x and vertices $y \in V(D_n) - \{x\}$. The set of arcs $xy \in E(D_n)$ has a Uniform probability of appearing and there are a maximum of $\binom{n-1}{k}$ arcs. This is known as the *k-out random graph model.*

In the special case where $k = 1$, the $k-$out random graph model is called a *random mapping.*

Other random graph models include the random dot product graph model where the probability of an edge occurring between two vertices is a dot product of the two input variables being vectors [2]. The latent position model has probabilities assigned to both the edges and vertices of a graph [26]. There are many other random graph models and more are still to be discovered in the future.

2.3 Asymptotic Equivalence of Models

In this dissertation, we consider only the Binomial model. Occasionally the Uniform model is mentioned as the two models, while different, are closely related. Two conditions for when the two models are asymptotically equivalent are as follows.

Theorem 2.1. *[1]*

The probability of choosing a random graph $G \in \mathbb{G}(n,p)$ with m edges is the same probability of choosing a random graph $G' \in \mathbb{G}(n,M)$.

Proof

Let G be a random graph under the Binomial model with m edges. Note that the graphs in the Binomial model have different sizes as the model has no restriction on the number of edges. Therefore, graphs with m edges are part of the set of all possible graphs under the Uniform random graph model as well. The probability of choosing a graph G with m edges is

$$\frac{P[G \text{ has } m \text{ edges}]}{P[\text{Expected number of graphs with } m \text{ edges}]} \\ = \frac{p^m(1-p)^{\binom{n}{2}-m}}{\binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}} \\ = \binom{\binom{n}{2}}{m}^{-1}.$$

By the Uniform model definition, the result follows. □

The following theorem is a trivial condition that we state without proof.

Theorem 2.2. *[1]*

Let $G_1 \in \mathbb{G}(n,p)$ and $G_2 \in \mathbb{G}(n,M)$ for n large. Then the expected number of edges in G_1 is equivalent to the number of edges in G_2 when $m \sim \binom{n}{2}p$.

In [3], Bollobás proved the relation between the Uniform random graph model and the Binomial random graph model in detail. Essentially, the Uniform model converts to the Binomial model, however, the converse is not always true. An exact condition for the Binomial model to be equivalent to the Uniform model is found in [1]. Knowing how these two models are related helps perform calculations in one model and, where applicable, convert results to results in the other model. This is specifically useful in cases where the Uniform model is more complicated to compute as the Binomial model is usually easier to work with as there is no dependency on the number of edges in the graph.

2.4 Conclusion

In Chapter 2, we define random graphs for various random graph models. In this dissertation we only consider the Binomial model, denoted $G \in \mathbb{G}(n,p)$, which is defined as follows.

Let $F_{n,p}$ be the set of all graphs with n labelled vertices and a probability $p = p(n)$ of an edge occurring independently where $0 \leq p \leq 1$. Each graph $G \in F_{n,p}$ has probability

$$P[G] = p^m(1-p)^{\binom{n}{2}-m}$$

of occurring where m is the number of edges occurring in G.

The Uniform model and the Binomial model are asymptotically equivalent if the conditions of Theorem 2.1 and Theorem 2.2 are met, respectively.

In the next chapter we deviate from random graphs to investigate a graph property called rainbow coloring.

Chapter 3

Rainbow Coloring

In this chapter, we study the rainbow coloring of graphs. The concept of rainbow coloring of graphs was introduced in 2008 with paper [14] written by Chartrand, Johns, Mckoen, and Zhang. Since then, many new results have been published. In Chapter 3, some important aspects of the rainbow connection number of graphs are looked at before delving into finding sharp thresholds for rainbow connected graphs in Chapter 5. As rainbow coloring is a fairly new topic in Graph Theory, we dedicate a chapter to its definition and state a few main results.

3.1 What is the Rainbow Coloring of Graphs?

Let G be a connected graph and $x, y \in V(G)$ with $x \neq y$. Define a *rainbow path* as a path in G where no two edges on the path between x and y are colored the same, as illustrated in Figure 3.1. A connected graph G is *rainbow connected* if any two distinct vertices in G are connected by a rainbow path.

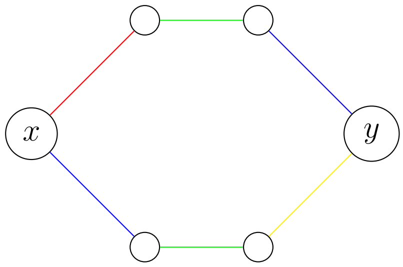

Figure 3.1: Rainbow path.

From these two definitions it is deduced quite easily that if G is rainbow connected, then G is connected. The converse is true. If each edge of a connected graph G is colored with a unique color, then G is rainbow connected. However, we are interested in the minimum number of colors required for G to be rainbow connected, which is called the *rainbow connection number*, denoted $rc(G)$.

Chartrand et al [14] expanded on the definition of rainbow connectedness. If the shortest path between any two vertices in G is also a rainbow path (called a *rainbow geodesic*), then we say that G is *strongly rainbow connected*. The minimum number of colors required for G to be strongly rainbow connected is called the *strong rainbow connection number*, denoted $src(G)$.

From the definitions above, we trivially have that

$$diam(G) \leq rc(G) \leq src(G) \leq |E(G)|.$$

Furthermore, Theorem 3.1 below relates the strong rainbow connection number and the rainbow connection number to specific graph properties.

Theorem 3.1. *[14]*

Let G be a connected non-trivial graph with m edges. Then,

a) $src(G) = 1$ if and only if G is complete,

b) $rc(G) = 2$ if and only if $src(G) = 2$, and

c) $rc(G) = m$ if and only if G is a tree.

In paper [15], published a year after [14], the authors defined the rainbow coloring problem for k-connected graphs by integrating the well known Menger's Theorem. In 1927, Menger [31] proved what is known today as Menger's Theorem stated below.

Theorem 3.2. ***Menger's Theorem*** *[31]*

Let $1 \leq k \leq z$, and G a z-connected graph. Further, let $x, y \in V(G)$, with $x \neq y$, then there are k internally disjoint paths between x and y.

We now apply Menger's Theorem to the concept of the rainbow connection number of a graph. Let G be a connected graph. As mentioned before, the definition of a graph being rainbow connected requires only one rainbow path between every two vertices in the graph. Hence, if there are at least k internally disjoint rainbow paths between every two vertices in G, then G is *rainbow k-connected*. Further, the minimum number of colors needed for G to be rainbow k-connected is then called the *rainbow k-connection number*, denoted $rc_k(G)$. Trivially,

$$rc_1(G) = rc(G).$$

In this dissertation we study rainbow colorings in a random graph setting instead of simple graphs. There are many more concepts that the interested reader is referred to such as the total rainbow k-connection number. The *total rainbow k-connection number* is the minimum number of colors needed for every two vertices in a connected graph G to be connected by k internally disjoint paths where both the edges and vertices are colored by distinct colors. This concept was introduced by Liu, Mestre, and Sousa in 2014 [33], which was an extension of the work done by Uchizawa, Aoki, Ito, Suzuki, and Zhou in 2012. They studied the not k-connected version [45].

The rainbow vertex-connection number, denoted $rvc(G)$, is defined analogously to the rainbow connection number. This was first introduced by Krivelevich and Yuster in 2009 [34]. Similar to the edge colored version, strong rainbow connection and k-connectedness are defined for vertex colored graphs.

3.2 A Few Rainbow Coloring Results

The following results for rainbow connected graphs speak to the origin of the concept rainbow connection and how it evolved into what it is today.

Rainbow connection graphs were studied for the first time in 2008 when Caro, Lev, Roditty, Tuza, and Yuster [9] established two fundamental results.

Theorem 3.3. *[9]*

For a connected graph G with n vertices and $\delta(G) \geq 3$, we have

$$rc(G) < \frac{5n}{6}.$$

Theorem 3.4. *[9]*

If G is a connected graph with n vertices and $\delta = \delta(G)$, then

$$rc(G) \leq min\left\{\frac{nln\delta}{\delta}(1 + o_{\delta}(1)), \frac{4nln\delta}{\delta}\right\}.$$

These bounds are not optimal but have been proven to be very difficult to improve and we consider it for future work. In the same paper, the authors gave the following conjecture.

Conjecture 3.1. *[9]*

If G is a connected graph with n vertices and $\delta(G) \geq 3$, then,

$$rc(G) < \frac{3n}{4}.$$

A year later, this conjecture by Cairo et al was proved by Schiermeyer [28]. In addition to the proof of the conjecture, the author demonstrated that the bound in Theorem 3.4 is sharper than the bound in Theorem 3.3. Another year later, Krivelevich and Yuster in [34] proved the following theorem.

Theorem 3.5. *[34]*

For a connected graph G, with n vertices and minimum degree δ,

$$rc(G) < 20n/\delta.$$

With Schiermeyer proving the conjecture, and taking into account the new result above from Krivelevich and Yuster, Schiermeyer conjectured the following from his observations.

Conjecture 3.2. *[34]*

Let G be a connected graph with n vertices. If $rc(G) \leq c_k n$ for all graphs G with $\delta(G) \geq k$, and c_k is a minimum constant with $0 < c_k \leq 1$, then $c_k = 3/(k+1)$ for all $k \geq 2$.

This conjecture was not proved as is, but Chandron et al in 2011 [11] found a tighter bound than $rc(G) < 20n/\delta$ (see Theorem 3.5). Chandron et al used Conjecture 3.2 as inspiration for the following two results.

Theorem 3.6. *[11]*

Let G be a connected graph with n vertices and let $\delta = \delta(G)$. Then,

$$rc(G) \leq \frac{3n}{\delta + 1} + 3.$$

Theorem 3.7. *[11]*

Let G be a connected graph with n vertices and let $\delta = \delta(G) \geq 2$. There are infinitely many graphs G that satisfy

$$rc(G) \geq \frac{3(n-2)}{\delta+1} + 1.$$

The next natural question here is to ask about graphs with higher connectivity. In this regard, we mention the following two results.

Theorem 3.8. *[32]*

For all $k \geq 1$, if G is a k-connected graph with n vertices, then

$$rc(G) \leq \frac{3n}{(k+1)} + 3.$$

Theorem 3.9. *[20]*

If G is a 2-connected graph with $n \geq 4$ vertices, then

$$rc(G) \leq \lceil \frac{n}{2} \rceil.$$

Another interesting approach is to relate the rainbow connection number to other graph properties. So far the diameter of a graph has been a fundamental motive in the study of rainbow connected graphs. There is not much published on the relation between the rainbow connection number and other graph properties. In this regard, we state the next three theorems.

Theorem 3.10. *[7]*

Let G be a connected bridge-less graph and let r be the radius of G. Then,

$$rc(G) \leq r(r+2).$$

Theorem 3.11. *[29]*

Let G be a connected graph with $\delta(G) \geq 2$. Then,

$$rc(G) \leq 2\alpha(G) - 1$$

where $\alpha(G)$ is the independence number of G.

Theorem 3.12. *[27]*

Let G be a connected graph. Then,

$$rc(G) \leq 2\chi(\bar{G}) - 1$$

where $\chi(G)$ is the chromatic number of G and $\bar{G}$ is the complement of G.

For more results in the space of rainbow coloring with and without the consideration of random graphs, the reader is referred to the survey by Li and Sun [46] where many more results are given.

3.3 Conclusion

Let G be a connected graph and $x, y \in V(G)$ with $x \neq y$. Define a rainbow path as a path in G where no two edges on the path between x and y are colored the same. Further, a connected graph G is rainbow connected if any two distinct vertices $x, y \in V(G)$ are connected by a rainbow path.

If there are at least k internally disjoint rainbow paths between any two vertices in G, then G is rainbow k-connected. Further, the minimum number of colors needed for G to be rainbow k-connected is then called the rainbow k-connection number, denoted $rc_k(G)$.

We observe that the diameter and the minimum degree of a graph is fundamental in studying rainbow graphs and the rainbow connection number.

We trivially have that

$$diam(G) \leq rc(G) \leq src(G) \leq |E(G)|.$$

A few results on rainbow coloring are given in this chapter that the interested reader can pursue. In the next chapter, we return to random graphs and analyse threshold functions.

Chapter 4

Threshold Functions

4.1 Introduction

Informally, a threshold function is a point where an event almost surely begins to occur, or begins to not occur.

Let G be a graph with an infinite number of isolated vertices. Adding one edge at a time, connecting any two vertices at random to form a random graph. During this process, various graph properties occur, for example, G has tree components. Threshold functions show us when these graph properties occur or not occur for a given probability of an edge occurring in the Binomial model (or any other random graph model).

In [5], Bollobas showed that every monotone graph property has a threshold function. The same is not necessarily true for non-monotone graph properties. This is something that is yet to be established and we consider this for future work.

In 1960, Erdös and Rényi [40] defined threshold functions, sharp threshold functions, and regular threshold functions. Threshold functions are considered here in Chapter 4, and Sharp threshold functions are considered in Chapter 5. Regular threshold functions are considered for future work.

4.1.1 Threshold Functions for the Binomial Model

Threshold functions for the Binomial model is defined as follows.

1. First, let $G \in \mathbb{G}(n, p)$ with $|V(G)| = n$. Second, let Q be a graph property that is increasing. Last, let $p = p(n)$ be the probability of an edge occurring in G with $p \in [0, 1]$. A function $t = t(n)$ is called the *threshold function* for Q

   - if $p << t$, then $\lim_{n\to\infty} P[G \text{ has property } Q] = 0$, and
   - if $p >> t$, then $\lim_{n\to\infty} P[G \text{ has property } Q] = 1$.

   Here, the symbol $<<$ is defined to be conceptually much less than, and $>>$ is conceptually much greater than. In other words, if $p << t$, then $p/t \to 0$, and if $p >> t$, then $p/t \to \infty$.

2. Using the same conditions as in (1), let Q now be a decreasing graph property. Then, $t = t(n)$ is the threshold function for Q

   - if $p << t$, then $\lim_{n\to\infty} P[G \text{ has property } Q] = 1$, and
   - if $p >> t$ then, $\lim_{n\to\infty} P[G \text{ has property } Q] = 0$.

For $P[G$ has property $Q] = 1$, G is said to asymptotically almost surely have Q. Further, for $P[G$ has property $Q] = 0$, G is said to asymptotically almost surely not have Q.

In the definition above, the threshold function is defined in terms of increasing and decreasing graph properties; implying the definition is only valid for graphs with monotone graph properties. In this dissertation we only consider graphs with monotone graph properties.

4.1.2 Threshold Functions for the Uniform Model

The following definition of the threshold function considers a graph under the Uniform model.

1. First, let $G \in \mathbb{G}(n, M)$ with $|V(G)| = n$ and M as the number of edges in G. Second, let Q be a graph property that is increasing. Last, let $r = r(n)$ be a function denoting the number of edges in G. Then, a function $t = t(n)$ with $r \in [0, \binom{n}{2}]$ is the threshold function for Q

   - if $r << t$, then $\lim_{n\to\infty} P[G$ has propertry $Q] = 0$, and

   - if $r >> t$, then $\lim_{n\to\infty} P[G$ has propertry $Q] = 1$.

2. Using the same conditions as in (1), to define the threshold function t for Q a decreasing graph property, we interchange 0 and 1 with the limits in (1).

4.2 Method to Determine Threshold Functions

In literature, there are two methods used to find the threshold function of a monotone graph, other than using analysis. These two methods are known as the Grading method, and the First and Second Moments method.

The grading method was established by Bollobas in [6]. The Grading method is applied in [6] to a specific threshold function on small subgraphs and is more complicated to use than the First and Second Moments method. The First and Second Moments method is used for any monotone graph property and, as such, is the method we are using in this dissertation.

4.2.1 The First and Second Moments Method

The First and Second Moments method was first mentioned by Chebyshev in 1887. The general understanding behind the First and Second Moments method is to find upper bounds and lower bounds of the probability of where a random variable might exist.

Theorem 4.1. ***First Moment*** *[43]*

Let X be a non-negative integer-valued random variable. If

$$\lim_{n\to\infty} E[X] = 0$$

then,

$$\lim_{n\to\infty} P[X = 0] = 1.$$

Proof

Let X be a non-negative integer-valued random variable and $a = 1$. Then, from Markov's inequality (Theorem 1.1),

$$P[X \geq 1] \leq E[X].$$

If

$$\lim_{n\to\infty} E[X] = 0,$$

then

$$\lim_{n\to\infty} P[X \geq 1] \leq \lim_{n\to\infty} E[X] = 0.$$

For $i = 1, 2, \ldots,$ we have $P = p_i \in [0, 1]$ and the total probability is given by

$$P = \sum_i p_i = 1.$$

Hence,

$$P[X \geq 1] + P[X = 0] = 1$$

and so, if $P[X \geq 1] \to 0$, then $P[X = 0] \to 1$. □

Theorem 4.2. ***Second Moment*** *[43]*

Let X be a non-negative integer-valued random variable. For $E[X] \neq 0$, if

$$\lim_{n\to\infty} E[X^2] = (E[X])^2,$$

then

$$\lim_{n\to\infty} P[X \geq 1] = 1.$$

Proof

Let X be a non-negative integer-valued random variable. From Chebyshev's inequality (Theorem 1.2), with $c = E[X] = \mu$, it follows that

$$P[|X-\mu| \geq E[X]] \leq \frac{Var[X]}{(E[X])^2}.$$

Considering the left-hand side of the inequality, we have

$$\begin{aligned}
P[|X-\mu| \geq E[X] = \mu] &= P[X-\mu \leq -\mu] \text{ or } P[X-\mu \geq \mu] \\
&= P[X \leq 0] \text{ or } P[X \geq 2\mu] \\
&\geq P[X \geq 0] \\
&= P[X = 0].
\end{aligned}$$

Hence, from Chebyshev's inequality,

$$\begin{aligned}
P[X=0] &\leq P[|X-\mu| \geq E[X]] \\
&\leq \frac{Var[X]}{(E[X])^2} \\
&= \frac{E[X^2]-(E[X])^2}{(E[X])^2} \\
&= \frac{E[X^2]}{(E[X])^2} - 1.
\end{aligned}$$

From the theorem's assumptions,

$$\lim_{n\to\infty} E[X^2] = (E[X])^2$$

and so, $E[X^2] \to (E[X])^2$ and $P[X = 0] = 0$. Consequently, if $P[X = 0] \to 0$, then $P[X \geq 1] \to 1$. □

4.2.2 Steps to Determine the Threshold Function

For a more detailed guide to the First and Second Moments method, the interested reader is referred to [43] by Maartens. The following step-by-step approach for the First and Second Moment method is given in [43] and is used to determine the threshold function for any random graph with a monotone graph property.

