5-2021

On Problems in Random Structures

Ryan Cushman
Western Michigan University, ryan.j.cushman@gmail.com

ON PROBLEMS IN RANDOM STRUCTURES

by

Ryan Cushman

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Mathematics
Western Michigan University
May 2021

Doctoral Committee:

Patrick Bennett, Ph.D., Co-Chair
Andrzej Dudek, Ph.D., Co-Chair
Alan Frieze, Ph.D.
Allen Schwenk, Ph.D.

ON PROBLEMS IN RANDOM STRUCTURES

Ryan Cushman, Ph.D.

Western Michigan University, 2021

This work addresses two problems in optimizing substructures within larger random structures. In the first, we study the triangle-packing number $\nu(G)$, which is the maximum size of a set of edge-disjoint triangles in a graph. In particular we study this parameter for the random graph $G(n, m)$. We analyze a random process called the online triangle packing process in order to bound $\nu(G)$. The lower bound on $\nu(G(n, m))$ that this produces allows for the verification of a conjecture of Tuza for $G(n, m)$. This conjecture states that in any graph G, there is a set of edges intersecting every triangle in the graph, and such that the size of this set of edges is bounded above by twice the triangle packing number $\nu(G)$. This work is a refinement of the methodology employed by Bennett, Dudek and Zerbib by establishing dynamic concentration of key random variables using the differential equation method. Tuza's conjecture has been independently verified in the case $G(n, p)$ by Kahn and Park using a very different approach.

In the second problem, we seek to study the number of paths in the r-edge-colored random graph $G(n, p)$, where adjacent edges have different colors. This question is inspired by a problem in coding theory and the work of Espig, Frieze, and Krivelevich, who found conditions under which a random graph with randomly 2-colored edges has a path that alternates between the two colors. Here the parameter alternat-

ing connectivity, $\kappa_{r,\ell}(G)$ is studied for random graphs. The parameter $\kappa_{r,\ell}(G)$ is the maximum t such that there is an r-edge-coloring of G such that any pair of vertices is connected by t internally disjoint and alternating (i.e. no consecutive edges of the same color) paths of length ℓ. We track this parameter's behavior in $G(n,p)$ over various ranges of p by utilizing different strategies and results for each range.

ACKNOWLEDGMENTS

I wish to express my gratitude to the many people who have helped me get to this point. For the plentiful advice and guidance of my advisors Dr. Patrick Bennett and Dr. Andrzej Dudek, who were instrumental to my development as a mathematician. And to my other dissertation committee members, Dr. Alan Frieze and Dr. Allen Schwenk, for their time and advice. To the many friends, classmates, and colleagues who have helped me learn, teach, and grow. To my parents and siblings for their continued support. And to my partner, Man Wing Chan, for her unending encouragement and blueberries.

Ryan Cushman

Contents

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Random Graphs

A *graph* is an ordered pair of sets (V, E), where members of V are called *vertices* and members of the set E, which consists of two-element subsets of V, are called *edges*. In 1959, Erdős and Rényi [28], and independently Gilbert [39], introduced models of graphs on n vertices where the edge set was determined randomly. The Erdős-Rényi model $G(n, m)$ is a graph with n vertices chosen uniformly from all graphs on $m = m(n)$ edges. Gilbert's model $G(n, p)$ is a graph on n vertices where each edge is present with probability $p = p(n)$, independently from the other edges.

These models have been extensively studied. The first results established in these models concerned *connectivity*. Two vertices u and v are connected if there exists a path from u to v, and a graph is connected if every pair of vertices is connected. Both in [28] and in [39], the authors studied the probability that $G(n, m)$ and $G(n, p)$ are connected for various ranges of p and m. Erdős and Rényi also proved more precise results concerning the probability of $G(n, m)$ having a certain number of connected components and the probability of the largest component being a certain size and in 1960 and 1961, they initiated the study of the "evolution" of random graphs [29, 30]. The goal of this pursuit was to determine the behavior of $G(n, m)$ for various functions $m(n)$. For a given property, such as the existence of a subgraph, they sought to show that the property was present in $G(n, m)$ with probability approaching 1 as n approaches infinity. When this is the case, the property is said to be present *with high probability* (abbreviated as w.h.p.). A remarkable discovery was the existence of "threshold functions" for certain properties. When a function is asymptotically larger than a threshold function for a property, the property holds with high probability; when asymptotically smaller, it does not hold with high probability. Thresholds were established for various properties. For connectivity, they established that the threshold is $\frac{1}{2}n\log n$, which was later generalized in [31] by Erdős and Rényi in the case of k-connectivity. A graph is called *k-connected* if it is connected and removing at most $k-1$ vertices leaves the graph connected. Thresholds were also proved for more local properties such as the existence of a particular subgraph [29], thus establishing the values of m for which the probability of finding that subgraph goes to 1 as n goes to infinity (and the values for which the probability approaches 0). Bollobás and Thomason underscored the centrality of these threshold functions by showing that they always exist when the property is monotone and nontrivial [24]. (A property is *monotone increasing* if adding an edge preserves the property and *monotone decreasing* if deleting an edge preserves the property.) This early work, especially the study of connectivity in random graphs, provided a foundation for this work's search for alternating paths in edge-colored random graphs. While connectivity seeks a path between every pair of vertices, we will seek as many paths as possible between pairs of vertices in edge-colored graphs in which adjacent edges do not share colors. We call this parameter alternating connectivity, which we will discuss in Chapter 3.

1.2 Random Processes

One of the most important questions in combinatorics is the following: how large does a structure have to be to guarantee a certain substructure? This is most classically applied to the complete graph K_N, which is the graph consisting of N vertices and every possible edge between them. This gives rise to the elusive Ramsey numbers, whose existence was established by Ramsey in 1930 [52]. The *Ramsey number $R(s,t)$* is the smallest N such that any red-blue coloring of the edges of K_N yields either a blue K_s or a red K_t as a subgraph. Determining these numbers, even for relatively small values, has been difficult. Hence, most progress has been in determining the order of magnitude of these numbers. A particularly helpful framework has been looking at random graphs and random processes. In particular, the search for precise asymptotic bounds on $R(3,t)$ resulted in the development and analysis of random processes run on random graphs.

The best-known of these is the *triangle-free process* introduced by Bollobás and Erdős (see [23]), which maintains a triangle-free subgraph $G_T(i) \subset G(n,i)$. Here $G(n,m)$ is the Erdős-Rényi random graph that assigns equal probability to all graphs on a fixed set V of n vertices with exactly $m = m(n)$ edges. At each step in the triangle-free process, an edge is revealed and added to $G_T(i)$ only if it does not create a triangle in $G_T(i)$. Famously used to study the Ramsey numbers $R(3,t)$, the triangle-free process has continued to yield results in this area; indeed, this can be seen from the recent results of Bohman and Keevash [18] and independently of Fiz Pontiveros, Griffiths and Morris [51], which gives the best-known lower bound of $R(3,t) \geq (1/4 - o(1))\, t^2/\log t$.

Another important process, called *random triangle removal* and also introduced by Bollobás and Erdős, creates a triangle-free graph from a complete graph instead of an empty graph (see [20, 19]). Here we start with $G_R(0) = K_n$ and remove all three edges from a triangle chosen uniformly at random from the triangles in $G_R(i)$ until there are no triangles remaining. Note that the edges of the removed triangles form a triangle packing in K_n. Although originally motivated by the study of $R(3,t)$, this process has not yet resulted in any good bound. Bohman, Frieze and Lubetzky [17] provided the best-known upper and lower bounds of $n^{3/2+o(1)}$ on the number of edges with high probability remaining at the end of this process.

A third process introduced by Bollobás and Erdős is the *reverse triangle-free process*. Similar to the previous process, this one starts with $G_{RT}(0) = K_n$ but instead removes *one* edge that is in a triangle from $G_{RT}(i)$. As before, the process terminates when the graph is triangle-free. The number of edges in the final graph is w.h.p. $(1+o(1))\sqrt{\pi}n^{3/2}/4$ due to Erdős, Suen and Winkler [33] and the final number of edges is concentrated about its expectation due to Makai [48] and independently Warnke [56].

In this work, we consider the *online packing process* defined as follows. Start with $G(n,0)$ and at step i we choose a nonexistent edge uniformly at random to add to $G(n,i)$. Define the sequences $M(i)$, the *matched graph at step* i, and $U(i)$, the *unmatched graph at step* i, such that $G(n,i) = M(i) \cup U(i)$ in the following manner. (For simplicity, we will frequently identify a graph H with its edge set $E(H)$.) When an edge is added, if this edge forms a triangle that is edge-disjoint from $M(i)$, then many triangles might be created as well. Choose one of these triangles uniformly at random from the set of created triangles and add its edges to $M(i+1)$. Set $U(i+1) = G(n,i+1) - M(i+1)$, which will have no triangles by induction on i. Note that $M(i)$ forms a triangle packing. We will investigate the online packing process to find a triangle packing in Chapter 2.

1.3 Preliminaries

In this section, we introduce several important tools that will be used extensively. For more information, see [5, 21, 38, 44].

1.3.1 Notation

We will take all asymptotics in n unless noted otherwise. Beside the standard *Big-O* and *Little-o* notation we will sometimes write $\sim f(n)$ instead of $(1+o(1))f(n)$ and $f(n) \ll g(n)$ instead of $f(n) = o(g(n))$. For simplicity, we do not round numbers that are supposed to be integers either up or down; this is justified since these rounding errors are negligible to the asymptomatic calculations we will make. All logarithms are natural unless written explicitly.

1.3.2 Random Graph Models

We define the Erdős-Rényi random graph $G(n,m)$, also known as the *uniform random graph*, as a graph on n vertices where m edges are inserted into the empty graph in such as way that all $\binom{\binom{n}{2}}{m}$ choices are equally likely.

We define the *binomial random graph* $G(n, p)$ as the graph on n vertices and edge set determined by successful trials in a Bernoulli experiment consisting of $\binom{n}{2}$ potential edges and p as the probability of success. In many cases, these two models can be shown to be equivalent.

1.3.3 With High Probability

We say that a sequence of events $X(n)$ happens *with high probability* (abbreviated w.h.p.) if $P(X(n)) \to \infty$ as $n \to \infty$.

1.3.4 Threshold Function

A graph property is *monotone increasing* if adding an edge preserves the property. Let $\mathscr{P}$ be a monotone increasing graph property. A function $f(n)$ is a *threshold function* of property $\mathscr{P}$ in $G(n, m)$ if

$$\lim_{n\to\infty} \Pr(G(n, m(n)) \text{ has property } \mathscr{P}) = \begin{cases} 0 & \text{if} \quad \lim_{n\to\infty} \frac{m(n)}{f(n)} = 0 \\ 1 & \text{if} \quad \lim_{n\to\infty} \frac{m(n)}{f(n)} = 1. \end{cases}$$

In a similar manner, a function $f(n)$ is a *threshold function* of property $\mathscr{P}$ in $G(n, p)$ if

$$\lim_{n\to\infty} \Pr(G(n, p(n)) \text{ has property } \mathscr{P}) = \begin{cases} 0 & \text{if} \quad \lim_{n\to\infty} \frac{p(n)}{f(n)} = 0 \\ 1 & \text{if} \quad \lim_{n\to\infty} \frac{p(n)}{f(n)} = 1. \end{cases}$$

1.3.5 Chernoff's Bound

Throughout, we will be using the following forms of Chernoff's bound (see, e.g., [44]). Let $X \sim \text{Bin}(n, p)$ and $\mu = \text{E}(X)$. Then, for all $0 < \delta < 1$

$$\Pr(X \geq (1+\delta)\mu) \leq \exp(-\mu\delta^2/3) \tag{1.1}$$

and

$$\Pr(X \leq (1-\delta)\mu) \leq \exp(-\mu\delta^2/2). \tag{1.2}$$

1.3.6 Other Inequalities

Several times we will also use the following well-known inequalities:

$$1 - x \leq e^{-x} \text{ for any } x \text{ and } 1 - x/2 \geq e^{-x} \text{ for } 0 \leq x \leq 1. \tag{1.3}$$

1.3.7 The Differential Equation Method

The random processes described above, although very straightforward to outline, are almost impossible to track without the proper framework. The collection of techniques and tools that has been successful in this regard is called the *differential equations method* for determining dynamic concentration of random variables (see [13, 57]). When studying an iterative process consisting of small steps, such as the addition of an edge in a graph, we seek to know how key statistics of certain random variables behave. If we establish *dynamic concentration*, then these random variables would remain close to their expectation, even as this expectation changes with each step. This allows us to make powerful conclusions about the process, such as (with high probability) how many edges are present at the end of the process. As the name suggests, a differential equation is crucial in the process. If each step is small enough, we may treat the process as continuous without too much error, allowing us to estimate the one-step change of a random variable with a derivative. This allows us to successfully apply probabilistic tools such as martingale concentration inequalities to establish dynamic concentration.

This approximation is formalized using Taylor's Theorem with the Lagrange form of the remainder, which we state here for convenience.

Theorem 1.3.1 (Taylor's Theorem [49]). *Let n be a natural number and suppose that a function f is $n+1$ times differentiable in $[a,b]$. If $c \in [a,b]$, then for any $x \in [a,b]$,*

$$f(x) = f(c) + \frac{df(c)}{dx}(x-c) + \frac{1}{2!}\frac{d^2 f(c)}{dx^2}(x-c)^2 + \ldots + \frac{1}{n!}\frac{d^n f(c)}{dx^2}(x-c)^n + r_n(x)$$

where

$$r_n(x) = \frac{1}{(n+1)!}\frac{d^{n+1}f(\omega)}{dx^{n+1}}(x-c)^{n+1}$$

and ω is between x and c.

In particular, we will estimate the change in a function H from $i-1$ to i as

$$H(i) - H(i-1) = \frac{dH(i-1)}{di} + \frac{1}{2}\frac{d^2 H(\omega)}{di^2}.$$

One of the first examples of methods resembling the current differential equation method is found in the famous 1981 matching algorithm of Karp and Sipser [46]. This algorithm outputs a near-maximum matching for the sparse random graph $G(n, c/n)$ (for some constant c) and runs in linear time. This was an important constructive approach to finding matchings in random graphs, contrasting with the traditional nonconstructive approaches (e.g. used by Erdős and Rényi in [32]). Their innovative analysis of the algorithm used a differential equation to establish tight bounds on random variables that changed over time, a key concept in what would become the differential equation method.

Another important step in the development of this process is Wormald and Ruciński's [53] work in analyzing the d-process. This processes starts with the empty graph and adds edges chosen uniformly at random from all edges not violating the following condition: all vertices must have degree at most d. Their analysis—particularly their use of a martingale concentration inequality—more closely resembles the techniques used today and provided a foundation for future work. This work was continued by Wormald, who was instrumental in the development of this method. He was able to study many processes and formulated a "black box" theorem that could be used in some cases to achieve dynamic concentration immediately [57]. Further innovations came from Bohman, who creatively applied previously unused inequalities such as in his initial analysis of the triangle-free process [16]. These innovations allowed for the groundbreaking results that $R(3,t) \geq (1/4 - o(1))\, t^2/\log t$ by Bohman and Keevash [18]. In Chapter 2, we will utilize the differential equation method to analyze the online triangle packing process.

1.3.8 Martingale Concentration Inequalities

We will require the following two martingale concentration inequalities in order to establish dynamic concentration. We will use Freedman's inequality for the bulk of our analysis of the online triangle packing process. We will also need to consider the triangle-free process. In that case, the Azuma-Hoeffding inequality will be utilized. Freedman's inequality is, in some senses, a refinement of the Azuma-Hoeffding. In particular, the loosen restriction that $\Delta Y(k) \leq C$ is more helpful than $|\Delta Y(k)| \leq C$ in cases where $\Delta Y(k)$ is a large negative number at some k, thus allowing for the choice of a smaller C.

Lemma 1.3.2 (Azuma [7], Hoeffding [43]). *If $Y_0, Y_1, \ldots$ variables are supermartingales and w.h.p. $|Y_j - Y_{j-1}| \leq C$, then for all positive integers m and λ*

$$\Pr(Y_m - Y_0 \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{2C^2 m}\right).$$

Lemma 1.3.3 (Freedman [37]). *Let $Y(i)$ be a supermartingale with $\Delta Y(i) \leq C$ for all i, and let*

$V(i) := \sum_{k \le i} \mathbf{Var}[\Delta Y(k) | \mathscr{F}_k]$. *Then,*

$$\Pr[\exists i : V(i) \le b, Y(i) - Y(0) \ge \lambda] \le \exp\left(-\frac{\lambda^2}{2(b + C\lambda)}\right).$$

1.4 Triangle Packings and Tuza's Conjecture

We define the *triangle packing number* of a graph G, denoted $\nu(G)$, as the maximal size of a set of edge-disjoint triangles. We employ an analysis of the online triangle packing process to establish dynamic concentration using the differential equation method to obtain the following theorem. This work is based on a paper with Bennett and Dudek [11].

Theorem 1.4.1. *Let $G = G(n, m)$ be a random graph of order n and size $m = kn^{3/2}$, where $0 \le k < \frac{1}{10000} \log\log n$. Then, w.h.p.*

$$\nu(G) \ge (1 + o(1))\frac{1}{3}\left(k - \frac{y(k)}{2}\right)n^{3/2},$$

where $y(t)$ is the solution of $dy/dt = 6e^{-y^2} - 4$ with initial condition $y(0) = 0$.

The proof of Theorem 1.4.1 can be found in Section 2.1. The version of the online triangle packing process we use here is a refinement of the one used by Bennett, Dudek, and Zerbib in [15]. Recall that the creation of a triangle in $U(i)$ potentially coincides with the creation of many triangles. More specifically, it coincides with the creation of a copy of the tripartite graph $K_{1,1,s}$ for some $s \ge 1$ (called a *book*). In [15], instead of choosing a triangle uniformly at random from the set of created triangles, the edges of $K_{1,1,s}$ were moved to $M(i+1)$ for maximal $s \ge 1$. Then U is triangle-free and a triangle packing can be obtained by choosing one triangle from each copy of $K_{1,1,s}$ in M. This modification was done to aid in technical details of the analysis, but also resulted in more edges being moved to the matched set at each step in the greedy algorithm than was necessary to obtain a triangle packing. In contrast, our form of the process moves only what is necessary to have U be triangle-free and M be a triangle packing. This allows the triangle-free graph at each step to have less edges, allowing the process to continue for longer. In addition, we were able to simplify the troublesome technical details, resulting in a more streamlined analysis.

We give an application of Theorem 1.4.1 in order to show that a conjecture of Zsolt Tuza holds for random graphs. This is done though proving the conjecture for a range of m that was previously in question. Tuza's conjecture relates the triangle packing number and the triangle covering number. The *triangle covering number* $\tau(G)$ is the minimal size of a set of edges intersecting all triangles. It is easy to see that $\nu(G) \le \tau(G) \le 3\nu(G)$ for any graph G. Tuza, however, conjectured that this trivial upper bound could be lowered.

Conjecture 1.4.2 (Tuza [55]). *For every graph G, $\tau(G) \le 2\nu(G)$.*

The conjecture is tight for the complete graphs of order 4 and 5. The best-known upper bound is $\tau(G) \le \frac{66}{23}\nu(G)$ from Haxell [41]. A recent development is due to Baron and Kahn [9], who showed that, in a strong sense, the multiplicative constant 2 in Tuza's conjecture cannot be improved in general. They demonstrated that for any $\alpha > 0$ there are arbitrarily large graphs G of positive density satisfying $\tau(G) > (1-o(1))|G|/2$ and $\nu(G) < (1+\alpha)|G|/4$, disproving a conjecture of Yuster [58]. See [47, 42, 1] for related results.

We will augment the analysis of the triangle-free process from Bohman [16] in order to obtain an upper bound on $\tau(G(n,m))$ and pair this with our lower bound on $\nu(G(n,m))$. Note that we are concerned here with a small range of m for which Tuza's conjecture is still open. Using their result on $\nu(G(n,m))$, Bennett, Dudek and Zerbib [15] proved the conjecture for $G(n,m)$, with the exception of a small range of m.

Theorem 1.4.3 (Bennett, Dudek, and Zerbib [15]). *There exist absolute constants $0 < c_1 < c_2$ such that if $m \le c_1 n^{3/2}$ or $m \ge c_2 n^{3/2}$, then w.h.p. Tuza's conjecture holds for $G = G(n,m)$.*

Note that we are only concerned with the sparse case of $G(n,kn)$ for $k < (\log n)^2$ since Frankl and Rödl [35] gave a bound on $\nu(G)$ that is optimal in order for $k \ge (\log n)^2$. (This was slightly improved by Pippenger by decreasing the bound on k, see [6].)

We will show in Chapter 2 that the following theorem.

Theorem 1.4.4. *Tuza's conjecture holds w.h.p. for $G(n,m)$ for any range of m.*

1.5 Alternating Paths

An alternating path is a path with adjacent edges having distinct colors. Originally studied by Bollobás and Erdős [22] in the context of finding alternating Hamiltonian cycles in complete graphs, alternating paths have been widely studied (e.g. [3, 4, 8, 25, 26, 54]). Along with being interesting objects in their own right, alternating paths can also be used to encapsulate certain parameters of codes in coding theory.

One such fundamental parameter of a code $C \subseteq [r]^m$ is the minimum *Hamming distance* between codewords $\mathbf{x}$ and $\mathbf{y}$ in C. This distance between codewords is defined as the number of positions where they differ:

$$\mathrm{dist}(\mathbf{x},\mathbf{y}) = \Big|\{i : 1 \le i \le m,\ \mathbf{x}(i) \ne \mathbf{y}(i)\}\Big|.$$

In the noisy channel model of coding theory, we wish to detect errors introduced during transmission by optimizing various parameters of the code. Codes with large minimum Hamming distance between codewords allow for greater detection and correction of errors. At the same time, one wants to be able to send many messages across this channel. Thus, one may wish to ask what the maximum n is such that there exists a code $C \subseteq [r]^m$ of n codewords and minimum Hamming distance t. We call this number $\alpha_r(m,t)$. (See [50] for more details concerning coding theory.)

We now encode this problem in terms of alternating paths. Let $K_{m,n}$ be a complete bipartite graph on vertex set $[m] \cup [n]$ and $c : E(K_{m,n}) \to [r]$ be an r-edge-coloring of $K_{m,n}$ with the property that every pair of vertices in $[n]$ is connected by at least t alternating paths of length 2 (with 3 vertices). We can represent this coloring as a collection of n vectors of length m with entries in $[r]$ in the following way: for a vertex $v \in [n]$, define the vector $\mathbf{c}_v \in [r]^m$ by $\mathbf{c}_v(u) = c(\{v,u\})$ for $u \in [m]$. Then $C = \{\mathbf{c}_v : v \in [n]\}$ fully encodes the edge coloring of $K_{m,n}$ since every edge will belong to a unique vector which will determine its color.

