Carnegie Mellon University
Department of Mathematical Sciences

Doctoral Dissertation

Random Graphs and Algorithms

Tony Johansson

January 2017

Submitted to the Department of Mathematical Sciences
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Algorithms, Combinatorics and Optimization

Dissertation Committee

Alan Frieze, chair
Tom Bohman
Andrzej Dudek
Wesley Pegden

Acknowledgements and Thanks

First and foremost, I would like to thank my advisor Alan Frieze for support, training, and belief in my abilities from the day I came to Pittsburgh, even when I doubted them myself.

Thanks to my committee; Tom Bohman, Andrzej Dudek, Alan Frieze and Wesley Pegden. Thanks also to the faculty and staff at the mathematics department at Carnegie Mellon, in particular Tom Bohman, Boris Bukh, Alan Frieze and Po-Shen Loh for teaching my combinatorics classes.

Thanks to Joseph Briggs, Xiao Chang, Francisco Cisternas, Tarek Elgindy, Giovanni Gravina, Wony Hong, Clive Newstead, Hân Nguyen, Youngmin Park, Ananya Uppal, Son Van, Shanning Wan, Wilson Ye and Andy Zucker for making life in Pittsburgh enjoyable.

Thanks also to Robert Andersson, Emilia Ingemarsdotter, Jenny Larsson, Isak Lyckberg, Ester Sandström and Emma Wijkmark for making the trip from Sweden to visit me during my years in the United States.

Lastly, thank you to my sister Caroline, my brother Jimmie and my parents Märit and Tomas, and my grandfather Lennart.

Large parts of the text in this thesis also appear in:

- *On random k-out sub-graphs of large graphs* (with Alan Frieze)
  To appear in *Random Structures and Algorithms.*
- *On the insertion time of random walk cuckoo hashing* (with Alan Frieze)
  To be presented at *ACM-SIAM Symposium on Discrete Algorithms*
- *Deletion of oldest edges in a preferential attachment graph*
  Pre-print.
- *Minimum-cost matching in a random graph with random costs* (with Alan Frieze)
  To appear in *SIAM Journal of Discrete Mathematics*
- *On edge disjoint spanning trees in a randomly weighted complete graph* (with Alan Frieze)
  Pre-print.

Abstract

This thesis is concerned with the study of random graphs and random algorithms. There are three overarching themes. One theme is sparse random graphs, i.e. random graphs in which the average degree is bounded with high probability. A second theme is that of finding spanning subsets such as spanning trees, perfect matchings and Hamilton cycles. A third theme is solving optimization problems on graphs with random edge costs. The research contributions of the thesis are separated into five chapters. The topics of the chapters are similar but separate, and can be read in any order. Each chapter fits at least one of the themes, while each theme fails to feature in at least one chapter.

In Chapter 2 we consider random k-out subgraphs G_k of general graphs G with minimum degree $\delta(G) \geq m$ for some m that tends to infinity with the size of G. We show that if $k \geq 2$ then G_k is k-connected with high probability. For a fixed $\varepsilon > 0$ we show that if k is large enough then G_k contains a cycle of length $(1-\varepsilon)m$ with high probability. When $m \geq (1/2+\varepsilon)n$ we strengthen this to showing that G_k contains a Hamilton cycle with high probability.

In Chapter 3 we analyze the random walk cuckoo hashing algorithm for finding L-saturating matchings in a random bipartite graph on vertex set $L \cup R$. It is shown that the algorithm has expected insertion time $O(1)$.

In Chapter 4 we introduce a variation on the Barabási-Albert preferential attachment graph in which edges are removed in an on-line fashion. The asymptotic behaviour of the degree sequence is determined, as well as conditions for the existence of a giant component.

In Chapter 5 we consider the following optimization problem. Let $G = G_{n,p}$ or $G = G_{n,n,p}$, and after generating G assign random costs to each edge, independently exponentially distributed with mean 1. We show that the expected minimum-cost perfect matching converges to $\pi^2/(12p)$ for $G = G_{n,p}$ and $\pi^2/(6p)$ for $G = G_{n,n,p}$ when $np \gg \log^2 n$. This generalizes a well-known result for the case $p = 1$.

Finally, in Chapter 6 we consider the complete graph K_n in which each edge is independently assigned a uniform $[0,1]$ cost. We exactly determine the expected minimum total cost of two edge-disjoint spanning trees, and show that the minimum total cost of k edge-disjoint spanning trees is proportional to k^2 for large k.

Contents

Chapter 1

Introduction

Complex networks are found in many areas of modern life. Devices connected to the internet form a massive network spanning the entire globe, while roads, electrical grids and airline travel form other infrastructural networks which humans depend on. More abstractly, human interaction and relations can be seen as enormous complex networks, to which the term "social network" refers. These networks are typically too large and complex to study directly, and we turn to mathematical models which imitate the behavior of real-world networks. A mathematical description which approximates the real network simplifies calculations, and may help in understanding *how* a network obtained its shape. Real-world networks are often formed by interactions between its members which can be viewed as random, and random graphs provide a tool with which networks are modelled.

A random graph is a graph which is chosen randomly from some large class $\mathcal{G}$ of graphs according to some distribution specified by the random graph model used. In random graph theory, we answer the question "what are the properties of a typical member of $\mathcal{G}$?", as opposed to the more classical mathematical question "what are the properties shared by every member of $\mathcal{G}$?". Rather than pure enumeration, we are interested in using probabilistic tools and asymptotic approximations to estimate the proportion of $\mathcal{G}$ which has a certain property.

Random graphs have applications outside of modelling real-world networks. Graph theoretic algorithms in theoretical computer science are plentiful, and random graph problems arise with the use of modern randomized algorithms. In other areas of combinatorics, some random graphs have properties which are difficult to construct "by hand". Some of the early applications of probabilistic ideas in graph theory appeared in Ramsey Theory, e.g. Szele [85] who showed the existence of tournaments on n vertices with $n!2^{1-n}$ Hamilton cycles and Erdős [24] who showed that the diagonal Ramsey number $R(k)$ is greater than $2^{k/2}$.

The study of random graphs was initiated by Erdős and Rényi [25], and independently by Gilbert [54], in the late 1950's. The former introduced the graph $G_{n,m}$, a graph chosen uniformly at random from all simple graphs on n vertices containing exactly m edges, while Gilbert considered the closely related graph $G_{n,p}$, a graph on n vertices in which each possible edge is included with probability $p \in (0, 1)$. The models $G_{n,m}$ and $G_{n,p}$ are now known as the Erdős-Rényi or Erdős-Rényi-Gilbert graph, and has been studied extensively since its introduction, most notably by Erdős and Rényi [26], [27], [28], [29]. The field was later dominated by Béla Bollobás, and his 1985 book [12] would help solidify the field as a branch of mathematics and attract new researchers to the field.

1.1 Preliminaries

The theory of random graphs lies at the intersection of graph theory and probability theory. While the reader is expected to have some familiarity with these branches of mathematics we now briefly introduce the subjects, establishing conventions and notation used throughout the thesis.

1.1.1 Graph theory

We define an *graph* G as a pair of sets (V, E) where V is the *vertex* set and $E \subseteq V \times V$ is the *edge* set. The graph may be *undirected* in which order is ignored, and we write $\{u, v\}$ for the edge (u, v) and (v, u), or *directed*. In an undirected graph we view (u, v) and (v, u) as one edge. Throughout this thesis, the sets V, E will be assumed to be finite. We allow E to be a multiset, i.e. we allow it to contain more than one copy of the same edge (*parallel edges*). If E contains no parallel edges and no edge of the form (v, v) (a *self-loop*), we say that G is *simple*. In most applications we are only interested in simple graphs. We write $e(G) = |E(G)|$ and $v(G) = |V(G)|$ for the number of edges and vertices in G, respectively.

In an undirected graph we say that $u, v \in V$ are *adjacent* or *neighbors* if $\{u, v\} \in E$, in which case we write $u \sim v$. The *neighborhood* of $v \in V$ is the set $N(v) = \{u \in V : \{u, v\} \in E\}$, and the *degree* of v, denoted $d(v)$ or $\deg v$, is the size of $N(v)$. In a directed graph we let $N^+(v) = \{u \in V : (v, u) \in E\}$ denote the *out-neighborhood* of v, and $d^+(v) = |N^+(v)|$ is the *out-degree* of v. The in-neighborhood $N^-(v)$ and in-degree $d^-(v)$ of v are defined similarly.

A *path* in an undirected graph is a sequence of distinct vertices $v_0, v_1, \ldots, v_k$ where $e_i = \{v_{i-1}, v_i\} \in E$ for $i = 1, \ldots, k$. Where convenient we may also view the path as the edge sequence $e_1, \ldots, e_k$. A *cycle* is a path $v_0, v_1, \ldots, v_k$ along with the edge $\{v_k, v_0\}$. In a directed graph we require orientations to be consistent; we require $(v_{i-1}, v_i) \in E$ for $i = 1, \ldots, k$ in paths and cycles, and $(v_k, v_0) \in E$ in cycles. A *walk* is a path in which repetitions of vertices and edges are allowed. A *Hamilton cycle* is a cycle which covers all vertices of the graph.

Given a graph $G = (V, E)$ we say that H is a *subgraph* of G, denoted $H \subseteq G$, if $H = (V, E')$ where $E' \subseteq E$. Note in particular that H is defined on the same vertex set as G. Given $W \subseteq V$, the *induced subgraph* of G on W, denoted $G[W]$, is the graph on vertex set W whose edge set is $\{(u, v) \in E : u, v \in W\}$.

We define a *tree* as a connected graph which contains no cycles. A tree may be viewed as a minimal connected graph, in the sense that removing any edge disconnects the graph. A *forest* is a graph with no cycles, i.e. a collection of disjoint trees. Given a graph G, a *spanning tree* of G is a subgraph of G which is a tree on all vertices of G.

A *matching* is a set of edges M such that each vertex is incident to at most one edge of M, or in other words no two edges of M meet at any vertex. A *perfect matching* in a graph $G = (V, E)$ is a matching $M \subseteq E$ which is incident to each vertex exactly once.

1.1.2 Discrete probability theory

The random graphs considered in this thesis are of large finite size. We now set up the necessary probabilistic and asymptotic notation required to study random graphs.

We are typically concerned with the asymptotic behaviour of a sequence $\{G_n : n \in \mathbb{N}\}$ of random graphs as n tends to infinity. To be precise, let $\mathcal{G}_1, \mathcal{G}_2, \dots$ be an infinite sequence of finite collections of finite graphs. The size of a typical element in $\mathcal{G}_n$ should increase with n; in many cases all graphs in $\mathcal{G}_n$ have exactly n vertices, but this is not always the case (e.g. Chapter 4). Let G_n be a random member of $\mathcal{G}_n$ chosen according to some distribution specified by a model, typically with one or more real numbers as parameters. Define a *property* to be a collection $\mathcal{P} = \{P_n \subseteq \mathcal{G}_n : n \in \mathbb{N}\}$. We say that the random sequence $\{G_n : n \in \mathbb{N}\}$ has property $\mathcal{P}$ with high probability (w.h.p.) if

$$\lim_{n\to\infty} \Pr\{G_n \in P_n\} = 1.$$

The sequence $\{G_n : n \in \mathbb{N}\}$ is typically understood to be implicit, and we say that G_n has property $\mathcal{P}$ with high probability.

As statements are made about the asymptotic behaviour of graphs, asymptotic notation is frequently used. For functions $f(n)$ and $g(n) \neq 0$ we write $f(n) = O(g(n)), f(n) = o(g(n))$ if there exists a constant $C > 0$ such that

$$\lim_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq C, \quad \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0,$$

respectively. Define $f(n) = \Omega(g(n)), f(n) = \omega(g(n))$ if $f(n) \neq 0$ and $g(n) = O(f(n))$ and $g(n) = o(f(n))$ respectively. At some points we prefer to write $f(n) \ll g(n)$ for $f(n) = o(g(n))$ and $f(n) \gg g(n)$ for $f(n) = \omega(g(n))$.

Unless a probability distribution is specified, the phrase "a random $x \in X$" should be interpreted as a member x of the (finite) set X chosen uniformly at random from X.

1.1.3 Random graph models

We now define three well-known models of random graphs which will be referenced in the thesis. The most well-known graph is the Erdős-Rényi graph $G_{n,p}$ (introduced by Gilbert [54]), the random graph obtained by deleting each edge of K_n independently with probability $1-p$. Here $p = p(n)$ typically depends on n. Let $\mathcal{G}_{n,m}$ be the class of graphs on n vertices with $m = m(n)$ edges and define $G_{n,m}$ as a uniformly random member of $\mathcal{G}_{n,m}$. The graph $G_{n,m}$ is closely related to $G_{n,p}$ via the following identity:

$$\forall G \in \mathcal{G}_{n,m} : \quad \Pr\{G_{n,p} = G \mid e(G_{n,p}) = m\} = \Pr\{G_{n,m} = G\}.$$

For a fixed positive integer k we define the *k-out random graph* $K_n(k-\text{out})$ on n vertices as follows. Each vertex v chooses a set N_v of k vertices uniformly at random, with or without replacement, and $K_n(k-\text{out})$ includes the edge $e = \{v, w\}$ if and only if $v \in N_w$ or $w \in N_v$. A bipartite version of this graph was first considered by Walkup [89]. The k-out graph achieves strong connectivity properties such as being k-connected [31] and containing spanning subgraphs such as perfect matchings and Hamilton cycles with only $O(n)$ edges (e.g. [10], [43]), while $G_{n,p}$ requires $\Omega(n \log n)$ edges to achieve the same properties.

One class of random graphs which have gained popularity in recent years are *preferential attachment graphs*. For a fixed $m \geq 1$ we will consider the following definition of such a graph, first rigorously defined in [7]. Given a graph G_t, we define G_{t+1} by adding a single vertex along with m edges. The m edges are attached to vertices in G_t with probabilities proportional to the degrees of the vertices. Starting at some fixed graph G_0, this defines a sequence $G_0, G_1, \dots$, and we consider some G_n where $n \gg v(G_0)$. Preferential attachment graphs were proposed by Barabási and Albert [2] as a way of modelling real-world networks, as such networks frequently have a degree sequence which follows a power law.

All random graphs considered in this thesis are related to the above three models, while models such as the geometric graph, random regular graphs and infinite graphs such as the $\mathbb{Z}^n$ lattice are not considered.

1.2 History and contributions

In this thesis we present five separate contributions to the fields of random graphs and random algorithms.

1.2.1 Long cycles in k-out subgraphs of large graphs

Traditionally, most results on (finite) random graphs are based on studying properties of random subgraphs of K_n or $K_{n,n}$. Recently however, Krivelevich, Lee and Sudakov [65] considered Erdős-Rényi subgraphs of some more general graph G. They define G_p by deleting any edge of G with probability $1-p$, independently. Their main result is that if G has minimum degree $\delta(G) \geq m$ and $p \gg 1/m$ then G_p contains a cycle of length $(1 - o_m(1))m$ with probability $1 - o_m(1)$. Riordan [81] gave a simpler proof of this result.

In [40] we consider k-out subgraphs of some arbitrary graph G on n vertices, assuming $\delta(G) \geq m$ for some $m = m(n)$ that tends to infinity with n. Define G_k, the k-out subgraph of G, to be the random graph given by each vertex protecting k of its incident edges uniformly at random, and removing any unprotected edge. We show that there exists a k_ε such that if $k \geq k_\varepsilon$, the k-out random subgraph of G has a cycle of length at least $(1-\varepsilon)m$ with high probability.

When $m \geq (1/2 + \varepsilon)n$ for some constant $\varepsilon > 0$, we also prove that k-out subgraphs of G are k-connected with high probability for any $k \geq 2$, and show that there exists a k_ε such that $k \geq k_\varepsilon$ implies that the k-out subgraph of G contains a Hamilton cycle with high probability. This is inspired by a result from Bohman and Frieze [10], who showed that if $k \geq 3$ is enough for k-out subgraphs of K_n to contain a Hamilton cycle with high probability, and by Krivelevich, Lee and Sudakov [66] who showed that there exists a $C > 0$ such that if $\delta(G) \geq n/2$ and $p \geq Cn^{-1} \log n$ then G_p is Hamiltonian with high probability.

In Chapter 2 we present [40] in full.

1.2.2 Random walk cuckoo hashing

Suppose $G = (L + R, E)$ is a bipartite graph which contains a matching of $L = \{v_1, \dots, v_n\}$ into $R = \{w_1, \dots, w_m\}$, $m \geq n$. Define a sequence of random induced subgraphs of G as follows. Let $\emptyset = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_n = L$ be some sequence with $|L_k \setminus L_{k-1}| = 1$ for all $k > 0$ and define $G_k = G[L_k \cup R]$ for $k = 0, 1, \dots, n$. Let v_k be the unique element of $L_k \setminus L_{k-1}$ for all k. We consider the problem of building and maintaining a matching on-line. Presented with $G_1, G_2, \dots, G_n$ in order with no knowledge of G, the aim is to maintain a matching M_k of L_k into R as k increases from 1 to n. We obtain M_k from M_{k-1} via augmenting paths from v_k to some vertex in R which is not incident to M_{k-1}. We wish to bound the total length of augmenting paths in this process. Chaudhuri, Daskalakis, Kleinberg, Lin [17] considered the case $|L| = |R| = n$, and showed that if the vertices of L arrive in a random order then the total length of augmenting paths needed to building a perfect matching, if the shortest possible augmenting path is used at each stage, is at most $n \log n + O(n)$. The problem has also been discussed by Bosek, Leniowski, Sankowski and Zych [15] and Gupta, Kumar and Sten [56].

In [41] (Chapter 3) we consider the problem on a random graph with $n = (1 - \varepsilon)m$ where the neighborhood of each $v \in L$ is a random set of d vertices in R for some constant d, and where augmenting paths are found via random walks. This corresponds to cuckoo hashing, introduced by Pagh and Rodler [77]. We show that if $d \gg 1/\varepsilon$, the expected length of an augmenting path from M_{k-1} to M_k is $O(1)$, for all k. This is the first result with constant expected insertion time, and improves on bounds by Frieze, Melsted and Mitzenmacher [50], Fountoulakis, Panagiotou and Steger [37], Fotakis, Pagh, Sanders and Spirakis [35]. The size of d required for a matching to exist is $d \gg \log(1/\varepsilon)$ ([49], [36]), so there is some room for improvement.

1.2.3 A preferential attachment graph with oldest-edge deletion

The preferential attachment graph, introduced by Albert and Barabási [2], is a popular random graph model for modelling real-world networks whose degree sequence follows a power law. In other words, the proportion of vertices having degree k is proportional to $k^{-\eta}$ for some constant η. A large number of real-world networks including the World Wide Web [30] follow a power law. A long list of real-world networks exhibiting power law behaviour can be found in [7].

Chapter 4 covers [60]. Here we consider a variation of the Barabási-Albert model with fixed parameters $m \in \mathbb{N}$ and $1/2 < p < 1$. A random graph G_n is generated in an on-line manner starting with some small graph G_{t_0}. With probability p a vertex is added along with m edges, each of whose endpoint is randomly chosen with probability proportional to vertex degrees. With probability $1 - p$ we locate the oldest vertex still holding its original m edges, and remove those m edges from the graph. We find a $p_0 \approx 0.83$ such that if $p > p_0$, the resulting graph G_n has a degree sequence that follows a power law, while if $p < p_0$ the degree sequence has an exponential tail. We also show that the graph has a unique giant component with high probability if and only if $m \geq 2$.

While this variation of the preferential attachment model has not been studied before, other preferential attachment models with vertex and edge deletion have previously been considered by Bollobás and Riordan [13], Cooper, Frieze and Vera [20], Flaxman, Frieze and Vera [33]. Chung and Lu [18] considered a general growth-deletion model for random power law graphs.

1.2.4 Minimum-cost matchings in a random graph with random costs

The final two chapters deal with optimization problems on some graph G_n, in which each edge e is assigned a cost $c(e)$, here random, independent and identically distributed for all e. For a class $\mathcal{H}_n$ of subgraphs of G_n we define a random variable

$$Z_n = \min \left\{ \sum_{e \in H} c(e) : H \in \mathcal{H}_n \right\},$$

and we are interested in the limiting behaviour of Z_n as $n \to \infty$. Most often $c(e)$ is uniform $[0,1]$ distributed and $G_n = K_n$ or $G_n = K_{n,n}$. The most studied instances of this problems are (i) spanning trees in K_n, see Frieze [42], (ii) paths in K_n, see Janson [59] (iii) perfect matchings in $K_{n,n}$ (see below) and (iv) Hamilton cycles, see Karp [61], Frieze [44] and Wästlund [93].

In [38] (Chapter 5) we consider perfect matchings in a bipartite graph G where costs are independently exponentially distributed with mean 1. The case $G = K_{n,n}$ has been extensively studied. Walkup [88] and Karp [62] bounded $\mathbf{E}[Z_n]$ before Aldous [3], [4] proved that $\mathbf{E}[Z_n] \to \zeta(2) = \sum_{k \geq 1} k^{-2}$ as $n \to \infty$. Parisi [78] conjectured that $\mathbf{E}[Z_n] = \sum_{k=1}^{n} k^{-2}$ exactly, and this was proved independently by Linusson and Wästlund [68] and by Nair, Prabhakar and Sharma [74]. An elegant proof was given by Wästlund [90], [92], who also extended the proof idea to K_n [91].

In [38] we extend the proof idea of Wästlund, replacing $K_{n,n}$ by the bipartite Erdős-Rényi graph $G_{n,n,p}$. Generating $G = G_{n,n,p}$ first and then assigning random costs, we show that with high probability G is such that $\mathbf{E}[Z_n \mid G_{n,n,p} = G] = p^{-1}\zeta(2) + o(p^{-1})$ as $n \to \infty$ when $p \gg n^{-1}\log^2 n$. For $G_{n,p}$ we similarly show that the expected minimum-cost perfect matching has cost $p^{-1}\zeta(2)/2 + o(p^{-1})$ when $p \gg n^{-1}\log^2 n$.

1.2.5 Minimum-cost disjoint spanning trees in the complete graph with random costs

Our final contribution also falls into the category of optimization problems defined above. It is motivated by the minimum-cost spanning tree problem in K_n with uniform $[0,1]$ edge costs, for which Frieze [42] showed that $\mathbf{E}[Z_n] \to \zeta(3) = \sum_{k \geq 1} k^{-3}$ as $n \to \infty$. Generalizations and refinements of this have been given by Steele [83], Frieze and McDiarmid [48], Janson [58], Penrose [79], Beveridge, Frieze and McDiarmid [9], Frieze, Ruszinko and Thoma [51] and Cooper, Frieze, Ince, Janson and Spencer [19]

In Chapter 6 we are interested in generalizing this to the problem of finding a minimum-weight basis in an element weighted matroid. In the language of matroids, a minimum-cost spanning tree is a minimum-weight basis in the cycle matroid with uniform $[0,1]$ weights. Kordecki and Lyczkowska-Hanćkowiak [64] have shown a general result for this optimization problem on general matroids, but the formulae obtained are somewhat difficult to penetrate. In [39], which Chapter 6 covers, we consider the union of k cycle matroids, for which a basis is given by k edge disjoint spanning trees. For large k we show that the weight of the minimum-cost basis is proportional to k^2, and for $k = 2$ we show that the expected minimum converges to an integral which approximately evaluates to 4.17.

1.3 Future considerations

As mentioned above, while our result on the cuckoo hashing algorithm [41] brings the expected insertion time down to a constant, it does require $d \gg 1/\varepsilon$, which is above the requirement $d \gg \log(1/\varepsilon)$ by a significant margin. More research should be focused on bringing the value of d down.

The preferential attachment graph with oldest-edge deletion has not been studied outside of [60], and many questions remain. Natural properties to investigate are the maximum degree of the graph, the exact size and diameter of the giant component, and connectivity properties. It is also natural to ask what happens when the parameter p is close to the power-law threshold $p_0 \approx 0.83$, as [60] does not address this. A closely variation on the preferential attachment graph is given by removing the oldest *vertex* in the graph with probability $1 - p$, rather than removing the oldest edges. We expect this variation to have properties similar to those of the edge-deletion model.

The generalization from $K_{n,n}$ to $G_{n,n,p}$ in [38] was made in the aim of generalizing from $K_{n,n}$ to general d-regular graphs. We conjecture that the minimum-cost perfect matching in a d-regular graph is $nd^{-1}\zeta(2) + o(nd^{-1})$, likely with some lower bound on d. The calculation of the minimum cost of two disjoint spanning trees in [39] was done in the hope of generalizing the elegant formulae available for spanning subsets such as spanning trees and perfect matchings. The formula obtained for $k = 2$ spanning trees is somewhat complicated, and it might be difficult to find a useful formula for general k. Nevertheless, one direction of future research in random optimization problems is toward simple formulae for general minimum-cost matroids.

Chapter 2

Long cycles in k-out subgraphs of large graphs

This chapter corresponds to [40].

Abstract

We consider random subgraphs of a fixed graph $G = (V, E)$ with large minimum degree. We fix a positive integer k and let G_k be the random subgraph where each $v \in V$ independently chooses k random neighbors, making kn edges in all. When the minimum degree $\delta(G) \geq (\frac{1}{2} + \varepsilon)n$, $n = |V|$ then G_k is k-connected w.h.p. for $k = O(1)$; Hamiltonian for k sufficiently large. When $\delta(G) \geq m$, then G_k has a cycle of length $(1 - \varepsilon)m$ for $k \geq k_\varepsilon$. By w.h.p. we mean that the probability of non-occurrence can be bounded by a function $\phi(n)$ (or $\phi(m)$) where $\lim_{n \to \infty} \phi(n) = 0$.

2.1 Introduction

The study of random graphs since the seminal paper of Erdős and Rényi [26] has by and large been restricted to analysing random subgraphs of the complete graph. This is not of course completely true. There has been a lot of research on random subgraphs of the hypercube and grids (percolation). There has been less research on random subgraphs of arbitrary graphs G, perhaps with some simple properties.

In this vain, the recent result of Krivelevich, Lee and Sudakov [66] brings a refreshing new dimension. They start with an arbitrary graph G which they assume has minimum degree at least k. For $0 \leq p \leq 1$ we let G_p be the random subgraph of G obtained by independently keeping each edge of G with probability p. Their main result is that if $p = \omega/k$ then G_p has a cycle of length $(1 - o_k(1))k$ with probability $1 - o_k(1)$. Here $o_k(1)$ is a function of k that tends to zero as $k \to \infty$. Riordan [81] gave a much simpler proof of this result. Krivelevich and Samotij [67] proved the existence of long cycles for the case where $p \geq \frac{1+\varepsilon}{k}$ and G is $\mathcal{H}$-free for some fixed set of graphs $\mathcal{H}$. Frieze and Krivelevich [47] showed that G_p is non-planar with probability $1 - o_k(1)$ when $p \geq \frac{1+\varepsilon}{k}$ and G has minimum degree at least k. In related works, Krivelevich, Lee and Sudakov [65] considered a random subgraph of a "Dirac Graph" i.e. a graph with n vertices and minimum degree at least $n/2$. They showed that if $p \geq \frac{C \log n}{n}$ for sufffficently large n then G_p is Hamiltonian with probability $1 - o_n(1)$.

The results cited above can be considered to be generalisations of classical results on the random graph $G_{n,p}$, which in the above notation would be $(K_n)_p$. In this paper we will consider generalising another model of a random graph that we will call $K_n(k - out)$. This has vertex set $V = [n] = \{1, 2, \ldots, n\}$ and each $v \in V$ independently chooses k random vertices as neighbors. Thus this graph has kn edges and average degree $2k$. This model in a bipartite form where the two parts of the partition restricted their choices to the opposing half was first considered by Walkup [89] in the context of perfect matchings. He showed that $k \geq 2$ was sufficient for bipartite $K_{n,n}(k - out)$ to contain a perfect matching. Matchings in $K_n(k - out)$ were considered by Shamir and Upfal [82] who showed that $K_n(5 - out)$ has a perfect matching w.h.p., i.e. with probability $1 - o(1)$ as $n \to \infty$. Later, Frieze [43] showed that $K_n(2 - out)$ has a perfect matching w.h.p. Fenner and Frieze [31] had earlier shown that $K_n(k - out)$ is k-connected w.h.p. for $k \geq 2$. After several weaker results, Bohman and Frieze [10] proved that $K_n(3 - out)$ is Hamiltonian w.h.p. To generalise these results and replace K_n by an arbitrary graph G we will define $G(k - out)$ as follows: We have a fixed graph $G = (V, E)$ and each $v \in V$ independently chooses k random neighbors, from its neighbors in G. It will be convenient to assume that each v makes its choices with replacement. To avoid cumbersome notation, we will from now on assume that G has n vertices and we will refer to $G(k - out)$ as G_k. We implicitly consider G to be one of a sequence of larger and larger graphs with $n \to \infty$. We will say that events occur w.h.p. if their probability of non-occurrence can be bounded by a function that tends to zero as $n \to \infty$.

For a vertex $v \in V$ we let $d_G(v)$ denotes its degree in G. Then we let $\delta(G) = \min_{v \in V} d_G(v)$. We will first consider what we call Strong Dirac Graphs (SDG) viz graphs with $\delta(G) \geq \left(\frac{1}{2} + \varepsilon\right) n$ where ε is an arbitrary positive constant.

Theorem 2.1. *Suppose that G is an SDG. Suppose that the k neighbors of each vertex are chosen without replacement. Then w.h.p. G_k is k-connected for $2 \leq k = o(\log^{1/2} n)$.*

If the k neighbors of each vertex are chosen with replacement then there is a probability, bounded above by $1 - e^{-k^2}$ that G_k will have minimum degree $k - 1$, in which case we can only claim that G_k will be $(k - 1)$-connected.

Theorem 2.2. *Suppose that G is an SDG. Then w.h.p. there exists a constant k_ε such that if $k \geq k_\varepsilon$ then G_k is Hamiltonian.*

We get essentially the same result if the k neighbors of each vertex are chosen with replacement.

Note that we need $\varepsilon > 0$ in order to prove these results. Consider for example the case where G consists of two copies of $K_{n/2}$ plus a perfect matching M between the copies. In this case there is a probability greater than or equal to $\left(1 - \frac{2k}{n}\right)^{n/2} \sim e^{-k}$ that no edge of M will occur in G_k.

We note the following easy corollary of Theorem 2.2.

Corollary 2.1. *Let k_ε be as in Theorem 2.2. Suppose that G is an SDG and we give each edge of G a random independent uniform $[0, 1]$ edge weight. Let Z denote the length of the shortest travelling salesperson tour of G. Then $\boldsymbol{E}[(] Z) \leq \frac{2(k_\varepsilon + 1)}{1 + 2\varepsilon}$.*

We will next turn to graphs with large minimum degree, but not necessarily SDG's. Our proofs use Depth First Search (DFS). The idea of using DFS comes from Krivelevich, Lee and Sudakov [66].

Theorem 2.3. *Suppose that G has minimum degree m where $m \to \infty$ with n. For every $\varepsilon > 0$ there exists a constant k_ε such that if $k \geq k_\varepsilon$ then w.h.p. G_k contains a path of length $(1-\varepsilon)m$.*

Using this theorem as a basis, we strengthen it and prove the existence of long cycles.

Theorem 2.4. *Suppose that G has minimum degree m where $m \to \infty$ with n. For every $\varepsilon > 0$ there exists a constant k_ε such that if $k \geq k_\varepsilon$ then w.h.p. G_k contains a cycle of length $(1-\varepsilon)m$.*

We finally note that in a recent paper, Frieze, Goyal, Rademacher and Vempala [45] have shown that G_k is useful in the construction of sparse subgraphs with expansion properties that mirror those of the host graph G.

2.2 Connectivity: Proof of Theorem 2.1

In this section we will assume that each vertex makes its choices without replacement. Let $G = (V, E)$ be an SDG. Let $c = 1/(8e)$. We need the following lemma.

Lemma 2.1. *Let G be an SDG and let $C = 48/\varepsilon$. Then w.h.p. there exists a set $L \subseteq V$, where $|L| \leq C \log n$, such that each pair of vertices $u, v \in V \setminus L$ have at least $12 \log n$ common neighbors in L.*

Proof. Define $L_p \subseteq V$ by including each $v \in V$ in L_p with probability $p = C \log n / 2n$. Since $\delta(G) \geq (1/2 + \varepsilon)n$, each pair of vertices in G has at least $2\varepsilon n$ common neighbors in G. Hence, the number of common neighbors in L_p for any pair of vertices in $V \setminus L_p$ is bounded from below by a $\text{Bin}(2\varepsilon n, p)$ random variable.

$$\begin{aligned}
& \Pr\{\exists u, v \in V \setminus L_p \text{ with less than } 12 \log n \text{ common neighbors in } L\} \\
\leq\ & n^2 \Pr\{\text{Bin}(2\varepsilon n, p) \leq 12 \log n\} \\
=\ & n^2 \Pr\{\text{Bin}(2\varepsilon n, p) \leq \varepsilon n p\} \\
\leq\ & n^2 e^{-\varepsilon n p/8} \\
=\ & o(1).
\end{aligned}$$

The expected size of L_p is $\frac{1}{2} C \log n$ and so the Chernoff bounds imply that w.h.p. $|L_p| \leq C \log n$. Thus there exists a set L, $|L| \leq C \log n$, with the desired property. □

Let L be a set as provided by the previous lemma, and let G'_k denote the subgraph of G_k induced by $V \setminus L$.

Lemma 2.2. *Let $c = 1/(8e)$. If $k \geq 2$ then w.h.p. all components of G'_k are of size at least cn. Furthermore, removing any set of $k-1$ vertices from G'_k produces a graph consisting entirely of components of size at least cn, and isolated vertices.*

Proof. We first show that w.h.p. G'_k contains no isolated vertex. The probability of G'_k containing an isolated vertex is bounded by

$$\Pr\{\exists v \in V \setminus L \text{ which chooses neighbors in } L \text{ only}\} \leq n\left[\frac{C\log n}{\frac{1}{2}n}\right]^k = o(1),$$

where L and C are as in Lemma 2.1.

We now consider the existence of small non-trivial components S after the removal of at most $k-1$ vertices A. Then,

$$\begin{aligned}
&\Pr\{\exists S, A,\ 2 \leq |S| \leq cn,\ |A| = k-1, \text{ such that } S \text{ only chooses neighbors in } S \cup L \cup A\} \\
&\leq \sum_{l=2}^{cn}\sum_{|S|=l}\sum_{|A|=k-1}\left[\frac{l+k-2+C\log n}{\left(\frac{1}{2}+\varepsilon\right)n}\right]^{lk} \\
&\leq \sum_{l=2}^{cn}\binom{n}{l}\binom{n-l}{k-1}\left[\frac{l+C\log n}{\frac{1}{2}n}\right]^{lk} \\
&\leq \sum_{l=2}^{cn}\left(\frac{ne}{l}\right)^l n^{k-1}\left[\frac{l+C\log n}{\frac{1}{2}n}\right]^{lk} \\
&= 2^k e\sum_{l=2}^{cn}\left[\frac{2^k e(l+C\log n)^k}{n^{k-1}l}\right]^{l-1}\frac{(l+C\log n)^k}{l}.
\end{aligned}$$

Now when $2 \leq l \leq \log^2 n$ we have

$$2^k e(l+C\log n)^k \leq \log^{3k} n \text{ and } \frac{(l+C\log n)^k}{l} \leq \log^{3k} n.$$

And when $\log^2 n \leq l \leq cn$ we have

$$2^k e(l+C\log n)^k \leq (2+o(1))^k e l^k \text{ and } \frac{(l+C\log n)^k}{l} = (1+o(1))l^{k-1},$$

which implies that

$$\left[\frac{2^k e(l+C\log n)^k}{n^{k-1}l}\right]^{l-1}\frac{(l+C\log n)^k}{l} \leq \frac{((2+o(1))^k e)^{l-1} l^{l(k-1)}}{n^{(k-1)(l-1)}} \leq$$
$$((2+o(1))^k e)^{l-1} c^{l(k-1)} n^{k-1} = ((2+o(1))^k e)^{l-1} c^{l(k-1)},$$

since $n^{k-1} = (n^{(k-1)/(l-1)})^{l-1} = (1+o(1))^{l-1}$.

Continuing, we get a bound of

$$2^k e\left(\sum_{l=2}^{\log^2 n}\left(\frac{\log^{6k} n}{n^{k-1}}\right)^{l-1} + \sum_{l=\log^2 n}^{cn}((2+o(1))^k e c^{k-1})^{l-1}\right) = o(1).$$

□

This proves that w.h.p. G'_k consists of $r \leq 1/c$ components $J_1, J_2, ..., J_r$ and that removing any $k-1$ vertices will only leave isolated vertices and components of size at least cn.

Lemma 2.3. *W.h.p., for any $i \neq j$, there exist k vertex-disjoint paths (of length 2) from J_i to J_j in G_k.*

Proof. Let X be the number of vertices in L which pick at least one neighbor in J_1 and at least one in J_2. Furthermore, let X_{uvw} be the indicator variable for $w \in L$ picking $u \in J_1$ and $v \in J_2$ as its neighbors. Note that these variables are independent of G'_k. Let $c = 1/(8e)$ as in Lemma 2.2 and let $C = 24/\varepsilon$ as in Lemma 2.1. For $w \in L$ we let

$$X_w = \sum_{\substack{(u,v)\in J_1\times J_2 \\ w\in N_G(J_1)\cap N_G(J_2)}} X_{uvw}.$$

These are independent random variables with values in $\{0, 1, \ldots, k\}$ and $X = \sum_{w\in L} X_w$. Then,

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{u\in J_1}\sum_{v\in J_2}\sum_{\substack{w\in L \\ w\in N(J_1)\cap N(J_2)}} \mathbf{E}[X_{uvw}] \\
&= \sum_{u\in J_1}\sum_{v\in J_2}\sum_{\substack{w\in L \\ w\in N(J_1)\cap N(J_2)}} \frac{\binom{d_G(w)}{k-2}}{\binom{d_G(w)}{k}} \\
&\geq \sum_{u\in J_1}\sum_{v\in J_2}\sum_{\substack{w\in L \\ w\in N(J_1)\cap N(J_2)}} \frac{1}{n^2} \\
&\geq \frac{24(cn)^2\log n}{n^2} \\
&= 24c^2\log n.
\end{aligned}$$

We apply the following inequality, Theorem 1 of Hoeffding [57]: Let $Z_1, Z_2, \ldots, Z_M$ be independent and satisfy $0 \leq Z_i \leq 1$ for $i = 1, 2, \ldots, M$. If $Z = Z_1 + Z_2 + \cdots + Z_M$ then for all $t \geq 0$,

$$\Pr\{|Z - \mathbf{E}[Z]| \geq t\} \leq e^{-2t^2/M}. \tag{2.1}$$

Putting $Z_w = X_w/k$ for $w \in L$ and $Z = X/k$ and applying (2.1), we get

$$\begin{aligned}
\Pr\{X \leq k\} = \Pr\{Z \leq 1\} \leq \Pr\left\{Z \leq \frac{\mathbf{E}[Z]}{2}\right\} &\leq \exp\left\{-\frac{(\mathbf{E}[Z])^2}{2|L|}\right\} \\
&= \exp\left\{-\frac{(\mathbf{E}[X])^2}{2k^2|L|}\right\} = o(1). \qquad (2.2)
\end{aligned}$$

Now for $w_1 \neq w_2 \in L$ let $\mathcal{E}(w_1, w_2)$ be the event that w_1, w_2 make a common choice. Then

$$\Pr\{\exists w_1, w_2 : \mathcal{E}(w_1, w_2)\} = O\left(\frac{k^2\log^2 n}{n}\right) = o(1). \tag{2.3}$$

To see this, observe that for a fixed w_1, w_2 and a choice of w_2, the probability this choice is also one of w_1's is at most $\frac{k}{n/2}$. Now multiply by the number k of choices for w_2. Finally multiply by $|L|^2$ to account for the number of possible pairs w_1, w_2.

Equations (2.2) and (2.3) together show that w.h.p., there are k node-disjoint paths from J_1 to J_2. Since the number of linear size components is bounded by a constant, this is true for all pairs J_i, J_j w.h.p. □

We can complete the proof of Theorem 2.1. Suppose we remove l vertices from L and $k-1-l$ vertices from the remainder of G. We know from Lemma 2.1 that $V \setminus L$ induces components $C_1, C_2, \dots, C_r$ of size at least cn. There cannot be any isolated vertices in $V \setminus L$ as G_k has minimum degree at least k. Recall that each vertex makes k choices without replacement. Lemmas 2.1, 2.2 and 2.3 imply that $r = 1$ and that every vertex in L is adjacent to C_1. □

2.3 Hamilton cycles: Proof of Theorem 2.2

Let G be a graph with $\delta(G) \geq (1/2 + \varepsilon)n$, and let k be a positive integer.

Let $\mathcal{D}(k,n) = \{D_1, D_2, ..., D_M\}$ be the $M = \prod_{v \in V} \binom{d_G(v)}{k} \leq \binom{n-1}{k}^n$ directed graphs obtained by letting each vertex x of G choose k G-neighbors $y_1, ..., y_k$, and including in D_i the k arcs (x, y_i). Define $\vec{N}_i(x) = \{y_1, ..., y_k\}$ and for $S \subseteq V$ let $\vec{N}_i(S) = \bigcup_{x \in S} \vec{N}_i(x) \setminus S$. For a digraph D we let $G(D)$ denote the graph obtained from D by ignoring orientation and coalescing multiple edges, if necessary. We let $\Gamma_i = G(D_i)$ for $i = 1, 2, \dots, M$. Let $\mathcal{G}(k,n) = \{\Gamma_1, \Gamma_2, ..., \Gamma_M\}$ be the set of k-out graphs on G. Below, when we say that D_i is Hamiltonian we actually mean that Γ_i is Hamiltonian. (It will occasionally enable more succint statements).