Step 1: Define a random variable X.

Let X be a non-negative integer-valued random variable. Define $X = X(G)$ in one of two ways.

1. Let X count the number of times the occurrence of property Q occurs in a random graph G.

2. Let X be a Boolean variable where it takes on the value 1 or 0 depending on whether property Q occurs in G or not.

Step 2: Determine $\boldsymbol{E[X]}$.

Using the model under consideration, determine the expected value of X, that is, $E[X]$. In particular, the Binomial model has

$$E[X] = \sum_{G} X(G)P[G].$$

Step 3: Guess a potential threshold function.

Guess a potential threshold function $t = t(n)$ which satisfies both of the following two requirements:

1. if $t \to 0$, then $\lim_{n\to\infty} E[X] \to 0$, and

2. if $t \to \infty$, then $\lim_{n\to\infty} E[X] \to \infty$.

Step 4: Apply the First Moment.

Apply the First Moment (Theorem 4.1) by substituting the potential threshold function from Step 3 into the expected function in Step 2. If the potential threshold function works, then, for the given graph property Q, we have that $P[X = 0] \to 1$, implying that Q almost surely does not exist in G.

Step 5: Apply the Second Moment.

Apply the Second Moment (Theorem 4.2) by substituting the potential threshold function from Step 3 into the expected function in Step 2. If the potential threshold function works, then for the given graph property Q, we have that $P[X \geq 1] \to 1$, implying that Q almost surely exists in G.

Step 4 and Step 5 verifies that the potential threshold function found in Step 3 is in fact the threshold function for the graph property Q in G. If Step 4 or Step 5 does not work, then the potential threshold function is not one.

Note that the process given above does not produce a unique threshold function. In the following section we address the question of whether threshold functions are unique.

4.3 Uniqueness of the Threshold Function

Note that threshold functions for both the Binomial model and Uniform model are not unique, that is, $t = t(n)$ has many possibilities that satisfy the definition of a threshold function. However, in some cases, threshold functions are unique in the sense that they differ by the product of a constant. It is customary in the field of Random Graph Theory to talk about "the" threshold function instead of "a" threshold function.

Theorem 4.3. *[43]*

If $t = t(n)$ is the threshold function for a graph property in some random graph model with $t \approx s = s(n)$, then s is also the threshold function for the same graph property.

Proof

Let $G \in \mathbb{G}(n,p)$ with $p = p(n)$ and $p \in [0,1]$. Further, let Q be a monotone graph property of G and let $t(n)$ be the threshold function for Q. Assume that $t(n) \approx s(n)$ with $t(n) \neq s(n)$, for some function $s(n)$. Then, there exists two positive constants, say c and C, such that

$$cs(n) \leq t(n) \leq Cs(n).$$

Since $t(n)$ is the threshold function, it follows that $p >> t(n)$ or $p << t(n)$ by definition.

Case 1: $p >> t(n)$.

If $p >> t(n)$, then

$$\lim_{n\to\infty} \frac{p}{t(n)} \to \infty.$$

From $cs(n) \leq t(n)$ it follows that

$$\lim_{n\to\infty} \frac{p}{cs(n)} \geq \lim_{n\to\infty} \frac{p}{t(n)} \to \infty$$

and so, $p >> s(n)$.

Case 2: $p << t(n)$.

Similarly, if $p << t(n)$, then

$$\lim_{n\to\infty} \frac{p}{t(n)} \to 0.$$

From $t(n) \leq Cs(n)$ it follows that

$$\lim_{n\to\infty} \frac{p}{Cs(n)} \leq \lim_{n\to\infty} \frac{p}{t(n)} \to 0$$

and so, $p << s(n)$.

Combining the two cases, it follows that $s(n) \approx t(n)$. Hence, $s(n)$ is also the threshold function for Q, implying that the threshold function is not unique.

A similar proof follows for the threshold function under the Uniform model.

□

4.4 Illustration of the First and Second Moments Method

The following theorem was proved by Bollobas [4], however, we provide our own proof here. While the method given in Section 4.2.2 seems easy to follow, note that some proofs quickly become computational intense. Finding the relevant expected function is also quite challenging in some instances and is usually a hurdle for many of the more complex graph properties.

Before proceeding to the main result, we prove the following lemma.

Lemma 4.4.

For positive constant c,

$$\lim_{n\to\infty}\left(1 - c\sqrt{\frac{ln(n)}{n}}\right) = 1.$$

Proof

From basic limit properties we have that

$$\lim_{n\to\infty}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)=\left(1-c\sqrt{\lim_{n\to\infty}\frac{ln(n)}{n}}\right).$$

By applying L'Hopital's rule (Theorem 1.6), we have that

$$\lim_{n\to\infty}\frac{ln(n)}{n}=\lim_{n\to\infty}\frac{1}{n}=0.$$

The result then follows. □

Theorem 4.5. *[4]*

The threshold function for a graph $G \in \mathbb{G}(n,p)$ with diameter two is

$$t=\sqrt{\frac{ln(n)}{n}}. \tag{4.1}$$

Proof

Let $G \in \mathbb{G}(n,p)$ with graph property Q as $diam(G) = 2$.

Step 1: Define a random variable X.

Define a bad pair of vertices in G as a pair (i,j) with distinct $i,j \in V(G)$ such that there is no other vertex in $V(G)$ that is adjacent to both i and j. This implies that $d(i,j) > 2$. Thus, i and j are not adjacent.

Let X count the number of bad pairs of vertices in $V(G)$ where

$$X = \sum_{i} \sum_{j, i \neq j} X_{ij}.$$

Therefore, let

$$X_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ a bad pair,} \\ 0 & \text{otherwise.} \end{cases}$$

Hence, if $X = 1$, then there is a bad pair in G and the diameter is larger than two. However, if $X = 0$, then the diameter is two.

Step 2: Determine $E[X]$.

The expected value of X is given by

$$E[X] = \sum_{G} X(G)P[G].$$

Let p be the probability that a vertex in G is adjacent to another vertex in G. The probability that i and j are not adjacent is then $(1 - p)$. Further, as $|V(G)| = n$, there are $\binom{n}{2}$ ways to choose the two distinct vertices i and j. The probability that another vertex, say k, of G is adjacent to both i and j is then p^2. Hence, the probability that vertex k is not adjacent to both i and j is $(1 - p^2)$. Lastly, if $X = 0$, then we know the diameter is two. As we have $(n - 2)$ vertices left in G after choosing vertices i and j, to ensure there are no bad pairs, we count $(n - 2)$ instances of $(1 - p^2)$. Thus, in total we have $(1 - p^2)^{n-2}$.

As the remaining part of G is unknown, and we need to consider all possible events that occur, we have that the probability of the remainder of G is one. Substituting into the equation for $E[X]$, we have

$$\begin{aligned} E[X] &= \binom{n}{2}(1-p)(1-p^2)^{n-2} \cdot P[\text{remainder of } G] \\ &= \binom{n}{2}(1-p)(1-p^2)^{n-2}. \end{aligned}$$

Step 3: Find a potential threshold function.

Let $t = c\sqrt{ln(n)/n}$ be a potential threshold function. Then,

$$\begin{aligned} E[X] &= \binom{n}{2}\Big(1-p\Big)\Big(1-p^2\Big)^{n-2} \\ &= \frac{n(n-1)}{2}\Big(1-p\Big)\Big(1-p^2\Big)^{n}\Big(1-p^2\Big)^{-2}. \end{aligned} \tag{4.2}$$

Consider

$$\frac{(n-1)(1-p^2)^{-2}(1-p)}{n}$$

where $a_n = (n-1)(1-p^2)^{-2}(1-p)$, $b_n = n$ and $p = c\sqrt{ln(n)/n}$. We show that $a_n/b_n \to 1$ as $n \to \infty$, and then $a_n \sim b_n$. Note that

$$\begin{aligned} \frac{(n-1)(1-p^2)^{-2}(1-p)}{n} &= \left(1-\frac{1}{n}\right)\left(1-\frac{c^2 ln(n)}{n}\right)^{-2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right) \\ &= \left(1-\frac{1}{n}\right)\left(\frac{n-c^2 ln(n)}{n}\right)^{-2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right). \end{aligned}$$

Therefore,

$$\begin{aligned}
&\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(\frac{n-c^2ln(n)}{n}\right)^{-2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\\
&=(1)\cdot\lim_{n\to\infty}\left(\frac{n-c^2ln(n)}{n}\right)^{-2}\cdot(1) \qquad (4.3)\\
&=\lim_{n\to\infty}\left(\frac{n}{n-c^2ln(n)}\right)^{2}\\
&=\lim_{n\to\infty}\left(\frac{1}{1-\frac{c^2}{n}}\right)\left(\frac{1}{1-\frac{c^2}{n}}\right) \text{ by applying L'Hopital's rule}\\
&=1
\end{aligned}$$

where in Line (4.3), Lemma 4.4 is applied.

Consider now Equation 4.2. If $p = c\sqrt{ln(n)/n}$ and we substitute n in place of $(n-1)(1-p^2)^{-2}(1-p)$, then

$$\begin{aligned}
E[X] &= \frac{n(n-1)}{2}\left(1-p\right)\left(1-p^2\right)^{n}\left(1-p^2\right)^{-2}\\
&= \frac{n(n-1)}{2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-\frac{c^2ln(n)}{n}\right)^{n-2}\\
&\sim \frac{n^2}{2}\left(1-\frac{c^2ln(n)}{n}\right)^{n}. \qquad (4.4)
\end{aligned}$$

Let $a = c^2ln(n)/n$. Note that

$$(1-a)^n = exp(ln(1-a)^n) = exp(nln(1-a)),$$

and applying the Maclaurin Series expansion of the natural logarithmic function, we obtain

$$\begin{aligned} exp(nln(1-a)) &= exp(n(-a - \frac{a^2}{2} - \frac{a^3}{3} - \dots)) \\ &= exp(-n(a + \frac{a^2}{2} + \frac{a^3}{3} + \dots)) \\ &= exp(-na - n(\frac{a^2}{2} + \frac{a^3}{3} + \dots)) \\ &= exp(-na)exp(-na^2(\frac{1}{2} + \frac{a}{3} + \dots)). \end{aligned}$$

If $na^2 \to 0$, then $exp(-na^2(\frac{1}{2} + \frac{a}{3} + \dots)) \to 1$. To demonstrate this, consider

$$\begin{aligned} \lim_{n\to\infty} na^2 &= \lim_{n\to\infty} n\frac{c^2ln(n)}{n}\frac{c^2ln(n)}{n} \\ &= \lim_{n\to\infty} c^4\frac{(ln(n))^2}{n} \\ &= 2c^4 \lim_{n\to\infty} \frac{ln(n)}{n} \qquad (4.5) \\ &= 2c^4 \lim_{n\to\infty} \frac{1}{n} \\ &= 0 \end{aligned}$$

where L'Hoptial's Rule is applied to Equation 4.5. Then,

$$exp(nln(1-a)) = exp(-na)exp(-na^2(\frac{1}{2} + \frac{a}{3} + \dots)) = exp(-na),$$

and hence,

$$(1-a)^n = exp(-na). \qquad (4.6)$$

We now have a way of simplifying functions of the form $(1-a)^n$ to $exp(-na)$

where $a = c^2 ln(n)/n$. Substituting Equation 4.6 into Equation 4.4 with $p = c\sqrt{ln(n)/n}$ yields

$$\begin{aligned} E[X] &= \frac{n^2}{2}\left(1 - \frac{c^2 ln(n)}{n}\right)^n \\ &\sim \frac{n^2}{2} e^{-c^2 ln(n)} \\ &= \frac{n^2}{2} (e^{-ln(n)})^{c^2} \\ &= \frac{n^2}{2} n^{-c^2} \\ &= \frac{n^{2-c^2}}{2}. \end{aligned}$$

Hence, $E[X] \sim (n^2/2)(n^{-c^2})$. For $c > \sqrt{2}$, note that $E[X] \to 0$ as $n \to \infty$. If $c < \sqrt{2}$, then $E[X] \to \infty$ as $n \to \infty$. That is,

$$E[X] = \begin{cases} 0 & \text{if } c > \sqrt{2}, \\ \infty & \text{if } c < \sqrt{2}. \end{cases}$$

Step 4: First Moment.

If $c > \sqrt{2}$, then $E[X] \to 0$ as $n \to \infty$, and so $P[X = 0] \to 1$. By the First Moment (Theorem 4.1), G asymptotically almost surely has no bad pair, implying that the diameter is two.

Step 5: Second Moment.

We already have that $c < \sqrt{2}$ implies $E[X] \to \infty$ as $n \to \infty$. Note that showing that G has at least one bad pair implies that the diameter is larger than two. Now, the expected function of X^2 is given by

$$E[X^2] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right]. \tag{4.7}$$

This is essentially a sum over all possible values for X^2. Here, the aim is to have two distinct pairs of vertices (i, j) and (k, l) distance two from each other. In Step 3 above, only one pair of vertices is considered. To find all possible combinations where two pairs of vertices are distance two apart, consider the three cases below.

Case 1: $a, b \in N(i) \cap N(j)$.

Consider the pairs (i, j) and (j, i). Let $a, b \in V(G) - \{i, j\}$. In this case, vertices i and j share common neighbours a and b. Since each pair of vertices has all the same common neighbours, the graph diameter is two. This is represented in Figure 4.1.

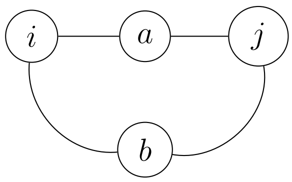

FIGURE 4.1: Case One.

There are $\binom{n}{2}$ ways of choosing i and j in $V(G)$. The probability that vertex a is adjacent to i or j is p. The probability that a is adjacent to both i and j is p^2. Similarly, the probability that vertex b is adjacent to both i an j is p^2. The probability that either a or b is adjacent to both i and j is $p^2 + p^2 = 2p^2$. Then, the probability that both a and b are not adjacent to both i and j is $(1-2p^2)$. There are $(n-3)$ such instances of $(1-2p^2)$, producing $(1-2p^2)^{n-3}$.

All together, the expected function yields

$$E\left[\sum_{i<j} X_i X_j\right] = \binom{n}{2}(1-p)(1-2p^2)^{n-3}. \tag{4.8}$$

Since $p = c\sqrt{ln(n)/n}$, Equation 4.8 becomes,

$$\binom{n}{2}(1-p)(1-2p^2)^{n-3} = \frac{n(n-1)}{2}\left(1 - c\sqrt{\frac{ln(n)}{n}}\right)\left(1 - \frac{2c^2 ln(n)}{n}\right)^{n-3}.$$

As before,

$$\begin{aligned} &\frac{(n-1)\left(1 - c\sqrt{\frac{ln(n)}{n}}\right)\left(1 - \frac{2c^2 ln(n)}{n}\right)^{-3}}{n} \\ &= \left(1 - \frac{1}{n}\right)\left(\frac{n - 2c^2 ln(n)}{n}\right)^{-3}\left(1 - c\sqrt{\frac{ln(n)}{n}}\right) \end{aligned} \tag{4.9}$$

so that

$$\begin{aligned} &\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)\left(\frac{n - 2c^2 ln(n)}{n}\right)^{-3}\left(1 - c\sqrt{\frac{ln(n)}{n}}\right) \\ &= (1)\cdot\lim_{n\to\infty}\left(\frac{n - 2c^2 ln(n)}{n}\right)^{-3}\cdot\left(1 - c\sqrt{\lim_{n\to\infty}\frac{ln(n)}{n}}\right) \qquad (4.10) \\ &= \lim_{n\to\infty}\left(\frac{n - 2c^2 ln(n)}{n}\right)^{-3}\cdot(1) \\ &= \lim_{n\to\infty}\left(1 - \frac{2c^2}{n}\right)^{-3} \qquad (4.11) \\ &= 1 \end{aligned}$$

by applying L'Hopital's rule (Theorem 1.6) to Equation 4.11, and Lemma 4.4 to Equation 4.10. Hence, the numerator of Equation 4.9 is replaced by n

in $E[X]$ and applying Equation 4.6 yields,

$$\begin{aligned} E\Big[\sum_{i<j} X_i X_j\Big] &= \frac{n(n-1)}{2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-\frac{2c^2ln(n)}{n}\right)^{n-3} \\ &\sim \frac{n^2}{2}\left(1-\frac{2c^2ln(n)}{n}\right)^{n} \\ &\sim \frac{n^2}{2}e^{-2c^2ln(n)}. \end{aligned} \tag{4.12}$$

In conclusion,

$$E\Big[\sum_{i<j} x_i x_j\Big] = \binom{n}{2}(1-p)(1-2p^2)^{n-3} \sim \frac{n^2}{2}e^{-2c^2ln(n)}. \tag{4.13}$$

Case 2: $c \in N(i) \cap N(j)$ and $d \in N(k) \cap N(l)$.