But if we take a path of length 2 of the form $v - u - w$ for $v, w \in [n]$, notice that this is alternating if and only if $c(\{v,u\}) \neq c(\{u,w\})$, that is, $\mathbf{c}_v(u) \neq \mathbf{c}_w(u)$. So the number of alternating paths between v and w is exactly the Hamming distance between $\mathbf{c}_v$ and $\mathbf{c}_w$. And since c has the property that every pair of vertices in the same partite set is connected by at least t alternating paths of length 2, it follows that C has minimum Hamming distance t. Therefore we may conclude that $|C| = n \leq \alpha_r(m,t)$.

So determining $\alpha_r(m,t)$ is equivalent to finding the largest n such that there is an r-coloring of the edge set of $K_{m,n}$ with the property that any pair of vertices in $[n]$ has at least t alternating paths of length 2 connecting them. Also notice that all such paths are internally disjoint.

Bennett, Dudek and LaForge [14] used this characterization to study a slightly different problem. Instead of fixing the alphabet, word length, and t and asking for the largest possible code, they fixed the alphabet, word length, and code size and asked for the largest possible t. In addition, they generalized the length of the alternating paths from 2 to any constant length $2k$. Their object of study was $\kappa_{r,2k}(m,n)$, the maximum t such that there is an r-coloring of the edges of $K_{m,n}$ such that any pair of vertices in class of size n is connected by t internally disjoint and alternating paths of length $2k$.

They showed that for fixed $r \geq 2$ and $n \geq m \gg \log n$

$$\kappa_{r,2}(m,n) \sim \left(1 - \frac{1}{r}\right) m$$

They also proved that for any fixed $r \geq 2$, $k \geq 2$ and $n \geq m \gg 1$

$$\kappa_{r,2k}(m,n) \sim \frac{m}{k}.$$

Besides these results, the authors also proposed the concept of alternating connectivity, which is the main concern of Chapter 2. This parameter was in part inspired by the work of Espig, Frieze, and Krivelevich [34], who found conditions under which a random graph with randomly two-colored edges has an alternating path joining any pair of vertices. *Alternating connectivity* $\kappa_{r,\ell}(G)$ was defined to be the maximum t such that there is an r-edge-coloring of G such that any pair of vertices is connected by t internally disjoint and alternating paths of length ℓ. For complete graphs, it was shown by Bennett, Dudek and LaForge [14] that

$$\kappa_{r,2}(K_n) \sim (1-1/r)n \quad \text{and} \quad \kappa_{r,\ell}(K_n) \sim n/(\ell-1)$$

for any $r \geq 2$ and $\ell \geq 3$. Motivated by this progress, we study the alternating connectivity of the random graph $G(n,p)$.

We have three main results concerning this parameter for $G(n,p)$, which are based on a paper with Bennett and Dudek [12]. The first result is concerned with the case when we seek paths of length two. Here, our answer is essentially the same as the expected codegree between two vertices in $G(n,p)$.

Theorem 1.5.1. *Let $p \gg \sqrt{\log n/n}$ and $G = G(n,p)$ and let r be an integer. Then, w.h.p.*

$$\kappa_{r,2}(G) \sim \left(1 - \frac{1}{r}\right) np^2.$$

Observe that if $p \ll \sqrt{\log n/n}$ then the diameter of G is at least three. Therefore, in this sense the above theorem is optimal.

After this, we consider $\ell \geq 3$ by looking at a "dense" case and a "sparse" case. For the "dense" case, when p is a constant, we have found that we have two main obstructions. Clearly, we cannot have more paths between two vertices than the minimum degree. However, the neighborhoods of two vertices w.h.p. have an intersection of size $\sim np^2$, so the most we could hope for here is $np(1-p/2)$. On the other hand, the total number of paths of length ℓ between two vertices is bounded above by $n/(\ell-1)$.

Theorem 1.5.2. *Let $0 < p < 1$ be a constant and $G = G(n,p)$. Then, for any integer $\ell \geq 3$ w.h.p.*

$$\kappa_{r,\ell}(G) \sim \min\left\{\frac{n}{\ell-1}, np\left(1-\frac{p}{2}\right)\right\}.$$

Finally, we consider the "sparse" case when $p = o(1)$. Here, we have two main obstructions: in parts (i) and (ii), we are limited by the minimum degree; in part (iii) we are limited by the total number of paths of length k. Here, we analyze $\kappa_{r,\ell}(G)$ in based on how close we are to the threshold for the diameter in $G(n,p)$.

Theorem 1.5.3. *Suppose $G = G(n,p)$ with $p = o(1)$ and $r \geq 2$ is an integer.*

(i) Let $k \geq 2$ be a positive integer such that $n^{1/k} \leq np \leq n^{1/(k-1)}$. If $\ell \geq k+2$, then w.h.p. we have $\kappa_{r,\ell}(G) \sim np$.

(ii) Let $k \geq 2$ be a positive integer such that $(n\log n)^{1/k} \ll np \leq n^{1/(k-1)}$. If $\ell = k+1$, then w.h.p. we have $\kappa_{r,\ell}(G) \sim np$.

(iii) Let $k \geq 3$ be a positive integer such that $(n\log n)^{1/k} \ll np \ll n^{1/(k-1)}$. If $\ell = k$, then w.h.p. we have $\kappa_{r,\ell}(G) = \Theta(n^{k-1}p^k)$.

Note that we obtain the most satisfying result for $\ell \geq k+2$ (in part (i)), in two senses: first there are no "gaps" between the values of p for which it applies, and second we obtain an estimate of $\kappa_{r,\ell}$ that is accurate up to a multiplicative $(1+o(1))$ error. Intuitively this is because the diameter is w.h.p. either k or $k+1$ (see, for example, Corollary 10.12 in [21]) and $\ell \geq k+2$ is safely larger than that. Roughly speaking, it becomes more difficult when our path length is closer to the diameter. In parts (ii) and (iii) the diameter is w.h.p. k and the results become weaker as we seek shorter paths.

Finally, we close with some remarks about the current state of $\lambda_{r,\ell}(G)$ (where we no longer require that paths are internally disjoint) and extending our results to pseudorandom graphs.

CHAPTER 2

TRIANGLE PACKINGS AND TUZA'S CONJECTURE IN RANDOM GRAPHS

2.1 Finding a Triangle Packing Through the Random Process

2.1.1 Introduction

We now turn to finding a maximal triangle packing in the random graph $G(n.m)$ and employ this to show that Tuza's conjecture holds in $G(n,m)$. Recall that the *triangle packing number* of a graph G, denoted $\nu(G)$, is the maximum size of a set of edge-disjoint triangles. Also, the *triangle covering number* $\tau(G)$ is the minimum size of a set of edges intersecting all triangles. One can show that these two parameters satisfy the trivial inequality $\nu(G) \leq \tau(G) \leq 3\nu(G)$ for all G. However, Tuza conjectured that these two were related in the following non-trivial way.

Conjecture 1.4.2 (Tuza [55]) *For every graph G, $\tau(G) \leq 2\nu(G)$.*

There has been some progress on this conjecture since its beginnings in 1981. The constant 3 in the trivial upper bound on the triangle covering above has been lowered by Haxell to about 2.87 [41]. In addition, Baron and Kahn found graphs that prove that, in many cases, the multiplicative constant 2 in Tuza's conjecture cannot be lowered. Further progress was made by Krivelevich [47] in 1995, who showed that Tuza's conjecture holds in $K_{3,3}$-free graphs. There have also been fractional versions of Tuza's conjecture and other generalizations [42, 1].

In addition, Bennett, Dudek and Zerbib [15] studied Tuza's conjecture for $G(n,m)$. They proved the following.

Theorem 2.1.1 (Bennett, Dudek, and Zerbib [15]). *There exist absolute constants $0 < c_1 < c_2$ such that if $m \leq c_1 n^{3/2}$ or $m \geq c_2 n^{3/2}$, then w.h.p. Tuza's conjecture holds for $G = G(n,m)$.*

The proof of Theorem 2.1.1 gives that one can take $c_1 := 0.2403$ and $c_2 := 2.1243$. The existence of the constant c_1 was recently also proved by Basit and Galvin [10]. Our turn to this problem in this chapter, seeking to remove this gap in certainty. We will be able to show the following.

Theorem 1.4.4. *Tuza's conjecture holds w.h.p. for $G(n,m)$ for any range of m.*

This will be proved in Section 2.2. A proof of Theorem 1.4.4 for $G(n,p)$ was independently proved by Kahn and Park [45] using a different approach. These results can be shown to be equivalent (see Proposition 1.12 in [44]).

Theorem 1.4.4 is obtained by finding a large triangle packing in $G(n,m)$ using the online triangle packing process, allowing for the lower bound on $\nu(G(n,m))$ given below. We restrict our concern to $G(n,kn)$ with $k < (\log n)^2$, as much progress has been made on $\nu(G)$ when $k \geq (\log n)^2$ [6, 35].

Theorem 1.4.1. *Let $G = G(n,m)$ be a random graph of order n and size $m = kn^{3/2}$, where $0 \leq k < \frac{1}{10000} \log\log n$. Then, w.h.p.*

$$\nu(G) \geq (1+o(1))\frac{1}{3}\left(k - \frac{y(k)}{2}\right) n^{3/2},$$

where $y(t)$ is the solution of $dy/dt = 6e^{-y^2} - 4$ with initial condition $y(0) = 0$.

This result follows from an analysis of the online triangle packing process. As with other random processes, we track key statistics of the process through random variables that change over time. Then we use the differential equation method to establish dynamic concentration of these random variables. To show Theorem 1.4.4, we use Theorem 1.4.1 and then find an upper bound on $\tau(G(n,m))$. This will involve considering the triangle-free process used in [16]. After this, we conjecture that Tuza's conjecture can be improved in the case of the random graph $G(n,m)$.

2.1.2 Outline of the Algorithm

Our process reveals one edge of $G(n,m)$ at each step. So for step i we have $G(n,i)$, whose edges we partition into subgraphs: the *matched graph* $M(i)$ and the *unmatched graph* $U(i)$. We will maintain the property that $U(i)$ is triangle-free, and $M(i)$ is the union of disjoint triangles. Let e_i be the edge we add at step i. If $U(i)\cup\{e_i\}$ is triangle-free then we let $U(i+1) := U(i)\cup\{e_i\}$ and $M(i+1) := M(i)$. Otherwise $U(i)\cup\{e_i\}$ has at least one triangle (any such triangle must use e_i since $U(i)$ is triangle-free), and we choose one such triangle uniformly at random. If from among the set of such triangles, we choose say T (a set of three edges), then we set $M(i+1) := M(i)\cup T$ and $U(i+1) = U(i)\setminus T$. We remark that the edge $e_i = \{u,v\}$ creates a triangle precisely when the codegree of u and v in $U(i)$ is positive. Recall that the *codegree* of two vertices u and v in a graph H, written $\text{codeg}_H(u,v)$, is the number vertices w such that both uw and vw are edges of H.

2.1.3 Heuristics

Along with the codegree of pairs of vertices in $U(i)$, we are also concerned with the degree of each vertex in both $M(i)$ and $U(i)$. Fix a vertex v. Then we write $d_U(v,i) = \deg_{U(i)}(v)$ and $d_M(v,i) = \deg_{M(i)}(v)$ to represent the unmatched and matched degree at step i, respectively. We will also write $d_G(v,i) = \deg_{G(n,i)}(v) = d_U(v,i) + d_M(v,i)$. For convenience, we will sometimes suppress "i" in this notation when it is clear from context.

Now define the scaled time parameter as

$$t = t(i) := \frac{i}{n^{3/2}}$$

for $0 \le i \le \frac{1}{10000}n^{3/2}\log\log n$. At each step $i+1$ we choose a random edge without replacement.

Thus, the probability of choosing any particular edge that has not been chosen yet is

$$\frac{1}{\binom{n}{2} - i} = \frac{2}{n^2}(1 + \tilde{O}(n^{-1/2})),$$

where $a(n) = \tilde{O}(b(n))$ if there exists $k \geq 0$ such that $a(n) \in O(b(n) \log^k n)$.

Next we describe some important heuristics which will be formally justified later. The first of these is that at each step i (excluding steps near the start), we have

$$d_{U(i)}(v) + d_{M(i)}(v) = deg_{G(n,i)}(v) = \frac{2i}{n}(1 + o(1)) = 2tn^{1/2}(1 + o(1))$$

for sufficiently large m. This is due to the concentration of vertex degrees in $G(n, m)$. Assuming heuristically that $d_U(v) \approx y(t)n^{1/2}$ and that the codegrees in $U(i)$ are distributed Poisson with expectation $n(yn^{-1/2})^2 = y^2$, we may say that $d_M(v) \approx (2t - y(t))n^{1/2}$ and that the number of unmatched edges is about $\frac{1}{2}n^{3/2}y$.

With this framework, we calculate the one-step change in the number of unmatched edges $e(U(i))$. If $e_i = \{u, v\}$, we gain one unmatched edge when $\text{codeg}_U(u, v) = 0$, which occurs with probability e^{-y^2}. If the unmatched codegree is positive, then we lose two unmatched edges; they move into the matched graph $M(i)$ along with e_i. Thus, approximating the one-step change as a derivative, we arrive at the following:

$$\Delta\left(\frac{1}{2}y(t)n^{3/2}\right) \approx \left(\frac{1}{2}\frac{dy(t)}{dt}n^{3/2}\right)\Delta t = \frac{1}{2}\frac{dy}{dt} \approx 1 \cdot e^{-y^2} - 2 \cdot (1 - e^{-y^2}).$$

Here we use the fact that $\Delta t = n^{-3/2}$. Hence, we get the differential equation $dy/dt = 6e^{-y^2} - 4$ with the initial condition $y(0) = 0$. (We discuss this differential equation further in Section 2.1.4.)

To conclude this heuristic, we recall that the number of matched edges $e(M(kn^{3/2}))$ is $kn^{3/2} - \frac{y(k)}{2}n^{3/2}$ after $kn^{3/2}$ edges have been revealed. Thus the number of edge-disjoint triangles we obtain at the end of our process should be

$$\left(\frac{1}{3}k - \frac{1}{6}y(k)\right)n^{3/2}.$$

We show that with high probability this is very close to the actual situation.

2.1.4 Preliminaries

Let $y = y(t)$ for $t \geq 0$ be such that the following autonomous differential equation holds:

$$\frac{dy}{dt} = 6e^{-y^2} - 4,$$

subject to $y(0) = 0$. Then y is an increasing function of t and y approaching the unique positive root of the equation $6e^{-x^2} - 4 = 0$ (as t goes to infinity), which is $\zeta = \sqrt{\log\left(\frac{3}{2}\right)} \approx 0.6367$. Hence, $0 \leq y \leq \zeta$. This also implies that $y'(t) \geq 0$.

Furthermore, note that

$$\frac{d^2y}{dt^2} = -12e^{-y^2}y\frac{dy}{dt} \leq 0 \tag{2.1}$$

and consequently $0 \leq dy/dt \leq dy/dt(0) = 2$.

For integers $b, c \geq 0$ let us define the following random variables for every step $i \geq 0$:

- $Q_{b,c}(u,v) = Q_{b,c}(u,v,i)$ is the set of vertices w such that the unmatched codegree of w and u is b and the unmatched codegree of w and v is c.

- $R_c(v) = R_c(v,i)$ is the set of vertices u such that the unmatched codegree of u and v is c.

- $S_c(u,v) = S_c(u,v,i)$ is the set of vertices $w \in N_U(v)$ such that the unmatched codegree of w and u, excluding v, is c.

Next we wish to define deterministic counterparts to these random variables. Here we use our heuristic that the unmatched graph is almost regular with degree $yn^{1/2}$ and that the codegrees are almost independent Poisson variables with expectation y^2. So for example for $Q_{b,c}$, we have

$$\sum_w \Pr\Big(\mathrm{codeg}_U(w,u)=b \text{ and } \mathrm{codeg}_U(w,v)=c\Big) \approx \left(\frac{e^{-y^2}y^{2b}}{b!}\right)\left(\frac{e^{-y^2}y^{2c}}{c!}\right)n = \frac{e^{-2y^2}y^{2b+2c}}{b!c!}n.$$

And for $S_c(u,v)$, note that

$$\sum_{w\in N_U(v)} \Pr\Big(\mathrm{codeg}_U(w,u)=c \text{ (excl. } v)\Big) \approx \left(\frac{e^{-y^2}y^{2c}}{c!}\right) yn^{1/2} = \frac{e^{-y^2}y^{2c+1}}{c!}n^{1/2}.$$

Reasoning in this way, we may define the following functions:

$$q_{b,c} = q_{b,c}(t) := \frac{e^{-2y^2}y^{2b+2c}}{b!c!}, \qquad r_c = r_c(t) := \frac{e^{-y^2}y^{2c}}{c!}, \qquad s_c = s_c(t) := \frac{e^{-y^2}y^{2c+1}}{c!}.$$

(Here we scale by an appropriate power of n.)

Observe that when $b=c=0$ we have $q_{0,0}=e^{-2y^2}$, $r_0=e^{-y^2}$, and $s_0=e^{-y^2}y$. Moreover, since for any $k\ge 0$ and $0\le x\le 1$, we have $e^{-x^2}x^k\le 1$, we obtain

$$q_{b,c} \le \frac{1}{b!c!}, \qquad r_c \le \frac{1}{c!}, \qquad s_c \le \frac{1}{c!}. \tag{2.2}$$

Define an “error function"

$$f(t) := \exp\left\{\frac{1000\log n}{\log\log n}\cdot t\right\} n^{-1/5}$$

and observe that for $0\le t\le \frac{1}{10000}\log\log n$ we have $n^{-1/5}\le f(t)\le n^{-1/10}$.

Now we define the “good event” at step i. For a given step i, let $\mathcal{E}_i$ be the event such that in $G=G(n,i)$ we have:

(i) *No huge codegree:* for all $u,v\in V$ we have

$$\mathrm{codeg}_G(u,v) \le \frac{3\log n}{\log\log n} =: \gamma(n).$$

(ii) *Dynamic concentration:* for all $u, v \in V$ and for every $j \leq i$,

- $d_G(v, j) \in \left(2t \pm n^{-1/4}\log^2 n\right) n^{1/2}$,
- $d_U(v, j) \in (y \pm f)\, n^{1/2}$,
- $|Q_{b,c}(u, v, j)| \in \left(q_{b,c} \pm f\right) n$,
- $|R_c(v, j)| \in (r_c \pm f)\, n$,
- $|S_c(u, v, j)| \in \left(s_c \pm (c+1)^{-1} f\right) n^{1/2}$,

where $a \pm b$ denotes the interval $[a - b, a + b]$, and the functions $y, f, q_{b,c}, r_c$, and s_c are evaluated at the point $t(j)$.

Note that if the event $\mathcal{E}_i$ fails (no matter if it fails due to condition (i) or (ii)), then $\mathcal{E}_{i'}$ also fails for all $i' > i$. We now show that the first condition of the event $\mathcal{E}_i$ holds w.h.p. for every i under consideration. We use the asymptotic equivalence of the models $G(n, m)$ and $G(n, p)$ (where $p = m/\binom{n}{2}$) and the fact that (i) is a monotone graph property (see Proposition 1.15 in [44]). Now to see that this holds w.h.p. we calculate the expected number of pairs u, v with at least $\gamma(n)$ common neighbors. At step i the number of edges we have added is at most $n^{3/2}(\log\log n)/10000$. Thus it is enough to show that (i) holds w.h.p. in $G(n, p)$ where $p \leq n^{-1/2}(\log\log n)/5000$. Now, the expected number of pairs of vertices in $G(n, p)$ with codegree at least $\gamma(n)$ is at most

$$n^2\binom{n}{\gamma(n)}p^{2\gamma(n)} \leq n^2\left(\frac{enp^2}{\gamma(n)}\right)^{\gamma(n)} \leq n^2\left(\frac{(\log\log n)^3}{\log n}\right)^{\gamma(n)} = e^{2\log n}\left(\frac{(\log\log n)^3}{\log n}\right)^{\gamma(n)}$$
$$\leq e^{2\log n}\left(\frac{1}{(\log n)^{5/6}}\right)^{\gamma(n)} = e^{-(\log n)/2} = o(1).$$

In Sections 2.1.5-2.1.8 we prove that (ii) also holds w.h.p.

Since unmatched codegrees are so important to this process, in our analysis we will frequently need to know, for a given pair of vertices u, v, how many possible choices for the next edge e_i would increase $\text{codeg}_U(u, v)$. We denote by $A(u, v) = A(u, v, i)$ the set of such possibilities for e_i, which we will now estimate with the assumption that the good event holds. Suppose w is a neighbor of u (resp. v). If we add the edge $\{v,w\}$ (resp. $\{u,w\}$), it may actually be removed in the same step since it might create a triangle. So as long as we ignore the $\tilde{O}(1)$ vertices in codeg(u,v), the number of w such that vw is not removed is $|S_0(u,v)|$ (resp. $|S_0(v,u)|$). Thus we have

$$|A(u,v)| := |S_0(v,u)| + |S_0(u,v)| - \tilde{O}(1).$$

So for $\alpha(t) := 2s_0$ and all $j \le i$ we have

$$|A(u,v)| \in (\alpha \pm 3f)n^{1/2}$$

for evaluation at $t(j)$. We use $3f$ as the error function here so we can ignore the $\tilde{O}(1)$ term from the definition of $A(u,v)$ (recall that $f = \Omega(n^{-1/5})$). We also wish to ignore any edges that are in M already. But any such candidates would be neighbors with both u and v in G, meaning that we can ignore these $\tilde{O}(1)$ edges. Note that since $s_0 \le 1$ we have $\alpha \le 2$.

In our analysis we will also need to estimate, for a given unmatched edge $e = \{u,v\} \in U(i)$, a “count" related to the possibility that e becomes matched in the next step. We put quotes around “count" because actually we need a weighted count: for each possibility for $e_i = \{v,w\}$ (or $\{u,w\}$) that might result in e becoming matched, we weight it by the probability $1/(c+1)$ that the triangle selected to go into $M(i+1)$ is the triangle containing e, where c is the unmatched codegree of w and v (resp. u), excluding u (resp. v). To this end, we define $B(u,v)$ to be the random variable such that

$$B(u,v) := |A(u,v)| + \sum_{c=1}^{\gamma(n)} \frac{1}{c+1}(|S_c(u,v)| + |S_c(v,u)|).$$

Here we want to avoid counting the $\tilde{O}(1)$ edges that might be in M already. This, however, is accounted for in the $\tilde{O}(1)$ term in $A(u,v)$. In addition, notice that for $e = \{u,v\}$, the probability that e becomes matched in step i is

$$B(u,v)\cdot\frac{2}{n^2}(1\ +\ \tilde{O}(n^{-1/2})).$$

(Recall that the probability of adding e_i is $2/n^2(1\ +\ \tilde{O}(n^{-1/2}))$.) Now define

$$\beta(y) := 2e^{-y^2}\sum_{c=0}^{\infty}\frac{y^{2c+1}}{(c+1)!} = \begin{cases} 0, & y=0 \\ 2y^{-1}(1-e^{-y^2}), & \text{otherwise.}\end{cases}$$

It is easy to check that β is continuous and twice differentiable for $y \geq 0$. In addition, for $y > 0$,

$$\beta(y) = 2y^{-1}(1-e^{-y^2}) \leq 2y^{-1}(1-(1-y^2)) = 2y;$$

therefore when $0 \leq y \leq \zeta$, we have

$$\beta(y) \leq 2\zeta < 2. \tag{2.3}$$

Then from the dynamic concentration of $S(u,v)$ and $\sum_{c=0}^{\infty}(c+1)^{-2} = \pi^2/6 < 2$, we have

$$\begin{aligned} B(u,v) &= \sum_{c=0}^{\gamma(n)}\frac{1}{c+1}(|S_c(u,v)|+|S_c(v,u)|)-\tilde{O}(1) \\ &\leq \sum_{c=0}^{\gamma(n)}\left(\frac{2e^{-y^2}y^{2c+1}}{(c+1)!}+\frac{2f}{(c+1)^2}\right)n^{1/2}-\tilde{O}(1) \\ &= n^{1/2}\sum_{c=0}^{\infty}\frac{2e^{-y^2}y^{2c+1}}{(c+1)!}+n^{1/2}\sum_{c=0}^{\gamma(n)}\frac{2f}{(c+1)^2}+\tilde{O}(1) \qquad (2.4) \\ &\leq \kappa(y)n^{1/2}+4fn^{1/2}. \end{aligned}$$

A similar argument using the lower bound for $|S_c(u,v)|$ gives us a lower bound for

$B(u,v)$. Thus

$$B(x,y) \in (\beta \pm 4f)n^{1/2},$$

with evaluation at $t(j)$.

