This is to certify that the

dissertation entitled

PROBABILISTIC THRESHOLD FOR COLLAPSIBILITY IN RANDOM GRAPHS

presented by

Joseph J. Spencer

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Mathematics

Major professor

Date 20 April 93

PROBABILISTIC THRESHOLD FOR COLLAPSIBILITY IN RANDOM GRAPHS

By

Joseph J. Spencer

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1993

ABSTRACT

PROBABILISTIC THRESHOLD FOR COLLAPSIBILITY

IN RANDOM GRAPHS

BY

Joseph J. Spencer

A graph G is *collapsible* if for every subset S of $V(G)$ with even cardinality, there is a connected spanning subgraph H of G whose vertices of odd degree are precisely the vertices of S. Collapsible graphs form an important class which satisfy the well-known double cycle conjecture. We have determined the random graph threshold function for this monotone property, i. e. we have found that a random graph with minimum degree two is almost surely collapsible. Our approach makes use of the theorem of Nash-Williams and Tutte which characterizes graphs with k edge-disjoint spanning trees. This method can also be used to estimate the minimum number of edges whose removal from a random graph leaves an eulerian subgraph.

DEDICATION

To my parents, Roy and Esther

ACKNOWLEDGEMENTS

There are many people that I need to thank. They have all in one way or another helped me through the years I have spent here at Michigan State. First and foremost I would like to thank my advisor, Professor Ed Palmer, for all the guidance and encouragement he has so patiently given me. I also would like to thank the members of my thesis committee, Professor Sagan, Professor Plotkin, Professor Rotthaus and Professor Esfahanian. Other faculty members that have had a positive influence on me are Professor R. Phillips, Professor Sledd, Professor S. Schuur, and Professor Weil. Next I would like to thank Barb Miller and Judi Miller who have done more than just perform their jobs efficiently, they have become my friends. Finally I would like to thank many of the graduate student who, over the years, have been my classmates and my friends, Jose Giraldo, Peter Springer, Kathy Dempsey, Rick Hensh, Rick Reynolds, Laura Reynolds, Bob Mckenzie, Rory Smith, Dawn Elzinga, Patti Bielak, Lisa Hansen, Steve Smith, Jack Jordan, Margret Readdy, Mirjana Jovovic, Joan Remski, John Clifford and anyone else that I have forgotten..

Contents

List of Tables

Introduction

The theory of random graphs was founded over thirty years ago by Paul Erdös and Alfred Rényi. In their fundamental papers [ErR59], [ErR60] and [ErR61] they brought the theory of probability into play with graph theory and opened up an area of investigation that has resulted in a thousand research papers, several books, two journals and many special conferences. It is now regarded as a major component of the field known as probabilistic combinatorics and has many important applications especially in the analysis of algorithms from computer science. One of their most interesting discoveries was the notion of the threshold function. The idea is that as a graph accumulates edges at random, certain properties occur rather abruptly. Erdös and Rényi found that they could predict quite accurately when such phenomena transpired. It is much the same as chemical phase transition that occurs when a solid such as ice is subjected to a gradual temperature increase and passes from solid to liquid and then to gas. In this thesis we will determine sharp thresholds for graph properties that arise in the study of the double cycle conjecture (DCC). See Jaeger [Ja85] for a survey of this topic.

The DCC constitutes one of the best-known unsolved problems in graph theory. It was first raised by G. Szekeres [Sz73] and asserts that if G is a graph with no bridges, then there is a multiset $\mathcal{C}$ of cycles in G such that each edge of G occurs in exactly two members of the family $\mathcal{C}$. The collapsible graphs originate in the work of Paul Catlin and are an important set of graphs which satisfy the DCC. So too are the supereulerian graphs. We first provide alternate definitions for collapsibility, relevant characterization theorems, descriptions of some families of collapsible graphs and a compulsory threshold. Then we introduce several probability models, discuss their relationships and give detailed proofs of important structural properties of random graphs. Next the decomposition theorem of Nash-Williams is presented. Together with our edge boundary lemma, it can be applied to obtain the threshold for edge-disjoint spanning trees and consequently the threshold for collapsibility. Finally we are able to use our techniques to estimate the size of large eulerian subgraphs in random graphs. During the course of our investigation several difficult unsolved problems arise. First there is the enumeration problem for collapsible graphs. Then we have the problem of determining an efficient algorithm for finding edge-disjoint spanning trees in random graphs with minimum degree two. Finally there is the question of the asymptotic probability of collapsibility for cubic graphs.

Here are some of the basic definitions we need from graph theory. Those not included may be found in the books [Bo85] and [Pa85]. A *graph* G consists of a finite set of *vertices*, $V(G)$, together with a set of *edges*, $E(G)$, which are unordered pairs of vertices. The cardinality of $V(G)$ is the *order* of G and the cardinality of $E(G)$ is the *size* of G. We will use the convention that $n = |V(G)|$. If $u, v \in V(G)$ and the edge $e = \{u, v\}$ we say u and v are *incident* with e or that e is *incident* with v. If u and v are both incident with the edge e, we say that u and v are *adjacent.* We also say that e *joins* u and v. The *degree* of a vertex v, denoted by $deg_G(v)$, is the number of edges that are incident with v in G. A *multigraph* is a graph in which any finite number of edges between two vertices is allowed.

The vertex neighborhood of a set $S \subseteq V(G)$ is defined by

$$N(S) = \{v \mid v \notin S \text{ is adjacent to some } u \in S\}.$$

For a singleton set, we will use $N(v) = N(\{v\})$. A *walk* in a graph G is a sequence of vertices $v_1, v_2, \ldots, v_m$ such that v_i is adjacent to v_{i+1}, $i = 1$ to $m-1$. If $v_1 = v_m$, the walk is *closed*. A *path* in G is a walk in which no vertex is repeated. The *length* of a path is number of edges in it. A graph G is *connected* if there is a path between every pair of vertices. The *distance* between two vertices is the length of a shortest path between them.

A *cycle* is a walk with at least three vertices in which the first and last vertices are the only ones repeated. A graph is *acyclic* if it contains no cycles. A *tree* is a connected, acyclic graph. A graph that is a path with n vertices and $n-1$ edges will be denoted by P_n. We denote a cycle of order n by C_n. A graph with n vertices and all $\binom{n}{2}$ possible edges is called the *complete graph* and is denoted by K_n. The *cartesian product* or *product* of the graphs G and H, $G \times H$, is the graph whose vertices are all ordered pairs (u, v), where $u \in V(G)$ and $v \in V(H)$. The edges of $G \times H$ join (u_1, v_1) and (u_2, v_2) if (i) $u_1 = u_2$ and v_1 is adjacent to v_2 or (ii) $v_1 = v_2$ and u_1 is adjacent to u_2. For $m \geq 1$ the *m-cube* Q_m is $K_2 \times \ldots \times K_2$, where the number of factors is m. We see that $Q_1 = P_1$ and $Q_2 = C_4$.

A *subgraph* H of G is a graph with $V(H) \subseteq V(G)$ and $E(H)$ is a subset of those edges in $E(G)$ that are incident with with only the vertices in $V(H)$. We will use the notation $H \subseteq G$ to mean that H is a subgraph of G and $V(H) = V(G)$. An *induced subgraph* H of G is a subgraph such that if $u, v \in V(H)$ and $\{u, v\}$ is and edge of G then $\{u, v\}$ is also and edge of H. For a set $X \subset V(G)$, X^* will represent the induced subgraph whose vertex set is X. A subgraph H of G *spans* G if $V(H) = V(G)$. A maximal, connected subgraph is called a *component*. An edge whose removal from a graph increases the number of components is called a *bridge*. A graph G is *eulerian* if there exists a closed spanning walk that includes each edge of G exactly once. A graph is *supereulerian* if it has a spanning eulerian subgraph.

Let $\mathcal{G}_n$ be the set of all $2^{\binom{n}{2}}$ labeled graphs of order n. A subset $\mathcal{Q}$ of $\mathcal{G}_n$ is a *graph property* if it is closed under isomorphism. A graph property $\mathcal{Q}$ is a *monotone graph property* or *monotone* if for any graph H that has $\mathcal{Q}$, all graphs G such that $H \subseteq G$ also have $\mathcal{Q}$.

We will often use a right arrow "$\rightarrow$" mean the limit as n goes to ∞.

Chapter 1

Collapsibility

1.1 Definitions

A graph G is *collapsible* (see [Ca88a]) if for every subset S of $V(G)$ with even cardinality, there is a connected spanning subgraph H of G in which S is the set of odd vertices in H. We shall refer to this as "definition one" for collapsibility. It is equivalent that a graph G is collapsible if for every subset S of $V(G)$ with even cardinality, there is a subgraph Γ such that $G - E(\Gamma)$ is connected and $v \in S$ if and only if the degree of v in Γ is odd. We call this "definition two". We will also consider K_1 to be collapsible. The cycle C_3 is an example of a collapsible graph while C_4 is not collapsible. More examples will follow later. In this section we will describe some of the properties of collapsible graphs.

For a graph G with a subgraph H, the *contraction* G/H is the multigraph obtained from G by identifying H with a single vertex v_H. Thus G/H has vertices $V(G/H) = [V(G) - V(H)] \cup \{v_H\}$ and $|E(G/H)| = |E(G)| - |E(H)|$. Note that all edges connecting $V(G) - V(H)$ to $V(H)$ are included in G/H. Multiple edges can arise when a vertex in $N(V(H))$ is adjacent to two or more vertices of H. We are now ready to state a very useful result for the identification of collapsible graphs.