For each D_i, let $D_{i1}, D_{i2}, ..., D_{i\kappa}$ be the $\kappa = k^n$ different edge-colorings of D_i in which each vertex has $k-1$ outgoing green edges and one outgoing blue edge. Define Γ_{ij} to be the colored (multi)graph obtained by ignoring the orientation of edges in D_{ij}. Let Γ_{ij}^g be the subgraph induced by green edges.

$\vec{N}(S)$ refers to $\vec{N}_i(S)$ when i is chosen uniformly from $[M]$, as it will be for G_k.

Lemma 2.4. *Let $k \geq 5$. There exists an $\alpha > 0$ such that the following holds w.h.p.: for any set $S \subseteq V$ of size $|S| \leq \alpha n$, $|\vec{N}(S)| \geq 3|S|$.*

Proof. The claim fails if there exists an S with $|S| \leq \alpha n$ such that there exists a T, $|T| = 3|S| - 1$ such that $\vec{N}(S) \subseteq T$. The probability of this is bounded from above by

$$\begin{aligned}
& \sum_{l=1}^{\alpha n} \binom{n}{l} \binom{n-l}{3l-1} \prod_{v \in S} \left[\binom{4l-2}{k} \Big/ \binom{d_G(v)}{k} \right] \\
\leq \quad & \sum_{l=1}^{\alpha n} \left(\frac{ne}{l} \right)^l \left(\frac{ne}{3l-1} \right)^{3l-1} \left[\frac{4le}{n/2} \right]^{kl} \\
\leq \quad & \sum_{l=1}^{\alpha n} \left[e^4 (8e)^k \left(\frac{l}{n} \right)^{k-4} \right]^l \\
= \quad & o(1),
\end{aligned}$$

for $\alpha = 2^{-16} e^{-9}$. □

We say that a digraph D_i *expands* if $|\vec{N}_i(S)| \geq 3|S|$ whenever $|S| \leq \alpha n$, $\alpha = 2^{-16}e^{-9}$. Since almost all D_i expand, we need only prove that an expanding D_i almost always gives rise to a Hamiltonian Γ_i. Write $\mathcal{D}'(k,n)$ for the set of expanding digraphs in $\mathcal{D}(k,n)$ and let $\mathcal{G}'(k,n) = \{\Gamma_i : D_i \in \mathcal{D}'(k,n)\}$.

Let H be any graph, and suppose $P = (v_1, ..., v_k)$ is a longest path in H. If $t \neq 1, k-1$ and $\{v_k, v_t\} \in E(H)$, then $P' = (v_1, ..., v_t, v_k, v_{k-1}, ..., v_{t+1})$ is also a longest path of H. Repeating this rotation for P and all paths created in the process, keeping the endpoint v_1 *fixed*, we obtain a set $EP(v_1)$ of other endpoints.

For $S \subseteq V(H)$ we let $N_H(S) = \{w \notin S : \exists v \in S \ s.t. \ vw \in E(H)\}$.

Lemma 2.5 (Pósa)**.** *For any endpoint x of any longest path in any graph H, $|N_H(EP(x))| \leq 2|EP(x)| - 1$.*

We say that an undirected graph expands if $|N_H(S)| \geq 2|S|$ whenever $|S| \leq \alpha n$, assuming $|V(H)| = n$. Note that the definition of expanding slightly differs from the digraph case.

Lemma 2.6. *Consider a green subgraph Γ_{ij}^g. W.h.p., there exists an $\alpha > 0$ such that for every longest path P in Γ_{ij}^g and endpoint x of P, $|EP(x)| > \alpha n$.*

Proof. Let $H = \Gamma_{ij}^g$. We argue that if D_i expands then so does H. If $|\vec{N}_i(S)| \geq 3|S|$, then $|N_H(S)| \geq 2|S|$, since each vertex of S picks at most one blue edge outside of S. Thus H expands. In particular, Lemma 2.4 implies that if $|S| \leq \alpha n$, then $|\vec{N}(S)| \geq 3|S|$ and hence $|N_H(S)| \geq 2|S|$. By Lemma 2.5, this implies that $|EP(x)| > \alpha n$ for any longest path P and endpoint x. □

Define a_{ij} to be 1 if $G(\Gamma_{i,j})$ is connected and Γ_{ij}^g contains a longest path of Γ_{ij}, $1 \leq i \leq M_1$ (i.e. Γ_{ij} is not Hamiltonian), and 0 otherwise.

Let M_1 be the number of expanding digraphs D_i among $D_1, ..., D_M$ for which $G(D_i)$ is connected and Γ_i is not Hamiltonian. We aim to show that $M_1/M \to 0$ as n tends to infinity. W.l.o.g. suppose $\mathcal{N}(k,n) = \{D_1, ..., D_{M_1}\}$ are the connected expanding digraphs which are not Hamiltonian.

Lemma 2.7. *For $1 \leq i \leq M_1$, we have $\sum_{j=1}^{\kappa} a_{ij} \geq (k-1)^n$.*

Proof. Fix $1 \leq i \leq M_1$ and a longest path P_i of Γ_i. Uniformly picking one of $D_{i1}, ..., D_{i\kappa}$, we have

$$\begin{aligned}
\Pr\{a_{ij} = 1\} &\geq \Pr\left\{E(P_i) \subseteq E(\Gamma_{ij}^g)\right\} \\
&\geq \left(1 - \frac{1}{k}\right)^{|E(P_i)|} \\
&\geq \left(1 - \frac{1}{k}\right)^{n}
\end{aligned}$$

The lemma follows from the fact that there are k^n colorings of D_i. □

Let $\Delta \in \mathcal{D}(k-1,n)$ be expanding and non-Hamiltonian and for the purposes of exposition consider its edges to be colored green. Let $D \in \mathcal{D}(k,n)$ be the random digraph obtained by letting each vertex of Δ randomly choose another edge, which will be colored blue. Let $\overline{B_\Delta}$ be the event (in the probability space of randomly chosen blue edges to be added to Δ):

> D has an edge between the endpoints of a longest path of $G(\Delta)$, or
>
> D has an edge from an endpoint of a longest path of Δ to the complement of the path.

Note that the occurrence of $\overline{B_\Delta}$ implies that the corresponding $a_{ij} = 0$. If $a_{ij} = 1$ then the connectivity of Γ_{ij} imlies that $G(D)$ has a longer path than $G(\Delta)$. Let B_Δ be the complement of $\overline{B_\Delta}$ and for Hamiltionian Δ let $B_\Delta = \emptyset$.

Let N_Δ be the number of i, j such that $\Gamma^g_{ij} = \Delta$. We have

$$\sum_{i,j:\Gamma^g_{ij}=\Delta} a_{ij} = N_\Delta \Pr\{B_\Delta\}$$

The number of non-Hamiltonian graphs is bounded by

$$\begin{aligned} M_1 &\leq \sum_{i=1}^{M}\sum_{j=1}^{\kappa} \frac{a_{ij}}{(k-1)^n} \\ &\leq \frac{\sum_\Delta N_\Delta \Pr\{B_\Delta\}}{(k-1)^n} \\ &\leq \frac{Mk^n \max_\Delta \Pr\{B_\Delta\}}{(k-1)^n} \\ &= M\frac{\max_\Delta \Pr\{B_\Delta\}}{(1-1/k)^n} \end{aligned}$$

Fix a $\Delta \in \mathcal{N}(k-1, n)$ and a longest path P_Δ of $G(\Delta)$. Let EP be the set of vertices which are endpoints of a longest path of $G(\Delta)$ that is obtainable from P_Δ by rotations. For $x \in EP$, say x is of Type I if x has at least $\varepsilon n/2$ neighbors outside P_Δ, and Type II otherwise. Let E_1 be the set of Type I endpoints, and E_2 the set of Type II endpoints.

Partition the set of expanding green graphs by

$$\mathcal{D}'(k-1, n) = \mathcal{H}(k-1, n) \cup \mathcal{N}_1(k-1, n) \cup \mathcal{N}_2(k-1, n)$$

where $\mathcal{H}(k-1, n)$ is the set of Hamiltonian graphs, $\mathcal{N}_1(k-1, n)$ the set of non-Hamiltonian graphs with $|E_1| \geq \alpha n/2$ and $\mathcal{N}_2(k-1, n)$ the set of non-Hamiltonian graphs with $|E_1| < \alpha n/2$. Here $\alpha > 0$ is provided by Lemma 2.6.

Lemma 2.8. *For $\Delta \in \mathcal{N}_1(k-1, n)$, $Pr\{B_\Delta\} \leq e^{-\varepsilon\alpha n/4}$.*

Proof. Let each $x \in E_1$ choose a neighbor $y(x)$. The event B_Δ is included in the event $\{\forall x \in E_1 : y(x) \in P_\Delta\}$. We have

$$\begin{aligned} \Pr\{B_\Delta\} &\leq \Pr\{\forall x \in E_1 : y(x) \in P_\Delta\} \\ &= \prod_{x\in E_1} \frac{d_{P_\Delta}(x)}{d_G(x)} \\ &\leq \left(1 - \frac{\varepsilon}{2}\right)^{\alpha n/2} \end{aligned}$$

where $d_{P_\Delta}(x)$ denotes the number of neighbors of x inside P_Δ. □

Lemma 2.9. *For $\Delta \in \mathcal{N}_2(k-1,n)$, $Pr\{B_\Delta\} \le e^{-\varepsilon\alpha^2 n/129}$.*

Proof. Let $X \subseteq E_2$ be a set of $\alpha n/4$ Type II endpoints. X exists because $|EP| \ge \alpha n$ and at most $\alpha n/2$ vertices in EP are of type I. For each $x \in X$, let P_x be a path obtained from P_Δ by rotations that has x as an endpoint. Let $A(x)$ be the set of Type II vertices $y \notin X$ such that a path from x to y in Δ can be obtained from P_x by a sequence of rotations with x fixed. By Lemma 2.6 we have $|A(x)| \ge \alpha n/4$ for each x, since $A(x) = EP(x) \setminus (E_1 \cup X)$.

Let $P_{x,y}$ be a path with endpoints $x \in X, y \in A(x)$ obtained from P_x by rotations with x fixed, and label the vertices on $P_{x,y}$ by $x = z_0, z_1, ..., z_l = y$. Suppose y chooses some z_i on the path with its blue edge. If $\{z_{i+1}, x\} \in E(G)$, let $B_y(x) = \{z_{i+1}\}$. Write $v(y)$ for z_{i+1}. If $\{z_{i+1}, x\} \notin E(G)$, or if y chooses a vertex outside P, let $B_y(x) = \emptyset$.

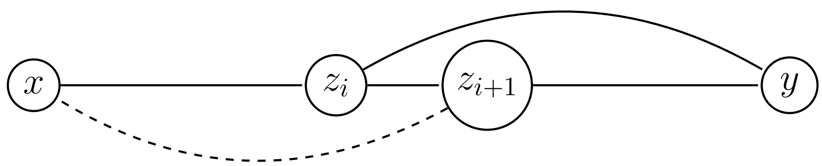

Figure 2.1: Suppose y chooses z_i. The vertex z_{i+1} is included in $B(x)$ if and only if $\{x, z_{i+1}\} \in E(G)$.

There will be at least $2\left(\frac{1}{2}+\frac{\varepsilon}{2}\right)n - n = \varepsilon n$ choices for i for which $\{x, z_{i+1}\} \in E(G)$. Let Y_x be the number of $y \in A(x)$ such that $B_y(x)$ is nonempty. This variable is bounded stochastically from below by a binomial $\text{Bin}(\alpha n/4, \varepsilon)$ variable, and by a Chernoff bound we have that

$$\Pr\left\{\exists x : Y_x \le \frac{\varepsilon\alpha n}{8}\right\} \le n\exp\left\{-\frac{\varepsilon\alpha n}{32}\right\}$$

Define $B(x) = \bigcup_{y\in A(x)} B_y(x)$. If x chooses a vertex in $B(x)$ then $\overline{B_\Delta}$ occurs. Conditional on $Y_x \ge \varepsilon\alpha n/8$ for all $x \in X$, let $y_1, y_2, ..., y_r$ be $r = \varepsilon\alpha n/8$ vertices whose choice produces a nonempty $B_y(x)$. Let $Z_x = |B(x)|$, and for $i = 1, ..., r$ define Z_i to be 1 if $v(y_i)$ is distinct from $v(y_1), ..., v(y_{i-1})$ and 0 otherwise. We have $Z_x = \sum_{i=1}^r Z_i$, and each Z_i is bounded from below by a Bernoulli variable with parameter $1 - \alpha/8$. To see this, note that y_i has at least εn choices resulting in a nonempty $B_{y_i}(x)$ since x and y_i are of Type II, so

$$\Pr\{\exists j < i : v(y_j) = v(y_i)\} \le \frac{i-1}{\varepsilon n} \le \frac{\varepsilon\alpha n/8}{\varepsilon n} = \frac{\alpha}{8}$$

Since $\alpha/8 < 1/2$, Z_x is bounded stochastically from below by a binomial $\text{Bin}(\varepsilon\alpha n/8, 1/2)$ variable, and so

$$\Pr\left\{\exists x : Z_x < \frac{\varepsilon\alpha n}{32}\right\} \le n\exp\left\{-\frac{\varepsilon\alpha n}{128}\right\}$$

Each x for which $Z_x \ge \varepsilon\alpha n/32$ will choose a vertex in $B(x)$ with probability

$$\frac{|B(x)|}{d_G(x)} \ge \frac{\varepsilon\alpha n/32}{n} = \frac{\varepsilon\alpha}{32}$$

Hence we have

$$\Pr\{B_\Delta\} \le \left(1 - \frac{\varepsilon\alpha}{32}\right)^{\alpha n/4} + n\exp\left\{-\frac{\varepsilon\alpha n}{32}\right\} + n\exp\left\{-\frac{\varepsilon\alpha n}{128}\right\} \le e^{-\varepsilon\alpha^2 n/129}.$$

□

We can now complete the proof of Theorem 2.2. From Lemmas 2.8 and 2.9 we have

$$\Pr\{B_\Delta\} \le \max\left\{e^{-\varepsilon\alpha n/4}, e^{-\varepsilon\alpha^2 n/129}\right\}.$$

Going back to (2.4) with $k = C/\varepsilon$ we have

$$\begin{aligned}
\Pr\{G_k \text{ is non-Hamiltonian}\} &= o(1) + \frac{M_1}{M} \\
&\le o(1) + \frac{\max_\Delta \Pr\{B_\Delta\}}{(1-1/k)^n} \\
&= o(1) + \left[\frac{e^{-\varepsilon\alpha^2/129}}{1-\varepsilon/C}\right]^n \\
&\le o(1) + \exp\left\{-\varepsilon\left(\frac{\alpha^2}{129} - \frac{2}{C}\right)n\right\} \\
&= o(1),
\end{aligned}$$

for $C = 259/\alpha^2$. □

We can now prove Corollary 2.1. We follow an argument based on Walkup [88]. If X_e is the *length* of edge $e = uv$ of G then we can write $X_e = \min\{Y_{uv}, Y_{vu}\}$ where Y_{uv}, Y_{vu} are independent copies of the random variable Y where $\Pr\{Y \ge y\} = (1-y)^{1/2}$. The density of Y is close to $y/2$ for y close to zero. Now consider G_{k_ε} where the choices $\{v_1, v_2, \ldots, v_{k_\varepsilon}\}$ of vertex u are the k_ε edges of lowest weight Y_{uv} among $uv \in E(G)$. Now consider the total weight of the Hamilton cycle H posited by Theorem 2.2. The expected weight of an edge of H is at most $2 \times \frac{k_\varepsilon+1}{2(\frac{1}{2}+\varepsilon)n}$ and the corollary follows.

2.4 Long Paths: Proof of Theorem 2.3

Let D_k denote the directed graph with out-degree k defined by the vertex choices. Consider a Depth First Search (DFS) of D_k where we construct D_k as we go. At all times we keep a stack U of vertices which have been visited, but for which we have chosen fewer than k out-edges. T denotes the set of vertices that have not been visited by DFS. Each step of the algorithm begins with the top vertex u of U choosing one new out-edge. If the other end of the edge v lies in T (we call this a *hit*), we move v from T to the top of U.

When DFS returns to $v \in U$ and at this time v has chosen all of its k out-edges, we move v from U to S. In this way we partition V into

S - Vertices that have chosen all k of its out-edges.

U - Vertices that have been visited but have chosen fewer than k edges.

T - Unvisited vertices.

Key facts: Let h denote the number of hits at any time and let κ denote the number of times we have re-started the search i.e. selected a vertex in T after the stack S empties.

P1 $|S \cup U|$ increases by 1 for each hit, so $|S \cup U| \ge h$.

P2 More specifically, $|S \cup U| = h + \kappa - 1$.

P3 At all times $S \cup U$ contains a path which contains all of U.

The goal will be to prove that $|U| \geq (1 - 2\varepsilon)m$ at some point of the search, where ε is some arbitrarily small positive constant.

Lemma 2.10. *After εkm steps, i.e. after εkm edges have been chosen in total, the number of hits $h \geq (1 - \varepsilon)m$ w.h.p.*

Proof. Since $\delta(G_k) \geq k$, each tree component of G_k has at least k vertices, and at least k^2 edges must be chosen in order to complete the search of the component. Hence, after εkm edges have been chosen, at most $\varepsilon km/k^2 \leq \varepsilon m/2$ tree components have been found. This means that if $h \leq (1-\varepsilon)m$ after εkm edges have been sent out, then **P2** implies that $|S \cup U| \leq (1 - \varepsilon/2)m$.

So if $h \leq (1 - \varepsilon)m$ each edge chosen by the top vertex u has probability at least $\frac{d(u)-|S\cup U|}{d(u)} \geq \varepsilon/2$ of making a hit. Hence,

$$\Pr\{h \leq (1 - \varepsilon)m \text{ after } \varepsilon km \text{ steps}\} \leq \Pr\{\text{Bin}(\varepsilon km, \varepsilon/2) \leq (1 - \varepsilon)m\} = o(1),$$

for $k \geq 2/\varepsilon^2$, by the Chernoff bounds. □

We can now complete the proof of Theorem 2.3. By Lemma 2.10, after εkm edges have been chosen we have $|S \cup U| \geq (1 - \varepsilon)m$ w.h.p. For a vertex to be included in S, it must have chosen all of its edges. Hence, $|S| \leq \varepsilon km/k = \varepsilon m$, and we have $|U| \geq (1 - 2\varepsilon)m$. Finally observe that U is the set of vertices of a path of G_k. □

2.5 Long Cycles: Proof of Theorem 2.4

Suppose now that we consider $G_{4k} = LR_k \cup DR_k \cup LB_k \cup DB_k$ where each for each vertex v and for each $c \in \{$"light red", "dark red", "light blue","dark blue"$\}$ the vertex v makes k choices of neighbor $N_c(v)$, distinct from any previous choices for this vertex. The edges $\{v, w\}, w \in N_c(v)$ are given the color c. Let LR_k, DR_k, LB_k, DB_k respectively be the graphs induced by the differently colored edges. We have by Theorem 2.3 that w.h.p. there is a path P of length at least $(1 - \varepsilon)m$ in the light red graph LR_k. At this point we start using a modification of DFS (denoted by $\Delta\Phi\Sigma$) and the differently colored choices to create a cycle.

We divide the steps into epochs $T_0, T_{00}, T_{01}, \ldots$, indexed by binary strings. We stop the search immediately if there is a high chance of finding a cycle of length at least $(1 - 20\varepsilon)m$. If executed, epoch $T_{\boldsymbol{\iota}}, \boldsymbol{\iota} = 0{*}{*}{*}$ will extend the exploration tree by at least $(1-5\varepsilon)m$ vertices, unless an unlikely failure occurs. Theorem 2.3 provides T_0. In the remainder, we will assume $\boldsymbol{\iota} \neq 0$.

Epoch $T_{\boldsymbol{\iota}}$ will use light red colors if $\boldsymbol{\iota}$ has odd length and ends in a 0, dark red if $\boldsymbol{\iota}$ has even length and ends in a 0, light blue if $\boldsymbol{\iota}$ has odd length and ends in a 1, and dark blue if $\boldsymbol{\iota}$ has even length and ends in a 1. Epochs $T_{\boldsymbol{\iota}0}$ and $T_{\boldsymbol{\iota}1}$ (where $\boldsymbol{\iota}j$ denotes the string obtained by appending j to the end of $\boldsymbol{\iota}$) both start where $T_{\boldsymbol{\iota}}$ ends, and this coloring ensures that every vertex discovered in an epoch will initially have no adjacent edges in the color of the epoch.

During epoch $T_{\boldsymbol{\iota}}$ we maintain a stack of vertices $S_{\boldsymbol{\iota}}$. When discovered, a vertex is placed in one of the three sets $A_{\boldsymbol{\iota}}, B_{\boldsymbol{\iota}}, C_{\boldsymbol{\iota}}$, and simultaneously placed in $S_{\boldsymbol{\iota}}$ if it is placed in $A_{\boldsymbol{\iota}}$. Once placed, the vertex remains in its designated set even if it is removed from S_ι. Let $d_T(v,w)$ be the length of the unique path in the exploration tree T from v to w. We designate the set for v as follows.

A_ι - v has less than $(1-2\varepsilon)d(v)$ G-neighbors in T.

B_ι - v has at least $(1-2\varepsilon)d(v)$ G-neighbors in T, but less than $\varepsilon d(v)$ G-neighbors w such that $d_T(v,w) \geq (1-19\varepsilon)m$.

C_ι - v has at least $(1-2\varepsilon)d(v)$ G-neighbors in T, and at least $\varepsilon d(v)$ G-neighbors w such that $d_T(v,w) \geq (1-19\varepsilon)m$.

At the initiation of epoch T_ι, a previous epoch will provide a set T_ι^0 of $3\varepsilon m$ vertices, as described below. Starting with $A_\iota = B_\iota = C_\iota = \emptyset$, each vertex of T_ι^0 is placed in A_ι, B_ι or C_ι according to the rules above. Let $S_\iota = A_\iota$, ordered with the latest discovered vertex on top.

If at any point during T_ι we have $|B_\iota| = \varepsilon m$ or $|C_\iota| = \varepsilon m$, we immediately interrupt $\Delta\Phi\Sigma$ and use the vertices of B_ι or C_ι to find a cycle, as described below.

An epoch T_ι consists of up to εkm steps, and each step begins with a $v \in A_\iota$ at the top of the stack S_ι. This vertex is called *active*. If v has chosen k neighbors, remove v from the stack and perform the next step. Otherwise, let v randomly pick one neighbor w from $N_G(v)$. If $w \notin T$, then w is assigned to A_ι, B_ι or C_ι as described above. If $w \in A_\iota$, perform the next step with w at the top of S_ι. If $w \in B_\iota \cup C_\iota$ perform the next step with the same v. If $w \in T$, perform the next step without placing w in S_ι.

The exploration tree T is built by adding to it any vertex found during $\Delta\Phi\Sigma$, along with the edge used to discover the vertex.

Note that unless $|B_\iota| = \varepsilon m$ or $|C_\iota| = \varepsilon m$, we initially have $|A_\iota| \geq \varepsilon m$, guaranteeing that εkm steps may be executed. Epoch T_ι *succeeds* and is ended (possibly after fewer than εkm steps) if at some point we have $|A_\iota| = (1-2\varepsilon)m$. If all εkm steps are executed and $|A_\iota| < (1-2\varepsilon)m$, the epoch fails.

Lemma 2.11. *Epoch T_ι succeeds with probability at least $1-e^{-\varepsilon^2 m/8}$, unless $|B_\iota| = \varepsilon m$ or $|C_\iota| = \varepsilon m$ is reached.*

Proof. An epoch fails if less than $(1-3\varepsilon)m$ steps result in the active vertex choosing a neighbor outside T. Since the active vertex is always in A_ι, we have

$$\Pr\{T_\iota \text{ finishes with } |A_\iota| < (1-2\varepsilon)m\} \leq \Pr\{\text{Bin}(\varepsilon km, 2\varepsilon) < (1-3\varepsilon)m\} \leq e^{-\varepsilon^2 m/8}$$

for $k \geq 1/2\varepsilon^2$, by Hoeffding's inequality. This proves the lemma. □

Ignoring the colors of the edges, an epoch produces a tree which is a subtree of T. Let P_ι be the longest path of vertices in A_ι, and let R_ι be the set of vertices discovered during T_ι which are not in P_ι. If the epoch succeeds, P_ι has length at least $(1-6\varepsilon)m$, and at most $3\varepsilon m$ vertices discovered during T_ι are not on the path. Indeed, a vertex of A_ι is outside P_ι if and only if it has chosen all its k neighbors. Thus, the number of vertices not on the path is bounded by

$$|R_\iota| \leq \frac{\varepsilon km}{k} + |B_\iota| + |C_\iota| < 3\varepsilon m.$$

If the epoch fails, the path P_ι may be shorter, but $|R_\iota|$ is still bounded by $3\varepsilon m$.

If $T_{\boldsymbol{\iota}}$ succeeds, the epochs $T_{\boldsymbol{\iota}0}$ and $T_{\boldsymbol{\iota}1}$ will be initiated at the end of $T_{\boldsymbol{\iota}}$, by letting $T^0_{\boldsymbol{\iota}0}$ and $T^0_{\boldsymbol{\iota}1}$ be the last $3\varepsilon m$ vertices discovered during $T_{\boldsymbol{\iota}}$. If $T_{\boldsymbol{\iota}}$ fails, $T_{\boldsymbol{\iota}0}$ and $T_{\boldsymbol{\iota}1}$ will not be initiated. The exploration tree T will resemble an unbalanced binary tree, in which each successful epoch gives rise to up to two new epochs. Epochs are ordered and $T_{\boldsymbol{\iota}_1}$ is initiated before $T_{\boldsymbol{\iota}_2}$ if and only if $\boldsymbol{\iota}_1 < \boldsymbol{\iota}_2$. Here let $\boldsymbol{\iota}_i = \boldsymbol{x}\boldsymbol{y}_i, i = 1, 2$ where $\boldsymbol{x}$ is the longest common substring of $\boldsymbol{\iota}_1, \boldsymbol{\iota}_2$. We will have $\boldsymbol{\iota}_1 < \boldsymbol{\iota}_2$ if either $\boldsymbol{y_1}$ is the empty string or if $\boldsymbol{y_1}$ starts with 0 and $\boldsymbol{y_2}$ starts with 1.

Lemma 2.12. *W.h.p., $\Delta\Phi\Sigma$ will discover an epoch $T_{\boldsymbol{\iota}}$ having $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$.*

Proof. Suppose that no epoch ends with $|B_{\boldsymbol{\iota}}| = \varepsilon m$ or $|C_{\boldsymbol{\iota}}| = \varepsilon m$. Under this assumption, each successful epoch $T_{\boldsymbol{\iota}}$ gives rise to $X'_{\boldsymbol{\iota}}$ new epochs. By Lemma 2.11, $X'_{\boldsymbol{\iota}}$ can be stochastically bounded from below by $X_{\boldsymbol{\iota}}$, where for some $c > 0$, $X_{\boldsymbol{\iota}} = 0$ with probability e^{-2cm}, $X_{\boldsymbol{\iota}} = 1$ with probability $2e^{-cm}(1 - e^{-cm})$ and $X_{\boldsymbol{\iota}} = 2$ with probability $(1 - e^{-cm})^2$. The number of successful epochs is then bounded from below by the total number of offspring in a Galton-Watson branching process with offspring distribution described by $X_{\boldsymbol{\iota}}$. The offspring distribution for this lower bound has generating function

$$G_m(s) = e^{-2cm} + 2se^{-cm}(1 - e^{-cm}) + s^2(1 - e^{-cm})^2.$$

Let s_m be the smallest fixed point $G_m(s_m) = s_m$. We have, with $\xi = e^{-cm}$,

$$s_m = \frac{1 - 2\xi(1-\xi) - [(1 - 2\xi(1-\xi))^2 - 4(1-\xi)^2\xi^2]^{1/2}}{2(1-\xi)^2} \to 0, \text{ as } m \to \infty.$$

Hence, the probability that the branching process never expires is at least $1 - s_m$, which tends to 1.

The number of epochs is bounded by a finite number. Hence, the branching process cannot be infinite. This contradiction finishes the proof. ☐

We may now finish the proof of the theorem. Condition first on $\Delta\Phi\Sigma$ being stopped by an epoch $T_{\boldsymbol{\iota}}$ having $|C_{\boldsymbol{\iota}}| = \varepsilon m$. In this case, let each $v \in C_{\boldsymbol{\iota}}$ choose k neighbors using edges with the epoch's color. Each choice has probability at least ε of finding a cycle of length at least $(1 - 19\varepsilon)m$, by choosing a neighbor w such that $d_T(v, w) \geq (1 - 19\varepsilon)m$. The probability of not finding a cycle of length at least $(1 - 19\varepsilon)m$ is bounded by

$$(1 - \varepsilon)^{\varepsilon km} \to 0.$$

Now condition on $\Delta\Phi\Sigma$ being stopped by an epoch $T_{\boldsymbol{\iota}}$ having $|B_{\boldsymbol{\iota}}| = \varepsilon m$. Note that we must have $\boldsymbol{\iota} = \boldsymbol{\iota}'1$ for some $\boldsymbol{\iota}'$. Indeed, if $\boldsymbol{\iota} = \boldsymbol{\iota}'0$, then any v discovered in $\boldsymbol{\iota}$ must have at least $11\varepsilon d(v)$ G-neighbors at distance at least $(1 - 19\varepsilon)m$, at its time of discovery. If not, and $v \notin A_{\boldsymbol{\iota}}$ then it has at most $2\varepsilon d(v)$ G-neighbors outside T, at most $3\varepsilon d(v) + 3\varepsilon d(v)$ G-neighbors in $R_{\boldsymbol{\iota}} \cup R_{\boldsymbol{\iota}'}$. There are at most $(1 - 19\varepsilon)d(v)$ G-neighbors in $T \setminus (R_{\boldsymbol{\iota}} \cup R_{\boldsymbol{\iota}'})$ at distance less than $(1 - 19\varepsilon)d(v)$ and so there are at least $11\varepsilon d(v)$ G-neighbors in T at distance at least $(1 - 19\varepsilon)d(v)$ from v, which implies that $v \in C_{\boldsymbol{\iota}}$, contradiction.

Since the epoch produces a tree with at most m vertices, using the pigeonhole principle we can choose a $W \subseteq B_{\boldsymbol{\iota}}$ such that $|W| = \varepsilon^2 m$ and $d_T(v, w) \leq \varepsilon m$ for any $v, w \in W$.

Note also that $d(v) \leq 2m$ for any $v \in B_{\boldsymbol{\iota}}$. This can be seen as follows: For any $v \in W$ let $\rho_v \in T^0_{\boldsymbol{\iota}}$ be the vertex which minimizes $d_T(v, \rho_v)$. Note that we may have $\rho_v = v$. There are at most $|Q|$

G-neighbors of v on the path Q from v to ρ_v. Then note that there are at most $2((1-19\varepsilon)m-|Q|)$ G-neighbors of v on $T \setminus (Q \cup R_\iota \cup R_{\iota'} \cup R_{\iota'0})$ that are within $(1-19\varepsilon)m$ of v. Here the factor 2 comes from counting G-neighbors in T_ι and $T_{\iota'0}$. So the maximum number of $w \in N_G(v) \cap T$ such that $d_T(v,w) \le (1-19\varepsilon)m$ is bounded by

$$|Q| + 2((1-19\varepsilon)m - |Q|) + |R_\iota| + |R_{\iota'}| + |R_{\iota'0}| \le (2-29\varepsilon)m \tag{2.5}$$

Equation (2.5) then implies that $d(v) \le (2-29\varepsilon)m + 3\varepsilon d(v)$.

Define an ordering on T by saying that $t_1 \le t_2$ if t_1 was discovered before t_2 during $\Delta\Phi\Sigma$, or if $t_1 = t_2$. If $S \subseteq T'$, and $t \le s$ for all $s \in S$, write $t \le S$. Similarly define $\ge, >$ and $<$.

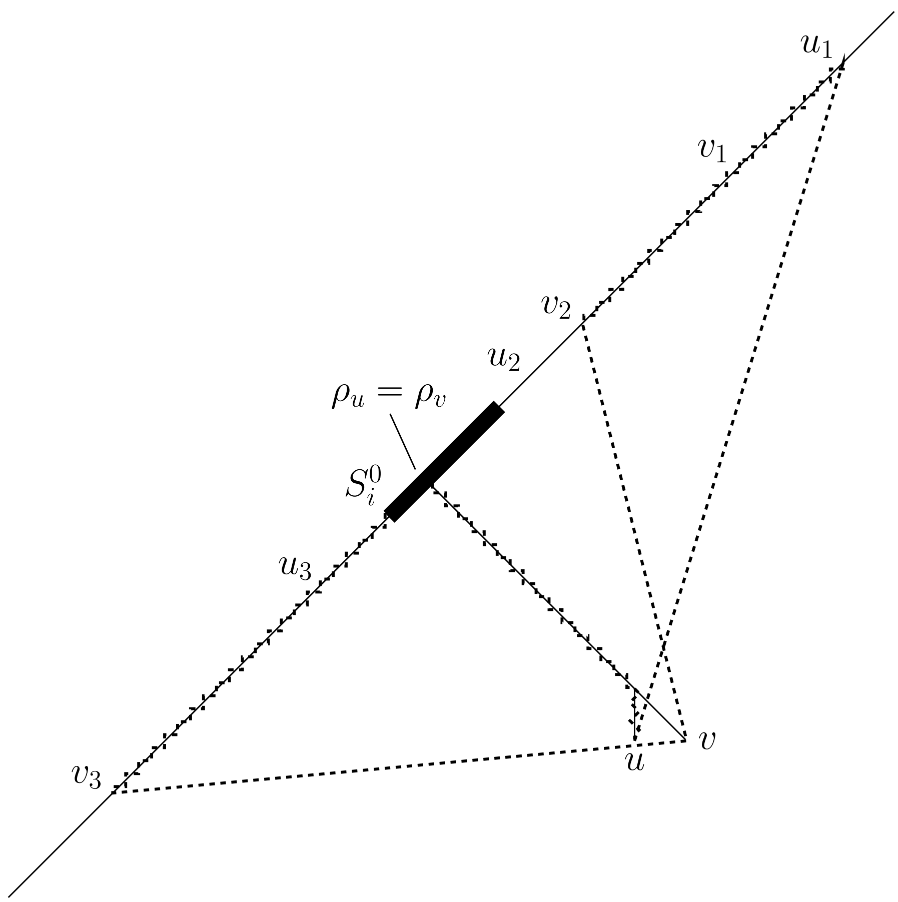

Figure 2.2: Example depiction of cycle found when $|B_\iota| = \varepsilon m$.

Let each $v \in W$ choose k neighbors in the color of epoch T_ι. We say that v is *good* if it chooses $v_1, v_2 \in P_{\iota'}$ and $v_3 \in P_{\iota'0}$ such that

$$d_T(v_1,v_2) + d_T(v_3, T_\iota^0) + d_T(\rho_v, v) \ge (1-17\varepsilon)m$$

where $d_T(v_3,S) = \min_{s\in S} d_T(v_3,s)$. For each $v \in W$ define $n_0(v) = |N_G(v) \cap P_\iota \setminus T_\iota^0|$, $n_1(v) = |N_G(v) \cap P_{\iota'} \setminus T_\iota^0|$ and $n_2(v) = |N_G(v) \cap P_{\iota'0} \setminus T_\iota^0|$. Since $v \in B_\iota$ we have

$$n_0(v) + n_1(v) + n_2(v) = |(N_G(v) \cap T) \setminus (R_{\iota'} \cup R_{\iota'0} \cup R_\iota \cup T_\iota^0)| \ge (1-14\varepsilon)m.$$

Since the $n_0(v)+n_1(v)$ vertices of $N_G(v)\cup P_{\boldsymbol{\iota}}\cup P_{\boldsymbol{\iota}'}\setminus T^0_{\boldsymbol{\iota}}$ are on a path, we must have $n_0(v)+n_1(v)\le(1-16\varepsilon)m$, otherwise v has $2\varepsilon m\ge\varepsilon d(v)$ neighbors at distance at least $(1-18\varepsilon)m$, contradicting $v\in B_{\boldsymbol{\iota}}$. This implies $n_2(v)\ge 2\varepsilon m$. Similarly, $n_1(v)\ge 2\varepsilon m$.

Fix a vertex $v\in W$ and define $V_1,V_2\subseteq(N_G(v)\cap P_{\boldsymbol{\iota}'})\setminus T^0_{\boldsymbol{\iota}}$ and $V_3\subseteq(N_G(v)\cap P_{\boldsymbol{\iota}'0})\setminus T^0_{\boldsymbol{\iota}}$, $|V_1|=|V_2|=|V_3|=\varepsilon m$ as follows. V_1 is the set of the first εm vertices of $N_G(v)\cap P_{\boldsymbol{\iota}'}$ discovered during $\Delta\Phi\Sigma$. V_2 is the set of the last εm vertices of $N_G(v)\cap P_{\boldsymbol{\iota}'}$ discovered before any vertex of $T^0_{\boldsymbol{\iota}}$. Lastly, V_3 consists of the εm last vertices discovered in $N_G(v)\cap P_{\boldsymbol{\iota}'0}$. Since $n_1(v)\ge 2\varepsilon m$ and $n_2(v)\ge 2\varepsilon m$, the sets V_1,V_2,V_3 exist and are disjoint.

Since $d(v)\le 2m$, the probability that v chooses $v_1\in V_1, v_2\in V_2$ and $v_3\in V_3$ is at least $(\varepsilon/2)^3$. If this happens, we have

$$d_T(v_1,v_2)+d_T(v_3,T^0_{\boldsymbol{\iota}})+d_T(\rho_v,v)\ge n_1(v)-2\varepsilon m+n_2(v)-\varepsilon m+n_3(v)\ge(1-17\varepsilon)m.$$

In other words, $v\in W$ is good with probability at least $(\varepsilon/2)^3$. Since $|W|=\varepsilon^2 m$, w.h.p. there exist two good vertices $u,v\in W$. By choice of W we have $d_T(\rho_u,u)\ge d_T(\rho_v,v)-2\varepsilon m$. Suppose u and v pick $u_1\le u_2\le u_3$ and $v_1\le v_2\le v_3$. We have $d_T(v_1,v_2)\le d_T(u_1,v_2)+|V_1|$. The cycle $(u,u_1,...,v_2,v,v_3,...,\rho_u,...,u)$ has length

$$\begin{aligned}
& 1+d_T(u_1,v_2)+1+1+d_T(v_3,\rho_u)+d_T(\rho_u,u)\\
\ge\ & d_T(v_1,v_2)+d_T(v_3,T^0_{\boldsymbol{\iota}})+d_T(\rho_v,v)-3\varepsilon m\\
\ge\ & (1-20\varepsilon)m.
\end{aligned}$$

Chapter 3

Random walk cuckoo hashing

This chapter corresponds to [41].

Abstract

Cuckoo Hashing is a hashing scheme invented by Pagh and Rodler [77]. It uses $d \geq 2$ distinct hash functions to insert items into the hash table. It has been an open question for some time as to the expected time for Random Walk Insertion to add items. We show that if the number of hash functions $d = O(1)$ is sufficiently large, then the expected insertion time is $O(1)$ per item.

3.1 Introduction

Our motivation for this paper comes from Cuckoo Hashing (Pagh and Rodler [77]). Briefly each one of n items $x \in L$ has d possible locations $h_1(x), h_2(x), \dots, h_d(x) \in R$, where d is typically a small constant and the h_i are hash functions, typically assumed to behave as independent fully random hash functions. (See [71] for some justification of this assumption.)

We assume each location can hold only one item. Items are inserted consecutively and when an item x is inserted into the table, it can be placed immediately if one of its d locations is currently empty. If not, one of the items in its d locations must be displaced and moved to another of its d choices to make room for x. This item in turn may need to displace another item out of one of its d locations. Inserting an item may require a sequence of moves, each maintaining the invariant that each item remains in one of its d potential locations, until no further evictions are needed.

We now give the formal description of the mathematical model that we use. We are given two disjoint sets $L = \{v_1, v_2, \dots, v_n\}, R = \{w_1, w_2, \dots, w_m\}$. Each $v \in L$ independently chooses a set $N(v)$ of $d \geq 2$ uniformly random neighbors in R. We assume for simplicity that this selection is done with replacement. This provides us with the bipartite cuckoo graph Γ. Cuckoo Hashing can be thought of as a simple algorithm for finding a matching M of L into R in Γ. In the context of hashing, if $\{x, y\}$ is an edge of M then $y \in R$ is a hash value of $x \in L$.

Cuckoo Hashing constructs M by defining a sequence of matchings $M_1, M_2, \dots, M_n$, where M_k is a matching of $L_k = \{v_1, v_2, \dots, v_k\}$ into R. We let Γ_k denote the subgraph of Γ induced by $L_k \cup R$.

We let R_k denote the vertices of R that are covered by M_k and define the function $\phi_k : L_k \to R_k$ by asserting that $M_k = \{\{v, \phi_k(v)\} : v \in L_k\}$. We obtain M_k from M_{k-1} by finding an augmenting path P_k in Γ_k from v_k to a vertex in $\bar{R}_{k-1} = R \setminus R_{k-1}$.

This augmenting path P_k is obtained by a random walk. To begin we obtain M_1 by letting $\phi_1(v_1)$ be a uniformly random member of $N(v_1)$, the neighbors of v_1. Having defined M_k we proceed as follows: Steps 1 – 4 constitute round k.