Then straightforward, but somewhat tedious, calculations show that the above functions satisfy the following differential equations, where $q'_{b,c}$, r'_c and s'_c denote derivatives of $q_{b,c}$, r_c and s_c as functions of t:

$$q'_{b,c} = 2q_{b-1,c}\alpha + 2q_{b,c-1}\alpha + 4(b+1)\beta q_{b+1,c} + 4(c+1)\beta q_{b,c+1} - 4q_{b,c}(\alpha + b\beta + c\beta), \tag{2.5}$$

$$r'_c = 2r_{c-1}\alpha + 4(c+1)\beta r_{c+1} - (2\alpha + 4c\beta)r_c, \tag{2.6}$$

$$s'_c = 2s_{c-1}\alpha + 4(c+1)\beta s_{c+1} + 2q_{c,0} - 2(\alpha + 2c\beta + \beta)s_c. \tag{2.7}$$

These differential equations can be viewed as idealized one-step changes in the random variables $Q_{b,c}(u,v)$, $R_c(v)$, and $S_c(u,v)$. Each of these variables counts copies of some type of substructure, and these copies can be created or destroyed by the process when we add or remove edges. Equations (2.5)–(2.7) can be understood as expressing the one-step changes in the random variables in terms of these creations and deletions, on average. We will ultimately use these differential equations to argue that the random variables stay close to their deterministic counterparts.

2.1.5 Tracking $d_U(v,j)$

First observe that Chernoff's bound implies that w.h.p.

$$d_G(v,j) \in \left(2t \pm n^{-1/4}\log^2 n\right)n^{1/2}.$$

Moreover, in order to estimate $d_U(v,j)$ it suffices to track $d_M(v,j)$.

We define the natural filtration $\mathscr{F}_i$ to be the history of the process up to step i. In particular, conditioning on $\mathscr{F}_i$ tells us the current state of the process. Assuming we are in the event $\mathscr{E}_{i-1}$, we calculate the expected one-step change of the matched degree, conditional on $\mathscr{F}_{i-1}$, namely,

$$\mathrm{E}[\Delta d_M(v,i)|\mathscr{F}_{i-1}] = \mathrm{E}[d_M(v,i) - d_M(v,i-1)|\mathscr{F}_{i-1}].$$

We have already revealed $i-1$ edges. Now we reveal a new edge e_i. Note that $d_M(v)$ is nondecreasing. If $e_i \subseteq N_U(v)$, then $d_M(v)$ could increase by 2 or not increase at all. For a fixed vertex u in $N_U(v)$, if the edge $e_i = \{u,w\}$ for some $w \in N_U(v)$ with $\mathrm{codeg}_U(u,w) = c+1$, then $d_M(v)$ increases by 2 with probability $1/(c+1)$. The number of such w can be counted with $|S_c(u,v)|$. Notice that, by taking a sum over $u \in N_U(v)$, we double count such w. Finally, if e_i is the edge $\{v,u\}$ for some vertex u not in $N_U(v)$ such that $\mathrm{codeg}_U(u,v) > 0$, then $d_M(v)$ increases by 2. Hence, we have

$$\begin{aligned}
&\mathrm{E}[\Delta d_M(v,i)|\mathscr{F}_{i-1}]\\
&\quad= \left[\sum_{u\in N_U(v)}\sum_{c=0}^{\gamma(n)} 2\cdot\frac{1}{2}(c+1)^{-1}\cdot|S_c(u,v,i-1)| + \sum_{c=1}^{\gamma(n)} 2\cdot|R_c(v,i-1)|\right]\frac{2}{n^2}(1+\tilde{O}(n^{-1/2}))\\
&\quad\le \left[2(y+f)\sum_{c=0}^{\gamma(n)}(c+1)^{-1}(s_c+(c+1)^{-1}f) + 4\sum_{c=1}^{\gamma(n)}(r_c+f)\right]n^{-1}(1+\tilde{O}(n^{-1/2}))\\
&\quad= \left[2\sum_{c=0}^{\gamma(n)}(c+1)^{-1}ys_c + 4\sum_{c=1}^{\gamma(n)} r_c + \left(-4+\sum_{c=0}^{\gamma(n)}\left(\frac{2y}{(c+1)^2}+\frac{2s_c}{c+1}+4\right)\right)f\right.\\
&\qquad\qquad \left. + f^2\cdot\sum_{c=0}^{\gamma(n)}\frac{2}{(c+1)^2}\right]n^{-1}(1+\tilde{O}(n^{-1/2}))
\end{aligned}$$

where the functions y and f are evaluated at point $t(i-1)$. Now,

$$\sum_{c=0}^{\gamma(n)}\frac{ys_c}{c+1} = \sum_{c=0}^{\gamma(n)}\frac{y}{c+1}\left(\frac{e^{-y^2}y^{2c+1}}{c!}\right) = e^{-y^2}\sum_{c=0}^{\infty}\frac{y^{2(c+1)}}{(c+1)!} + O(n^{-2}) = 1-e^{-y^2}+O(n^{-2})$$

where the second equality uses the fact that for $c \geq \gamma(n)$ we have

$$c! = \exp\{(1+o(1))c\log c\} \geq \exp\{(3+o(1))\log n\},$$

and so

$$\sum_{c=\gamma(n)+1}^{\infty} \frac{y^{2(c+1)}}{(c+1)!} < n^{-3+o(1)} = O(n^{-2}).$$

In a similar manner,

$$\sum_{c=1}^{\gamma(n)} r_c = \sum_{c=1}^{\gamma(n)} \frac{e^{-y^2} y^{2c}}{c!} = e^{-y^2}\left(\sum_{c=0}^{\infty} \frac{y^{2c}}{c!} - 1\right) + O(n^{-2}) = 1 - e^{-y^2} + O(n^{-2}).$$

Since $y \leq 1$ and $s_c \leq 1/c!$, we may estimate the coefficient of f as follows:

$$\begin{aligned} -4 + \sum_{c=0}^{\gamma(n)}\left(\frac{2y}{(c+1)^2} + \frac{2s_c}{c+1} + 4\right) &\leq -4 + 2\sum_{c=0}^{\infty}\frac{1}{(c+1)^2} + \sum_{c=0}^{\infty}\frac{2}{(c+1)!} + 4\gamma(n) \\ &= -4 + 2\cdot\frac{\pi^2}{6} + 2(e-1) + 4\gamma(n) \leq -4+4+4+4\gamma(n) = 4 + 4\gamma(n). \end{aligned}$$

Further, since $2(1-e^{-y^2}) = y\beta(y) \leq 2$ from (2.3) and $f^2 \leq n^{-1/10} f$, we may write

$$\mathrm{E}[\Delta d_M(v,i)|\mathscr{F}_{i-1}] \leq \left[6 - 6e^{-y^2} + 4\gamma(n)f + O(f)\right]n^{-1} + \tilde{O}(n^{-3/2}). \tag{2.8}$$

Define variables

$$D^{\pm}(v) = D^{\pm}(v,i) := \begin{cases} d_M(v,i) - (2t(i) - y(t(i)) \pm f(t(i)))n^{1/2} & \text{if } \mathscr{E}_{i-1} \text{ holds} \\ D^{\pm}(v, i-1) & \text{otherwise.} \end{cases}$$

We will show that the probability of $D^+(v)$ (resp. $D^-(v)$) becoming positive (resp. negative) is very small, thus verifying the upper (resp. lower) bound of the matched degree of v at step i. This implies that we obtain the desired bound on the unmatched degree of v in (ii) of $\mathscr{E}_i$. To do this, we show that the sequence $D^+(v)$ is a supermartingale and use an appropriate martingale concentration inequality. Symmetric calculations can be used for $D^-(v)$.

We first apply Taylor's theorem 1.3.1 to approximate the change in the deterministic function by its derivative. We extend $t(i) := \frac{i}{n^{3/2}}$ to the real numbers. Then set $h(t) := 2t - y(t) + f(t)$ and $H(i) := h(t(i))$. Then,

$$H(i) - H(i-1) = \frac{dH(i-1)}{di} + \frac{1}{2}\frac{d^2H(\omega)}{di^2} = \frac{dH(i-1)}{dt}n^{-3/2} + \frac{1}{2}\frac{d^2H(\omega)}{di^2}$$

where $\omega \in [i-1, i]$. But

$$\frac{d^2H(i)}{di^2} = \frac{d}{di}\left(\frac{dH(i)}{dt}n^{-3/2}\right) = \left(-\frac{d^2y(t(i))}{dt^2} + \frac{d^2f(t(i))}{dt^2}\right)n^{-3}$$

Furthermore, by (2.1) we get that $|d^2y(t)/dt^2| \le 24$. Also,

$$\frac{d^2f(t)}{dt^2} = \left(\frac{1000\log n}{\log\log n}\right)^2 \exp\left\{\frac{1000\log n}{\log\log n}\cdot t\right\} n^{-1/5} = \left(\frac{1000\log n}{\log\log n}\right)^2 f(t).$$

Thus, $d^2H(\omega)/dt^2 = O(n^{-2})$ and

$$H(i) - H(i-1) = \left(2 - \frac{dy(t(i-1))}{dt} + \frac{df(t(i-1))}{dt}\right)n^{-3/2} + O(n^{-2}). \tag{2.9}$$

Now if we are in $\mathscr{E}_{i-1}$, then (2.8) and (2.9) for $t = t(i-1)$ imply

$$\begin{aligned}\mathrm{E}\left[\Delta D^+(v,i)|\mathscr{F}_{i-1}\right] &\le \left(-\frac{df}{dt} + 4\gamma(n)f + O(f)\right)n^{-1} + \tilde{O}(n^{-3/2}) \\ &= \left[-\frac{1000\log n}{\log\log n} + \frac{12\log n}{\log\log n} + O(1)\right]f n^{-1} + \tilde{O}(n^{-3/2}) \le 0,\end{aligned}$$

showing that the sequence $D^+(v,i)$ is a supermartingale.

We now show that the probability of $D^+(v)$ becoming positive is small, implying that there is only a small probability that $d_M(v)$ strays from our desired bounds. We use the martingale concentration inequality due to Freedman, Lemma 1.3.3. Observe that $|\Delta d_M(v,i)| = O(1)$, since at most two edges adjacent to v can become matched at step i. Moreover, due to (2.9), $|\Delta(2t(i) - y(t(i)) + f(t(i)))n^{1/2}| = O(n^{-1})$ trivially. The triangle inequality thus implies that $|\Delta D^+(v,i)| = O(1)$. Also since the variable $d_M(v,i)$ is nondecreasing we have $\mathrm{E}[|\Delta d_M(v,i)||\mathscr{F}_{i-1}] = \mathrm{E}[\Delta d_M(v,i)|\mathscr{F}_{i-1}] = O(n^{-1})$ by (2.8). Hence, the triangle inequality yields $\mathrm{E}[|\Delta D^+||\mathscr{F}_{i-1}] = O(n^{-1})$. So the one-step variance is

$$\mathbf{Var}[\Delta D^+|\mathscr{F}_{i-1}] \leq \mathrm{E}[(\Delta D^+)^2|\mathscr{F}_{i-1}] \leq O(1) \cdot \mathrm{E}[|\Delta D^+||\mathscr{F}_{i-1}] = O(n^{-1}).$$

Therefore, for Freedman's inequality we use $b = O(n^{-1}) \cdot O(n^{3/2} \log\log n) = \tilde{O}(n^{1/2})$. The "bad" event here is the event that we have $D^+(v,i) > 0$, and since $D^+(v,0) = -n^{3/10}$ we set $\lambda = n^{3/10}$. Then, applying Lemma 1.3.3 with $\lambda = n^{3/10}$, $b = \tilde{O}(n^{1/2})$ and $C = \tilde{O}(1)$ yields that the failure probability is at most

$$\exp\left\{-\frac{n^{3/5}}{\tilde{O}(n^{1/2}) + \tilde{O}(1) \cdot n^{3/10}}\right\},$$

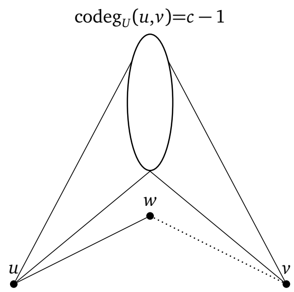
which is small enough to beat a union bound over all vertices.

Using symmetric calculations one can apply Freedman's inequality to the supermartingale $-D^-(v,i)$ to show that the "bad" event $D^-(v,i) < 0$ does not occur w.h.p.

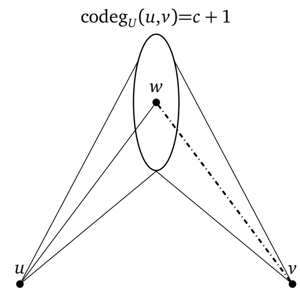
2.1.6 Tracking $|R_c(v)|$

Since the tracking of $|R_c(v)|$ seems especially illustrative, we eschew alphabetical order and estimate $\mathrm{E}[\Delta|R_c(v,i)||\mathscr{F}_{i-1}]$ next. Because $|R_c(v,i)|$ counts the number of vertices u such that $\mathrm{codeg}_U(u,v) = c$, we are interested to know how these codegree functions can increase or decrease.

We first focus on the cases that occur most often. Note that $\mathrm{codeg}_U(u,v)$ can increase by at most one at any step. Hence, we can increase $\mathrm{codeg}_U(u,v)$ by one if $e_i = \{x,y\}$ such that $x = u$ (resp. $x = v$), y is adjacent to v (resp. u), and e_i does not create a triangle with other edges in U (see Figure 2.1a). In the event $\mathscr{E}_i$, the number of such edges e_i is $|A(u,v)|$. Thus, for the positive contribution, we would take the sum of $|A(u,v)|$ over all u in $R_{c-1}(v)$. But



(a) Ways $u \in R_{c-1}(v, i-1)$ could join $R_c(v, i)$ with the addition of the dotted edge.



(b) Ways $u \in R_{c+1}(v, i-1)$ could join $R_c(v, i)$ with the removal of the dashed edge.

Figure 2.1: Cases Considered in the Expected One-Step Change of $|R_c(v, i)|$.

$|R_c(v)|$ can also decrease if $\text{codeg}_U(u, v)$ increases in this way for $u \in R_c(v)$. Thus the negative contribution can be described by taking the sum over $|A(u, v)|$ for all u in $R_c(v)$.

The unmatched codegree can also decrease. This can occur when an edge $\{u, w\}$ or $\{v, w\}$ becomes matched, for some w in the unmatched codegree of u and v (see Figure 2.1b). Recall from our previous discussion that $B(w, v)$ and $B(w, u)$ can be used to track these changes. So we obtain a positive contribution by taking the sum of $B(w, v) + B(w, u)$ over all u in $R_{c+1}(v)$ and all w in the unmatched codegree of u and v. If u comes from $R_c(v)$, though, the contribution to the unmatched codegree would be negative.

There are two cases affecting $R_c(v, i-1)$ that occur less frequently. The first such way is to create a triangle by having e_i be in the common neighborhood of u and v. The second way is for the edge e_i to contain u and v. Either way when we sum over all u, the total change is $\tilde{O}(n^{-1})$.

Thus, we obtain

$$
\begin{aligned}
&\mathrm{E}\left[\Delta |R_c(v,i)| | \mathscr{F}_{i-1}\right] \\
=&\Bigg[\sum_{u\in R_{c-1}(v)} |A(u,v)| + \sum_{\substack{u\in R_{c+1}(v)\\ w\in \mathrm{codeg}_U(u,v)}} [B(w,v)+B(w,u)] \\
&- \sum_{u\in R_c(v)} |A(u,v)| - \sum_{\substack{u\in R_c(v)\\ w\in \mathrm{codeg}_U(u,v)}} [B(w,v)+B(w,u)]\Bigg]\cdot \frac{2}{n^2}(1+\tilde{O}(n^{-1/2}))+\tilde{O}(n^{-1}). \quad (2.10)
\end{aligned}
$$

Estimating this gives

$$
\begin{aligned}
&\mathrm{E}\left[\Delta |R_c(v,i)| | \mathscr{F}_{i-1}\right] \\
&\le \Big[(r_{c-1}+f)(\alpha+3f)+2(r_{c+1}+f)(c+1)(\beta+4f) \\
&\qquad -(r_c-f)(\alpha-3f)-2(r_c-f)\cdot c\cdot(\beta-4f)\Big]2n^{3/2}\cdot n^{-2}+\tilde{O}(n^{3/2}\cdot n^{-5/2})+\tilde{O}(n^{-1}) \\
&= \Big[2r_{c-1}\alpha-(2\alpha+4c\beta)r_c+4(c+1)\beta r_{c+1} \qquad (2.11) \\
&+8c\beta f+\Big(6r_{c-1}+2(8c+3)r_c+16(c+1)r_{c+1}+4\alpha+4\beta\Big)f+16f^2\Big]n^{-1/2}+\tilde{O}(n^{-1})
\end{aligned}
$$

where all functions are evaluated at point $t(i-1)$.

Observe that $8c\beta f \le 20cf$ and from the bounds on r_c, α and β in (2.2) and (2.3), we get that all other terms with f are $O(f)$. Further the f^2 terms are also $O(f)$ since $f^2 \le n^{-1/10} f$. Lastly observe that (2.11) is r_c' by (2.6). Thus, $\mathrm{E}[\Delta|R_c(v,i)||\mathscr{F}_{i-1}]$ is at most

$$
\Big[r_c'+20cf+O(f)\Big]n^{-1/2}+\tilde{O}(n^{-1}). \qquad (2.12)
$$

Now we define variables

$$
R_c^{\pm}(v)=R_c^{\pm}(v,i):=\begin{cases} |R_c(v,i)|-(r_c(t(i))\pm f(t(i)))n & \text{if } \mathscr{E}_{i-1} \text{ holds} \\ R_c^{\pm}(v,i-1) & \text{otherwise.}\end{cases}
$$

We show that $R^+(v)$ are supermartingales. We extend $t(i) := i/n^{3/2}$ to the real numbers. Set $h(t) := r_c(t) + f(t)$ and $H(t) := h(t(i))$. Then Taylor's theorem with the Lagrange remainder gives us

$$H(i) - H(i-1) = \frac{dH(i-1)}{di} + \frac{1}{2}\frac{d^2H(\omega)}{di^2} = \frac{dH(i-1)}{dt}n^{-3/2} + \frac{1}{2}\frac{d^2H(\omega)}{di^2}$$

where $\omega \in [i-1, i]$. But

$$\frac{d^2H(i)}{di^2} = \frac{d}{di}\left(\frac{dH(i)}{dt}n^{-3/2}\right) = \left(\frac{d^2r_c(t(i))}{dt^2} + \frac{d^2f(t(i))}{dt^2}\right)n^{-3}$$

Furthermore, this is $O(n^{-2})$. To see this, note that

$$\frac{d^2r_c}{dt^2} = -\frac{(2y^4 - (4c+1)y^2 + 2c^2 - c)y^{2c-2}(8e^{-y^2} - 12e^{-2y^2})}{c!},$$

and so $|d^2r_c/dt^2| = O(1)$ along with $|d^2f/dt^2| = O(n^{-2})$. Hence,

$$(H(i) - H(i-1))n = \left[\frac{dr_c(t(i-1))}{dt} + \frac{df(t(i-1))}{dt}\right]n^{-1/2} + O(n^{-2}).$$

Therefore for $t = t(i-1)$ due to (2.12) we get

$$\begin{aligned}\mathrm{E}\left[\Delta R_c^+(v,i)|\mathscr{F}_{i-1}\right] &\leq \left[-\frac{df}{dt} + 20cf + O(f)\right]n^{-1/2} + \tilde{O}(n^{-1}) \\ &\leq \left[-\frac{1000\log n}{\log\log n} + \frac{60\log n}{\log\log n} + O(1)\right]fn^{-1/2} + \tilde{O}(n^{-1}) \leq 0.\end{aligned}$$

Now observe that $|\Delta|R_c(v)|| = \tilde{O}(n^{1/2})$. Indeed, if the new edge e_i has one vertex at v and the other at say x, then this only affects the codegree of v with the $\tilde{O}(n^{1/2})$ many neighbors of x. On the other hand if e_i is not incident with v then v loses at most two unmatched edges, say $\{v,x\}$ and $\{v,y\}$, in which case only the codegree of v with the $\tilde{O}(n^{1/2})$ neighbors of x and y can be affected. Thus, we also have $|\Delta R_c^+(v)| = \tilde{O}(n^{1/2})$, since r_c and f have much smaller one-step changes. Now we would like to bound $\mathrm{E}[|\Delta|R_c(v)|||\mathscr{F}_{i-1}]$, so we will re-examine (2.10). There are positive and negative contributions to $\Delta R_c(v)$, and of course (2.10) represents the expected positive contributions minus the expected negative contributions. Now by the triangle inequality $|\Delta|R_c(v)||$ is at most the sum of the positive and negative contributions, and so

$$\begin{aligned}\mathrm{E}[|\Delta|R_c(v)|||\mathscr{F}_{i-1}] &\leq \Bigg[\sum_{u\in R_{c-1}(v)} |A(u,v)| + \sum_{\substack{u\in R_{c+1}(v)\\ w\in \mathrm{codeg}_U(u,v)}} [B(w,v)+B(w,u)] \\ &\quad + \sum_{u\in R_c(v)} |A(u,v)| + \sum_{\substack{u\in R_c(v)\\ w\in \mathrm{codeg}_U(u,v)}} [B(w,v)+B(w,u)]\Bigg]\cdot \frac{2}{n^2} + \tilde{O}(n^{-1}) \\ &= O(n^{-1/2}), \end{aligned} \tag{2.13}$$

since each term in (2.11) is $O(n^{-1/2})$ as demonstrated in (2.12). Thus,

$$\mathrm{E}[|\Delta R_c^+(v)||\mathscr{F}_{i-1}] \leq \mathrm{E}[|\Delta|R_c(v)|||\mathscr{F}_{i-1}] + |\Delta(r_c(t)+f(t))|n = O(n^{-1/2}),$$

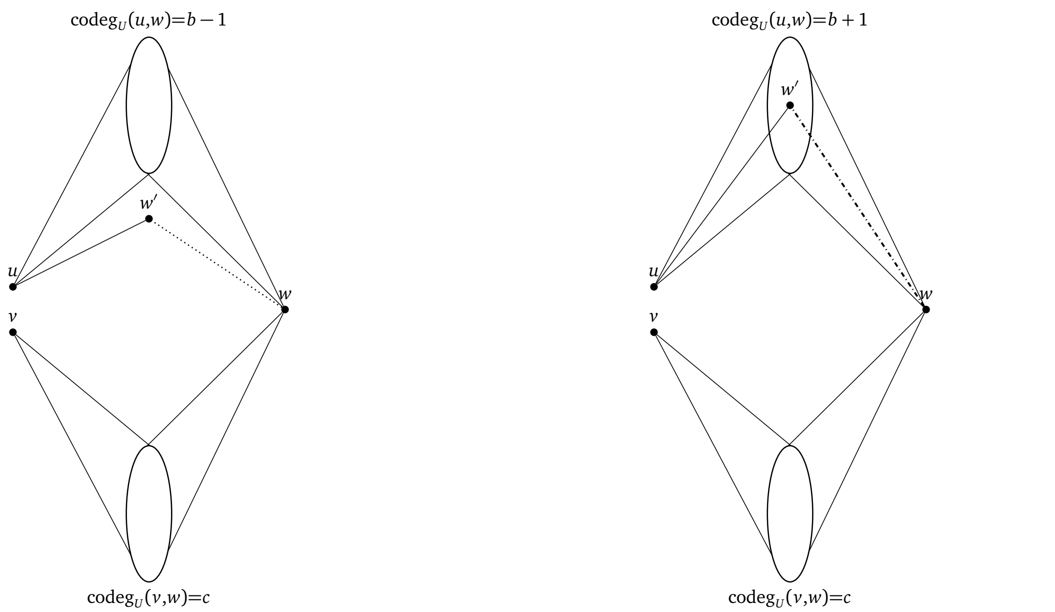
and hence the one-step variance is

$$\mathbf{Var}[\Delta R_c^+(v)|\mathscr{F}_{i-1}] \leq \mathrm{E}[(\Delta R_c^+(v))^2|\mathscr{F}_{i-1}] = \tilde{O}(n^{1/2})\cdot \mathrm{E}[|\Delta R_c^+(v)||\mathscr{F}_{i-1}] = \tilde{O}(1).$$

The “bad" event here is the event that $R_c^+(v,i) > 0$. Since $R_c^+(v,0) = n^{4/5}$ we set $\lambda = n^{4/5}$. Then, Lemma 1.3.3 applied with $\lambda = n^{4/5}$, $b = \tilde{O}(n^{3/2})$ and $C = \tilde{O}(n^{1/2})$ yields that the failure probability is at most

$$\exp\left\{-\frac{\tilde{O}(n^{8/5})}{\tilde{O}(n^{3/2}) + \tilde{O}(n^{1/2})\cdot n^{4/5}}\right\},$$

which is small enough to beat a union bound over all vertices as well as possible values of c.