Theorem 1.1.1 ([Ca88a]) *Let H be a collapsible subgraph of G. Then G is col-*

lapsible if and only if G/H is collapsible.

As a result of this theorem, we can sometimes determine whether or not a graph G is collapsible by performing a series of contractions. Thus G is collapsible if it can be reduced to a single vertex. But we also have the following important sufficient condition for a graph to be collapsible.

Theorem 1.1.2 ([Ca88a]) *If a graph has two edge-disjoint spanning trees then it is collapsible.*

Proof: This can be shown directly from "definition one" of collapsibility. Let G be a graph with two edge disjoint spanning trees and let S be an even ordered subset of $V(G)$. We wish to construct a spanning subgraph H where vertices of odd degree are precisely the vertices in S. There are four kinds of vertices depending on their parity in G and their desired parity in H. Let

$$A = \{v \mid deg_G(v) \text{ is even and } deg_H(v) \text{ is even}\},$$

$$B = \{v \mid deg_G(v) \text{ is even and } deg_H(v) \text{ is odd}\},$$

$$C = \{v \mid deg_G(v) \text{ is odd and } deg_H(v) \text{ is even}\}$$

and

$$D = \{v \mid deg_G(v) \text{ is odd and } deg_H(v) \text{ is odd}\}.$$

Note that $B \bigcup D = S$. The vertices in B and C must have their parities changed. Since $|B| + |D|$ and $|C| + |D|$ count the number of odd vertices in a graph, they are both even integers. It follows that $|B| + |C| + 2|D|$ is also an even integer. Thus $|B| + |C|$, the number of vertices that must have their parities different in H and G, is even. We arbitrarily pair up the vertices that must have their parity changed and consider the paths that connect them in one of the spanning trees. By removing the symmetric differences of these paths, we create a graph H in which the vertices of S are exactly those of odd degree. The existence of the second spanning tree of G guarantees that H is a connected subgraph. ■

We can see directly from the "definition one" that all collapsible graphs are supereulerian. One simply chooses S to be the empty set. Here is another interesting trait of collapsible graphs.

Theorem 1.1.3 ([Ca88a]) *Let G be a graph with a collapsible subgraph H. Then G is supereulerian if and only if G/H is supereulerian.*

This theorem shows how collapsibility can be used to help us identify supereulerian graphs. Supereulerian graphs are an important class of graphs which are known to satisfy the DCC.

1.2 Families of Collapsible Graphs

In this section we will give examples of collapsible graphs as well as examples of non-collapsible graphs. Certain families of collapsible graphs will be described.

Let G be any collapsible graph with $|V(G)| \geq 2$. Consider the cartesian product of G with any connected graph H. Since G is collapsible, we may contract each copy of G in $G \times H$ to a single vertex, resulting in a graph we shall call H_G. Clearly the graph H_G can be obtained from the graph H by replacing every single edge in H by $|V(G)|$ multiple edges. Thus we have $|V(H_G)| = |V(H)|$ and $|E(H_G)| = |E(H)||V(G)|$. Since two vertices with two or more edges joining them form a collapsible graph, we may, by a series of contractions, reduce H_G to a single vertex. Thus it is collapsible and so ultimately $G \times H$ is also collapsible.

We now need another reduction method which Catlin has described in [Ca88b]. Let G be a graph containing the induced four cycle C which has vertices w, x, y and z and edges $\{w,x\}$, $\{x,y\}$, $\{y,z\}$, and $\{z,w\}$. The reduction π is made by eliminating the edges of the cycle and then identifying the non-adjacent pairs of vertices and adding an edge between the two resulting new vertices. If v_1 is the identification of w and y and v_2 is the identification of x and z, we keep all edges not in C, including multiple edges that may arise, and add the edge $\{v_1, v_2\}$. We will call the reduced graph G/π.

Theorem 1.2.1 ([Ca88b]) *Let C be an induced four cycle in the graph G. If G/π is collapsible then G is collapsible.*

Now we may consider the m-cube Q_m. The graphs Q_1 and Q_2 are not collapsible. But we can show Q_3 is collapsible by performing two π reductions and then use contractions to obtain a single vertex. It can easily be shown by induction that Q_m for $m \geq 4$ is also collapsible.

Next let us next consider the cartesian product of paths, which we will call a *lattice*. We find that the lattices of the form $P_1 \times P_m$, $m \geq 1$, are not collapsible. These lattices have no spanning connected subgraphs with the four corners as the only vertices of odd degree. Also the lattice $P_2 \times P_2$ is not collapsible. If we take S to be a set of eight vertices, leaving out only one of the middle side vertices, we cannot find a connected spanning graph with the vertices in S having odd degree. It is easily shown by a series of π reductions that $P_2 \times P_3$ is collapsible. Using induction, it follows that all lattices $P_l \times P_m$ with $l \geq 2$ and $m \geq 3$ are collapsible. Here we have found two non-collapsible graphs whose cartesian product is collapsible.

Another family of collapsible graphs consists of the cartesian products of cycles, $C_l \times C_m$. If $l = 3$, then one of the cycles is a triangle, so the cartesian product is collapsible, as shown above. If both $l, m \geq 4$, then the lattice $P_{l-1} \times P_{m-1}$ is a collapsible spanning subgraph, so the graph $C_l \times C_m$ is collapsible. Using π reductions it is easily seen that the set of graphs of the form $K_2 \times C_n$ are also collapsible.

The well-known Petersen graph may described as a five pointed star inside a pentagon with each point of the star connected to the corresponding vertex of the pentagon. Here is a more precise description. Let the vertex set be

$$\{1, 2, 3, 4, 5, a, b, c, d, e\}$$

and the edge set be

$$\Big\{ \ \{1,2\}, \{2,3\}, \{3,4\}, \{4,5\}, \{5,1\}, \{a,b\}, \{b,c\}, \{c,d\},$$
$$\{d,e\}, \{e,a\}, \{1,a\}, \{2,d\}, \{3,b\}, \{4,e\}, \{5,c\} \ \Big\}.$$

This is the smallest bridgeless graph with minimum degree three that is not collapsible.

The Augmented Petersen graph is constructed from the Petersen graph by adding one vertex on the edge between 1 and 2 and one vertex on the edge between b and c and then connecting these two new vertices with an edge. According to Zhi-Hong Chen of Wayne State University (email communication) it is not known if this graph is collapsible.

1.3 Compulsory Threshold

Consider a monotone graph property Q. Suppose every graph with n vertices and at least M_n edges has property Q. Let us assume that M_n is the least such value. Then M_n is called a *compulsory threshold* for property Q. In this section we will find the compulsory threshold for collapsibility. First we need two lemmas.

Lemma 1.3.1 *Let Q be the graph property that every edge is in a triangle. Then the compulsory threshold for Q is*

$$\binom{n-1}{2}+2$$

for $n \geq 3$.

Proof: Consider the graph with vertex set $\{1,2,\ldots,n\}$ which consists of a complete graph K_{n-1} on the first $n-1$ vertices and also has the edge e joining vertex $n-1$ with vertex n. This graph has $\binom{n-1}{2}+1$ edges but e is not in a triangle. Therefore the compulsory threshold is at least $\binom{n-1}{2}+2$.

Consider a complete graph K_n, $n \geq 3$. Each edge of K_n belongs to $n-2$ triangles, so removal from K_n of any $n-3$ edges will result in a graph that has property Q. Since

$$\binom{n}{2}-(n-3)=\binom{n-1}{2}+2,$$

any graph with $\binom{n-1}{2}+2$ edges will have property Q. ■

Our second lemma follows from a theorem of Catlin.

Theorem 1.3.2 ([Ca88a]) *Let H_1 and H_2 be subgraphs of H such that $V(H) = V(H_1)\bigcup V(H_2)$, $E(H)=E(H_1)\bigcup E(H_2)$ and $V(H_1)\bigcap V(H_2)\neq\emptyset$. If H_1 and H_2 are collapsible, then so is H.*

Lemma 1.3.3 *If G is a connected graph in which every edge is in a collapsible subgraph of G, then G is collapsible.*

Proof: By repeatedly applying Theorem 1.3.2, we see that G is collapsible.■

Theorem 1.3.4 *The compulsory threshold for collapsibility for a graph with n vertices is*

$$\binom{n-1}{2}+2$$

for $n \geq 3$.

Proof: Consider the graph G of order n and size $\binom{n-1}{2}+1$ described in the proof of Lemma 1.3.1. Let H be the subgraph isomorphic to K_{n-1}. Then $G/H = K_2$. Since H is collapsible but K_2 is not, it follows from Theorem 1.1.1 that G is not collapsible.

We know from Lemma 1.3.1 that every edge in a graph with $\binom{n-1}{2}+2$ edges is in a triangle, which is a collapsible graph. So by Lemma 1.3.3 a graph with $\binom{n-1}{2}+2$ edges is always collapsible. ∎

The enumeration of collapsible and and supereulerian graphs constitutes an unsolved problem. For graphs with a small number of vertices we have generated the following numbers for the labeled and unlabeled cases.