Algorithm INSERT:

Step 1 $x \leftarrow v_k$; $M \leftarrow M_{k-1}$;

Step 2 **If** $S_k(x) = N(x) \cap \bar{R}_{k-1} \neq \emptyset$ **then** choose y uniformly at random from $S_k(x)$ and let $M_k = M \cup \{\{x, y\}\}$, **else**

Step 3 Choose y uniformly at random from $N(x)$;

Step 4 $M \leftarrow M \cup \{\{x, y\}\} \setminus \{y, \phi_{k-1}^{-1}(y)\}$; $x \leftarrow \phi_{k-1}^{-1}(y)$; **goto** Step 2.

This algorithm was first discussed in the conference version of [35]. Our interest here is in the expected time for INSERT to complete a round. Our results depend on d being large. In this case we will improve on the bounds on insertion time given in Frieze, Melsted and Mitzenmacher [50], Fountoulakis, Panagiotou and Steger [37], Fotakis, Pagh, Sanders and Spirakis [35]. The paper [35] studied the efficiency of insertion via Breadth First Search and also carried out some experiments with the random walk approach. The papers [50] and [37] considered insertion by random walk and proved that the expected time to complete a round can be bounded by $\log^{2+o_d(1)} n$, where $o_d(1)$ tends to zero as $d \to \infty$. The paper [37] improved on the space requirements in [50]. They showed that given ε, their bounds hold for any d large enough to give the existence of a matching w.h.p. Mitzenmacher [70] gives a survey on Cuckoo Hashing and Problem 1 of the survey asks for the expected insertion time.

Frieze and Melsted [49], Fountoulakis and Panagiotou [36] give information on the relative sizes of L, R needed for there to exist a matching of L into R w.h.p.

We will prove the following theorem: it shows that the expected insertion time is $O(1)$, but only for a large value of d. The theorem focusses on the more interesting case where the load factor n/m is close to one. When the load factor is small enough i.e. when $nd \leq (1-\varepsilon)m$ the components of Γ will be bounded in expectation and so it is straightforward to show an $O(1)$ expected insertion time.

Theorem 3.1. *Suppose that $n = (1-\varepsilon)m$ where ε is a fixed positive constant, assumed to be small. Let $0 < \theta < 1$ also be a fixed positive constant and let*

$$\gamma = 5(1-\varepsilon)^{d/2}. \tag{3.1}$$

If $d^2\gamma \leq (1-\theta)(d-1)$ then w.h.p. the structure of Γ is such that over the random choices in Steps 2,3,

$$\boldsymbol{E}[|P_k|] \leq 1 + \frac{2}{\theta} \text{ for } k = 1, 2, \dots, n. \tag{3.2}$$

Here $|P_k|$ is the length (number of edges) of P_k.

When d is large the value of θ in (3.2) is close to $d\varepsilon/2$. It can be seen from the proof that as $\varepsilon \to 0$, the value of d needed is of the order $\left(\frac{\log 1/\varepsilon}{\varepsilon}\right)$. This is larger than the value $O(\log(1/\varepsilon))$ needed for there to be a perfect matching from L to R and finding an $O(1)$ bound on the expected insertion time for small d remains as an open problem. We note that the BFS algorithm of [35], which requires $\Omega(n^\delta)$ space for constant $\delta > 0$, has an expected insertion time of $d^{\Omega(\log 1/\varepsilon)}$.

The problem here bears some relation to the *On-line bipartite matching problem* discussed for example in Chaudhuri, Daskalakis, Kleinberg and Lin [17], Bosek, Leniowski, Sankowski and Zych [15] and Gupta, Kumar and Stein [56]. In these papers the bipartite graph is arbitrary and has a perfect matching and vertices on one side A of the bipartition arrive in some order, along with their choice of neighbors in the other side B. As each new member of A arrives, a current matching is updated via an augmenting path. The aim is to keep the sum of the lengths of the augmenting paths needed to be as small as possible. It is shown, among other things, in [17] that this sum can be bounded by $O(n \log n)$ in expectation and w.h.p. This requires finding a shortest augmenting path each time. Our result differs in that our graph is random and $|A| = (1-\varepsilon)|B|$ and we only require a matching of A into B. On the other hand we obtain a sum of lengths of augmenting paths of order $O(n)$ in expectation via a random choice of path.

3.2 Proof of Theorem 3.1

3.2.1 Outline of the main ideas

Let

$$B_k = \left\{v \in L_k : N(v) \cap \bar{R}_{k-1} = \emptyset\right\}.$$

If $x \notin B_k$ in Step 2 of INSERT then we will have found P_k.

Let $P = (x_1, \xi_1, x_2, \xi_2, \ldots, x_\ell)$ be a path in Γ, where $x_1, x_2, \ldots, x_\ell \in L_{k-1}$ and $\xi_1, \xi_2, \ldots, \xi_{\ell-1} \in R_{k-1}$. We say that P is *interesting* if $x_1, x_2, \ldots, x_\ell \in B_k$. We note that if the path $P_k = (x_1 = v_n, \xi_1, x_2, \xi_2, \ldots, x_\ell, \xi_\ell, x_{\ell+1}, \xi_{\ell+1})$ then $Q_k = (x_1, \xi_1, x_2, \xi_2, \ldots, x_\ell)$ is interesting. Indeed, we must have $x_i \in B_k$, $1 \le i \le \ell$, else INSERT would have chosen $\xi_i \in \bar{R}_{k-1} \subseteq \bar{R}_{x_i}$ and completed the round.

Our strategy is simple. We show that w.h.p. there are relatively few long interesting paths and because our algorithm (usually) chooses a path at random, it is unlikely to be long and interesting. One caveat to this approach is that while all augmenting paths yield interesting sub-paths, the reverse is not the case. In which case, it would be better to estimate the number of possible long augmenting paths. The problem with this approach is that we then need to control the distribution of the matching M_k. This has been the difficulty up to now and we have avoided the problem by focussing on interesting paths. Of course, there is a cost in that d is larger than one would like, but it is at least independent of n.

To bound the number of interesting paths, we bound $|B_k|$ and use this to bound the number of paths.

3.2.2 Detailed proof

Fix $1 \leq k \leq n$. We observe that if $R_{k-1} = \{y_1, y_2, \ldots, y_{k-1}\}$ then

$$y_k \text{ is chosen uniformly from } \bar{R}_{k-1} \tag{3.3}$$

and is independent of the graph Γ_{k-1} induced by $L_{k-1} \cup R_{k-1}$. This is because we can expose Γ along with the algorithm. When we start the construction of M_k we expose the neighbors of v_k one by one. In this way we either determine that $S_k(v_k) = \emptyset$ or we expose a uniformly random member of $S_k(v_k)$ without revealing any more of $N(v_k)$. In general, in Step 2, we have either exposed all the neighbors of x and these will necessarily be in R_{k-1}. Or, we can proceed to expose the unexposed neighbors of x until either (i) we determine that $S_k(x) = \emptyset$ and we choose a uniformly random member of $N(x)$ or (ii) we find a neighbor of x that is a uniformly random member of $\bar{R}_{k-1}$. Thus

$$R_{k-1} \text{ is a uniformly random subset of } R.$$

We need to show that B_k is small. It is clear that $v_1 \notin B_1$ i.e. $B_1 = \emptyset$ and so we deal next with $2 \leq k \leq d-2$. If $v_k \in B_k$ then v_k must choose some vertex in R three times. But,

$$\mathbf{Pr}(\exists 2 \leq k \leq d-2, w \in R : v_k \text{ chooses } w \text{ three times}) \leq (d-2)m \times m^{-3} = o(1).$$

This implies that w.h.p. $B_k = \emptyset$ for $2 \leq k \leq d-2$. We deal next with $d-1 \leq k \leq n^{9/10}$. Since $N(v_k)$ is uniformly random, we see that

$$\mathbf{Pr}(\exists k \leq n^{9/10} : \ v_k \in B_k) \leq n^{1-d/10} = o(1) \text{ for } d > 10.$$

Assume from now on that $n^{9/10} \leq k \leq n-1$. Let $\nu_{k,\ell}$ denote the number of interesting paths with $2\ell - 1$ vertices. Let $\theta, \gamma > 0$ be as in the statement of Theorem 3.1.

Lemma 3.1. *Given A_0 and d sufficiently large,*

$$\mathbf{Pr}\left(\exists 2 \leq \ell \leq A_0 \log\log n : \nu_{k,\ell} \geq (1+\theta)k\gamma(d^2\gamma)^{\ell-1}\right) = o(n^{-2}). \tag{3.4}$$

The bound $o(n^{-2})$ is sufficient to deal with the insertion of n items.

Before proving the lemma, we show how it can be used to prove Theorem 3.1. We will need the following claims:

Claim 3.1. *Let Δ denote the maximum degree in Γ. Then for any $t \geq \log n$ we have $\mathbf{Pr}(\Delta \geq t) \leq e^{-t}$.*

Proof of Claim: If $v \in L$ then its degree $\deg(v) = d$. Now consider $w \in R$. Then for $t \geq \log n$,

$$\mathbf{Pr}(\exists w \in R : \deg(w) \geq t) \leq m\binom{dn}{t}\frac{1}{m^t} \leq m\left(\frac{de}{t}\right)^t \leq e^{-t}.$$

End of proof of Claim

We will make use of the following simple modifcation of the Azuma-Hoeffding inequality.

Lemma 3.2. *Let $Z = Z(X_1, X_2, \ldots, X_M) \geq 0$ where $X_1, X_2, \ldots, X_M$ are independent random variables. Let $\mathcal{E} = \mathcal{E}(X_1, X_2, \ldots, X_M)$ be an event. Suppose that if $\mathcal{E}$ occurs, then changing a single X_i can only change Z by at most A_1. Then, for any $t > A_0 \mathbf{Pr}(\bar{\mathcal{E}})$ we have*

$$\mathbf{Pr}(Z \geq \boldsymbol{E}[Z] + t) \leq \exp\left\{-\frac{t^2}{2MA_1^2}\right\} + \mathbf{Pr}(\bar{\mathcal{E}}).$$

Proof. We have

$$\begin{aligned}\mathbf{Pr}(Z \geq \mathbf{E}[Z] + t) &= \mathbf{Pr}(Z1_{\mathcal{E}} \geq \mathbf{E}[Z] + t) + \mathbf{Pr}(Z1_{\bar{\mathcal{E}}} \geq \mathbf{E}[Z] + t) \\ &\leq \mathbf{Pr}(Z1_{\mathcal{E}} \geq \mathbf{E}[Z1_{\mathcal{E}}] + u) + \mathbf{Pr}(\bar{\mathcal{E}}) \end{aligned} \tag{3.5}$$

where

$$u = \mathbf{E}[Z] - \mathbf{E}[Z1_{\mathcal{E}}] + t \geq t. \tag{3.6}$$

Applying the Azuma-Hoeffding inequality (more precisely, the special case referred to as McDiarmid's inequality) we get

$$\mathbf{Pr}(Z1_{\mathcal{E}} \geq \mathbf{E}[Z1_{\mathcal{E}}] + u) \leq \exp\left\{-\frac{u^2}{2MA_1^2}\right\}. \tag{3.7}$$

The lemma follows after using (3.6) and (3.7) in (3.5). □

Claim 3.2. *With probability $1 - o(n^{-2})$, Γ contains at most $n^{1/2+o(1)}$ cycles of length at most $\Lambda = (\log\log n)^2$.*

Proof of Claim: The expected number of cycles of length at most $2\ell = \Lambda$ is bounded by

$$\sum_{s=2}^{\ell} \binom{n}{s}\binom{m}{s}(s!)^2\left(\frac{d}{m}\right)^{2s} \leq \sum_{s=2}^{\ell} d^{2s} = n^{o(1)}.$$

Let C denote the number of cycles of length at most Λ. We apply Lemma 3.2 to C with $M = dn$, $\mathcal{E} = \left\{\Delta \leq \lambda = (\log n)^2\right\}$ and $A_1 = \lambda^\Lambda$ and $t = n^{1/2} A_0 \log n = n^{1/2+o(1)}$. We use Claim 3.1 to bound $\mathbf{Pr}(\bar{\mathcal{E}})$.
End of proof of Claim

These two claims imply the following:

$$\text{With probability } 1 - o(n^{-2}) \text{ there are at most } n^{1/2+o(1)}(3\log n)^{2\ell} = n^{1/2+o(1)} \text{ vertices within distance at most } 2A_0 \log\log n \text{ of a cycle of length at most } \Lambda = (\log\log n)^2. \tag{3.8}$$

Now let $p_{k,\ell}$ denote the probability that INSERT requires at least ℓ steps to insert v_k.

We finish the proof of the theorem by showing that

$$\mathbf{E}[|P_k|] = 1 + 2\sum_{\ell=2}^{\infty} p_{k,\ell} \leq 1 + \frac{2}{\theta}. \tag{3.9}$$

We observe that if v_k has no neighbor in $\bar{R}_{k-1}$ and has no neighbor in a cycle of length at most Λ then for some $\ell \leq A_0 \log\log n$, the first $2\ell - 1$ vertices of P_n follow an interesting path. Hence, if

$d^2\gamma \leq (1-\theta)(d-1)$ then

$$\sum_{\ell=2}^{A_0 \log\log n} p_{k,\ell} \leq O(n^{-1/2+o(1)}) + \sum_{\ell=2}^{A_0 \log\log n} \frac{\nu_{k,\ell}}{k(d-1)^\ell}$$
$$\leq O(n^{-1/2+o(1)}) + \sum_{\ell=2}^{A_0 \log\log n} \frac{k\gamma(d^2\gamma)^{\ell-1}}{k(d-1)^\ell}$$
$$\leq o(1) + (1+\theta)\sum_{\ell=2}^{\infty}(1-\theta)^{\ell-1} = o(1) + \frac{1-\theta^2}{\theta}. \quad (3.10)$$

Explanation of (3.10)**:** Following (3.8), we find that the probability v_k is within $2A_0 \log\log n$ of a cycle of length at most Λ is bounded by $n^{-1/2+o(1)}$. The $O(n^{-1/2+o(1)})$ term accounts for this and also absorbs the error probability in (3.4). Failing this, we have divided the number of interesting paths of length $2\ell-1$ by the number of equally likely walks $k(d-1)^\ell$ that INSERT could take. To obtain $k(d-1)^\ell$ we argue as follows. We carry out the following thought experiment. We run our walk for ℓ steps regardless. If we manage to choose $y \in \bar{R}_{k-1}$ then instead of stopping, we move to v_k and continue. In this way there will in fact be $k(d-1)^\ell$ equally likely walks. In our thought experiment we choose one of these walks at random, whereas in the execution of the algorithm we only proceed as far the first time we reach $\bar{R}_{k-1}$. Finally, for the algorithm to take at least ℓ steps, it must choose an interesting path of length at least $2\ell-1$.

Note next that

$$p_{k,A_0 \log\log n} \leq O(n^{-1/2+o(1)}) + 3^{-A_0 \log\log n}.$$

It follows that

$$\sum_{\ell=A_0 \log\log n}^{(\log n)^{A_0}} p_{k,\ell} \leq \sum_{\ell=A_0 \log\log n}^{(\log n)^{A_0}} p_{k,A_0 \log\log n} = o(1). \quad (3.11)$$

We will use the result of [50]: We phrase Claim 10 of that paper in our current terminology.

Claim 3.3. *There exists a constant $a > 0$ such that for any $v \in L_{k-1}$, the expected time for* INSERT *to reach $\bar{R}_{k-1}$ is $O((\log k)^a)$.*

It follows from Claim 3.3 that for any integer $\rho \geq 1$,

$$\mathbf{Pr}(|P_k| \geq \rho(\log k)^{2a}) \leq \frac{1}{(\log k)^{\rho a}}. \quad (3.12)$$

Indeed, we just apply the Markov inequality every $(\log k)^{2a}$ steps to bound $|P_k|$ by a geometric random variable.

It follows from (3.12) that

$$\sum_{\ell \geq 3(\log k)^{2a}} p_{k,\ell} \leq \sum_{\rho=3}^{\infty} \sum_{\ell/(\log k)^{2a} \in [\rho,\rho+1]} p_{k,\ell} \leq \sum_{\rho=3}^{\infty} \frac{1}{(\log k)^{\rho a - 2a}} = o(1). \quad (3.13)$$

Theorem 3.1 now follows from (3.9), (3.10), (3.11) and (3.13), if we take $A_0 > 2a$.

3.2.3 Proof of Lemma 3.1

We apply Lemma 3.2 to $\nu_{k,\ell}$ to argue that

$$\mathbf{Pr}(\nu_{k,\ell} \geq \mathbf{E}[\nu_{k,\ell}] + n^{3/4}) \leq 2e^{-(\log n)^2}. \tag{3.14}$$

We let $M = kd$, $\mathcal{E} = \left\{\Delta \leq \lambda = (\log n)^2\right\}$ as before, $A_1 = \lambda^{2\ell}$ and $t = n^{3/4}$. We use Claim 3.1 to bound $\mathbf{Pr}(\bar{\mathcal{E}})$. The bound on A_1 follows from the fact that an edge can be in at most $\Delta^{2\ell}$ interesting paths.

It follows from (3.14) that to finish the proof, all we need to show is that if $\theta > 0$ is an arbitrary positive constant

$$\mathbf{E}[\nu_{k,\ell}] \leq (1+\theta)k\gamma(d^2\gamma)^{\ell-1}, \tag{3.15}$$

where γ is as in (3.1).

Claim 3.4. *Let*

$$\mathcal{B}_k = \{|B_k| \geq k\gamma\}.$$

Then

$$\mathbf{Pr}(\mathcal{B}_k) = O(e^{-\Omega(n^{1/2})}).$$

Proof of Claim:
Let $B_{k,1}$ denote the set of vertices $v_i \in L_k$ such that round i exposes at least $d/2$ edges incident with v_i. Then

$$\mathbf{Pr}(v_i \in B_{k,1}) \leq (1-\varepsilon)^{d/2}.$$

It then follows from the Chernoff bounds that

$$\mathbf{Pr}\left(|B_{k,1}| \geq 2k(1-\varepsilon)^{d/2}\right) = O(e^{-\Omega(n^{1/2})}). \tag{3.16}$$

Next let

$$B_{k,2} = \{s \leq k : \text{ round } s \text{ does not end immediately in Step 2 with } x = v_s.\}$$

Then, $\mathbf{Pr}(s \in B_{k,2}) = \left(\frac{s-1}{m}\right)^d$ and this holds for each value of s independently and so

$$\mathbf{E}[|B_{k,2}|] \leq \sum_{s=1}^{k} \left(\frac{s-1}{m}\right)^d \leq \frac{k^{d+1}}{(d+1)m^d}.$$

Now $|B_{k,2}|$ is the sum of independent $\{0,1\}$ random variables and so Hoeffding's theorem [57] implies that for a constant $\theta > 0$,

$$\mathbf{Pr}(|B_{k,2}| \geq (1+\theta)\frac{k^{d+1}}{(d+1)m^d} = O(e^{-\varepsilon_1 k}) \text{ for some constant } \varepsilon_1 = \varepsilon_1(d,\varepsilon,\theta) > 0. \tag{3.17}$$

Now if $B_{k,3} = \{s \in B_k : \exists \ell \leq k, \ell \neq s \text{ s.t. round } \ell \text{ ends with } x = v_s\}$ then $|B_{k,3}| \leq |B_{k,2}|$. Define $B_{k,4} = B_k \setminus (B_{k,1} \cup B_{k,2} \cup B_{k,3})$. Let $t > s$ be the first time that v_s is re-visited by INSERT or let $t = k$ if v_s is not re-visited. Then $s \in B_{k,4}$ only if in round t, at least $d/2$ unexposed edges incident to s are found to be in R_{t-1}. It follows that

$$E(|B_{k,4}|) \leq k(1-\varepsilon)^{d/2}.$$

Since membership of s in $B_{k,4}$ is determined by the random choices of v_s, $|B_{k,4}|$ is the sum of independent random variables and so

$$\mathbf{Pr}\left(|B_{k,4}| \geq 2k(1-\varepsilon)^{d/2}\right) = O(e^{-\Omega(n^{1/2})}). \tag{3.18}$$

The claim follows from (3.16), (3.17) and (3.18).
End of proof of Claim

Given Claim 3.4, we have

$$\begin{aligned}\mathbf{E}\left[\nu_{k,\ell}\right] &= \mathbf{E}\left[\nu_{k,\ell} \mid \neg\mathcal{B}_k\right]\mathbf{Pr}(\neg\mathcal{B}_k) + \mathbf{E}\left[\nu_{k,\ell} \mid \mathcal{B}_k\right]\mathbf{Pr}(\mathcal{B}_k) \\ &\leq k^\ell\gamma^\ell k^{\ell-1}\cdot\left((1+o(1))\frac{d}{k}\right)^{2\ell-2} + O(k^{2\ell-1}\cdot e^{-\Omega(n^{1/4})}), \\ &\leq (1+o(1))k\gamma(d^2\gamma)^{\ell-1} + o(1).\end{aligned} \tag{3.19}$$

This proves (3.15).

Explanation of (3.19)**:** We can choose the vertex sequence $\sigma = (x_1, \xi_1, \ldots, \xi_{\ell-1}, x_\ell)$ of an interesting path P in at most $|B_k|^\ell k^{\ell-1}$ ways, and we apply Claim 3.4. Having chosen σ we see that $((1+o(1))d/k)^{2\ell-2}$ bounds the probability that the edges of P exist. To see this, condition on $\bar{R}_{k-1}$ and the random choices for vertices not on P. In particular, we can fix $R_{k-1} = \{y_1, y_2, \ldots, y_{k-1}\}$ from the beginning and this simply constrains the sequence of choices $y_1, y_2, \ldots, y_{k-1}$ to be a uniformly random permutation of R_{k-1}. Let $\mathcal{M}_k$ be the property that Γ has a matching from L_k to R. It is known that $\mathbf{Pr}(\mathcal{M}_k) = 1 - O(n^{4-d})$. This will also be true conditional on the value of $\bar{R}_{k-1}$. This follows by symmetry. The conditional spaces will be isomorphic to each other. So for large d, we can assume that our conditioning is such that with probability $1 - O(1/n^3)$ the edge choices by $x_1, x_2, \ldots, x_\ell$ are such that Γ_k has property $\mathcal{M}_k$ with probability $1 - O(n^{7-d})$. Recall from (3.3) that the disposition of the edges of Γ_{k-1} is independent of $\bar{R}_{k-1}$. Now each edge adjacent to a given $x \in \sigma \cap L_k$ is a uniform choice over those edges consistent with x being in B_k. But there will be at least $k-1$ such choices for such an x viz. the vertices of R_{k-1}. Thus

$$\mathbf{Pr}(P \text{ exists} \mid \mathcal{M}_k) \leq \frac{\mathbf{Pr}(P \text{ exists})}{\mathbf{Pr}(\mathcal{M}_k)} \leq (1+o(1))\left(\frac{d}{k}\right)^{2\ell-2}.$$

Note that $\mathbf{Pr}(\bar{\mathcal{M}}_k)$ is only inflated by at most $\frac{1}{(1-\varepsilon)^{d\ell}} = o(n^{o(1)})$ if we condition on $x_1, x_2, \ldots, x_\ell$ making their choices in $\bar{R}_{k-1}$. This has to be compared with the unconditional probability of $O(n^{7-d})$.

This completes the proof of Theorem 3.1. □

Remark 3.1. *Along with an upper bound, we can prove a simple lower bound:*

$$\boldsymbol{E}\left[|P_k|\right] \geq \frac{2}{1-(1-\varepsilon)^d}.$$

This follows from the fact that Step 2 of INSERT *ends the procedure with probability* $1-(1-\varepsilon)^{|S_k(x)|}$ *and* $|S_k(x)| \leq d$.

3.3 Final Remarks

There is plenty of room for improvement in the bounds on d in Theorem 3.1. It would be most interesting to prove an $O(1)$ bound on the expected insertion time for small d, e.g. $d = 3, 4, 5$. This no doubt requires an understanding of the evolution of the matching M.

Acknowledgement: We thank Wesley Pegden and the reviewers for their comments.

Chapter 4

Preferential attachment with edge deletion

This chapter corresponds to [60].

Abstract

We consider a variation on the Barabási-Albert random graph process with fixed parameters $m \in \mathbb{N}$ and $1/2 < p < 1$. With probability p a vertex is added along with m edges, randomly chosen proportional to vertex degrees. With probability $1 - p$, the oldest vertex still holding its original m edges loses those edges. It is shown that the degree of any vertex either is zero or follows a geometric distribution. If p is above a certain threshold, this leads to a power law for the degree sequence, while a smaller p gives exponential tails. It is also shown that the graph contains a unique giant component with high probability if and only if $m \geq 2$.

4.1 Introduction

In recent years, considerable attention has been paid toward real-world networks such as the World Wide Web (e.g. [30]) and social networks such as Facebook [87] and Twitter [73]. Many but not all of these networks exhibit a so-called power law, and are sometimes referred to as scale free, meaning that the number of elements of degree k is asymptotically $k^{-\eta}$ for some constant $\eta > 0$. In [7] it is shown that the social network of scientific collaborations is scale free. For a number of real-world scale free networks see [7].

As a means of describing such networks with a random graph, Barabási and Albert [2] introduced a class of models, commonly called preferential attachment graphs, and argued that its degree distribution has a tail that decreases polynomially, a claim that was subsequently proved by Bollobás, Riordan, Spencer and Tusnády [14]. This is in contrast to many well-known random graph models such as the Erdős-Rényi model where the degree distribution has an exponential tail. While the Barabási-Albert model in its basic form falls short as a description of the World Wide Web [1], the model has become popular for modelling scale free networks.

The base principle of preferential attachment graphs is the following: vertices are added sequentially to the graph, along with edges that attach themselves to previously existing vertices with probability proportional to their degree. This principle is susceptible to many variations, and can be combined with other random graph models. See for example Flaxman, Frieze and Vera [32], [34], who introduced a random graph model combining aspects of preferential attachment graphs and random geometric graphs.

Real-world networks will encounter both growth and deletion of vertices and edges. Bollobás and Riordan [13] considered the effect of deleting vertices from the graph after it has been generated. Cooper, Frieze and Vera considered random vertex deletion [20], and Flaxman, Frieze and Vera considered adversarial vertex deletion [33], where vertices are deleted while the graph is generated. Chung and Lu [18] considered a general growth-deletion model for random power law graphs.

In this paper, we consider a preferential attachment model in which the oldest edges are regularly removed while the graph is generated. There are two fixed parameters, an integer $m \geq 1$ and a real number $1/2 < p < 1$. As the graph is generated, with probability p we add a vertex along with m edges to random endpoints proportional to their degree. Choices are made with or without replacement. The vertices are ordered by time of insertion, and with probability $1-p$ we remove all edges that were added along with a vertex, where the vertex is the oldest for which this has not already been done. This is a new variation of the preferential attachment model, and the focus on the paper is to find the degree sequence of this graph. The proof method also leads to a partial result on the existence of a giant component.

In Theorem 4.2 we find the degree sequence of the graph, and show that it exhibits a phase transition at $p = p_0 \approx 0.83$, independently of m. If $p > p_0$ then the degree sequence follows a power law, while $p < p_0$ gives exponential tails. A real-world example of this behaviour is given by family names; in [72] it is shown that the frequency of family names in Japan follow a power law, while [63] shows that family names in Korea decay exponentially.

We prove three theorems. The first deals with the degree distribution of any fixed vertex, show that it is the sum of m independent variables that are either zero or geometrically distributed. We let G_n denote the nth member of the graph sequence described above. The notation given here is imprecise at this point, but the theorems will be restated with precise notation below.

Let $\mathcal{D}$ be the event that at some point of the graph process, the graph contains no edges. The probability of $\mathcal{D}$ is addressed in Lemma 4.1, and we will be conditioning on $\mathcal{D}$ not occurring. At this point we remark that if the process starts with a graph with ω_H edges where $\omega_H = \omega_H(n) \to \infty$ as $n \to \infty$, then $\Pr\{\mathcal{D}\} = o(1)$. Note that the ω in the following theorem is different from ω_H.

Theorem 4.1. *Suppose $\omega = o(\log n)$ tends to infinity with n. Let $d(n,v)$ denote the degree of vertex v in G_n. Conditioning on $\overline{\mathcal{D}}$, there exist functions $p(n,v), q(n,v)) \in [0,1]$ and a constant $0 < c < 1/2$ such that $d(n,v)$ is distributed as the sum $d_1(n,v) + d_2(n,v) + \cdots + d_m(n,v)$ of independent random variables with*

$$Pr\{d_i(n,v) = k\} = \begin{cases} 1 - q(n,v) + O(n^{-c}), & k = 0, \\ q(n,v)p(n,v)(1-p(n,v))^{k-1} + O(n^{-c}), & k > 0, \end{cases}$$

for $i = 1, \ldots, m$, for all $v \geq n/\omega$.

We do not address the degrees of vertices $v < n/\omega$. In particular, we present no bounds for the maximum degree of G_n. We have instead focused on finding the degree sequence and connected components of G_n.

The second theorem translates the degree distribution into a degree sequence for G_n. It shows that the graph follows a power law if and only if p is above a certain threshold.

Theorem 4.2. *Let $p_0 \approx 0.83113$ be the unique solution in $(1/2, 1)$ to $p/(4p-2) = \ln(p/(1-p))$. Let $X_k(n)$ denote the number of vertices of degree k in G_n. Conditioning on $\overline{\mathcal{D}}$, there exists a sequence $\{x_k : k \geq 0\}$ such that*

(i) if $\alpha < 1$ then $x_k = \alpha^{k(1+o_k(1))}$ and if $\alpha > 1$ then there exist constants $a, b > 0$ such that $x_k = ak^{-\eta-1} + O_k(k^{-\eta-2}\log^b k)$, where $\eta = \eta(p) > 2$ is defined for $p > p_0$, and

(ii) for any fixed $k \geq 0$, $X_k(n) = x_k n(1 + o_n(1))$ with high probability[1].

The third theorem shows that G_n has a giant component if and only if $m \geq 2$. This is in contrast to the classical Barabási-Albert model which is trivially connected. Let $B(n) = \lambda \ln n$ when $p < p_0$ and $B(n) = \lambda n^{1/\eta} \ln n$ when $p > p_0$ for some constants $\lambda > 0$, $\eta > 2$, explicitly defined later. Note that when $p > p_0$ and m is large, Theorem 4.3 states that the number of vertices which are not in the largest component is $O_m(c^m n)$ for some $0 < c < 1$, since the total number of vertices in G_n will be shown to be $pn(1 + o_n(1))$ whp.

Theorem 4.3. *Condition on $\overline{\mathcal{D}}$.*

(i) If $m = 1$, the largest component of G_n has size $O(\Delta \log n)$ with high probability, where Δ is the maximum degree of G_n.

(ii) If $m \geq 2$, there exists a constant $\xi > 0$ such that with high probability the number of isolated vertices is ξpn, the largest component contains at least $(1-\xi)(1-(13/14)^{m-1})pn$ vertices, and all other components have size $O(\log n)$. If $p > p_0$ then $\xi = O_m(c^m)$ for some $0 < c < 1$.

4.1.1 Proof outline

The paper is laid out as follows. In Section 4.2 we define the graph process precisely and define constants and functions that are central to the main results. Section 4.3 is devoted to Crump-Mode-Jagers processes, which will be the central tool in studying the graph process. Sections 4.4, 4.5 and 4.6 are devoted to proving Theorem 4.1, 4.2 and 4.3 respectively.

We will now outline the proof of Theorem 4.1. Theorem 4.2 is an elementary consequence of Theorem 4.1, and the proof of Theorem 4.3 is heavily based on the method used to prove Theorem 4.1.

In Section 4.2.1 we will define a *master graph* Γ which contains G_t for all t. We will mainly be proving results for Γ and show how they transfer to G_n, but for this informal outline we will avoid the somewhat technical definition of Γ and show the idea behind the proofs.

We begin by describing the Crump-Mode-Jagers process (or CMJ process). The name Crump-Mode-Jagers applies to a more general class than what is considered here, but we will mainly be talking about the special case described as follows. Fix a constant $\alpha > 0$ and consider a Poisson process $\mathcal{P}_0$ with rate α on $[0, 1)$. Suppose $\mathcal{P}_0$ has arrivals at time $\tau_{01} < \tau_{02} < \cdots < \tau_{0k}$. The jth arrival gives rise to a Poisson process $\mathcal{P}_{0j}$ on $[\tau_{0j}, \tau_{0j} + 1)$, $j = 1, \ldots, k$, independent of all other

[1] We say that a sequence of events $\mathcal{E}_n$ occur with high probability (whp) if $\Pr\{\mathcal{E}_n\} \to 1$ as $n \to \infty$

Poisson processes. In general, let $s = 0\dots$ be a string of integers starting with 0 and suppose $\mathcal{P}_s$ is a Poisson process on $[\tau_s, \tau_s + 1)$. Then the jth arrival in $\mathcal{P}_s$, at time τ_{sj}, gives rise to a Poisson process $\mathcal{P}_{sj}$ on $[\tau_{sj}, \tau_{sj} + 1)$. Here sj should be interpreted as appending j to the end of the string s. Let $d(\tau)$ be the number of processes alive at time τ, i.e. the number of processes $\mathcal{P}_s$ with $\tau_s \in (\tau - 1, \tau]$. Lemma 4.3 will show that for fixed τ, $d(\tau)$ is either zero or geometrically distributed.

We will now explain how the degree of a vertex in the graph process relates to a CMJ process. Firstly, note that choosing a random vertex with probability proportional to degrees is equivalent to choosing an edge e uniformly at random, and choosing one of the two endpoints of e uniformly at random. We will refer to this as choosing a *half-edge* (e, ℓ) where $\ell \in \{1, 2\}$. If $e = \{v, w\}$ is added along with vertex v, we say that choosing $(e, 1)$ corresponds to choosing w *via* e, and choosing $(e, 2)$ corresponds to choosing v via e. This is described in detail in Section 4.2.

For the purpose of demonstration consider the case $m = 1$, i.e. the case in which exactly one edge is added along with any vertex added to the graph. It will follow from Lemma 4.1 that if a vertex v_0 is added along with an edge e_0 at time t_0 then with high probability e_0 is removed at time $\gamma t_0 + o(t)$, where $\gamma = p/(1-p)$. Note that the degree of v_0 may still be non-zero after the removal of e_0. If the degree of v_0 is to increase from its initial value 1, then there must exist a time T_{01} with $t_0 < T_{01} < \gamma t_0 + o(t_0)$ at which a vertex v_{01} is added along with edge e_{01}, where e_{01} is randomly assigned to $(e_0, 2)$. The time T_{01} is random and we will see (equation (4.1)) that $\log_\gamma(T_{01}/t_0) \in (0, 1 + o(1))$ is approximately exponentially distributed with rate $\alpha = \alpha(p)$. Furthermore, if $T_{01} < T_{02} < \cdots < T_{0k}$ denote the times at which a vertex is added that chooses v_0 via e_0, then the sequence $(\log_\gamma(T_{0i}/t_0))$ can be approximated by a Poisson process with rate α on the interval $(0, 1)$. Let e_{01} denote the edge that is added at time T_{01} and chooses $(e_0, 2)$ (if such an edge exists). Then the degree of v_0 may increase by some edge e_{011} added at time T_{011} with $T_{01} < T_{011} < \gamma T_{01} + o(T_{01})$ choosing $(e_{01}, 1)$, i.e. choosing v_0 via e_{01}. As above, the sequence of times $T_{011}, T_{012}, \dots, T_{01\ell}$ at which a vertex is added that chooses v_0 via e_{01} are such that $(\log_\gamma(T_{01i}/T_{01}))$ approximately follows a Poisson process on $(\log_\gamma T_{01}, 1 + \log_\gamma T_{01})$. Repeating the argument, any edge incident to v_0 gives rise to a Poisson process, and as long as the degree of v_0 is not too large the processes are "almost independent". Under the time transformation $\tau(t) = \log_\gamma(t/t_0)$, the times at which the degree of v_0 increases or decreases can be approximated by the times at which $d(\tau)$ increases or decreases in a CMJ process with rate α. This approximation is made precise in the proof of Theorem 4.1, and shows that the degree of a vertex is either zero or approximately geometrically distributed.

Now suppose $m > 1$. Then each of the m edges added along with v gives rise to a CMJ process by the argument above, and the processes are "almost independent". The degree of v will be approximated by a sum of m independent random variables that are each either zero or geometrically distributed.

4.2 The model

Fix $m \in \mathbb{N}$ and $1/2 < p < 1$. Let $\mathcal{G}_m$ be the class of undirected graphs on $[\nu_G] = \{1, \dots, \nu_G\}$ for some integer ν_G such that if edges are oriented from larger integers to smaller, there exists some integer 1_G with $m < 1_G \le \nu_G$ such that a vertex v has out-degree m if $v \ge 1_G$ and out-degree zero if $v < 1_G$. All graphs we deal with will be in $\mathcal{G}_m$. In some places it will be convenient to think of graphs as being directed, in which case we always refer to the orientation from larger to smaller integers. We will allow parallel edges but no self-loops.

Our graph G will be defined by $G = G_n$ for some graph sequence (G_t) and some n that grows to infinity. Each G_t will be in $\mathcal{G}_m$, and we write $1_t = 1_{G_t}, \nu_t = \nu_{G_t}$. Given G_t, we randomly generate G_{t+1} as follows. With probability $1-p$, remove all m edges oriented out of 1_t, so that $1_{t+1} = 1_t + 1$. Note that edges oriented into 1_t remain in G_{t+1}. With probability p, add vertex $\nu_{t+1} = \nu_t + 1$ along with m edges to distinct vertices, where vertices are picked with probability proportional to their degree with replacement. In other words, if $d(t, v)$ denotes the degree of vertex v in G_t, then ν_{t+1} is added along with edges (ν_{t+1}, v_i) where $v_1, \ldots, v_m$ are independent with

$$\Pr\{v_i = v\} = \frac{d(t-1, v)}{e(G_{t-1})}$$

where $e(G_{t-1})$ denotes the number of edges in G_{t-1}. Rather than using the terminology of ν_{t+1} choosing $v_1, \ldots, v_m$, we will say that the m edges $e_1, \ldots, e_m$ added at time $t+1$ choose $v_1, \ldots, v_m$ respectively. Let $d^+(t, v), d^-(t, v)$ denote the out- and in-degree of v in G_t in the natural orientation. Write $d(t, v) = 0$ if $v \notin G_t$. The issue of the empty graph appearing in the process is addressed shortly.

We will assume that the graph process starts with some graph $H \in \mathcal{G}_m$ on $\nu = o(n^{1/2})$ vertices, and we label this graph G_{t_0} where $t_0 = 1_H + \nu_H$ in order to maintain the identity $1_t + \nu_t = t$ for every G_t, $t_0 \le t \le n$. Let $\sigma \in \{0, 1\}^{n-t_0}$ be such that $\sigma(u)$ is the indicator for if a vertex and m edges are added at time $u + t_0$, or if m edges are removed at time $u + t_0$. Then $\nu_t = \nu_H + \sum_{u=t_0+1}^{t} \sigma(u)$ for all $t > t_0$, and $1_t = t - \nu_t$. The entries $\sigma(u)$ are independent and $\sigma(u) = 1$ with probability p. Say that σ is *feasible* if it is such that $\nu_t > 1_t$ for all $t > t_0$, noting that $\{\sigma \text{ is feasible}\} = \overline{\mathcal{D}}$ with $\mathcal{D}$ as in Section 4.1. For a function $\omega = \omega(n)$ such that $\omega \to \infty$ as $n \to \infty$, We say that σ is *ω-concentrated* if $|\nu_t - pt| \le t^{1/2} \ln t$ for all $t \ge n/\omega$. Note that if σ is ω-concentrated then $|1_t - (1-p)t| \le t^{1/2} \ln t$ and $|e(G_t) - m(2p-1)t| \le mt^{1/2} \ln t$ for all $t \ge n/\omega$. Furthermore, if an edge e is added at time $t \ge n/\omega$ then it is removed at time $pt/(1-p) + O(t^{1/2} \ln t)$.

Lemma 4.1. *Let $\omega = \omega(n) \to \infty$ with n. If the graph process is initiated at $H \in \mathcal{G}_m$ on $\nu_H \le \omega^{-1} n^{1/2}$ vertices and $\nu_H - 1_H = N$, then σ is feasible with probability $1 - O(c^N)$, i.e. $Pr\{\mathcal{D}\} = O(c^N)$, for some constant $c \in (0, 1)$. Furthermore, σ is ω-concentrated with probability $1 - O(n^{-C})$ for any $C > 0$.*

Proof. Recall that $\sigma(t) = 1$ with probability p and $\sigma(t) = 0$ otherwise. The difference $\nu_t - 1_t$ is a random walk, and the fact that $\nu_t \ge 1_t$ for all $t \ge t_0$ with probability $1 - O(c^N)$ for some $c \in (0, 1)$ is well known (see e.g. [55, Section 5.3]).

Suppose $t \ge n/\omega$. Then $t - t_0 \ge n/2\omega$ and by Hoeffding's inequality [57], since $pt_0 - \nu_H = o((n/\omega)^{1/2}) = o(t^{1/2} \ln t)$,

$$\Pr\left\{\nu_t - pt > t^{1/2} \ln t\right\} = \Pr\left\{\left(\sum_{u=t_0+1}^{t} \sigma(u)\right) - p(t - t_0) > pt_0 - \nu_H + t^{1/2} \ln t\right\} = e^{-\Omega(\ln^2 t)}$$

Summing over $t = n/\omega, \ldots, n$ shows that $\nu_t \le pt + t^{1/2} \ln t$ for all $t \ge n/\omega$ whp, and similarly $\nu_t \ge pt - t^{1/2} \ln t$ for all $t \ge n/\omega$ whp. □

4.2.1 The master graph

The above description of G_t is limited in that it forces one to generate the graph on-line, i.e. vertices need to make their random choices in a fixed order. Conditioning on σ we can define an off-line graph Γ which contains G_t for all t. This graph enables us to generate small portions of the graph without revealing a large part of the probability space.