(a) Ways $w \in Q_{b-1,c}(u,v,i-1)$ Could Join $Q_{b,c}(u,v,i)$ with the Addition of the Dotted Edge.

(b) Ways $w \in Q_{b+1,c}(u,v,i-1)$ Could Join $Q_{b,c}(u,v,i)$ with the Removal of the Dashed Edge.

Figure 2.2: Cases Considered in the Expected One-Step Change of $|Q_{b,c}(u,v,i)|$.

2.1.7 Tracking $|Q_{b,c}(u,v)|$

We continue by calculating $\mathrm{E}\left[\Delta|Q_{b,c}(u,v,i)||\mathscr{F}_{i-1}\right]$. As before, we consider positive and negative contributions, beginning with the positive ones. First, for $w \in Q_{b-1,c}(u,v)$, e_i could increase $\mathrm{codeg}_U(u,w)$ up to b from $b-1$ (see Figure 2.2a). This situation is accounted with $|A(u,w)|$. Second, we could have a symmetric case for $w \in Q_{b,c-1}(u,v)$ tracked by $|A(v,w)|$. Third, for $w \in Q_{b+1,c}(u,v)$ an edge $\{u,w'\}$ (or $\{w,w'\}$) could become matched for $w' \in \mathrm{codeg}_U(u,w)$. This would result in $\mathrm{codeg}_U(u,w) = b$ (see Figure 2.2b). This can be tracked with $B(u,w') + B(w,w')$. Fourth, a similar case to the third can occur but with $w \in Q_{b,c+1}(u,v)$.

For the negative contribution, all cases are symmetric. The only difference is that all w are taken from $Q_{b,c}(u,v)$ and thus these gains or losses of edges destroy the existing structure.

There are also three unlikely cases that affect $|Q_{b,c}(u,v)|$. In each case, we consider $w \in Q_{b',c'}(u,v)$ for with b' and c' appropriately chosen from $\{b-1,b,b+1\}$ and $\{c-1,c,c+1\}$, respectively. There are $\tilde{O}(n)$ such w. First, the edge e_i could be in the common neighborhood of u, w or of v, w. Second, e_i could be either the edge $\{v, w\}$ or $\{u, w\}$, reducing the codegree of the vertices in that edge. Third, e_i could have one vertex in the common neighborhood of u, w and the other in the common neighborhood of v, w. In all of these cases, we have a change of $\tilde{O}(1)$ and hence together with the edge probability, we have a change of $\tilde{O}(n^{-1})$.

Consequently,

$$\begin{aligned}
&\mathrm{E}\left[\Delta|Q_{b,c}(u,v,i)||\mathscr{F}_{i-1}\right] \\
&\quad = \Bigg[\sum_{w\in Q_{b-1,c}(u,v)} |A(u,w)| + \sum_{w\in Q_{b,c-1}(u,v)} |A(v,w)| + \sum_{\substack{w\in Q_{b+1,c}(u,v)\\ w'\in \mathrm{codeg}(u,w)}} (B(u,w')+B(w,w')) \\
&\qquad + \sum_{\substack{w\in Q_{b,c+1}(u,v)\\ w'\in \mathrm{codeg}(v,w)}} (B(v,w')+B(w,w')) - \sum_{w\in Q_{b,c}(u,v)} |A(u,w)| - \sum_{w\in Q_{b,c}(u,v)} |A(v,w)| \\
&\qquad - \sum_{\substack{w\in Q_{b,c}(u,v)\\ w'\in \mathrm{codeg}(u,w)}} (B(u,w')+B(w,w')) - \sum_{\substack{w\in Q_{b,c}(u,v)\\ w'\in \mathrm{codeg}(v,w)}} (B(v,w')+B(w,w')) \Bigg] \\
&\qquad\qquad \times \frac{2}{n^2}(1+\tilde{O}(n^{-1/2})) + \tilde{O}(n^{-1}) \qquad (2.14) \\
&\quad \le \Big[(q_{b-1,c}+f)(\alpha+3f) + (q_{b,c-1}+f)(\alpha+3f) + 2(\beta+4f)(q_{b+1,c}+f)(b+1) \\
&\qquad + 2(\beta+4f)(q_{b,c+1}+f)(c+1) - 2(q_{b,c}-f)(\alpha-3f) \\
&\qquad - 2(b+c)(\beta-4f)(q_{b,c}-f) \Big] \cdot 2n^{-1/2} + \tilde{O}(n^{-1})
\end{aligned}$$

$$\begin{aligned}
= \Big[& 2q_{b-1,c}\alpha + 2q_{b,c-1}\alpha + 4q_{b+1,c}(b+1)\beta + 4q_{b,c+1}(c+1)\beta - 4q_{b,c}(\alpha+b\beta+c\beta) \qquad (2.15) \\
& + 8\beta(b+c)f + \Big(6q_{b-1,c} + 6q_{b,c-1} + 16(b+1)q_{b+1,c} + 16(c+1)q_{b,c+1} \\
& \quad + 16q_{b,c}(b+c) + 12q_{b,c} + 8(\alpha+\beta) \Big) f + 32f^2 \Big] \cdot n^{-1/2} + \tilde{O}(n^{-1}).
\end{aligned}$$

We simplify this. First notice that $8\beta(b+c)f \le 20(b+c)f$ and from the bounds $q_{b,c} \le 1/(b!c!)$

and $\alpha, \beta \leq 3$, we see that all other multiples of f are $O(f)$. Also recall that $f^2 \leq n^{-1/10} f = O(f)$. Finally, observe that by (2.5), line (2.15) is $q'_{b,c}$. Therefore, $\mathrm{E}\big[\Delta|Q_{b,c}(u,v,i)||\mathscr{F}_{i-1}\big]$ is at most

$$\Big[q'_{b,c} + 20(b+c)f + O(f)\Big]n^{-1/2} + \tilde{O}(n^{-1}). \tag{2.16}$$

Now we define variables

$$Q^{\pm}_{b,c}(u,v) = Q^{\pm}_{b,c}(u,v,i) := \begin{cases} |Q_{b,c}(u,v,i)| - (q_{b,c}(t(i)) \pm f(t(i)))n & \text{if } \mathscr{E}_{i-1} \text{ holds} \\ Q^{\pm}_{b,c}(u,v,i-1) & \text{otherwise.} \end{cases}$$

Now we demonstrate that $Q^{+}_{b,c}(u,v)$ is a supermartingale. Extend $t(i) := i/n^{3/2}$ to the real numbers. Then let $h(t) := q_{b,c}(t) + f(t)$ and $H(i) := h(t(i))$. By applying Taylor's theorem, we obtain

$$H(i) - H(i-1) = \frac{dH(i-1)}{di} + \frac{1}{2}\frac{d^2H(\omega)}{di^2} = \frac{dH(i-1)}{dt}n^{-3/2} + \frac{1}{2}\frac{d^2H(\omega)}{di^2},$$

where $\omega \in [i-1, i]$. But

$$\frac{d^2H(i)}{di^2} = \frac{d}{di}\left(\frac{dH(i)}{dt}n^{-3/2}\right) = \left(\frac{d^2q_{b,c}(t(i))}{dt^2} + \frac{d^2f(t(i))}{dt^2}\right)n^{-3}.$$

Furthermore,

$$\frac{d^2q_{b,c}}{dt^2} = -\frac{16\left(e^{y^2} - \frac{3}{2}\right)\left(4y^4 - 4(b+c+1)y^2 + (b+c)\left(b+c-\frac{1}{2}\right)\right)e^{-3y^2}y^{2b+2c-2}}{b!c!},$$

so $|d^2q_{b,c}/dt^2| = O(1)$. Also, we have $|d^2f(t)/dt^2| = O(n^{-2})$. Hence,

$$(H(i) - H(i-1))n = \left[\frac{dq_{b,c}(t(i-1))}{dt} + \frac{df(t(i-1))}{dt}\right]n^{-1/2} + O(n^{-2}).$$

Therefore from (2.16) with $t = t(i-1)$, we have

$$\mathrm{E}\Big[\Delta Q^+_{b,c}(u,v,i)|\mathscr{F}_{i-1}\Big] \le \Big[-\frac{df}{dt} + 20(b+c)f + O(f)\Big]n^{-1/2} + \tilde{O}(n^{-1})$$
$$\le \Big[-\frac{1000\log n}{\log\log n} + \frac{120\log n}{\log\log n} + O(1)\Big]f n^{-1/2} + \tilde{O}(n^{-1}) \le 0.$$

Let us consider the effect on $|Q_{b,c}(u,v)|$ by removing one edge e from the unmatched graph. If e is incident with u, say $e = \{u,x\}$, then the only vertices $w \in Q_{b,c}(u,v)$ that could be affected are in the set $\{x\} \cup N(x)$ which has size $\tilde{O}(n^{1/2})$. Similarly if e is incident with v. If e is not incident with u,v then the only affected $w \in Q_{b,c}(u,v)$ would be the endpoints of e. Thus we have $|\Delta Q_{b,c}(u,v)| = \tilde{O}(n^{1/2})$, and also $|\Delta Q^+_{b,c}(u,v)| = \tilde{O}(n^{1/2})$ because the deterministic terms in $Q^+_{b,c}(u,v)$ have much smaller one-step changes. We can also see that $\mathrm{E}[|\Delta|Q_{b,c}(u,v)|||\mathscr{F}_{i-1}] = O(n^{-1/2})$ by another argument analogous to the one used to justify (2.13). Indeed, $\mathrm{E}[|\Delta|Q_{b,c}(u,v)|||\mathscr{F}_{i-1}]$ is at most the sum of the absolute values of the terms in (2.14), all of which are $O(n^{-1/2})$. Thus,

$$\mathrm{E}[|\Delta Q^+_{b,c}(u,v)||\mathscr{F}_{i-1}] \le \mathrm{E}[|\Delta|Q_{b,c}(u,v)|||\mathscr{F}_{i-1}] + |\Delta(q_{b,c}(t) + f(t))n| = O(n^{-1/2}),$$

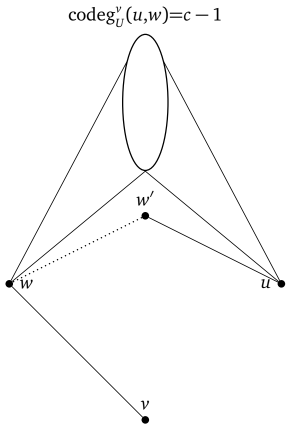
and the one-step variance is

$$\mathbf{Var}[\Delta Q^+_{b,c}(u,v)|\mathscr{F}_{i-1}] \le \mathrm{E}[(\Delta Q^+_{b,c}(u,v))^2|\mathscr{F}_{i-1}] = \tilde{O}(n^{1/2}) \cdot \mathrm{E}[|\Delta Q^+_{b,c}(u,v)||\mathscr{F}_{i-1}] = \tilde{O}(1).$$

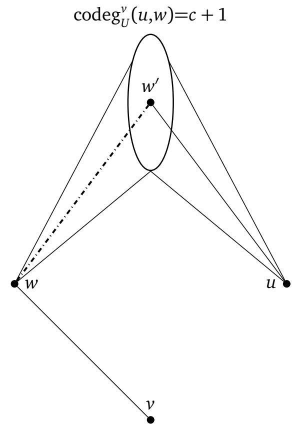
Thus, Lemma 1.3.3 applied with $\lambda = n^{4/5}$, $\tilde{b} = \tilde{O}(n^{3/2})$ and $C = \tilde{O}(n^{1/2})$ (using $\tilde{b}$ instead of b to avoid notational collision) yields that the failure probability is at most

$$\exp\left\{-\frac{n^{8/5}}{\tilde{O}(n^{3/2}) + \tilde{O}(n^{1/2}\cdot n^{4/5})}\right\}$$

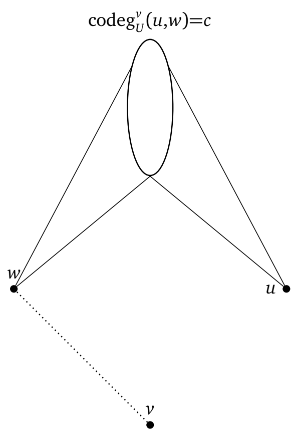
which is again small enough to beat a union bound over all pairs of vertices and values of b,c.



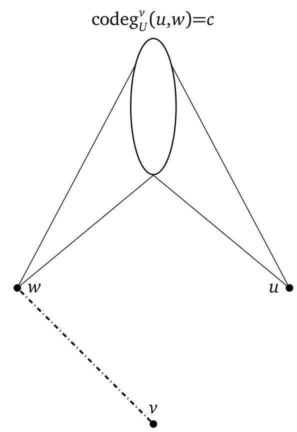
(a) Ways $w \in S_{c-1}(u,v,i-1)$ Could Join $S_c(u,v,i)$ with the Addition of the Dotted Edge.



(b) Ways $w \in S_{c+1}(u,v,i-1)$ Could Join $S_c(u,v,i)$ with the Removal of the Dashed Edge.



(c) Ways $w \in Q_{c,0}(u,v,i-1)$ Could Join $S_c(u,v,i)$.



(d) Ways to Lose $w \in S_c(u,v,i-1)$.

Figure 2.3: Cases Considered in the Expected One-Step Change of $|S_c(u,v,i)|$. Note that $\mathrm{codeg}_U^v(u,w)$ Denotes the Unmatched Codegree of u and w, Excluding v.

2.1.8 Tracking $|S_c(u,v)|$

Finally, we calculate $\mathrm{E}[\Delta|S_c(u,v,i)||\mathscr{F}_{i-1}]$, starting with positive contributions. We denote by $\mathrm{codeg}_U^v(u,w)$ the unmatched codegree of w and u, excluding v. Thus, the addition of $\{u,v\}$. First, for $w \in S_{c-1}(u,v)$, we could increase $\mathrm{codeg}_U^v(u,w)$ by the addition of an edge, a situation accounted for by $|A(u,w)|$ (see Figure 2.3a). Second, for $w \in S_{c+1}(u,v)$, an edge $\{w,w'\}$ or $\{u,w'\}$ could become matched, thereby reducing $\mathrm{codeg}_U^v(u,w)$ from $c+1$ to c. This situation is handled with $B(w,w') + B(u,w')$ and shown in Figure 2.3b. Third, e_i could be an edge $\{v,w\}$ for some w with $\text{codeg}_U^v(w,u) = c$. We may use $Q_{c,0}(u,v)$ for this case (see Figure 2.3c).

Next, we consider negative contributions. The first two cases are identical to the positive contribution with the exception that we consider $w \in S_c(u,v)$, and hence lose such a w instead of gaining one. For the third case, the edge $\{w,v\}$ could become matched for $w \in S_c(u,v)$ and thus w no longer meets the requirements for membership in $S_c(u,v)$. This is shown in Figure 2.3d.

We are also concerned with more unlikely cases. Take $w \in S_{c'}(u,v)$ for appropriate c'. There are $\tilde{O}(n^{1/2})$ of such w. Notice that if $e_i = \{u,v\}$, $|S_{c'}(u,v)|$ would remain unchanged. On the other hand, e_i could be in the common neighborhood of u,w or e_i could be the edge $\{u,w\}$, which would form triangles with the common neighborhood of these vertices. Both of these cases are a change of $\tilde{O}(1)$, and so the overall change for these unlikely cases is $\tilde{O}(n^{-3/2})$.

Therefore,

$$
\begin{aligned}
&\mathrm{E}[\Delta|S_c(u,v,i)||\mathscr{F}_{i-1}] \\
&= \Bigg[\sum_{w\in S_{c-1}(u,v)} |A(u,w)| + |Q_{c,0}(u,v)| + \sum_{\substack{w\in S_{c+1}(u,v)\\ w'\in \mathrm{codeg}_U(u,w)}} \big[B(w,w')+B(u,w')\big] - \sum_{w\in S_c(u,v)} |A(u,w)| \\
&- \sum_{\substack{w\in S_c(u,v)\\ w'\in \mathrm{codeg}_U(u,w)}} \big[B(w,w')+B(u,w')\big] - \sum_{w\in S_c(u,v)} B(v,w)\Bigg] \cdot \frac{2}{n^2}(1+\tilde{O}(n^{-1/2})) + \tilde{O}(n^{-3/2}) \qquad (2.17)\\
&\le \Big[(s_{c-1}+c^{-1}f)(\alpha+3f) + (q_{c,0}+f) + 2(c+1)(s_{c+1}+(c+2)^{-1}f)(\beta+4f) \\
&\qquad - (s_c-(c+1)^{-1}f)(\alpha-3f) - 2c(s_c-(c+1)^{-1}f)(\beta-4f) \\
&\qquad - (s_c-(c+1)^{-1}f)(\beta-4f)\Big] \cdot 2n^{-1} + \tilde{O}(n^{-3/2}) \\
&\le \Big[2s_{c-1}\alpha + 4(c+1)\beta s_{c+1} + 2q_{c,0} - 2(\alpha+2c\beta+\beta)s_c \qquad (2.18)\\
&\quad + \big(2+4c\beta(c+2)^{-1}+4c\beta(c+1)^{-1}\big)f \\
&\quad + \big(2\alpha c^{-1} + 16(c+1)s_{c+1} + 2(8c+7)s_c + 6s_{c-1} + 4\beta(c+2)^{-1} + 2(\alpha+\beta)(c+1)^{-1}\big)f \qquad (2.19)\\
&\quad + \big(16(c+1)(c+2)^{-1} - 16c(c+1)^{-1} + 6c^{-1} - 14(c+1)^{-1}\big)f^2\Big] n^{-1} + \tilde{O}(n^{-3/2}).
\end{aligned}
$$

Recall from (2.7) that (2.18) is s_c'. We also remark that

$$2+4c\beta(c+2)^{-1}+4c\beta(c+1)^{-1} \le 50.$$

From the bounds on α, s_c and β in (2.2) and (2.3) we can easily conclude that (2.19) is $O((c+1)^{-1}f)$. Furthermore, since $f^2 \le n^{-1/10}f$, we can say that the f^2 terms are also $O((c+1)^{-1}f)$. Thus,

$$\mathrm{E}[\Delta|S_c(u,v,i)||\mathscr{F}_{i-1}] \le \big[s_c' + 50f + O((c+1)^{-1}f)\big]n^{-1} + \tilde{O}(n^{-3/2}). \qquad (2.20)$$

Now we define variables

$$S_c^{\pm}(u,v) = S_c^{\pm}(u,v,i) := \begin{cases} |S_c(u,v,i)| - (s_c(t(i)) \pm (c+1)^{-1} f(t(i)))n^{1/2} & \text{if } \mathscr{E}_{i-1} \text{ holds} \\ S_c^{\pm}(u,v,i-1) & \text{otherwise.} \end{cases}$$

Extend the function $t(i) := i/n^{3/2}$ to the real numbers. Then set with $h(t) := s_c(t) + (c+1)^{-1} f(t)$ and $H(i) = h(t(i))$. Then Taylor's theorem yields

$$H(i) - H(i-1) = \frac{dH(i-1)}{di} + \frac{1}{2}\frac{d^2H(\omega)}{di^2} = \frac{dH(i-1)}{dt} n^{-3/2} + \frac{1}{2}\frac{d^2H(\omega)}{di^2},$$

where $\omega \in [i-1, i]$. But

$$\frac{d^2H(i)}{di^2} = \frac{d}{di}\left(\frac{dH(i)}{dt} n^{-3/2}\right) = \left(\frac{d^2 s_c(t(i))}{dt^2} + (c+1)^{-1}\frac{d^2 f(t(i))}{dt^2}\right) n^{-3}.$$