Table 1.1: Numbers of Collapsible and Supereulerian Graphs

	Collapsible		Supereulerian	
n	Unlabeled	Labeled	Unlabeled	Labeled
1	1	1	1	1
2	0	0	0	0
3	1	1	1	1
4	2	7	3	10
5	9	231	10	243

Chapter 2

Random Graphs

2.1 Probability Models

In this section we will describe three probability models commonly used in the study of random graphs and give a simple illustration of how each model is used. They will be called Model A, Model B and Model C.

In Model A, the sample space consists of all labeled graphs with n vertices. Given a number $0 < p < 1$, the probability of a graph G with M edges is defined by

$$P(G) = p^M (1-p)^{N-M}$$

where $N = \binom{n}{2}$, the number of unordered pairs of vertices. The number p is the *edge probability*. Thus the sample space consists of Bernoulli trials and the edges are selected independently with probability p. Note that when $p = 1/2$, every graph with n vertices has equal probability. In Model A, the expected number of vertices of degree two is

$$n\binom{n-1}{2}p^2(1-p)^{n-3}.$$

In Model B, the sample space consists of all labeled graphs with n vertices and M edges. For a given value M, $0 \leq M \leq N$, the probability of a graph G is defined by

$$P(G) = \frac{1}{\binom{N}{M}}.$$

In Model B the expected number of isolated vertices in a graph is

$$\frac{n\binom{\binom{n-1}{2}}{M}}{\binom{\binom{n}{2}}{M}}.$$

In Model C the sample space for a given n and r, $1 \leq r < n$, consists of all digraphs with every vertex having outdegree r. This model is sometimes called "r–out". Each digraph D has probability

$$P(D) = \frac{1}{\binom{n-1}{r}^n}.$$

For a digraph in Model C, the expected number of vertices with indegree zero is

$$\frac{n\binom{n-2}{r}^{n-1}}{\binom{n-1}{r}^n}.$$

By ignoring the orientation of the edges and consolidating multiple edges another sample space consisting of graphs is created.

In the study of random graphs, we can conclude nothing about any one graph, what we do study are properties of sets of graphs. If Q is some property of graphs and $\mathcal{A}$ is the set of graphs of order n with property Q, and $P(\mathcal{A}) \to 1$ as $n \to \infty$, then we say *almost all graphs have property Q* or *the random graph has property Q almost surely (a. s.)*. We are studying a sequence of sample spaces and the limit of a sequence of probabilities.

An important discovery of Erdös is the notion of the *threshold function* for certain properties of random graphs. In Model A we express the edge probability p as a function of the number of vertices n, i. e. $p = p(n)$. The function $p^*(n)$ is a threshold function for a graph property Q if

$$p(n)/p^*(n) \to 0 \text{ implies that almost no } G \text{ has } Q$$

and

$$p(n)/p^*(n) \to \infty \text{ implies that almost every } G \text{ has } Q.$$

We may similarly define threshold functions in Model B by considering $M = M(n)$. It has been observed by Bollobás and Thomason that every monotone graph property has a threshold function [BoT86].

The first moment method is an important tool from probability theory that is used frequently in the study of random graphs. Suppose X is a nonnegative integer valued random variable, then $E[X] \geq P(X \geq 1)$. Thus if $E[X] \to 0$, then $P(X \geq 1) \to 0$ and therefore $P(X = 0) \to 1$. So, for example, if

$$X = \text{ the number of vertices of degree 0 in a random graph}$$

and

$$p = 2\frac{\log n}{n},$$

then

$$\begin{aligned} E[X] &= n(1-p)^{n-1} && (2.1) \\ &\leq n\exp(-p(n-1)) && (2.2) \\ &\leq n\exp(-2\log n)\exp(p) && (2.3) \\ &\leq n(n^{-2})\exp(p) \to 0. && (2.4) \end{aligned}$$

Thus $P(X = 0) \to 0$ and we may conclude that if $p = 2\frac{\log n}{n}$, then a random graph in Model A almost surely has no vertices of degree zero.

2.2 Degree Distribution

Having minimum degree at least k is a monotone graph property. It has a well known sharp threshold function. If

$$p = \frac{\log n + (k-1)\log\log n + \omega_n}{n}, \tag{2.5}$$

where $\omega_n \to \infty$ and $\omega_n = o(\log\log n)$, then a random graph almost surely has minimum degree k.

Let us look at equation (2.5) with $k = 2$ and see how each term affects the degree of a random graph. Let X_i be the random variable that counts the number of vertices of degree i in a random graph. Then the expected number of vertices of degree two is given by

$$E[X_2] = n\binom{n-2}{2}p^2(1-p)^{n-2}. \tag{2.6}$$

First let us look at $p = \frac{\log n}{n}$. We have

$$E[X_2] = \left(1+o(1)\right)\frac{n^3}{2}\left(\frac{\log n}{n}\right)^2 e^{-pn} \tag{2.7}$$

$$= \left(1+o(1)\right)\frac{n(\log n)^2}{2}\frac{1}{n} \to \infty. \tag{2.8}$$

But this is not the threshold for minimum degree two because the expected number of vertices with degree one is

$$E[X_1] = n\binom{n-1}{1}p(1-p)^{n-1} \tag{2.9}$$

$$= \left(1+o(1)\right)n^2\frac{\log n}{n}\frac{1}{n} \to \infty. \tag{2.10}$$

In fact we find that $E[X_i] \to \infty$ for all $i \geq 1$.

If

$$p = \frac{\log n + \log\log n}{n}$$

then

$$E[X_2] = \left(1+o(1)\right)\frac{n(\log n)^2}{2}\frac{1}{n\log n} \tag{2.11}$$

$$= \left(1+o(1)\right)\frac{\log n}{2} \to \infty. \tag{2.12}$$

but

$$E[X_1] = \left(1 + o(1)\right)n^2 \frac{\log n}{n} \frac{1}{n \log n} \not\to \infty.$$

Adding the factor ω_n with $\omega_n \to \infty$ and $\omega_n = o(\log\log n)$ forces $E[X_1] \to 0$ while not interfering with the behavoir of $E[X_2]$.

2.3 Relations Between Models

Usually we find that it is more straightforward to work with Model A than Model B. Fortunately these two models are very closely related and it is often easy to convert results from one to the other. In this section a random graph in Model A with n vertices and edge probability p will be represented by $G_{n,p}$ and its set of edges by $E_{n,p}$. A random graph in Model B with n vertices and M edges will be represented by $G_{n,M}$. A graph property Q is *convex* if $G_1 \subseteq G_2$ and G_1 and G_2 both have Q implies that any H such that $G_1 \subseteq H \subseteq G_2$ also has Q. For a convex graph property we have the following well-known conversion theorem.

Theorem 2.3.1 ([Bo79]) *Let $0 < p = p(n) < 1$ be such that $pn^2 \to \infty$ and $(1-p)n^2 \to \infty$ as $n \to \infty$ and let Q be a property of graphs.*

(i) Suppose $\epsilon > 0$ is fixed and, if $(1-\epsilon)pN < M < (1+\epsilon)pN$, then $G_{n,M}$ has Q almost surely. Then $G_{n,p}$ has Q almost surely.

(ii) If Q is a convex property and $G_{n,p}$ has Q almost surely then $G_{n,\lfloor pN \rfloor}$ has Q almost surely.

For a graph property Q that is not convex we have the following method of conversion from Model A to Model B that is presented in [BoFF87]. Note that at times we will use non-integral quantities that ought to be rounded up or down. It should be clear that this practice will not affect the results.

Lemma 2.3.2 ([BoFF87]) *Let Q be any graph property and suppose $M = pN$. If*

$$nP(G_{n,p} \text{ has } Q) \to 0 \text{ as } n \to \infty$$

then

$$P(G_{n,M} \text{ has } Q) \to 0 \text{ as } n \to \infty.$$

Proof: First note that

$$P(|E_{n,p}| = M) = \binom{N}{M} p^M (1-p)^{N-M}. \tag{2.13}$$

Now using Stirling's formula and the fact that $M = pN$, we can show that

$$P(|E_{n,p}| = M) \geq \left(1 - o(1)\right) \frac{1}{\sqrt{2\pi N}}. \tag{2.14}$$

Also, using equation (2.13) we find that

$$\begin{aligned} P(G_{n,p} \text{ has } Q \mid |E_{n,p}| = M) &= \frac{P(G_{n,p} \text{ has } Q \text{ and } |E_{n,p}| = M)}{P(|E_{n,p}| = M)} && (2.15) \\ &= \frac{|\{G_{n,M} \text{ has } Q\}| p^M (1-p)^{N-M}}{\binom{N}{M} p^M (1-p)^{N-M}} && (2.16) \\ &= P(G_{n,M} \text{ has } Q). && (2.17) \end{aligned}$$

Now clearly it follows from the definition of Model A and from (2.17) that

$$\begin{aligned} P(G_{n,p} \text{ has } Q) &= \sum_{M'} P(G_{n,p} \text{ has } Q \mid |E_{n,p}| = M') P(|E_{n,p}| = M') && (2.18) \\ &= \sum_{M'} P(G_{n,M'} \text{ has } Q) P(|E_{n,p}| = M'). && (2.19) \end{aligned}$$

By using a single summand of the right hand side of (2.19), say when $M' = M$, we see

$$P(G_{n,p} \text{ has } Q) \geq P(G_{n,M} \text{ has } Q) P(|E_{n,p}| = M) \tag{2.20}$$

But now with equation (2.14) and using that N is asymptotic to $n^2/2$,

$$P(G_{n,p} \text{ has } Q) \geq (1 - o(1))\sqrt{\frac{1}{\pi n^2}}P(G_{n,M} \text{ has } Q). \tag{2.21}$$

Thus

$$nP(G_{n,p} \text{ has } Q) \geq (1 - o(1))\sqrt{\frac{1}{\pi}}P(G_{n,M} \text{ has } Q)$$

and our result is proven. ∎

2.4 Structural Properties

We are interested in random graphs with edge probability p roughly equal to $\log n/n$. The average degree of the vertices of such a graph is about $\log n$. We need to know that vertices of low degree are not too close to each other in a random graph and that there are not too many of these. These structural properties will be needed to help us show that in random graphs, small sets of vertices have large edge boundaries and to show the existence of edge-disjoint trees in random graphs. We let $d = pn$, and we say that v is *small* its degree is less than $\frac{d}{20}$, otherwise we say v is *large*. In [BoFF87] the authors gave only a sketch of the proof of the next theorem, but here we shall include more details. As in [BoFF87] the proof will be given for both Models A and B.