Consider the distinct pairs (i,j) and (k,l), where i and j and also k and l are distance two from each other. Let $c, d \in V(G) - \{i,j,k,l\}$. This is illustrated in Figure 4.2.

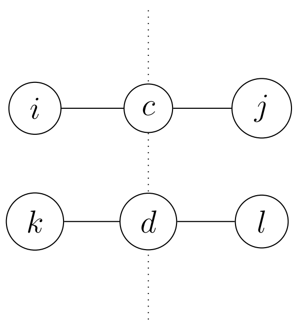

FIGURE 4.2: Case Two.

For Case 2 to occur, choose two vertices of the n vertices in $V(G)$ for the first pair (i,j), and then, of the remaining $n-2$ vertices, choose two again for the other pair (k,l). Thus, we have

$$\binom{n}{2}\binom{n-2}{2} = \binom{n}{4}$$

possible ways of choosing the two distinct pairs. The probability that i and j are adjacent is p, and so, the probability that i and j are not adjacent is $(1-p)$. Likewise for vertices k and l. The probability that $ij \notin E(G)$ and $kl \notin E(G)$ is $(1-p)^2$. The probability of a vertex adjacent to both i and j is p^2, and so, the probability of a vertex not adjacent to both i and j is $(1-p^2)$. Likewise for vertices k and l. Since four of the n vertices are being used, and we have two of these good pairs, there are in total $2(n-4)$ such instances of $(1-p^2)$. Thus,

$$E\left[\sum_{i<j,k<l} X_{ij}X_{kl}\right] = \binom{n}{2}\binom{n-2}{2}(1-p)^2(1-p^2)^{2(n-4)}. \tag{4.14}$$

Now, evaluate Equation 4.14 with

$$\begin{aligned}&\binom{n}{2}\binom{n-2}{2}(1-p)^2(1-p^2)^{2(n-4)}\\ &= \frac{n(n-1)}{2}\frac{(n-2)(n-3)}{2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{2n-8}.\end{aligned} \tag{4.15}$$

As

$$\frac{(n-1)(n-2)(n-3)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{-8}}{n^3} \tag{4.16}$$

$$= (1-\frac{1}{n})(1-\frac{2}{n})(1-\frac{3}{n})\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{-8},$$

we have,

$$\begin{aligned}
&\lim_{n\to\infty}(1-\frac{1}{n})(1-\frac{2}{n})(1-\frac{3}{n})\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{-8}\\
&= (1)(1)(1)(1)\lim_{n\to\infty}\left(1-c^2\frac{ln(n)}{n}\right)^{-8} &(4.17)\\
&= \left(1-c^2\lim_{n\to\infty}\frac{ln(n)}{n}\right)^{-8}\\
&= \left(1-\lim_{n\to\infty}\frac{1}{n}\right)^{-8} &(4.18)\\
&= 1
\end{aligned}$$

where Lemma 4.4 is applied to obtain Equation 4.17, and L'Hopital's Rule (Theorem 1.6), is applied to obtain Equation 4.18. Hence, the numerator of Equation 4.16 is replaced by n^3 in Equation 4.15, yielding

$$\begin{aligned}
E\Bigg[\sum_{i<j,k<l} X_{ij}X_{kl}\Bigg] &= \frac{n(n-1)}{2}\frac{(n-2)(n-3)}{2}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{2n-8}\\
&\sim \frac{n^4}{4}\left(1-c^2\frac{ln(n)}{n}\right)^{2n-3}\\
&\sim \frac{n^4}{4}e^{-2c^2ln(n)}.
\end{aligned}$$

Hence,

$$E\left[\sum_{i<j,k<l} X_{ij}X_{kl}\right] = \binom{n}{2}\binom{n-2}{2}(1-p)^2(1-p^2)^{2(n-4)} \sim \frac{n^4}{4}e^{-2c^2ln(n)}.$$

Case 3: $i, k \in N(e)$ and $i, j \in N(f)$.

Lastly, consider the case where the pairs (i, j) and (i, k) both are distance two apart. However, vertex i has multiple vertices distance two around it. Let $e, f \in V(G) - \{i, j, k\}$. This is represented in Figure 4.3.

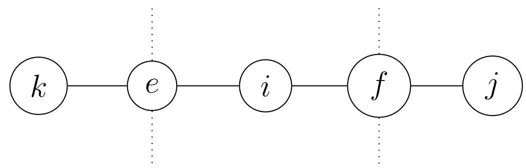

FIGURE 4.3: Case Three.

In this case, select the two pairs of vertices (i, j) and (i, k) with distance two apart, where i is common in both pairs. This requires three vertices and so choose three vertices from the available $\binom{n}{3}$ combinations. The probability that i and j are adjacent is p, and so, the probability that i and j are not adjacent is $(1-p)$. Likewise for vertices i and k. The probability $ij \notin E(G)$ and the probability that $ik \notin E(G)$ is $(1-p)^2$. The probability of a vertex adjacent to both i and j is p^2, and so, the probability of a vertex not adjacent to both i and j is $(1-p^2)$. Likewise for vertices k and i. Since three of the n vertices are being used and two of these are good pairs, there are in total $2(n-3)$ such instances where $ij, ik \notin E(G)$. Thus,

$$E\left[\sum_{i<j,i<k} X_{ij}X_{ik}\right] = \binom{n}{3}(1-p)^2(1-p^2)^{2(n-3)}. \tag{4.19}$$

Now, evaluate Equation 4.19 with

$$\begin{aligned}&\binom{n}{3}(1-p)^2(1-p^2)^{2(n-3)}\\ &= \frac{n(n-1)(n-2)}{6}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{2n-6}.\end{aligned} \tag{4.20}$$

Since

$$\begin{aligned}&\frac{(n-1)(n-2)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-c^2\frac{ln(n)}{n}\right)^{-6}}{n^2}\\ &=\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-c^2\frac{ln(n)}{n}\right)^{-6},\end{aligned} \tag{4.21}$$

we have

$$\begin{aligned}&\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-c\sqrt{\frac{ln(n)}{n}}\right)\left(1-c^2\frac{ln(n)}{n}\right)^{-6}\\ &= (1)(1)(1)(1)\lim_{n\to\infty}\left(1-c^2\frac{ln(n)}{n}\right)^{-6} &(4.22)\\ &= \left(1-c^2\lim_{n\to\infty}\frac{ln(n)}{n}\right)^{-6}\\ &= \left(1-c^2\lim_{n\to\infty}\frac{1}{n}\right)^{-6} &(4.23)\\ &= 1\end{aligned}$$

where Equation 4.22 is obtained by applying Lemma 4.4, and Equation 4.23

by applying L'Hopital's Rule (Theorem 1.6). The numerator of Equation 4.21 is replaced by n^2 in Equation 4.20, yielding

$$\begin{aligned} E\left[\sum_{i<j,i<k} X_{ij}X_{ik}\right] &= \frac{n(n-1)(n-2)}{6}\left(1-c\sqrt{\frac{ln(n)}{n}}\right)^2\left(1-c^2\frac{ln(n)}{n}\right)^{2n-6} \\ &\sim \frac{n^3}{6}\left(1-c^2\frac{ln(n)}{n}\right)^n\left(1-c^2\frac{ln(n)}{n}\right)^n \\ &= \frac{n^3}{6}e^{-c^2ln(n)}e^{-c^2ln(n)} \\ &= \frac{n^3}{6}e^{-2c^2ln(n)}. \end{aligned}$$

Therefore,

$$E\left[\sum_{i<j,i<k} X_{ij}X_{ik}\right] = \binom{n}{3}(1-p)^2(1-p^2)^{2(n-3)} \sim \frac{n^3}{6}e^{-2c^2ln(n)}.$$

Note that $E[X^2]$, as defined in Equation 4.7, is a sum of Equation 4.8, Equation 4.14, and Equation 4.19:

$$E[X^2] = \left(\frac{n^2}{2}e^{-2c^2ln(n)}\right) + \left(\frac{n^3}{6}e^{-2c^2ln(n)}\right) + \left(\frac{n^4}{4}e^{-2c^2ln(n)}\right).$$

To conclude this proof, we compute $E[X^2]/(E[X])^2$. Thus,

$$
\begin{aligned}
\frac{E[X^2]}{(E[X])^2} &= \frac{\left(\frac{n^2}{2}e^{-2c^2 ln(n)}\right) + \left(\frac{n^3}{6}e^{-2c^2 ln(n)}\right) + \left(\frac{n^4}{4}e^{-2c^2 ln(n)}\right)}{\left(\frac{n^2}{2}e^{-c^2 ln(n)}\right)^2} \\
&= \frac{\left(\frac{n^2}{2}n^{-2c^2}\right) + \left(\frac{n^3}{6}n^{-2c^2}\right) + \left(\frac{n^4}{4}n^{-2c^2}\right)}{\left(\frac{n^4}{4}n^{-2c^2}\right)} \\
&= \left(\frac{n^2}{2n^{2c^2}} + \frac{n^3}{6n^{2c^2}} + \frac{n^4}{4n^{2c^2}}\right)\left(\frac{4n^{2c^2}}{n^4}\right) \\
&= \frac{2}{n^2} + \frac{2}{3n} + 1,
\end{aligned}
$$

and so,

$$\lim_{n\to\infty} \frac{E[X^2]}{(E[X])^2} = 1.$$

Hence, $E[X^2]/(E[X])^2 \to 1$ as $n \to \infty$.

By the Second Moment Method (Theorem 4.2), asymptotically almost surely at least one bad pair exists. Therefore, the threshold function for $diam(G) = 2$ is $t = \sqrt{2}\sqrt{ln(n)/n}$. □

In Theorem 4.6 below, we generalize the result in Theorem 4.5 to any diameter d. The following theorem was originally proved by Bollobas in [4], however, we provide our own proof here.

Note that if $d = 0$, then G is an empty graph, and if $d = 1$, then G is the complete graph.

Theorem 4.6. *[4]*

The threshold function for a graph $G \in \mathbb{G}(n,p)$ with $diam(G) = d \geq 2$ is

$$t^d = \frac{ln(\frac{n^2}{c})}{n^{d-1}}. \tag{4.24}$$

Proof

We proceed by strong mathematical induction on diameter d. Let $G \in \mathbb{G}(n,p)$ with $diamg(G) = d \geq 2$. For $d = 2$ in Equation 4.24,

$$t^2 = \frac{ln(n^2/c)}{n^{2-1}}$$

so that

$$t = \sqrt{\frac{ln(\frac{n^2}{c})}{n}} = \sqrt{\frac{2ln(\frac{n}{c})}{n}} = \sqrt{2}\sqrt{\frac{ln(\frac{n}{c})}{n}}.$$

Note that

$$\begin{aligned}
\lim_{n\to\infty} \left(\frac{ln(n) - ln(c)}{n}\right)^{\frac{1}{2}} \Big/ \left(\frac{ln(n)}{n}\right)^{\frac{1}{2}} &= \lim_{n\to\infty} \left(\left(\frac{ln(n) - ln(c)}{n}\right) \Big/ \left(\frac{ln(n)}{n}\right)\right)^{\frac{1}{2}} \\
&= \lim_{n\to\infty} \left(\frac{n(ln(n) - ln(c))}{nln(n)}\right)^{\frac{1}{2}} \\
&= \lim_{n\to\infty} \left(\frac{(ln(n) - ln(c))}{ln(n)}\right)^{\frac{1}{2}} \\
&= \lim_{n\to\infty} \left(\frac{ln(n)}{ln(n)} - \frac{ln(c)}{ln(n)}\right)^{\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned}\lim_{n\to\infty}\left(\frac{ln(n)}{ln(n)}-\frac{ln(c)}{ln(n)}\right)^{\frac{1}{2}} &= \lim_{n\to\infty}\left(1-\frac{ln(c)}{ln(n)}\right)^{\frac{1}{2}}\\ &= \left(1-\lim_{n\to\infty}\frac{ln(c)}{ln(n)}\right)^{\frac{1}{2}}\\ &= \sqrt{1}\\ &= 1.\end{aligned}$$

Therefore, by the definition of asymptotic equivalence,

$$t=\sqrt{\frac{ln(\frac{n^2}{c})}{n}}\sim\sqrt{2}\sqrt{\frac{ln(n)}{n}}.$$

This is the case we proved in Theorem 4.5 above. Thus, Equation 4.24 holds for $d = 2$. Now, assume that Equation 4.24 holds for any d where $2 \leq d \leq k$ and k is a positive integer. Consider, $d = k + 1$. Trivially we have $t^k = tt^{k-1}$ and $t^{k+1} = tt^k$. Therefore,

$$\begin{aligned}t^{k+1} &= \frac{t^k}{t^{k-1}}t^k\\ &= \left(\frac{ln(\frac{n^2}{c})}{n^{k-1}}\right)\left(\frac{ln(\frac{n^2}{c})}{n^{k-1}}\right)\Big/\left(\frac{ln(\frac{n^2}{c})}{n^{k-2}}\right)\\ &= \left(\frac{ln(\frac{n^2}{c})}{n^{k-1}}\right)\left(\frac{ln(\frac{n^2}{c})}{n^{k-1}}\right)\left(\frac{n}{n}\right)\Big/\left(\frac{ln(\frac{n^2}{c})}{n^{k-2}}\frac{n}{n}\right)\\ &= \left(\frac{nln(\frac{n^2}{c})}{n^{k-1}}\right)\left(\frac{ln(\frac{n^2}{c})}{n^{k}}\right)\Big/\left(\frac{nln(\frac{n^2}{c})}{n^{k-1}}\right)\\ &= \frac{ln(\frac{n^2}{c})}{n^k}\end{aligned}$$

and so, the case for $d = k + 1$ is true.

The result holds, implying that the threshold for $G \in \mathbb{G}(n,p)$ with $diam(G) = d$ is

$$t^d = \frac{ln(\frac{n^2}{c})}{n^{d-1}}.$$

□

4.5 Conclusion

The threshold functions for the Binomial model are defined as follows.

1. Let Q be a graph property that is increasing. Let $G \in \mathbb{G}(n,p)$ with $|V(G)| = n$. Further, let $p = p(n)$ be the probability of an edge occurring in G with $p \in [0,1]$. A function $t = t(n)$ is called the threshold function for Q

   - if $p << t$, then $\lim_{n\to\infty} P[G \text{ has property } Q] = 0$, and
   - if $p >> t$, then $\lim_{n\to\infty} P[G \text{ has property } Q] = 1$.

   Here, the symbol $<<$ is defined to be conceptually much less than, and the symbol $>>$ is conceptually much greater than. In other words, if $p << t$, then $p/t \to 0$, and if $p >> t$, then $p/t \to \infty$.

2. Using the same conditions as in (1), let Q now be a decreasing graph property. Then, $t = t(n)$ is the threshold function for Q

   - if $p << t$, then $\lim_{n\to\infty} P[G \text{ has property } Q] = 1$, and
   - if $p >> t$ then, $\lim_{n\to\infty} P[G \text{ has property } Q] = 0$.

For $P[G \text{ has property } Q] = 1$, G is said to asymptotically almost surely have Q. Further, for $P[G \text{ has property } Q] = 0$, G is said to asymptotically almost surely not have Q.

A fundamental result of threshold functions is that every monotone graph property has a threshold function as proved by Bollobas. The same is not necessarily true for non-monotone graph properties.

A method for determining the threshold function is investigated and proven using the First Moment Method (Theorem 4.1), and the Second Moment method (Theorem 4.2). The method is as follows.

Step 1: Define a non-negative integer-valued random variable X for the graph property in question.
Step 2: Determine $E[X]$ using the chosen random graph model.
Step 3: Guess a potential threshold function and apply it to $E[X]$.
Step 4: Apply the First Moment.
Step 5: Apply the Second Moment.

In addition, the uniqueness up to a constant product of the threshold function is proved in Theorem 4.3.

The First and Second Moments method is illustrated with the following two theorems.

Theorem 4.7.

The threshold function for a graph $G \in \mathbb{G}(n,p)$ with $diam(G) = 2$ is

$$t = \sqrt{\frac{ln(n)}{n}}. \tag{4.25}$$

Theorem 4.8.

The threshold function for a graph $G \in \mathbb{G}(n,p)$ with $diam(G) = d \geq 2$ is

$$t^d = \frac{ln(\frac{n^2}{c})}{n^{d-1}}. \tag{4.26}$$

In the next chapter we study a sharper version of the threshold function.