Fixing a feasible σ we define a *master graph* $\Gamma = \Gamma_n^\sigma(H)$ which has G_t as a subgraph (in the sense that G_t can be obtained from Γ by removing edges and possibly vertices) for all $t_0 \le t \le n$. There are two key observations that allow the construction. Firstly, if σ is fixed, then $\nu_t = \nu_H + \sum_{u=1}^{t-t_0} \sigma(u)$ is known for all $t_0 \le t \le n$. This means that $1_t = t - \nu_t$ is known, and we know that the $m(\nu_t - 1_t + 1)$ edges in G_t are those added along with $1_t, 1_t + 1, \ldots, \nu_t$, for all $t_0 \le t \le n$. Secondly, suppose a vertex v is added along with edges $e_1, \ldots, e_m$ at time $t > t_0$. Rather than using the terminology of v choosing m vertices $v_1, \ldots, v_m \in G_{t-1}$ with probability proportional to their degrees, we will adopt the terminology of the edges e_i independently choosing edges $f_i \in G_{t-1}$ uniformly at random, then choosing one of the two endpoints of f_i uniformly at random. To make this formal, let E_e^σ be the edges that are in the graph when the edge e is added, noting that if e is added at time t then $E_e^\sigma = \{m(1_{t-1} - 1) + 1, \ldots, m\nu_{t-1}\}$ with $1_{t-1}, \nu_{t-1}$ determined by σ. Then each e_i independently chooses an $f_i \in E_{e_i}^\sigma$ uniformly at random, along with $j_i \in [2]$ chosen uniformly at random. If $f_i = \{u, u'\}$ with $u' < u$, then e_i choosing $(f_i, 1)$ means e_i chooses u', and $(f_i, 2)$ means choosing u. We say that f_i chooses u (or u') *via* f_i. We call a pair (e, j) with $j \in [2]$ a *half-edge*.

Suppose the graph process is initiated with some graph $G_{t_0} = H \in \mathcal{G}_m$ on $[\nu_H]$ with $1_H + \nu_H = t_0$. We will introduce an integer labelling $L(e)$ for the edges e in Γ. The L will be dropped from calculations and we write $e_1 \ge e_2$ to mean $L(e_1) \ge L(e_2)$ and $f(e) = f(L(e))$ whenever f is a function on the integers. The labelling is defined by labelling the m edges along with $v > \nu_H$ by $m(v-1)+1, m(v-1)+2, \ldots, mv$. The edges in the initial graph H can be oriented in such a way that vertices $1, \ldots, 1_H - 1$ have out-degree zero, and $1_H, \ldots, \nu_H$ have out-degree m. We can then label the edges in H by $m(1_H - 1) + 1, \ldots, m\nu_H$ in such a way that $1_H \le v \le \nu_H$ is incident with edges $m(v-1)+1, \ldots, mv$. Note that under this labelling, every edge e is incident with vertex $\lceil e/m \rceil$ while its other endpoint $v(e)$ will satisfy $v(e) < \lceil e/m \rceil$.

Definition of Γ: Fix a feasible σ and a graph $H \in \mathcal{G}_m$. We define $\Gamma = \Gamma_n^\sigma(H)$ as follows. The vertex set is $[\nu_n]$ where $\nu_n = \nu_H + \sum_{i=1}^{n-t_0} \sigma(i)$. The graph Γ contains H as an induced subgraph on $[\nu_H]$. Every edge $e > m\nu_H$ is associated with a set $\Omega(e) = E_e^\sigma \times [2]$, and makes a random choice $\phi(e) = (f(e), j(e)) \in \Omega(e)$ uniformly at random, independent of all other edges. One endpoint of e is $\lceil e/m \rceil$ (the fixed endpoint) and one is $v(e)$ (the random endpoint). If $j(e) = 2$ then $v(e) = \lceil f(e)/m \rceil$. If $j(e) = 1$ then $v(e) = v(f(e))$.

Note the recursion in defining the random endpoint $v(e)$ of an edge e. If $j(e) = 1$ and $j(f(e)) = 1$ then $v(e) = v(f(e)) = v(f(f(e)))$, and so on until either $j(f^{(k)}(e)) = 2$ for some k, or $f^{(k)}(e) \le m\nu_H$ for some k, in which case $v(e) = v(f^{(k)}(e))$ is determined by H. Here $f^{(k)}$ denotes k-fold composition of f.

We will generate Γ carefully by keeping a close eye on the sets $\Omega(e)$. Let Γ_0 be the graph in which no randomness has been revealed; in Γ_0 only the edges in H are known, all other edges are free, and all sets $\Omega(e) = \Omega_0(e) = E_e^\sigma \times [2]$. For sets $A \subseteq \{m\nu_H + 1, m\nu_H + 2, \ldots, m\nu_n\}$ of free edges and $R \subseteq \{m(1_H - 1) + 1, m(1_H - 1) + 2, \ldots, m\nu_n\} \times [2]$ of half-edges, define a class $\mathcal{G}(A, R)$ of *partially generated* graphs as follows. We say that $\widetilde{\Gamma} \in \mathcal{G}(A, R)$ if (i) for $e > m\nu_H$, $\phi(e)$ is known if and only if $e \in A$, and (ii) for all $e \notin A$ we have $\Omega(e) \supseteq \Omega_0(e) \setminus R$. In other words, if $(f, j) \in R$ then for each e with $(f, j) \in \Omega_0(e)$, we may have determined that $\phi(e) \ne (f, j)$.

Given a partially generated $\widetilde{\Gamma} \in \mathcal{G}(A, R)$, we define two operations that reveal more information about Γ. We say that we *assign* $e \notin A$ when we choose $\phi(e)$ uniformly at random from $\widetilde{\Omega}(e) =$

$\Omega_0(e) \setminus R$. For any (f, j) we can *reveal* (f, j) to find the $\phi^{-1}(\{f, j\}) \setminus A$ of edges $e \notin A$, free in $\widetilde{\Gamma}$, that choose (f, j). We reveal (f, j) as follows. For every edge $e \notin A$ with $(f, j) \in \widetilde{\Omega}(e)$, set $\phi(e) = (f, j)$ with probability $1/|\widetilde{\Omega}(e)|$, and otherwise remove (f, j) from $\widetilde{\Omega}(e)$.

Starting with $\Gamma_0 \in \mathcal{G}(\emptyset, \emptyset)$ (this class contains only one graph), we can generate Γ by a sequence of *assigns* and *reveals*. Given $\Gamma_i \in \mathcal{G}(A_i, R_i)$, we can assign $e \notin A_i$ to form $\Gamma_{i+1} \in \mathcal{G}(A_i \cup \{e\}, R_i)$, and we set $A_{i+1} = A_i \cup \{e\}$ and $R_{i+1} = R_i$. If $(f, j) \notin R_i$ is revealed and $e_1, \ldots, e_k$ are the edges that choose (f, j), we get $\Gamma_{i+1} \in \mathcal{G}(A_i \cup \{e_1, e_2, \ldots, e_k\}, R_i \cup \{(f, j)\})$, and we set $A_{i+1} = A_i \cup \{e_1, \ldots, e_k\}$ and $R_{i+1} = R_i \cup \{(f, j)\}$. We get a sequence $\Gamma_0, \Gamma_1, \ldots$ where $\Gamma_i \in \mathcal{G}(A_i, R_i)$ and $A_i \subseteq A_{i+1}$ and $R_i \subseteq R_{i+1}$ for all i.

Note that in a partially generated graph, if $\phi(e) = (f(e), 1)$ where $f(e)$ is free, then we know that $v(e) = v(f(e))$, but $v(f(e))$ is not yet determined. We say that e is committed to $f(e)$. This can be pictured by gluing the free end of e to the free end of $f(e)$. At a later stage when $f(e)$ is attached to the its random endpoint $v(f(e))$, the edge e will follow and be attached to the same vertex.

We will condition on σ being ω-concentrated for some ω in the proofs to follow. In Γ, this translates to each e with $e \geq mn/\omega$ having $E_e^\sigma = \{e/\gamma + O(n^{1/2} \ln n), \ldots, m(\lceil e/m \rceil - 1)\}$, where $\gamma = p/(1-p)$. Note in particular that $|E_e^\sigma| = e(1 - 1/\gamma) + O(n^{1/2} \ln n)$. Note also that for any edge $e \geq mn/\omega$, the largest f for which $e \in E_f^\sigma$ is $f = \gamma e + O(n^{1/2} \ln n)$.

4.2.2 Constants and functions

In this section we collect constants and functions that will be used throughout the remaining sections. Fixing p and m, we define

$$\mu = m(2p - 1), \quad \gamma = \frac{p}{1 - p}, \quad \alpha = \frac{pm}{2\mu} \ln \gamma = \frac{p}{4p - 2} \ln \gamma.$$

The constant α will play a central role in what follows. We note that it is independent of m, and viewed as a function of $p \in (1/2, 1)$ it is continuously increasing and takes values in $(1/2, \infty)$. Let $p_0 \approx 0.83113$ be the unique p for which $\alpha = 1$. When $\alpha \neq 1$ define ζ as the unique solution in $\mathbb{R} \setminus \{1\}$ to

$$\zeta e^{\alpha(1-\zeta)} = 1.$$

Also let $\eta = -\ln \gamma / \ln \zeta$ if $\alpha > 1$. If $\alpha < 1$ then η is undefined.

Define a sequence a_k by $a_0 = 1$ and

$$a_k = \left(-\frac{e^\alpha}{\alpha}\right)\left(\frac{a_0}{(k-1)!} + \frac{a_1}{(k-2)!} + \cdots + a_{k-1}\right) = \left(-\frac{e^\alpha}{\alpha}\right)\sum_{j=0}^{k-1} \frac{a_j}{(k-j-1)!}, \ k \geq 1.$$

For $k \geq 0$ define functions $Q_k : [k, k+1) \to [0, 1]$ by

$$Q_k(\tau) = \sum_{j=0}^{k} \frac{a_j}{(k-j)!}(\tau - k)^{k-j},$$

and for $\tau \geq 0$ let $Q(\tau) = Q_{\lfloor \tau \rfloor}(\tau)$. We note that $Q(\tau)$ is discontinuous at integer points k with

$$Q(k) = a_k \quad \text{and} \quad Q(k^-) = -\alpha e^{-\alpha} a_k$$

where $Q(k^-)$ denotes the limit of $Q(\tau)$ as $\tau \to k$ from below. Define

$$q(\tau) = 1,\ 0 \leq \tau < 1, \quad q(\tau) = 1 + \frac{Q(\tau-1)}{\alpha Q(\tau)},\ \tau \geq 1.$$

Finally, define

$$p(\tau) = \exp\left\{-\alpha \int_0^\tau q(x)dx\right\}.$$

For $\tau < 0$ we define $Q(\tau) = q(\tau) = p(\tau) = 0$.

In Section 4.6 we will need explicit formulae for $q(\tau)$ for $0 \leq \tau \leq 3$. We have $a_0 = 1$, $a_1 = -e^\alpha/\alpha$ and $a_0 = e^{2\alpha}/\alpha^2 - e^\alpha/\alpha$, so if $0 \leq \tau < 1$ then $Q(\tau) = 1$, $Q(\tau+1) = \tau - e^\alpha/\alpha$ and $Q(\tau+2) = \frac{1}{2}\tau^2 - \alpha^{-1}e^\alpha\tau + e^{2\alpha}/\alpha^2 - e^\alpha/\alpha$, and

$$q(\tau) = 1, \quad q(\tau+1) = 1 - \frac{1}{e^\alpha - \alpha\tau}, \quad q(\tau+2) = 1 - \frac{e^\alpha - \alpha\tau}{e^{2\alpha} - (\tau+1)\alpha e^\alpha + \frac{1}{2}\alpha^2\tau^2},\ 0 \leq \tau < 1.$$

The following lemma collects properties of the constants and functions presented here. Its proof is postponed to Section 4.7.

Lemma 4.2. *(i) If $\alpha > 1$ then $\zeta < \alpha^{-1}$ and if $\alpha < 1$ then $\zeta > 1 - \alpha^{-1} + \alpha^{-2} > \alpha^{-1}$.*

(ii) If $\alpha > 1$ then $\eta > 2$.

(iii) The functions $p(\tau), q(\tau)$ are decreasing and take values in $[0,1]$.

(iv) For any non-integer $\tau > 0$,

$$Q'(\tau) = Q(\tau-1) \quad \text{and} \quad q(\tau) = \frac{1}{\alpha}\frac{(Q(\tau)e^{\alpha\tau})'}{Q(\tau)e^{\alpha\tau}}.$$

(v) If $\alpha < 1$ then there exist constants $\lambda_1, \lambda_2 > 0$ where $\lambda_1 < \alpha$ such that for all $\tau \geq 0$,

$$p(\tau) = 1 - \alpha + \frac{\lambda_1}{\zeta^\tau} + O(\zeta^{-2\tau}) \quad \text{and} \quad q(\tau) = \frac{\lambda_2}{\zeta^\tau} + O(\zeta^{-2\tau}).$$

(vi) If $\alpha > 1$ then there exist constants $\lambda_3, \lambda_4 > 0$ and a constant $C > 0$ such that for all $\tau \geq 0$,

$$\lambda_3\zeta^\tau \leq p(\tau) \leq \lambda_3\zeta^\tau + C\zeta^{2\tau} \quad \text{and} \quad q(\tau) = 1 - \zeta + \lambda_4\zeta^\tau + O(\zeta^{2\tau}).$$

The proof of Lemma 4.2 is postponed to Section 4.7.

4.3 A Poisson branching process

We now define a process $\mathcal{C}$, called a Crump-Mode-Jagers (or CMJ) process. The name Crump-Mode-Jagers applies to a more general class than what is considered here, but we will mainly be talking about the special case described as follows. Fix a constant $\alpha > 0$ and consider a Poisson process $\mathcal{P}_0$ on $[0,1)$. Suppose $\mathcal{P}_0$ has arrivals at time $\tau_{01} < \tau_{02} < \cdots < \tau_{0k}$. The jth arrival gives rise to a Poisson process $\mathcal{P}_{0j}$ on $[\tau_{0j}, \tau_{0j}+1)$, $j = 1, \ldots, k$, independent of all other Poisson processes.

In general, let $s = 0 * **$ be a string of integers and suppose $\mathcal{P}_s$ is a Poisson process on $[\tau_s, \tau_s + 1)$. Then the jth arrival in $\mathcal{P}_s$, at time τ_{sj}, gives rise to a Poisson process $\mathcal{P}_{sj}$ on $[\tau_{sj}, \tau_{sj} + 1)$. Here sj should be interpreted as appending j to the end of the string s. Let $d(\tau)$ be the number of processes alive at time τ, i.e. the number of processes $\mathcal{P}_s$ with $\tau_s \in (\tau - 1, \tau]$, and define $b(\tau)$ to be the number of processes born before τ, i.e. the number of s for which $\tau_s \le \tau$.

For a random variable X and $p, q \in [0, 1]$, say that $X \sim G(p, q)$ if

$$\Pr\{X = k\} = \begin{cases} 1 - q, & k = 0, \\ qp(1-p)^{k-1}, & k \ge 1. \end{cases}$$

Lemma 4.3. *For all $\tau \ge 0$, $d(\tau) \sim G(p(\tau), q(\tau))$.*

The proofs of Lemmas 4.3 and 4.4 are postponed to Section 4.8.

Lemma 4.4. *There exists a constant $\lambda > 0$ such that for $0 \le \tau \le \log_\gamma n$, as $n \to \infty$*

(i) if $\alpha < 1$,

$$Pr\{b(\tau) > \lambda \ln n\} = o\left(\frac{1}{n}\right).$$

(ii) if $\alpha > 1$,

$$Pr\left\{b(\tau) > \lambda n^{1/\eta} \ln n\right\} = o\left(\frac{1}{n}\right)$$

where $\eta = -\ln \gamma / \ln \zeta > 2$.

(iii) If $\alpha \ne 1$ then $d(\tau) \ge \lfloor b(\tau)/(\lambda \log_\gamma^2 n)\rfloor$ for all $0 \le \tau \le \log_\gamma n$ with probability $1 - o(n^{-1})$.

Let $\lambda > 0$ be as provided by Lemma 4.4 and define

$$B(n) = \begin{cases} \lambda \ln n, & \alpha < 1, \\ \lambda n^{1/\eta} \ln n, & \alpha > 1. \end{cases}$$

Given a time $\tau > 0$ we can calculate $b(\tau), d(\tau)$ by the following algorithm, based on the breadth-first-search algorithm. Here S, S' are sets of integer strings. The numbers i, j count the number of times $\mathcal{E}, \mathcal{L}$ have been called, respectively.

0. Let $S = \{0\}$, $S' = \{0\}$ and $\tau_0 = 0$.
1. If S' is empty, stop and output S and $T = \{\tau_s : s \in S\}$. Otherwise choose the smallest $s \in S'$ (ordered lexicographically) and remove it from S'. Let $L_s = 1$ be the lifetime of process $\mathcal{P}_s$.
2. Let $X_{s1}, X_{s2}, \ldots, X_{s(k+1)}$ be independent $\mathrm{Exp}(\alpha)$ variables where $k \ge 0$ is the smallest integer for which $X_{s1} + \cdots + X_{s(k+1)} > L_s$. If $k \ge 1$, set

$$\begin{aligned} \tau_{s1} &= \tau_s + X_{s1}, \\ \tau_{s2} &= \tau_s + X_{s1} + X_{s2}, \\ &\vdots \\ \tau_{sk} &= \tau_s + X_{s1} + X_{s2} + \cdots + X_{sk}. \end{aligned}$$

   Add $s1, s2, \ldots, sk$ to S and S'.

4.4 The degree distribution

This section is devoted to proving the following theorem. Suppose the graph process starts at $H = G_{t_0}$ where $t_0 = o(n^{1/2})$. As in Section 4.2, let $\mathcal{D}$ denote the event that $1_t > \nu_t$ for some $t \geq t_0$.

Let $G^m(p, q)$ denote the distribution of $X = X_1 + X_2 + \cdots + X_m$ where $X_1, \ldots, X_m$ are independent $G(p, q)$ distributed variables. Let $g^m(k; p, q)$ denote the probability mass function of X. Note that we define $d(n, v) = 0$ if v is not in G_n. The functions $p(\tau), q(\tau)$ are defined in Section 4.2.2.

Theorem 4.1. *Let $\omega = o(\log n)$ be such that $t_0 \leq n^{1/2}/\omega$ and $\omega \to \infty$ as $n \to \infty$. Let $v \geq n/\omega$, $\delta = n^{-1/2} \ln n$, and $\tau = \log_\gamma(pn/v)$. There exists a function $\widetilde{q}(\tau) \in [0, 1]$ with $\widetilde{q}(\tau) = q(\tau)$ for all $\tau \notin (-\delta, \delta) \cup (1 - \delta, 1 + \delta)$, such that the degree $d(n, v)$ of v satisfies*

$$Pr\left\{d(n, v) = k \mid \overline{\mathcal{D}}\right\} = g^m(k; p(\tau), \widetilde{q}(\tau)) + O(B(n)n^{-1/2} \ln^2 n), \ k \geq 0.$$

In Section 4.1.1, the idea behind this proof is outlined in the notation of the process G_t. The full proof presented here is based on the master graph Γ and will be rather technical, but the idea is the same. Condition on a feasible and ω-concentrated $\sigma \in \{0, 1\}^{n-t_0}$, see Lemma 4.1. We will be considering the master graph $\Gamma = \Gamma_n^\sigma(H)$ defined in Section 4.2.1. Let E_n be the set of edges in Γ with at least one endpoint in $\{1_n, \ldots, \nu_n\}$, so that G_n is obtained from Γ by removing all edges not in E_n.

Consider the graph $\Gamma_0 \in \mathcal{G}(\emptyset, \emptyset)$ in which all edges $e > m\nu_H$ are free. Fix a vertex $v > n/\omega$, and let $e_\ell = m(v - 1) + \ell, \ell = 1, \ldots, m$ denote the m edges adjacent to v. Suppose an edge $e > mv$ is adjacent to v in Γ. Then e must choose $\phi(e) = (f, j)$ for some edge f which is also adjacent to v. Here j must be 2 if $f \in \{e_1, \ldots, e_m\}$ and 1 otherwise. In words, for an edge to be adjacent to v in Γ but not in Γ_0, it must choose the appropriate endpoint of some other edge that is adjacent to v in Γ.

We will now make this idea more precise. Consider a partially generated graph $\widetilde{\Gamma} \in \mathcal{G}(A, R)$ for some sets A, R. For $(e_0, j_0) \notin R$, we define an operation called *exposing* (e_0, j_0), as a sequence of *reveals* (as defined in Section 4.2.1). Let $Q_0 = \{(e_0, j_0)\}$. For $i \geq 1$ define $Q_i = \{(e, 1) : e \notin A, \phi(e) \in Q_{i-1}\}$. Consider the following algorithm for finding the edges in $\cup_{i \geq 0} Q_i$. The parts labelled Setup are not essential to the running of the algorithm, but are included to emphasize the similarity to the algorithm in Section 4.3, to which it will later be compared.

The algorithm takes as input sets A, R, a partially generated $\widetilde{\Gamma} \in \mathcal{G}(A, R)$ and a half-edge $(e_0, j_0) \notin R$.

0. Let $S = \{0\}, S' = \{0\}$. Let $Q = \{(e_0, j_0)\}$.
1. If S' is empty, stop and output S and Q. Otherwise, let s be the smallest member of S' (in the lexicographical order) and remove s from S'.

   Setup: Let $L'_s = \log_\gamma(f/e_s)$ where f is the largest edge with $e_s \in E_f^\sigma$.
2. Reveal (e_s, j_s) to find $\phi^{-1}(\{e_s, j_s\}) \setminus A$. Label the edges in $\phi^{-1}(\{e_s, j_s\}) \setminus A$ by $e_{s1} < e_{s2} < \cdots < e_{sk}$ (where si denotes string concatenation). Add $(e_{s1}, 1), \ldots, (e_{sk}, 1)$ to Q, and add $s1, s2, \ldots, sk$ to S and S'.

   The partially generated graph is now in $\mathcal{G}(A \cup \{e_{s1}, \ldots, e_{sk}\}, R \cup \{(e_s, j_s)\})$. Set $A \leftarrow A \cup \{e_{s1}, \ldots, e_{sk}\}$ and $R \leftarrow R \cup \{(e_s, j_s)\}$. Go to step 1.

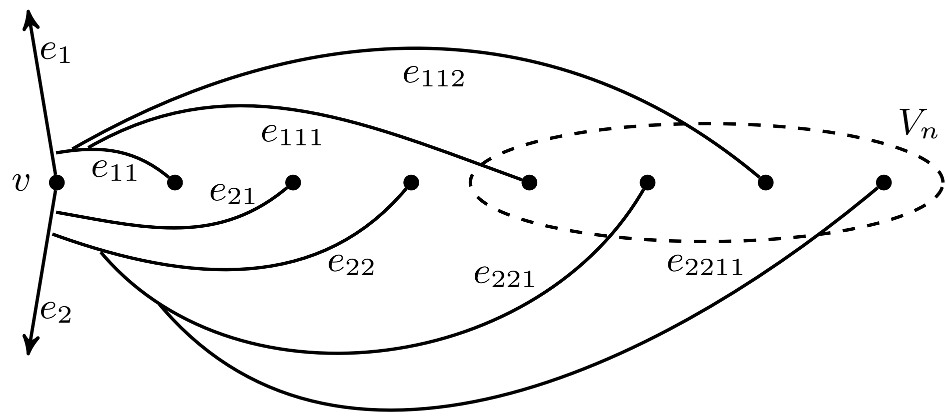

Figure 4.1: One outcome of the expose algorithm for $m = 2$. Here v has degree 9 in Γ and degree 4 in G_n. All edges in the figure are adjacent to v, but are drawn to indicate which half-edge was chosen, e.g. $\phi(e_{221}) = (e_{22}, 1)$. Free edges are drawn as arrows.

Setup: Let $X'_{s1} = \log_\gamma(e_{s1}/e_s)$ and $X'_{s\ell} = \log_\gamma(e_{s\ell}/e_{s(\ell-1)})$ for $\ell = 1, 2, \ldots, k$. Set $X'_{s(k+1)} = \infty$ and $e_{s(k+1)} = \infty$.

With input (e_0, j_0) and $\widetilde{\Gamma} \in \mathcal{G}(A, R)$, let $E((e_0, j_0), \widetilde{\Gamma})$ be the set of edges $e \in E_n$ (i.e. edges in G_n) such that $e = e_s$ for some $s \in S$.

Lemma 4.5. *Let $\omega = o(\log n)$ tend to infinity with n. Suppose either $\alpha < 1$ and $0 < \varepsilon < 1/2$, or $\alpha > 1$ and $0 < \varepsilon < 1/2 - 1/\eta$. Let $\widetilde{\Gamma} \in \mathcal{G}(A, R)$ where $|A|, |R| = O(n^{1/2+\varepsilon} \log^k n)$ for some $k \geq 1$. Let $(e_0, j_0) \notin R$ satisfy $e_0 \geq mn/\omega$, and let $\tau = \log_\gamma(pmn/e_0)$. There exists a $\delta = O(n^{-1/2} \ln n)$ and a function $\widetilde{q}(\tau) \in [0, 1]$ such that*

$$Pr\Big\{|E((e_0, j_0), \widetilde{\Gamma})| = k\Big\} = \begin{cases} 1 - \widetilde{q}(\tau) + O(B(n) n^{-1/2+\varepsilon} \ln^2 n), & k = 0 \\ \widetilde{q}(\tau) p(\tau)(1 - p(\tau))^{k-1} + O(B(n) n^{-1/2+\varepsilon} \ln^2 n), & k \geq 1. \end{cases}$$

where $\widetilde{q}(\tau) = q(\tau)$ for all $\tau \notin (-\delta, \delta) \cup (1 - \delta, 1 + \delta)$.

Before proving the lemma, we show how it is used to finish the proof of Theorem 4.1. Consider the graph $\Gamma_0 \in \mathcal{G}(\emptyset, \emptyset)$ in which no assignments or reveals have been made. We expose $(e_1, 2)$ to find that $|E((e_1, 2), \Gamma_0)|$ is asymptotically $G(p(\tau), \widetilde{q}(\tau))$ distributed. Exposing $(e_1, 2)$ gives a partially generated graph $\Gamma_1 \in \mathcal{G}(A_1, R_1)$ where A_1 is the set of edges assigned while exposing $(e_1, 2)$ and R_1 consists of $(e_1, 2)$ and $(e, 1)$ for all $e \in A_1$. By Lemma 4.4 we have $|A_1|, |R_1| = O(B(n)) = o(n^{1/2})$ whp. Apply Lemma 4.5 to Γ_1 to find that $|E((e_2, 2), \Gamma_1)|$ is asymptotically $G(p(\tau), \widetilde{q}(\tau))$ distributed, and consider $\Gamma_2 \in \mathcal{G}(A_2, R_2)$, where $A_2 \setminus A_1$ and $R_2 \setminus R_1$ consist of the edges assigned and revealed when exposing $(e_2, 2)$. Repeating this m times keeps the sets A_i, R_i of size $o(n^{1/2})$, and we find that $|E((e_i, 2), \Gamma_{i-1})|$ is asymptotically $G(p(\tau), \widetilde{q}(\tau))$ distributed for $i = 1, 2, \ldots, m$. Then

$$d(n, v) = \sum_{i=1}^{m} |E((e_i, 2), \Gamma_{i-1})|$$

and the theorem follows.

The above assumes that each vertex makes its m random choices with replacement. In the process of determining $d(n, v)$, $O(B(n))$ edges are revealed. The probability for any edge e to be adjacent to v is $O(B(n)/n)$, and it follows that the probability that two edges e_1, e_2 with $\lceil e_1/m \rceil = \lceil e_2/m \rceil$ are adjacent to v is $O(B(n)^2/n) = o(B(n)n^{-1/2})$. This shows that $d(n, v)$ has the same asymptotic distribution when choices are made with or without replacement.

4.4.1 Proof of Lemma 4.5

For $i \geq 1$ write $X'_{si} = \log_\gamma(e_{si}/e_{s(i-1)})$, where we say $e_{s0} = e_s$. We will show that the collection $\{X'_s : s \in S\}$ can be coupled to a collection $\{X_s : s \in S\}$ of independent $\text{Exp}(\alpha)$ variables in such a way that $X'_s = X_s + O(n^{-1/2+\varepsilon} \ln n)$ for all s with high probability. The lemma will then follow from arguing that a CMJ process on $[0, \tau]$ with $\tau \leq \log_\gamma n$ is robust with high probability, in the sense that changing interarrival times by $O(n^{-1/2+\varepsilon} \ln n)$ does not change the value of $d(\tau)$.

The set of edges e with $e_0 \in E_e^\sigma$ is $\{e_0 + i, e_0 + i + 1, \ldots, e'_0\}$ for some $i \in [m]$ and some e'_0. Since σ is ω-concentrated, there exists a constant $C > 0$ such that for all edges $e \geq mn/\omega$, the largest edge that may choose e is e' where $e(\gamma - Cn^{-1/2} \ln^2 n) < e' < e(\gamma + Cn^{-1/2} \ln^2 n)$. Fix such a C and let $\delta_1 = Cn^{-1/2} \ln^2 n$, and let $\delta = O(n^{-1/2} \ln^2 n)$ be such that $1 - \delta < \log_\gamma(\gamma + \delta_1) < 1 + \delta$. Let $\tau = \log_\gamma(pmn/e_0)$. If $\tau \leq -\delta$ then $e_0 \notin \Gamma$, if $\delta \leq \tau \leq 1 - \delta$ then $e_0 \in E_n$, and if $\tau \geq 1 + \delta$ then $e_0 \in \Gamma$ but $e_0 \notin E_n$. We will be assuming that $\tau \notin (-\delta, \delta) \cup (1 - \delta, 1 + \delta)$, and leave the cases $\tau \in (-\delta, \delta)$ and $\tau \in (1 - \delta, 1 + \delta)$ until the end of the proof.

Now, consider the first edge e_{01} that chooses (e_0, j_0), taken to be ∞ if no edge chooses (e_0, j_0). Since σ is ω-concentrated and $|R| = O(n^{1/2+\varepsilon} \log^k n)$ for some $k \geq 1$, we have $|\widetilde{\Omega}(e)| = 2\mu e/pm + O(n^{1/2} \ln n) - O(n^{1/2+\varepsilon} \log^k n) = 2\mu e/pm + O(n^{1/2+\varepsilon} \log^k n)$ for all $e > mn/\omega$. Since $e_0 > mn/\omega$, if $(e_0, j_0) \in \widetilde{\Omega}(e)$ then

$$\Pr\{e \text{ chooses } (e_0, j_0)\} = \frac{pm}{2\mu e} + O(n^{-3/2+\varepsilon} \ln^k n),$$

independently of the random choice of all other edges. Let $i \in [m]$ be the smallest number for which $e_0 \in E^\sigma_{e_0+i}$, and suppose $y > 1$ is such that $e_0 \in E^\sigma_{\lfloor ye_0 \rfloor}$. Then if $x = \log_\gamma y$,

$$\begin{aligned}
\Pr\{e_{01} > ye_0\} &= \prod_{\substack{e=e_0+i \\ e \notin A}}^{\lfloor ye_0 \rfloor} \left(1 - \frac{pm}{2\mu e} + O(n^{-3/2+\varepsilon} \ln^k n)\right) \\
&= \exp\left\{-\frac{pm}{2\mu} \sum_{\substack{e=e_0+i \\ e \notin A}}^{\lfloor ye_0 \rfloor} \left(\frac{1}{e} + O(n^{-3/2+\varepsilon} \ln^k n)\right)\right\} \\
&= \exp\left\{-\alpha x \left(1 + O\left(\frac{|A|}{n} + n^{-1/2+\varepsilon} \ln^k n\right)\right)\right\} \\
&= \exp\left\{-\alpha x \left(1 + O(n^{-1/2+\varepsilon} \ln^k n)\right)\right\}. \qquad (4.1)
\end{aligned}$$

This suggests that $X'_{01} = \log_\gamma(e_{01}/e_0)$ is approximately exponentially distributed, in the range of y for which $e_0 \in E^\sigma_{\lfloor ye_0 \rfloor}$. We will couple X'_{01} to an exponentially distributed random variable, and the coupling technique will depend on whether or not $e_0 \in E_n$. Define $\tau = \log_\gamma(pmn/e_0)$. As noted above, $\tau > 1+\delta$ implies $e_0 \in \Gamma$ and $e_0 \notin E_n$, while $\delta < \tau < 1-\delta$ implies $e_0 \in E_n$.

Case 1, $e_0 \notin E_n$**.**
Suppose $\tau > 1+\delta$, so that $e_0 \notin E_n$ under our choice of σ. By choice of δ_1, there exists a $y \in (\gamma-\delta_1, \gamma+\delta_1)$ such that ye_0 is the largest edge for which $e_0 \in E^\sigma_{ye_0}$. Applying (4.1) with this y, we have $\Pr\{e_{01} = \infty\} = \exp\left\{-\alpha(1+O(n^{-1/2}\ln^k n))\right\}$, since $\log_\gamma(\gamma+\delta) = 1 + O(n^{-1/2}\ln^2 n)$. For $L > 0$ we define a distribution $\mathrm{Exp}(\alpha, L)$ by saying that $X \sim \mathrm{Exp}(\alpha, L)$ if $\Pr\{X > x\} = e^{-\alpha x}$ for $0 < x < L$ and $\Pr\{X = \infty\} = e^{-\alpha L}$. We will couple X'_{01} to an $\mathrm{Exp}(\alpha, 1)$ variable, as described below.

Condition on e_{01} and consider e_{02}, the second edge that chooses e_0. Repeating (4.1) shows that $\Pr\{e_{02} > ye_{01}\} = \exp\{-\alpha x(1+O(n^{-1/2+\varepsilon}\ln^k n))\}$ where $x = \log_\gamma y$, for all y such that $e_0 \in E^\sigma_{ye_{01}}$. The largest such y is $\gamma e_0/e_{01} + O(n^{-1/2}\ln^2 n)$, and

$$\log_\gamma\left(\frac{\gamma e_0}{e_{01}} + O(n^{-1/2}\ln^2 n)\right) = 1 - X'_{01} + O(n^{-1/2}\ln^2 n).$$

We will couple X'_{02} to an $\mathrm{Exp}(\alpha, 1 - X'_{01})$ variable. In general, X'_{0i} will be coupled to an $\mathrm{Exp}(\alpha, 1 - X'_{01} - \cdots - X'_{0(i-1)})$ variable, conditioning on $X'_{01}, \ldots, X'_{0(i-1)}$.

Case 2, $e_0 \in E_n$**.**
In the case $\delta < \tau < 1-\delta$, where $e_0 \in E_n$, we instead couple X'_{01} to an $\mathrm{Exp}(\alpha, \tau)$ variable, since the largest edge that may choose (e_0, j_0) is $m\nu_n = pmn + O(n^{-1/2}\ln n)$, the largest edge in Γ. We will couple X'_{0i} to an $\mathrm{Exp}(\alpha, \tau - X'_{01} - \cdots - X'_{0(i-1)})$ variable.

Coupling the variables: Let $\tau' = \min\{1, \tau\}$. In terms of $\mathrm{Exp}(\alpha, L)$ variables, we can define a Poisson process on $[0, \tau']$ as follows. Let $X_{01} \sim \mathrm{Exp}(\alpha, \tau')$. Conditioning on $X_{01} = x_{01} < 1$ we define $X_{02} \sim \mathrm{Exp}(\alpha, \tau' - x_{01})$. In general let $X_{0i} \sim \mathrm{Exp}(\alpha, \tau' - x_{01} - \cdots - x_{0(i-1)})$ until $X_{0k} = \infty$. Then $X_{01}, \ldots, X_{0(k-1)}$ are the interarrival times for a Poisson process of rate α on $[0,1]$.

We will now describe the coupling explicitly. Let $U_{01}, U_{02}, \ldots$ be a sequence of independent uniform $[0,1]$ variables. The variable $X_{01} \sim \mathrm{Exp}(\alpha, 1)$ is given by

$$X_{01} = \begin{cases} -\alpha^{-1}\ln U_{01}, & e^{-\alpha} < U_{01} < 1, \\ \infty, & 0 < U_{01} < e^{-\alpha}. \end{cases}$$

and for $i \geq 1$, conditioning on $X_{01} = x_{01}, \ldots, X_{0i} = x_{0i}$ where $x_{01} + \cdots + x_{0i} < 1$,

$$X_{0(i+1)} = \begin{cases} -\alpha^{-1}\ln U_{0(i+1)}, & e^{-\alpha(1-x_{01}-\cdots-x_{0i})} < U_{0(i+1)} < 1, \\ \infty, & 0 < U_{0(i+1)} < e^{-\alpha(1-x_{01}-\cdots-x_{0i})}. \end{cases}$$

Define $X'_{01} = \min\{\log_\gamma y : \Pr\{e_{01} > ye_0\} \leq U_{01}\}$, taken to be ∞ if the set is empty. Recall that $\delta_1 = O(n^{-1/2}\ln^2 n)$ is such that $1 - \delta_1 < \log_\gamma(\gamma+\delta) < 1 + \delta_1$. Then by (4.1) and the choice of δ,

$$\text{if } \; U_{01} > e^{-\alpha(1-\delta_1)} \; \text{ then } \; X'_{01} = \frac{-1}{\alpha + O(n^{-1/2+\varepsilon}\ln^k n)}\ln U_{01} = X_{01} + O(n^{-1/2+\varepsilon}\ln^k n),$$

and if $U_{01} < \exp\{-\alpha(1+\delta_1)\}$ then $X'_{01} = \infty$. Say that this coupling of X_{01}, X'_{01} is *good* if either X_{01}, X'_{01} are both infinite, or $X'_{01} = X_{01} + O(n^{-1/2+\varepsilon}\ln^k n)$, and *bad* otherwise. The above shows that

$$\begin{aligned} \Pr\{\text{the coupling of } X_{01}, X'_{01} \text{ is bad}\} &= \Pr\left\{e^{-\alpha(1+\delta_1)} < U_{01} < e^{-\alpha(1-\delta_1)}\right\} \\ &= O(n^{-1/2}\ln^2 n). \end{aligned}$$

Suppose $i > 1$ and condition on the couplings of X_{0j}, X'_{0j} being good with $X_{0j}, X'_{0j} < \infty$ for all $1 \le j < i$. Define $X'_{0i} = \min\left\{\log_\gamma y : \Pr\left\{e_{0i} > y e_{0(i-1)} \middle| X'_{01}, \ldots, X'_{0(i-1)}\right\} \le U_{0i}\right\}$. We repeat the above argument to show that the coupling of X'_{0i}, X_{0i}, conditioning on previous couplings, is bad with probability $O(n^{-1/2} \ln^2 n)$.

Let $\mathcal{C}_0$ be the event that the coupling of the X_{0i} making up the Poisson process $\mathcal{P}_0$ is good for all i. Since the process has $O(\log n)$ arrivals with probability $1 - o(n^{-1})$, we have $\Pr\{\mathcal{C}_0\} = 1 - O(n^{-1/2+\varepsilon} \ln^3 n)$. After revealing e_0, the partially generated graph $\widetilde{\Gamma}$ is in $\mathcal{G}(A', R')$, where $|A' \setminus A| = O(\log n)$ whp and $|R' \setminus R| = 1$. Thus, the coupling argument can be applied to $O(n^{1/2+\varepsilon})$ Poisson processes with $|A|, |R| = O(n^{1/2+\varepsilon} \ln^k n)$ being maintained.

Let $S'(t)$ be the state of the set S' after Step 2 of the algorithm has been executed t times, and let $S'_c(t)$ be the corresponding set in the CMJ generating algorithm of Section 4.3. We just showed that $S'(1) = S'_c(1)$ with probability $1 - O(n^{-1/2+\varepsilon} \ln^2 n)$. For any process $\mathcal{P}_s$ that appears, we apply a coupling using the technique above, and we have $S'(t) = S'_c(t)$ for all $1 \le t < B(n)$ with probability $1 - O(B(n) n^{-1/2+\varepsilon} \ln^2 n) = 1 - o(1)$. We also have $S'_c(B(n)) = \emptyset$ with probability $1 - o(n^{-1})$, by Lemma 4.4, so

$$\Pr\left\{S'(t) \neq S'_c(t) \text{ for some } t \ge 1\right\} = o(1).$$

Condition on the two algorithms producing the same set S of strings. For $s = 0s_1 \ldots s_j \in S$ we have

$$\tau_s = \sum_{i=1}^{s_1} X_{0i} + \sum_{i=1}^{s_2} X_{0s_1 i} + \cdots + \sum_{i=1}^{s_j} X_{0s_1 \ldots s_{j-1} i},$$

and the same identity holds with τ_s, X_r replaced by τ'_s, X'_r. If $s = 0s_1 \ldots s_j$ let $|s| = j$ be the *generation* of s. With probability $1 - o(n^{-1})$, each Poisson process has $O(\log n)$ arrivals, so each $s_i = O(\log n)$. Thus τ_s is a sum of $O(|s| \log n)$ variables X_r, and if all couplings are good then $\tau'_s = \tau_s + O(|s| n^{-1/2} \ln^{k+1} n)$ for all $s \in S$. We need to bound $|s|$.