Theorem 2.4.1 ([BoFF87]) *In Model A with edge probability*

$$p \geq \frac{\log n}{n} \tag{2.22}$$

or in Model B with

$$M \geq \frac{1}{2}n \log n \tag{2.23}$$

the following statements for a random graph G hold almost surely:

(a) G contains no more than $n^{\frac{1}{3}}$ small vertices.

(b) G does not contain two small vertices at a distance of four or less apart.

Proof: (a) Let $\mathcal{Q}_a$ be the property that a random graph G has more than $n^{\frac{1}{3}}$ small vertices. We would like to show that almost all graphs do not have this property. Let X count the number of sets containing $n^{\frac{1}{3}}$ small vertices. We know from the first moment method that

$$P(G_{n,p} \text{ has } \mathcal{Q}_a) \leq E[X]. \tag{2.24}$$

For each vertex v in a set S of order s consisting of only small vertices, the number of edges joining v to vertices not in S will be $\binom{n-s}{i}$ for $0 \leq i \leq \frac{d}{20}$. Each of the edges occurs with probability p and each non-edge with probability $1-p$. So the probability that a given S consists of small vertices is bounded by

$$\left[\sum_{i=0}^{\frac{d}{20}} \binom{n-s}{i} p^i (1-p)^{n-s-i}\right]^s. \tag{2.25}$$

So the expected number of sets S of small vertices with $|S| = s$ will be

$$E[X] \leq \binom{n}{s} \left[\sum_{i=0}^{\frac{d}{20}} \binom{n-s}{i} p^i (1-p)^{n-s-i}\right]^s. \tag{2.26}$$

Now using the facts that $|S| = n^{\frac{1}{3}}$ and $p \geq \frac{\log n}{n}$ along with the bound for the tail of a binomial distribution (see [Pa85] page 133) we see

$$E[X] \leq \binom{n}{s} \left[\binom{n-s}{d/20} p^{\frac{d}{20}} (1-p)^{n-s-\frac{d}{20}} \frac{p(n-s-\frac{d}{20}+1)}{p(n-s+1)-\frac{d}{20}}\right]^s \tag{2.27}$$

$$\leq (\frac{en}{s})^s \left[\mathcal{O}(1)(20e\frac{n-s}{d}p)^{d/20} \exp(-pn+ps+\frac{dp}{20})\right]^s \tag{2.28}$$

$$\leq \left[\mathcal{O}(1)\frac{n}{s} \exp(\frac{4}{20}d) \exp(-d+ps+dp/20)\right]^s \tag{2.29}$$

$$\leq \left[\mathcal{O}(1)n^{\frac{2}{3}}\exp(-\frac{15}{20}d)\right]^{n^{\frac{1}{3}}} \tag{2.30}$$

$$\leq \left[\mathcal{O}(1)n^{\frac{2}{3}}\exp(-2d/3)\exp(-5d/60)\right]^{n^{\frac{1}{3}}} \tag{2.31}$$

$$\leq \mathcal{O}(1)\exp(-\frac{1}{12}dn^{\frac{1}{3}}). \tag{2.32}$$

Because $d = np \geq \log n$, clearly $\exp(-\frac{1}{12}dn^{\frac{1}{3}}) \to 0$ as $n \to 0$. Thus $P(G_{n,p}$ has $\mathcal{Q}_a) \to 0$ and so almost surely $G_{n,p}$ has fewer than $n^{\frac{1}{3}}$ small vertices.

This result is easily converted to Model B as follows. Multiply both sides of inequality (2.32) by n and we see that $n\exp(-\frac{1}{12}dn^{\frac{1}{3}}) \to 0$ as $n \to \infty$. It follows from Lemma 2.3.2 that almost surely $\mathcal{Q}_a$ does not hold in Model B.

(b) Let $\mathcal{Q}_b$ be the property that there exist two small vertices at a distance of four of less apart. We would like to show that this property does not hold almost surely. Let X count the number of pairs of small vertices at a distance of four or less apart. Then

$$P(G_{n,p} \text{ has } \mathcal{Q}_b) \leq E[X]. \tag{2.33}$$

The expected value of X is bounded by the expected number sets of two small vertices multiplied by the probability that the two vertices have a path of length four or less between them. So we multiply inequality (2.26) by $n^3p^4 + n^2p^3 + np^2 + p$ to obtain

$$E[X] \leq \binom{n}{2}\left[\sum_{i=0}^{d/20}\binom{n-2}{i}p^i(1-p)^{n-i-2}\right]^2\left(n^3p^4 + n^2p^3 + np^2 + p\right). \tag{2.34}$$

Using the same estimates as in part (a) for $s = 2$, we find

$$P(G_{n,p} \text{ has } \mathcal{Q}_b) \leq n^5p^4e^{-1.5d}. \tag{2.35}$$

Since $d \geq \log n$ we see that $P(G_{n,p}$ has $\mathcal{Q}_b) \to 0$. Therefore in Model A we know that almost surely there are not two small vertices at a distance of four or less apart.

Now for Model B. If $p \geq \frac{2\log n}{n}$ then $d \geq 2\log n$. Multipling inequality (2.35) by n we find that

$$n(n^5 p^4 e^{-1.5d}) \leq n^6 p^4 n^{-3} \to 0 \tag{2.36}$$

and thus by Lemma 2.3.2 we know $P(G_{n,M}$ has $Q_b) \to 0$. This gives us our result for Model B if $M \geq \lfloor n \log n \rfloor$.

For smaller M, corresponding to the range $\frac{\log n}{n} \leq p \leq 2\frac{\log n}{n}$, more work is required. First let $M = \lfloor pN \rfloor$. Now we will show that there exists an M', with $M - (n\log^3 n)^{\frac{1}{2}} \leq M' \leq M$ such that

$$P(G_{n,M'} \text{ has } Q_b) \leq 3n^5 p^4 e^{-1.5d}. \tag{2.37}$$

Suppose that no such M' exists. Then using equation (2.19) we have

$$\begin{aligned} P(G_{n,p} \text{ has } Q_b) &\geq \sum_{i=M-M'}^{M} P(G_{n,M'} \text{ has } Q_b) P(|E_{n,p}| = M') && (2.38) \\ &\geq \sum_{i=M-M'}^{M} (3n^5 p^4 e^{-1.5d}) \binom{N}{2} p^i (1-p)^{N-i}. && (2.39) \end{aligned}$$

Then using inequalities (2.35) and (2.39) we see

$$\begin{aligned} n^5 p^4 e^{-1.5d} &\geq P(G_{n,p} \text{ has } Q_b) && (2.40) \\ &\geq \sum_{i=M-M'}^{M} (3n^5 p^4 e^{-1.5d}) \binom{N}{2} p^i (1-p)^{N-i}. && (2.41) \end{aligned}$$

So we have

$$\frac{1}{3} \geq \sum_{i=M-M'}^{M} \binom{N}{2} p^i (1-p)^{N-i}. \tag{2.42}$$

It follows from the normal approximation to the binomial distribution that the right hand side of inequality (2.42) approaches $\frac{1}{2}$ as $n \to \infty$, so we have a contradiction. Hence there exists an M', with $M - (n\log^3 n)^{\frac{1}{2}} \leq M' \leq M$ such that

$$P(G_{n,M'} \text{ has } Q_b) \leq 3n^5 p^4 e^{-1.5d}. \tag{2.43}$$

Now let

$$P_1 = P(G_{n,M'} \text{ has } Q_b \text{ and } G_{n,M} \text{ has } Q_b)$$

and

$$P_2 = P(G_{n,M'} \text{ does not have } Q_b \text{ and } G_{n,M} \text{ has } Q_b).$$

Clearly

$$P(G_{n,M} \text{ has } Q_b) = P_1 + P_2. \tag{2.44}$$

Next we show that both $P_1 \to 0$ and $P_2 \to 0$ as $n \to \infty$.