Chapter 5

Sharp Threshold Functions

5.1 Introduction

In this chapter we investigate sharp threshold functions which are a natural continuation of the threshold functions from Chapter 4. First, a few definitions are given.

Let Q be a graph property. Recall that a graph $G \in \mathbb{G}(n, p)$ satisfies Q almost surely if $P[G \in Q] \to 1$ as $n \to \infty$. Likewise, G almost surely does not satisfy Q if $P[G \notin Q] \to 0$ as $n \to \infty$.

For a random graph G, a function $f(n)$ is called the *sharp threshold function* for the graph property Q if there exists constants C and c, both positive, such that G satisfies Q almost surely for $p \geq Cf(n)$, and G almost surely does not satisfy Q for $p \leq cf(n)$.

In other words, the definition of a sharp threshold function for an increasing graph property is,

$$\lim_{n\to\infty} P[G \in Q] = \begin{cases} 1 & p \geq Cf(n), \\ 0 & p \leq cf(n). \end{cases}$$

For a decreasing graph property, the definition for an increasing graph property holds but with zero and one interchanged.

Let G be a random graph with $V(G) = N_1 \cup N_2$ such that $N_1 \cap N_2 = \emptyset$, $|N_1| = n_1$ and $|N_2| = n_2$. If each vertex in N_1 is joined with the vertices in N_2 with probability p at random, then G is called a *random bipartite graph*, denoted $G(n_1, n_2, p)$.

We now investigate the sharp threshold function of the rainbow connection number which was introduced earlier in Chapter 3 above. The results below are stated without proof. One of the fundamental results proved by Caro et al in [9] is stated first as an introduction.

Theorem 5.1. *[9]*

Let $G \in \mathbb{G}(n, p)$ where n is the number of vertices. Then,

$$f(n) = \sqrt{\frac{ln(n)}{n}}$$

is the sharp threshold function for the graph property $rc(G) \leq 2$.

Theorem 5.1 is an important sharp threshold function as the form "$\sqrt{ln(n)/n}$"

appears in many other sharp threshold functions when dealing with the rainbow connection number. Hence, when trying to find a sharp threshold function for a different graph property linked to the rainbow connection number, it is advised to consider variations of the form "$\sqrt{ln(n)/n}$" first.

Two years later, He and Liang [25] generalized Theorem 5.1 by considering $rc_k(G) \leq d$, which is the rainbow k-connected version of Theorem 5.1.

Theorem 5.2. *[25]*

Let $G \in \mathbb{G}(n,p)$ have $diam(G) = d \geq 2$, and let $k = k(n) \leq O(ln(n))$ be a positive integer. Then,

$$f(n) = \frac{(ln(n))^{\frac{1}{d}}}{n^{\frac{(d-1)}{d}}}$$

is the sharp threshold function for the graph property $rc_k(G) \leq d$.

Fujita, Liu, and Magnant considered other random graph models in [22] and expanded on the result of Theorem 5.2.

Theorem 5.3. *[22]*

Let $k \geq 1$ and let $G \in \mathbb{G}(n_1, n_2, p)$. Then,

$$f(n) = \sqrt{\frac{ln(n)}{n}}$$

is the sharp threshold function for the graph property $rc_k(G) \leq 3$.

Analogously we have the next Theorem for the Uniform Model.

Theorem 5.4. *[22]*

Let $G \in \mathbb{G}(n, M)$ and $k \geq 1$ be a positive integer. Then,

$$f(n) = \sqrt{n^3 ln(n)}$$

is the sharp threshold function for the graph property $rc_k(G) \leq 2$.

In [22], the authors further proposed the following open problem.

Open Problem 5.1. *[22]*

For $k \geq 1$ and $d \geq 2$, can the sharp threshold function be found for the graph property $rc_k(G) \leq d$ where G is a random graph in a model other than the Binomial model?

In 2020, this open problem was partialy solved by Chen, Li, and Lian in [16] where the authors solved the problem for random bipartite graphs.

Theorem 5.5. *[16]*

Let $G \in \mathbb{G}(n_1, n_2, p)$, $d \geq 2$ and $k = k(n) \leq O(ln(n))$. If d is odd, then

$$f(n) = \frac{(ln(mn))^{\frac{1}{d}}}{m^{\frac{(d-1)}{(2d)}} \cdot n^{\frac{(d-1)}{2d}}}$$

is the threshold function for the graph property $rc_k(G) \leq d+1$, where $pn \geq pm \geq (ln(n))^4$. If d is even, then

$$f(n) = \frac{(ln(n))^{\frac{1}{d}}}{(m^{\frac{1}{2}} \cdot n^{\frac{(d-2)}{2d}})}$$

is the sharp threshold function for the graph property $rc_k(G) \leq d+1$ *where there exists* $0 < \epsilon < 1$ *such that* $pn_1^{1-\epsilon} \geq pn_2^{1-\epsilon} \geq (ln(n))^4$.

Other random graph models are still to be considered for Open Problem 5.1.

The NP-hardness of $rc(G)$ where G is a random graph is shown in [10]. In [30], it is shown how Graph Theory problems become simpler when making use of random graphs instead of simple graphs. Thus, a motivation for using random graphs as opposed to simple graphs is to decrease the complexity and computational time of these problems.

Recall that all monotone graph properties have a threshold function [5] which was proved in 1987. Later, in 1996, it was proved by the authors in [17] that every monotone graph property has a sharp threshold function. It is still an open problem if this result holds for non-monotone properties.

The rainbow connection number is a monotone graph property. To see this, note that if edges are added to a graph, then the rainbow connection number does not change. This implies that there are numerous tools at our disposal when dealing with the rainbow connection number as it is monotone. Consequently, the sharp threshold function exists for monotone graph properties, such as the rainbow connection number. However, the strong rainbow connection number, which is so much more difficult to work with, is a non-monotone property. The strong rainbow connection number is left for future studies.

To find the sharp threshold function, it is natural to first find the threshold function as the sharp threshold function is a sharper bound of the threshold function. By first finding the threshold function, it is simpler to estimate what the sharp threshold function looks like. However, it is possible to find the sharp threshold function directly.

Earlier, in Theorem 4.5, the First and Second Moments method is used. This proof is long and grows in complexity with more challenging graph properties, such as rainbow coloring properties. When modifying the method of finding the sharp threshold function from the method of finding the threshold function, in Section 4.2.2, we note that this method becomes tedious. Therefore, we make use of the Chernoff Bounds to simplify the process.

5.2 The Chernoff Bounds

The Chernoff Bounds are first used in [24], however, the underlying mathematics is inspired by the works of Sergei Bernstein and apparently Herman Rubin. Today, many variations of the Chernoff Bounds exist depending on the application. In this dissertation, the multiplicative form of the Chernoff Bounds is used. Note that this is a generalization of the Binomial distribution.

Theorem 5.6. ***Chernoff Bounds - Lower Tail*** *[24]*

Let $X = \{x_1, x_2 \dots x_n\}$ *be an independent random variable. Further, let* $\delta \in (0, 1]$, $\mu = E[X] = \sum_{i=1}^{n} P_i$, *and* $P[x_i = 1] = P_i$. *Then,*

$$P[X < (1 - \delta)\mu] \leq e^{\frac{-\mu\delta^2}{2}}.$$

Proof

Consider

$$P[X < (1-\delta)\mu] = P[e^{-tX} > e^{-t(1-\delta)\mu}].$$

Let $t > 0$. From Markov's inequality in Theorem 1.1, we have

$$\begin{aligned} P[e^{-tX} > e^{-t(1-\delta)\mu}] &\leq \frac{E[e^{-tX}]}{exp(-t(1-\delta)\mu)} \\ &= \frac{E[exp(-tx_1 - tx_2 - \cdots - tx_n)]}{exp(-t(1-\delta)\mu)} \\ &= \frac{E[e^{-tx_1} \cdot e^{-tx_2} \cdot \cdots \cdot e^{-tx_n}]}{exp(-t(1-\delta)\mu)} \\ &= \frac{E[\prod_{i=1}^{n}(exp(-tx_i))]}{exp(-t(1-\delta)\mu)}. \end{aligned}$$

Since the $x_i's$ are independent, the product is moved outside the expected function to produce

$$\begin{aligned} P[X < (1-\delta)\mu] &< \frac{E[\prod_{i=1}^{n}(exp(-tx_i))]}{exp(-t(1-\delta)\mu)} \\ &= \frac{\prod_{i=1}^{n} E[(exp(-tx_i))]}{exp(-t(1-\delta)\mu)}. \end{aligned}$$

Here every x_i has a different probability, are independent, and is called a Poisson trial by definition, as opposed to a Bernoulli trial where every x_i has the same probability. Hence, a proof using Poisson trials also holds for the Bernoulli trials. Note that the sum of Bernoulli trials has a Binomial distribution, and the Binomial distribution becomes the Poisson distribution when $n \to \infty$ and $E[X]$ remains constant.

For Poisson trials $P[X_i = 1] = P_i$ and $P[X_i = 0] = (1 - P_i)$, we have,

$$\begin{aligned} P[X < (1-\delta)\mu] &< \frac{\prod_{i=1}^{n} E[(exp(-tx_i))]}{exp(-t(1-\delta)\mu)} \\ &= \frac{\prod_{i=1}^{n}(P_i e^{-t} + (1-P_i))}{exp(-t(1-\delta)\mu)} \\ &= \frac{\prod_{i=1}^{n}(1 - P_i(1-e^{-t}))}{exp(-t(1-\delta)\mu)}. \end{aligned}$$

Now, let $r \geq 1$ and $0 \leq x \leq 1$ where

$$r = 1 + 1 + \cdots + 1 \text{ (r times)} \tag{5.1}$$

$$\geq 1 + (1-x) + (1-x)^2 + \cdots + (1-x)^{r-1}, \tag{5.2}$$

which is a geometric series. Hence,

$$\begin{aligned} r &\geq \frac{1-(1-x)^r}{1-(1-x)} \\ xr &\geq 1 - (1-x)^r, \end{aligned}$$

and so, $(1-x)^r \geq 1 - xr$ which is an alternative form of the Bernoulli inequality in Theorem 1.3. By definition,

$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n,$$

and so,

$$e^{-x} = \lim_{n\to\infty}\left(1 - \frac{x}{n}\right)^n \geq 1 - x.$$

Thus,

$$\begin{aligned}
E[exp(-tX_i)] &= 1 - P_i(1 - e^{-t}) \\
&\leq exp(-P_i(1 - e^{-t})) \\
&= exp(P_i(e^{-t} - 1))
\end{aligned}$$

which is used to obtain

$$\begin{aligned}
P[X < (1-\delta)\mu] &< \frac{\prod_{i=1}^{n}(1 - P_i(1 - e^{-t}))}{exp(-t(1-\delta)\mu)} \\
&\leq \frac{\prod_{i=1}^{n} exp(P_i(e^{-t} - 1))}{exp(-t(1-\delta)\mu)} \\
&= \frac{exp\Big(\sum_{i=1}^{n} P_i(e^{-t} - 1)\Big)}{exp(-t(1-\delta)\mu)} \\
&= \frac{exp((e^{-t} - 1)\sum_{i=1}^{n} P_i)}{exp(-t(1-\delta)\mu)} \\
&= \frac{exp((e^{-t} - 1)\mu)}{exp(-t(1-\delta)\mu)} \qquad (5.3) \\
&= exp(\mu(e^{-t} - 1) - \mu(-t(1-\delta))) \\
&= exp(\mu(e^{-t} - 1 + t - t\delta)). \qquad (5.4)
\end{aligned}$$

Hence,

$$P[X < (1-\delta)\mu] < exp(\mu(e^{-t} - 1 + t - t\delta)).$$

Note that, in Equation 5.3, $\mu = \sum_{i=1}^{n} P_i = E[X]$. The proof still holds for $\mu \leq E[X]$.

We now determine t which gives the tightest bound. Consider the expression $(e^{-t} - 1 + t - t\delta)$ with

$$\frac{d}{dt}(e^{-t} - 1 + t - t\delta) = -e^{-t} + 1 - \delta. \tag{5.5}$$

Setting Equation 5.5 equal to zero, and solving for t, we have

$$\begin{aligned} -e^{-t} + 1 - \delta &= 0 \\ e^{-t} &= 1 - \delta \\ lne^{-t} &= ln(1-\delta) \\ -t &= ln(1-\delta) \\ t &= -ln(1-\delta) \\ t &= ln(\frac{1}{1-\delta}). \end{aligned} \tag{5.6}$$

Substitute Equation 5.6 back into Equation 5.4 to obtain

$$\begin{aligned} P[X < (1-\delta)\mu] &< exp\left(\mu\left(e^{-ln(\frac{1}{1-\delta})} - 1 + ln\left(\frac{1}{1-\delta}\right) - \delta ln\left(\frac{1}{1-\delta}\right)\right)\right) \\ &= exp(\mu((1-\delta) + ln\left(\frac{1}{1-\delta}\right)(1-\delta) - 1)) \\ &= exp(\mu(-\delta + ln\left(\frac{1}{1-\delta}\right)(1-\delta))) \\ &= exp(-\delta\mu + \mu ln\left(\frac{1}{1-\delta}\right)(1-\delta)) \\ &= e^{-\delta\mu} e^{\mu^{ln(\frac{1}{1-\delta})(1-\delta)}} \\ &= \left(e^{-\delta}\right)^{\mu} \left(e^{ln(\frac{1}{1-\delta})(1-\delta)}\right)^{\mu} \end{aligned}$$

$$\left(e^{-\delta}\right)^{\mu}\left(e^{ln(\frac{1}{1-\delta})^{(1-\delta)}}\right)^{\mu} = \left(e^{-\delta}\right)^{\mu}\left(\frac{1}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$
$$= \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}. \tag{5.7}$$

Here, Equation 5.7 is the sharpest form of the Chernoff Bounds. However, the bound in this form is usually difficult to compute. Therefore, an alternative but weaker form, of Equation 5.7 is determined below. We proceed by simplifying the term $(1-\delta)^{(1-\delta)}$. Taking the natural logarithm of this term, we have

$$(1-\delta)ln(1-\delta).$$

The Taylor Series expansion of $ln(x)$ (Theorem 1.7) is represented by

$$\sum_{n=1}^{\infty}\frac{(-1)^{n-1}(x-1)^n}{n}.$$

Hence,

$$\begin{aligned}
(1-\delta)ln(1-\delta) &= (1-\delta)\sum_{n=1}^{\infty}(-1)^{n-1}\frac{(1-\delta-1)^n}{n} \\
&= (1-\delta)[-\delta - \frac{-\delta^2}{2} - \frac{-\delta^3}{3} - \dots] \\
&= (-\delta - \frac{-\delta^2}{2} - \frac{-\delta^3}{3} - \frac{-\delta^4}{4} - \dots) + (\delta^2 + \frac{\delta^3}{2} + \frac{\delta^4}{3} + \frac{\delta^5}{4} + \dots) \\
&= -\delta + \left(\frac{-\delta^2}{2} + \delta^2\right) + \left(\frac{-\delta^3}{3} + \frac{\delta^3}{2}\right) + \left(\frac{-\delta^4}{4} + \frac{\delta^4}{3}\right) + \dots \\
&= \delta + \frac{\delta^2}{2} + (\text{positive terms}) \\
&> -\delta + \frac{\delta^2}{2}.
\end{aligned}$$

Raising both sides of this last inequality by e yields

$$(1-\delta)^{1-\delta} > exp(-\delta + \frac{\delta^2}{2}).$$

Proceeding from Equation 5.7 yields

$$\begin{aligned} P[X < (1-\delta)\mu] &< \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \\ &< \left(\frac{e^{-\delta}}{exp(-\delta + \frac{\delta^2}{2})}\right)^{\mu} \\ &= [exp(-\delta + \delta - \frac{\delta^2}{2})]^{\mu} \\ &= e^{\frac{-\mu\delta^2}{2}}. \end{aligned}$$

Hence, the proof is concluded and

$$P[X < (1-\delta)\mu] \leq e^{\frac{-\mu\delta^2}{2}}.$$

□

Lastly, we state the upper tail of the Chernoff Bound. The proof of the lower tail of the Chernoff Bounds is similar to the proof of the upper tail, and thus, the proof is omitted here.

Theorem 5.7. ***Chernoff Bounds - Upper Tail*** *[24]*

Let $X = \{x_1, x_2, \ldots, x_n\}$ *be an independent random variable and let* $\delta \in (0, 1]$, $\mu = E[X] = \sum_{i=1}^{n} P_i$, *and* $P[x_i = 1] = P_i$. *Then,*

$$P[X > (1 + \delta)\mu] < 2^{-\delta\mu}.$$

5.3 Method to Determine the Sharp Threshold Function

Below is a four step process to find the sharp threshold function for random graphs with a monotone graph property. This methodology is as a result of studying numerous journal article proofs and combining those learnings into one streamlined approach to find sharp threshold functions. As an application, four sharp threshold functions are found in this chapter to illustrate this methodology.

Let $f(n)$ be a proposed sharp threshold function for a monotone graph property Q. Further, let C and c be two real constants.

The method to determine the sharp threshold function is illustrated below.

Step 1: Define X.

Define X to be a non-negative integer-valued random variable. The random variable X is defined in one of two ways. Let X count the number of times property Q occurs in a random graph G. Alternatively, let X be a Boolean variable where X takes the value one or zero depending on whether property Q occurs in G or not.

Step 2: Determine $E[X]$.