Claim: Consider a CMJ process with rate $\alpha > 0$ and lifetime 1. Let $0 \le \tau \le \log_\gamma n$ and $S(\tau) = \max\{|s| : \tau_s \le \tau\}$. Then $\Pr\left\{S(\tau) > \log_\gamma^2 n\right\} = o(n^{-1})$.

Proof of claim: Let $P_k(\tau)$ denote the number of processes $\mathcal{P}_s$ with $|s| = k$ and $\tau_s < \tau$. Condition on $\mathcal{P}_0$ having arrivals at time $x_1, \ldots, x_\ell$. Then $\mathcal{C}$ can be seen as $\mathcal{P}_0$ together with ℓ independent CMJ processes $\mathcal{C}^1, \ldots, \mathcal{C}^\ell$ on $[x_1, \tau], \ldots, [x_\ell, \tau]$ respectively. Then

$$P_k(\tau) = \sum_{j=1}^{\ell} P^j_{k-1}(\tau - x_j)$$

where $P^j_{k-1}(\tau - x_j)$ counts the number of $(k-1)$th generation processes started before $\tau - x_j$ in $\mathcal{C}^j$. Let U denote a uniform $[0,1]$ random variable. Removing the conditioning and taking expectations, we have

$$\mathbf{E}\left[P_k(\tau)\right] = \sum_{\ell \ge 0} \frac{e^{-\alpha}\alpha^\ell}{\ell!} \sum_{j=1}^{\ell} \mathbf{E}\left[P^j_{k-1}(\tau - U)\right] = \mathbf{E}\left[P_{k-1}(\tau - U)\right] \sum_{\ell \ge 1} \frac{e^{-\alpha}\alpha^\ell}{(\ell-1)!} = \alpha \mathbf{E}\left[P_{k-1}(\tau - U)\right].$$

Here we use the fact that if we condition on a Poisson process on $[0,1]$ having ℓ arrivals, the arrival times are independently uniformly distributed. Note that $P_k(\tau) = 0$ for all $\tau < 0$.

We show by induction over k that $P_k(\tau) \le (\alpha\tau)^k/k!$ for all integers $k \ge 0$ and all $\tau \ge 0$. For the base case we have $P_0(\tau) = 1$ for $\tau \ge 0$. If $P_k(\tau) \le (\alpha\tau)^k/k!$ for all $\tau \ge 0$ then if $\tau \ge 0$,

$$\begin{aligned}\mathbf{E}\left[P_{k+1}(\tau - U)\right] &= \alpha\mathbf{E}\left[\mathbf{E}\left[P_k(\tau - U) \mid U\right]\right] \\ &\le \alpha\mathbf{E}\left[\frac{\alpha^k(\tau-U)^k}{k!}\right] = \frac{\alpha^{k+1}}{(k+1)!}(\tau^{k+1} - \max\{0, \tau-1\}^{k+1}) \le \frac{(\alpha\tau)^{k+1}}{(k+1)!},\end{aligned}$$

where we use the fact that $P_k(\tau) = 0$ for $\tau < 0$, and the induction is complete.

Let $k = \log_\gamma^2 n$. Then by Markov's inequality and the bound $k! \ge (k/e)^k$,

$$\Pr\left\{\exists s : \tau_s < \tau \text{ and } |s| = k\right\} \le \mathbf{E}\left[P_k(\tau)\right] \le \left(\frac{e\alpha\tau}{k}\right)^k \le \left(\frac{e\alpha}{\log_\gamma n}\right)^{\log_\gamma n} = O(n^{-C})$$

for any $C > 0$. Since $P_{k'}(\tau) \le P_k(\tau)$ for any $k' \ge k$, the claim follows.
End of proof of claim.

We have shown that with high probability, the graph algorithm produces a set S and a set $\{\tau_s' : s \in S\}$ that matches the set $\{\tau_s : s \in S\}$ of a CMJ process in the sense that $\tau_s' = \tau_s + O(n^{-1/2+\varepsilon} \ln^{k+3} n)$ for all s. Since $d(\tau)$ counts the number of τ_s in the interval $(\tau - 1, \tau)$, we can finish the proof by arguing that

$$\{s \in S : \tau - 1 < \tau_s < \tau\} = \{s \in S : e_s \in E_n\}. \tag{4.2}$$

Since σ is ω-concentrated, every edge $e \in E_n$ satisfies $\log_\gamma(e/e_0) \in (\tau - 1 - O(n^{-1/2} \ln n), \tau + O(n^{-1/2} \ln n))$ where $\tau = \log_\gamma(pmn/e_0)$. Condition on $\tau_s' = \tau_s + O(n^{-1/2+\varepsilon} \ln^{k+3} n)$ for all $s \in S$. If (4.2) is false, there must exist some $s \in S$ such that either $\tau_s = \tau - 1 + O(n^{-1/2+\varepsilon} \ln^{k+3} n)$ or $\tau_s = \tau + O(n^{-1/2+\varepsilon} \ln^{k+3} n)$. The probability of this is $O(B(n)n^{-1/2} \ln^{k+3} n)$, since a CMJ process with at most $B(n)$ active processes locally behaves like a Poisson process with rate at most $\alpha B(n)$. This finishes the proof of the lemma for $\tau \notin (-\delta, \delta) \cup (1 - \delta, 1 + \delta)$.

If $\tau \in (-\delta, \delta) \cup (1 - \delta, 1 + \delta)$, then (4.2) is false with some significant probability, since one set may contain $s = 0$ while the other one does not. The function $\widetilde{q}(\tau)$ accounts for this event.

4.5 The degree sequence

For $k \ge 0$ let $X_k(n)$ denote the number of vertices of degree k in G_n. In this section we prove the following. Recall $\eta = -\ln\gamma/\ln\zeta > 2$, defined when $\alpha > 1$, see Section 4.2.2. Recall that $\mathcal{D}$ denotes the event that at some point, the graph process contains no edges. The probability of $\mathcal{D}$ depends on the initial graph $H = G_{t_0}$, see Lemma 4.1.

Theorem 4.2. *Condition on $\overline{\mathcal{D}}$. There exists a sequence $\{x_k : k \ge 0\}$ such that*

(i) if $\alpha < 1$ then $x_k = \alpha^{k(1+o_k(1))}$ and if $\alpha > 1$ then there exist constants $a, b > 0$ such that $x_k = ak^{-\eta-1} + O_k(k^{-\eta-2}\log^b k)$, and

(ii) for any fixed $k \ge 0$, $X_k(n) = x_k n(1 + o_n(1))$ with high probability as $n \to \infty$.

Proof. Fix $k \ge 0$. We begin by showing that $X_k(n) = (1 + o_n(1))\mathbf{E}\left[X_k(n)\right]$ whp. We will use Azuma's inequality in the general exposure martingale setting in [5, Section 7.4]. To do this, fix a feasible σ and consider the master graph $\Gamma = \Gamma_n^\sigma(H)$ for a fixed starting graph H (see Section 4.2). Let Γ_0 be the unexplored graph, as defined in Section 4.2.1, and define a sequence $\Gamma_1, \Gamma_2, \ldots, \Gamma_M = \Gamma$ of partially generated graphs. Here Γ_i is obtained from Γ_{i-1} by letting edge $m\nu_H + i$ make its random choice. Consider the edge exposure martingale $Y_i^\sigma = \mathbf{E}[X_k(n) \mid \Gamma_i]$. If E_n denotes the edge set of G_n, then $X_k(n)$ is given as a function of Γ by counting the number of vertices which are incident to exactly k edges of E_n. This martingale satisfies the Lipschitz inequality $|Y_i^\sigma - Y_{i-1}^\sigma| \leq 3$, since the degrees of at most 3 vertices are affected by changing the choice of one edge (see e.g. Theorem 7.4.1 of [5]). By Azuma's inequality, conditioning on σ we have $|X_k(n) - \mathbf{E}[X_k(n)]| < n^{1/2} \ln n$ with probability $1 - e^{-\Omega(\ln^2 n)}$, noting that $M = m(\nu_n - \nu_H)$ is of order n. We will show that $\mathbf{E}[X_k(n)]/n$ has essentially the same limit for all feasible and ω-concentrated σ, setting $\omega = \log\log n$, and the result will follow since σ is ω-concentrated with probability $1 - o(n^{-1})$ (Lemma 4.1) and $X_k(n) \leq n$. Fix a feasible and ω-concentrated σ for the remainder of the proof.

Recall that $G^m(p,q)$ denotes the distribution of the sum of m independent $G(p,q)$ variables. If $X \sim G^m(p,q)$ and $k \geq m$ then

$$\begin{aligned}\Pr\{X = k\} &= \sum_{\ell=1}^{m} \binom{m}{\ell} \sum_{\substack{k_1+\cdots+k_\ell=k \\ k_1,\ldots,k_\ell>0}} (1-q)^{m-\ell} \prod_{i=1}^{\ell} qp(1-p)^{k_i-1} \\ &= \sum_{\ell=1}^{m} \binom{m}{\ell}\binom{k-1}{\ell-1}(1-q)^{m-\ell} q^\ell p^\ell (1-p)^{k-\ell}. \end{aligned} \tag{4.3}$$

Here ℓ represents the number of nonzero terms in the sum $X = X_1 + \cdots + X_m$, and $\binom{k-1}{\ell-1}$ is the number of ways to write k as a sum of ℓ positive integers. By linearity of expectation,

$$\mathbf{E}[X_k(n)] = \sum_{v=1}^{n} \Pr\{d(n,v) = k\}.$$

Let $\omega = \log\log n$. By Theorem 4.1 we have

$$\sum_{v=1}^{n} \Pr\{d(n,v) = k\} = O\left(\frac{n}{\omega}\right) + \sum_{v=n/\omega}^{n} \left(\Pr\{G^m(p(\tau), \widetilde{q}(\tau)) = k\} + O\left(\frac{B(n)\ln^3 n}{n^{1/2}}\right)\right).$$

Summing the $O(n^{-1/2} B(n) \ln^3 n)$ terms gives a cumulative error of $O(n^{1/2} B(n) \ln^3 n) = o(n)$, since either $B(n) = O(\log n)$ or $B(n) = O(n^{1/\eta} \ln n)$ (see Section 4.3) and $\eta > 2$ (see Lemma 4.2 (ii)). So if $k \geq m$ and $\tau_v = \log_\gamma(pn/v)$, by (4.3),

$$\mathbf{E}[X_k(n)] = O\left(\frac{n}{\omega}\right) + \sum_{v=n/\omega}^{n} \sum_{\ell=1}^{m} \binom{m}{\ell}\binom{k-1}{\ell-1}(1-\widetilde{q}(\tau_v))^{m-\ell} \widetilde{q}(\tau_v)^\ell p(\tau_v)^\ell (1-p(\tau_v))^{k-\ell}, \tag{4.4}$$

where $\widetilde{q}(\tau) \in (0,1)$ and $\widetilde{q}(\tau) = q(\tau)$ outside $(-\delta, \delta) \cup (1-\delta, 1+\delta)$ for some $\delta = O(n^{-1/2} \ln n)$. For $n/\omega \leq v \leq n$ we have $\log_\gamma p \leq \tau_v \leq \log_\gamma(p\omega)$ (note that $\log_\gamma p < 0$), and for any τ in the interval, the number of v for which $\tau \leq \tau_v \leq \tau + \varepsilon$ is $pn\varepsilon \ln(\gamma)\gamma^{-\tau} + O(\varepsilon^2)$. Viewing the sum as a Riemann sum, we have

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\sum_{v=n/\omega}^{n}(1-\widetilde{q}(\tau_v))^{m-\ell}\widetilde{q}(\tau_v)^{\ell}p(\tau_v)^{\ell}(1-p(\tau_v))^{k-\ell}\\
=&p\ln\gamma\int_{\log_\gamma p}^{\infty}\frac{(1-\widetilde{q}(\tau))^{m-\ell}\widetilde{q}(\tau)^{\ell}p(\tau)^{\ell}(1-p(\tau))^{k-\ell}}{\gamma^{\tau}}d\tau\\
=&O(n^{-1/2}\ln n)+p\ln\gamma\int_0^{\infty}\frac{(1-q(\tau))^{m-\ell}q(\tau)^{\ell}p(\tau)^{\ell}(1-p(\tau))^{k-\ell}}{\gamma^{\tau}}d\tau. \qquad (4.5)
\end{aligned}$$

The last identity comes from (i) the fact that $\widetilde{q}(\tau)=q(\tau)$ outside a set of total length $O(n^{-1/2}\ln n)$, (ii) the fact that the integral converges since the integrand is dominated by $\gamma^{-\tau}$ where $\gamma>1$, and (iii) the fact that $q(\tau)=0$ for $\tau<0$.

Plugging (4.5) into (4.4) we have

$$\lim_{n\to\infty}\frac{\mathbf{E}\left[X_k(n)\right]}{n}=p\ln\gamma\sum_{\ell=1}^{m}\binom{m}{\ell}\binom{k-1}{\ell-1}\int_0^{\infty}\frac{(1-q(\tau))^{m-\ell}q(\tau)^{\ell}p(\tau)^{\ell}(1-p(\tau))^{k-\ell}}{\gamma^{\tau}}d\tau. \qquad (4.6)$$

Let

$$f_\ell(\tau)=\frac{(1-q(\tau))^{m-\ell}q(\tau)^{\ell}p(\tau)^{\ell}(1-p(\tau))^{k-\ell}}{\gamma^{\tau}}.$$

Our aim is to calculate $\int_0^\infty f_\ell(\tau)d\tau$.

Case 1: $\alpha>1$.
By Lemma 4.2 (vi) we have $p(\tau)\geq\lambda_3\zeta^\tau$ for all $\tau\geq 0$, where $\lambda_3>0$. Let $\psi(k)=-\log_\zeta((k-\ell)/(C\ln k))$ for some constant $C>0$, noting that $\psi(k)\to\infty$ when $k\to\infty$. Making C large enough,

$$\int_0^{\psi(k)}f_\ell(\tau)\leq\psi(k)(1-\lambda_3\zeta^{\psi(k)})^{k-\ell}\leq\psi(k)e^{-\lambda_3 C\ln k}=O(k^{-\eta-2}). \qquad (4.7)$$

Here we used the fact that $f_\ell(\tau)\leq(1-p(\tau))^{k-\ell}$.

Again by Lemma 4.2 (vi) we have $p(\tau)=\lambda_3\zeta^\tau+O(\zeta^{2\tau})$ and $q(\tau)=1-\zeta+O(\zeta^\tau)$. Suppose $\tau\geq\psi(k)$. Then $k\zeta^{2\tau}=o_k(1)$ and

$$f_\ell(\tau)=\frac{\zeta^{m-\ell}(1-\zeta)^{\ell}(\lambda_3\zeta^\tau)^{\ell}(1-\lambda_3\zeta^\tau)^{k-\ell}}{\gamma^\tau}\left(1+O(m\zeta^\tau)+O_k(k\zeta^{2\tau})\right).$$

Indeed, each of the m factors involving $q(\tau)$ contributes an error factor of $1+O(\zeta^\tau)$ and each of the k factors involving $p(\tau)$ contributes an error factor of $1+O(\zeta^{2\tau})$. We have $m\zeta^\tau=O(\ln k/k)$ and $k\zeta^{2\tau}=O(\ln^2 k/k)$, so

$$f_\ell(\tau)=\frac{\zeta^{\tau\ell}(1-\lambda_3\zeta^\tau)^{k-\ell}}{\gamma^\tau}\left(\lambda_3^{\ell}\zeta^{m-\ell}(1-\zeta)^{\ell}+O\left(\frac{\ln^2 k}{k}\right)\right). \qquad (4.8)$$

Note that λ_3,ζ,m and ℓ are independent of k and τ.

Claim: If $\alpha>1$ there exists a constant c_ℓ such that

$$\int_{\psi(k)}^{\infty}\frac{\zeta^{\tau\ell}(1-\lambda_3\zeta^\tau)^{k-\ell}}{\gamma^\tau}=c_\ell k^{-\eta-\ell}+O(k^{-\eta-\ell-1}).$$

It will follow from the claim and (4.8) that for some constant c'_ℓ,

$$\int_{\psi(k)}^{\infty} f_\ell(\tau)d\tau = c'_\ell k^{-\eta-\ell}\left(1+O\left(\frac{\ln^2 k}{k}\right)\right). \tag{4.9}$$

Proof of claim: We make the integral substitution $u = \lambda_3\zeta^\tau$, noting that $\tau = \log_\zeta(u/\lambda_3)$ so (recalling that $\eta = -\ln\gamma/\ln\zeta$, see Section 4.2.2)

$$\gamma^{-\tau} = \exp\left\{-\ln\gamma\frac{\ln(u/\lambda_3)}{\ln\zeta}\right\} = \left(\frac{u}{\lambda_3}\right)^\eta.$$

This implies that the integral equals (up to a multiplicative constant)

$$\begin{aligned}\int_0^{\frac{C\lambda_3\ln k}{k-\ell}} u^{\eta+\ell-1}(1-u)^{k-\ell}du &= \int_0^1 u^{\eta+\ell-1}(1-u)^{k-\ell}du - \int_{\frac{C\lambda_3\ln k}{k-\ell}}^1 u^{\eta+\ell-1}(1-u)^{k-\ell}du\\ &= B(\eta+\ell, k-\ell+1) + O\left(k^{-C\lambda_3}\right)\end{aligned}$$

where $B(x,y) = \int_0^1 u^{x-1}(1-u)^{y-1}du$ denotes the Beta function. Here the $O(k^{-C\lambda_3})$ term comes from bounding $u^{\eta+\ell-1} \le 1$ and $1-u \le e^{-C\lambda_3\ln k/(k-\ell)}$. Taking C to be large enough makes the error $O(k^{-\eta-m-1})$ (recall that $\ell \le m$). As $k\to\infty$, Stirling's formula provides an asymptotic expression for $B(\eta, k+1)$:

$$B(\eta+\ell, k-\ell+1) = \Gamma(\eta+\ell)k^{-\eta-\ell} + O(k^{-\eta-\ell-1}),$$

where Γ denotes the Gamma function. **End of proof of claim.**

We finish the proof for $\alpha > 1$ by noting that by Stirling's formula, for some constant s_ℓ

$$\binom{k-1}{\ell-1} = s_\ell k^{\ell-1} + O(k^{\ell-2}) \tag{4.10}$$

Plugging (4.7), (4.9) and (4.10) into (4.6) shows that

$$\begin{aligned}\frac{\mathbf{E}\left[X_k(n)\right]}{n} &\to p\ln\gamma\sum_{\ell=1}^m\binom{m}{\ell}\binom{k-1}{\ell-1}\int_0^\infty f_\ell(\tau)dt\\ &= p\ln\gamma\sum_{\ell=1}^m\binom{m}{\ell}(s_\ell k^{\ell-1}+O(k^{\ell-2}))(c'_\ell k^{-\eta-\ell}+O(k^{-\eta-\ell-1}\ln^2 k))\\ &= \left(p\ln\gamma\sum_{\ell=1}^m\binom{m}{\ell}s_\ell c'_\ell\right)k^{-\eta-1}+O(k^{-\eta-2}\ln^{\eta+m+3}k).\end{aligned}$$

Here the expression in brackets depends only on p, m, and this is the constant a in the statement of the theorem.

Case 2: $\alpha < 1$.
In this case we need not be as careful. By Lemma 4.2 (v) we have $1-p(\tau) = \alpha - \lambda_1/\zeta^\tau + O(\zeta^{2\tau})$ where $0 < \lambda_1 < \alpha$ and $\zeta > 1$, so we can write

$$f_\ell(\tau) = \alpha^{k-\ell}\frac{(1-q(\tau))^{m-\ell}q(\tau)^\ell p(\tau)^\ell\left(\frac{1-p(\tau)}{\alpha}\right)^{k-\ell}}{\gamma^\tau}$$

and the calculation of $\alpha^{-(k-\ell)}\int_0^\infty f_\ell(\tau)d\tau$ proceeds much like the $\alpha > 1$ case. We find that

$$\int_0^\infty f_\ell(\tau)d\tau = \alpha^{k-\ell}O(k^C) = \alpha^{k(1+o_k(1))}$$

for some constant $C > 0$. Summing over $\ell = 1, \dots, m$ does not affect this expression. □

4.6 The largest component

This section deals with connectivity properties of G_n. Note that G_n is disconnected whp since one can show that the number of isolated vertices is $\Omega(n)$ whp. It is also the case that the set of non-isolated vertices is disconnected whp, since the probability that a vertex v shares a component only with its m older neighbors is a nonzero constant, as can be seen by methods similar to those used in the proof of Lemma 4.6 below.

In the following theorem, the size of a component refers to the number of vertices in the component. Recall that $B(n) = \lambda \ln n$ if $\alpha < 1$ and $B(n) = \lambda n^{1/\eta} \ln n$ if $\alpha > 1$, for a constant $\lambda > 0$. Recall also that $\mathcal{D}$ denotes the event that the graph process contains zero edges at some point (see Lemma 4.1).

Note that the number of vertices in G_n is $pn + O(n^{1/2}\ln n)$ whp, so when $m \geq 2$ and $\alpha > 1$, Theorem 4.3 states that whp the number of vertices outside the giant component is $O_m(c^m n)$ for some $0 < c < 1$.

Theorem 4.3. *Condition on $\overline{\mathcal{D}}$.*

(i) There exists a $\xi = \xi(m,p) \in (0,p)$ such that the number of isolated vertices in G_n is $\xi n(1 + o_n(1))$ whp. If $\alpha > 1$ then $\xi = O_m(c^m)$ for some $0 < c < 1$.

(ii) If $m = 1$, all components in G_n have size $O(\Delta \log n)$ whp, where Δ denotes the maximum degree of G_n.

(iii) If $m \geq 2$, whp there exists a component containing at least $p(1-\xi)(1-(13/14)^{m-1})n$ vertices while all other components have size $O(\log n)$.

The remainder of the section is devoted to the proof of this theorem. Let $\omega = \log\log n$. We fix a feasible and ω-concentrated σ, see Lemma 4.1. We also fix $\varepsilon > 0$ with $1/2 - \varepsilon > 1/\eta$ if $\alpha > 1$ and $\varepsilon < 1/2$ if $\alpha < 1$.

We first prove (i). The existence of ξ is provided by Theorem 4.2 (ii), so we need only prove that $\xi = O_m(c^m)$ for some $0 < c < 1$ when $\alpha > 1$. Fix a vertex $v \geq n/\omega$. By Theorem 4.1 the probability for v to be isolated is $(1 - q(\tau))^m$ for some τ. By Lemma 4.2 (vi), $\alpha > 1$ implies $1 - q(\tau) \leq \zeta < 1$ for all τ, so the probability of being isolated is at most ζ^m. By linearity of expectation we expect at most $\zeta^m pn + O(n/\omega)$ vertices to be isolated, accounting for the n/ω vertices for which Theorem 4.1 does not apply. Theorem 4.2 shows that the number of vertices of degree zero is within $O(n^{1/2}\ln n)$ of its mean with high probability, so the number of isolated vertices is at most $2p\zeta^m n$ whp. This finishes part (i), and the remainder of the section is devoted to proving (ii), (iii).

The proof will rely heavily on the master graph Γ defined in Section 4.2.1. We will define an algorithm that searches for a large connected edge set in Γ, which remains connected when restricting to the edge set E_n of G_n.

Orient each edge $\{u, v\}$ in Γ from larger to smaller, i.e. $v \to u$ if $v > u$. Then $d^+(v) = m$ for all $v \geq 1_H$ and $d^+(v) = 0$ for $v < 1_H$. When $m = 1$, this implies that Γ is a forest in which each tree is rooted in $\{1, \ldots, 1_H - 1\}$, and any edge is oriented towards the root in its tree. Restricting to E_n breaks the trees into smaller trees. Let $v \in V_n$. Then there exists a unique vertex $u \notin V_n$ that is reachable from v via directed edges in E_n. The connected component of v is T_u, where T_u is the tree rooted at u of vertices which can reach u via a directed path. This shows that the connected components in G_n are $\{T_u : u \notin V_n\}$ when $m = 1$.

We now show that $|T_u| = O(d(n, u) \log n)$ for all u whp. Let $u \notin V_n$ and let $v_1, \ldots, v_k$ be the neighbors of u in V_n, and let e_i be the unique edge oriented out of v_i for $i = 1, \ldots, k$. Expose $(e_1, 2), \ldots, (e_k, 2)$. For any edge e found, we expose $(e, 1)$ and $(e, 2)$. Repeating the coupling argument of Lemma 4.5 one can show that the of descendants of e_1 can modelled by a CMJ process of rate 2α. The number of descendants of e_1 is geometrically distributed with rate $e^{-2\alpha\tau_1}$ for $\tau_1 = \log_\gamma(pmn/e_1) \leq 1 + O(n^{-1/2} \ln n)$. With high probability each e_i has $O(\log n)$ descendants, and it follows that whp $|T_u| = O(d(n, u) \log n)$ for all $u \notin V_n$. In particular, the largest component has size $O(\Delta \log n)$ where Δ denotes the maximum degree of G_n. In this paper we make no attempt to bound Δ.

Let $m \geq 2$ for the remainder of the section. We now loosely describe the intuition that will help us prove the theorem. Suppose $e_1, \ldots, e_m$ are the m edges oriented out of $v \in V_n$ in G_n. We imagine splitting v into m smaller vertices $v_1, \ldots, v_m$ with $d^+(v_i) = 1$ for each i. In Section 4.4 we saw that each edge e directed into v can be traced back to a unique e_i, in that e either directly chooses $(e_i, 2)$ or chooses $(e', 1)$ for some e' that chooses $(e_i, 2)$, and so on. If e can be traced back to e_i, we make it point to v_i. Let G'_n be the graph in which all vertices in V_n are split into m parts in this fashion. In G'_n vertices have out-degree 0 or 1, and we can define trees T_u as above for $u \notin V_n$. Then each $v \in V_n$ is associated with m trees, namely the m connected components of $v_1, \ldots, v_m$ in G'_n.

We now make this precise. Let $u \notin V_n$. In Section 4.4 we saw how to find the neighbors of u in V_n by exposing $(e_1, 2), \ldots, (e_m, 2)$ for the m edges $e_1, \ldots, e_m$ oriented out of u in Γ. We start building T_u by letting u be the root, and the children of u each vertex $v \in V_n$ that is adjacent to u. For such a v, let e_v be an edge that was found when exposing $(e_1, 2), \ldots, (e_m, 2)$. Expose $(e_v, 2)$ to find all neighbors of v that can be traced back to the edge e_v. The children of v in T_u will be all neighbors of v that are incident to some edge that can be traced back to the edge e_v. Repeat this for all $v \in V_n$ in T_u. Note that T_u may not be a tree, since two edges adjacent to the same vertex may be found when exposing edges.

With this definition, we can partition the edges of G_n into $\{T_u : u \notin V_n\}$. In particular, for each $e \in E_n$ there is a unique vertex $u \notin V_n$ such that $e \in T_u$. Write $T_e = T_u$. The idea behind the algorithm described in detail below is to do a "breadth-first search on the T_u". Starting with a free edge $x_0 \in E_n$, we determine (part of) T_{x_0}. For any edge $f \in T_{x_0}$, we expect the other $m - 1$ edges oriented out of the same vertex as f to be free. These $m - 1$ edges provide the starting point for $m - 1$ future rounds of the algorithms, and in each round a new T_u is determined.

For a vertex v_0 let $C_\Gamma(v_0), C_G(v_0)$ be the set of edges in the connected component of v_0 in Γ, G, respectively. Starting with a vertex v_0 and the graph $\Gamma_0 \in \mathcal{G}(\emptyset, \emptyset)$, we use the following algorithm to find a set $C(v_0) \subseteq C_G(v_0)$. An explanation of the algorithm follows immediately after its description. See Figure 4.2 for an example outcome of one round of the algorithm.

0. If $v_0 \in V_n$ let $C = X = \{m(v_0 - 1) + 1, \ldots, mv_0\}$, and $A = R = \emptyset$. If $v_0 \notin V_n$, set $C = X = A = R = \emptyset$ and $Q(x_0) = \{(m(v_0 - 1) + 1, 2), \ldots, (mv_0, 2)\}$ and go to step 3.

1. If $X = \emptyset$, stop. If $X \neq \emptyset$ choose an edge $x_0 \in X$ and remove it from X. Set $Q(x_0) = \{(x_0, 1), (x_0, 2)\}$, $X_1(x_0) = \emptyset$ and $Y_1(x_0) = \emptyset$.

2. Choose $(x_1, j_1) \in \Omega(x_0) \setminus R$ uniformly at random.

   (2.1.) If $x_1 \in A$, do nothing.

   (2.2.) If $x_1 \in E_n$, add $D(x_1)$ to $X_1(x_0)$.

   (2.3.) If $x_1 \notin E_n$ and $j_1 = 2$, add $(x_1, 2)$ to H.

   (2.4.) If $x_1 \notin E_n$ and $j_1 = 1$, choose $(x_2, j_2) \in \widetilde{\Omega}(x_1)$ uniformly at random. Repeat until one of the following holds:

   (2.4.1.) $j_1 = j_2 = \cdots = j_{k-1} = 1$ and $j_k = 2$. Add $(x_1, 1), \ldots, (x_{k-1}, 1)$ and $(x'_k, 2)$ for all $x'_k \in D(x_k)$ to $Q(x_0)$. Set $\Omega(x_i) = \emptyset$ for $i = 1, 2, \ldots, k-1$.

   (2.4.2.) $j_1 = j_2 = \cdots = j_{k-1} = 1$ and $x_k < m\nu_H$. Let v be the vertex (in H) corresponding to (x_k, j_k). Add $(x_1, 1), \ldots, (x_{k-1}, 1)$ to $Q(x_0)$, along with (x', j') for all edges x' incident to v in H, for the proper choice of j'.

   Add $x_0, x_1, \ldots, x_{k-1}$ to A.

3. While $Q(x_0)$ is nonempty, repeat the following.

   (3.1.) Pick $(h, j) \in Q(x_0)$ and remove it from $Q(x_0)$. Let $Y' = \{(h, j)\}$. Add h to $Y_1(x_0)$. While $Y' \neq \emptyset$ repeat the following:

   (3.1.1) Choose $(y, i) \in Y'$ and remove it from Y'. For each $e \notin X \cup A$ with $(y, i) \in \Omega(e)$, query whether e chooses (y, i), i.e. set $\phi(e) = (y, i)$ with probability $1/|\Omega(e)|$ and remove (y, i) from $\Omega(e)$ otherwise. If e chooses (y, i) then add $(e, 1)$ to Y' and $Y_1(x_0)$, and add all edges $f \neq e$ with $\lceil f/m \rceil = \lceil e/m \rceil$ to $X_1(x_0)$ and X. If $e \in E_n$ then also add $(e, 2)$ to Y' and $Y_1(x_0)$.

4. Set $C \leftarrow C \cup X_1(x_0) \cup (Y_1(x_0) \cap E_n)$. Go to step 1.

Explanation of algorithm: We call steps 1–4 a *round* of the algorithm. At the beginning of each round, we choose some free edge $x_0 \in E_n$ that has been determined to be in $C \subseteq C_G(e_0)$. The objective of the round is to build the set T_{x_0} in order to find free edges $X_1(x_0)$ which share a component with x_0. See Figure 4.2 for a typical outcome of a round in which $x_1 \notin E_n$. Note that part of T_{x_0} may have been found in a past round.

Step 0 is a preliminary step; if $v_0 \in V_n$ then we feed the m free edges adjacent to v_0 into X, and if $v_0 \notin V_n$ then we find T_{v_0} in step 3 and feed any free edges adjacent to T_{v_0} into X in step 4. We call this round 0.

The edge x_0 makes a random choice (x_1, j_1). If $x_1 \in E_n$ then $T_{x_0} = T_{x_1}$ and we cut the search short and find all of T_{x_0} in a future round. The reason for this is mainly to make calculations easier in Lemma 4.6. In the current round we will find the part of T_{x_0} that can be traced back to x_0.

The edge x_0 has a fixed endpoint $\lceil x_0/m \rceil$ and a random endpoint $v(x_0)$. If $x_1 \notin E_n$ then $v(x_0) \notin V_n$, and we will have $T_{x_0} = T_{v(x_0)}$. In step 2 we determine $v(x_0)$. We assign x_0 to (x_1, j_1), and if $j_1 = 1$ we assign x_1 to (x_2, j_2), and so on until one of two things happen. If $j_1 = j_2 = \cdots = j_{k-1} = 1$ and $j_k = 2$ for some k then $v(x_0) = \lceil x_k/m \rceil$. If $j_1 = \cdots = j_k = 1$ and $x_k \leq m\nu_H$, then $v(x_0) = v(x_1) = \cdots = v(x_k)$, noting that $v(x_k)$ is not random when $x_k \leq m\nu_H$.

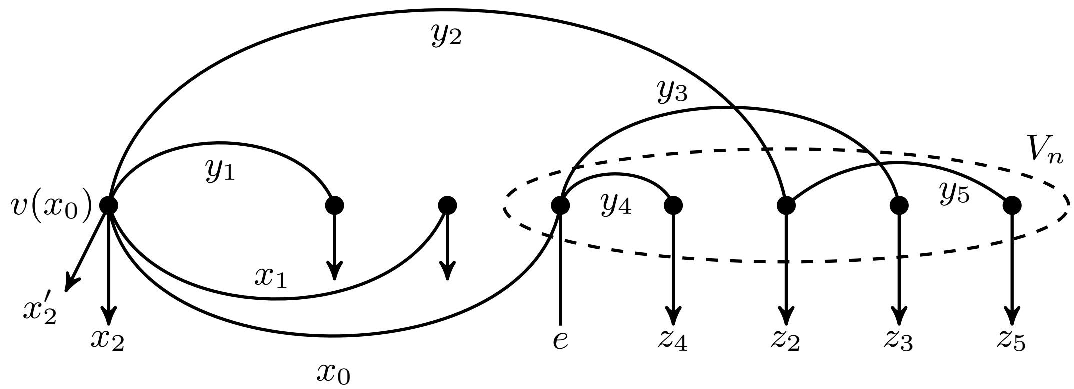

Figure 4.2: A typical round of the algorithm when $m = 2$ and $x_1 \notin E_n$. Free edges are denoted by arrows, and e is an edge with $\lceil e/m \rceil = \lceil x_0/m \rceil$ found in a previous round which may or may not be free. In this example, $Y_1(x_0) = \{y_1, y_2, y_3, y_4, y_5\}$ and $X_1(x_0) = \{z_2, z_3, z_4, z_5\}$. Note that each member of $Y_1(x_0) \cap E_n = \{y_2, y_3, y_4, y_5\}$ contributes exactly $(m-1)$ free edge(s) to $X_1(x_0)$. Edges $y_2, y_3, y_4, y_5, z_2, z_3, z_4, z_5$ are added to C, which already contains x_0 and e. Half-edges $(x_0, 1), (x_0, 2), (x_1, 1), (x_2, 2), (x_2', 2), (y_1, i), (y_2, i), (y_3, i), (y_4, i), (y_5, i)$, $i = 1, 2$, are added to R, and edges $x_0, x_1, y_1, y_2, y_3, y_4, y_5$ to A. Edges x_0, y_2, y_3, y_4, y_5 are in T_{x_0}. Note that T_{x_0} may contain more edges, not pictured, if some edge randomly chose $(x_0, 1)$ in a previous round.

At the start of any round, we have sets A, R, X and a partially generated graph $\widetilde{\Gamma} \in \mathcal{G}(A, R)$ such that if $e \in A$ then $(e, 1) \in \Omega(x)$ only if $x \in X$. For this reason, it is not possible that $x_j \in A$ for any $j \geq 2$, since we only consider $j \geq 2$ when $x_1 \notin E_n$, so $x_1 \notin X$.

Assuming $v(x_0)$ was found, in step 3 we find T_{x_0} using a modification of the *expose* algorithm in Section 4.4, noting that part of the tree has already been built. We do this by exposing (i) $(e, 2)$ for the m free edges e adjacent to $v(x_0)$, (ii) $(e, 1)$ for all edges determined to be in T_{x_0}, and (iii) $(e, 2)$ for the edges in T_{x_0} that are in E_n. We take care not to include edges in X, and in particular if one edge e is determined to be in X then we immediately place the other $m-1$ edges adjacent to $\lceil e/m \rceil$ in X. These rules are included to avoid X decreasing in size.

Entering step 4 we have a set $Y_1(x_0)$ of non-free edges that are in T_{x_0} and a set $X_1(x_0)$ of free edges whose fixed endpoint is also the fixed endpoint of some edge in $Y_1(x_0) \cap E_n$. If $x_1 \in E_n$ we have $|X_1(x_0)| = m + (m-1)|Y_1(x_0) \cap E_n|$, and if $x_1 \notin E_n$ then $|X_1(x_0)| = (m-1)|Y_1(x_0) \cap E_n|$.
End of explanation.

If the algorithm terminates, i.e. $X = \emptyset$ at some point, then $C = C_G(v_0)$. By estimating the round T at which the algorithm terminates, we can estimate the size of $C_G(e_0)$ via Lemma 4.7 (ii) below. Let $E_c = \{e : e > mn/\omega\}$ be the set of edges for which Lemma 4.5 applies. In Lemma 4.6 we estimate T by showing that if $R \cap E_c$ (taken to mean $\{e \in E_c : (e, 1) \in R \text{ or } (e, 2) \in R\}$) is not too large then $\{|X_t| : t \geq 0\}$ is bounded below by a random walk with positive drift.

Lemma 4.6. *Suppose $m \geq 2$ and let Z be a random variable taking values in $\{0, 1, 2\}$ with $Pr\{Z = 0\} = 0.26$ and $Pr\{Z = 1\} = 0.46$. Suppose a round starts at $x_0 \in X$ and with $|R \cap E_c| \leq n^{1/2+\varepsilon} \log_\gamma^3 n$. Then $|X_1(x_0)|$ is stochastically bounded below by Z.*

The following lemma shows that if $R \cap E_c$ is too large for the bounds in Lemma 4.6 to apply, then we have found a large component.

Lemma 4.7. *Let C_t, R_t, X_t denote the states of C, R, X after t rounds of the algorithm.*

(i) There exists a constant $\lambda > 0$ such that $|R_t| \leq \lambda |C_t| \log_\gamma^3 n$ for all t with probability $1 - o(n^{-1})$.

(ii) For all t, $\frac{1}{2}|C_t| \leq |X_t| + t \leq |C_t|$.

The proofs of Lemmas 4.6 and 4.7 are postponed to the end of this section. Suppose the algorithm is run starting at some vertex v_0. If at some point $|R \cap E_c| \geq n^{1/2+\varepsilon} \log_\gamma^3 n$ then we conclude that $|C_G(v_0)| \geq \lambda^{-1} n^{1/2+\varepsilon}$, and say that the component (and every edge and vertex in it) is *large*. If the algorithm terminates with $|X| = 0$ then we say that the component is *small*.

As long as $|R \cap E_c| < n^{1/2+\varepsilon} \log_\gamma^3 n$ we will bound $|X_t|$ below by a random walk $|X_0| + \sum_{i=1}^t (Z_i - 1)$ where the Z_i are independent copies of Z as defined in Lemma 4.6. Here X_0 is the state of X after round 0, and we have $|X_0| = m$ if $v_0 \in V_n$, $|X_0| = 0$ if $v_0 \notin V_n$ is isolated in G_n, and $|X_0| \geq m - 1$ if $v_0 \notin V_n$ is non-isolated in G_n.

The rest of the proof follows from four separate claims.

Claim 1: Small components have size $O(\log n)$. Let X_t, R_t denote the states of the sets X, R after t rounds of the algorithm, i.e. when steps 1–4 have been executed t times. Let T denote the minimum $t > 0$ for which $X_t = \emptyset$. We have $|C_G(e_0)| = |C_T|$, so by Lemma 4.7 (ii), $\frac{1}{2}|C_G(e_0)| \leq T \leq |C_G(e_0)|$ with probability $1 - o(n^{-1})$. We bound the probability that $c \log n \leq T \leq n^{1/2+\varepsilon}$ for some $c > 0$ to be chosen.

Suppose $t < T$. Since $|X_{t+1}| \geq |X_t| - 1$ for all t, we must have $0 = |X_T| \geq |X_t| - (T - t)$, so $T \geq |X_t| + t$. Conditioning on Lemma 4.7, $T \leq n^{1/2+\varepsilon}$ implies that for all $t < T$,

$$|R_t| \leq 2\lambda(|X_t| + t) \log_\gamma^3 n \leq 2\lambda n^{1/2+\varepsilon} \log_\gamma^3 n,$$

so

$$\Pr\left\{c \log n \leq T \leq n^{1/2+\varepsilon}\right\} \leq \Pr\left\{c \log n \leq T \leq n^{1/2+\varepsilon} \middle| |R_t| \leq 2\lambda n^{1/2+\varepsilon} \log_\gamma^3 n \text{ for } t \leq T\right\}.$$

Conditioning on $|R_t| \leq 2\lambda n^{1/2+\varepsilon} \log_\gamma^3 n$, Lemma 4.6 applies. We couple $|X_t| - |X_{t-1}|$ to independent copies $Z_t - 1$ of $Z - 1$, so if $|X_0|$ denotes the size of X after round 0,

$$|X_t| = |X_0| + \sum_{i=1}^t (|X_i| - |X_{i-1}|) \geq m - 1 + \sum_{i=1}^t (Z_i - 1).$$

Here $|X_0| \geq m - 1$ whenever $T > 0$.