First we see

$$\begin{aligned} P_1 &= P(G_{n,M'} \text{ has } Q_b \text{ and } G_{n,M} \text{ has } Q_b) \\ &\leq P(G_{n,M'} \text{ has } Q_b) \\ &\leq 3n^5 p^4 e^{-1.5d} \to 0. \end{aligned}$$

Now we focus on P_2. Let π be the probability that $G_{n,M'}$ does not have Q_b but $G_{n,M}$ obtained from $G_{n,M'}$ by adding $M - M'$ edges has at least one of its new edges incident with a small vertex or a neighbor of a small vertex. Clearly $P_2 \leq \pi$. To show that $P_2 \to 0$, we will show that $1 - \pi \to 1$. Now $1 - \pi$ is the probability that $G_{n,M'}$ does not have Q_b and $G_{n,M}$ obtained from $G_{n,M'}$ by adding $M - M'$ edges has none of its new edges incident with a small vertex or a neighbor of a small vertex. The number of positions where these new edges may be placed is at least $N - M' - n^{\frac{4}{3}}(1 + \frac{d}{20})$. To see this, observe that the number of small vertices and their neighbors is at most $n^{\frac{1}{3}} + n^{\frac{1}{3}}\frac{d}{20}$. There could be at most $n - 1$ edges incident with each of these. And so $n(n^{\frac{1}{3}} + n^{\frac{1}{3}}\frac{d}{20})$ is an upper bound on the number of possible edges that are incident with a small vertex or a neighbor of a small vertex. Thus

$$(1-\pi) \geq \frac{\binom{N-M'-n^{\frac{4}{3}}\left(1+\frac{d}{20}\right)}{M-M'}}{\binom{N-M'}{M-M'}}. \tag{2.45}$$

The right side of (2.45) is of the form

$$\frac{\binom{a}{k}}{\binom{b}{k}}. \tag{2.46}$$

The behavior of such an expression was considered in [Pa85]. It is shown that if

$$k^2\frac{a-b}{b^2} \to 0 \tag{2.47}$$

and

$$k\left(\frac{b-a}{b-k}\right)^2 \to 0 \tag{2.48}$$

then

$$\frac{\binom{a}{k}}{\binom{b}{k}} = \left(1+(1)\right)\exp(-k\frac{b-a}{b}). \tag{2.49}$$

Now since $M-M' \leq (n\log^3 n)^{\frac{1}{2}}$ and $n^{\frac{4}{3}}(1+\frac{d}{20}) = \mathcal{O}(n^{\frac{4}{3}}\log n)$, we see that conditions (2.47) and (2.48) hold and that the binomial expression in (2.45) is asymptotic to 1. Therefore $\pi \to 0$ and then so must P_2. This shows that both $P_1 \to 0$ and $P_2 \to 0$ and it follows from equation (2.44) that

$$P(G_{n,M} \text{ has } \mathcal{Q}_b) \to 0.$$

Hence property $\mathcal{Q}_b$ almost surely does not hold for Model B when $M = pN \geq \left\lfloor\frac{1}{2}n\log n\right\rfloor$. That is, almost surely the distance between any two small vertices is at least five. ∎.

Chapter 3

Edge-Disjoint Trees

3.1 The Theorem of Nash-Williams

We will make use of a beautiful theorem found independently by Nash-Williams and Tutte, which provides necessary and sufficient conditions for a graph to have k edge-disjoint spanning trees. We use $\mathcal{P}$ to denote a partition of the vertex set of a graph and $|\mathcal{P}|$ is the number of parts in $\mathcal{P}$. For a graph G with partition $\mathcal{P}$, contracting each part of $\mathcal{P}$ to a single vertex leaves the graph which we denote as $G_{\mathcal{P}}$. We will call the edges of $G_{\mathcal{P}}$ *external edges* and denote them by either $E(G_{\mathcal{P}})$ or $E_{\mathcal{P}}(G)$.

Theorem 3.1.1 ([N-W61] , [Tu61]) *A graph G has k edge-disjoint spanning trees if and only if*

$$|E_{\mathcal{P}}(G)| \geq k(|\mathcal{P}| - 1) \tag{3.1}$$

for every partition $\mathcal{P}$ of $V(G)$.

The necessity is easy to show. Let G be a graph with k edge-disjoint spanning trees T_1, T_2, $\ldots T_k$ and let $\mathcal{P}$ be a partition of $V(G)$. Then each $(T_i)_{\mathcal{P}}$ is connected and so it has at least $|\mathcal{P}| - 1$ edges. Hence

$$|E_{\mathcal{P}}(G)| \geq \sum_{i=1}^{k} |E_{\mathcal{P}}(T_i)| \geq k(|\mathcal{P}| - 1) \tag{3.2}$$

as required.

The proof of the sufficiency is more complicated. We will provide an algorithmic proof for the case when $k = 2$ that is implicit in [N-W61]. The algorithm will explicitly find two edge-disjoint spanning trees in any graph that satisfies condition (3.1). We will prove results for arbitrary k when it is no more complicated than the case in which $k = 2$.

We need the following definitions. In a graph G for any non-empty subset X of $V(G)$ let $\Delta_G(X) = k(|X| - 1) - |E(X^*)|$. The set X is called *critical* if $\Delta_G(X) = 0$. A partition $\mathcal{P}$ of $V(G)$ is *admissible* if it satisfies inequality (3.1) and *critical* if the inequality can be replaced with an equality. A graph G is called *admissible* if all partitions of $V(G)$ are admissible and *critical* if $|E(G)| = k(|V(G)| - 1)$

The algorithm is recursive. We wish to reduce a graph to one of the base cases when $|V(G)| = 1, 2$ or 3.

Start the algorithm by considering a graph G. Every partition of $V(G)$ must be examined to make sure G is admissible and to find which partitions are critical. Checking every partition will mean that this algorithm cannot be done in polynomial time. Every graph returned by the algorithm will be admissible. If $G = K_1$, then we will consider G to be admissible and to have two edge-disjoint spanning trees. If G has two or three vertices and is admissible, it will be clear that there are two edge-disjoint spanning trees and they will be easily found.

For $|V(G)| \geq 4$, every graph G falls into one of three classes:

(I) The only critical partition is $\mathcal{P} = \{V(G)\}$, the trivial partition (which is always critical).

(II) There is a critical partition that is neither the trivial partition, $\mathcal{P} = \{V(G)\}$, nor the singleton partition, $\mathcal{P} = V(G)$.

(III) The singleton partition and the trivial partition are the only critical partitions.

For each of the three classes the algorithm will send a graph or a set of graphs that have fewer edges and/or fewer vertices than the original graph back to the start for treatment.

Class I graphs:

In this case, for every partition $\mathcal{P} \neq \{V(G)\}$

$$|E_{\mathcal{P}}(G)| > k(|\mathcal{P}| - 1).$$

So we may pick any $e \in E(G)$ and remove it, creating an new graph $H = G - \{e\}$. Now every partition $\mathcal{P}$ of $V(H)$ has

$$|E_{\mathcal{P}}(H)| \geq k(|\mathcal{P}| - 1)$$

and thus H is admissible and may be returned to the beginning of the algorithm.

Class II graphs:

Here there is a critical partition $\mathcal{P}$ of $V(G)$ other than the singleton or trivial partition, so

$$|E_{\mathcal{P}}(G)| = k(|\mathcal{P}| - 1)$$

with $|\mathcal{P}| \neq 1$ and $|\mathcal{P}| \neq |V(G)|$. Let $\mathcal{P} = \{X_1, X_2, \ldots, X_t\}$. For any $X_i \in \mathcal{P}$, let $\mathcal{Q}$ be a partition of X_i and $\mathcal{R} = (\mathcal{P} - \{X_i\}) \bigcup \mathcal{Q}$. Then since G is admissible,

$$k(|\mathcal{R}| - 1) \leq |E_{\mathcal{R}}(G)| = |E_{\mathcal{P}}(G)| + |E_{\mathcal{Q}}(X_i^*)| \tag{3.3}$$

and so

$$\begin{aligned} |E_{\mathcal{Q}}(X_i^*)| &\geq k(|\mathcal{R}| - 1) - |E_{\mathcal{P}}(G)| = && (3.4) \\ & k(|\mathcal{R}| - 1) - k(|\mathcal{P}| - 1) = && (3.5) \\ & k(|\mathcal{R}| - |\mathcal{P}|) = k(|\mathcal{Q}| - 1). && (3.6) \end{aligned}$$

Thus we see that X_i^* is admissible since Q is arbitrary and that every X_i^* is admissible since i is arbitrary.

Now consider the graph $G_{\mathcal{P}}$. Any partition $\mathcal{S}$ of $V(G_{\mathcal{P}})$ is also a partition of $V(G)$. Since G is admissible, $G_{\mathcal{P}}$ is also admissible. We can then return $|\mathcal{P}|+1$ graphs, $X_1^*, X_2^*, \ldots, X_t^*$ and $G_{\mathcal{P}}$, to the beginning of the algorithm. In each of these graphs, 2 edge-disjoint spanning trees will be constructed and then put together to form 2 edge-disjoint spanning trees on G.

Class III graphs:

We need the following lemmas.