Here we determine $E[X]$, making use of the model under consideration. Note that $E[X]$ needs to be defined in such a way that taking the probability of $E[X]$ occurring or not occurring results in the graph property Q being true or not.

Step 3: Show $P[G \in Q] \to 0$.

In this step apply the potential sharp threshold function to $E[X]$ and show that $P[G \in Q] \to 0$ as $n \to \infty$ for a sufficiently small (large) constant C.

Step 4: Show $P[G \in Q] \to 1$

Apply the potential sharp threshold function again to show $P[G \in Q] \to 1$ as $n \to \infty$ for sufficiently large (small) c.

All four steps combined will determine if the potential sharp threshold function is true or not.

Note that the random variable X in Step 4 needs to not be the same random variable as in Step 1. In which case we define a new random variable as Z, for example.

Further, note that in Step 3 and Step 4 there are two constants C and c that are either sufficiently large or sufficiently small. If C is sufficiently large in

Step 3 then c is sufficiently small in Step 4 and vice versa. Further, Step 3 and Step 4 are applied in any order.

In this dissertation we make use of the Chernoff Bounds, Theorem 5.6 and Theorem 5.7, in Step 3 and Step 4. The Chernoff Bounds are significantly sharper than the Markov's Inequality (Theorem 1.1) and Chebychev's Inequality (Theorem 1.2) that are used in the method for determining a threshold function. As seen in the proofs to follow, the Chernoff Bounds significantly decreases computational time as we bypass the need to calculate the challenging $E[X^2]$, as seen in the proof of Theorem 4.1.

In Chapter 4, the method of determining the threshold function can be modified to find the sharp threshold function. However, as seen from Theorem 4.5, the proof is long and tedious. In addition, the proofs are quite complex for random graphs with rainbow coloring. In particular, calculating $E[X]$ and $E[X^2]$ is quite difficult when rainbow coloring of graphs are considered. Thus, the new methodology in Chapter 5 provides a shorter proof with a sharper result and without the need to calculate $E[X^2]$.

This proposed methodology from Chapter 5 is applied below to illustrate how this method is applied.

5.4 Sharp Threshold Function for the Graph Property $rc(G) \leq 3$

We consider the proof by Yilun Shang in [47] for Theorem 5.9 in order to illustrate the method of finding the sharp threshold function. First, a lemma is stated. The purpose of the lemma is to illustrate that previously proved results are used in conjunction with the methodology in Chapter 5 to simplify other proofs. Note that results applying to the rainbow connection number are applied to random graphs to simplify the proofs.

Lemma 5.8. *[9]*

Let $G \in \mathbb{G}(n_1, n_2, p)$ be a non-complete bipartite graph. Let N and M be two distinct partite sets of G where n and m are the number of vertices in each partite set, respectively. Further, let any two vertices in N have greater than or equal to $2\alpha ln(n+m)$ common neighbors in M, where $\alpha = \frac{1}{ln(9/7)}$. Then, $rc(G) = 3$.

Theorem 5.9. *[47]*

Let $ln(n) << m < n$. For a random bipartite graph $G \in \mathbb{G}(n_1, n_2, p), p = \sqrt{ln(n)/m}$ is the sharp threshold function for the property $rc(G) \leq 3$.

Proof

If $diam(G) = 1$, then $G = K_{n,m}$ and so, $rc(G) = 1$.

First, we show that G meets the requirements of Lemma 5.8 so that $rc(G) = 3$. In this regard, let $p = c\sqrt{ln(n)/m}$ and c be an arbitrary constant.

Let N and M be two separate partite sets of $V(G)$ and $|E(G)| = |N||M|$. Then, p is the probability that an edge between N and M occurs. So, the probability that G is a complete graph is p^{nm}. Thus,

$$\begin{aligned} P[G = K_{n,m}] &= p^{nm} \\ &= \left(c\sqrt{\frac{ln(n)}{m}}\right)^{nm} \\ &= \left(\frac{c^2 ln(n)}{m}\right)^{\frac{nm}{2}}. \end{aligned}$$

Applying the natural logarithm on both sides of the equation yields

$$\begin{aligned} ln(p^{nm}) &= ln\left(\frac{c^2 ln(n)}{m}\right)^{\frac{nm}{2}} \\ nmln(p) &= \frac{nm}{2} ln(c^2 \frac{ln(n)}{m}) \\ ln(p) &= \frac{1}{2}[ln(c^2) + ln(ln(n)) - ln(m)]. \end{aligned}$$

Thus,

$$\begin{aligned} \lim_{n,m\to\infty} \frac{1}{2}[ln(c^2) + ln(ln(n)) - ln(m)] &= \frac{1}{2}[ln(c^2) + \lim_{n,m\to\infty} ln(ln(n)) - \lim_{n,m\to\infty} ln(m)] \\ &= \lim_{n,m\to\infty} \frac{1}{2nln(n)} - \lim_{n,m\to\infty} \frac{1}{2m} \\ &= 0. \end{aligned}$$

Therefore, the probability that G is a complete bipartite graph is zero, implying that G almost surely is a non-complete bipartite graph.

Now apply the methodology in Chapter 5.

Step 1: Define X.

Let X be a random variable which counts the number of common neighbours of a pair of vertices in G.

Step 2: Determine $E[X]$.

For $x, y \in N$ and $z \in M$ with $x \neq y$, we have

$$P[z \text{ a common neighbor of } x \text{ and } y] = p^2 = C^2 \frac{ln(n)}{m}.$$

Since $m < n$ by the assumptions of the theorem, the expectation of X is give by

$$E[X] = mp^2 = m(C^2 ln(n)/m) = C^2 ln(n).$$

Step 3: Show $P[G \in Q] \to 0$.

With $m \leq n$ and for a constant C sufficiently large,

$$\begin{aligned}
(n+m)^{16} &\leq n^{\frac{C^2}{2}} \\
e^{ln(n+m)^{16}} &\leq e^{ln(n)^{\frac{C^2}{2}}} \\
16ln(n+m) &\leq \frac{C^2}{2} ln(n) \\
16ln(n+m) &\leq C^2 ln(n) - \frac{C^2}{2} ln(n).
\end{aligned}$$

With $2\alpha ln(9/7) < 16$,

$$2\alpha ln(n+m) \leq C^2 ln(n) - \frac{C^2}{2} ln(n).$$

From the Chernoff's Bound (Theorem 5.6) and the fact that $E[X] = C^2 ln(n)$, it follows that

$$\begin{aligned} P[G \in Q] &= P[X < 2\alpha ln(n+m)] \\ &\leq P[X < E[X] - \frac{C^2 ln(n)}{2}] \\ &\leq e^{-(\frac{1}{2})^2 \cdot \frac{C^2 ln(n)}{2}} \\ &= e^{-\frac{C^2 ln(n)}{8}}. \end{aligned}$$

Show that $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$.

Consider,

$$\begin{aligned} \lim_{n\to\infty} \frac{e^{-\frac{C^2 ln(n)}{8}}}{\frac{1}{n^2}} &= \lim_{n\to\infty} n^2 e^{\frac{-C^2 ln(n)}{8}} \\ &= \lim_{n\to\infty} n^2 e^{ln(n)^{-(\frac{C^2}{8})}} \\ &= \lim_{n\to\infty} n^2 n^{-\frac{C^2}{8}} \\ &= \lim_{n\to\infty} \frac{n^2}{n^{\frac{C^2}{8}}} = 0, \end{aligned}$$

for $c > 4$. We have that

$$\lim_{n\to\infty} \frac{e^{-\frac{C^2 ln(n)}{8}}}{\frac{1}{n^2}} = 0, \tag{5.8}$$

meaning that

$$e^{-\frac{C^2 ln(n)}{8}} = o(n^{-2}).$$

Hence,

$$P[X < 2\alpha ln(n+m)] \leq e^{(-c^2 ln(n))/8} = o(n^{-2})$$

and

$$P[G \in Q] = P[X < 2\alpha ln(n+m)] \leq 0,$$

as $n \to \infty$.

The probability that a pair of vertices $x, y \in N$ has $X < 2\alpha ln(n+m)$ common neighbours with a vertex in M is zero. Hence, $X \geq 2\alpha ln(n+m)$. Note that x, y are arbitrary chosen and the conditions for Lemma 5.8 are met.

Since G is a bipartite graph and Lemma 5.8 does not specify the size of the two partite sets in G, it is required to repeat Step 2 and Step 3 in this proof above with partite sets M and N swapped around, as M and N have different orders.

Step 2: Determine $E[X]$.

Let $a, b \in M, a \neq b$. Then,

$$P[w \text{ a common neighbor of } a \text{ and } b] = p^2 = C^2 \frac{ln(n)}{m}.$$

We get $E[X] = np^2 = n(C^2 ln(n)/m)$.

Step 3: Show $P[G \in Q] \to 0$.

Make the initial assumption that $m < n$ and let C be sufficiently large. Assuming the same logic as Step 2 above with $(n+m)^{16} \leq n^{nC^2/(2m)}$, then

$$\begin{aligned} P[G \in Q] &= P[X < 2\alpha ln(n+m)] \\ &\leq P[X < \frac{nC^2 ln(n)}{2m}] \\ &= P[X < E[X] - \frac{nC^2 ln(n)}{2m}]. \end{aligned}$$

Using the Chernoff Bounds we have

$$\begin{aligned} P[X < 2\alpha ln(n+m)] &\leq P[X < E[X] - \frac{nC^2 ln(n)}{2m}] \\ &\leq e^{\frac{-(\frac{1}{2})^2 nC^2 ln(n)}{m^2}} \\ &= e^{\frac{-nC^2 ln(n)}{8m}}. \end{aligned}$$

Consider,

$$\begin{aligned} \lim_{n\to\infty} \frac{e^{-\frac{nC^2 ln(n)}{8m}}}{\frac{1}{n^2}} &= \lim_{n\to\infty} n^2 e^{\frac{-nC^2 ln(n)}{8m}} \\ &= \lim_{n\to\infty} n^2 e^{ln(n)^{\frac{-nC^2}{8m}}} \\ &= \lim_{n\to\infty} n^2 n^{\frac{-nC^2}{8m}} \\ &= \lim_{n\to\infty} \frac{n^2}{n^{\frac{nC^2}{8m}}} = 0, \end{aligned}$$

for $C > 4\sqrt{m/n}$, which is possible as m and n are finite.

We have that
$$\lim_{n\to\infty} \frac{e^{-\frac{nC^2 ln(n)}{8m}}}{\frac{1}{n^2}} = 0,$$
meaning that
$$e^{-\frac{nC^2 ln(n)}{8m}} = o(n^{-2}).$$

Hence,
$$P[G \in Q] = P[X < 2\alpha ln(n+m)] \leq 0,$$

as $n \to \infty$. In general, any two vertices of the same vertex partite set have at least $2\alpha ln(n+m)$ common neighbors. All the conditions for Lemma 5.8 have been met, and so $rc(G) = 3$.

Step 4: Show $P[G \in Q] \to 1$.

Now consider the other direction with sufficiently small c. Let A be the event that there exist vertices $x, y \in N, x \neq y$, where the pair of vertices x and y have no common neighbor in M.

For not event A, let p^2 be the probability that x and y have a neighbor $z \in M$. Then,
$$P[G \in Q] = P[A] = \left(1 - \frac{c^2 ln(n)}{m}\right)^m$$

and so the probability that $n/2$ pairs of vertices x and y have a common neighbor $z \in M$ is

$$1 - P[A] = \left(1 - \left(1 - \frac{c^2 ln(n)}{m}\right)^m\right)^{\frac{n}{2}}.$$

Solve,

$$\begin{aligned}
1 - P[A] &= \left(1 - \left(1 - \frac{c^2 ln(n)}{m}\right)^m\right)^{\frac{n}{2}} \\
&\sim \left(1 - e^{\frac{-mc^2 ln(n)}{m}}\right)^{\frac{n}{2}} \text{ by } (4.6) \\
&= (1 - n^{-c^2})^{\frac{n}{2}} \\
&\sim e^{-(\frac{n}{2})(n^{-c^2})} \text{ by } (4.6) \\
&= e^{(\frac{-n}{2n^{c^2}})}.
\end{aligned}$$

Consider,

$$\begin{aligned}
n &> 2n^{c^2} \\
e^n &> e^{c^2 2n} \\
n &> c^2 2n \\
\sqrt{\frac{1}{2}} &> c.
\end{aligned}$$

For $c < \sqrt{1/2}$,

$$1 - P[A] \to 0,$$

as $n \to \infty$ and,

$$P[G \in Q] = P[A] \to 1,$$

as $n \to \infty$. Hence, diam $(G) \ngeq 4$ and so $rc(G) \leq 3$. Hence, $p = \sqrt{ln(n)/m}$ is the sharp threshold function for the graph property $rc(G) \leq 3$ where $ln(n) << m \leq n$. □

5.5 Sharp Threshold Function for the Graph Property $rc(G) \leq 2$

For $G \in \mathbb{G}(n,p)$, the following proof is related to Theorem 4.5 where the sharp threshold function for $diam(G) \leq 2$ is addressed. Now, a proof for the threshold function $rc(G) \leq 2$ is given below. Interestingly enough, the sharp threshold functions are the same for the two results.

Lemma 5.10. *[9]*

For $G \in \mathbb{G}(n,p)$, any disconnected graph with $\delta(G) \geq \frac{n}{2} + ln(n)$ has $rc(G) = 2$

Theorem 5.11. *[9]*

For $G \in \mathbb{G}(n,p)$, the graph property $rc(G) \leq 2$ has $p = \sqrt{ln(n)/n}$ as the sharp threshold function.

Proof

For sufficiently large c with $p = \sqrt{ln(n)/n}$, we show that the graph $G \in \mathbb{G}(n,p)$ almost surely has $rc(G) = 2$. Thereafter, for sufficiently small C we show that G does not have a diameter of length two almost surely when $p = c\sqrt{ln(n)/n}$.

Step 1: Define X.

Let X count the number of common neighbours between u and v where $u, v \in V(G)$.

Step 2: Determine $E[X]$.

It is sufficient to show that if all pairs of vertices in G have at least $2ln(n)$ common neighbours, then the requirements for Lemma 5.10 mentioned above are met and $rc(G) = 2$.

Let u and v be two non-adjacent vertices of G. Assume that both u and v have at least $n/2 + ln(n)$ vertices connected to them independently. If there are no vertices adjacent to both u and v, then there are at least $(n/2+ln(n))+(n/2 + ln(n)) = n + 2ln(n)$ vertices in total in G. This is a contradiction as the maximum number of vertices in G is n. Therefore, u and v have at least $2ln(n)$ common neighbours and we meet the requirements for Lemma 5.10.

Let u and v now be a fixed pair of vertices in G and let p be the probability that u or v is adjacent to $w \in V(G)$. The probability that w is adjacent to both u and v is then p^2. Since u and v are fixed, we sum p^2 over the $(n-2)$ remaining vertices to count for all possible combinations of common neighbours. The expected number of common neighbours is then,

$$E[X] = (n-2)p^2 = \frac{n-2}{n} \cdot c^2 ln(n) > \frac{c^2 ln(n)}{2}.$$

Step 3: Show $P[G \in Q] \to 0$.

For sufficiently large c,

$$P[G \in Q] = P[X < 2ln(n)] \leq P[X < \frac{1}{4}c^2 ln(n)].$$

Consider $(n-2)/(2n) = 1/2 - 1/n$. For large n, $1/2 - 1/n \to 1/2$, and so

$$\begin{aligned} P[X < 2ln(n)] &\leq P[X < \frac{1}{4}c^2 ln(n)] \\ &\leq P[X < \frac{n-2}{2n}c^2 ln(n)] \\ &= P[X < \frac{n-2}{n}c^2 ln(n) - \frac{n-2}{2n}c^2 ln(n)]. \end{aligned}$$

From the Chernoff Bounds,

$$P[X < 2ln(n)] \leq e^{\frac{-\frac{1}{2}(n-2)c^2 ln(n)}{2n}}.$$

Consider,

$$\begin{aligned} & -\lim_{n\to\infty}\left(\frac{\frac{1}{2}(n-2)c^2 ln(n)}{2n}\right) \\ &= -\lim_{n\to\infty}\left(\frac{nc^2 ln(n) - 2c^2 ln(n)}{4n}\right) \\ &= -\lim_{n\to\infty}\left(\frac{c^2 ln(n) + c^2 - \frac{2c^2}{n}}{4}\right) && (5.9) \\ &= -\lim_{n\to\infty}\left(\frac{c^2}{4}ln(n) + \frac{c^2}{4} - \frac{2c^2}{4n}\right) \\ &= -\lim_{n\to\infty}\left(\frac{n^2}{4}ln(n) + \frac{n^2}{4} - \frac{2n^2}{4n}\right) && (5.10) \\ &= -\lim_{n\to\infty}\left(\frac{n^2}{4}ln(n) + \frac{n^2}{4} - \frac{n}{2}\right) \\ &= -\lim_{n\to\infty}\left(\frac{1}{4}((n + 2nln(n)) + \frac{n}{2} - \frac{1}{2})\right) && (5.11) \\ &\to -\infty, \end{aligned}$$

where L'Hopital's rule (Theorem 1.6) is used in Equation 5.9 and Equation 5.11. Equation 5.10 is obtained by substituting $c = n$. It follows that

$e^{-\infty} \to 0$ as $n \to \infty$, and

$$P[G \in Q] = P\left[x < 2ln(n) \leq e^{\frac{-\frac{1}{2}(n-2)c^2 ln(n)}{2n}}\right] \to 0.$$

As the number of common neighbours almost surely is larger than $2ln(n)$, we have that $rc(G) = 2$ by Lemma 5.10 for $c = n$.