The process $W_t = m - 1 + \sum_{i=1}^t (Z_i - 1)$ is a random walk with $W_t - W_{t-1} \in \{-1, 0, 1\}$ and $\mathbf{E}[W_t - W_{t-1}] = \mathbf{E}[Z_t - 1] = 0.02$. Choosing $c > 0$ large enough, Hoeffding's inequality [57] shows that

$$\Pr\{\exists t \geq c \log n : W_t = 0\} \leq \sum_{t \geq c \log n} \Pr\{Z_1 + \cdots + Z_t < 1.01t\} = o(n^{-1}),$$

and since $|X_t| \geq W_t$, it follows that with probability $1 - o(n^{-1})$ the algorithm either terminates after at most $c \log n$ steps, or $T \geq n^{1/2+\varepsilon}$, in which case the component is large. Since $\frac{1}{2}|C_G(e_0)| \leq T \leq$

$|C_G(e_0)|$ with probability $1 - o(n^{-1})$, and the number of components is $O(n)$, all small components have size at most $c \log n$ with high probability.

Claim 2: The probability for a non-isolated v_0 to be in a small component is at most $(13/14)^m$. Recall that in a small component, $|X_t| \geq W_t$ for a random walk W_t as above. Since $W_0 \geq m - 1$, we have

$$\Pr\{\exists t \leq c \log n : |X_t| = 0\} \leq \Pr\{\exists t : W_t = 0\} = \left(\frac{0.26}{0.28}\right)^{m-1} = \left(\frac{13}{14}\right)^{m-1},$$

see e.g. [55, Exercise 5.3.1].

Claim 3: All large vertices are in the same connected component. Suppose v is a large edge and let X_v, R_v be the states of X, R at the point that $|R|$ hits $n^{1/2+\varepsilon} \log_\gamma^3 n$ when the algorithm is run starting at v. Then the above shows that $|X_v| \geq cn^{1/2+\varepsilon}$ whp for some $c > 0$. Similarly, if w is a large vertex then $|X_w| \geq cn^{1/2+\varepsilon}$. Assign all edges in $X_v \cup X_w$. For every pair $e \in X_v, f \in X_w$, either $e \in E_f^\sigma$ or $f \in E_e^\sigma$, since edges in X are required to be in the edge set E_n of G_n. In particular, either half the edges $e \in X_v$ have half of X_w in E_e^σ, or half the edges $f \in X_w$ have half of X_v in E_f^σ. In the former case, the probability that no edge $e \in X_v$ chooses any $f \in X_w$ is bounded above by

$$\left(1 - \frac{\Omega(n^{1/2+\varepsilon})}{n}\right)^{\Omega(n^{1/2+\varepsilon})} = \exp\left\{-\Omega(n^{2\varepsilon})\right\}$$

and in the other case, the same bound holds. So with high probability, any two large edges belong to the same component. In other words, there is a unique large component.

Claim 4: The large component contains $\Omega(n)$ vertices. The number of vertices in G_n is $pn + O(n^{1/2} \ln n)$ since σ is ω-concentrated. By part (i) of Theorem 4.3, the number of non-isolated vertices is$(1 - \xi)n + O(n^{1/2} \ln n)$ whp for some $\xi > 0$. By linearity of expectation and Claim 2, the number S of small, non-isolated vertices in G_n satisfies

$$\mathbf{E}[S] \leq (1 - \xi) \left(\frac{13}{14}\right)^{m-1} n + O(n^{1/2} \ln n).$$

We note that $\mathbf{E}[S] = \Omega(n)$: when the algorithm starts with $X = \{x_1, \ldots, x_m\}$, the m free edges adjacent to some $v_0 \in V_n$, the probability that $X_1(x_i) = \emptyset$ for $i = 1, 2, \ldots, m$ is bounded away from 0.

Write $S = \sum_{v \in V} S_v$ where S_v is the indicator variable for v being small. Then $\mathbf{E}[S(S-1)] = \sum_{u \neq v} \mathbf{E}[S_u S_v]$. Fix $u \neq v$. Suppose we run the process starting at v and find that the component is small. In the process of determining that the component is small, we assign some edges A_v and expose some half-edges R_v, where $|A_v| = O(\log n)$ and $|R_v| = O(\log^2 n)$. The probability that u is in the component is $O(\log^2 n/n)$. If u is not in the component, the algorithm is run starting at u on the partially generated $\widetilde{\Gamma} \in \mathcal{G}(A_v, R_v)$. In the statement of the algorithm we assumed that it is run on $\Gamma_0 \in \mathcal{G}(\emptyset, \emptyset)$, but it can be easily modified to accommodate for $\widetilde{\Gamma} \in \mathcal{G}(A_v, R_v)$, and it will follow that $\mathbf{E}[S_u \mid S_v = 1] = \mathbf{E}[S_u](1 + o(1))$. Hence $\mathbf{E}[S_u S_v] = \mathbf{E}[S_u]\mathbf{E}[S_v](1 + o(1))$, and Chebyshev's inequality shows that $S = \mathbf{E}[S] + o(n)$. Since $\mathbf{E}[S] = \Omega(n)$, this shows that with high probability,

$$S = \mathbf{E}[S] + o(n) \leq (1 - \xi) \left(\frac{13}{14}\right)^{m-1} n + o(n).$$

The theorem follows.

4.6.1 Proof of Lemma 4.6

We first note that $\Pr\{x_1 \in A\} = o(1)$, since if $x \in A$ then $(x, 1) \in R$, and $|R| = o(n)$. If $x_1 \in E_n \setminus A$ then $D(x_1) \subseteq X_1(x_0)$ so $|X_1(x_0)| \geq m \geq 2$. If $x_1 \notin E_n$ we have $|X_1(x_0)| = (m-1)|Y_1(x_0) \cap E_n|$. The lemma will follow from showing that for all $m \geq 2$,

$$\Pr\{|Y_1(x_0) \cap E_n| = 0 \text{ and } x_1 \notin E_n\} \leq 0.255$$

and for $m = 2$,

$$\Pr\{|Y_1(x_0) \cap E_n| = 1 \text{ and } x_1 \notin E_n\} \leq 0.455.$$

Throughout this proof we take $a \approx b$ to mean that $a = b + o_n(1)$. Let (x_1, j_1) be the random choice of x_0. We first note that if $\tau_0 = \log_\gamma(pmn/x_0) \in [0, 1]$ and $\tau_1 = \log_\gamma(pmn/x_1)$ then for $y \in [0, 1]$,

$$\Pr\{\tau_1 - \tau_0 \leq y\} \approx \frac{\ln \gamma}{1 - 1/\gamma} \int_0^y \gamma^{-x} dx. \tag{4.11}$$

Indeed, (x_1, j_1) is a uniformly random member of $\widetilde{\Omega}(x_0) = (E^\sigma_{x_0} \times [2]) \setminus R$, and since σ is ω-concentrated we have $E^\sigma_{x_0} = \{x_0/\gamma + O(n^{-1/2} \ln n), \ldots, x_0 - i\}$ for some $i \in [m]$. Since $|R| = o(|E^\sigma_{x_0}|)$ and $\widetilde{\Omega}(x_0) \supseteq (E^\sigma_{x_0} \times [2]) \setminus R$, we can view x_1 as essentially being a uniform member of $E^\sigma_{x_0}$. Then $\tau_1 - \tau_0 = \log_\gamma(x_0/x_1)$ is exponentially distributed, truncated to $[0, 1]$ as in (4.11). In particular, since $x_1 \notin E_n$ when $\tau_1 > 1 + \delta$ for some $\delta = O(n^{-1/2} \ln n)$.

$$\Pr\{x_1 \notin E_n\} \approx \frac{\ln \gamma}{1 - 1/\gamma} \int_{1-\tau_0}^1 \gamma^{-x} dx = \frac{\gamma^{\tau_0} - 1}{\gamma - 1} \leq \tau_0. \tag{4.12}$$

Claim A: Let $m \geq 2$ and $x_0 \in E_n$. Then $\Pr\{|Y_1(x_0) \cap E_n| = 0 \text{ and } x_1 \notin E_n\} < 0.255$.

Proof of claim A: Let $m \geq 2$ and fix an edge $x_0 \in E_n$. Suppose $x_1 \notin E_n$. In step 2 of the algorithm we then find a chain of edges $x_1, x_2, \ldots, x_K$ for some random K. Since $|R| = o(n^{3/4})$ and $|\Omega(x_i)| = \Omega(n/\omega)$ for all $x_i \geq mn/\omega$, we have $\Pr\{j_i = 1\} = 1/2 + o(n^{-1/4})$ for all i, and K is approximately geometric with mean 2. In particular, since $\log_\gamma(x_i/x_{i-1}) \leq 1 + o(1)$ for all i we have $x_K \geq mn/\omega$ with probability $1 - o_n(1)$. Condition on this.

We will consider two subsets of $Y_1(x_0)$. Let $\mathcal{R}(x_0)$ be the edges found when exposing $(x_0, 1)$ and $(x_0, 2)$, and let $\mathcal{L}(x_0)$ be the set of edges in E_n found by exposing $(x_1, 1), (x_2, 1), \ldots, (x_{K-1}, 1)$ and $(x'_K, 2)$ for all $x'_K \in D(x_K)$. Then

$$\Pr\{|X_1(x_0) = 0|\} = \Pr\{|\mathcal{R}(x_0)| = |\mathcal{L}(x_0)| = 0\},$$

and we now argue that $|\mathcal{R}(x_0)|, |\mathcal{L}(x_0)|$ are essentially independent. We find $\mathcal{R}(x_0)$ by exposing $(x_0, 1)$ and $(x_0, 2)$. By Lemma 4.5, the number of edges found is asymptotically geometric, and in particular is $O(\log n)$ whp. Initially $|\Omega(e)|$ is of order n for all $e > x_0$, so exposing $O(\log n)$ edges only shrinks $\Omega(e)$ to $\widetilde{\Omega}(e)$ of size $|\widetilde{\Omega}(e)| = |\Omega(e)|(1 - o(1))$. When $|\mathcal{L}(x_0)|$ is calculated, starting with $\widetilde{\Omega}(e)$ instead of $\Omega(e)$ for $e > x_0$ makes an insignificant difference to the result, and we have

$$\Pr\{|\mathcal{R}(x_0)| = j_1 \text{ and } |\mathcal{L}(x_0)| = j_2\} \approx \Pr\{|\mathcal{R}(x_0) = j_1\}\Pr\{|\mathcal{L}(x_0)| = j_2\}.$$

Let $\tau_0 = \log_\gamma(pmn/x_0)$. Since σ is ω-concentrated we have $\tau_0 \in (-\delta, 1+\delta)$ for some $\delta = O(n^{1/2} \ln n)$, see Lemma 4.1. Assume for now that $\tau_0 \in [0, 1]$. Let $E(x_0, i)$ denote the set of edges in E_n found by exposing (x_0, i). By Lemma 4.5, $|E(x_0, i)|$ is asymptotically geometrically distributed (nonzero since $x_0 \in E_n$) with rate $e^{-\alpha\tau_0}$ for $i = 1, 2$ so

$$\Pr\{|\mathcal{R}(x_0)| = j\} \approx \begin{cases} e^{-2\alpha\tau_0}, & j = 0, \\ 2e^{-2\alpha\tau_0}(1 - e^{-\alpha\tau_0}), & j = 1. \end{cases} \tag{4.13}$$

Now consider the chain $x_0 > x_1 > \cdots > x_K$ where x_{i-1} chooses $(x_i, 1)$ for $1 \le i \le K-1$ and x_{K-1} chooses $(x_K, 2)$. If $K = 1$ and $x_1 \notin E_n$, then $\Pr\{|\mathcal{L}(x_0)| = 0\} \approx (1 - q(\tau_1))^m \le (1 - q(\tau_1))^2$ by Lemma 4.5, where $\tau_1 - \tau_0$ can be approximated by a truncated exponential as above, so

$$\Pr\{|\mathcal{L}(x_0)| = 0,\ K = 1 \text{ and } x_1 \notin E_n\} \le \frac{1}{2}\frac{\ln\gamma}{1 - 1/\gamma}\int_{1-\tau_0}^{1} \frac{(1 - q(\tau_0 + x))^2}{\gamma^x}dx.$$

In Claim C we show that for all $\alpha > 1/2$ and $\tau_0 \in [0, 1]$,

$$\frac{1}{2}\frac{\ln\gamma}{1 - 1/\gamma}\int_{1-\tau_0}^{1} \frac{(1 - q(\tau_0 + x))^2}{\gamma^x}dx \le \frac{\tau_0}{2e - e^{1/2}}.$$

If $K > 1$, then $\mathcal{L}(x_0) = \emptyset$ only if $E(x_1, j_1) = E(x_2, j_2) = \emptyset$. If $\tau_i = \log_\gamma(pmn/x_i)$ denotes the age of x_i then the probability of $E(x_i, j_i)$ being empty is $1 - q(\tau_i) \le 1 - q(\tau_0 + i)$ for $i = 1, 2$. Here we used the fact that $q(\tau)$ is decreasing, see Lemma 4.2 (iii). Since $\Pr\{K \ge 2\} = 1/2$, we have by (4.12),

$$\Pr\{|\mathcal{L}(x_0)| = 0,\ K \ge 2 \text{ and } x_1 \notin E_n\} \le \frac{\tau_0}{2}(1 - q(\tau_0 + 1))(1 - q(\tau_0 + 2)).$$

The function $q(\tau)$ is defined in Section 4.2.2, and we have

$$(1 - q(\tau_0 + 1))(1 - q(\tau_0 + 2)) = \frac{1}{e^\alpha - \alpha\tau_0}\frac{e^\alpha - \alpha\tau_0}{e^{2\alpha} - (\tau_0 + 1)\alpha e^\alpha + \frac{1}{2}\alpha^2\tau_0^2}.$$

We show in Claim C that this is at most $1/(e - e^{1/2} + 1/8)$. So

$$\begin{aligned} &\Pr\{|Y_1(x_0) \cap E_n| = 0 \text{ and } x_1 \notin E_n\} \\ \le& e^{-2\alpha\tau_0}\left(\frac{1}{2}\frac{\ln\gamma}{1 - 1/\gamma}\int_{1-\tau_0}^{1} \frac{(1 - q(\tau_0 + x))^2}{\gamma^x}dx + \frac{1}{2}(1 - q(\tau_0 + 1))(1 - q(\tau_0 + 2))\right) \\ \le& \tau_0 e^{-2\alpha\tau_0}\left(\frac{1}{2e - e^{1/2}} + \frac{1}{2(e - e^{1/2} + \frac{1}{8})}\right). \end{aligned}$$

Let L_0 denote the expression in brackets, and note that $L_0 < 0.69$. We have $\tau_0 e^{-2\alpha\tau_0} \le e^{-1}$ for $\alpha > 1/2$ and $\tau_0 \in [0, 1]$, so

$$\Pr\{|Y_1(x_0) \cap E_n| = 0 \text{ and } x_1 \notin E_n\} < e^{-1} \cdot 0.69 < 0.255.$$

End of proof of claim A.

Claim B: Let $m = 2$ and $x_0 \in E$. Then $\Pr\{|Y_1(x_0) \cap E_n| = 1 \text{ and } x_1 \notin E_n\} < 0.455$.

Proof of claim B: We note that while $\mathcal{L}(x_0)$ and $\mathcal{R}(x_0)$ do not necessarily partition $Y_1(x_0) \cap E_n$, it is the case that

$$\begin{aligned} \Pr\{|Y_1(x_0) \cap E_n| = 1\} &\le \Pr\{|\mathcal{L}(x_0)| = 1, |\mathcal{R}(x_0)| = 0\} + \Pr\{|\mathcal{L}(x_0)| = 0, |\mathcal{R}(x_0)| = 1\} \\ &\approx \Pr\{|\mathcal{L}(x_0)| = 1\}\Pr\{|\mathcal{R}(x_0)| = 0\} + \Pr\{|\mathcal{L}(x_0)| = 0\}\Pr\{|\mathcal{R}(x_0)| = 1\}. \end{aligned}$$

We calculated the probability that $|\mathcal{R}(x_0)| = 0, 1$ in (4.13). For the probability that $|\mathcal{L}(x_0)| = 1$, let x_{K+1} denote the edge added along with x_K (so $|x_{K+1} - x_K| = 1$). Let $\tau_i = \log_\gamma(pmn/x_i)$ for $i = 0, 1, \dots, K+1$. Then

$$\Pr\{|\mathcal{L}(x_0)| = 1 \mid x_1 \notin E_n\} \le \sum_{k\ge 1} \Pr\{K = k\} \sum_{i=1}^{k+1} q(\tau_i)p(\tau_i) \prod_{\substack{1\le j\le k+1 \\ j\ne i}} (1 - q(\tau_j))$$

where i denotes the edge whose exposure contributes to $\mathcal{L}(x_0)$. We use the bound $1 - q(\tau_1) \le 1 - q(\tau_0 + 1)$ whenever $1 - q(\tau_1)$ is involved in the product (i.e. when $i > 1$), and bound $p(\tau_i)q(\tau_i) \le p(1)q(1)$ for all $i \ge 1$ (which follows from $p(\tau), q(\tau)$ being decreasing, see Lemma 4.2 (iii)) to get

$$\begin{aligned}\Pr\{|\mathcal{L}(x_0)| = 1 \mid x_1 \notin E_n\} &\le \sum_{k\ge 1} \frac{1}{2^k} p(1)q(1)\left(1 + k(1 - q(\tau_0 + 1))\right) \\ &= e^{-\alpha}(1 - e^{-\alpha})\left(1 + \frac{2}{e^\alpha - \alpha\tau_0}\right) \\ &\le \frac{1}{4}\left(1 + \frac{2}{e^{1/2} - 1/2}\right).\end{aligned}$$

This bound holds for all $\alpha > 1/2, \tau_0 \in [0, 1]$. Let $L_1 = 1/4 + 1/(2e^{1/2} - 1)$.

We now bound

$$\begin{aligned}\Pr\{|Y_1(x_0) \cap E_n| = 1 \text{ and } x_1 \notin E_n\} \le& \Pr\{|\mathcal{R}(x_0)| = 0\}\Pr\{|\mathcal{L}(x_0)| = 1 \mid x_1 \notin E_n\}\Pr\{x_1 \notin E_n\} \\ &+ \Pr\{|\mathcal{R}(x_0)| = 1\}\Pr\{|\mathcal{L}(x_0)| = 0 \text{ and } x_1 \notin E_n\} \\ \le& \tau_0 e^{-2\alpha\tau_0} L_1 + 2\tau_0 e^{-2\alpha\tau_0}(1 - e^{-\alpha\tau_0})L_0 \\ \le& \frac{1}{e} L_1 + \frac{1}{\alpha e}(1 - e^{-\alpha})L_0,\end{aligned}$$

where we used the fact that $\tau_0 e^{-2\alpha\tau_0}$ viewed as a function of τ_0 has a global maximum at $\tau_0 = 1/2\alpha$, so $\tau_0 e^{-2\alpha\tau_0} \le 1/(2\alpha e) \le 1/e$, and we also used $1 - e^{-\alpha\tau_0} \le 1 - e^{-\alpha}$. Finally, $(1 - e^{-\alpha})/(\alpha e)$ is decreasing in α, so

$$\Pr\{|Y_1(x_0) \cap E_n| = 1 \text{ and } x_1 \notin E_n\} \le \frac{L_1}{e} + \frac{2}{e}(1 - e^{-1/2})L_0 < 0.455$$

End of proof of claim B.

Claim C: The following two inequalities hold for all $\alpha > 1/2$ and $\tau_0 \in [0, 1]$:

$$(1 - q(\tau_0 + 1))(1 - q(\tau_0 + 2)) \le \frac{1}{e - e^{1/2} + 1/8}, \tag{4.14}$$

and

$$\frac{\ln\gamma}{2 - 2/\gamma} \int_{1-\tau_0}^{1} \frac{(1 - q(\tau_0 + x))^2}{\gamma^x} dx \le \frac{\tau_0}{2e - e^{1/2}}. \tag{4.15}$$

Proof of claim C: To emphasize the dependence on α we briefly write $q(\alpha, \tau) = q(\tau)$. For $\tau_0 \in [0, 1]$ we have

$$q(\alpha, \tau_0 + 1) = \frac{1}{e^\alpha - \alpha\tau_0}, \quad q(\alpha, \tau_0 + 2) = \frac{e^\alpha - \alpha\tau_0}{e^{2\alpha} - (\tau_0 + 1)\alpha e^\alpha + \frac{1}{2}\alpha^2\tau_0^2}.$$

Suppose $\alpha_1 > \alpha_2$ and let $\mathcal{C}_1$ be a CMJ process with rate α_1. Mark any arrival red with probability α_2/α_1, and consider the CMJ process $\mathcal{C}_r$ on the red arrivals. This will have rate α_2, and if $\mathcal{C}_r$ is active at time τ_0 then so is $\mathcal{C}$. This implies $q(\alpha_1, \tau) \geq q(\alpha_2, \tau)$ for all τ, since $q(\alpha, \tau)$ is the probability that a CMJ process of rate α is active at time τ. So for any $\alpha > 1/2, \tau_0 \in [0, 1]$,

$$1 - q(\alpha, \tau_0 + 1) \leq 1 - q\left(\frac{1}{2}, \tau_0 + 1\right) = \frac{1}{e^{1/2} - \tau_0/2} \tag{4.16}$$

and

$$1 - q(\alpha, \tau_0 + 2) \leq 1 - q\left(\frac{1}{2}, \tau_0 + 2\right) = \frac{e^{1/2} - \tau_0/2}{e - \frac{\tau_0+1}{2}e^{1/2} + \tau_0^2/8}. \tag{4.17}$$

Consider multiplying (4.16) and (4.17). It is easy to confirm that $e - \frac{\tau_0+1}{2}e^{1/2} + \tau_0^2/8$ is decreasing for $\tau_0 \in [0, 1]$, and (4.14) follows.

Now consider (4.15). First note that $\alpha = \frac{p}{4p-2}\ln\gamma = \frac{1}{2-2/\gamma}\ln\gamma$. We have

$$\frac{\ln\gamma}{2 - 2/\gamma}\int_{1-\tau_0}^{1}\frac{(1 - q(\tau_0 + x))^2}{\gamma^x}dx = \int_{1-\tau_0}^{1}\frac{\alpha}{\gamma^x(e^\alpha - \alpha(x + \tau_0 - 1))^2}dx = \int_0^{\tau_0}\frac{\alpha}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^2}dx.$$

Fix τ_0 and let $f(\alpha, x) = \alpha/(\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^2)$ for $0 < x < \tau_0$. We will show that $f(\alpha, x) \leq \lim_{\alpha\to 1/2} f(\alpha, x)$ for $\alpha > 1/2$ by showing that $f(\alpha, x)$ is decreasing in α. To calculate the derivative of $\gamma^{-(x+1-\tau_0)}$ with respect to α, we note that since $\alpha = \frac{1}{2-2/\gamma}\ln\gamma$,

$$\frac{d\gamma}{d\alpha} = \frac{(2\gamma - 2)^2}{2\gamma - 2 - 2\ln\gamma} = \frac{2\gamma - 2}{1 - \frac{1}{\gamma-1}\ln\gamma} = \frac{2\gamma - 2}{1 - 2\alpha/\gamma}.$$

Since $\ln\gamma < \gamma - 1$ we have $1 < 2\alpha = \ln\gamma/(1 - 1/\gamma) < \gamma$, so

$$\frac{d\gamma}{d\alpha} = 2\gamma\frac{\gamma - 1}{\gamma - 2\alpha} > 2\gamma.$$

In particular,

$$\frac{d}{d\alpha}\gamma^{-(x+1-\tau_0)} = -(x + 1 - \tau_0)\gamma^{-(x+1-\tau_0)}\frac{1}{\gamma}\frac{d\gamma}{d\alpha} < -2(x + 1 - \tau_0)\gamma^{-(x+1-\tau_0)}.$$

Now for $0 \leq x \leq \tau_0 \leq 1$ and $\alpha > 1/2$, since $e^\alpha > 1/2 + \alpha x$ we have

$$\begin{aligned}
\frac{\partial f}{\partial\alpha} &= \frac{1}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^2} + \frac{\alpha}{(e^\alpha - \alpha x)^2}\left(\frac{d}{d\alpha}\gamma^{-(x-\tau_0+1)}\right) - \frac{2\alpha(e^\alpha - x)}{\gamma^{x-\tau_0+1}(e^\alpha - \alpha x)^3} \\
&< \frac{1}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^2} - \frac{2(x + 1 - \tau_0)\alpha}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^2} - \frac{2\alpha(e^\alpha - x)}{\gamma^{x-\tau_0+1}(e^\alpha - \alpha x)^3} \\
&= \frac{1}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^3}\left(e^\alpha - \alpha x - 2(x + 1 - \tau_0)\alpha(e^\alpha - \alpha x) - 2\alpha(e^\alpha - x)\right) \\
&< \frac{1}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^3}\left(e^\alpha - \alpha x - 2x\alpha(e^\alpha - \alpha x) - 2\alpha(e^\alpha - x)\right) \\
&= \frac{1}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^3}\left(e^\alpha(1 - 2\alpha) - 2\alpha x(e^\alpha - \alpha x - 1/2)\right) \\
&< 0.
\end{aligned}$$

Noting that $\gamma \to 1$ as $\alpha \to 1/2$, this implies

$$\int_0^{\tau_0} \frac{\alpha}{\gamma^{x+1-\tau_0}(e^\alpha - \alpha x)^2} dx < \int_0^{\tau_0} \frac{1/2}{(e^{1/2} - x/2)^2} = \frac{1}{e^{1/2} - \tau_0/2} - \frac{1}{e^{1/2}} = \frac{\tau_0}{2e^{1/2}(e^{1/2} - \tau_0/2)}.$$

Then (4.15) follows from $e^{1/2} - \tau_0/2 \geq e^{1/2} - 1/2$.
End of proof of claim C.

4.6.2 Proof of Lemma 4.7

Recall Lemma 4.7:

Lemma 4.7. *Let C_t, R_t, X_t denote the states of C, R, X after t rounds of the algorithm.*

(i) There exists a constant $\lambda > 0$ such that $|R_t| \leq \lambda |C_t| \log_\gamma^3 n$ for all t with probability $1 - o(n^{-1})$.

(ii) For all t, $\frac{1}{2}|C_t| \leq |X_t| + t \leq |C_t|$.

Proof of (i). The key observation is that by Lemma 4.4 (iii) and Lemma 4.5, if $e > mn/\omega$ and we expose (e, j) then there exists a $\lambda > 0$ such that $|E(e, j) \cap E_n| \geq \lfloor |E(e, j)|/(\lambda \log_\gamma^2 n) \rfloor$ with probability $1 - o(n^{-1})$. Here $E(e, j)$ denotes the set of edges found when exposing (e, j). Condition on this being the case for all $O(n)$ half-edges exposed over the course of the algorithm. To avoid rounding, we note that if $|E(e, j) \cap E_n| = 0$ then $|E(e, j)| \leq \lambda \log_\gamma^2 n$ and if $|E(e, j) \cap E_n| > 0$ then $|E(e, j)| \leq 2\lambda |E(e, j) \cap E_n| \log_\gamma^2 n$.

The above holds if $e > mn/\omega$. If $e \leq mn/\omega$ and $(e, j) \in Q(x)$, Lemma 4.5 does not apply to exposing (e, j). In this case, reveal (e, j), i.e. find all f such that $\phi(f) = (e, j)$. Note that

$$E(e, j) = \{(e, j)\} \cup \bigcup_{(f,1): f \in \phi^{-1}(e,j)} E(f, 1).$$

Remove (e, j) from $Q(x)$ and replace it by $(f, 1)$ for all $f \in \phi^{-1}(e, j)$. Repeat this until all $(e, j) \in Q(x)$ have $e > mn/\omega$. Let $Q'(x)$ be the end result of this process.

Recall that E_c is the set of edges e with $e > mn/\omega$. We have

$$|R_t \cap E_c| \leq \sum_{i=1}^{t} 2|Y_1(x_i) \cap E_c|, \quad |C_t| \geq \sum_{i=1}^{t} |Y_1(x_i) \cap E_n|$$

and in round i,

$$|Y_1(x_i) \cap E_c| = \sum_{(e,j) \in Q'(x_i)} |E(e, j)|, \quad |Y_1(x_i) \cap E_n| = \sum_{(e,j) \in Q'(x_i)} |E(e, j) \cap E_n|.$$

Letting $(e_1, j_1), \ldots, (e_s, j_s) \in \cup_i Q'(x_i)$ be the half-edges exposed in the first t rounds of the algorithm, we then have

$$|R_t \cap E_c| \leq 2 \sum_{i=1}^{s} |E(e_i, j_i)|, \quad |C_t| \geq \sum_{i=1}^{s} |E(e_i, j_i) \cap E_n|.$$

Let I_i be the indicator variable for $|E(e_i, j_i) \cap E_n| > 0$, and let $I = I_1 + \cdots + I_s$. Then by the above,

$$|R_t \cap E_c| \leq 2(s-I)\lambda \log_\gamma^2 n + 2\lambda \log_\gamma^2 n \sum_{i:I_i=1} |E(e_i, j_i) \cap E_n| = 2(s-I)\lambda \log_\gamma^2 n + 2\lambda |C_t| \log_\gamma^2 n,$$

and we will show that $s \leq I \log n \leq |C_t| \log n$.

Every edge exposed in the process is in E_c, so the probability that $I_i = 1$ is, by Lemma 4.5, $q(\tau_i)$ where $\tau_i = \log_\gamma(pmn/e_i)$. For all i, $\tau_i \leq \log_\gamma \omega$, and $q(\tau)$ is decreasing by Lemma 4.2 (iii), so $I_i = 1$ with probability at least $q(\log_\gamma \omega) \geq \lambda_2 \zeta^{-\log_\gamma \omega}$ where $\lambda_2 > 0$, see Lemma 4.2. Let $c > 0$ be such that $q(\tau_i) \geq \omega^{-c}$ for all i. Then I can be bounded below by a binomial random variable $J \sim \text{Bin}(s, \omega^{-c})$. Suppose $s > 4\omega^{2c} \log n$. Then Hoeffding's inequality [57] implies

$$\Pr\left\{I < s\omega^{-c}\right\} \leq \Pr\left\{J < s\omega^{-c}\right\} \leq \exp\left\{-2\left(\frac{\omega^{-c}}{2}\right)^2 s\right\} \leq n^{-2}.$$

Since $|C_t| \geq I$, This shows that with high probability, if $s > 4\omega^{2c} \log n$ then $s \leq I\omega^c \leq |C_t|\omega^c$ and

$$|R_t \cap E_c| \leq 2(s-I)\lambda \log_\gamma^2 n + 2\lambda |C_t| \log_\gamma^2 n \leq 3\lambda |C_t| \omega^c \log_\gamma^2 n.$$

If $s \leq 4\omega^{2c} \log n$ then $|C_t| \geq 0$ implies

$$|R_t \cap E_c| \leq 4\lambda\omega^{2c} \log_\gamma^3 n \leq 4\lambda\omega^{2c}(|C_t| + 1) \log_\gamma^3 n,$$

and since $\omega^{2c} = (\log\log n)^{2c} \leq \log_\gamma n$ for n large enough, this finishes the proof of (i).

Proof of (ii). In each round we have $|X_1(x)| = m + (m-1)|Y_1(x) \cap E_n|$ if $x_1 \in E_n$ and $|X_1(x)| = (m-1)|Y_1(x) \cap E_n|$ if $x_1 \notin E_n$. In particular, $|Y_1(x) \cap E_n| \leq |X_1(x)|/(m-1) \leq |X_1(x)|$. If x_i denotes the starting edge of round i then

$$|C_t| = m + \sum_{i=1}^t |X_1(x_i)| + |Y_1(x_i) \cap E_n|, \quad |X_t| = m + \sum_{i=1}^t |X_1(x_i)| - 1,$$

so

$$|C_t| - |X_t| - t = \sum_{i=1}^t |Y_1(x_i) \cap E_n| \leq \sum_{i=1}^t |X_1(x_i)| = |X_t| + t.$$

It follows immediately that $\frac{1}{2}|C_t| \leq |X_t| + t \leq |C_t|$.

4.7 Proof of Lemma 4.2

In this section we prove Lemma 4.2, in which we collect useful properties of the central constants and functions defined in Section 4.2.2. We will restate the definitions here to make this section self-contained. Firstly, the integer $m \geq 1$ and the real number $1/2 < p < 1$ are the parameters for the graph process, and we define

$$\mu = m(2p-1), \quad \gamma = \frac{p}{1-p}, \quad \alpha = \frac{pm}{2\mu} \ln \gamma = \frac{p}{4p-2} \ln \gamma.$$

We let $p_0 \approx 0.83113$ be the unique p for which $\alpha = 1$, and when $\alpha \neq 1$ we define ζ as the unique solution in $\mathbb{R} \setminus \{1\}$ to

$$\zeta e^{\alpha(1-\zeta)} = 1. \tag{4.18}$$

If $\alpha > 1$ define $\eta = -\ln\gamma/\ln\zeta$. If $\alpha < 1$ then η is undefined.

We define a sequence a_k by $a_0 = 1$ and

$$a_k = \left(-\frac{e^\alpha}{\alpha}\right)\sum_{j=0}^{k-1}\frac{a_j}{(k-j-1)!},\ k \geq 1. \tag{4.19}$$

For $k \geq 0$ define functions $Q_k : [k, k+1) \to [0, 1]$ by

$$Q_k(\tau) = \sum_{j=0}^{k}\frac{a_j}{(k-j)!}(\tau-k)^{k-j},$$

and for $\tau \geq 0$ we let $Q(\tau) = Q_{\lfloor\tau\rfloor}(\tau)$. We note that $Q(\tau)$ is discontinuous at integer points k with

$$Q(k) = a_k \quad \text{and} \quad Q(k^-) = -\alpha e^{-\alpha}a_k \tag{4.20}$$

where $Q(k^-)$ denotes the limit of $Q(\tau)$ as $\tau \to k$ from below. Define

$$q(\tau) = 1, 0 \leq \tau < 1,\ q(\tau) = 1 + \frac{Q(\tau-1)}{\alpha Q(\tau)},\ \tau \geq 1.$$

Finally, define

$$p(\tau) = \exp\left\{-\alpha\int_0^\tau q(x)dx\right\}.$$

For $\tau < 0$ we define $Q(\tau) = q(\tau) = p(\tau) = 0$.

Lemma 4.2. *(i) If $\alpha > 1$ then $\zeta < \alpha^{-1}$ and if $\alpha < 1$ then $\zeta > 1 - \alpha^{-1} + \alpha^{-2} > \alpha^{-1}$.*

(ii) If $\alpha > 1$ then $\eta > 2$.

(iii) The functions $p(\tau), q(\tau)$ are decreasing and take values in $[0, 1]$.

(iv) For any non-integer $\tau > 0$,

$$Q'(\tau) = Q(\tau-1) \quad \textit{and} \quad q(\tau) = \frac{1}{\alpha}\frac{(Q(\tau)e^{\alpha\tau})'}{Q(\tau)e^{\alpha\tau}}.$$

(v) If $\alpha < 1$ then there exist constants $\lambda_1, \lambda_2 > 0$ where $\lambda_1 < \alpha$ such that for all $\tau \geq 0$,

$$p(\tau) = 1 - \alpha + \frac{\lambda_1}{\zeta^\tau} + O(\zeta^{-2\tau}) \quad \textit{and} \quad q(\tau) = \frac{\lambda_2}{\zeta^\tau} + O(\zeta^{-2\tau}).$$

(vi) If $\alpha > 1$ then there exist constants $\lambda_3, \lambda_4 > 0$ and a constant $C > 0$ such that for all $\tau \geq 0$,

$$\lambda_3\zeta^\tau \leq p(\tau) \leq \lambda_3\zeta^\tau + C\zeta^{2\tau} \quad \textit{and} \quad q(\tau) = 1 - \zeta + \lambda_4\zeta^\tau + O(\zeta^{2\tau}).$$

Proof. **Proof of (i).** Let $\alpha \neq 1$. The function $x \mapsto xe^{\alpha(1-x)}$ is strictly increasing for $x < \alpha^{-1}$ and strictly decreasing for $x > \alpha^{-1}$, and its global maximum at $x = \alpha^{-1}$ is $\alpha^{-1}e^{\alpha-1} > 1$. The two solutions x_1, x_2 of $xe^{\alpha(1-x)} = 1$ must satisfy $x_1 < \alpha^{-1} < x_2$, and $\zeta < \alpha^{-1}$ for $\alpha > 1$ follows from the fact that ζ is the solution which is not 1. When $\alpha < 1$, it is straightforward to plug in $x = 1 - \alpha^{-1} + \alpha^{-2}$ and confirm that $xe^{\alpha(1-x)} > 1$, which shows that $\zeta > 1 - \alpha^{-1} + \alpha^{-2} > \alpha^{-1}$.

Proof of (ii). Let $\alpha > 1$, so $p > p_0 \approx 0.83$. To see that $\eta > 2$, we first note that the definition of α givesln $\gamma = \alpha(4 - 2/p)$ and the definition of ζ gives $\ln \zeta = -\alpha(1-\zeta)$, so

$$\eta = -\frac{\ln \gamma}{\ln \zeta} = \frac{4 - \frac{2}{p}}{1-\zeta} > 1$$

since $4 - \frac{2}{p} > 1$ for $p > p_0 \approx 0.83$ and $1 - \zeta < 1 - \alpha^{-1} < 1$ by (i). Now, $(4-2/p)/(1-\zeta) > 2$ is equivalent to $\zeta + 1 - 1/p > 0$, and $\eta > 1$ and $\gamma > 1$ implies

$$\zeta + 1 - \frac{1}{p} = \gamma^{\ln \zeta / \ln \gamma} - \frac{1-p}{p} = \gamma^{-1/\eta} - \gamma^{-1} > 0.$$

Proof of (iii). Lemma 4.3 shows that $q(\tau) = \Pr\{X > \tau\}$ for a random variable X, namely $X = \min\{x > 0 : d(x) = 0\}$ in the notation of Lemma 4.3, and (iii) follows immediately.

Proof of (iv). Suppose $k \geq 1$ is an integer such that $k < \tau < k+1$. Then (iv) follows from the fact that

$$Q'(\tau) = \frac{d}{d\tau} \sum_{j=0}^{k} \frac{a_j}{(k-j)!}(\tau - k)^{k-j} = \sum_{j=0}^{k-1} \frac{a_j}{(k-j-1)!}(\tau-k)^{k-j-1} = Q(\tau - 1).$$

The case $\tau < 1$ follows from the fact that $Q(x) = 0$ for all $x < 0$.

Proof of (v), (vi). We now need to look closer at the sequence $\{a_k\}$. Let $A(z)$ denote its generating function. From (4.19) we have

$$\begin{aligned} A(z) = 1 + \sum_{k \geq 0} z^k \left(-\frac{e^\alpha}{\alpha}\right) \sum_{j=0}^{k-1} \frac{a_j}{(k-j-1)!} &= 1 - \frac{e^\alpha}{\alpha} \sum_{j=0}^{\infty} a_j z^{j+1} \sum_{k=j+1}^{\infty} \frac{z^{k-j-1}}{(k-j-1)!} \\ &= 1 - \frac{e^\alpha}{\alpha} z e^z A(z) \end{aligned}$$

so $A(z) = 1/(1 + \alpha^{-1} z e^{\alpha + z})$. The sequence $b_k = (-\alpha)^k a_k$ then has generating function

$$B(z) = A(-\alpha z) = \frac{1}{1 - z e^{\alpha(1-z)}}.$$

This has simple poles at $z = 1$ and $z = \zeta$, with residues $1/(\alpha - 1)$ and $\zeta/(\alpha\zeta - 1)$ respectively. Then

$$\overline{B}(z) = \frac{1}{1 - z e^{\alpha(1-z)}} - \frac{1}{(1-\alpha)(1-z)} - \frac{\zeta}{(1-\alpha\zeta)(\zeta - z)}$$

is analytic. Writing $\beta = (1-\alpha\zeta)^{-1}$, the power series representation of $\overline{B}(z)$ is $b_k - 1/(1-\alpha) - \beta/\zeta^k$. Since $\overline{B}(z)$ is analytic, Cauchy's integral formula shows that for any $\varepsilon > 0$,

$$b_k = \frac{1}{1-\alpha} + \frac{\beta}{\zeta^k} + O_k(\varepsilon^k).$$

In the remainder of the proof, fix $0 < \varepsilon < \zeta^{-1}$.

Using (iv) and (4.20) we have, for any integer $k \geq 0$,

$$p(k) = \exp\left\{-\alpha \int_0^k q(x)dx\right\} = \prod_{j=1}^{k} \frac{Q(j-1)e^{\alpha(j-1)}}{Q(j^-)e^{\alpha j}} = \frac{1}{(-\alpha)^k a_k} = \frac{1}{b_k}.$$

and $q(k) = 1 + Q(k-1)/(\alpha Q(k)) = 1 - b_{k-1}/b_k$. If $\alpha < 1$ then $\zeta > 1$ so for integers k,

$$q(k)\zeta^k = \zeta^k\left(1 - \frac{b_{k-1}}{b_k}\right) = \frac{\zeta^k b_k - \zeta^k b_{k-1}}{b_k} = \frac{\beta - \zeta\beta + O(\zeta^{-k})}{\frac{1}{1-\alpha} + O(\zeta^{-k})} = \beta(1-\alpha)(1-\zeta) + O(\zeta^{-k})$$

and we set $\lambda_2 = \beta(1-\alpha)(1-\zeta) = (1-\alpha)(\zeta-1)/(\alpha\zeta-1)$. Recall that $p(k) = 1/b_k$. By Taylor's theorem we have $1/(a+bx) = a^{-1} - ba^{-2}x + O(x^2)$ for any constants $a, b \neq 0$, so with $x = \zeta^{-k}$

$$(p(k) - (1-\alpha))\zeta^k = \left(\frac{1}{\frac{1}{1-\alpha} + \frac{\beta}{\zeta^k} + O(\varepsilon^k)} - (1-\alpha)\right)\zeta^k = -\beta(1-\alpha)^2 + O(\zeta^{-k})$$

and we set $\lambda_1 = -\beta(1-\alpha)^2 = (1-\alpha)^2/(\alpha\zeta-1)$. Here $\zeta > 1 - \alpha^{-1} + \alpha^{-2}$ (from (i)) implies $\lambda_1 < \alpha$.