Lemma 3.1.2 ([N-W61]) *Let G be a critical graph. Then G is admissible if and only if $\Delta_G(X) \geq 0$ for every $\emptyset \neq X \subset V(G)$.*

Proof: Let us start by showing the necessity. If G is admissible, let $X \subset V(G)$ and the complement of X be $\overline{X} = \{v_1, v_2, \ldots, v_r\}$. Also let $\mathcal{P} = \{X, \{v_1\}, \{v_2\}, \ldots, \{v_r\}\}$. Then

$$\begin{aligned}
\Delta_G(X) &= k(|X|-1) - |E(X^*)| && (3.7)\\
&= k(|V(G)| - 1 - |\mathcal{P}|) - (|E(G)| - |E_{\mathcal{P}}(G)|) && (3.8)\\
&= k|V(G)| - |E(G)| - k|\mathcal{P}| + |E_{\mathcal{P}}(G)| && (3.9)\\
&= [k(|V(G)|-1) - |E(G)|] + [E_{\mathcal{P}}(G) - k(|\mathcal{P}|-1)] && (3.10)\\
&\geq 0 && (3.11)
\end{aligned}$$

using that G is critical and admissible. Now let us show the sufficiency. We assume $\Delta_G(X) \geq 0$ for all $X \subset V(G)$. For any partition $\mathcal{P}$ of $V(G)$,

$$\begin{aligned}
|E_{\mathcal{P}}(G)| &= |E(G)| - \sum_{X_i \in \mathcal{P}} |E(X_i)| && (3.12)\\
&\geq k(|V(G)|-1) - \sum_{X_i \in \mathcal{P}} k(|X_i|-1) = k(|\mathcal{P}|-1) && (3.13)
\end{aligned}$$

so G is admissible. This ends the proof of the lemma.

We also need some results about critical graphs with $|V(G)| \geq 4$ and $k = 2$ that are summarized in the following lemma.

Lemma 3.1.3 *Suppose a graph G is critical with $|V(G)| \geq 4$ and $\Delta_G(X) \geq 0$ for every $\emptyset \neq X \subset V(G)$. For $k = 2$ the following hold:*

(a) There are no vertices of degree two.

(b) There exists a vertex of degree three.

(c) For any vertex v of degree three, $|N(v)| > 1$.

(d) If a vertex v has degree three and X is a critical set not containing v, then $|N(v) \cap X| \leq 1$

Proof: (a) If v is a vertex of degree 2, then the partition $\{\{v\}, V(G) - \{v\}\}$ is critical. But we are assuming that G has only the singleton and trivial partitions as critical partitions.

(b) The average degree of the vertices in G is

$$\frac{2|E(G)|}{|V(G)|} = \frac{2(2(|V(G)| - 1))}{|V(G)|} = 4 - \frac{4}{|V(G)|} < 4.$$

Since there are no vertices of degree 2, then a vertex of degree 3 must exist.

(c) If v has degree three and $N(v) = \{u\}$, a single vertex then

$$\Delta_G(\{v, u\}) = 2(2 - 1) - 3 = -1$$

which contradicts Lemma 3.1.2.

(d) First suppose the vertex v has degree three and that $N(v) \subset X$. Then since $\Delta_G(X) = 0$, we see that $\Delta_G(X \cup \{v\}) = 0$ which contradicts Lemma 3.1.2. So one of v's neighbors is not contained in X.

Next suppose X contains two neighbors of v. Let $\overline{X} = \{u_1, u_2, \ldots, u_r\}$ and consider the partition

$$\mathcal{P} = \big\{X \cup \{v\}, \{u_1\}, \ldots, \{u_r\}\big\}.$$

Using that G is a critical graph, X is a critical set $|\mathcal{P}| = |\overline{X}| + 1$, we see

$$\begin{aligned} |E_{\mathcal{P}}(G)| &= |E(G)| - |E(X^*)| - 2 && (3.14)\\ &= 2(|X| + |\overline{X}| + 1 - 1) - 2(|X| - 1) - 2 && (3.15)\\ &= 2(|\overline{X}| + 1 - 1) && (3.16)\\ &= 2(|\mathcal{P}| - 1) && (3.17)\end{aligned}$$

But this contradicts the assumption that G only has the singleton and trivial partitions as critical partitions. This ends the proof of the second lemma.

Let v be a vertex with degree three and with neighbors u and w. Let H be the graph constructed from G by adding an edge e between u and w while deleting v and the edges incident to v. Clearly H is critical and we must show that H is admissible. For any set $X \subset V(H)$, if u or w are not in X, then

$$\Delta_H(X) = \Delta_G(X) \geq 0.$$

If both u and w are in X, then because X cannot be critical and contain both u and w we know $\Delta_G(X) > 0$. From this we see that $\Delta_H(X) = \Delta_G(X) - 1 \geq 0$. Thus by Lemma 3.1.2 we know that H is admissible. We then may return H to the beginning of the algorithm.

From H we construct two edge-disjoint spanning trees in G as follows: When H returns it will have two edge-disjoint spanning trees and the edge $e = \{u, w\}$ will be in one of them. For that tree, replace e with the edges $\{v, u\}$ and $\{v, w\}$, and add the third edge incident to v to the other tree.

Thus by reducing the graph G to one of the base cases we may construct two edge disjoint trees on G. ■

The computational complexity of this algorithm is, of course, exponential because every partition of the vertex set must be examined. On the other hand, the proof seems easier to follow and more convincing than traditional proofs when presented in this style. Furthermore it is apparent that the traditional proofs do not lend themselves easily to adaptation for efficient algorithms. Dave Johnson has told us that finding edge-disjoint spanning trees can actually be achieved in polynomial time (email communication). The best known treatment makes use of the matroid greedy algorithm. See [RoT85] for related references.

3.2 Edge Boundary Lemma

To implement the theorem of Nash-Williams, we must have some knowledge of the edge boundary of a set of vertices. We need to show that a vertex set of small order will have a large number of edges joining it to the rest of the graph. For a subset S of $V(G)$, let us define the *edge boundary of* S to be

$$\delta_1(S) = \big\{\{u,v\} \mid u \in S, v \in V(G)\backslash S\big\}. \tag{3.18}$$

Lemma 3.2.1 (Edge Boundary Lemma) *Let the edge probability p be defined by equation (2.5) so that almost all graphs have minimum degree k and let $0 < \epsilon < 1$ be fixed. Then almost every graph G has the following property: For every subset S of V(G) with $|S| \leq \epsilon n$ we have*

$$|\delta_1(S)| \geq k|S|. \tag{3.19}$$

Proof: The expected number of sets with cardinalities between the integers a and b inclusive for which condition (3.19) does not hold is

$$F(a,b) = \sum_{s=a}^{b} \binom{n}{s} \sum_{j=1}^{ks-1} \binom{s(n-s)}{j} p^j (1-p)^{s(n-s)-j}. \tag{3.20}$$

We would like to find the largest interval possible on which $F(a,b) \to 0$ as $n \to \infty$. Now by using the fact that the inner sum is a binomial tail and applying Stirling's formula we see

$$F(a,b) \leq \sum_{s=a}^{b} \binom{n}{s} \binom{s(n-s)}{ks-1} p^{ks-1} (1-p)^{s(n-s)-ks+1} \left[\frac{p\{s(n-s)-(ks-1)+1\}}{p\{s(n-s)+1\}-(ks-1)} \right] \tag{3.21}$$

$$\leq \sum_{s=a}^{b} o(1) \binom{n}{s} \binom{s(n-s)}{ks-1} p^{ks-1} (1-p)^{s(n-s)-ks-1} \tag{3.22}$$

$$\leq \sum_{s=a}^{b} \mathcal{O}(1) \frac{1-p}{p} \left(\frac{ne}{s}\right)^s \left(\frac{s(n-s)e}{ks-1}\right)^{ks-1} (1-p)^{s(n-s)-ks} \tag{3.23}$$

$$\leq \sum_{s=a}^{b} \mathcal{O}(1) \frac{1-p}{p} \frac{ks-1}{s(n-s)} \left[\frac{n}{s} \left(\frac{s(n-s)e}{(ks-1)n} \right)^k (np)^k \exp[-p(n-s-k)] \right]^s \tag{3.24}$$

$$\leq \sum_{s=a}^{b} \mathcal{O}(\frac{1}{\log}) \left[o(1) \frac{n}{s} (\log n)^k \frac{e^{sp} e^{pk}}{n (\log n)^{k-1} e^{\omega_n}} \right]^s \tag{3.25}$$

$$\leq \sum_{s=a}^{b} \mathcal{O}(\frac{1}{\log n}) \left[o(1) \frac{\log n}{s} \{n (\log n)^{k-s} e^{\omega_n}\}^{s/n} e^{-\omega_n} \right]^s. \tag{3.26}$$

We now split the sum in (3.26) into two parts. First we would like to have $c_0 \frac{\log n}{s} < 1$ and $\{n(\log n)^{k-1} e^{\omega_n}\}^{s/n} < L$ where c_0 and L are constants. For this we need

$$a = c_0 \log n < s$$

and

$$s \leq \frac{c_1 n}{\log + (k-1) \log \log n + \omega_n} = \frac{c_1}{p} = b$$

where $c_1 = \log L$. It then follows that $F(c_0 \log n, \frac{c_1}{p}) \to 0$.