Step 4: Show $\boldsymbol{P[G \in Q] \to 1}$**.**

For the next part, let R be a set containing $n^{1/5}$ vertices with the assumption that $n^{1/5}$ is an even positive integer. Let $Y = V(G) - R$ and we have $|y| = n - n^{1/5}$. Further, let A be the event that R and Y are the two partite sets of a bipartite graph. Let B be the event that there exists vertices in R with no common neighbors in Y. We now determine the probability of two independent events, A and B.

If p is the probability that $x \in R$ is adjacent to $w \in R$, then $1 - p$ is the probability that x and w are not adjacent. The number of combinations x and w that are chosen from R is $\binom{n^{1/5}}{2}$. So, the probability that all pairs are not adjacent is $(1-p)^{\binom{n^{1/5}}{2}}$. As

$$\binom{n^{\frac{1}{5}}}{2} = \frac{n^{\frac{1}{5}}!}{2!(n^{\frac{1}{5}} - 2)!} = \frac{n^{\frac{1}{5}}(n^{\frac{1}{5}} - 1)}{2},$$

it follows that

$$P[A] = (1-p)^{\binom{n^{\frac{1}{5}}}{2}}.$$

For $p = C\sqrt{ln(n)/n}$, we have

$$\begin{aligned} P[A] &= (1-p)^{\binom{n^{\frac{1}{5}}}{2}} \\ &= \left(1 - C\sqrt{\frac{ln(n)}{n}}\right)^{\binom{n^{\frac{1}{5}}}{2}} \\ &= exp\left(1 - \binom{n^{\frac{1}{5}}}{2} C\sqrt{\frac{ln(n)}{n}}\right) \text{ by } (4.6) \\ &= exp\left(\frac{-n^{\frac{1}{5}}(n^{\frac{1}{5}}-1)}{2} C\sqrt{\frac{ln(n)}{n}}\right). \end{aligned}$$

As

$$\lim_{n\to\infty}\left[\frac{\sqrt{ln(n)}}{\sqrt{n}}\right] = 0,$$

it follows that

$$\lim_{n\to\infty}\frac{n^{\frac{1}{5}}(n^{\frac{1}{5}}-1)}{2} C\sqrt{\frac{ln(n)}{n}} \to 0$$

implying that

$$P[A] = \frac{1}{exp^{\frac{n^{\frac{1}{5}}(n^{\frac{1}{5}}-1)}{2} C\sqrt{\frac{ln(n)}{n}}}} \to 1$$

Let p^2 be the probability that a pair $x, w \in R$ has a common neighbour $y \in Y$. Assuming there are $n^{1/5}/2$ pairs of vertices in R, it follows that $1-p^2$ is the probability that there is no common neighbour $y \in Y$. Thus,

$$(1-p^2)^{n-n^{1/5}}$$

is the probability that the pair (x, w) has no common neighbours. Then,

$$(1-(1-p^2)^{n-n^{1/5}})^{n^{1/5}/2}$$

is the probability that all pairs $x, w \in R$ have common neighbours in Y.

Consider

$$(1 - (1 - \frac{C^2 ln(n)}{n})^{n-n^{\frac{1}{5}}})^{\frac{n^{\frac{1}{5}}}{2}}.$$

As $n \to \infty$, note that for $C = n^{1/5}$, $(C^2 ln(n))/n \to 0$ and $(1-(C^2 ln(n))/n)^{n-n^{1/5}} \to 1$. Consequently,

$$(1 - (1 - \frac{C^2 ln(n)}{n})^{n-n^{\frac{1}{5}}})^{\frac{n^{\frac{1}{5}}}{2}} \to 0,$$

implying that at least one pair (x, w) has almost surely no common neighbors in Y, and $P[B] \to 1$.

We have that $P[A] \to 1$ and $P[B] \to 1$ independently. Hence, R and Y are two parties of a bipartite graph and there are no vertices between $x, q \in R$ and $y \in Y$. Hence, $diam(G) \geq 3$ and $P[G \in Q] \to 1$.

Whence, for $p = \sqrt{ln(n)/n}$ and sufficiently small C, G almost surely has $rc(G) > 2$, and for sufficiently large c, G almost surely has $rc(G) \leq 2$. Thus, the sharp threshold function $p = \sqrt{ln(n)/n}$ holds. □

5.6 Sharp Threshold Function for the Graph Property $diam(G) \leq 2$

We prove Theorem 4.5 again in the following result to illustrate the simplicity of the new methodology in comparison to its proof in Chapter 4. The following proof is our own.

Theorem 5.12. *[4]*

Let $G \in \mathbb{G}(n,p)$. Then,

$$p = \sqrt{\frac{ln(n)}{n}}$$

is the sharp threshold function for the graph property $diam(G) \leq 2$.

Proof

Step 1: Define X.

Let X count the number of common neighbors.

Step 2: Determine $E[X]$.

Let $G \in \mathbb{G}(n,p)$ and $x, y \in V(G)$, $x \neq y$. Further, let p be the probability that x is adjacent to a vertex of $V(G) - \{x\}$. Then p^2 is the probability that a vertex is adjacent to both x and y. The maximum possible number of common neighbors for $p = c\sqrt{ln(n)/n}$ is then given by

$$E[X] = np^2 = \frac{nc^2 ln(n)}{n}.$$

Step 3: Show $P[G \in Q] \to 0$.

The probability that X is less than the maximum number of common neighbors is

$$\begin{aligned} P[G \in Q] &= P[X < c^2 ln(n)] \\ &\leq P\left[X < E[X] - \frac{c^2 ln(n)}{2}\right]. \end{aligned}$$

By applying the Chernoff Bound (Theorem 5.6), it follows that

$$P[X < c^2 ln(n)] \leq e^{-\frac{1}{8}\frac{c^2 ln(n)}{n}}.$$

Equation 5.8 showed

$$\lim_{n\to\infty} \frac{e^{-\frac{c^2 ln(n)}{8}}}{\frac{1}{n^2}} = 0.$$

For $c > 4$, it follows that

$$e^{\frac{-c^2 ln(n)}{8}} = o(n^{-2})$$

and

$$P[G \in Q] = P[X < c^2 ln(n)] \to 0.$$

Reading the implication in words, we have that the probability of the number of common neighbors X is less than the maximum possible number of common neighbors, tends to zero as n tends to infinity. That is, X is equal to the maximum number of common neighbors and $diam(G) \leq 2$.

Step 4: Show $P[G \in Q] \to 1$.

For the other direction, let A be the event that there are no common neighbors between x and y. Then,

$$P[A] = (1 - p^2)^n.$$

The event that there are only common neighbors between x and y is given by

$$\begin{aligned} P[G \in Q] &= 1 - P[A] \\ &= 1 - (1 - p^2)^n \\ &= 1 - \left(1 - \frac{C^2 ln(n)}{n}\right)^n \\ &= 1 - (1 - e^{\frac{-nC^2 ln(n)}{n}}) \text{ by } (4.6) \\ &= 1 - (1 - e^{-C^2 ln(n)}). \end{aligned} \tag{5.12}$$

For $C = \sqrt{1/n}$,

$$\lim_{n\to\infty} (1 - P[A]) = \lim_{n\to\infty} (1 - (1 - e^{-C^2 ln(n)})) \to 1.$$

Hence, $P[G \in Q] \to 1$ and there are only common neighbors between x and y, where x and y are arbitrary. Therefore, $diam(G) \leq 2$. $\square$

The proof above is substantially shorter than the proof in Theorem 4.5. In addition, the proof in Theorem 4.5 found a threshold function, whereas the proof above found a sharp threshold function which is a sharper result. The method discovered in Chapter 5 is better to use than the method in Chapter 4 when dealing with colorings of random graphs.

The remainder of this chapter is dedicated to more challenging theorems that combine random graphs with rainbow coloring.

5.7 Sharp Threshold Function for the Graph Property $rc_k(G) \leq d$

We first state a few lemmas and assumptions before proceeding to the last proof. These results appear in the last proof, but the author omitted the proofs, thus, we provide our own proofs. For the remainder of Chapter 5, the logarithm with base e is denoted by ln, and the logarithm with base 2 is denoted by log. This distinction is important as it is later shown that the small difference between log and ln aids in finding the sharp threshold function. The difference between ln and log is equivalent to using large C and small c as part of Step 3 and Step 4 in our method to find a sharp threshold function. Also, we define the following constants which are used throughout the remainder of Chapter 5. Let $c_0 \geq 1$ be a constant at least one and $k = k(n) \leq c_0 log(n)$ for large n. Further, let $c = 2^{20}c_0$ and $c_1 = 2^{10d}c_0$ where d is the diameter of $G \in \mathbb{G}(n, p)$.

We define the binary entropy function as

$$H(z) = zlog\left(\frac{1}{z}\right) + (1-z)log\left(\frac{1}{1-z}\right).$$

Lemma 5.13. *[25]*

For $z = m/n$ in the binary entropy function,

$$\binom{n}{m} \leq 2^{nH(m/n)}.$$

Proof

To begin, consider the formula

$$\sum_{m=0}^{n} \binom{n}{m} p^m (1-p)^{n-m} = 1$$

which is the sum of individual probabilities using the Binomial model. Let $p = m/n$ and consider only the m^{th} term which is at most one. Then,

$$\begin{aligned} 1 &\geq \binom{n}{m} \left(\frac{m}{n}\right)^m \left(1 - \frac{m}{n}\right)^{n-m} \\ &= \binom{n}{m} e^{log(m/n)^m} \cdot e^{log(1-m/n)^{n-m}} \\ &= \binom{n}{m} e^{mlog(m/n)} \cdot e^{(n-m)log(1-m/n)}. \end{aligned}$$

It is known that $2^x \leq e^x$ for all $x \geq 0$. Therefore,

$$\begin{aligned} 1 &\geq \binom{n}{m} 2^{mlog(m/n)} \cdot 2^{(n-m)log(1-m/n)} \\ &= \binom{n}{m} 2^{mlog(m/n)+(n-m)log(1-m/n)} \\ &= \binom{n}{m} 2^{n(\frac{m}{n}log(m/n)+\frac{n-m}{n}log(\frac{n-m}{n}))}. \end{aligned}$$

Let variable $z = m/n$ so that

$$\begin{aligned} 1 &\geq \binom{n}{m} 2^{n(zlog(z)+(1-z)log(1-z))} \\ &= \binom{n}{m} 2^{-n(zlog(1/z)+(1-z)log(1/(1-z)))}. \end{aligned}$$

From the definition of the binary entropy function, we have

$$1 \geq \binom{n}{m} 2^{-nH(z)} = \binom{n}{m} 2^{-nH(m/n)}$$

Consequently,

$$\binom{n}{m} \leq 2^{nH(m/n)}.$$

□

Lemma 5.14. *[25]*

For all $x \geq 0$, $1 - x \leq e^{-x} \leq 2^{-x}$.

Proof

Consider the first inequality, $1-x \leq e^{-x}$, and let $f(x) = e^{-x} - (1-x)$. Then, $f'(x) = -e^{-x} + 1$, and $f'(x) = 0$ gives $e^x = 1$. This is true if and only if $x = 0$.

Now, consider the second derivative of $f(x)$ which is $f''(x) = e^{-x}$. As $e^{-x} > 0$ for all $x \in \mathbb{R}$, $f''(0) = e^0 = 1 \geq 0$. Hence, there is a global minimum for $f(x)$, resulting in $e^{-x} - (1 - x) \geq 0$ and $1 - x \leq e^{-x}$.

For the other inequality, $e^{-x} \leq 2^{-x}$, it is known that $e^{x} \geq 2^{x}$ for all $x \geq 0$. Therefore, $e^{-x} \leq 2^{-x}$ for all $x \geq 0$.

In conclusion for all $x \geq 0$, $1 - x \leq e^{-x} \leq 2^{-x}$. □

Lemma 5.15. *[25]*

For $n > 0$, $2^{log(n)} = n$.

Proof

For $x = 2^{log(n)}$, we have

$$x = 2^{log(n)}$$

$$log(x) = log(2^{log(n)})$$

$$log(x) = log(n)log(2)$$

$$log(x) = log(n).$$

Hence $x = n$ and $x = 2^{log(n)}$. The result then follows. □

Lemma 5.16. *[25]*

For sufficiently large $n > 0$, $log(n) \leq \sqrt{n}$.

Proof

Divide the two terms and take their limit as n tends to infinity to establish which term is largest. Applying L'Hopital's rule,

$$\begin{aligned}
\frac{\frac{d}{dn}log_2(n)}{\frac{d}{dn}\sqrt{n}} &= \frac{\frac{d}{dn}\frac{log(n)}{log(2)}}{\frac{d}{dn}\sqrt{n}} \\
&= \frac{\frac{1}{nlog(2)}}{\frac{1}{2\sqrt{n}}} \\
&= \frac{2\sqrt{n}}{nlog(2)} \cdot \frac{\sqrt{n}}{\sqrt{n}} \\
&= \frac{2}{\sqrt{n}log(2)}.
\end{aligned}$$

Consequently,

$$\lim_{n\to\infty} \frac{2}{\sqrt{n}log(2)} = 0.$$

Hence, $\sqrt{n}$ is larger than $log(n)$ for sufficiently large n. □

Lemma 5.17. *[25]*

For all $x \geq 0$, $1 + x \leq e^x \leq 2^{2x}$.

Proof

As $e < 2^2 = 4$, we have $e^x \leq 2^{2x}$. For $e^x \geq 1 + x$, let $f(x) = e^x - (1 + x)$. Then,

- $f(x) = e^x - (1 + x)$,
- $f'(x) = e^x - 1$,
- $f'(x) = 0 \iff x = 0$,
- $f''(x) = e^x$,
- $f''(0) = e^0 = 1 > 0$.

Hence, there is a minimum for $f(x)$, and $e^x - (1+x) \geq 0$ for all $x \geq 0$. Thus, $e^x \geq 1 + x$, and the result is proved. □

Lemma 5.18. *[25]*

For all $x \geq 0$, $log(1 + x) \leq 2x$.

Proof

Applying log to $1 + x \leq 2^{2x}$ from Lemma 5.17 yields

$$log(1 + x) \leq log(2^{2x}) = 2xlog(2) = 2x.$$

Hence, $log(1 + x) \leq 2x$. □

Lemma 5.19. *[25]*

For c_0 and c_1 as defined on Page 93, $H(c_0/c_1) \leq 3\sqrt{c_0/c_1}$.

Proof

By substituting $z = c_0/c_1$ in to the binary entropy function, we have

$$H\left(\frac{c_0}{c_1}\right) = \left(\frac{c_0}{c_1}\right) log\left(\frac{c_1}{c_0}\right) + \left(1 - \frac{c_0}{c_1}\right) log\left(1 + \frac{c_0}{c_1 - c_0}\right).$$

Applying Lemma 5.16 and Lemma 5.18 yields

$$\begin{aligned}
H\left(\frac{c_0}{c_1}\right) &\leq \left(\frac{c_0}{c_1}\right)\sqrt{\frac{c_1}{c_0}} + \left(1-\frac{c_0}{c_1}\right) log\left(1+\frac{c_0}{c_1-c_0}\right) \\
&\leq \left(\frac{c_0}{c_1}\right)\sqrt{\frac{c_1}{c_0}} + \left(1-\frac{c_0}{c_1}\right) 2\left(\frac{c_0}{c_1-c_0}\right) \\
&= \sqrt{\left(\frac{c_0}{c_1}\right)^2}\sqrt{\frac{c_1}{c_0}} + \left(1-\frac{c_0}{c_1}\right) 2\left(\frac{c_0}{c_1-c_0}\right) \\
&= \sqrt{\frac{c_1 c_0^2}{c_0 c_1^2}} + \left(\frac{c_1-c_0}{c_1}\right) 2\left(\frac{c_0}{c_1-c_0}\right) \\
&= \sqrt{\frac{c_0}{c_1}} + 2\frac{c_0}{c_1}.
\end{aligned}$$

Recall that $c_1 = 2^{10d}c_0$, resulting in

$$\sqrt{\frac{c_0}{c_1}} + 2\left(\frac{1}{2^{10d}}\right).$$

Since the second term is a proper fraction, we have

$$\frac{1}{2^{10d}} \leq \sqrt{\frac{1}{2^{10d}}}$$

and so,

$$\begin{aligned}
\sqrt{\frac{c_0}{c_1}} + 2\cdot\frac{1}{2^{10d}} &\leq \sqrt{\frac{c_0}{c_1}} + 2\cdot\sqrt{\frac{1}{2^{10d}}} \\
&= \sqrt{\frac{c_0}{c_1}} + 2\cdot\sqrt{\frac{c_0}{c_1}} \\
&= 3\cdot\sqrt{\frac{c_0}{c_1}}.
\end{aligned}$$

□

Before stating the main result, we state the following lemma without proof. Recall, the omega notation is given by $f(n) = \Omega(g(n))$, if $\lim_{n\to\infty} f(n)/g(n) > 0$ (otherwise known as a lower bound).