Suppose $\alpha > 1$. Then $0 < \zeta < \alpha^{-1}$ and

$$\frac{q(k) - (1-\zeta)}{\zeta^k} = \frac{\zeta b_k - b_{k-1}}{\zeta^k b_k} = \frac{\frac{\zeta-1}{1-\alpha} + O(\varepsilon^{k-1})}{\beta + O(\zeta^k)} = \frac{(1-\zeta)(1-\alpha\zeta)}{\alpha-1} + O(\zeta^k)$$

and we set $\lambda_4 = (1-\zeta)(1-\alpha\zeta)/(\alpha-1)$. From the definition (4.18) of ζ we have

$$\frac{p(\tau)}{\zeta^\tau} = \exp\left\{-\alpha\int_0^\tau q(x)dx - \tau\ln\zeta\right\} = \exp\left\{-\alpha\int_0^\tau (q(x) - (1-\zeta))dx\right\}$$

and since $q(\tau)$ decreases toward $1-\zeta$ at an exponential rate, the integral converges as $\tau \to \infty$ and $p(\tau)\zeta^{-\tau}$ is decreasing. Again considering integer values k, we have

$$\frac{p(k)}{\zeta^k} = \frac{1}{b_k\zeta^k} = \frac{1}{\beta + O(\zeta^k)} = 1 - \alpha\zeta + O(\zeta^k)$$

and we set $\lambda_3 = 1 - \alpha\zeta$.

The above shows the asymptotic behaviour of $p(\tau), q(\tau)$ for integer values of τ. Since both functions are monotone, the same asymptotics apply to non-integer values of τ. From (i) it follows that $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ all are positive. □

4.8 Proof of Lemmas 4.3, 4.4

Recall Lemma 4.3.

Lemma 4.3. *For all $\tau \geq 0$, $d(\tau) \sim G(p(\tau), q(\tau))$.*

Proof of Lemma 4.3. The process is a Crump-Mode-Jagers process, a class of processes which were studied in general in companion papers [21], [22]. Define

$$F(s,\tau) = \sum_{k\geq 0} \Pr\{d(\tau) = k\} s^k.$$

In [22] it is shown that the probability generating function satisfies

$$F(s,\tau) = s\exp\left\{\alpha\int_0^\tau (F(s,u) - 1)\ du\right\}, 0 \leq \tau < 1 \tag{4.21}$$

$$F(s,\tau) = \exp\left\{\alpha\int_{\tau-1}^\tau (F(s,u) - 1)\ du\right\}, \tau > 1. \tag{4.22}$$

We will show that $F(s,\tau) = \widetilde{F}(s,\tau)$ where

$$\widetilde{F}(s,\tau) = 1 - q(\tau) + \frac{p(\tau)q(\tau)s}{1 - s(1-p(\tau))} = 1 + \frac{q(\tau)(s-1)}{1-s(1-p(\tau))}$$

with $p(\tau), q(\tau)$ defined in Section 4.2.2. This is the probability generating function of $G(p(\tau), q(\tau))$.

Firstly, when $0 \le \tau < 1$ we plug $q(\tau) = 1$, $p(\tau) = e^{-\alpha\tau}$ into (4.21), and via the integral substitution $w = e^{\alpha u}$,

$$\begin{aligned} s\exp\left\{\alpha\int_0^\tau \left(1 + \frac{s-1}{1-s(1-e^{-\alpha u})} - 1\right)\, du\right\} &= s\exp\left\{\alpha\int_0^\tau \frac{(s-1)e^{\alpha u}}{s+(1-s)e^{\alpha u}}du\right\} \\ &= s\exp\left\{\int_1^{e^{\alpha\tau}} \frac{s-1}{s+w(1-s)}dw\right\} \\ &= s\exp\left\{-\ln(s-(s-1)e^{\alpha\tau})\right\} \\ &= \frac{se^{-\alpha\tau}}{1-s(1-e^{-\alpha\tau})} \end{aligned}$$

confirming that $\widetilde{F}(s,\tau)$ satisfies (4.21).

For $\tau > 1$ we have

$$\exp\left\{\alpha\int_{\tau-1}^{\tau} (\widetilde{F}(s,u) - 1)\, du\right\} = \exp\left\{\alpha\int_{\tau-1}^{\tau} \frac{q(u)(s-1)}{1-s+sp(u)}\right\} du$$

and since $p(u) = \exp\left\{-\alpha\int_0^u q(x)dx\right\}$, the substitution $v(u) = \ln p(u)$ with $dv/du = -\alpha q(u)$ yields

$$\begin{aligned} \alpha\int_{\tau-1}^{\tau} \frac{q(u)(s-1)}{1-s+sp(u)}du &= \int_{v(\tau-1)}^{v(\tau)} \frac{1-s}{1-s+se^v}dv \\ &= \int_{v(\tau-1)}^{v(\tau)} \left(1 - \frac{se^v}{1-s+se^v}\right) dv \end{aligned}$$

and substituting $w = e^v$ gives, as above,

$$\begin{aligned} \int_{v(\tau-1)}^{v(\tau)} \left(1 - \frac{se^v}{1-s+se^v}\right) dv &= v(\tau) - v(\tau-1) + \int_{e^{v(\tau-1)}}^{e^{v(\tau)}} \frac{s}{1-s+sw}dw \\ &= v(\tau) - v(\tau-1) + \ln\left(\frac{1-s+se^{v(\tau-1)}}{1-s+se^{v(\tau)}}\right). \end{aligned}$$

So since $v(u) = \ln p(u)$,

$$\exp\left\{\alpha\int_{\tau-1}^{\tau} (\widetilde{F}(s,u) - 1)\, du\right\} = \frac{p(\tau)}{p(\tau-1)}\frac{1-s+sp(\tau-1)}{1-s+sp(\tau)}. \tag{4.23}$$

We have $1 - q(\tau) = p(\tau)/p(\tau-1)$ (see (4.25)), so

$$\frac{p(\tau)}{p(\tau-1)}\frac{1-s+sp(\tau-1)}{1-s+sp(\tau)} = \frac{(1-s)(1-q(\tau)) + sp(\tau)}{1-s+sp(\tau)} = 1 + \frac{q(\tau)(s-1)}{1-s+sp(\tau)} = \widetilde{F}(s,\tau). \tag{4.24}$$

Now (4.23) and (4.24) imply that $\widetilde{F}(s,\tau)$ satisfies (4.22).

To see that $1 - q(\tau) = p(\tau)/p(\tau - 1)$, recall from (4.20) that $Q(k)/Q(k^-) = -1/(\alpha e^\alpha)$ for integers k, and from Lemma 4.2 (iv) we have $q(\tau) = \alpha^{-1}(Q(\tau)e^{\alpha\tau})'/(Q(\tau)e^{\alpha\tau})$ for non-integer values of τ. So the integral of $q(\tau)$ is $\alpha^{-1}\ln(Q(\tau)e^{\alpha\tau})$, and

$$\begin{aligned}
\frac{p(\tau)}{p(\tau-1)} &= \exp\left\{-\alpha \int_{\tau-1}^{\tau} q(x)dx\right\} \\
&= \exp\left\{-\alpha \int_{\tau-1}^{\lfloor\tau\rfloor} q(x)dx\right\} \exp\left\{-\alpha \int_{\lfloor\tau\rfloor}^{\tau} q(x)dx\right\} \\
&= \frac{Q(\tau-1)e^{\alpha(\tau-1)}}{Q(\lfloor\tau\rfloor^-)e^{\alpha\lfloor\tau\rfloor}} \frac{Q(\lfloor\tau\rfloor)e^{\alpha\lfloor\tau\rfloor}}{Q(\tau)e^{\alpha\tau}} \\
&= \frac{-Q(\tau-1)}{\alpha Q(\tau)} \\
&= 1 - q(\tau).
\end{aligned} \tag{4.25}$$

The last equality comes form the definition of $q(\tau)$. □

Recall Lemma 4.4.

Lemma 4.4. *There exists a constant $\lambda > 0$ such that for $0 \le \tau \le \log_\gamma n$, as $n \to \infty$*

(i) if $\alpha < 1$,

$$Pr\{b(\tau) > \lambda \ln n\} = o\left(\frac{1}{n}\right).$$

(ii) if $\alpha > 1$,

$$Pr\left\{b(\tau) > \lambda n^{1/\eta} \ln n\right\} = o\left(\frac{1}{n}\right)$$

where $\eta = -\ln\gamma/\ln\zeta > 2$.

(iii) If $\alpha \ne 1$ then $d(\tau) \ge \lfloor b(\tau)/(\lambda \log_\gamma^2 n)\rfloor$ for all $0 \le \tau \le \log_\gamma n$ with probability $1 - o(n^{-1})$.

Proof of Lemma 4.4. Each Poisson process has lifetime exactly 1, so

$$\sum_{k=0}^{\lfloor\tau\rfloor} d(k) \le b(\tau) \le \sum_{k=0}^{\lceil\tau\rceil} d(k)$$

and in particular,

$$b(\tau) \le \lceil\tau\rceil \max_{0\le k\le\lceil\tau\rceil} d(k). \tag{4.26}$$

From Lemma 4.3 we have

$$\Pr\{d(\tau) > \ell\} = q(\tau)(1 - p(\tau))^\ell.$$

For $\alpha < 1$, Lemma 4.2 (i), (v) imply that $1 - p(\tau) \le \alpha$, so

$$\Pr\left\{\max_{0\le k\le\lceil\tau\rceil} d(k) > -2\log_\alpha n\right\} \le \lceil\tau\rceil \alpha^{-2\log_\alpha n} = o(n^{-1}).$$

For $\alpha > 1$, Lemma 4.2 (i), (vi) imply

$$\Pr\left\{\max_{0\le k\le\lceil\tau\rceil} d(k) > \lambda n^{1/\eta}\ln n\right\} \le \lceil\tau\rceil(1-\lambda_3\zeta^{\lceil\tau\rceil})^{\lambda n^{1/\eta}\ln n} \le \lceil\tau\rceil\exp\left\{-\lambda\lambda_3\zeta^{\lceil\tau\rceil}n^{1/\eta}\ln n\right\}$$

and since $\tau \le \log_\gamma n$ and $\zeta^{\log_\gamma n}n^{1/\eta} = 1$, this is $o(n^{-1})$ for λ large enough.

Assertion (iii) follows from (i) for $\alpha < 1$. Suppose $\alpha > 1$. The claim will follow from showing that we can choose $A, B > 0$ so that if $\tau \le \log_\gamma n$,

$$\Pr\left\{\exists x \in [0,\tau] : d(x) \ge A\log_\gamma n \text{ and } d(\tau) \le d(x)/B\right\} = o(n^{-1}). \tag{4.27}$$

Indeed, suppose $b(\tau) \ge A(\log_\gamma n)^2$. Then by (4.26) there exists some $x < \tau$ for which $d(x) \ge b(\tau)/\tau \ge A\log_\gamma n$. It will follow from (4.27) that $d(\tau) \ge b(\tau)/(B\tau) \ge AB^{-1}b(\tau)/\log_\gamma n$ with probability $1 - o(n^{-1})$. If $b(\tau) < A(\log_\gamma n)^2$ we choose $\lambda > A$ so that $d(\tau) \ge 0 = \lfloor b(\tau)/(C\log_\gamma^2 n)\rfloor$.

If $x' < \tau$ is such that $d(x') \ge A\log_\gamma n$ Poisson processes are active, then either (i) at least $d(x')/2$ of the processes are still active at time $x' + 1/2$, or (ii) at least $d(x')/2$ of the processes were active at time $x' - 1/2$. In either case, there exists an $x < \tau$ such that $d(x) \ge \frac{A}{2}\log_\gamma n$ and at least $d(x)/2$ processes are active at time $x + 1/2$. If $x \ge \tau - 1/2$ then $d(\tau) \ge d(x)/2$, so suppose $x < \tau - 1/2$.

Suppose $\mathcal{P}_i$ is a process which is active at times x and $x + 1/2$. The probability that $\mathcal{P}_i$ has at least one arrival in $(x, x + 1/2)$ is $1 - e^{-\alpha/2}$. Suppose $\mathcal{P}_i$ has an arrival at time $x_i \in (x, x + 1/2)$. Then the process starting at time x_i can be seen as the initial process of a CMJ process $\mathcal{C}_i$ on $[x_i, \tau]$. Since $\alpha > 1$, the probability that $\mathcal{C}_i$ is active at time τ is $q(\tau - x_i) \ge 1 - \zeta$ (see Lemma 4.2 (iii) and (vi)). In other words, if X_i is the indicator variable for $\mathcal{P}_i$ having an active descendant at time τ, then $\Pr\{X_i = 1\} \ge (1 - e^{-\alpha/2})(1 - \zeta)$. This is true independently for the $d(x)/2$ processes $\mathcal{P}_1, \ldots, \mathcal{P}_{d(x)/2}$ active at time x and $x + 1/2$, and we have $d(\tau) \ge X_1 + \cdots + X_{d(x)/2}$. Choosing A, B large enough, Hoeffding's inequality [57] shows that $d(\tau) \ge d(x)/B$ with probability $1 - o(n^{-1})$. This finishes the proof. □

4.9 Concluding remarks

The main computational task in improving the results of this paper is in estimating integral involving $p(\tau), q(\tau)$ and $\gamma^{-\tau}$. To find the exact number of vertices of degree k for $k = O(1)$, one needs to calculate integrals involving terms of the form $\gamma^{-\tau}q(\tau)p(\tau)(1 - p(\tau))^{k-1}$, and this is difficult to do in any generality. Integrals involving $p(\tau), q(\tau)$ and $\gamma^{-\tau}$ also appear when looking for small components, which prevented us from finding the exact size of the giant component.

Chapter 5

Minimum matching in a random graph with random costs

This chapter corresponds to [38].

Abstract

Let $G_{n,p}$ be the standard Erdős-Rényi-Gilbert random graph and let $G_{n,n,p}$ be the random bipartite graph on $n+n$ vertices, where each $e \in [n]^2$ appears as an edge independently with probability p. For a graph $G = (V, E)$, suppose that each edge $e \in E$ is given an independent exponential rate one cost X_e. Let $C(G)$ denote the random variable equal to the length of the minimum cost perfect matching if G contains at least one perfect matching and let $C(G) = 0$ otherwise. Let $\mu(G) = \mathbf{E}[C(G)]$. We show that if $np \gg \log^2 n$ and $G = G_{n,n,p}$ then w.h.p. $\mu(G) \approx \frac{\pi^2}{6p}$. This generalises the well-known result for the case $G = K_{n,n}$, where $p = 1$. We also show that if $G = G_{n,p}$ then $\mu(G_{n,p}) \approx \frac{\pi^2}{12p}$ w.h.p. along with concentration results for both types of random graph.

5.1 Introduction

There are many results concerning the optimal value of combinatorial optimization problems with random costs. Sometimes the costs are associated with n points generated uniformly at random in the unit square $[0,1]^2$. In which case the most celebrated result is due to Beardwood, Halton and Hammersley [8] who showed that the minimum length of a tour through the points a.s. grew as $\beta n^{1/2}$ for some still unknown β. For more on this and related topics see Steele [84].

The optimisation problem in [8] is defined by the distances between the points. So, it is defined by a random matrix where the entries are highly correlated. There have been many examples considered where the matrix of costs contains independent entries. Aside from the Travelling Salesperson Problem, the most studied problems in combinatorial optimization are perhaps, the shortest path problem; the minimum spanning tree problem and the matching problem. As a first example, consider the shortest path problem in the complete graph K_n where the edge lengths are independent exponential random variables with rate 1. We denote the exponential random variable with rate λ by $E(\lambda)$. Thus $\mathbf{Pr}(E(\lambda) \geq x) = e^{-\lambda x}$ for $x \in \boldsymbol{R}$. Janson [59] proved (among other things) that if $X_{i,j}$ denotes the shortest distance between vertices i, j in this model then $\mathbf{E}[X_{1,2}] = \frac{H_n}{n}$ where $H_n = \sum_{i=1}^{n} \frac{1}{i}$.

As far as the spanning tree problem is concerned, the first relevant result is due to Frieze [42]. He showed that if the edges of the complete graph are given independent uniform $[0,1]$ edge weights, then the (random) minimum length of a spanning tree L_n satisfies $\mathbf{E}[L_n] \to \zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3}$ as $n \to \infty$. Further results on this question can be found in Steele [83], Janson [58], Beveridge, Frieze and McDiarmid [9], Frieze, Ruszinko and Thoma [52] and Cooper, Frieze, Ince, Janson and Spencer [19].

In the case of matchings, the nicest results concern the the minimum cost of a matching in a randomly edge-weighted copy of the complete bipartite graph $K_{n,n}$. If C_n denotes the (random) minimum cost of a perfect matching when edges are given independent exponential $E(1)$ random variables then the story begins with Walkup [88] who proved that $\mathbf{E}[C_n] \leq 3$. Later Karp [62] proved that $\mathbf{E}[C_n] \leq 2$. Aldous [3, 4] proved that $\lim_{n\to\infty} \mathbf{E}[C_n] = \zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2}$. Parisi [78] conjectured that in fact $\mathbf{E}[C_n] = \sum_{k=1}^{n} \frac{1}{k^2}$. This was proved independently by Linusson and Wästlund [68] and by Nair, Prabhakar and Sharma [74]. A short elegant proof was given by Wästlund [90, 92].

In this paper we replace the complete graphs $K_{n,n}$ and K_n by random graphs. More precisely each graph G gives rise to a probability space $\Omega_G = \prod_{e \in G} X_e$ where for each $e \in E(G)$ we have $\mathbf{Pr}(X_e \geq x) = e^{-x}$ for $x \geq 0$ i.e. each edge has an independent exponential mean one cost. For each $\mathbf{X} \in \Omega_G$ we let $G_{\mathbf{X}}$ denote the corresponding edge-weighted graph. The random variable $C(G, \mathbf{X})$ is equal to the length of the minimum cost perfect matching in $G_{\mathbf{X}}$, if G contains at least one perfect matching and is equal to zero otherwise. Let $\mu(G) = \mathbf{E}_{\mathbf{X}}[C(G, \mathbf{X})]$.

Let $G_{n,n,p}$ be the random bipartite graph on $n+n$ vertices, where each $e \in [n]^2$ appears as an edge independently with probability p. The next theorem deals with the random variable $\mu(G_{n,n,p})$.

Theorem 5.1. *If $d = np = \omega(n) \log^2 n$ where $\omega(x) \to \infty$ with x, then $\boldsymbol{E}[\mu(G_{n,n,p})] \approx \frac{\pi^2}{6p}$.*

To be clear, the expectation in the theorem is taken over random $G = G_{n,n,p}$ and random $\mathbf{X}$.

In addition we will in fact show that $\mu(G_{n,n,p})$ will be highly concentrated around $\frac{\pi^2}{6p}$. In this case the probability is over the graph G alone. Here $\mu(G)$ is a function of G alone.

Here $A_n \approx B_n$ iff $A_n = (1 + o(1))B_n$ as $n \to \infty$ and the event $\mathcal{E}_n$ occurs with high probability (w.h.p.) if $\mathbf{Pr}(\mathcal{E}_n) = 1 - o(1)$ as $n \to \infty$.

In the case of $G_{n,p}$ we prove

Theorem 5.2. *If $d = np = \omega(n) \log^2 n$ where $\omega(x) \to \infty$ with x, then $\boldsymbol{E}[\mu(G_{n,p})] \approx \frac{\pi^2}{12p}$.*

Applying results of Talagrand [86] we can prove the following concentration result.

Theorem 5.3. *Let $\varepsilon > 0$ be fixed, then*

$$\mathbf{Pr}\left(\left|\mu(G_{n,n,p}) - \frac{\pi^2}{6p}\right| \geq \frac{\varepsilon}{p}\right) \leq n^{-K}, \quad \mathbf{Pr}\left(\left|\mu(G_{n,p}) - \frac{\pi^2}{12p}\right| \geq \frac{\varepsilon}{p}\right) \leq n^{-K}$$

for any constant $K > 0$ and n large enough.

In this theorem the probabilities are with respect to random $G_{n,n,p}$ or $G_{n,p}$ alone.

In the paper [9] on the minimum spanning tree problem, the complete graph was replaced by a d-regular graph G. Under some mild expansion assumptions, it was shown that if $d \to \infty$ then $\zeta(3)$ can be replaced asymptotically by $n\zeta(3)/d$.

Now consider a d-regular bipartite graph G on $2n$ vertices. Here $d = d(n) \to \infty$ as $n \to \infty$. Each edge e is assigned a cost $w(e)$, each independently chosen according to the exponential distribution $E(1)$. Denote the total cost of the minimum-cost perfect matching by $C(G)$.

We conjecture the following (under some possibly mild restrictions):

Conjecture 5.1. *Suppose $d = d(n) \to \infty$ as $n \to \infty$. For any n-vertex, d-regular bipartite G,*

$$\mu(G) \approx \frac{n}{d}\frac{\pi^2}{6}.$$

In this paper we prove the conjecture for random graphs and random bipartite graphs.

5.2 Proof of Theorem 5.1

We find that the proofs in [90], [92] can be adapted to our current situation. Suppose that the vertices of $G = G_{n,n,p}$ are denoted $A = \{a_i, i \in [n]\}$ and $B = \{b_j, j \in [n]\}$. We will use the notation (a, b) for edges of G, where $a \in A$ and $b \in B$. Let $C(n, r)$ denote the cost of the minimum cost matching

$$M_r = \{(a_i, \phi_r(a_i)) : i = 1, 2, \ldots, r\} \text{ of } A_r = \{a_1, a_2, \ldots, a_r\} \text{ into } B.$$

We will prove that w.h.p.

$$\mathbf{E}\left[C(n, r) - C(n, r-1)\right] \approx \frac{1}{p}\sum_{i=0}^{r-1}\frac{1}{r(n-i)}.$$

for $r = 1, 2, \ldots, n - m$ where

$$m = \frac{n}{\omega^{1/2}\log n}.$$

Using this we argue that w.h.p.

$$\mathbf{E}\left[C(G)\right] = \mathbf{E}\left[C(n, n)\right] = \mathbf{E}\left[C(n, n) - C(n, n-m+1)\right] + \frac{1+o(1)}{p}\sum_{r=1}^{n-m}\sum_{i=0}^{r-1}\frac{1}{r(n-i)}. \tag{5.1}$$

We will then show that

$$\sum_{r=1}^{n-m}\sum_{i=0}^{r-1}\frac{1}{r(n-i)} \approx \sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}. \tag{5.2}$$

$$\mathbf{E}\left[C(n, n) - C(n, n-m+1)\right] = o(p^{-1}) \quad w.h.p. \tag{5.3}$$

Theorem 5.1 follows from these two statements.

5.2.1 Outline of the proof

We first argue (Lemma 5.1) that w.h.p. vertices $v \in A_r$ have aproximately $(n-r)p$ neighbors in $B \setminus B_r$, where $B_r = \phi_r(A_r)$. Then comes the beautiful idea of adding a vertex b_{n+1} and joining it to every vertex in A by an edge of cost $E(\lambda)$. The heart of the proof is in Lemma 5.2 that relates $\mathbf{E}\left[C(n,r) - C(n,r-1)\right]$ in a precise way to the probability that b_{n+1} is covered by M_r^*, the minimum cost matching of A_r into $B^* = B \cup \{b_{n+1}\}$. The proof now focuses on estimating this probability $P(n,r)$. If r is not too close to n then this probability can be estimated (see (5.6)) by careful conditioning and the use of properties of the exponential distribution. From thereon, it is a matter of analysing the consequences of the estimate for $\mathbf{E}\left[C(n,r) - C(n,r-1)\right]$ in (5.7). The final part of the proof involves showing (Lemma 5.5) that $\mathbf{E}\left[C(n,n) - C(n-m+1)\right]$ is insignificant. This essentially boils down to showing that w.h.p. no edge in the minimum cost matching has cost more than $O(\log n/(np))$.

5.2.2 Proof details

We use the Chernoff bounds to bound degrees. For reference we use the following: Let $B(n,p)$ denote the binomial random variable with parameters n, p. Then for $0 \leq \varepsilon \leq 1$ and $\alpha > 0$,

$$\mathbf{Pr}(B(n,p) \leq (1-\varepsilon)np) \leq e^{-\varepsilon^2 np/2}. \tag{5.4}$$

$$\mathbf{Pr}(B(n,p) \leq (1+\varepsilon)np) \leq e^{-\varepsilon^2 np/3}. \tag{5.5}$$

$$\mathbf{Pr}(B(n,p) \geq \alpha np) \leq \left(\frac{e}{\alpha}\right)^{\alpha np}.$$

For $v \in A$ let $d_r(v) = |\{w \in B \setminus B_r : (v,w) \in E(G)\}|$. Then we have the following lemma:

Lemma 5.1.

$$|d_r(v) - (n-r)p| \leq \omega^{-1/5}(n-r)p \text{ w.h.p. for } v \in A,\ 0 \leq r \leq n-m.$$

Proof. For the purposes of this lemma, we construct $G_{\mathbf{X}}$ by generating, for $i = 1, 2, \ldots, n$, the edges incident with a_i, along with their costs. Note that we can now determine the function ϕ_r and hence B_r, as soon as we have dealt with $a_1, a_2, \ldots, a_r$. At this point we have not exposed any of the edges of $G_{n,n,p}$ incident with $A \setminus A_r$. We condition on the set B_r. It follows then that $d_r(v), v \notin A_r$ is distributed as $B(n-r,p)$. Applying the Chernoff bounds (5.4), (5.5) with $\varepsilon = \omega^{-1/5}$ we obtain

$$\begin{aligned}\mathbf{Pr}(\exists v : |d_r(v) - (n-r)p| \geq \omega^{-1/5}(n-r)p) &\leq 2ne^{-\omega^{-2/5}(n-r)p/3} \\ &\leq 2n^{1-\omega^{1/10}/3}.\end{aligned}$$

□

We can now use the ideas of [90], [92]. We add a special vertex b_{n+1} to B, with edges to all n vertices of A. Each edge adjacent to b_{n+1} is assigned an $E(\lambda)$ cost independently, $\lambda > 0$. We now consider M_r to be a minimum cost matching of A_r into $B^* = B \cup \{b_{n+1}\}$. We denote this matching by M_r^* and we let B_r^* denote the corresponding set of vertices of B^* that are covered by M_r^*.

Define $P(n,r)$ as the normalized probability that b_{n+1} participates in M_r^*, i.e.

$$P(n,r) = \lim_{\lambda \to 0} \frac{1}{\lambda} \mathbf{Pr}(b_{n+1} \in B_r^*).$$

Its importance lies in the following lemma:

Lemma 5.2.

$$\boldsymbol{E}[C(n,r) - C(n,r-1)] = \frac{1}{r}P(n,r).$$

Proof. For the argument that follows, we need to argue that the vertices $\widehat{B}$ of B covered by the cheapest $(A_r \setminus \{a_i\})$-assignment are a subset of B_r. Using the symmetry of the problem, we argue instead that $B_{r-1} \subseteq B_r$ with probability one. So, consider the symmetric difference $D = M_{r-1} \oplus M_r$. In general, it consists of a collection of alternating paths and cycles from A_r to B. The cost of a path/cycle is the difference between the cost of its M_r-edges and its M_{r-1}-edges. Now with probability one these will be non-zero. If a cycle has positive cost then it means that M_r can be improved by reversing the inclusion of edges in the cycle and if the cost is negative then M_{r-1} can be improved. Hence with probability one there are no cycles in D. An alternating path from $a_j, j \neq r$ to $B \setminus B_{r-1}$ leads to the same conclusion about the optimality of M_{r-1}, M_r. Thus with probability one, D consists of an augmenting path from a_r to $B \setminus B_{r-1}$ and this implies that $B_{r-1} \subseteq B_r$.

Let $X = C(n,r)$ and let $Y = C(n,r-1)$. Fix $i \in [r]$ and let w be the cost of the edge (a_i, b_{n+1}), and let I denote the indicator variable for the event that the cost of the cheapest A_r-assignment that contains this edge is smaller than the cost of the cheapest A_r-assignment that does not use b_{n+1}. In other words, I is the indicator variable for the event $\{\widehat{Y} + w < X\}$, where $\widehat{Y}$ is the minimum cost of a matching from $A_r \setminus \{a_i\}$ to B. This uses the fact that $\widehat{B} \subseteq B_r$ by assuming that after deleting the edge (a_i, b_{n+1}) from M_r we have a matching from $A_r \setminus \{a_i\}$ to $\widehat{B}$ of cost $\widehat{Y}$. Note that by symmetry Y and $\widehat{Y}$ have the same distribution. For this symmetry argument to be valid, we need to be dealing with the probability space of G and $\mathbf{X}$.

If $(a_i, b_{n+1}) \in M_r^*$ then $w < X - \widehat{Y}$. Conversely, if $w < X - \widehat{Y}$ and no other edge from b_{n+1} has cost smaller than $X - \widehat{Y}$, then $(a_i, b_{n+1}) \in M_r^*$, and when $\lambda \to 0$, the probability that there are two distinct edges from b_{n+1} of cost smaller than $X - \widehat{Y}$ is of order $O(\lambda^2)$. Indeed, let $\mathcal{F}$ denote the existence of two distinct edges from b_{n+1} of cost smaller than X and let $\mathcal{F}_{i,j}$ denote the event that (a_i, b_{n+1}) and (a_j, b_{n+1}) both have cost smaller than X. Then by symmetry,

$$\mathbf{Pr}(\mathcal{F}) \leq n^2 \mathbf{E}_X[\mathbf{Pr}(\mathcal{F}_{1,2} \mid X)] = n^2 \mathbf{E}[(1 - e^{-\lambda X})^2] \leq n^2 \mathbf{E}[X^2]\lambda^2,$$

and since $\mathbf{E}\left[X^2\right]$ is finite and independent of λ, this is $O(\lambda^2)$.

Since w is $E(\lambda)$ distributed, as $\lambda \to 0$ we have,

$$\mathbf{E}[X - Y] = \mathbf{E}\left[X - \widehat{Y}\right] = \frac{d}{d\lambda}\mathbf{E}[I]\bigg|_{\lambda=0} = \lim_{\lambda \to 0} \frac{1}{\lambda}\mathbf{Pr}(w < X - \widehat{Y}) = \frac{1}{r}P(n,r).$$

The factor $1/r$ comes from each $i \in [r]$ being equally likely to be incident to the matching edge containing b_{n+1}, if it exists. $\square$

We now proceed to estimate $P(n,r)$.

Lemma 5.3. *Suppose $r < n - m$. Then*

$$\mathbf{Pr}(b_{n+1} \in B_r^* \mid b_{n+1} \notin B_{r-1}^*) = \frac{\lambda}{p(n-r+1)(1+\varepsilon_{r,n}) + \lambda} \tag{5.6}$$

where $|\varepsilon_{r,n}| \leq \omega^{-1/5}$.

Proof. Assume that $b_{n+1} \notin B^*_{r-1}$. M^*_r is obtained from M^*_{r-1} by finding an augmenting path $P = (a_r, \ldots, a_\sigma, b_\tau)$ from a_r to $B^* \setminus B^*_{r-1}$ of minimum additional cost. Let $\alpha = W(\sigma, \tau)$ denote the cost of (a_σ, b_τ). We condition on (i) σ, (ii) the lengths of all edges other than $(a_\sigma, b_j), b_j \in B^* \setminus B^*_{r-1}$ and (iii) $\min\left\{W(\sigma, j) : b_j \in B^* \setminus B^*_{r-1}\right\} = \alpha$. With this conditioning $M_{r-1} = M^*_{r-1}$ will be fixed and so will $P' = (a_r, \ldots, a_\sigma)$. We can now use the following fact: Let $X_1, X_2, \ldots, X_M$ be independent exponential random variables of rates $\alpha_1, \alpha_2, \ldots, \alpha_M$. Then the probability that X_i is the smallest of them is $\alpha_i/(\alpha_1 + \alpha_2 + \cdots + \alpha_M)$. Furthermore, the probability stays the same if we condition on the value of $\min\{X_1, X_2, \ldots, X_M\}$. Thus

$$\mathbf{Pr}(b_{n+1} \in B^*_r \mid b_{n+1} \notin B^*_{r-1}) = \frac{\lambda}{d_{r-1}(a_\sigma) + \lambda}.$$

□

Corollary 5.1. *If $r \leq n - m$ then*

$$P(n, r) = \frac{1}{p}\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-r+1}\right)(1 + \varepsilon_{r,n}) \tag{5.7}$$

where $|\varepsilon_{r,n}| \leq \omega^{-1/5}$.

Proof. Let $\nu(j) = p(n-j)(1 + \varepsilon_{j,n})$, $|\varepsilon_{j,n}| \leq \omega^{-1/5}$. Then

$$\begin{aligned}
\mathbf{Pr}(b_{n+1} \in B^*_r) &= 1 - \frac{\nu(0)}{\nu(0) + \lambda} \cdot \frac{\nu(1)}{\nu(1) + \lambda} \cdots \frac{\nu(r-1)}{\nu(r-1) + \lambda} \\
&= 1 - \left(1 + \frac{\lambda}{\nu(0)}\right)^{-1} \cdots \left(1 + \frac{\lambda}{\nu(r-1)}\right)^{-1} \\
&= \left(\frac{1}{\nu(0)} + \frac{1}{\nu(1)} + \cdots + \frac{1}{\nu(r-1)}\right)\lambda + O(\lambda^2) \\
&= \frac{1}{p}\left(\frac{1}{n(1+\varepsilon_{0,n})} + \frac{1}{(n-1)(1+\varepsilon_{1,n})} + \cdots + \frac{1}{(n-r+1)(1+\varepsilon_{r-1,n})}\right)\lambda + O(\lambda^2)
\end{aligned}$$

and each error factor satisfies $|1 - 1/(1 + \varepsilon_{j,n})| \leq \omega^{-1/5}$. Letting $\lambda \to 0$ gives the lemma. □

Lemma 5.4. *If $r \leq n - m$ then*

$$\boldsymbol{E}\left[C(n, r) - C(n, r-1)\right] = \frac{1 + o(1)}{rp} \sum_{i=0}^{r-1} \frac{1}{n-i}.$$

Proof. This follows from Lemma 5.2 and Corollary 5.1. □

This confirms (5.1) and we turn to (5.2). We use the following expression from Young [95].

$$\sum_{i=1}^{n} \frac{1}{i} = \log n + \gamma + \frac{1}{2n} + O(n^{-2}), \qquad \text{where } \gamma \text{ is Euler's constant.} \tag{5.8}$$

Let $m_1 = \omega^{1/4} m$. Observe first that

$$\begin{aligned}
\sum_{i=0}^{m_1} \frac{1}{n-i} \sum_{r=i+1}^{n-m} \frac{1}{r} &\leq O\left(\frac{\log n}{n^{1/4}}\right) + \sum_{i=n^{3/4}}^{m_1} \frac{1}{n-i} \sum_{r=i+1}^{n-m} \frac{1}{r} \\
&\leq o(1) + \frac{1}{n-m_1} \sum_{i=n^{3/4}}^{m_1} \left(\log\left(\frac{n}{i}\right) + \frac{1}{2(n-m)} + O(n^{-3/2})\right) \\
&\leq o(1) + \frac{2}{n} \log\left(\frac{n^{m_1}}{m_1!}\right) \\
&\leq o(1) + \frac{2m_1}{n} \log\left(\frac{ne}{m_1}\right) \\
&= o(1).
\end{aligned}$$

Then,

$$\begin{aligned}
\sum_{r=1}^{n-m} \sum_{i=0}^{r-1} \frac{1}{r(n-i)} &= \sum_{i=0}^{n-m-1} \frac{1}{n-i} \sum_{r=i+1}^{n-m} \frac{1}{r}, \\
&= \sum_{i=m_1}^{n-m-1} \frac{1}{n-i} \sum_{r=i+1}^{n-m} \frac{1}{r} + o(1), \\
&= \sum_{i=m_1}^{n-m-1} \frac{1}{n-i} \left(\log\left(\frac{n-m}{i}\right) + \frac{1}{2(n-m)} - \frac{1}{2i} + O(i^{-2})\right) + o(1), \\
&= \sum_{i=m_1}^{n-m-1} \frac{1}{n-i} \log\left(\frac{n-m}{i}\right) + o(1), \\
&= \sum_{j=m+1}^{n-m_1} \frac{1}{j} \log\left(\frac{n-m}{n-j}\right) + o(1), \qquad (5.9) \\
&= \int_{x=m+1}^{n-m_1} \frac{1}{x} \log\left(\frac{n-m}{n-x}\right) dx + o(1).
\end{aligned}$$

We can replace the sum in (5.9) by an integral because the sequence of summands is unimodal and the terms are all $o(1)$.

Continuing, we have

$$\begin{aligned}
&\int_{x=m+1}^{n-m_1} \frac{1}{x} \log\left(\frac{n-m}{n-x}\right) dx \\
&= -\int_{x=m+1}^{n-m_1} \frac{1}{x} \log\left(1 - \frac{x-m}{n-m}\right) dx \\
&= \sum_{k=1}^{\infty} \int_{x=m+1}^{n-m_1} \frac{1}{x} \frac{(x-m)^k}{k(n-m)^k} dx \\
&= \sum_{k=1}^{\infty} \int_{y=1}^{n-m-m_1} \frac{1}{y+m} \frac{y^k}{k(n-m)^k} dy. \qquad (5.10)
\end{aligned}$$

Observe next that for every $k \geq 1$

$$\int_{y=1}^{n-m-m_1} \frac{1}{y+m} \frac{y^k}{k(n-m)^k} dy \leq \int_{y=1}^{n-m-m_1} \frac{y^{k-1}}{k(n-m)^k} dy \leq \frac{1}{k^2}.$$

So,

$$0 \leq \sum_{k=\log n}^{\infty} \int_{x=m+1}^{n-m_1} \frac{1}{x} \frac{(x-m)^k}{k(n-m)^k} dx \leq \sum_{k=\log n}^{\infty} \frac{1}{k^2} = o(1). \tag{5.11}$$

If $1 \leq k \leq \log n$ then we write

$$\int_{y=1}^{n-m-m_1} \frac{1}{y+m} \frac{y^k}{k(n-m)^k} dy = \int_{y=1}^{n-m-m_1} \frac{(y+m)^{k-1}}{k(n-m)^k} dy + \int_{y=1}^{n-m-m_1} \frac{y^k-(y+m)^k}{(y+m)k(n-m)^k} dy.$$

Now

$$\int_{y=1}^{n-m-m_1} \frac{(y+m)^{k-1}}{k(n-m)^k} dy = \frac{1}{k^2} \frac{(n-m_1)^k-(m+1)^k}{(n-m)^k} = \frac{1}{k^2} + O\left(\frac{1}{k\omega^{1/4}\log n}\right). \tag{5.12}$$

If $k=1$ then our choice of m implies that

$$\int_{y=1}^{n-m-m_1} \frac{(y+m)^k-y^k}{(y+m)k(n-m)^k} dy \leq \frac{m\log(n-m_1)}{n-m} = o(1).$$

And if $2 \leq k \leq \log n$ then

$$\begin{aligned}
\int_{y=1}^{n-m-m_1} \frac{(y+m)^k-y^k}{(y+m)k(n-m)^k} dy &= \sum_{l=1}^{k} \int_{y=1}^{n-m-m_1} \binom{k}{l} \frac{y^{k-l}m^l}{(y+m)k(n-m)^k} dy \\
&\leq \sum_{l=1}^{k} \int_{y=0}^{n-m-m_1} \binom{k}{l} \frac{y^{k-l-1}m^l}{k(n-m)^k} dy \\
&= \sum_{l=1}^{k} \binom{k}{l} \frac{m^l(n-m-m_1)^{k-l}}{k(k-l)(n-m)^k} \qquad (5.13)\\
&= O\left(\frac{km}{k(k-1)n}\right) = O\left(\frac{1}{k\omega^{1/2}\log n}\right). \qquad (5.14)
\end{aligned}$$

To go from (5.13) to (5.14) we argue that if the summand in (5.13) is denoted by u_l then $u_{l+1}/u_l = O(1/\omega^{1/2})$ for $1 \leq l \leq \log n$. Hence the sum is $O(u_1)$.

It follows that

$$0 \leq \sum_{k=1}^{\log n} \int_{y=1}^{n-m-m_1} \frac{(y+m)^k-y^k}{(y+m)k(n-m)^k} dy = o(1) + O\left(\sum_{k=2}^{\log n} \frac{1}{k\omega^{1/2}\log n}\right) = o(1). \tag{5.15}$$

Equation (5.2) now follows from (5.10), (5.11), (5.12) and (5.15).

Turning to (5.3) we prove the following lemma:

Lemma 5.5. *If $r \geq n-m$ then $0 \leq C(n,r+1) - C(n,r) = O\left(\frac{\log n}{np}\right)$.*

This will prove that

$$0 \leq \mathbf{E}\left[C(n,n) - C(n-m+1)\right] = O\left(\frac{m\log n}{np}\right) = O\left(\frac{n}{\omega^{1/2}np}\right) = o\left(\frac{1}{p}\right)$$

and complete the proof of (5.3) and hence Theorem 5.1.