However,

$$\frac{d^2 s_c}{dt^2} = -\frac{-4y^{2c-1}e^{-2y^2}(2y^4 - (4c+3)y^2 + 2c^2 + c)(2e^{y^2} - 3)}{c!},$$

so $|d^2 s_c/dt^2| = O(1)$. Also, we have $|(c+1)^{-1} d^2 f/dt^2| = O(n^{-2})$. Hence,

$$(H(i) - H(i-1))n^{1/2} = \left[s_c'(t(i-1)) + (c+1)^{-1} f'(t(i-1))\right] n^{-1} + O(n^{-3/2}).$$

So as a consequence of (2.20) for $t = t(i-1)$ we have

$$\begin{aligned} \mathrm{E}\left[\Delta S_c^{+}(u,v,i) | \mathscr{F}_{i-1}\right] &\le \left[-(c+1)^{-1}\frac{df}{dt} + 50f + O((c+1)^{-1}f) \right] n^{-1} + \tilde{O}(n^{-3/2}) \\ &\le \left[-\frac{1000\log n}{\log\log n} + 50(c+1) + O(1) \right] (c+1)^{-1} f n^{-1} + \tilde{O}(n^{-3/2}) \\ &\le \left[-\frac{1000\log n}{\log\log n} + \frac{200\log n}{\log\log n} + O(1) \right] (c+1)^{-1} f n^{-1} + \tilde{O}(n^{-3/2}) \le 0. \end{aligned}$$

Next we demonstrate that $|\Delta |S_c(u,v)|| = O(\log n)$. First, an edge e in in the unmatched graph G_U might be removed. Indeed, if e is incident with u, say $e = \{u, x\}$, then the removal of e can only affect vertices $w \in S_c(u, v)$ such that $w \in \{x\} \cup (N(x) \cap N(v))$ of which there are only $O(\log n)$. Similarly if e is incident with v then at most $O(1)$ vertices $w \in S_c(u, v)$ are affected. Finally, if e is not incident with u, v then the only vertices $w \in S_c(u, v)$ that could be affected are the endpoints of e. Thus, zero or three edges are removed at any step and each one affects $O(\log n)$ vertices w. We can also make symmetrical observations for adding an edge to G_U. Hence, $|\Delta|S_c(u, v)|| = O(\log n)$. Also $|\Delta S_c^+(u, v)| = O(\log n)$, since the deterministic terms have much smaller one-step changes. We can also see that $\mathrm{E}[|\Delta|S_c(u, v)|||\mathscr{F}_{i-1}] = O(n^{-1})$ by an argument analogous to the one used to justify (2.13). Indeed, $\mathrm{E}[|\Delta|S_c(u, v)|||\mathscr{F}_{i-1}]$ is at most the sum of the absolute values of the terms in (2.17), all of which are $O(n^{-1})$. Thus,

$$\mathrm{E}[|\Delta S_c^+(u, v)||\mathscr{F}_{i-1}] \leq \mathrm{E}[|\Delta|S_c(u, v)|||\mathscr{F}_{i-1}] + |\Delta(s_c(t) + f(t))n^{1/2}| = O(n^{-1})$$

and

$$\mathbf{Var}[\Delta S_c^+(u, v)|\mathscr{F}_{i-1}] \leq \mathrm{E}[(\Delta S_c^+(u, v))^2|\mathscr{F}_{i-1}] = O(\log n) \cdot \mathrm{E}[|\Delta S_c^+(u, v)||\mathscr{F}_{i-1}] = \tilde{O}(n^{-1}).$$

Therefore, using Lemma 1.3.3 with $\lambda = (c+1)^{-1} n^{3/10}$, $b = \tilde{O}(n^{1/2})$ and $C = \tilde{O}(1)$ our failure probability is at most

$$\exp\left\{-\frac{(c+1)^{-2} n^{3/5}}{\tilde{O}(n^{1/2}) + \tilde{O}(n^{3/10})}\right\},$$

which is small enough to beat a union bound over all pairs of vertices and values of c.

2.1.9 Proof of Theorem 1.4.1

At the end of our process, after revealing $kn^{3/2}$ edges, the number of unmatched edges is at most $\frac{n^{3/2}}{2}(y(k) + f(k))$, and thus the number of matched edges is at least

$$kn^{3/2} - \frac{n^{3/2}}{2}(y(k) + f) \geq kn^{3/2} - \frac{y(k) n^{3/2}}{2} - n^{7/5}.$$

Recall that the only edges of M are those of edge-disjoint triangles. Hence, the number of edge-disjoint triangles at the end of our process is w.h.p. at least

$$(1+o(1))\frac{1}{3}\left(k-\frac{y(k)}{2}\right)n^{3/2}.$$

2.2 Proof of Theorem 1.4.4

We bound $\tau(G)$ using two different approaches for different ranges of k. First, we consider the triangle-free process for $k = O(n^{3/2})$ (and thus $t = O(1)$). This process outputs a set of unmatched edges that form a triangle-free subgraph of $G(n,m)$ and the remaining, matched edges that form a triangle cover. We will refer to Bohman's original triangle-free paper [16]. Recall that in this process one maintains a triangle-free subgraph $G_T(i) \subseteq G(n,i)$ by revealing one edge at a time, and adding that edge to $G_T(i)$ only if it does not create a triangle in $G_T(i)$.

In order to use the results about the triangle-free process, we must reconcile the differences between our step parameter and the one used by Bohman. In [16], one step was counted as the addition of an unmatched edge whereas we count each step as the presenting of an edge, which might become matched. To avoid notational confusion with Bohman's paper, all of the variables in [16] will appear here with a circumflex above them, e.g. $\hat{i}, \hat{t}, \hat{Q}$, etc. With this notation, the number of unmatched edges produced by the process after $i = tn^{3/2}$ edges are proposed is $\hat{i} = \hat{t}n^{3/2}$.

Bohman proved that w.h.p. for all $\hat{i} \leq kn^{3/2}$ the number $\hat{Q}(\hat{i})$ of edges eligible to be inserted into the triangle-free graph (i.e. edges that would remain unmatched if proposed) is

$$\hat{Q}(\hat{i}) \in (1 \pm n^{-\delta})\binom{n}{2}e^{-4\hat{t}^2}. \tag{2.21}$$

where the constant $\delta > 0$ is derived from Bohman's original error function. Since (2.21) holds for some $\delta > 0$ we may take δ to be arbitrarily small. We will assume that $0 < \delta < 1/2$. We also note that Bohman proved this for all $\hat{i}$ at most some constant times $n^{3/2}\log^{1/2} n$ but we will not fully use that here.

Let the random variable $A(i)$ be the number of unmatched edges that have been added after i edges have been presented. Of course Bohman's step number is $\hat{i} = A(i)$ and so $\hat{t} = A(i)n^{-3/2}$. We will track $A(i)$ using the differential equation method. We will show that $A(i) \approx n^{3/2}a(t)$ for some function $a(t)$.

Let $\tilde{\mathcal{E}}_i$ be the event that for all $i^* \leq i$ we have:

(i) $\hat{Q}(A(i^*)) \in (1 \pm n^{-\delta})\binom{n}{2}e^{-4\left(A(i^*)n^{-3/2}\right)^2}$,

(ii) $A(i^*) \in (1 \pm f_A(t^*))n^{3/2}a(t^*)$,

where $t^* = i^*/n^{3/2}$. Since in this section we have $t = O(1)$, we also have $a(t) = O(1)$. Assume the constant c is an upper bound on $a(t)$. We define the error function f_A to be

$$f_A(t) := n^{-\delta}e^{(2+13c^2)t} = O(n^{-\delta}).$$

Since Bohman showed that the first condition holds w.h.p., we verify that the second condition also w.h.p. holds. Note that in the event $\tilde{\mathcal{E}}_i$ we have

$$\begin{aligned}
\hat{Q}(A(i)) &\in (1 \pm n^{-\delta})\binom{n}{2}e^{-4(1\pm 3f_A(t))a(t)^2} \\
&\subseteq (1 \pm n^{-\delta}) \cdot (1 \pm 12f_A(t)a^2(t) + O(f_A^2(t)a^4(t))) \cdot \binom{n}{2}e^{-4a(t)^2} \\
&\subseteq \left(1 \pm n^{-\delta} \pm 12c^2 f_A(t) + O(n^{-2\delta})\right)\binom{n}{2}e^{-4a(t)^2} \\
&\subseteq \left(1 \pm (1 + 12c^2)f_A(t) + O(n^{-2\delta})\right)\binom{n}{2}e^{-4a(t)^2}
\end{aligned}$$

where the second inclusion follows from the Taylor series for $e^{12f_A(t)a^2(t)}$, and the last inclusion follows from $f_A(t) \geq n^{-\delta}$. Therefore we have

$$\mathrm{E}[\Delta A(i) | \mathscr{F}_i] = \frac{\hat{Q}(A(i))}{\binom{n}{2} - i}$$

and we may derive the following differential equation:

$$\frac{da}{dt} = e^{-4a(t)^2} \quad \text{and} \quad a(0) = 0.$$

Now estimate $\mathrm{E}[\Delta A(i) | \mathscr{F}_i]$ as follows, noting that the lower bound can be attained with symmetric calculations:

$$\begin{aligned}
\mathrm{E}[\Delta A(i) | \mathscr{F}_i] &= \frac{\hat{Q}(A(i))}{\binom{n}{2} - i} \\
&\leq \frac{(1 + (1 + 12c^2) f_A(t)) \binom{n}{2} e^{-4a(t)^2}}{\binom{n}{2} - i} \\
&\leq (1 + O(n^{-1/2}))(1 + (1 + 12c^2) f_A(t)) e^{-4a^2} \\
&\leq e^{-4a^2} \big(1 + (1 + 12c^2) f_A(t)\big) + O(n^{-1/2}).
\end{aligned}$$

Next define variables

$$A^{\pm} = A^{\pm}(i) := \begin{cases} A(i) - (a(t(i)) \pm f_A(t(i))) n^{3/2} & \text{if } \tilde{\mathscr{E}}_{i-1} \text{ holds} \\ A^{\pm}(i-1) & \text{otherwise.} \end{cases}$$

Then A^+ is a supermartingale and A^- is a submartingale. The calculation to verify A^+ is shown. Applying Taylor's theorem with $g(t) := a(t) + f_A(t)$ and $t(i) := \frac{i}{n^{3/2}}$ yields

$$(g \circ t)(i) - (g \circ t)(i-1) = (g \circ t)'(i-1) + \frac{(g \circ t)''(\omega)}{2} = g'(t(i-1)) n^{-3/2} + \frac{(g \circ t)''(\omega)}{2},$$

where $\omega \in [i-1, i]$. But

$$(g \circ t)''(i) = (g'(t(i))n^{-3/2})' = g''(t(i))n^{-3} = \left(a''(t) + f_A''(t)\right) n^{-3}$$

as well as

$$a''(t) = 2a(t) \cdot a'(t) \cdot e^{-4a^2(t)} = 2a(t)e^{-8a^2(t)} = O(1)$$

and

$$f_A''(t) = (2+13c^2)^2 n^{-\delta} e^{(2+13c^2)t} = O(n^{-\delta}).$$

This gives us $(g \circ t)''(i) = O(n^{-3})$, so

$$\Delta(a(t(i)) + f_A(t(i)))n^{3/2} = a'(t(i-1)) + f_A'(t(i-1)) + O(n^{-3/2}).$$

Thus for $t = t(i-1)$ we have

$$\begin{aligned} \mathrm{E}\left[\Delta A^+(i) | \mathscr{F}_{i-1}\right] &\leq (1+12c^2)e^{-4a^2} f_A - f_A' + O(n^{-3/2}) \\ &\leq (1+12c^2) f_A - f_A' + O(n^{-3/2}) \\ &= -(1+c^2)e^{(2+13c^2)t} n^{-\delta} + O(n^{-3/2}) \leq 0. \end{aligned}$$

Now we use the Azuma-Hoeffding inequality (Lemma 1.3.2) to bound the failure probability. Clearly $|\Delta A(i)| = O(1)$ and

$$|\Delta(a(t(i)) + f_A(t(i)))n^{3/2}| \leq |a'(t(i-1))| + |f_A'(t(i-1))| + O(n^{-3/2}) = O(1),$$

and thus for $t = t(i)$, $\lambda = f_A(0)n^{3/2} = n^{13/10}$ and $C = O(1)$,

$$\begin{aligned} \mathbb{P}(A(i) > n^{3/2}(a(t)) + f_A(t)) &= \mathbb{P}(A^+(i) > 0) \\ &= \mathbb{P}(A^+(i) - A^+(0) > n^{13/10}) \leq \exp\left(-\frac{n^{13/5}}{2C^2 t n^{3/2}}\right) = o(1). \end{aligned}$$

Hence w.h.p. $A(i^*) = (1+o(1))n^{3/2}a(t^*)$ for $t^* \le t$. So we may say that $\tau(G(n, tn^{3/2})) \le (1+o(1))(t-a(t))n^{3/2}$.

To optimize our upper bound, we combine what we just obtained with another bound on $\tau(G)$ which is better for certain values of k. We observe that one can cover all triangles in any graph G by using at most half of the edges. To demonstrate this, let H be the largest bipartite subgraph of G. It is well-known that $|E(H)| \ge \frac{1}{2}|E(G)|$ (see e.g., [27]). Therefore $E(G) \setminus E(H)$ cover all triangles and we have

$$\tau(G(n,m)) \le m/2.$$

Therefore, we can conclude that w.h.p.

$$\tau(G(n, kn^{3/2})) \le (1+o(1))U_\tau(k)n^{3/2},$$

where

$$U_\tau(k) := \min\{k - a(k), k/2\}. \tag{2.22}$$

And from Theorem 1.4.1, we have

$$\nu(G(n, kn^{3/2})) \ge (1+o(1))L^*_\nu(k)n^{3/2},$$

where

$$L^*_\nu(k) := \frac{1}{3}\left(k - \frac{y(k)}{2}\right). \tag{2.23}$$

By the proof of Theorem 2.1.1, to verify Tuza's conjecture for $G(n, kn^{3/2})$ it suffices to check k in the range of $0.2 \le k \le 3$. However, our bounds (2.22) and (2.23) are enough to show that the quotient $U_\tau(k)/L^*_\nu(k) \le 2$ for this range (see Appendix for details).

2.3 Remarks and Further Directions

As a possible area of future study, we conjecture that we can improve Tuza's conjecture for the random graph $G(n,m)$. Recall that in general, Baron and Kahn [9] showed that we cannot decrease the multiplicative constant 2 since for any $\alpha > 0$ there are arbitrarily large graphs G of positive density satisfying $\tau(G) > (1-o(1))|G|/2$ and $\nu(G) < (1+\alpha)|G|/4$. However, the graphs used in this are significantly different than $G(n,m)$, being a subgraph of a blowup construction.

Frankl and Rödl proved in [36] that w.h.p. for every $\varepsilon_1 > 0$ there is some k such that the largest triangle-free subgraph of $G(n,p)$ for $p = kn^{-1/2}$ has at least $1/2-\varepsilon_1$ proportion of the edges. In addition, for $\varepsilon_2 > 0$ the triangle packing number $\nu(G(n,m))$ with $m = kn^{3/2}$ and $k \geq (\log n)^2$ is w.h.p. $1/3(1-\varepsilon_2)kn^{3/2}$. Then by the asymptotic equivalence of $G(n,p)$ and $G(n,m)$ we see that for $G = G(n,m)$, we have

$$\frac{\tau(G)}{\nu(G)} \leq \frac{(1/2-\varepsilon_1)kn^{3/2}}{1/3(1-\varepsilon_2)kn^{3/2}} \approx \frac{3}{2}$$

On the other hand we have $\tau(G) \geq \nu(G)$ for all G and the ratio $\tau(G)/\nu(G)$ is about 1 when triangles start emerging in G. We believe this ratio starts at about 1 and grows to about 3/2. This leads us to the following conjecture.

Conjecture 2.3.1. *For all $C > 3/2$ and $G = G(n,m)$, w.h.p. $\tau(G) \leq C \cdot \nu(G)$ for any range of m.*

2.4 Calculations

Here we verify that $U_\tau(k)/L_\nu^*(k) \leq 2$ for $0.2 \leq k \leq 3$. First define $g(k) := (k/2)/L_\nu^*(k)$ and $h(k) := (k-a(k))/L_\nu^*(k)$. We show that

(i) h is increasing for $.2 \leq k \leq 1.29$,

(ii) g is decreasing for $1.28 \leq k \leq 3$, and

(iii) $h(1.29), g(1.28) < 2$.

We first demonstrate (ii) and (iii) for g. Recall that $y(k) \leq 0.6368$ and hence

$$L_\nu^*(1.28) \geq \frac{1}{3}\left(1.28 - \frac{0.6368}{2}\right) \geq 0.3205.$$

Thus

$$g(1.28) = \frac{(1.28/2)}{L_\nu^*(1.28)} \leq \frac{0.64}{0.3205} = 1.9969,$$

verifying (iii) for g. To show (ii), using $dy/dt = 6e^{-y^2} - 4$ gives us

$$\frac{dg}{dk} = \frac{18ke^{-y(k)^2} - 3y(k) - 12k}{(-2k + y(k))^2}.$$

Then the numerator of dg/dk is 0 when $k = 0$. Now taking the derivative of this numerator tells us that

$$18e^{-y^2} - 36kye^{-y^2}(6e^{-y^2} - 4) - 3(6e^{-y^2} - 4) - 12 = -36kye^{-y^2}(6e^{-y^2} - 4) \leq 0$$

for $k \geq 0$ since $dy/dt, y \geq 0$. Thus, dg/dk is nonpositive for $1.28 \leq k \leq 3$.

Verifying (i) and (iii) for h analytically, however, seems more complicated. Therefore, to avoid tedious calculations we use Maple and verify the required conditions numerically. For a, we obtain

```
DE_A := {diff(a(t), t) = exp(-4*a(t)^2), a(0) = 0};
a_sol_list := dsolve(DE_A, numeric, output = listprocedure);
a_sol := rhs(a_sol_list[2]);
```

With this in hand, we can see that (iii) holds since $h(1.29) \leq 1.987$. Now plotting the graph of h (see Figure 2.4) verifies (i).

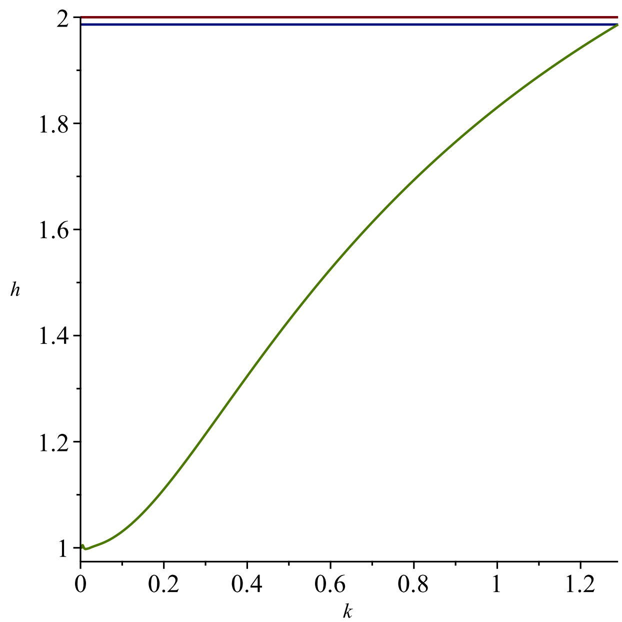


Figure 2.4: The Graph of h (in Green) Compared with $h(1.29)$ (in Blue).

CHAPTER 3

ALTERNATING PATHS IN RANDOM GRAPHS

3.1 Introduction

In this chapter, we study alternating connectivity of the random graph $G(n,p)$. Suppose the edges of a graph are colored with r colors. Then we say a path in the graph is an *alternating path* if adjacent edges have different colors. Then we define the the alternating connectivity $\kappa_{r,\ell}(G)$ as the maximum t such that there is an r-edge-coloring of G such that any pair of vertices is connected by t internally disjoint and alternating paths of length ℓ.

Concepts similar to alternating connectivity were previously studied in [34]. Here, the edges of random graphs were randomly colored with two colors. Then conditions under which the graph had every pair of vertices connected by an alternating path were determined. In addition the concept of alternating connectivity was studied for complete graphs in [14]. In this case,

$$\kappa_{r,2}(K_n) \sim (1-1/r)n \quad \text{and} \quad \kappa_{r,\ell}(K_n) \sim n/(\ell-1)$$

for any $r \geq 2$ and $\ell \geq 3$.

In this chapter, we will prove the following three theorems on alternating connectivity of $G(n,p)$. Our first theorem concerns paths of length 2.

Theorem 1.5.1. *Let $p \gg \sqrt{\log n/n}$ and $G = G(n,p)$ and let r be an integer. Then, w.h.p.*

$$\kappa_{r,2}(G) \sim \left(1 - \frac{1}{r}\right) np^2.$$

The strategy for this theorem will be to establish a stronger result for graphs that are nearly d-regular and have every pair of vertices with codegree about d^2/n, conditions satisfied by $G(n,p)$ with high probability. And for the range of p considered, the result is essentially optimal. A random coloring is employed to obtain the lower bound.

The next two theorems concerns paths of length at least three, covering a dense case (Theorem 1.5.2) and a sparse case (Theorem 1.5.3).

Theorem 1.5.2. *Let $0 < p < 1$ be a constant and $G = G(n, p)$. Then, for any integer $\ell \geq 3$ w.h.p.*

$$\kappa_{r,\ell}(G) \sim \min\left\{\frac{n}{\ell - 1}, np\left(1 - \frac{p}{2}\right)\right\}.$$

Theorem 1.5.3. *Suppose $G = G(n, p)$ with $p = o(1)$ and $r \geq 2$ is an integer.*

(i) Let $k \geq 2$ be a positive integer such that $n^{1/k} \leq np \leq n^{1/(k-1)}$. If $\ell \geq k + 2$, then w.h.p. we have $\kappa_{r,\ell}(G) \sim np$.

(ii) Let $k \geq 2$ be a positive integer such that $(n \log n)^{1/k} \ll np \leq n^{1/(k-1)}$. If $\ell = k + 1$, then w.h.p. we have $\kappa_{r,\ell}(G) \sim np$.

(iii) Let $k \geq 3$ be a positive integer such that $(n \log n)^{1/k} \ll np \ll n^{1/(k-1)}$. If $\ell = k$, then w.h.p. we have $\kappa_{r,\ell}(G) = \Theta(n^{k-1}p^k)$.

Both of these theorems are shown to achieve a natural upper bound of either the total number of paths of a certain length between two vertices or the size of a neighborhood in $G(n, p)$ (adjusting for any intersection). The lower bound is where the majority of the work is done. In both theorems, we employ a uniform random 2-coloring, allowing our results to be independent of r. The general strategy is to take two vertices and consider their red and blue neighborhoods and then link these up with sets of the same or larger size from both ends, and then find an appropriately-sized matching in the middle. For Theorem 1.5.3, this strategy requires more careful analysis.