Lemma 5.20. *[25]*

For $G \in \mathbb{G}(n, C(log(n))^{1/d}/n^{(d-1)/d})$, every pair of vertices $(x, y) \in V(G)$, is connected by at least $C_1 log(n)$ internally vertex-disjoint paths of length d with probability at least $1 - n^{-\Omega(1)}$.

The complexity increases as we proceed to the main result on the k-connectivity of the rainbow coloring of a graph $G \in \mathbb{G}(n, p)$. The proof thereof is an amalgamation of many concepts discussed throughout this dissertation.

Theorem 5.21. *[25]*

For $G \in \mathbb{G}(n, p)$ the equation

$$p = \frac{(log(n))^{\frac{1}{d}}}{n^{\frac{d-1}{d}}}$$

is the sharp threshold function for the graph property $rc_k(G) \leq d$ where $d \geq 2$ is the graph diameter and $k = k(n) \leq O(log(n))$.

Proof

From Lemma 5.20, $G \in \mathbb{G}(n, p)$ has at least $C_1 log(n)$ possible rainbow paths being the number of internally disjoint paths of length d as $k \leq C_1 log(n)$.

Step 1: Define X.

Let X count the number of non-rainbow paths.

Step 2: Determine E[X].

The number of ways the colors of one path between two fixed vertices, where no color is to be repeated, is $d!$. With d paths of length d, the number of colors chosen is d^d. Therefore, the probability that a rainbow path occurs is $q = d!/d^d$. We have for large n, using the Sterling Approximation (Theorem 1.5),

$$\begin{aligned} q = \frac{d!}{d^d} &\sim \frac{\sqrt{2\pi d}(\frac{d}{e})^d}{d^d} \\ &\geq \frac{(\frac{d}{e})^d}{d^d} \\ &= \frac{1}{e^d} \\ &\geq \frac{1}{4^d}. \end{aligned}$$

Since there are at least $C_1 log(n)$ rainbow paths, G contains $(k-1)$ rainbow paths as an upper bound. The probability then of no paths being rainbow colored is given by

$$E[X] = \binom{c_1 log(n)}{k-1}(1-q)^{c_1 log(n)-(k-1)}.$$

Step 3: Show $P[G \in Q] \to 1$.

Consider,

$$
\begin{aligned}
&\binom{c_1 log(n)}{k-1}(1-q)^{c_1 log(n)-(k-1)} \\
&\leq \binom{c_1 log(n)}{c_0 log(n)}(1-4^{-d})^{c_1 log(n)-c_0 log(n)+1} && (5.13)\\
&\leq \binom{c_1 log(n)}{c_0 log(n)}(1-4^{-d})^{log(n)(c_1-c_0)} && (5.14)\\
&\leq 2^{c_1 log(n) H(\frac{c_0 log(n)}{c_1 log(n)})}(1-4^{-d})^{log(n)(c_1-c_0)} && (5.15)\\
&\leq 2^{c_1 log(n) H(\frac{c_0 log(n)}{c_1 log(n)})} 2^{-4^d log(n)(c_1-c_0)} && (5.16)\\
&= 2^{c_1 log(n) H(\frac{c_0 log(n)}{c_1 log(n)}) - 4^{-d} log(n)(c_1-c_0)} \\
&= 2^{c_1 log(n) H(\frac{c_0}{c_1}) - 4^{-d} log(n)(c_1-c_0)} \\
&= 2^{log(n)(c_1 H(c_0/c_1) - 4^{-d}(c_1-c_0))} \\
&= n^{-(4^{-d}(c_1-c_0) - c_1 H(c_0/c_1))}. && (5.17)
\end{aligned}
$$

Above, Equation 5.13 is produced by replacing $k = c_0 log(n)$ as stipulated by the conditions for this proof. For Equation 5.14, note that

$$(1-4^{-d})^{c_1 log(n)-c_0 log(n)+1}$$

is between zero and one as $d > 1$ and $c_1 > c_0$ are positive constants. Equation 5.15 follows by applying Lemma 5.13, and Equation 5.16 by using Lemma 5.14. Lastly, Equation 5.17 follows from Lemma 5.15.

From Lemma 5.19, the exponent of Equation 5.17 becomes

$$\begin{aligned} 4^{-d}(c_1 - c_0) - c_1 H(c_0/c_1) &\geq 4^{-d}(c_1 - c_0) - c_1 \cdot 3\sqrt{c_0/c_1} \\ &= 4^{-d}(c_1 - c_0) - 3\sqrt{c_1^2}\sqrt{c_0/c_1} \\ &= 4^{-d}(c_1 - c_0) - 3\sqrt{c_0 c_1}. \end{aligned}$$

Recall that $c_1 = 2^{10d}c_0$, implying that $c_0 = c_1/(2^{10d})$ and

$$\begin{aligned} 4^{-d}(c_1 - c_0) - 3\sqrt{c_0 c_1} &= 4^{-d}\left(c_1 - \frac{c_1}{2^{10d}}\right) - 3\sqrt{\frac{c_1 c_1}{2^1 0d}} \\ &= 4^{-d}c_1(1 - 2^{-10d}) - 3\sqrt{2^{-10d}c_1^2} \\ &= 2^{-2d}c_1(1 - 2^{-10d}) - 3\sqrt{2^{-10d}c_1^2} \\ &\geq 2^{-2d}c_1(1 - 2^{-5d+2}) - 3c_1\sqrt{2^{-5d+2}}. \end{aligned} \tag{5.18}$$

Note that $(1 - 2^{-5d+2}) \geq 2^{-1}$ as $1/(2^{-5d+2}) < 1/2$.

Further, note that $\sqrt{1/(2^{-5d+2})} < 1$ as $d \geq 2$. Hence,

$$c_1 2^{-5d+2} \leq 3c_1 2^{-5d+2} \leq 3c_1\sqrt{2^{-5d+2}}.$$

Applying these two results back into Equation 5.18 yields

$$\begin{aligned} &2^{-2d}c_1(1 - 2^{-5d+2}) - 3c_1\sqrt{2^{-5d+2}} \\ &\geq 2^{-2d}2^{-1}c_1 - 2^{-5d+2}c_1 \\ &= 2^{-2d-1}c_1 - 2^{-5d+2}c_1 \\ &= 2^{-2d-1}c_1 - 2^{-2d-1}2^{-3d+3}c_1 \end{aligned}$$

$$2^{-2d-1}c_1 - 2^{-2d-1}2^{-3d+3}c_1 = c_1 2^{-2d-1}(1 - 2^{-3d+3}) \quad (5.19)$$
$$\geq c_1 2^{-2d-1}2^{-1} \quad (5.20)$$
$$= c_1 2^{-2d-2} \quad (5.21)$$
$$= c_0 2^{10d} 2^{-2d-2} \quad (5.22)$$

as $c_1 = 2^{10d}c_0$. As $c_0 \geq 1$ and $d \geq 2$, it follows that

$$c_0 2^{10d} 2^{-2d-2} > 100.$$

The value 100 is chosen to be large enough for the inequality to hold. Returning to Equation 5.17, we now have

$$n^{(4^{-d}(C_1 - C_0) - C_1 H(C_0/C_1))} < n^{-100}.$$

Thus, the probability that there are no rainbow paths in G tends to zero as n tends to infinity. Since there are $\binom{n}{2}$ pairs of vertices in G the probability that every two vertices in G have k rainbow paths is given by

$$P[G \in Q] = 1 - \binom{n}{2} n^{-100}.$$

Consider

$$\begin{aligned}\binom{n}{2} n^{-100} &= \frac{n!}{2!(n-2)!n^{10}n^{90}} \\ &= \frac{n(n-1)}{2n^{10}n^{90}} \\ &= \frac{n-1}{2n^9 n^{90}}.\end{aligned}$$

Note that $(n-1)/(2n^9) > 1/n^{90}$ for all $n > 0$. Hence, the probability that every two vertices in G have k rainbow paths is given by

$$P[G \in Q] = 1 - \binom{n}{2} n^{-100} \geq 1 - n^{-90},$$

confirming that Lemma 5.20 holds with at least k internal disjoint rainbow paths. Since every pair of vertices in G have at least k internal disjoint rainbow paths of length d, G is rainbow k-connected and $rc_k(G) \leq d$.

Step 4: Show $P[G \in Q] \to 0$.

In Step 4, the methodology from Chapter 5 is not applied. Instead, an alternative method is provided to illustrate the relationship between many of the results proved in this dissertation. However, what we are doing is equivalent to showing $P[G \in Q] \to 0$.

In Theorem 4.5, we proved that $diam(G) \leq 2$ has threshold function $p = \sqrt{2}\sqrt{log(n)/n}$. The result of Theorem 4.5 is then generalised in Theorem 4.6 where we proved that $diam(G) \geq 2$ has threshold function $p = (ln(n^2/c))^{1/d}/n^{(d-1)/d}$.

Further, it was proved in [4] that the threshold function in Theorem 4.5 is in fact the sharp threshold function and the result in Theorem 4.6 is also the sharp threshold function. Hence,

$$p = \frac{ln(\frac{n^2}{c})^{\frac{1}{d}}}{n^{\frac{d-1}{d}}}$$

is the sharp threshold function for the graph property $diam(G) \geq d$.

Thus, the probability $diam(G) \leq d$ is zero. Let $p^{'} = (ln(n))^{1/d}/n^{(d-1)/d}$, then $p^{'} \leq p$ for all $n > c$. For $diam(G^{'}) \in \mathbb{G}(n, p^{'})$, we have that $diam(G^{'}) \leq d = 0$.

From the observation that the rainbow $k-$connectivity of a graph is greater than or equal to its diameter, we have

$$rc_k(G^{'}) \geq d + 1$$

with sharp threshold function

$$p^{'} = \frac{(ln(n))^{\frac{1}{d}}}{n^{\frac{d-1}{d}}}.$$

In Step 3, we showed that the sharp threshold function $p = (log(n))^{1/d}/n^{(d-1)/d}$ is true for $rc_k(G) \leq d$. We have now shown that if $p = (log(n))^{1/d}/n^{(d-1)/d}$ increased slightly by switching from log to ln, the graph property $rc_k(G) \leq d$ no longer holds. Thus, $p = (log(n))^{1/d}/n^{(d-1)/d}$ is the sharp threshold function for $rc_k(G) \leq d$. $\square$

5.8 Conclusion

Let Q be a graph property. Recall that a graph $G \in \mathbb{G}(n, p)$ satisfies Q almost surely if $P[G \in Q] \to 1$ as $n \to \infty$. Likewise, G almost surely does not satisfy Q if $P[G \notin Q] \to 0$ as $n \to \infty$.

For any random graph G, a function $f(n)$ is called the *sharp threshold function* for the graph property Q if there exists constants C and c, both positive, such that G satisfies Q almost surely for $p \geq Cf(n)$, and G almost surely does not satisfy Q for $p \leq cf(n)$. In other words, the definition of a sharp threshold function for an increasing graph property is,

$$\lim_{n\to\infty} P[G \in Q] = \begin{cases} 1 & p \geq Cf(n), \\ 0 & p \leq cf(n). \end{cases}$$

For a decreasing graph property, the same definition holds but with zero and one interchanged.

As mentioned before, a motivation for using random graphs as opposed to simple graphs is to decrease the complexity and computational time of these problems.

We have that all monotone graph properties have a threshold function as well as a sharp threshold function. It is still an open problem if these results hold for non-monotone properties.

The Chernoff Bounds simplified our calculations.

Theorem 5.22. ***Chernoff Bounds - Lower Tail***

Let $X = \{x_1, x_2 \ldots x_n\}$ *be an independent random variable. Further, let* $\delta \in (0, 1]$, $\mu = E[X] = \sum_{i=1}^{n} P_i$, *and* $P[x_i = 1] = P_i$. *Then,*

$$P[X < (1-\delta)\mu] \leq e^{\frac{-\mu\delta^2}{2}}.$$

Theorem 5.23. ***Chernoff Bounds - Upper Tail***

Let $X = \{x_1, x_2, \ldots, x_n\}$ *be an independent random variable and let* $\delta \in (0, 1]$, $\mu = E[X] = \sum_{i=1}^{n} P_i$, *and* $P[x_i = 1] = P_i$. *Then,*

$$P[X > (1 + \delta)\mu] < 2^{-\delta\mu}.$$

Using the Chernoff Bounds, a method for determining a sharp threshold function is stated as follows for any random graph G and monotone property Q.

Step 1: Define X.

Step 2: Determine $E[X]$.

Step 3: Show $P[G \in Q] \to 0$.

Step 4: Show $P[G \in Q] \to 1$.

It is established that the method for determining a sharp threshold function (Section 5.3) is more efficient than the method for determining a threshold function (Section 4.2), as there is no need to calculate $E[X^2]$. The latter proved to be computationally challenging in Chapter 4. Further, four results are proved using the method for determining a sharp threshold function.

Theorem 5.24.

Let $ln(n) << m < n$. *For a random bipartite graph* $G \in \mathbb{G}(n_1, n_2, p), p = \sqrt{ln(n)/m}$ *is the sharp threshold function for the property* $rc(G) \leq 3$.

Theorem 5.25.

For $G \in \mathbb{G}(n,p)$, the graph property $rc(G) \leq 2$ has $p = \sqrt{ln(n)/n}$ as the sharp threshold function.

Theorem 5.26.

Let $G \in \mathbb{G}(n,p)$. Then,

$$p = \sqrt{\frac{ln(n)}{n}}$$

is the sharp threshold function for the graph property $diam(G) \leq 2$.

Theorem 5.27.

For $G \in \mathbb{G}(n,p)$ the equation $p = (log(n))^{1/d}/n^{(d-1)/d}$ is the sharp threshold function for the graph property $rc_k(G) \leq d$ where $d \geq 2$ is the graph diameter and $k = k(n) \leq O(log(n))$.

We conclude our investigation of the threshold functions of colorings of random graphs in the next chapter.

Chapter 6

Closing Chapter

6.1 Closing Remarks

The aim of this dissertation is to investigate random graphs and rainbow coloring, and to use the results of this investigation to create a method for determining sharp threshold functions.

First, in Chapter 1, fundamental building blocks of Graph Theory, Probability Theory, Calculus, Combinatorics, and Asymptotic Theory is established.

Chapter 2 dives into the understanding of random graphs. A number of random graph models are discussed, but particular attention is given to the Binomial Model as we use this model in the rest of the dissertation.

Let $F_{n,p}$ be the set of all graphs with n labelled vertices and a probability $p = p(n)$ of an edge occurring independently where $0 \leq p \leq 1$. Each graph

$G \in F_{n,p}$ has probability

$$P[G] = p^m(1-p)^{\binom{n}{2}-m}$$

of occurring where m is the number of edges occurring in G. This model is known as the Binomial model or sometimes the Bernoulli model. A graph G which is part of the Binomial graph model is denoted by $G \in \mathbb{G}(n,p)$.

An important concept around random graphs is established in Theorem 2.1 that tells us that the Binomial model and Uniform model are asymptotically equivalent.

Rainbow coloring is investigated in Chapter 3, starting with the history of rainbow coloring, and ending with some important results.

Let G be a connected graph and $x, y \in V(G)$ with $x \neq y$. Define a rainbow path as a path in G where no two edges on the path between x and y are colored the same. Further, a connected graph G is rainbow connected if any two distinct vertices $x, y \in V(G)$ are connected by a rainbow path.

If there are at least k internally disjoint rainbow paths between any two vertices in G, then G is rainbow k-connected. Further, the minimum number of colors needed for G to be rainbow k-connected is then called the rainbow k-connection number, denoted $rc_k(G)$.

We observe that the diameter and the minimum degree of a graph is fundamental in working with rainbow graphs and the rainbow connection number.

We trivially have that

$$diam(G) \leq rc(G) \leq src(G) \leq |E(G)|.$$

We noticed throughout this investigation that a general motif appeared when investigating rainbow coloring; that is, the graph diameter or the minimum degree. The diameter of a graph plays a fundamental role in finding rainbow connection numbers. Logically this makes sense as a rainbow path is bounded below by the graph diameter. Most results for rainbow coloring feature the graph diameter or minimum degree.

With an understanding of rainbow coloring and random graphs, the study shifted to applying these concepts in determining threshold functions. This leads to Chapter 4 where threshold functions are defined and investigated for monotone graph properties.

1. Let Q be a graph property that is increasing. Let $G \in \mathbb{G}(n,p)$ with $|V(G)| = n$. Further, let $p = p(n)$ be the probability of an edge occurring in G with $p \in [0,1]$. A function $t = t(n)$ is called the threshold function for Q

    - if $p << t$, then $\lim_{n\to\infty} P[G \text{ has propertry } Q] = 0$, and
    - if $p >> t$, then $\lim_{n\to\infty} P[G \text{ has propertry } Q] = 1$.

    Here, the symbol $<<$ is defined to be conceptually much less than, and the symbol $>>$ is conceptually much greater than. In other words, if $p << t$, then $p/t \to 0$, and if $p >> t$, then $p/t \to \infty$.

2. Using the same conditions as in (1), let Q now be a decreasing graph property. Then, $t = t(n)$ is the threshold function for Q

   - if $p << t$, then $\lim_{n\to\infty} P[G$ has propertry $Q] = 1$, and

   - if $p >> t$ then, $\lim_{n\to\infty} P[G$ has propertry $Q] = 0$.