Secondly, for the upper part of the sum, if $a = \frac{c_1}{p}$ and $b = \epsilon n$ we see

$$F(\frac{c_1}{p}, \epsilon n) \leq \sum_{s=c_1/p}^{\epsilon n} O(\frac{1}{\log n}) \left[o(1) \frac{\log}{s} \{n(\log n)^{k-1} e^{\omega_n}\}^{s/n} e^{-\omega_n} \right]^s \tag{3.27}$$

$$\leq \sum_{s=c_1/p}^{\epsilon n} O(\frac{1}{\log n}) \left[o(1) \frac{(\log n)^{1+\epsilon(k-1)}}{c_1 n^{1-\epsilon}} \{\log n + (k-1)\log\log n + e^{\omega_n}\} e^{-\omega_n(1-\epsilon)} \right]^s \tag{3.28}$$

$$\to 0$$

because the term $n^{1-\epsilon}$ in the denominator will dominate. So $F(c_0 \log n, \epsilon n) \to 0$ as $n \to \infty$.

Showing $F(0, c_0 \log n) \to 0$ cannot be accomplished using the estimates we have employed here. In fact, we can easily see that the right hand side of (3.26) does not go to zero if $a = 1$ and $b = 2$. An entirely different approach must be used for sets of smaller order. First consider sets S with $|S| \leq \frac{1}{20(k+1)} \log n$. If S contains only small vertices then a consequence of the structural properties of Theorem 2.4.1 is that none of the vertices in S are adjacent to one another. Hence all of the edges incident with vertices in S are in $\delta_1(S)$, and since the minimum degree is k, inequality (3.19) holds. If S contains a large vertex, v, then at least $\frac{1}{20}\log n - \frac{1}{20(k+1)}\log n = \frac{k}{20(k+1)}\log n$ edges incident with v are in $\delta_1(S)$, and again inequality (3.19) is satisfied. ∎

3.3 Threshold Theorem

Now we apply the structural properties, the edge boundary lemma and the theorem of Nash-Williams to establish our main results. First we will find the threshold function for edge-disjoint spanning trees and as a consequence we also have the threshold function for collapsibility.

Theorem 3.3.1 *Let the edge probability be defined by*

$$p = \frac{\log n + (k-1)\log\log n + \omega_n}{n}$$

where $\omega_n \to \infty$ *and* $\omega_n = o(\log\log n)$ *so that almost all graphs have minimum degree* k. *Then almost all graphs have* k *edge-disjoint spanning trees.*

Proof: We wish to show that for every partition $\mathcal{P}$ of $V(G)$ we have

$$|E_{\mathcal{P}}(G)| \geq k(|\mathcal{P}| - 1).$$

We will use the convention that the $t = |\mathcal{P}|$ sets of $\mathcal{P}$ are ordered as follows

$$|V_1| \geq |V_2| \geq \ldots \geq |V_t|.$$

There are three cases. In the first two, ϵ is arbitrary.

Case 1: $|\mathcal{P}| \leq \frac{1}{2}n + 1$ and $|V_1| < \epsilon n$.

Since $\epsilon n \geq |V_1| \geq |V_2| \geq \ldots \geq |V_t|$, Lemma 3.2.1 will apply to every V_i. Hence

$$\begin{aligned} |E_{\mathcal{P}}(G)| &= \frac{1}{2}\sum_{i=1}^{|\mathcal{P}|} |\delta_1(V_i)| && (3.29) \\ &\geq \frac{1}{2}\sum_{i=1}^{|\mathcal{P}|} k|V_i| && (3.30) \\ &\geq \frac{nk}{2} && (3.31) \\ &\geq k(|\mathcal{P}| - 1). && (3.32) \end{aligned}$$

Case 2: $|\mathcal{P}| > \frac{1}{2}n + 1$ and $|V_1| < \epsilon n$.

It is sufficient to show that the average number of edges in the edge boundary of the sets of $\mathcal{P}$ is at least $2k$ almost surely, i. e.

$$\sum_{i=1}^{|\mathcal{P}|} |\delta_1(V_i)| \geq 2k|\mathcal{P}|. \tag{3.33}$$

Then it will follow that

$$|E_{\mathcal{P}}(G)| \geq k|\mathcal{P}|. \tag{3.34}$$

We call a set V_i of $\mathcal{P}$ *primary* if it consists entirely of large vertices. Otherwise a set of $\mathcal{P}$ is *secondary*. If a primary set consists of one or two vertices, its edge boundary has at least $\frac{\log n}{20}$ or $2\frac{\log n}{20} - 1$ edges respectively. In each case, the edge boundary will have at least $3k$ edges for n sufficiently large. Lemma 3.2.1 implies that any set with at least three vertices has at least $3k$ edges in its edge boundary. Hence all primary sets have at least $3k$ edges in their edge boundaries. Lemma 3.2.1 also implies that each secondary set has at least k edges in its edge boundary.

Theorem 2.4.1 states that there are at most $n^{1/3}$ small vertices, so it follows that $\mathcal{P}$ has at most $n^{1/3}$ secondary sets and at least $\frac{1}{2}n - n^{1/3}$ primary sets. Thus the number of primary sets exceeds the number of secondary sets and so the average number of edges in the edge boundaries of the sets of $\mathcal{P}$ is at least $2k$.

Case 3: $|V_1| > \epsilon n$.

Let $\epsilon > \frac{1}{2}$ so we have $|V_i| < \epsilon n$ for $i = 2$ to t. Let $R = \bigcup_{i=2}^{t} V_i$ and then clearly $|R| < \epsilon n$ and $\delta_1(V_1) = \delta_1(R)$. It follows from Lemma 3.2.1 that for $i = 2$ to t almost surely

$$|\delta_1(V_i)| \geq k|V_i| \tag{3.35}$$

and

$$|\delta_1(V_1)| = |\delta_1(R)| \geq k|R|. \tag{3.36}$$

So from these relations

$$\begin{aligned} |E_{\mathcal{P}}(G)| &= \frac{1}{2}\left(|\delta_1(V_1)| + \sum_{i=2}^{|\mathcal{P}|} |\delta_1(V_i)|\right) && (3.37)\\ &\geq \frac{1}{2}(k|R| + k|R|) = k|R| && (3.38)\\ &\geq k(|\mathcal{P}| - 1). && (3.39) \end{aligned}$$

Thus Theorem 3.1.1 can be applied to all three cases and so we know that a random graph of minimum degree k has k edge-disjoint spanning trees a. s. ∎

It follows from Theorem 3.3.1 that a random graph with minimum degree two has two edge-disjoint spanning trees almost surely. Hence Theorem 1.1.2 implies that with probability approaching 1 these are collapsible. Obviously a graph with a vertex of degree one is not collapsible. Thus, we have determined the threshold for collapsibility.

Corollary 3.3.2 *Let the edge probability p be defined by equation (2.5) with $k = 2$ so that almost all graphs have minimum degree two. Then almost all graphs are collapsible.*

Theorem 3.3.1 and Corollary 3.3.2 can be refined. Suppose p is defined by equation (2.5) but $\omega_n \to c$, for a constant c. No doubt it can be shown that the probability that the minimum degree is k and the probability that there are k edge-disjoint spanning trees are the same in the limit, namely $\exp[-\{\exp(-c)\}/(k-1)!]$, (see [Bo85], p. 61 for further details on the degree distribution).

Our proof of Theorem 3.3.1 is not algorithmic. On the other hand, if the minimum degree is at least $k = 2r + 1$, then there is a method for finding $r + 1$ edge-disjoint spanning trees in a random graph. First we need another consequence of the Edge Boundary Lemma 3.2.1.

Theorem 3.3.3 *Let the edge probability p be defined by equation (2.5) with $k \geq 2r+1$. Then a random graph G has r edge-disjoint spanning cycles, $C_1, C_2, \ldots, C_r$, and the subgraph H, which is obtained from G by deleting the edges of the r spanning cycles, is connected.*

Proof: Bollobás and Frieze have shown in [BoF85] that a random graph G with minimum degree $2r + 1$ has r edge-disjoint spanning cycles. Now fix $\epsilon > 1/2$. Then from Lemma 3.2.1 it follows that for any subset S of $V(G)$, with $|S| \leq n/2$, we have a.s.

$$|\delta_1(S)| \geq (2r + 1)|S|. \tag{3.40}$$

Removing the edges of r edge-disjoint spanning cycles from G means at most $2r|S|$ edges are removed from $\delta_1(S)$. Thus

$$|\delta_1(S)| - 2r|S| \geq |S| > 0, \tag{3.41}$$

so $\delta_1(S)$ in H is non-empty. Hence H is almost surely connected. ∎

Suppose we have a random graph G with minimum degree at least $2r + 1$. The algorithm of [BoFF87] can be used to find r edge-disjoint spanning cycles, and hence r edge-disjoint spanning paths. Let H be the graph obtained by deleting the edges of these paths from G. Theorem 3.3.3 implies that H is almost surely connected and so Depth-First Search on H will produce another spanning tree of G disjoint from the r paths.

All of the results in this section can be converted to Model B. Let us restate Corollary 3.3.2.

Corollary 3.3.4 *Let the number M of edges be given by*

$$M = \left\lfloor \frac{n}{2}(\log n + \log\log n + \omega_n) \right\rfloor, \tag{3.42}$$

where $\omega_n \to \infty$ but $\omega_n = o(\log\log n)$. Then almost all graphs in $\mathcal{G}(n, M)$ are collapsible.

We conclude by pointing out that Frieze and Luczak [FrL90] established a theorem corresponding to Theorem 3.3.1 for a slight variation of Model C. To obtain the graphs of their version, the orientation of the r arcs out of each vertex is ignored but a symmetric pair of arcs becomes a pair of multiple edges. They showed that these graphs obtained from r-out almost surely have r edge-disjoint spanning trees. They used the Nash-Williams theorem but their proof was significantly different because of the nature of the probability model.