After proving these theorems, we will study two generalizations. The first is the parameter $\lambda_{r,\ell}(G)$, which is the same as $\kappa_{r,\ell}(G)$ but without the restriction that only internally disjoint paths are allowed. We will provide some progress on this more difficult parameter for d-regular graphs. The second generalization is considering pseudorandom graphs instead of the random graph $G(n,p)$. With the modification of a key lemma, we will be able to show that Theorem 1.5.2 holds in certain pseudorandom graphs.

3.2 Auxiliary Results for $G(n,p)$

We will rely on several straightforward lemmas about matchings in random graphs. The first one is about almost perfect matchings in the dense random graph. A similar result for bipartite random graphs (with almost identical proof) was obtained in [14].

Lemma 3.2.1. *Let $0<\alpha,p<1$ be constants and $G(n,p)$ be a random graph on set of vertices V. Then, w.h.p. for any subsets $A,B\subseteq V$ with $|A|=|B|=\alpha n$, there exists a matching between them of size $\alpha n(1+o(1))$.*

Proof. Fix $A,B\subseteq V$ with $|A|=|B|=\alpha n$ and consider the random binomial graph $G=G(|A|,|B|,p)$. First we consider an auxiliary bipartite graph H on $U\cup W$ such that $U=A\cup A'$, $W=B\cup B'$, $H[A\cup B]=G$, and $H[A'\cup W]$ and $H[U\cup B']$ are complete bipartite graphs. Furthermore, let $s=\log n=|A'|=|B'|$. We show that H has a perfect matching. It suffices to show that the Hall condition holds, i.e.,

$$\text{if } S\subseteq U \text{ and } |S|\le |U|/2, \text{ then } |N(S)|\ge |S|, \tag{3.1}$$

and

$$\text{if } T\subseteq W \text{ and } |T|\le |W|/2, \text{ then } |N(T)|\ge |T|. \tag{3.2}$$

If $|S|<s$, then since $N(S)\supseteq B'$, $|N(S)|\ge |B'|=s\ge |S|$. Therefore, we assume that $s\le |S|\le |U|/2$. Furthermore, we may assume that $S\cap A'=\emptyset$. For otherwise, $N(S)=W$. We will show that already for $|S|=s$, $|N(S)|\ge |W|/2=|U|/2$.

Suppose not, that is, $|N(S)|<(\alpha n+\log n)/2$. That means $|B\cap N(S)|<(\alpha n-\log n)/2$, $e(S,B\setminus N(S))=0$ and $|B\setminus N(S)|\ge(\alpha n+\log n)/2>\alpha n/2$. Observe that the probability that there are sets $S \in A$ and $T \in B$ such that $|S| = s$ and $|T| = \alpha n/2$ and $e(S,T) = 0$ is at most

$$\binom{\alpha n}{s}\binom{\alpha n}{\alpha n/2}(1-p)^{s\alpha n/2} \leq 2^{2\alpha n}(1-p)^{s\alpha n/2}.$$

Thus, with probability at most $2^{2\alpha n} \cdot (1-p)^{s\alpha n/2}$ the graph H violates (3.1), and similarly (3.2). In other words, with probability at least $1 - 2 \cdot 2^{2\alpha n}(1-p)^{s\alpha n/2}$ the graph H has a perfect matching, and consequently, there is a matching of size $\alpha n - s$ between A and B.

Finally, by taking the union bound over all $A \in \binom{V}{\alpha n}$ and $B \in \binom{V}{\alpha n}$ we get that the probability that there exist A and B such that between A and B there is no matching of size $\alpha n - s$ is at most

$$\binom{n}{\alpha n}\binom{n}{\alpha n}2^{2\alpha n+1}(1-p)^{s\alpha n/2} \leq 2^n \cdot 2^n \cdot 2^{2\alpha n+1}(1-p)^{s\alpha n/2} = o(1),$$

since $s = \log n$. Also, clearly, we get that $\alpha n - s = \alpha n - o(n)$. Thus, w.h.p. for each A and B there is a matching between A and B of size $\alpha n - o(n)$. □

The next two lemmas deal with matchings with sparse random graphs. For a bipartite graph $G = (A \cup B, E)$ we say that G contains a *d-matching* from A to B of size t if there exists in G a subgraph of t vertex-disjoint stars $K_{1,d}$ such that each star is centered in A.

Lemma 3.2.2. *For a fixed integer k, let $\log n \ll np \leq n^{1/(k-1)}$. Suppose the set A satisfies $1 \leq |A| \leq (np)^{k-2}$ and the set B has order $\sim n$ and is disjoint from A. Consider the random bipartite graph on $A \cup B$ with edge probability p. Then with probability at least $1 - \frac{1}{n^3}$ there is a d-matching that saturates A with $d = np/4$.*

Proof. To prove this lemma, we verify that Hall's condition holds with high probability. Hence, we show that w.h.p. every set $S \subseteq A$ has $|N(S)| \geq d|S|$. For convenience, we will denote the orders of S, A, and B by their corresponding lowercase letters.

To that end, we will apply (1.2) to the random variable $|N(S)| \in \text{Bin}(b,q)$ with $q = 1-(1-p)^s$. But $sp \leq ap \leq (np)^{k-2}p = (np)^{k-1}/n \leq 1$. Therefore, by (1.3) q may be estimated as $q = 1-(1-p)^s \geq 1 - e^{-sp} \geq 1-(1-sp/2) = sp/2$, and so the expected value satisfies

$\mu = bq \geq bsp/2 \geq (1+o(1))snp/2 \gg s\log n$. Now invoking Chernoff's bound (1.2) with $\delta = 1/3$ implies together with the union bound that the probability of failure is at most

$$\sum_{S\subseteq A} \Pr(|N(S)| \leq (1-\delta)\mu) \leq \sum_{s=1}^{a} \binom{a}{s} \exp\left(-\delta^2\mu/2\right) \leq \sum_{s=1}^{a} \exp\left(s\log a - \mu/18\right)$$

and since trivially $a \leq n$ and $\mu \gg s\log n$ we can easily bound the above probability by $1/n^3$. Finally note that

$$|N(S)| \geq (1-\delta)\mu \geq (1+o(1))2snp/6 > snp/4 = d|S|,$$

as required. $\square$

Lemma 3.2.3. *For a fixed integer $k \geq 2$, let $\sqrt{n} \leq np \leq \sqrt{n\log n}$. Suppose the set A satisfies $|A| \leq np$ and the set B has order at least $n/2$ and is disjoint from A. Consider the random bipartite graph on $A \cup B$ with edge probability p. Then with probability at least $1-\frac{1}{n^3}$ there is a d-matching that saturates A with $d = 1/(6p)$.*

Proof. We show that w.h.p. every set $S \subseteq A$ has $|N(S)| \geq d|S|$. We apply (1.2) to $|N(S)| \in \mathrm{Bin}(b,q)$ with $q = 1-(1-p)^s$. We consider two cases. First assume that $ps \leq 1$. Then by (1.3) $q = 1-(1-p)^s \geq 1-e^{-sp} \geq 1-(1-sp/2) = sp/2$ and so the expected value satisfies $\mu = bq \geq bsp/2 \geq snp/4 \gg s\log n$. So invoking Chernoff's bound (1.2) with $\delta = 1/3$ implies together with the union bound (over all subsets of size at most $1/p$) that the probability of failure is at most

$$\sum_{S\subseteq A} \Pr(|N(S)| \leq (1-\delta)\mu) \leq \sum_{s=1}^{1/p} \binom{a}{s} \exp\left(-\delta^2\mu/2\right) \leq \sum_{s=1}^{a} \exp\left(s\log a - \mu/18\right)$$

and since trivially $a \leq n$ and $\mu \gg s\log n$ we can easily bound the above probability by $1/n^3$. Finally note that

$$|N(S)| \geq (1-\delta)\mu \geq snp/6 = s(np)^2/(6np) = d|S|(np)^2/n \geq d|S|,$$

as required. Now assume that $ps > 1$. But then $q \geq 1 - e^{-sp} \geq 1 - e^{-1} \geq 1/2$. Hence $\mu = bq \geq n/4$. So again using Chernoff's bound (1.2) and the union bound again (over all subsets of size at least $1/p$) with $\delta = 1/3$, we have that the failure probability is at most

$$\sum_{S \subseteq A} \Pr(|N(S)| \leq (1-\delta)\mu) \leq \sum_{s=1/p}^{a} \binom{a}{s} \exp\left(-\delta^2\mu/2\right) \leq \sum_{s=1}^{a} \exp\left(s \log a - \mu/18\right)$$

so since $s \log a \leq a \log a \leq np \log n \leq \sqrt{n}(\log n)^{3/2}$ and $\mu \geq n/4$ then we may again bound the failure probability by $1/n^3$. We also satisfy Hall's Condition since

$$|N(S)| \geq (1-\delta)\mu \geq n/6 = np/(6p) \geq d|S|.$$

□

The next lemma modifies the well-known result that asserts that w.h.p. $G(n,n,p)$ with $p = \frac{\log n + \omega}{n}$ has a perfect matching.

Lemma 3.2.4. *Consider the random bipartite graph $G(m,m,q)$ with $q = \frac{\log m}{m}$. Let $C > 0$ be an absolute constant. Then with probability at least $1 - \frac{1}{m^C}$ the graph has a matching of size $(1-o(1))m$.*

Proof. The proof goes along the lines of Theorem 6.1 from [38].

Let $G = G(m,m,q)$ the graph on the set of vertices $A \cup B$. Let $r = m/(\log m)$. We show that with the desired probability $|N(S)| \geq |S| - r$ for all $S \subseteq A$. This will prove that there is a matching consisting of at least $m - r = (1-o(1))m$ edges.

Note that if we have an "obstruction" $|N(S)| < |S| - r$ for $S \subseteq A$, $|S| > (m+r)/2$, then letting $T = B \setminus N(S)$ we have that $|T| = m - |N(S)|$ so $|T| \leq m - (|S| - r) < (m+r)/2$ and we have $|N(T)| \leq m - |S| < m - (|N(S)| + r) = |T| - r$.

Also, if $S \subseteq A$ is a minimal obstruction then every vertex in $N(S)$ has degree at least 2 into S (otherwise we could remove its neighbor from S to obtain a smaller obstruction). Also we may assume that $|N(S)|$ is exactly $s - r - 1$ or else we could find a smaller obstruction.

Thus we look only for obstructions of the following form: $S \subseteq A$ (or $T \subseteq B$) consisting of $r+2 \leq s \leq (m+r)/2$ vertices, and having a neighborhood consisting of exactly $s-r-1$ vertices, each of which has two neighbors in S (or T) and hence there are at least $2(s-r-1)$ edges between S and $N(S)$.

$$\begin{aligned}
&\sum_{s=r+2}^{(m+r)/2} \binom{m}{s}\binom{m}{s-r-1}\binom{s(s-r-1)}{2(s-r-1)} q^{2(s-r-1)}(1-q)^{s(m-s+r+1)} \\
&\quad \leq \sum_{s=r+2}^{(m+r)/2} \left(\frac{em}{s}\right)^s \left(\frac{em}{s-r-1}\right)^{s-r-1} \left(\frac{esq}{2}\right)^{2(s-r-1)} \exp\{-s(m-s+r+1)q\} \\
&\quad = \sum_{s=r+2}^{(m+r)/2} \left(\frac{em}{s}e^{-mq}\right)^{r+1} \left(\frac{em}{s}\frac{em}{s-r-1}\frac{e^2s^2q^2}{4}e^{-(m-s)q}\right)^{s-r-1} \\
&\quad \leq \sum_{s=r+2}^{(m+r)/2} \left(\frac{O(1)}{r}\right)^r \left(O(1)\cdot\frac{(\log m)^2 s}{s-r-1}e^{-(1-s/m)\log m}\right)^{s-r-1}.
\end{aligned}$$

Now if $r+2 \leq s \leq 2r$, then $\frac{s}{s-r-1} \leq 2r$ and

$$\left(O(1)\cdot\frac{(\log m)^2 s}{s-r-1}e^{-(1-s/m)\log m}\right)^{s-r-1} \leq \left((\log m)^2 2re^{-\frac{1}{2}\log m}\right)^{s-r-1} \leq \left((\log m)^2 2re^{-\frac{1}{2}\log m}\right)^{r}.$$

implying

$$\left(\frac{O(1)}{r}\right)^r \left(O(1)\cdot\frac{(\log m)^2 s}{s-r-1}e^{-(1-s/m)\log m}\right)^{s-r} = \left(O(1)\cdot(\log m)^2 e^{-\frac{1}{2}\log m}\right)^r = (o(1))^r.$$

Otherwise, if $s \geq 2r$, then $\frac{s}{s-r} \leq 2$ and so $\frac{(\log m)^2 s}{s-r}e^{-(1-s/m)\log m} = o(1)$ yielding again an upper bound of $(o(1))^r$.

Finally, since in the sum we have only $O(m)$ terms and $r = m/(\log m)$, we trivially get that the failure probability is at most $1/m^C$ for any positive constant C. □

The last lemma deals with colored degrees and codegrees. Let $G = (V, E)$. Denote by $N_i(v)$ the i-colored neighborhood of v, i.e., the set of vertices w such that $\{v, w\} \in E$ is colored by i. In particular, $N(v)$ is the union of $N_i(v)'s$ over all colors i. Also let $N_{ij}(u, v)$ be the set of all $w \in V$ such that $\{u, w\}$ and $\{v, w\}$ are edges and are colored i and j, respectively.

Lemma 3.2.5. *Let $0 < p < 1$ be a constant and $G = (V, E) = G(n, p)$. Let E be colored uniformly at random with the colors red and blue denoted by R and B, respectively. Then, w.h.p. for any two vertices $u, v \in V$ we have*

$$|N_R(u) \setminus N(v)| \sim |N_B(u) \setminus N(v)| \sim \frac{np}{2}(1-p) \text{ and } |N_{RB}(u,v)| \sim |N_{BR}(u,v)| \sim \frac{np^2}{4}.$$

Proof. Consider any two vertices $u, v \in V$. Note that $|N_R(u) \setminus N(v)|$ is binomially distributed with $n-2$ trials and probability of success $p(1-p)/2$, so the expected value μ is $(1+o(1))np(1-p)/2$. Set $\delta = \sqrt{7(\log n)/\mu}$ and apply Chernoff's bounds (1.1) and (1.2). Thus,

$$\Pr(||N_R(u) \setminus N(v)| - \mu| \geq \delta\mu) \leq 2\exp(-\delta^2\mu/3) = 2\exp(-(7/3 + o(1))\log n).$$

Finally, the union bound over all $\binom{n}{2}$ pairs of vertices $u, v \in V$ yields that w.h.p. for any u and v, $|N_R(u) \setminus N(v)| \sim \frac{np}{2}(1-p)$. By symmetry, the same holds for $|N_B(u) \setminus N(v)|$.

For $|N_{RB}(u,v)|$, we apply similar reasoning. Note that $|N_{RB}(u,v)| \in \mathrm{Bin}(n-2, p^2/4)$ and so $\mu \sim \frac{np^2}{4} \gg 1$. Applying Chernoff's bounds again with $\delta = \sqrt{7(\log n)/\mu}$ will imply the statement. □

3.3 Alternating Paths of Length Two

Here we prove Theorem 1.5.1. Actually we will prove a more general statement.

Theorem 3.3.1. *Let $G = (V, E)$ be a graph with $|V| = n$, where all vertices have degree $\sim d$ and every pair of vertices has codegree $\sim d^2/n \gg \log n$. Then, for any number of colors $r \geq 2$*

$$\kappa_{r,2}(G) \sim \left(1 - \frac{1}{r}\right)\frac{d^2}{n}.$$

The standard application of Chernoff's bound implies that $G(n, p)$ meets the requirements of Theorem 3.3.1 w.h.p. and hence Theorem 1.5.1 holds. Furthermore, observe that if

$p \ll \sqrt{\log n/n}$, then w.h.p. the diameter is at least three. Therefore, Theorem 1.5.1 is basically optimal.

We separate the proof of Theorem 3.3.1 into two lemmas. The upper bound follows from the Cauchy-Schwarz inequality and does not require the assumption on the codegrees. The lower bound is obtained by considering a random coloring along with Chernoff's bound.

Lemma 3.3.2. *Let $G = (V, E)$ be a graph where all vertices have degree $\sim d$ and $|V| = n$. Then, for any number of colors $r \geq 2$*

$$\kappa_{r,2}(G) \leq \left(1 - \frac{1}{r} + o(1)\right)\frac{d^2}{n}.$$

Proof. The total number of alternating paths of length 2 in G is at most (due to the Cauchy-Schwarz inequality)

$$\begin{aligned}\sum_{v\in V}\sum_{1\leq i<j\leq r}\deg_i(v)\deg_j(v) &= \sum_{v\in V}\frac{1}{2}\left(\left(\sum_{i=1}^{r}\deg_i(v)\right)^2 - \sum_{i=1}^{r}\deg_i(v)^2\right)\\ &\leq \sum_{v\in V}\frac{1}{2}\left(\deg(v)^2 - \frac{\deg(v)^2}{r}\right)\\ &\leq \frac{nd^2}{2}\left(1-\frac{1}{r}+o(1)\right),\end{aligned}$$

and so

$$\kappa_{r,2}(G) \leq \frac{nd^2}{2}\left(1-\frac{1}{r}+o(1)\right)\Big/\binom{n}{2} = \left(1-\frac{1}{r}+o(1)\right)\frac{d^2}{n}.$$

□

Lemma 3.3.3. *Let $G = (V, E)$ be a graph with $|V| = n$, where all vertices have degree $\sim d$ and every pair of vertices has codegree $\sim d^2/n \gg \log n$. Then, for any number of colors $r \geq 2$*

$$\kappa_{r,2}(G) \geq \left(1 - \frac{1}{r} + o(1)\right)\frac{d^2}{n}.$$

Proof. To each edge in E we assign a color from $\{1, \ldots, r\}$ uniformly at random. For $u, v \in V$, let $X_{u,v}$ be the random variable that counts the number of alternating paths between u and v of length 2. Clearly, $X_{u,v} \sim \text{Bin}((1+o(1))d^2/n, 1-1/r)$.

Notice that $\mu := \text{E}(X_{u,v}) = \left(1 - \frac{1}{r} + o(1)\right)\frac{d^2}{n} \gg \log n$ and $\delta := \sqrt{(5\log n)/\mu} = o(1)$. Thus, (1.2) yields

$$\Pr(X_{u,v} \leq (1-\delta)\mu) \leq \exp(-\mu\delta^2/2) \leq \exp((-5/2+o(1))\log n).$$

Thus, the union bound taken over all $\binom{n}{2} \leq \exp(2\log n)$ pairs of vertices in V yields the statement. □

3.4 Alternating Paths of Length at Least Three in Dense Random Graphs

We consider the case when p is constant and the length ℓ of the paths we seek is greater than two. Clearly the largest possible number of internally disjoint paths between two vertices is $\sim n/(\ell-1)$. On the other hand, two vertices cannot have more internally disjoint paths between them than one half of the size of the union of their neighborhoods. As it happens, there exists a coloring that allows for a matching lower bound of *alternating* paths, no matter which case occurs.

Here we prove Theorem 1.5.2, which we state below again for convenience.

Theorem 1.5.2. *Let $0 < p < 1$ be a constant and $G = G(n,p)$. Then, for any integer $\ell \geq 3$ w.h.p.*

$$\kappa_{r,\ell}(G) \sim \min\left\{\frac{n}{\ell-1}, np\left(1-\frac{p}{2}\right)\right\}.$$

Our strategy for the lower bound will be to color all edges uniformly at random with two colors. Then the main idea to find paths between u and v will be to consider the (slightly adjusted) red and blue neighborhoods of each and use Lemma 3.2.1 to find almost perfect, appropriately-colored matchings with linear-sized sets of vertices connecting neighborhoods of u to v (see Figure 3.1). For convenience, we split the proof of Theorem 1.5.2 between Lemmas 3.4.1 and 3.4.2.

Note that when $\ell = 3$, we have $\min\{n/2, np(1-p/2)\} = np(1-p/2)$ for any $0 < p < 1$.

Lemma 3.4.1. *Let $0 < p < 1$ be a constant and $G = G(n, p)$. Then for any integer $\ell \geq 4$ satisfying $n/(\ell-1) \leq np(1-p/2)$ for sufficiently large n, we have w.h.p.*

$$\kappa_{r,\ell}(G) \sim \frac{n}{\ell-1}.$$

Proof. The upper bound on $\kappa_{r,\ell}(G)$ is obvious. For a matching lower bound, for each pair of vertices $\{x, y\}$ we will have to find a collection of paths using almost all other vertices in the graph. We will accomplish this by first covering the mutual non-neighbors of x, y, and then covering the remaining vertices. We color each edge in $E = E(G)$ uniformly at random either red or blue. For any pair of vertices $x, y \in V$, consider the set $U = N(x) \cup N(y)$ and $S = V \setminus U$. Note that w.h.p. $|U| \sim 2np(1-p/2)$ and $|S| = s \sim n(1-p)^2$.

Define disjoint sets

$$X_B = (N_B(x) \setminus N(y)) \cup N_{BB}(x, y),$$

$$X_R = (N_R(x) \setminus N(y)) \cup N_{RR}(x, y),$$

$$Y_B = (N_B(y) \setminus N(x)) \cup N_{RB}(x, y),$$

$$Y_R = (N_R(y) \setminus N(x)) \cup N_{BR}(x, y).$$

and observe that by Lemma 3.2.5, we have that $|X_B| \sim |X_R| \sim |Y_B| \sim |Y_R| \sim \frac{np}{2}\left(1 - \frac{p}{2}\right)$.

Now notice that

$$np(1-p/2) \geq n/(\ell-1) \tag{3.3}$$

is equivalent to

$$(\ell-3)np(1-p/2) \geq n - 2np(1-p/2) \tag{3.4}$$

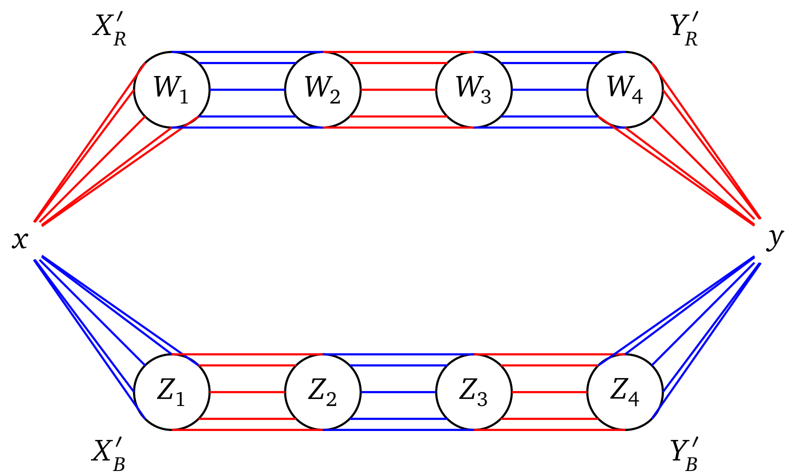


Figure 3.1: Finding the Sets W_i and Z_i and Corresponding Matchings for $\ell = 5$.

and hence to

$$np(1-p/2) \geq n(1-p)^2/(\ell-3) \sim s/(\ell-3). \tag{3.5}$$

Consequently, $\frac{np}{2}(1-\frac{p}{2}) \geq (1+o(1))\frac{s}{2(\ell-3)}$ and we can choose

$$X_B' \subset X_B,\ X_R' \subset X_R,\ Y_B' \subset Y_B,\ Y_R' \subset Y_R$$

such that $|X_B'| \sim |X_R'| \sim |Y_B'| \sim |Y_R'| \sim \frac{s}{2(\ell-3)}$.