For $P[G$ has property $Q] = 1$, G is said to asymptotically almost surely have Q. Similarly, for $P[G$ has property $Q] = 0$, G is said to asymptotically almost surely not have Q.

A fundamental result of threshold functions is that every monotone graph property has a threshold function. The same is not necessarily true for non-monotone graph properties.

A method for determining the threshold function is investigated and proved using the First Moment (Theorem 4.1) and the Second Moment (Theorem 4.2). The steps to finding the threshold functions are summarised as follows.

Step 1: Define X.

Step 2: Determine $E[X]$.

Step 3: Find a potential threshold function.

Step 4: Apply the First Moment method.

Step 5: Apply the Second Moment method.

This method was used to create our own proof for the fundamental results proved by Bollobas below.

Theorem 6.1.

The threshold function for a graph $G \in \mathbb{G}(n,p)$ with $diam(G) = 2$ is

$$t = \sqrt{\frac{ln(n)}{n}}.$$

Theorem 6.2.

The threshold function for a graph $G \in \mathbb{G}(n,p)$ with $diam(G) = d \geq 2$ is

$$t^d = \frac{ln(\frac{n^2}{c})}{n^{d-1}}.$$

The proof of Theorem 6.1 is long and made use of many results from real analysis.

Lastly, we proved in Theorem 4.3 that threshold functions are not unique and, in fact, many threshold functions could exist for the same graph property up to a constant product.

Chapter 5 expanded on Chapter 4 by investigating the sharp threshold functions.

Let Q be a graph property. Recall that a graph $G \in \mathbb{G}(n,p)$ satisfies Q almost surely if $P[G \in Q] \to 1$ as $n \to \infty$. Likewise, G almost surely does not satisfy Q if $P[G \notin Q] \to 0$ as $n \to \infty$.

For a random graph G, a function $f(n)$ is called the sharp threshold function for the graph property Q if there exists constants C and c, both positive, such that G satisfies Q almost surely for $p \geq Cf(n)$, and G almost surely does not satisfy Q for $p \leq cf(n)$.

In other words, the definition of a sharp threshold function for an increasing graph property is

$$\lim_{n \to \infty} P[G \in Q] = \begin{cases} 1 & p \geq Cf(n), \\ 0 & p \leq cf(n). \end{cases}$$

For a decreasing graph property, the definitions hold but with zero and one interchanged.

An important point here is that both threshold functions and sharp threshold functions have been proved to exist for all monotone graph properties. This is not necessarily true for non-monotone graph properties such as the strong rainbow connection number.

In addition to this, an in-depth study into sharp threshold functions with particular application to rainbow coloring is conducted. Through this investigation, we discovered that the sharp threshold function of the form $p = \sqrt{log(n)/n}$, or a variant of it, is common amongst many sharp threshold function results. This leads us to consider sharp threshold functions of this form first when in search of new sharp threshold functions for other graph properties relating to random graphs, especially when considering properties including rainbow coloring.

Through reviewing numerous proofs on sharp threshold functions, we define our own method for finding sharp threshold functions, specifically for rainbow colored random graphs. This method requires the use of the Chernoff Bounds.

A summary of the method to find the sharp threshold function is as follows.

Step 1: Define X.

Step 2: Determine $E[X]$.

Step 3: Show $P[G \in Q] \to 0$.

Step 4: Show $P[G \in Q] \to 1$

All four steps combined will determine if the potential sharp threshold function is true or not.

We established that the method for determining a sharp threshold function (Section 5.3) is more efficient than the method for determining a threshold function (Section 4.2), as there is no need to calculate $E[X^2]$. This proved to be computationally challenging. Further, four results are proved using the method for determining a sharp threshold function.

Theorem 6.3.

Let $ln(n) << m < n$. For a random bipartite graph $G \in \mathbb{G}(n_1, n_2, p), p = \sqrt{ln(n)/m}$ is the sharp threshold function for the property $rc(G) \leq 3$.

Theorem 6.4.

For $G \in \mathbb{G}(n, p)$, the graph property $rc(G) \leq 2$ has $p = \sqrt{ln(n)/n}$ as the sharp threshold function.

Theorem 6.5.

Let $G \in \mathbb{G}(n,p)$. Then,

$$p = \sqrt{\frac{ln(n)}{n}}$$

is the sharp threshold function for the graph property $diam(G) \leq 2$.

Theorem 6.6.

For $G \in \mathbb{G}(n,p)$, the equation $p = (log(n))^{1/d}/n^{(d-1)/d}$ is the sharp threshold function for the graph property $rc_k(G) \leq d$ where $d \geq 2$ is the graph diameter and $k = k(n) \leq O(log(n))$.

By investigating rainbow coloring, random graphs, threshold functions and sharp threshold functions, we are able to state our own method above.

6.2 Further Work

In this section, open problems and future work to this dissertation are presented.

The focus of this dissertation is on rainbow coloring. First, vertex coloring and total coloring is considered for future work as in this dissertation the focus was on edge coloring. Second, there are many other graph colorings which are all applied to random graphs. One such example is in *Ramsey Theory*, established by Frank Ramsey in [21]. The purpose of Ramsey Theory is to bring order to what could seemingly be chaos. An example of a Ramsey Theory problem is: if there is a party where a group of three people either all know each other or all not know each other, how many people need to be invited to the party to make this happen?

A *monochromatic graph* is a graph where all edges are the same color and a *2 coloring* is when all edges of a graph are one of two possible colors. Ramsey Theory tells us that if there is a sufficiently large complete graph G, then in any edge coloring of G, a monochromatic complete subgraph is found.

The *Ramsey number*, denoted $R(s,t)$, is the smallest possible integer such that every two coloring of K_n has either the subgraph K_t or K_s colored blue or red, respectively. The Ramsey number has been studied extensively and the application of it to random graphs is quite interesting. For more information on the application of the Ramsey number to random graphs, the reader is referred to [18] where a method to find sharp threshold functions for Ramsey properties is given.

In this dissertation the threshold function and the sharp threshold function are discussed. However, a third variation exists which is called the *regular threshold function*. This threshold function is thought of as a boundary where the probability of a graph property asymptotically almost surely occurs on one side of the boundary, and asymptotically almost surely does not occur on the other side of the boundary. The sharp threshold function improves on the bound of the threshold function and more accurately shows where a graph property occurs or not occurs. Finally, the regular threshold function tells us the exact probability where a graph property occurs or does not occurs.

Thus, the end goal is to find a regular threshold function. Being more challenging to work with, regular threshold functions offer a lot to be studied. More information on regular threshold functions is found in [40].

It is established that the rainbow connection number is a monotone graph property. Threshold functions and sharp threshold functions then exist for all monotone graph properties. This is not necessarily the case for non-monotone graph properties. The existence of the sharp threshold function or the threshold functions for all non-monotone graph properties is still unknown and is an open problem.

One such property is the strong rainbow connection number of a graph. More work understanding the strong rainbow connection number of a graph and its relation to random graphs needs to be done. In [14] a number of results for the strong rainbow connection number are proved for some special graph classes such as bipartite graphs or complete graphs. However, precise results of the rainbow connection number for the general graph G offer opportunity for future work, especially in relation to random graphs.

In this dissertation the focus has been on the rainbow connection number of a graph with mention of the strong rainbow connection number of a graph. However, other connection numbers for rainbow graphs exist. An example of this is the (k, l)-*rainbow index*, denoted $rx_{k,l}$, where k and l are positive integers with $k \geq 3$. Define a *rainbow tree* as a tree where each edge on the tree contains a unique color. Then, $rx_{k,l}(G)$ is the minimum number of colors required to color graph G such that there exist l internally disjoint rainbow trees colored by a set of k colors. A good starting point for more information on this is found in [8] and [12]. These other rainbow connection numbers in relation to random graphs are a consideration for future work.

The open problem presented by Fujita et al in [22] is also to be considered for future work. The open problem looks at establishing sharp threshold results for random graph models other than the Binomial model. Two other random graph models already considered is the Uniform model (Theorem 5.4) and the Bipartite Random Graph model (Theorem 5.3). Other random graph models are still considered for this open problem. In particular, the consideration of a graph model for the random regular graph is open and is an excellent starting point for future work in this topic as it is a natural expansion of this dissertation. Note that regular graphs are not necessarily related to regular threshold functions.

Let A be the probability space for all k-regular graphs. Let $3 \leq k \leq n$ and let nk be an even integer. Then, the *random regular graph* G is a graph selected from A.

A proposed problem to be tackled for future work is stated below.

Proposed Problem 6.1.

Let $k \geq 1$*,* $d \geq 2$*, and* $3 \leq r \leq n$*. Further, let* G *be a* r*-regular random graph under the Uniform model. Lastly, let* nr *be an even integer. Then, can one find the sharp threshold function to satisfy the graph property* $rc_k(G) \leq d$*?*

As an alternative, we consider Theorem 3.4 that states if G is a connected graph with n vertices and $\delta = \delta(G)$, then

$$rc(G) \leq min\left\{\frac{nln\delta}{\delta}(1 + o_\delta(1)), \frac{4nln\delta}{\delta}\right\}.$$

These bounds are not optimal but have proved very difficult to improve. This creates room for improvement.

Bibliography

[1] Frieze A. and Karoński M. *Introduction to Random Graphs*. Cambridge University Press, 2016.

[2] Tang M. Priebe C. park Y. Vogelstein J. Levin K. Lyzinski V. Athreya A. Fishkind D. and Qin Y. "Statistical Inference on Random Dot Product Graphs: a Survey". In: *The Journal of Machine Learning Research* 18.1 (2017), pp. 8393–8484.

[3] Bollobas B. *Graph Theory*. Springer, 1979.

[4] Bollobas B. "The Diameter of Random Graphs". In: *Transactions of the American Mathematical Society* 267.1 (1981), pp. 41–52.

[5] Bollobas B. and Thomason A. "Threshold Functions". In: *Combinatorica* 7.1 (1987), pp. 35–38.

[6] Bollobs B. "Threshold Functions for Small Subgraphs". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 90. 2. Cambridge University Press. 1981, pp. 197–206.

[7] Rajendraprasad D. Basavaraju M. Chandran L. and Ramaswamy A. "Rainbow Connection Number and Radius". In: *Graphs and Combinatorics* 30.2 (2014), pp. 275–285.

[8] Li X. Cai Q. and Song J. "The (k,l) - Rainbow Index of Random Graphs". In: *Bulletin of the Malaysian Mathematical Sciences Society* 39.2 (2016), pp. 765–771.

[9] Roditty Y. Tuza Z. Caro Y. Lev A. and Yuster R. "On Rainbow Connection". In: *The Electronic Journal of Combinatorics* 15.1 (2008), R57.

[10] Matsliah A. Chakraborty S. Fischer E. and Yuster R. "Hardness and Algorithms for Rainbow Connection". In: *Journal of Combinatorial Optimization* 21 (Oct. 2008).

[11] Rajendraprasad D. Chandran L. Das A. and Varma N. "Rainbow Connection Number and Connected Dominating Sets". In: *Electronic Notes in Discrete Mathematics* 38 (2011), pp. 239–244.

[12] Magnant C. and Qin Z. Chang H. Li X. "The (k,l) - Proper Index of Graphs". In: *Journal of Combinatorial Optimization* 36.2 (2018), pp. 458–471.

[13] Lesniak L. Chartrand G. and Zhang P. *Graphs and Digraphs*. Sixth. Boca Raton: Chapman and Hall, 2015.

[14] McKeon K. Chartrand G. Johns G. and Zhang P. "Rainbow Connection in Graphs". In: *Mathematica bohemica* 133.1 (2008), pp. 85–98.

[15] McKeon K. Chartrand G. Johns G. and Zhang P. "The Rainbow Connectivity of a Graph". In: *Networks: An International Journal* 54.2 (2009), pp. 75–81.

[16] Li X. Chen X. and Lian H. "Rainbow k-connectivity of Random Bipartite Graphs". In: *Acta Mathematicae Applicatae Sinica, English Series* 36.4 (2020), pp. 879–890.

[17] Friedgut E. and Kalai G. "Every Monotone Graph Property Has a Sharp Threshold". In: *Proceedings of the American Mathematical Society* 124.10 (1996), pp. 2993–3002.

[18] Friedgut E. and Krivelevich M. "Sharp Thresholds for Certain Ramsey Properties of Random Graphs". In: *Random Structures & Algorithms* 17.1 (2000), pp. 1–19.

[19] Gilbert E. "Random Graphs". In: *The Annals of Mathematical Statistics* 30.4 (1959), pp. 1141–1144.

[20] Kaiser T. Koch M. Camacho S. Ryjacek Z. Ekstein J. Holub P. and Schiermeyer I. "The Rainbow Connection Number of 2-connected Graphs". In: *Discrete Mathematics* 313.19 (2013), pp. 1884–1892.

[21] Ramsey F. "On a Problem of Formal Logic". In: *Classic Papers in Combinatorics*. Springer, 2009, pp. 1–24.

[22] Liu H. Fujita S. and Magnant C. "Rainbow k-connection in Dense Graphs". In: *Electronic Notes in Discrete Mathematics* 38 (2011), pp. 361–366.

[23] L'Hôpital G. *Analyse des Infiniment Petits Pour l'intelligence des Lignes Courbes*. Chez Francois Montalant, 1715.

[24] Chernoff H. "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations". In: *The Annals of Mathematical Statistics* (1952), pp. 493–507.

[25] Jing H. and Liang H. "On rainbow-k-connectivity of Random Graphs". In: *Information Processing Letters* 112.10 (2012), pp. 406–410.

[26] Raftery A. Hoff P. and Handcock M. "Latent Space Approaches to Social Network Analysis". In: *Journal of the American Statistical Association* 97.460 (2002), pp. 1090–1098.

[27] Schiermeyer I. "Bounds for the Rainbow Connection Number of Graphs". In: *Discussiones Mathematicae Graph Theory* 31.2 (2011), pp. 387–395.

[28] Schiermeyer I. "Rainbow Connection in Graphs With Minimum Degree Three". In: *International Workshop on Combinatorial Algorithms*. Springer. 2009, pp. 432–437.

[29] Dong J. and Li X. "Rainbow Connection Number and Independence Number of a Graph". In: *Graphs and Combinatorics* 32.5 (2016), pp. 1829–1841.

[30] He J. and Liang H. "On Rainbow-k-Connectivity of Random Graphs". In: *Information Processing Letters* 112.10 (2012), pp. 406 –410.

[31] Menger K. "Zur Allgemeinen Kurventheorie". In: *Fundamenta Mathematicae* 10 (1927), pp. 96–1159.

[32] Chandran L. Mathew R. Li X. Liu S. and Rajendraprasad D. "Rainbow Connection Number and Connectivity". In: *The Electronic Journal of Combinatorics* (2012), P20–P20.

[33] Mestre A. Liu H. and Sousa T. "Total Rainbow k-connection in Graphs". In: *Discrete Applied Mathematics* 174 (2014), pp. 92–101.

[34] Krivelevich M. and Yuster R. "The Rainbow Connection of a Graph is (At Most) Reciprocal to Its Minimum Degree". In: *Journal of Graph Theory* 63.3 (2010), pp. 185–191.

[35] Erdős P. and Rényi A. "Asymmetric Graphs". In: *Acta Mathematica Hungarica* 14.295-315 (1963), p. 3.

[36] Erdős P. and Rényi A. "On Random Matrices II". In: *Studia Scientiarum Mathematicarum Hungar* 3 (1968), pp. 459–464.

[37] Erdős P. and Rényi A. "On the Existence of a Factor of Degree One of a Connected Random Graph". In: *Acta Mathematica Hungarica* 17 (1966), pp. 359–368.

[38] Erdős P. and Rényi A. "On the Strength of Connectedness of a Random Graph". In: *Acta Mathematica Hungarica* 12.1 (1961), pp. 261–267.

[39] Erdős P. and Rényi A. "On Random Graphs I". In: *Publicationes Mathematicae Debreceen* 6.1 (1959), pp. 290–297.

[40] Erdős P. and Rényi A. "On the Evolution of Random Graphs". In: *Publication of the Mathematical Institute of the Hungarian Academy of Science* 5.1 (1960), pp. 17–60.

[41] Flajolet P. and Sedgewick R. *Analytic Combinatorics*. Cambridge University Press, 2013.

[42] Ash R. *Basic Probability Theory*. Courier Corporation, 2008.

[43] Maartens R. "Balanced Graphs, Threshold Functions of Random Graphs and Their Co-occurrence, MSc (Mathematics)". In: *University of Johannesburg* (2012).

[44] Fenner T. and Frieze A. "On the Connectivity of Random m-orientable Graphs and Digraphs". In: *Combinatorica* 2.4 (1982), pp. 347–359.

[45] Ito T. Suzuki A. Uchizawa K. Aoki T. and Zhou Z. "On the Rainbow Connectivity of Graphs: Complexity and FPT Algorithms". In: *Algorithmica* 67.2 (2013), pp. 161–179.

[46] Li X. and Sun Y. "An Updated Survey on Rainbow Connections of Graphs - a Dynamic Survey". In: *Theory and Applications of Graphs* 1 (2017), p. 3.

[47] Shang Y. "A Sharp Threshold for Rainbow Connection of Random Bipartite Graphs". In: *Intrnational Journal of Applied Math* 24.1 (2011), pp. 149–153.