3.4 Regular Graphs

In a *regular* graph all vertices have the same degree. A graph is *r-regular* if each vertex has degree r. Our sample space for random regular graphs consists of all labeled r-regular graphs on n vertices. We use the equiprobable model. In this section we will explore the collapsibility of regular graphs.

The *connectivity* $\kappa(G)$of a graph G is the minimum number of vertices whose removal from G results in a disconnected or trivial graph. The *edge connectivity* $\lambda(G)$ of a graph G is the minimum number of edges whose removal from G results in a disconnected or trivial graph. We will now determine the number of edge-disjoint spanning trees in a random regular graph.

Theorem 3.4.1 *For a fixed $r \geq 3$ almost all r-regular graphs have $\left\lfloor \frac{r}{2} \right\rfloor$ edge-disjoint spanning trees.*

Proof: Wormald has shown that for $r \geq 3$, almost all r-regular graphs have connectivity r (see [Wo81]). Now let G be an r-regular graph, $r \geq 4$, and let

$$\mathcal{P} = \{V_1, V_2, \dots, V_t\}$$

be a partition of $V(G)$. Using the fact that $\kappa(G) \geq \lambda(G)$ and the above mentioned result of Wormald, we see that $|\delta_1(V_i)| \geq r$ for all $0 \leq i \leq t$ almost surely. Therefore for almost all r-regular graphs,

$$|E_{\mathcal{P}}(G)| = \frac{1}{2}\sum_{i=1}^{|\mathcal{P}|}|\delta_1(V_i)| \tag{3.43}$$

$$\geq \frac{1}{2}\sum_{i=1}^{|\mathcal{P}|} r \tag{3.44}$$

$$\geq \left\lfloor\frac{r}{2}\right\rfloor(|\mathcal{P}|-1). \tag{3.45}$$

It then follows from Theorem 3.1.1 that almost all r-regular graphs will have $\lfloor\frac{r}{2}\rfloor$ edge-disjoint spanning trees. ■

Using Theorem 1.1.2, we can easily arrive at the following corollary.

Corollary 3.4.2 *For fixed $r \geq 4$, almost all r-regular graphs are collapsible.*

The question as to whether random 3-regular graphs (also called *cubic graphs*) are collapsible has yet to be answered. It is a much more difficult problem. As stated above, it is not even known if the Augmented Petersen graph is collapsible. The use of "definition two" of collapsibility will lead to a rather concise equivalent statement of the problem. Recall that a graph G is collapsible if for every subset S of $V(G)$ with even cardinality, there is a subgraph Γ such that $G - E(\Gamma)$ is connected and $v \in S$ if and only if the degree of v in Γ is odd. So if G is a cubic graph, in order for $G - E(\Gamma)$ to be connected the only odd degree allowed in Γ is one. This means Γ must be a collection of paths. Thus a cubic graph G is collapsible if and only if for every subset S of $V(G)$ with even cardinality there exist $|S|/2$ disjoint paths whose ends are the vertices of S and the removal of the edges of these paths does not disconnect G.

Let us refine the problem further. Suppose G is a cubic graph. Consider "definition one" of collapsibility and let the subset S of $V(G)$ be the empty set. We wish to find a spanning connected subgraph H in which the vertices of S have odd degree. Since

G is 3-regular, all vertices of H must have degree two. Hence, this case shows that if a cubic graph G does not have a spanning cycle, it is not collapsible. It has been shown by Robinson and Wormald that almost all cubic graphs have a spanning cycle (see [RoW92]), so it is possible that almost all cubic graphs are collapsible. However there are examples of cubic graphs that contain a spanning cycle and are not collapsible. A *diamond* is a K_4 with one edge removed. By arranging m diamonds in a circle and connecting each one to its successor by a new edge joining vertices of degree two, we create a *ring of m diamonds.* This ring is clearly hamiltonian but not collapsible if $m \geq 4$, because on contracting the triangles in the diamonds, a cycle C_m is left.

3.5 Large Eulerian Subgraphs

Let $\mu(G)$ be the minimum number of edges whose removal from a graph G leaves a spanning eulerian subgraph H. It follows from a result in [Pa88] that if $pn \geq \log n$ and $p \to 0$ as $n \to \infty$, then the number of vertices with odd degree is asymptotic to $n/2$. Furthermore if $pn \geq 2\log n$, then the subgraph induced by the vertices of odd degree has a perfect matching (a. s.). Hence for $pn \geq 2\log n$, μ aproaches $n/4$. Our object now is to estimate bounds on μ for smaller p.

Suppose p is defined by equation (2.5) with $k \geq 3$ so that the minimum degree is at least 3 (a. s.). From Theorem 3.3.3 we know that a random graph G has a spanning cycle C whose edges may be removed and a connected graph will remain. This cycle can be expressed as the edge-disjoint union of paths joining vertices of odd degree. On removing the edges of alternate paths of the cycle from G, we obtain an eulerian subgraph. Clearly no more than $n/2$ edges need to be removed. Hence $\mu \leq n/2$ almost surely.

On the other hand, suppose p is defined by equation (2.5) with $k = 2$ so that

a. s. the minimum degree is 2. We know from Theorem 3.3.1 that a random graph G will have two edge-disjoint spanning trees, T_1 and T_2. Now arrange the vertices of odd degree in pairs arbitrarily and consider the paths in T_1 that join each pair. By forming the symmetric difference of these paths, we obtain a set of edge-disjoint paths in T_1. The deletion of the edges in these paths leaves an eulerian graph and hence $\mu \leq n - 1$ (a. s.).

These results are summarized in the following theorem.

Theorem 3.5.1. *Suppose the edge probability p of a random graph is defined by equation (2.5) and $\epsilon > 0$ is arbitrary.*

If $k \geq 3$, then a. s.

$$(1 - \epsilon)n/4 \leq \mu \leq n/2,$$

and if $k = 2$, then a. s.

$$(1 - \epsilon)n/4 \leq \mu \leq n - 1.$$

Now there is the problem of refining these results. In the range where $pn = c \log n$, $1 < c < 2$, can the upper bound of $n/2$ be lowered? For p defined as in equation (2.5), a random graph will most likely not have a matching on the vertices of odd degree, so how may the lower bound of $(1 - \epsilon)n/4$ be raised? And finally, for the case $k = 2$ in Theorem 3.5.1, how much can the upper bound be improved? If a spanning cycle on the vertices of degree at least three can be found whose removal almost surely does not disconnect the graph, the upper bound can be lowed to $n/2$.

Bibliography

[Bo79] B. Bollobás, *Graph Theory, An Introductory Course*, Springer-Verlag, New York, (1979).

[Bo85] B. Bollobás, *Random Graphs*, Academic, London, (1985).

[BoF85] B. Bollobás and A. M. Frieze, On matchings and hamiltonian cycles in random graphs, *Annals of Discrete Math.* **28** (1985) 23-46.

[BoFF87] B. Bollobás, T.I. Fenner and A.M. Frieze, An algorithm for finding hamilton paths and cycles in random graphs, *Combinatorica* **7** (1987) 327-341.

[BoT86] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* **7** (1986) 35-38.

[Ca88a] P. A. Catlin, A reduction method to find eulerian subgraphs, *J. Graph Theory* **12** (1988) 29-45.

[Ca88b] P. A. Catlin, Supereulerian graphs, collapsible graphs, and four-cycles, *Congr. Numer.* **58** (1988) 233-246.

[ErR59] P. Erodös and A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6** (1959) 290-297.

[ErR60] P. Erodös and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Közl.* **5** (1960) 17-61.

[ErR61] P. Erodös and A. Rényi, On the strength of connectedness of random graphs, *Acta Math. Acad. Sci. Hungar.* **12** (1961) 261-267.

[FrL90] A. M. Frieze and T. Luczak, Edge disjoint spanning trees in random graphs, *Periodica Mathematica Hungarica* **21** (1990) 35-37.

[Ja85] F. Jaeger, A survey of the cycle double cover conjecture, *Annals of Discrete Math.* **27** (1985) 1-12.

[N-W61] C. St.J. A. Nash-Williams, Edge-disjoint spanning trees of finite graphs, *J. London Math. Soc.* **36** (1961) 445-450.

[Pa85] E. M. Palmer, *Graphical Evolution*, Wiley, New York (1985).

[Pa88] E. M. Palmer, Eulerian subgraphs of random graphs, *Congr. Numer.* **63** (1988) 139-145.

[RoW92] R. W. Robinson and N. C. Wormald, Almost all cubic graphs are hamiltonian, *Random Structures and Algorithms* **2** (1992) 117-125.

[RoT85] J. Roskind and R. E. Tarjan, A note on finding minimum-cost edge-disjoint spanning trees, *Math. Oper. Res.* **10** (1985) 701-708

[Sz73] G. Szekeres, Polyhedral decomposition of cubic graphs, *Bull. Austral. Math. Soc.* **8** (1973) 367-387.

[Tu61] W. T. Tutte, On the problem of decomposing a graph into n factors, *J. London Math. Soc.* **36** (1961) 221-230.

[Wo81] N. C. Wormald, The asymptotic connectivity of labelled regular graphs, *J. Combinatorial Theory, Ser. B* **31** (1981) 156-167.