Now we define a family of disjoint sets $\{W_i\}_{i=1}^{\ell-1}$ and $\{Z_i\}_{i=1}^{\ell-1}$ as follows. First, we set $W_1 = X_R'$ and $Z_1 = X_B'$. Then for ℓ even we define $W_{\ell-1} = Y_B'$ and $Z_{\ell-1} = Y_R'$. For ℓ odd we define $W_{\ell-1} = Y_R'$ and $Z_{\ell-1} = Y_B'$. Note that these sets are disjoint. Now for $2 \leq i \leq \ell-2$ we inductively and arbitrarily select $W_i,\ Z_i$ of size $s/2(\ell-3)$ from the remaining vertices of S such that W_i and Z_i are disjoint from the previously selected sets (see Figure 3.1).

Now by Lemma 3.2.1 for each $1 \leq i \leq \ell-2$ we have w.h.p. an almost perfect matching between W_i and W_{i+1} as well as Z_i and Z_{i+1} of the appropriate color. That means for odd (even) i we can w.h.p. find a blue (red) matching of size $s/(2(\ell-3)) + o(n)$ between W_i and W_{i+1} as well as a red (blue) matching of size $s/(2(\ell-3)) + o(n)$ between Z_i and Z_{i+1}. This gives us a total of $s/(\ell-3) + o(n)$ alternating paths between x and y, covering almost all vertices of S

(and also covering some neighbors of x, y).

Now we apply this same argument to the set $U' = (X_R \setminus X'_R) \cup (X_B \setminus X'_B) \cup (Y_R \setminus Y'_R) \cup (Y_B \setminus Y'_B)$. Let $k := |U'| = 2np(1-p/2) - 2s/(\ell-3) + o(n)$. If $k = o(n)$, then $2np(1-p/2) \sim 2s/(\ell-3)$ and by the previous calculations (3.3)-(3.5) we obtain that $s/(\ell-3) \sim n/(\ell-1)$ yielding $(1+o(1))n/(\ell-1)$ alternating paths between x and y.

Assume that $k = \Omega(n)$ and note that $|X_B \setminus X'_B| \sim |X_R \setminus X'_R| \sim |Y_B \setminus Y'_B| \sim |Y_R \setminus Y'_R| \sim k/4$. Hence, we can find sets

$$X''_B \subset X_B \setminus X'_B,\ X''_R \subset X_R \setminus X'_R,\ Y''_B \subset Y_B \setminus Y'_B,\ Y''_R \subset Y_R \setminus Y'_R$$

such that $|X''_B| = |X''_R| = |Y''_B| = |Y''_R| = \frac{k}{2(\ell-1)}$.

As before, we define a family of disjoint sets $\{W'_i\}_{i=1}^{\ell-1}$ and $\{Z'_i\}_{i=1}^{\ell-1}$. We set $W'_1 = X''_R$ and $Z'_1 = X''_B$. Then for ℓ even we define $W'_{\ell-1} = Y''_B$ and $Z'_{\ell-1} = Y''_R$ and for ℓ odd $W'_{\ell-1} = Y''_R$ and $Z'_{\ell-1} = Y''_B$. Now for $2 \le i \le \ell-2$ we inductively and arbitrarily select W'_i, Z'_i of size $m/2 = \frac{k}{2(\ell-1)}$ from the remaining vertices of U' such that W'_i and Z'_i are disjoint from the previously selected sets. Let us observe that we were able to find all disjoint sets W_i, Z_i, W'_i and Z'_i since

$$\begin{aligned} 2(\ell-1)\frac{s}{2(\ell-3)} + 2(\ell-1)\frac{k}{2(\ell-1)} &= s\left(1+\frac{2}{\ell-3}\right) + k \\ &= s + 2np(1-p/2) = n(1-p)^2 + 2np(1-p/2) + o(n) = n + o(n). \end{aligned} \tag{3.6}$$

Again invoking Lemma 3.2.1, for odd (even) i we can w.h.p. find a blue (red) matching of size $\frac{k}{2(\ell-1)} + o(n)$ between W'_i and W'_{i+1} as well as a red (blue) matching of size $\frac{k}{2(\ell-1)} + o(n)$ between Z'_i and Z'_{i+1}. So altogether we have found $\frac{k}{\ell-1} + o(n)$ additional alternating paths between x and y. Taken together with the other paths we have found

$$\frac{k}{\ell-1} + \frac{s}{\ell-3} + o(n) = \frac{n}{\ell-1} + o(n)$$

alternating paths between x and y. Note that the above equality holds by dividing (3.6) by $\ell - 1$. $\square$

Lemma 3.4.2. *Let* $0 < p < 1$ *be a constant and* $G = G(n, p)$*. Then for any integer* $\ell \geq 3$ *satisfying* $n/(\ell - 1) \geq np(1 - p/2)$ *for sufficiently large n, w.h.p.*

$$\kappa_{r,\ell}(G) \sim np\left(1 - \frac{p}{2}\right).$$

Proof. For the upper bound, take any pair of vertices $x, y \in V$ and consider the set $U = N(x) \cup N(y)$. Observe that w.h.p. $|U| = 2np(1 - p/2) + o(n)$. Furthermore, since the interior of every path of length ℓ from x to y must use two vertices from U and $\ell - 3$ vertices from elsewhere, the largest possible number of disjoint xy-paths possible is $np(1 - \frac{p}{2}) + o(n)$.

Now we show the lower bound. Color each edge in E uniformly at random either red or blue and define disjoint sets

$$X_B = (N_B(x) \setminus N(y)) \cup N_{BB}(x, y),$$
$$X_R = (N_R(x) \setminus N(y)) \cup N_{RR}(x, y),$$
$$Y_B = (N_B(y) \setminus N(x)) \cup N_{RB}(x, y),$$
$$Y_R = (N_R(y) \setminus N(x)) \cup N_{BR}(x, y).$$

Then by Lemma 3.2.5, we have that $|X_B| \sim |X_R| \sim |Y_B| \sim |Y_R| \sim \frac{np}{2}\left(1 - \frac{p}{2}\right)$. We now apply a similar argument to the proof of Lemma 3.4.1. Define $\{W_i\}_{i=1}^{\ell-1}$ and $\{Z_i\}_{i=1}^{\ell-1}$ of size $(np/2)(1 - p/2) + o(n)$. First, we set $W_1 = X_R$ and $Z_1 = X_B$. Then for ℓ even we define $W_{\ell-1} = Y_B$ and $Z_{\ell-1} = Y_R$; for ℓ odd we define $W_{\ell-1} = Y_R$ and $Z_{\ell-1} = Y_B$. Note that these sets are disjoint. Now for $2 \leq i \leq \ell - 2$ we inductively and arbitrarily select W_i, Z_i of size $\sim (np/2)(1 - p/2)$ from the remaining vertices such that W_i and Z_i are disjoint from the previously selected sets. We are able to find such sets W_i and Z_i, since

$$2(\ell-1)\frac{np}{2}\left(1-\frac{p}{2}\right)=(\ell-1)\cdot np\left(1-\frac{p}{2}\right)\leq(\ell-1)\cdot\frac{n}{\ell-1}=n.$$

We apply Lemma 3.2.1 to find matchings of the appropriate color between the sets we have defined. For each $1\leq i\leq\ell-2$ and odd (even) i we can w.h.p. find a blue (red) matching of size $(np/2)(1-p/2)+o(n)$ between W_i and W_{i+1} as well as a red (blue) matching of size $(np/2)(1-p/2)+o(n)$ between Z_i and Z_{i+1}. This gives us $np(1-p/2)+o(n)$ alternating paths between x and y.

□

3.5 Alternating Paths of Length at Least Three in Sparse Random Graphs

In this section we investigate the sparser case $p=o(1)$ and prove Theorem 1.5.3 (stated for convenience below).

Theorem 1.5.3 *Suppose $G=G(n,p)$ with $p=o(1)$ and $r\geq 2$ is an integer.*

(i) Let $k\geq 2$ be a positive integer such that $n^{1/k}\leq np\leq n^{1/(k-1)}$. If $\ell\geq k+2$, then w.h.p. we have $\kappa_{r,\ell}(G)\sim np$.

(ii) Let $k\geq 2$ be a positive integer such that $(n\log n)^{1/k}\ll np\leq n^{1/(k-1)}$. If $\ell=k+1$, then w.h.p. we have $\kappa_{r,\ell}(G)\sim np$.

(iii) Let $k\geq 3$ be a positive integer such that $(n\log n)^{1/k}\ll np\ll n^{1/(k-1)}$. If $\ell=k$, then w.h.p. we have $\kappa_{r,\ell}(G)=\Theta(n^{k-1}p^k)$.

Observe that for $k=2$ in parts (i) and (ii) condition $p=o(1)$ implies that $np\ll n$. Also notice that in part (iii) we may assume that $k\geq 3$; the case $k=2$ follows from Theorem 1.5.1.

Proof. Due to the low probabilities involved, our strategy is more delicate than in the dense

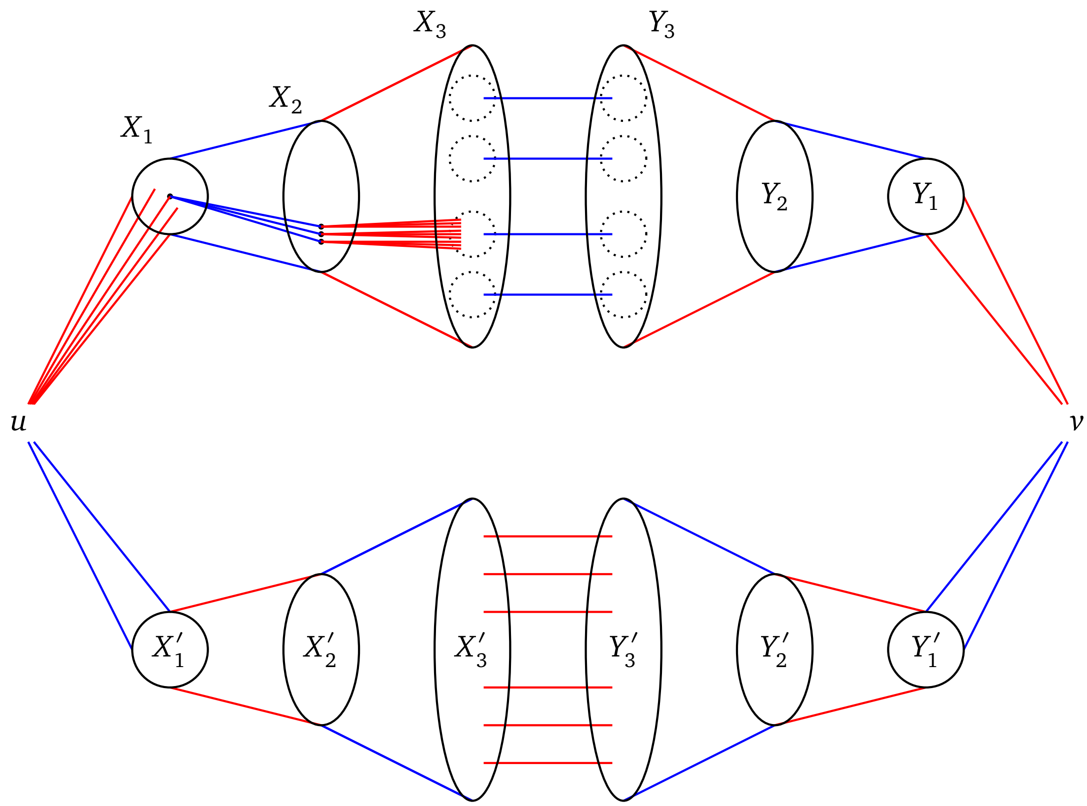


Figure 3.2: Finding Alternating Paths of Length 7 in the Sparse Case

case. Given two fixed vertices, our program is to progressively reveal their colored neighborhoods in an alternating fashion. Due to the sparseness of the graph, these sets are small and can be taken to be disjoint with only minor adjustments. Then we use auxiliary bipartite graphs to find a matching of the appropriate size between the last neighborhoods.

First we prove **part (ii)**. Let $k \geq 2$ be a positive integer such that $(n \log n)^{1/k} \ll np \leq n^{1/(k-1)}$ and $\ell = k+1$. Clearly, $\kappa_{r,\ell}(G)$ is bounded above by the minimum degree, but a Chernoff argument tells us that all degrees in G are w.h.p. concentrated around the mean of $np(1+o(1))$. So it suffices to establish the lower bound.

We color each edge of G uniformly and independently from the colors red and blue. Fix two vertices u and v. We estimate the probability of finding $np(1+o(1))$ disjoint, red-blue paths between u and v and then take the union bound over all choices for u, v. To that end, we construct the following sequences of disjoint subsets of V: for $1 \leq i \leq \lceil k/2 \rceil$ the sets X_i,

X_i'; and for $1 \le i \le \lfloor k/2 \rfloor$ the sets Y_i, Y_i'. We require these sequences to have the following properties:

- The sets X_1 and X_1' correspond to the red and blue neighborhoods of u and Y_1 and Y_1' correspond to the red and blue neighborhoods of v. Hence, by a simple Chernoff argument, their orders x_1, x_1', y_1, y_1' are $\sim np/2$.
- For $i \ge 2$ we have $x_i = x_1(np/8)^{i-1}$ and there is an $np/8$-matching from X_{i-1} to X_i consisting of blue edges for even i and red for odd i. Similarly, we have $x_i' = x_1'(np/8)^{i-1}$ and there is an $np/8$-matching from X_{i-1}' to X_i' consisting of red edges for even i and blue for odd i.
- For $i \ge 2$ we have $y_i = y_1(np/8)^{i-1}$ and there is an $np/8$-matching from Y_{i-1} to Y_i consisting of blue edges for even i and red for odd i. Similarly, we have $y_i' = y_1'(np/8)^{i-1}$ and there is an $np/8$-matching from Y_{i-1}' to Y_i' consisting of red edges for even i and blue for odd i.

To construct the sets X_i, X_i', Y_i, Y_i' we inductively use Lemma 3.2.2 by first finding the X_i's and then finding the Y_i's. First, we initialize X_1, X_1', Y_1, Y_1' as the appropriately colored neighborhoods of u and v. Then for $2 \le i \le \lceil k/2 \rceil - 1$ (this necessarily means that $k \ge 3$), define

$$A = X_{i-1}, \quad B = V \setminus \cup_{j=1}^{i-1} \left(X_j \cup X_j' \cup Y_1 \cup Y_1' \right).$$

Notice that A has order at most $(np)^{\lceil k/2 \rceil - 2} \le (np)^{k-2}$ and B has order $\sim n$ as required by Lemma 3.2.2. So viewing $A \cup B$ as a random bipartite graph with colored edge probability $p/2$, applying the lemma gives that there exists an $(np/8)$-matching from A to B that saturates A with a failure probability at most $1/n^3$. Set the image of this matching in B to be X_i, and observe that $x_i = x_{i-1}(np/8) = x_1(np/8)^{i-1}$. Repeat this process with $A' \cup B'$ for

$$A' = X_{i-1}', \quad B' = V \setminus \cup_{j=1}^{i-1} \left(X_j \cup X_j' \cup Y_1 \cup Y_1' \right) \cup X_i.$$

Thus, we obtain the sets X_i for $1 \leq i \leq \lceil k/2 \rceil$ with the desired property and with a failure probability of $O(1/n^3)$. In a similar manner, for $2 \leq i \leq \lfloor k/2 \rfloor$ we obtain Y_i, Y_i' with a failure probability of $O(1/n^3)$.

In order to complete our search for $np(1+o(1))$ disjoint, red-blue paths between u and v, we find $np(1+o(1))$ correctly colored edges between the sets $X_{\lceil k/2 \rceil}, X'_{\lceil k/2 \rceil}$ and $Y_{\lfloor k/2 \rfloor}, Y'_{\lfloor k/2 \rfloor}$. If k is odd, we find red edges between $X_{\lceil k/2 \rceil}$ and $Y'_{\lfloor k/2 \rfloor}$ as well as blue edges between $X'_{\lceil k/2 \rceil}$ and $Y_{\lfloor k/2 \rfloor}$. If k is even, we find red edges between $X_{\lceil k/2 \rceil}$ and $Y_{\lfloor k/2 \rfloor}$ as well as blue edges between $X'_{\lceil k/2 \rceil}$ and $Y'_{\lfloor k/2 \rfloor}$. For convenience, we will denote X and Y to be the sets with red edges between and X' and Y' to be the sets with blue edges between.

We now construct a random auxiliary bipartite graph $H(X,Y)$ by partitioning X into $\sim np/2$ disjoint sets of size $(np/8)^{\lceil k/2 \rceil - 1}$ and Y into $\sim np/2$ disjoint sets of size $(np/8)^{\lfloor k/2 \rfloor - 1}$, each set being identified as a vertex in $H(X,Y)$. This partition is done in such a way that each partition class consists of the leaves of a tree rooted at one of the $\sim np/2$ vertices of X_1, and internal vertices of this tree are the neighbors of this root in X_i for $2 \leq i \leq \lceil k/2 \rceil - 1$ (see Figure 3.2).

We say there is an edge in $H(X,Y)$ if there is at least one red edge between the corresponding sets of vertices in G. Observe that the edge probability is

$$q = 1 - (1-p/2)^{(np/8)^{\lceil k/2 \rceil - 1 + \lfloor k/2 \rfloor - 1}} = 1 - (1-p/2)^{(np/8)^{k-2}} \sim 1 - e^{-n^{k-2}p^{k-1}/(2 \cdot 8^{k-2})}.$$

Since $n^{k-2}p^{k-1} = (np)^{k-1}/n \leq 1$, we get by (1.3) that

$$q \sim 1 - e^{-n^{k-2}p^{k-1}/(2 \cdot 8^{k-2})} \geq 1 - (1 - n^{k-2}p^{k-1}/(4 \cdot 8^{k-2})) = n^{k-2}p^{k-1}/(4 \cdot 8^{k-2}).$$

We view $H(X,Y)$ as $G(m,m,q)$ where $m = np/2$. The expected degree of $H(X,Y)$ is given by

$$\Theta(mq) = \Theta(n^{k-1}p^k) = \Theta((np)^k/n) \gg \log n = \Omega(\log m).$$

Hence, by applying Lemma 3.2.4 with m, q and $C = 2k$ we get that there is an almost perfect matching of size $np/2(1 + o(1))$ between the bipartition in $H(X, Y)$, with a failure probability of at most $1/m^{2k}$. We follow the same process with $H(X', Y')$ to obtain an almost perfect matching of size $np/2(1 + o(1))$. So altogether we've found $np(1 + o(1))$ disjoint alternating paths between our choice of u and v. But now taking the union bound over all choices of u and v gives us a total failure probability of at most

$$\binom{n}{2} \cdot O\left(\frac{1}{n^3} + \frac{2}{(np)^{2k}}\right) \sim O\left(\frac{1}{n} + \frac{n^2}{(np)^{2k}}\right) = o(1),$$

since $(np)^{2k} \gg (n \log n)^2$, by assumption.

Let us summarize what we have found. For any choice of u and v, we build four tree structures: two rooted at u and two rooted at v. The first level of these trees consists of $\sim np/2$ edges that are completely red or completely blue. Each vertex of this level then has several neighbors (of the opposite color than the first level) that are disjoint from the rest of the tree structures. This pattern continues for each of these neighbors until we have the desired length. Then, looking at the leaves that can be traced back to a single neighbor of u, we find at least one edge between these leaves and the leaves in the corresponding tree that can be traced back to a single neighbor of v. This gives us our desired $\sim np$ alternating paths between any u and v.

Now we discuss how to prove **part (i)**. We first assume $k \geq 5$. Here, we may follow the proof of the previous case by fixing a u and v and finding the sets X_i, X_i' for $1 \leq i \leq \lceil k/2 \rceil$ and Y_i, Y_i' for $1 \leq i \leq \lfloor k/2 \rfloor$. We can achieve this with a failure probability of $O(1/n^3)$.

We must use Lemma 3.2.2 again to find sets $Y_{\lfloor k/2 \rfloor + 1}$ and $Y'_{\lfloor k/2 \rfloor + 1}$ of order $(np/8)y_{\lfloor k/2 \rfloor}$. Observe that for $k \geq 5$ we have

$$|Y_{\lfloor k/2 \rfloor + 1}| = |Y'_{\lfloor k/2 \rfloor + 1}| = O((np)^{\lfloor k/2 \rfloor + 1}) = O(n^{\frac{\lfloor k/2 \rfloor + 1}{k-1}}) \ll n,$$

so we have enough room. These sets take the role of Y and Y' in the previous part. We first deal with when $\ell = k + a$ for some integer $a > 2$. Here, we alter our construction: we still apply Lemma 3.2.2, which gives us that each vertex has $np/8$ correctly colored neighbors from $X_{\lceil k/2 \rceil+1}$ to the remaining vertices. But instead, we let $X_{\lceil k/2 \rceil+1}$ consist of just one vertex from each of the $np/8$-sized stars. Then $X_{\lceil k/2 \rceil+1}$ has the same order as $X_{\lceil k/2 \rceil}$. Our previous calculations allow us to continue in this manner, always being sure to alternate the color and accruing a failure probability of $1/n^3$ for each step, until we obtain $X_{\lceil k/2 \rceil+a-2}$, which has order $x_{\lceil k/2 \rceil} \ll n$. Notice that if we find an appropriate matching between this set and $Y_{\lfloor k/2 \rfloor+1}$ (or $Y_{\lfloor k/2 \rfloor+1}$ depending on the color), this will give us $\sim np/2$ alternating paths of length $\ell = k + a$.

So we may assume without loss of generality that $\ell = k+2$. Then we build an auxiliary bipartite graph $H(X, Y)$ as before, but this time we partition Y into $\sim np/2$ vertices of order $(np/8)^{\lfloor k/2 \rfloor}$ and X into $\sim np/2$ vertices of order $(np/8)^{\lceil k/2 \rceil -1}$. And since $n^{k-1}p^k = (np)^k/n \geq 1$, this allows us to say that the edge probability is

$$q = 1-(1-p/2)^{(np/8)^{k-1}} \geq 1-e^{-n^{k-1}p^k/(2\cdot 8^{k-1})} \geq 1-e^{-1/(2\cdot 8^{k-2})} = \Omega(1).$$

Set $m = np/2$ and observe that clearly $\Theta(mq) \gg \log m$. Hence, we may apply the same argument to find w.h.p. a matching of size $\sim np/2$ between X and Y. Similarly for $H(X', Y')$. This gives us the desired $\sim np$ disjoint alternating paths between u and v of length $\ell = k+2$.

We still need to address the case $2 \leq k \leq 4$. In these cases, we need to adjust our previous strategy to ensure that the sets in our construction are not too large. We do this first with $3 \leq k \leq 4$. Here if $(n \log n)^{1/k} \ll np \leq n^{1/(k-1)}$, then we proceed like in case (ii) so we have no need to create larger sets. Therefore, we may assume that $n^{1/k} \leq np \leq (n \log n)^{1/k}$. But now the additional sets we need to make are of order

$$|Y_{\lfloor k/2 \rfloor+1}| = |Y'_{\lfloor k/2 \rfloor+1}| = O((np)^{\lfloor k/2 \rfloor+1}) = O((n \log n)^{\frac{\lfloor k/2 \rfloor+1}{k}}) \ll n.$$

Now we consider the case $k = 2$. Recall that since $p = o(1)$ we have $\sqrt{n} \leq np \ll n$ and as before we may assume that $\sqrt{n} \leq np \leq \sqrt{n \log n}$. We initialize $X_1,\ X'_1,\ Y_1 := Y,\ Y'_1 := Y'$ as before. Then we construct a set X and X' of order at most $n/6$ each. We apply Lemma 3.2.3 with $A = X_1$, $B = V \setminus (X_1 \cup X_1' \cup Y_1 \cup Y_1')$, then $|A| \leq np$ and $|B| \geq n - 4np(1 + o(n)) \geq n/2$. And since $\sqrt{n} \leq np \leq \sqrt{n \log n}$, we have that there exists an $1/(3p)$-matching between A and B with failure probability $1/n^3$. We call the image of this matching X, which has order at most $n/3$. We do the same process with X_1' to obtain X'. Then we apply the same strategy as before by constructing $H(X, Y)$ where we have $\sim np/2$ partition classes of size $1/(3p)$ (where each class consists of $1/(3p)$ vertices that are the image of a single vertex in X). We partition Y as before. Then since $1/(3p) \geq 1$, the edge probability is

$$q = 1 - (1 - p/2)^{1/(3p)} \geq 1 - e^{-1/6} = \Omega(1).$$

And if we set $m = np/2$ then $\Theta(mq) \gg \log m$. Hence, we may apply the same argument (using Lemma 3.2.4 with $C = 5$) as before to find w.h.p. a matching of size $\sim np/2$ between X and Y. Similarly for $H(X', Y')$. If $\ell > 3$, then we apply the same strategy as above by using Lemma 3.2.3 instead of 3.2.2 to finding sets of the same size, before finally a matching between X and Y of the appropriate color. We finish the proof by taking the union bound over all choices of u and v. But again $\binom{n}{2} O(1/n^3 + 1/(np)^5) = O(1/n + 1/n^{1/2}) = o(1)$ giving us the desired number of alternating paths in this case.

We now turn our attention to **part (iii)**. For the upper bound, note that w.h.p. the number of paths of length k is $O(n^{k+1}p^k)$. Indeed, choose an arbitrary vertex v_0 and build a copy of P_{k+1} greedily choosing next vertex v_1 from $N(v_0)$, $v_2 \in N(v_1)$, etc. Since for every vertex v, w.h.p. $|N(v)| \sim np$, the number of all paths of length k is $O(n(np)^k) = O(n^{k+1}p^k)$ and so there is a pair of vertices with at most $O(n^{k+1}p^k)/\binom{n}{2} = O(n^{k-1}p^k)$ paths.

Now for the lower bound. Fix two vertices u and v. We continue as before and again search by applying Lemma 3.2.2 for X_i, X_i' for $1 \leq i \leq k-1$. (The assumptions of Lemma 3.2.2 are still satisfied since the last time we applied this lemma to $|A| = O((np)^{k-2})$ and $|B| \sim n$.) Again, we can achieve this with a failure probability of $O(1/n^3)$. As before, we construct stars

$H(X,\{v\})$ and $H(X',\{v\})$ by partitioning X into $\sim np/2$ subsets of order $\Theta((np)^{k-2})$ and X' also into $\sim np/2$ subsets of order $\Theta((np)^{k-2})$. Thus, since $p(np)^{k-2} = (np)^{k-1}/n = o(1)$ by assumption, the edge probability is given by

$$q = 1-(1-p/2)^{\Theta((np)^{k-2})} = \Theta(n^{k-2}p^{k-1}).$$

Now, set $m = \Theta(np)$ and observe that the number of edges in $H(X,\{v\})$ as well as in $H(X',\{v\})$ has binomial distribution $\text{Bin}(m,q)$ with the expected value

$$\mu = mq = \Theta(np \cdot n^{k-2}p^{k-1}) = \Theta((np)^k/n) \gg \log n,$$

by assumption. Thus, the number of edges in these stars is at least $\mu/2$ with probability at least $1-e^{-\Omega(\mu)}$. Since $\mu \gg \log n$, this probability suffices to overcome the union bound of $\binom{n}{2}$ choices, completing the proof. $\square$

3.6 Remarks and Further Directions

Here we provide two preliminary results that suggest further directions for study. One involves the parameter $\lambda_{r,\ell}(G)$, introduced in [14], in which we relax the requirement from $\kappa_{r,\ell}(G)$ that alternating paths between vertices be disjoint. The second result is a pseudorandom analog of Theorem 1.5.2.

3.6.1 Not Necessarily Disjoint Paths

Removing the restriction that alternating paths in G be internally disjoint gives us the number $\lambda_{r,\ell}(G)$, which is the maximum t such that there is an r-coloring of the edges of G such that any pair of vertices is connected by t alternating paths of length ℓ. A few results for this number were obtained in [14] where it was shown that

$$\lambda_{2,3}(K_{m,n}) \sim mn/4 \quad \text{and} \quad \lambda_{2,4}(K_{m,n}) \sim m^2n/8.$$

Determining $\lambda_{2,\ell}(K_{m,n})$ for general ℓ seems to be not an easy problem. Here we provide the following result for general G and $\ell = 3$.

Proposition 3.6.1. *Let G be a d-regular graph. Then $\lambda_{2,3}(G) \leq d^3/(4(n-1))$.*

Proof. Let the edges of G be 2-colored. Then we claim that the number of all alternating paths of length 3 is at most $nd^3/8$, which will yield the result since

$$\binom{n}{2}\lambda_{2,3}(G) \leq nd^3/8.$$

Let G be a d-regular graph with $G = ([n], E)$ and 2-colored edges. Then let $E = R \cup B$ where R and B are the preimage of red and blue under c, respectively. Then we have a red and a blue degree sequence

$$r_1 \leq r_2 \leq \ldots \leq r_n,$$
$$b_1 \geq b_2 \geq \ldots \geq b_n,$$

where $r_i, b_i = d - r_i$ is the red and blue degree of vertex i, respectively (under a possible reordering of the vertices). Thus the total number of alternating paths of length 3 in G is

$$\sum_{ij \in B} r_i r_j + \sum_{k\ell \in R} b_k b_\ell.$$

We claim that this is at most

$$\sum_{i \in [n]} \left(r_i^2 \cdot \frac{b_i}{2} + b_i^2 \cdot \frac{r_i}{2} \right).$$

To do this, we use the following version of the rearrangement inequality (see, for example, [40]): For $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_n$ and $y_1 \leq y_2 \leq y_3 \leq \ldots \leq y_n$ and any permutation π of $[n]$, we have

$$\sum_{i=1}^{n} x_i y_{\pi(i)} \leq \sum_{i=1}^{n} x_i y_i.$$

Define $B(i) := \{j \in [n] \setminus \{i\} : ij \in B\}$. Note that $|B(i)| = b_i$.

$$\sum_{ij \in B} r_i r_j \leq \frac{1}{2} \sum_{i=1}^{n} \sum_{j \in B(i)} r_i r_j.$$

We now think of this last sum as $\sum_{i=1}^{m} x_i y_{\pi(i)}$ where $m = b_1 + b_2 + \ldots + b_n$ and

$$\begin{aligned} x_1 = x_2 = x_3 = \ldots = x_{b_1} &:= r_1 \\ x_{b_1+1} = x_{b_1+2} = x_{b_1+3} := \ldots = x_{b_1+b_2} &:= r_2 \\ &\ldots \\ x_{b_1+\ldots+b_{n-1}+1} = \ldots = x_m &:= r_n, \end{aligned}$$

$$\begin{aligned} y_1 = y_2 = y_3 = \ldots = y_{b_1} &:= r_1 \\ y_{b_1+1} = y_{b_1+2} = y_{b_1+3} := \ldots = y_{b_1+b_2} &:= r_2 \\ &\ldots \\ y_{b_1+\ldots+y_{n-1}+1} = \ldots = y_m &:= r_n. \end{aligned}$$

And observe that $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_m$ and $y_1 \leq y_2 \leq y_3 \leq \ldots \leq y_m$ since $r_1 \leq r_2 \leq \ldots \leq r_n$.

But since for $1 \leq i \leq n$ we have that r_i appears in exactly b_i of the sums $\sum_{j \in B(i)} r_j$, then there exists a permutation of $[m]$ that takes the ordering of the indices we get in $\sum_{i=1}^{n} \sum_{j \in B(i)} r_i r_j$

to $[m]$. Let π be the inverse of this permutation. Then by the rearrangement inequality,

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j\in B(i)} r_i r_j = \frac{1}{2}\sum_{i=1}^{m} x_i y_{\pi(i)} \leq \frac{1}{2}\sum_{i=1}^{m} x_i y_i = \frac{1}{2}\sum_{i=1}^{n} r_i^2 b_i$$

We apply the same argument to the $\sum_{k\ell\in R} b_k b_\ell$ with

$$\begin{aligned} x_1 = x_2 = x_3 = \ldots = x_{r_n} &:= b_n \\ x_{r_n+1} = x_{r_n+2} = x_{r_n+3} := \ldots = x_{r_n+r_{n-1}} &:= b_{n-1} \\ &\ldots \\ x_{r_n+\ldots+r_2+1} = \ldots = x_{m'} &:= b_1 \end{aligned}$$

and

$$\begin{aligned} y_1 = y_2 = y_3 = \ldots = y_{r_n} &:= b_n \\ y_{r_n+1} = y_{r_n+2} = y_{r_n+3} := \ldots = y_{r_n+r_{n-1}} &:= b_{n-1} \\ &\ldots \\ y_{r_n+\ldots+r_2+1} = \ldots = y_{m'} &:= b_1, \end{aligned}$$

where $m' = r_n + \ldots + r_1$. Then $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_m$ and $y_1 \leq y_2 \leq y_3 \leq \ldots \leq y_m$ since $b_n \leq b_{n-1} \leq \ldots \leq b_1$.

So we obtain

$$\sum_{ij\in B} r_i r_j + \sum_{k\ell\in R} b_k b_\ell \leq \sum_{i\in[n]} \left(r_i^2 \cdot \frac{b_i}{2} + b_i^2 \cdot \frac{r_i}{2} \right) = \sum_{i\in[n]} \frac{r_i b_i}{2}(r_i + b_i) = \frac{d}{2}\sum_{i\in[n]} r_i b_i.$$

But this sum is maximized when $r_i = b_i = d/2$ so the total number of alternating paths of length three at most $\frac{d^3 n}{8}$. □

It is not difficult to show that the random two-coloring of the edges of $G = G(n,p)$

implies that w.h.p. $\lambda_{2,3}(G) \geq (1+o(1))n^2p^3/8$ for $np \gg (n\log n)^{1/3}$. This together with a slightly modified proof of Proposition 3.6.1 (where the d-regular requirement is replaced by almost d-regular) shows that w.h.p. $\lambda_{2,3}(G) \sim n^2p^3/8$. It is plausible to believe that the random two-coloring of $G(n,p)$ always maximizes the parameter $\lambda_{2,\ell}$ for any $\ell \geq 4$.

3.6.2 Alternating Paths of Length at Least Three in Pseudorandom Graphs

We say G is *d-pseudorandom* if all degrees are $\sim d$ and all codegrees are $\sim d^2/n$, for $d \gg n^{1/2}$. Recall that we have already determined the alternating connectivity for d-pseudorandom graphs when we have paths of length two (see Theorem 3.3.1). More generally, we say that G is a *(n,d,λ)-pseudorandom* if it has n vertices, all vertices have degree $\sim d$, and all eigenvalues except the largest have absolute value at most λ. It is well known that any d-pseudorandom graph G is an $(n,d,o(d))$-graph.

The interested reader can verify that the following result about $\kappa_{r,\ell}$(G) for G a d-pseudorandom.

Theorem 3.6.2. *Suppose G is a (n,d,λ)-graph, and $\lambda \ll d^2/n$. Then for all fixed $\ell \geq 3$ we have*

$$\kappa_{r,\ell}(G) \sim \min\left\{\frac{n}{\ell-1}, d-\frac{d^2}{2n}\right\}.$$

Here we follow the same strategy as the proof of Theorem 1.5.2 by first coloring each edge randomly and independently and then finding appropriate disjoint sets and colored matchings between them. The main difference is that instead of Lemma 3.2.1 one can use its analog, given below.

Lemma 3.6.3. *Let G be an (n,d,λ)-graph with $\lambda = o(d)$ and let A,B be disjoint sets of $m \gg (\lambda/d)n$ vertices each. Then G has a matching from A to B containing $\sim m$ edges.*

Proof. The proof is based on an easy application of the Expander Mixing Lemma [2] that asserts that if G is an (n,d,λ)-graph, then for any $S,T \subseteq V(G)$ we have

$$\left| e(S,T) - \frac{d|S||T|}{n} \right| \le \lambda\sqrt{|S||T|} + o\left(\frac{d|S||T|}{n}\right). \tag{3.7}$$

Let $\delta = \frac{2\lambda n}{dm}$. By assumption, $\delta = o(1)$. Notice that

$$|B \setminus N(S)| = m - |N(S) \cap B| \ge (1+\delta)m - |S|$$

and $e(S, B \setminus N(S)) = 0$. But by (3.7), we must have $e(S,T) > 0$ whenever $|S||T| \ge 2\left(\frac{\lambda}{d}\right)^2 n^2$. Letting $T = B \setminus N(S)$, we get

$$|S||T| \ge |S|[(1+\delta)m - |S|] \ge \delta^2 m^2 = 4\left(\frac{\lambda}{d}\right)^2 n^2$$

which is a contradiction. □

Here we have an answer for $d \gg (\lambda n)^{1/2}$. The main obstacle in obtaining an analogous result for smaller d is that we rely on the Expander Mixing Lemma which is “too strong" in the sense that it is a statement about all sets of vertices S,T, and this comes at a price of being “too weak" in the error term (the little-o term in (3.7)). This is in contrast to random graphs where we are able to handle the sparser cases because we rely only on a similar statement about relatively few (order n^2) pairs of sets S,T. For (n,d,λ)-graphs we do not have any analogous tool, i.e. one that has a smaller error term and still tells us what we need to know about the specific sets S,T that concern our proof.

3.6.3 Remark on the Windows Used in the Sparse Case

Here we comment on the windows of p used in Theorem 1.5.3 for parts (ii) and (iii). By stipulating that $(n \log n)^{1/k} \ll np$, we avoid when the diameter is changing. Analyzing $\kappa_{r,l}(G)$ for this range seems to require different strategies than the ones we have used here.

Thus, we have small gaps where we do not know what is happening. Further, although we have determined the order of magnitude in (iii), we still leave open the exact constant for this range. An avenue for further work, then, would be to make this result more precise in both of these areas.

Bibliography

[1] R. Aharoni and S. Zerbib. A generalization of tuza's conjecture. *J. Graph Theory*, 94(3):445–462, 2020.

[2] J. Akiyama, Y. Egawa, and H. Enomoto, editors. *Explicit construction of linear sized tolerant networks*. Elsevier Science B.V., Amsterdam, 1988. Held in Hakone, June 1–5, 1986, Discrete Math. **72** (1988), no. 1-3.

[3] M. Albert and A. Frieze. Multicoloured Hamilton cycles. *Electron. J. Combin.*, 2:Research Paper 10, approx. 13, 1995.

[4] N. Alon and G. Gutin. Properly colored hamilton cycles in edge-colored complete graphs. *Random Structures Algorithms*, 11(2):179–186, 1997.

[5] N. Alon and J. Spencer. *The Probabilistic Method*. Wiley Publishing, 4th edition, 2016.

[6] N. Alon and R. Yuster. On a hypergraph matching problem. *Graphs Combin.*, 21(4):377–384, 2005.

[7] K. Azuma. Weighted sums of certain dependent random variables. *Tohoku Math. J. (2)*, 19:357–367, 1967.

[8] J. Bang-Jensen and G. Gutin. Alternating cycles and paths in edge-coloured multigraphs: a survey. *Discrete Math.*, 165/166:39–60, 1997. Graphs and combinatorics (Marseille, 1995).

[9] J. D. Baron and J. Kahn. Tuza's conjecture is asymptotically tight for dense graphs. *Combin. Probab. Comput.*, 25(5):645–667, 2016.

[10] A. Basit and D. Galvin. Private communication.

[11] P. Bennett, R. Cushman, and A. Dudek. Closing the random graph gap in tuza's conjecture through the online triangle packing process, 2020. Available online: `arXiv:2007.04478`. Submitted to SIDMA.

[12] P. Bennett, R. Cushman, and A. Dudek. On the number of alternating paths in random graphs, 2020. Available online: `arXiv:2001.09369`. Submitted to Discrete Appl. Math.

[13] P. Bennett and A. Dudek. A gentle introduction to the differential equation method and dynamic concentration, 2020. Available online: `arXiv:2007.01994`.

[14] P. Bennett, A. Dudek, and E. Laforge. On the number of alternating paths in bipartite complete graphs. *Graphs Combin.*, 33(2):307–320, 2017.

[15] P. Bennett, A. Dudek, and S. Zerbib. Large triangle packings and Tuza's conjecture in sparse random graphs. *Combin. Probab. Comput.*, 29(5):757–779, 2020.

[16] T. Bohman. The triangle-free process. *Adv. Math.*, 221(5):1653–1677, 2009.

[17] T. Bohman, A. Frieze, and E. Lubetzky. Random triangle removal. *Adv. Math.*, 280:379–438, 2015.

[18] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. In *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, volume 16 of *CRM Series*, pages 489–495. Ed. Norm., Pisa, 2013.

[19] B. Bollobás. To prove and conjecture: Paul Erdős and his mathematics. *Amer. Math. Monthly*, 105(3):209–237, 1998.

[20] B. Bollobás. The life and work of paul Erdős. In S. S. Chern and F. Hirzebruch, editors, *Wolf Prize in Mathematics*, volume 1, pages 292–342. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[21] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001.

[22] B. Bollobás and P. Erdős. Alternating hamiltonian cycles. *Israel J. Math.*, 23(2):126–131, 1976.

[23] B. Bollobás and O. Riordan. Random graphs and branching processes. In *Handbook of large-scale random networks*, volume 18 of *Bolyai Soc. Math. Stud.*, pages 15–115. Springer, Berlin, 2009.

[24] B. Bollobás and A. Thomason. Hereditary and monotone properties of graphs. In *The mathematics of Paul Erdős, II*, volume 14 of *Algorithms Combin.*, pages 70–78. Springer, Berlin, 1997.

[25] C. C. Chen and D. E. Daykin. Graphs with hamiltonian cycles having adjacent lines different colors. *J. Combinatorial Theory Ser. B*, 21(2):135–139, 1976.

[26] A. Dudek, A. Frieze, and A. Ruciński. Rainbow hamilton cycles in uniform hypergraphs. *Electron. J. Combin.*, 19(1):Paper 46, 11, 2012.

[27] P. Erdős. On some extremal problems in graph theory. *Israel J. Math.*, 3:113–116, 1965.

[28] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[29] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[30] P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.

[31] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Math. Acad. Sci. Hungar.*, 12:261–267, 1961.

[32] P. Erdős and A. Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar.*, 17:359–368, 1966.

[33] P. Erdős, S. Suen, and P. Winkler. On the size of a random maximal graph. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6(2-3), pages 309–318, 1995.

[34] L. Espig, A. Frieze, and M. Krivelevich. Elegantly colored paths and cycles in edge colored random graphs. *SIAM J. Discrete Math.*, 32(3):1585–1618, 2018.

[35] P. Frankl and V. Rödl. Near perfect coverings in graphs and hypergraphs. *European J. Combin.*, 6(4):317–326, 1985.

[36] P. Frankl and V. Rödl. Large triangle-free subgraphs in graphs without K_4. *Graphs Combin.*, 2(2):135–144, 1986.

[37] David A. Freedman. On tail probabilities for martingales. *Ann. Probability*, 3:100–118, 1975.

[38] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

[39] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959.

[40] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.

[41] P. E. Haxell. Packing and covering triangles in graphs. *Discrete Math.*, 195(1-3):251–254, 1999.

[42] P. E. Haxell and V. Rödl. Integer and fractional packings in dense graphs. *Combinatorica*, 21(1):13–38, 2001.

[43] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[44] S. Janson, T. Luczak, and A. Rucinski. *Random Graphs*. Wiley Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2009.

[45] J. Kahn and J. Park. Tuza's conjecture for random graphs, 2020. Available online: `arXiv:2007.04351`.

[46] R. M. Karp and M. Sipser. Maximum matching in sparse random graphs. In *Proceedings of the Twenty-Second Annual IEEE Symposium on Foundations of Computing*, pages 364–375. SFCS '81, 1981.

[47] M. Krivelevich. On a conjecture of Tuza about packing and covering of triangles. *Discrete Math.*, 142(1-3):281–286, 1995.

[48] T. Makai. The reverse H-free process for strictly 2-balanced graphs. *J. Graph Theory*, 79(2):125–144, 2015.

[49] J. S. Petrovic. *Advanced Calculus: Theory and Practice*. Textbooks in Mathematics. Chapman and Hall/CRC, 2014.

[50] V. Pless. *Introduction to the theory of error-correcting codes*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 3 edition, 1998. A Wiley-Interscience Publication.

[51] G. F. Pontiveros, S. Griffiths, and R. Morris. *The Triangle-Free Process and the Ramsey Number $R(3,k)$*. Mem. Amer. Math. Soc. American Mathematical Society, 2020.

[52] F. P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[53] A. Ruciński and N. C. Wormald. Random graph processes with degree restrictions. *Combin. Probab. Comput.*, 1(2):169–180, 1992.

[54] J. Shearer. A property of the colored complete graph. *Discrete Math.*, 25(2):175–178, 1979.

[55] Zs. Tuza. A conjecture. In A. Hajnal, L. Lovász, and V.T. Sós, editors, *Proc. Colloq. Math. Soc. J. Bolyai*, volume 37, page 888. North-Holland, 1984. Held in Eger, July 1981.

[56] L. Warnke. On the method of typical bounded differences. *Combin. Probab. Comput.*, 25(2):269–299, 2016.

[57] N. C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, pages 73–155. Wydawnictwo Naukowe Pwn, 1999.

[58] R. Yuster. Dense graphs with a large triangle cover have a large triangle packing. *Combin. Probab. Comput.*, 21(6):952–962, 2012.