5.2.3 Proof of Lemma 5.5

Let $w(e)$ denote the weight of edge e in G. Let $V_r = A_{r+1} \cup B$ and let G_r be the subgraph of G induced by V_r. For a vertex $a \in A_{r+1}$ order the neighbors $u_1, u_2, \ldots,$ of a in B so that $w(a, u_i) \leq w(a, u_{i+1})$. Similarly, if $b \in B$ order the neighbors $u_1, u_2, \ldots,$ of b in A_{r+1} so that $w(u_i, b) \leq w(u_{i+1}, b)$. For $v \in V_r$, define the k-neighborhood $N_k(v) = \{u_1, u_2, \ldots, u_k\}$. This is defined independently from any matchings between A and B.

Let the k-neighborhood of a set be the union of the k-neighborhoods of its vertices. In particular, for $S \subseteq A_{r+1}$, $T \subseteq B$,

$$N_k(S) = \{b \in B :\ \exists a \in S : y \in N_k(a)\},$$
$$N_k(T) = \{a \in A_{r+1} :\ \exists b \in T : a \in N_k(b)\}.$$

Given a function ϕ_r defining a matching M of A_r into B, we define the following digraph: let $\vec{\Gamma}_r = (V_r, \vec{X})$ where $\vec{X}$ is an orientation of

$$X = \{\{a,b\} \in G : a \in A_{r+1}, b \in N_{40}(a)\} \cup \{\{a,b\} \in G : b \in B, a \in N_{40}(b)\} \cup \{(\phi_r(a_i), a_i) : i = 1, 2, \ldots, r\}.$$

An edge $e \in M$ is oriented from B to A and has weight $-w(e)$. The remaining edges are oriented from A to B and have weight equal to their weight in G.

The arcs of directed paths in $\vec{\Gamma}_r$ are alternately forwards $A \to B$ and backwards $B \to A$ and so they correspond to alternating paths with respect to the matching M. It helps to know (Lemma 5.6, next) that given $a \in A_{r+1}, b \in B$ we can find an alternating path from a to b with $O(\log n)$ edges. The ab-diameter will be the maximum over $a \in A_{r+1}, b \in B$ of the length of a shortest alternating path from a to b.

Lemma 5.6. *W.h.p., for every ϕ_r, the (unweighted) ab-diameter of $\vec{\Gamma}_r$ is at most $k_0 = \lceil 3 \log_4 n \rceil$.*

Proof. For $S \subseteq A_{r+1}$, $T \subseteq B$, let

$$\Lambda(S) = \{b \in B :\ \exists a \in S \text{ such that } (a,b) \in \vec{X}\},$$
$$\Lambda(T) = \{a \in A_{r+1} :\ \exists b \in T \text{ such that } (a,b) \in \vec{X}\}.$$

We first prove an expansion property: that w.h.p., for all $S \subseteq A_{r+1}$ with $|S| \leq \lceil n/5 \rceil$, $|\Lambda(S)| \geq 4|S|$. (Note that $\Lambda(S), \Lambda(T)$ are defined by edges oriented from A to B and so do not depend on ϕ_r.)

$$\begin{aligned}
\mathbf{Pr}(\exists S :\ |S| \leq \lceil n/5 \rceil,\, |\Lambda(S)| < 4|S|) &\leq o(1) + \sum_{s=1}^{\lceil n/5 \rceil} \binom{r+1}{s} \binom{n}{4s} \left(\frac{\binom{4s}{40}}{\binom{n}{40}} \right)^s \\
&\leq \sum_{s=1}^{\lceil n/5 \rceil} \left(\frac{ne}{s}\right)^s \left(\frac{ne}{4s}\right)^{4s} \left(\frac{4s}{n}\right)^{40s} \\
&= \sum_{s=1}^{\lceil n/5 \rceil} \left(\frac{e^5 4^{36} s^{35}}{n^{35}} \right)^s \\
&= o(1).
\end{aligned}$$

Explanation: *The $o(1)$ term accounts for the probability that each vertex has at least 40 neighbors in $\vec{\Gamma}_r$. Condition on this. Over all possible ways of choosing s vertices and $4s$ "targets", we take the probability that for each of the s vertices, all 40 out-edges fall among the $4s$ out of the n possibilities.*

Similarly, w.h.p., for all $T \subseteq B$ with $|T| \leq \lceil n/5 \rceil$, $|\Lambda(T)| \geq 4|T|$. Thus by the union bound, w.h.p. both these events hold. In the remainder of this proof we assume that we are in this "good" case, in which all small sets S and T have large vertex expansion.

Now, choose an arbitrary $a \in A_{r+1}$, and define $S_0, S_1, S_2, \dots$ as the endpoints of all alternating paths starting from a and of lengths $0, 2, 4, \dots$. That is,

$$S_0 = \{a\} \text{ and } S_i = \phi_r^{-1}(\Lambda(S_{i-1})).$$

Since we are in the good case, $|S_i| \geq 4|S_{i-1}|$ provided $|S_{i-1}| \leq n/5$, and so there exists a smallest index i_S such that $|S_{i_S-1}| > n/5$, and $i_S - 1 \leq \log_4(n/5) \leq \log_4 n - 1$. Arbitrarily discard vertices from S_{i_S-1} to create a smaller set S'_{i_S-1} with $|S'_{i_S-1}| = \lceil n/5 \rceil$, so that $S'_{i_S} = \Lambda(S'_{i_S-1})$ has cardinality $|S'_{i_S}| \geq 4|S'_{i_S-1}| \geq 4n/5$.

Similarly, for an arbitrary $b \in B$, define $T_0, T_1, \dots,$ by

$$T_0 = \{b\} \text{ and } T_i = \phi_r(\Lambda(T_{i-1})).$$

Again, we will find an index $i_T \leq \log_4 n$ whose modified set has cardinality $|T'_{i_T}| \geq 4n/5$.

With both $|S'_{i_S}|$ and $|T'_{i_T}|$ larger than $n/2$, there must be some $a' \in S'_{i_S}$ for which $b' = \phi_r(a') \in T'_{i_T}$. This establishes the existence of an alternating walk and hence (removing any cycles) an alternating path of length at most $2(i_S + i_T) + 1 \leq 2\log_4 n$ from a to b in $\vec{\Gamma}_r$. □

We will need the following lemma,

Lemma 5.7. *Suppose that $k_1 + k_2 + \cdots + k_M \leq a \log N$, and $X_1, X_2, \dots, X_M$ are independent random variables with X_i distributed as the k_ith minimum of N independent exponential rate one random variables. If $\mu > 1$ then*

$$\mathbf{Pr}\left(X_1 + \cdots + X_M \geq \frac{\mu a \log N}{N - a \log N} \right) \leq N^{a(1 + \log \mu - \mu)}.$$

Proof. Let $Y_1, Y_2, \dots, Y_N$ be independent exponentials with mean one and let $Y_{(k)}$ denote the kth smallest of these variables, where we assume that $k = O(\log N)$. We therefore have $X_i = Y_{(k_i)}$. The density function $f_k(x)$ of $Y_{(k)}$ is

$$f_k(x) = \binom{N}{k} k(1 - e^{-x})^{k-1} e^{-x(N-k+1)}$$

and hence the ith moment of $Y_{(k)}$ is given by

$$\begin{aligned}
\mathbf{E}\left[Y_{(k)}^i\right] &= \int_0^\infty \binom{N}{k} k x^i (1 - e^{-x})^{k-1} e^{-x(N-k+1)} dx \\
&\leq \int_0^\infty \binom{N}{k} k x^{i+k-1} e^{-x(N-k+1)} dx \\
&= \binom{N}{k} k \frac{(i+k-1)!}{(N-k+1)^{i+k}} \\
&\leq \left(1 + O\left(\frac{k^2}{N}\right)\right) \frac{k(k+1)\cdots(i+k-1)}{(N-k+1)^i}.
\end{aligned}$$

Thus, if $0 \le t < N-k+1$,

$$\mathbf{E}\left[e^{tY_{(k)}}\right] \le \left(1+O\left(\frac{k^2}{N}\right)\right)\sum_{i=0}^{\infty}\left(\frac{t}{N-k+1}\right)^i\binom{k+i-1}{i}$$
$$= \left(1+O\left(\frac{k^2}{N}\right)\right)\left(1-\frac{t}{N-k+1}\right)^{-k}.$$

If $Z = X_1+X_2+\cdots+X_M$ then if $0 \le t < N - a\log N$,

$$\mathbf{E}\left[e^{tZ}\right] = \prod_{i=1}^{M}\mathbf{E}\left[e^{tX_i}\right] \le \left(1-\frac{t}{N-a\log N}\right)^{-a\log N}.$$

It follows by the Markov inequality that

$$\mathbf{Pr}\left(Z \ge \frac{\mu a\log N}{N-a\log N}\right) \le \left(1-\frac{t}{N-a\log N}\right)^{-a\log N}\exp\left\{-\frac{t\mu a\log N}{N-a\log N}\right\}.$$

We put $t = (N-a\log N)(1-1/\mu)$ to minimise the above expression, giving

$$\mathbf{Pr}\left(Z \ge \frac{\mu a\log N}{N-a\log N}\right) \le (\mu e^{1-\mu})^{a\log N}.$$

□

Lemma 5.8. *W.h.p., for all ϕ_r, the weighted ab-diameter of $\vec{\Gamma}_r$ is at most $c_1\frac{\log n}{np}$ for some absolute contant $c_1 > 0$.*

Proof. Let

$$Z_1 = \max\left\{\sum_{i=0}^{k}w(x_i,y_i) - \sum_{i=0}^{k-1}w(y_i,x_{i+1})\right\},$$

where the maximum is over sequences $x_0,y_0,x_1,\ldots,x_k,y_k$ where (x_i,y_i) is one of the 40 shortest arcs leaving x_i for $i=0,1,\ldots,k \le k_0 = \lceil 3\log_4 n\rceil$, and (y_i,x_{i+1}) is a backwards matching edge.

We compute an upper bound on the probability that Z_1 is large. For any $\eta > 0$ we have

$$\mathbf{Pr}\left(Z_1 \ge \eta\frac{\log n}{np}\right) \le o(n^{-4}) + \sum_{k=0}^{k_0}((r+1)n)^{k+1}\left(\frac{1+o(1)}{n}\right)^{k+1}\times p^{k-1}\times$$
$$\int_{y=0}^{\infty}\left[\frac{1}{(k-1)!}\left(\frac{y\log n}{np}\right)^{k-1}\sum_{\rho_0+\rho_1+\cdots+\rho_k\le 40(k+1)}q(\rho_0,\rho_1,\ldots,\rho_k;\eta+y)\right]dy$$

where

$$q(\rho_0,\rho_1,\ldots,\rho_k;\eta) = \mathbf{Pr}\left(X_0+X_1+\cdots+X_k \ge \eta\frac{\log n}{np}\right),$$

$X_0,X_1,\ldots,X_k$ are independent and X_j is distributed as the ρ_jth minimum of r independent exponential random variables. (When $k=0$ there is no term $\frac{1}{(k-1)!}\left(\frac{y\log n}{n}\right)^{k-1}$).

Explanation: *The $o(n^{-4})$ term is for the probability that there is a vertex in V_r that has fewer than $(1-o(1))np$ neighbors in V_r. We have at most $((r+1)n)^{k+1}$ choices for the sequence*

$x_0, y_0, x_1, \ldots, x_k, y_k$. *The term* $\frac{1}{(k-1)!}\left(\frac{y\log n}{np}\right)^{k-1} dy$ *bounds the probability that the sum of* k *independent exponentials,* $w(y_0,x_1)+\cdots+w(y_{k-1},x_k)$, *is in* $\frac{\log n}{np}[y, y+dy]$. *(The density function for the sum of* k *independent exponentials is* $\frac{x^{k-1}e^{-x}}{(k-1)!}$.*) We integrate over* y.
$\frac{(1+o(1))p}{np}$ *is the probability that* (x_i, y_i) *is and edge of* G *and is the* ρ_i*th shortest edge leaving* x_i*, and these events are independent for* $0 \le i \le k$. *The factor* p^{k-1} *is the probability that the* B *to* A *edges of the path exist. The final summation bounds the probability that the associated edge lengths sum to at least* $\frac{(\eta+y)\log n}{np}$.

It follows from Lemma 5.7 with $a \le 3, N = (1+o(1))np, \mu = (\eta+y)/a$ that if η is sufficiently large then, for all $y \ge 0$,

$$q(\rho_1, \ldots, \rho_k; \eta+y) \le (np)^{-(\eta+y)\log n/(2\log np)} = n^{-(\eta+y)/2}.$$

Since the number of choices for $\rho_0, \rho_1, \ldots, \rho_k$ is at most $\binom{41k+40}{k+1}$ (the number of positive integral solutions to $a_0 + a_1 + \ldots + a_{k+1} \le 40(k+1)$) we have

$$\begin{aligned}
\mathbf{Pr}\left(Z_1 \ge \eta\frac{\log n}{np}\right) &\le o(n^{-4}) + 2n^{2-\eta/2}\sum_{k=0}^{k_0}\frac{(\log n)^{k-1}}{(k-1)!}\binom{41k+40}{k+1}\int_{y=0}^{\infty} y^{k-1}n^{-y/2}dy \\
&\le o(n^{-4}) + 2n^{2-\eta/2}\sum_{k=0}^{k_0}\frac{(\log n)^{k-1}}{(k-1)!}2^{41k+40}\left(\frac{2}{\log n}\right)^{k-2}\int_{z=0}^{\infty} z^{k-1}e^{-z}dz \\
&= o(n^{-4}) + 2^{39}n^{2-\eta/2}\log n\sum_{k=0}^{k_0}2^{42k} \\
&= o(n^{-4}),
\end{aligned}$$

for η sufficiently large. □

Lemma 5.8 shows that with probability $1 - o(n^{-4})$ in going from M_r to M_{r+1} we can find an augmenting path of weight at most $\frac{c_1\log n}{np}$. This completes the proof of Lemma 5.5 and Theorem 5.1. (Note that to go from w.h.p. to expectation we use the fact that w.h.p. $w(e) = O(\log n)$ for all $e \in A \times B$,) □

Notice also that in the proof of Lemmas 5.6 and 5.8 we can certainly make the failure probability less than n^{-K} for any constant $K > 0$.

5.3 Proof of Theorem 5.2

Just as the proof method for $K_{n,n}$ in [90], [92] can be modified to apply to $G_{n,n,p}$, the proof for K_n in [91] can be modified to apply to $G_{n,p}$.

5.3.1 Outline of the proof

This has many similarities with the proof of Theorem 5.1. The differences are subtle. The first is to let M_r^* be the minimum cost matching of size r after adding a special vertex v_{n+1}. We do this, because there is no natural way to choose a set of vertices of size $n/2 - o(n)$ that we can be sure contains no edge of the minimum cost matching. It is again important (Lemma 5.9) to estimate the probability that $v_{n+1} \in M_r^*$. (This will be our short-hand for v_{n+1} lies in an edge of M_r^*.) The approach is similar to that for Theorem 5.1, except that we now need to prove separate lower and upper bounds for this probability $P(n, r)$.

5.3.2 Proof details

Consider $G = G_{n,p}$, and denote the vertex set by $V = \{v_1, v_2, \ldots, v_n\}$. We will now use the notation $\{a, b\}$ for the edges of G. Add a special vertex v_{n+1} with $E(\lambda)$-cost edges to all vertices of V, and let G^* be the extended graph on $V^* = V \cup \{v_{n+1}\}$. Say that $v_1, \ldots, v_n$ are *ordinary*. Let M_r^* be the minimum cost r-matching (one of size r) in G^*, unique with probability one. (Note the change in definition.) Define $P(n, r)$ as the normalized probability that $v_{n+1} \in M_r^*$, i.e.

$$P(n, r) = \lim_{\lambda \to 0} \frac{1}{\lambda} \mathbf{Pr}(v_{n+1} \in M_r^*)$$

Let $C(n, r)$ denote the cost of the cheapest r-assignment of G. To estimate $C(n, r)$, we will again need to estimate $P(n, r)$, by the following lemma.

Lemma 5.9.

$$\boldsymbol{E}[C(n, r) - C(n-1, r-1)] = \frac{1}{n} P(n, r)$$

Proof. Let $X = C(n, r)$ and $Y = C(n-1, r-1)$. Fix $i \in [n]$ and let w be the cost of the edge (v_i, v_{n+1}), and let I denote the indicator variable for the event that the cost of the cheapest r-assignment that contains this edge is smaller than the cost of the cheapest r-assignment that does not use v_{n+1}. The rest of the proof is identical to the proof of Lemma 5.2, except that there are now n choices for i as opposed to r in the previous lemma. □

In this case, unlike the bipartite case, we are unable to directly find an asymptotic expression for $P(n, r)$, as we did in Lemma 5.3 and Corollary 5.1. Here we will have to turn to bounding $P(n, r)$ from below and above.

5.3.3 A lower bound for $P(n, r)$

We will consider an algorithm that finds a set $A_s \subseteq V^*$ which contains the set B_s of vertices participating in M_r^*, $s = |A_s| \geq |B_s| = 2r$. Call A_s the set of *exposed* vertices.

Initially let $A_s = B_s = \emptyset$ and $r = s = 0$. At stage s of the algorithm we condition on

> A_s, B_s and the existence and cost of all edges within A_s.
> In particular, we condition on r and the minimum r-matching M_r^*.

Given a minimum matching M_r^*, we decide how to build a proposed $(r+1)$-matching by comparing the following numbers and picking the smallest.

(a) z_a equals the cost of the cheapest edge between a pair of unexposed vertices.

(b) $z_b = \min\{c_1(v) : v \in A_s \setminus B_s\}$, where $c_1(v)$ is the cost of the cheapest edge between v and a vertex $\tau_1(v) \notin A_s$.

(c) $z_c = \min\{c_1(v) + c_1(u) + \delta(u,v) : u, v \in B_s\}$ where $\delta(u,v)$ denotes the cost of the cheapest alternating path from u to v with internal vertices in B_s, with the cost of edges in M_r^* taken as the negative of the actual value.

Let

$$z_{\min} = \min\{z_a, z_b, z_c\}.$$

If $z_{\min} = z_a$ then we reveal the edge $\{v, w\}$ and add it to M_r^* to form M_{r+1}^*. Once v, w have been determined, they are added to A_s and B_s, and we move to the next stage of the algorithm, updating $s \leftarrow s+2, r \leftarrow r+1$.

If $z_{\min} = z_b$ then let $v \in A_s \setminus B_s$ be the vertex with the cheapest $c_1(v)$. We reveal $w = \tau_1(v)$ and add w to A_s and to B_s while adding v to B_s. Now $M_{r+1}^* = M_r^* \cup \{v, w\}$. We move to the next stage of the algorithm, updating $s \leftarrow s+1, r \leftarrow r+1$.

If $z_{\min} = z_c$ then reveal $w_1 = \tau_1(u), w_2 = \tau_1(v)$. If $w_1 = w_2$, we say that we have a *collision*. In this case, the vertex w_1 is added to A_s (but not B_s), and we move to the next stage with $s \leftarrow s+1$. If there is no collision, we update M_r^* by the augmenting path $w_1, u, \ldots, v, w_2$ to form M_{r+1}^*. We add w_1, w_2 to A_s and B_s, and move on to the next stage with $s \leftarrow s+2$ and $r \leftarrow r+1$.

It follows that $A_s \setminus B_s$ consists of unmatched vertices that have been the subject of a collision.

It will be helpful to define A_s for all s, so in the cases where two vertices are added to A_s, we add them sequentially with a coin toss to decide the order.

The possibility of a collision is the reason that not all vertices of A_s participate in M_r^*. However, the probability of a collision at v_{n+1} is $O(\lambda^2)$, and as $\lambda \to 0$ this is negligible. In other words, as $\lambda \to 0$,

$$\mathbf{Pr}(v_{n+1} \in M_r^*) = \mathbf{Pr}(v_{n+1} \in B_{2r}) = \mathbf{Pr}(v_{n+1} \in A_{2r}) - O(\lambda^2)$$

and we will bound $\mathbf{Pr}(v_{n+1} \in A_{2r})$ from below.

Lemma 5.10. *Conditioning on $v_{n+1} \notin A_s$, A_s is a random s-subset of V.*

Proof. Trivial for $s = 0$. Suppose A_{s-1} is a random $(s-1)$-subset of V. Define $N_s(v) = \{w \notin A_s : (v, w) \in E\}$. In stage s, if we condition on $d_s(v) = |N_s(v)|$, then under this conditioning $N_s(v)$ is a random $d_s(v)$-subset of $V \setminus A_s$. This is because the constructon of A_s does not require the edges from A_s to $V \setminus A_s$ to be exposed. So, if $A_s \setminus A_{s-1} = \{w\}$ where w is added due to being the cheapest unexposed neighbor of an exposed v, then w is a random element of $N_s(v)$ and hence a random element of $V \setminus A_s$.

If we are in case (a), i.e. M_{r+1}^* is formed by adding an edge between two ordinary unexposed vertices v, w, then since we only condition on the size of the set $\{(v, w) : v, w \notin A_s\}$, all pairs $v, w \in V \setminus A_s$ are equally likely, and after a coin toss this can be seen as adding two random elements sequentially. We conclude that A_s is a random s-subset of V. □

Recall that $m = n/(\omega^{1/2} \log n)$.

Corollary 5.2. *W.h.p., for all* $0 \le s \le n-m$ *and all* $v \in V$,

$$|d_s(v) - (n-s)p| \le \omega^{-1/5}(n-s)p.$$

Proof. This follows from the Chernoff bounds as in Lemma 5.1. □

We now bound the probability that $A_s \setminus A_{s-1} = \{v_{n+1}\}$ from below. There are a few different ways this may happen.

We now have to address some cost conditioning issues. Suppose that we have just completed an iteration. First consider the edges between vertices not in A_s. For such an edge e, all we know is $w(e) \ge \eta$ where $\eta = z_{\min}$ of the just completed iteration. So the conditional cost of such an edge can be expressed as $\eta + E(1)$ or $\eta + E(\lambda)$ in the case where e is incident with v_{n+1}. The exponentials are independent. We only need to compare the exponential parts of each edge cost here to decide the probability that an edge incident with v_{n+1} is chosen.

We can now consider case (a). Suppose that an edge $\{u, v\}$ between unexposed vertices is added to A_{s-1}. By Corollary 5.2, there are at most $p\binom{n-s+1}{2}(1+\omega^{-1/5})$ ordinary such edges. There are $n-s$ edges between v_{n+1} and $V \setminus A_s$, each at rate λ. As $\lambda \to 0$, the probability that one of the endpoints of the edge chosen in case (a) is v_{n+1} is therefore at least

$$\frac{\lambda(n-s)}{\lambda(n-s) + p\binom{n-s+1}{2}(1+\omega^{-1/5})} \ge \frac{1}{p}\frac{2\lambda}{n-s}(1-\omega^{-1/5}) + O(\lambda^2)$$

We toss a fair coin to decide which vertex in the edge $\{u, v\}$ goes in A_s. Hence the probability that $A_s \setminus A_{s-1} = \{v_{n+1}\}$ in case (a) is at least

$$\frac{1}{p}\frac{\lambda}{n-s}(1-\omega^{-1/5}) + O(\lambda^2).$$

We may also have $A_s \setminus A_{s-1} = \{v_{n+1}\}$ if case (a) occurs at stage $s-2$ and v_{n+1} loses the coin toss, in which case the probability is at least

$$\frac{1}{p}\frac{\lambda}{n-s+1}(1-\omega^{-1/5}) + O(\lambda^2).$$

Now consider case (b). Here only one vertex is added to A_{s-1}, the cheapest unexposed neighbor w of some $v \in A_{s-1} \setminus B_{s-1}$. The cost conditioning here is the same as for case (a), i.e. that the cost of an edge is $\eta + E(1)$ or $\eta + E(\lambda)$. By Corollary 5.2, this v has at most $p(n-s+1)(1+\omega^{-1/5})$ ordinary unexposed neighbors, so the probability that $w = v_{n+1}$ is at least

$$\frac{\lambda}{p(n-s+1)(1+\omega^{-1/5}) + \lambda} = \frac{1}{p}\frac{\lambda}{n-s+1}(1-\omega^{-1/5}) + O(\lambda^2).$$

Finally, consider case (c). To handle the cost conditioning, we condition on the values $c_1(v)$ for $v \in B_s$. By well-known properties of independent exponential variables, the minimum is located with probability proportional to the rates of the corresponding exponential variables. A collision at v_{n+1} has probability $O(\lambda^2)$, so assume we are in the case of two distinct unexposed vertices w_1, w_2. Suppose that w_1 is revealed first. Exactly as in (b), the probability that $w_1 = v_{n+1}$ is at least

$$\frac{1}{p}\frac{\lambda}{n-s+1}(1-\omega^{-1/5}) + O(\lambda^2).$$

If $w_1 \neq v_{n+1}$, the probability that $w_2 = v_{n+1}$ (i.e. $A_{s+1} \setminus A_s = \{v_{n+1}\}$) is at least

$$\frac{1}{p}\frac{\lambda}{n-s}(1-\omega^{-1/5}) + O(\lambda^2),$$

so by considering the possibility that v_{n+1} is the second vertex added from A_{s-2}, we again have probability at least

$$\frac{1}{p}\frac{\lambda}{n-s+1}(1-\omega^{-1/5}) + O(\lambda^2).$$

We conclude that no matter which case occurs, the probability is at least

$$\frac{1}{p}\frac{\lambda}{n-s+1}(1-\omega^{-1/5}) + O(\lambda^2).$$

So

$$P(n,r) \geq \lim_{\lambda\to 0}\frac{1}{\lambda}\sum_{s=1}^{2r}\mathbf{Pr}(A_s \setminus A_{s-1} = \{v_{n+1}\}) \geq \frac{1-\omega^{-1/5}}{p}\sum_{s=1}^{2r}\frac{1}{n-s+1}. \tag{5.16}$$

Write

$$L(n,r) = \sum_{s=1}^{2r}\frac{1}{n-s+1}.$$

5.3.4 An upper bound for $P(n,r)$

We now alter the algorithm above in such a way that $A_{2r} = B_{2r}$. We do not consider A_s for odd s here. At a stage with $s = 2r$, we condition on

A_s, and the appearance and cost of all edges within A_s.
In particular, we condition on r and the minimum r-matching M_r^*.
A set $C_s \subseteq A_s$, where each $v \in C_s$ has been involved in a collision.
The cost $c_1(v), v \in C_s$ of the cheapest edge from v to a vertex not in A_s.

This changes how we calculate a candidate for M_{r+1}^*. We now take the minimum of

(a) z_a equals the cost of the cheapest edge between unexposed vertices.

(b) $z_b = \min\{c_1(u) + c_1(v) + \delta(u,v) : u, v \in A_s, |\{u,v\} \cap C_s| \leq 1\}$, where c_1 and δ are as defined in Section 5.3.3.

(c) $z_c = \min\{c_1(u) + c_2(v) + \delta(u,v) : u, v \in C_s, \tau_1(u) = \tau_1(v)\}$, where τ_1 is defined in Section 5.3.3 and $c_2(v)$ is the cost of the second cheapest edge between v and a vertex $\tau_2(v) \notin A_s$.

Let

$$z_{\min} = \min\{z_a, z_b, z_c\}.$$

If $z_{\min} = z_a$ then we reveal the edge $\{v, w\}$ and add it to M_r^* to form M_{r+1}^*. Once v, w have been determined, they are added to A_s and we move to the next stage of the algorithm, updating $s \leftarrow s + 2$.

If $z_{\min} = z_b$ then reveal $w_1 = \tau_1(u), w_2 = \tau_1(v)$. If $w_1 = w_2$ then we add u, v to C_s and go to the next stage of the algorithm without changing s. (The probability that $\tau_1(u) = \tau_1(v) = v_{n+1}$ is $O(\lambda^2)$, and we can safely ignore this as $\lambda \to 0$). If at some later stage w_1 is added to A_s and u say is still in C_s then we remove u from C_s. If $w_1 \neq w_2$ then we update M_r^* by the augmenting path $w_1, u, \ldots, v, w_2$ to form M_{r+1}^*. We add w_1, w_2 to A_s, and move on to the next stage with $s \leftarrow s+2$.

If $z_{\min} = z_c$ then we update M_r^* by the augmenting path $w_1 = \tau_1(u), u, \ldots, v, w_2 = \tau_2(v)$ to form M_{r+1}^*. We add w_1, w_2 to A_s, and move on to the next stage with $s \leftarrow s+2$.

Eventually we will construct M_{r+1}^* since case (b) with $\tau_1(u) = \tau_1(v)$ can happen at most s times before $C_s = A_s$.

The cost conditioning is the same as we had for computing the lower bound in Section 5.3.3, except for the need to deal with $c_2(v), v \in C_s$. For this we condition on $c_2(v)$ and argue that the probability $\tau_2(v) = x$ is proportional to the exponential rate for the edge (v, x). At this point we know that $\tau_1(v) \neq v_{n+1}$, since we are assuming λ is so small that this possibility can be ignored. So, in this case, we can only add v_{n+1} as $\tau_2(v)$ for some $v \in C_s$.

To analyze this algorithm we again need to show that A_{2r} is a uniformly random subset of V.

Lemma 5.11. *Conditioning on $v_{n+1} \notin A_{2r}$, A_{2r} is a random $2r$-subset of V.*

Proof. Let D denote the $n \times n$ matrix of edge costs, where $D(i,j) = w(v_i, v_j)$ and $D(i,j) = \infty$ if edge (v_i, v_j) does not exist in G. For a permutation π of V let D_π be defined by $D_\pi(i,j) = D(\pi(i), \pi(j))$. Let X, Y be two distinct $2r$-subsets of V and let π be any permutation of V that takes X into Y. Then we have

$$\mathbf{Pr}(A_{2r}(D) = X) = \mathbf{Pr}(A_{2r}(D_\pi) = \pi(X)) = \mathbf{Pr}(A_{2r}(D_\pi) = Y) = \mathbf{Pr}(A_{2r}(D) = Y),$$

where the last equality follows from the fact that D and D_π have the same distribution. This shows that A_{2r} is a random $2r$-subset of V. □

Let $d_{2r}(v) = |\{w \notin A_{2r} : (v, w) \in E\}|$.

Corollary 5.3. *W.h.p., for all $0 \leq r \leq (n-m)/2$ and all $v \in V$,*

$$|d_{2r}(v) - p(n-2r)| \leq \omega^{-1/5} p(n-2r).$$

Proof. The proof is again via Chernoff bounds, see Lemma 5.1. □

We bound the probability that $v_{n+1} \in A_{2r} \setminus A_{2r-2}$ from above. Suppose we are at a stage where a collisionless candidate for M_r^* has been found.

In case (a), as in the previous section the probability that v_{n+1} is one of the two unexposed vertices is at most

$$\frac{\lambda(n-2r+2)}{\lambda(n-2r+2) + p\binom{n-2r+2}{2}(1-\omega^{-1/5})} = \frac{1}{p}\frac{2\lambda}{n-2r+1}(1+\omega^{-1/5}) + O(\lambda^2)$$

Now suppose we are in case (b) with $u, v \notin C_{2r-2}$. If no collision occurs, the probability that one of $\tau_1(u), \tau_1(v)$ is v_{n+1} is at most

$$\frac{\lambda}{\lambda + d_s(u)} + \frac{\lambda}{\lambda + d_s(v) - 1} \leq \frac{1}{p}\frac{2\lambda}{n-2r+1}(1+\omega^{-1/5}) + O(\lambda^2)$$

Finally, if we find M^*_{r+1} by alternating paths where one exposed vertex uses its second-cheapest edge to an unexposed vertex, the probability of that vertex being v_{n+1} is even smaller at $\lambda/(n-2r+1)$. So,

$$P(n,r) = \lim_{\lambda\to 0}\sum_{s=1}^{r}\mathbf{Pr}(v_{n+1}\in A_{2s}\setminus A_{2s-2}) \le \frac{2(1+\omega^{-1/5})}{p}\sum_{s=1}^{r}\frac{1}{n-2s+1}$$

Write

$$U(n,r) = \sum_{s=1}^{r}\frac{2}{n-2s+1}.$$

5.3.5 Calculating $\mathbf{E}\left[C(G_{n,p})\right]$

From Lemma 5.9 and (5.16) we have

$$\begin{aligned}
&\mathbf{E}\left[C(n,(n-m)/2)\right]\\
&= \sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}P(n-r+1,(n-m)/2-r+1)\\
&\ge \frac{1+o(1)}{p}\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}L(n-r+1,(n-m)/2-r+1)\\
&= \frac{1+o(1)}{p}\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}\sum_{s=1}^{n-m-2r+2}\frac{1}{(n-r+1)-s+1}\\
&= \frac{1+o(1)}{p}\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}\left(\log\left(\frac{n-r+1}{m+r}\right)+\frac{1}{2(n-r)}-\frac{1}{2(m+r)}+O(m^{-2})\right), \quad (5.17)
\end{aligned}$$

by (5.8).

The correction terms are easily taken care of. First we have

$$\begin{aligned}
&\left|\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}\left(\frac{1}{2(n-r)}-\frac{1}{2(m+r)}+O(m^{-2})\right)\right|\\
=&O\left(\frac{1}{m}\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}\right)\\
=&O\left(\frac{n-m}{mn}\right)\\
=&o(1).
\end{aligned}$$

Now we want to replace the $(m+r)$ term in the logarithm in the RHS of (5.17) by r. For this we let $m_1 = n/(\omega^{1/4}\log n) = m\omega^{1/4}$. Then

$$\begin{aligned}
&\left|\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}\log\left(\frac{r}{m+r}\right)\right| \\
=&\sum_{r=1}^{m_1-1}\frac{1}{n-r+1}\log\left(1+\frac{m}{r}\right)+\sum_{r=m_1}^{(n-m)/2}\frac{1}{n-r+1}\log\left(1+\frac{m}{r}\right) \\
\leq&\log m\sum_{r=1}^{m_1-1}\frac{1}{n-r+1}+\log\left(1+\frac{m}{m_1}\right)\sum_{r=m_1}^{(n-m)/2}\frac{1}{n-r+1} \\
\leq&\log n\frac{m_1}{n-m_1}+\frac{m}{m_1}\frac{(n-m)/2}{n/2} \\
=&o(1).
\end{aligned}$$

So, using (5.17) we have

$$\begin{aligned}
p\times\mathbf{E}\left[C(n,(n-m)/2)\right] =&o(1)+\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}\log\left(\frac{n-r}{r}\right) \\
=&o(1)+\int_0^{1/2}\frac{1}{1-\alpha}\log\left(\frac{1-\alpha}{\alpha}\right)d\alpha.
\end{aligned}$$

Substituting $y=\log(1/\alpha-1)$ we have

$$\begin{aligned}
\int_0^{1/2}\frac{1}{1-\alpha}\log\left(\frac{1-\alpha}{\alpha}\right)d\alpha &= \int_0^{\infty}\frac{y}{e^y+1}dy \\
&= \int_0^{\infty}\frac{ye^{-y}}{1+e^{-y}}dy \\
&= \sum_{j=0}^{\infty}\int_0^{\infty} ye^{-y}(-e^{-y})^j dy \\
&= \sum_{j=1}^{\infty}(-1)^{j+1}\frac{1}{j^2} \\
&= \frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k^2} \\
&= \frac{\pi^2}{12}.
\end{aligned}$$

This proves a lower bound for $\mathbf{E}\left[C(n,(n-m)/2)\right]$. It also shows that

$$\sum_{r=1}^{(n-m)/2}\frac{1}{n-r+1}L(n-r+1,(n-m)/2-r+1)=(1+o(1))\frac{\pi^2}{12}. \tag{5.18}$$

For an upper bound, note that for $r \le (n-m)/2$,

$$\begin{aligned} U(n,r) - L(n,r) &= \sum_{s=1}^{r} \left(\frac{1}{n-2s+1} - \frac{1}{n-2s+2} \right) \\ &= \sum_{s=1}^{r} \frac{1}{(n-2s+1)(n-2s+2)} \\ &= O\left(\frac{r}{(n-2r)^2} \right) \end{aligned}$$

So,

$$\begin{aligned} &\mathbf{E}\left[C(n,(n-m)/2)\right] \\ &\le \frac{1+o(1)}{p} \sum_{r=1}^{(n-m)/2} \frac{1}{n-r+1} U(n-r+1,(n-m)/2-r+1) \\ &= \frac{1+o(1)}{p} \sum_{r=1}^{(n-m)/2} \frac{1}{n-r+1} \left(L(n-r+1,(n-m)/2-r+1) + O\left(\frac{r}{(n-2r)^2} \right) \right) \\ &= \frac{1+o(1)}{p} \left(o(1) + \sum_{r=1}^{(n-m)/2} \frac{1}{n-r+1} \left(L(n-r+1,(n-m)/2-r+1) \right) \right) \qquad (5.19) \\ &= \frac{\pi^2}{12p}(1+o(1)). \qquad (5.20) \end{aligned}$$

To get from (5.19) to (5.20) we use (5.18).

We show that for n even, $\mathbf{E}\left[C(n,n/2) - C(n,(n-m)/2)\right] = o(1/p)$ to conclude that

$$\mathbf{E}\left[C(G_{n,p})\right] = \mathbf{E}\left[C(n,n/2) - C(n,(n-m)/2)\right] + \mathbf{E}\left[C(n,(n-m)/2)\right] = \frac{\pi^2}{12p}(1+o(1)).$$

As above, this will follow from the following lemma.

Lemma 5.12. *Suppose n is even. If $(n-m)/2 \le r \le n/2$ then $0 \le \mathbf{E}[C(n,r+1) - C(n,r)] = O\left(\frac{\log n}{np}\right)$.*

5.3.6 Proof of Lemma 5.12

This section will replace Section 5.2.3. Let $M = \{v, \phi(v), v \in [n]\}, \phi_r^2(v) = v$ for all $v \in [n]$ be an arbitrary perfect matching of $G_{n,p}$. We let $\vec{M} = \{(u,v) : v = \phi_r(u)\}$ consist of two oppositely oriented copies of each edge of M. We then randomly orient the edges of $G_{n,p}$ that are not in M and then add $\vec{M}$ to obtain the digraph $\vec{G} = \vec{G}_{n,p}$. In the case of $G_{n,n,p}$ we only needed to orient $e \in M$ from B to A. Here we need an oriented copy of $e \in M$ in both directions because if $e, f \in M$, we cannot guarantee that an alternating path will traverse e and f consistently with a single given orientation. Because $np = \omega \log^2 n$ we have that w.h.p. the minimum in- or out-degree in $\vec{G}_{n,p}$ is at least $\omega \log^2(n)/3$. Let $\mathcal{D}$ be the event that all in- and out-degrees are at least this large. Let the *alternating diameter* of $\vec{G}$ be the maximum over pairs of vertices $u \neq v$ of the minimum length of an odd length alternating path w.r.t. M between u and v where (i) the edges are oriented along the path in the direction u to v, (ii) the first and last edges are not in M. Given this orientation, we define $\vec{\Gamma}_r$ to be the subdigraph of $\vec{G}$ consisting of the r cheapest non-$\vec{M}$ out-edges from each vertex together with $\vec{M}$. Once we can show that the alternating diameter of $\vec{\Gamma}_{20}$ is at most $\lceil 3\log_3 n\rceil$, the proof follows the proof of Lemma 5.8 more or less exactly.

Lemma 5.13. *W.h.p., the alternating diameter of* $\vec{\Gamma}_{20}$ *is at most* $k_0 = \lceil 3\log_3 n\rceil$.

Proof. We first consider the relatively simple case where $np \geq n^{1/3}\log n$. Let $N^+(u)$ be the set of out-neighbors of u in $\vec{G}$ and let $N^-(v)$ be the set of in-neighbors of v in $\vec{G}$. If there is an edge of $\vec{M}$ from $N^+(u)$ to $N^-(v)$ then this creates an alternating path of length three. Otherwise, let $N^{++}(u)$ be the other endpoints of the matching edges incident with $N^+(u)$ and define $N^{--}(v)$ analogously. Note that now we have $N^{++}(u)\cap N^{--}(v)=\emptyset$ and given $\mathcal{D}$, the conditional probability that there is no edge from $N^{++}(u)$ to $N^{--}(v)$ in $\vec{G}$ is at most $(1+o(1))(1-p)^{(np/3)^2}\leq e^{-(\log n)^3/10}=o(n^{-2})$. Thus in this case there will w.h.p. be an alternating path of length five between any pair of vertices $u,v\in V$.

Now assume that $np < n^{1/3}\log n$. For $S\subseteq V$, $N_{20}^+(S)=\left\{w\notin S:\exists v\in S,(v,w)\in E(\vec{\Gamma}_{20}\right\})$ is the set of out-neighbors of S in $\vec{\Gamma}_{20}$ and $N_{20}^-(S)$ is similarly defined.

Imitating Lemma 5.6, we prove an expansion property for $\vec{\Gamma}_{20}$: The $o(1)$ term in the first inequality accounts for conditioning on the event $\mathcal{D}$.

$$
\begin{aligned}
\mathbf{Pr}(\exists S:\ |S|\leq n^{2/3},\ |N_{20}^+(S)|<10|S|) &\leq o(1)+\sum_{s=1}^{n^{2/3}}\binom{n}{s}\binom{n-s}{10s}\left(\frac{\binom{11s}{20}}{\binom{n}{20}}\right)^s\\
&\leq \sum_{s=1}^{n^{2/3}}\left(\frac{ne}{s}\right)^s\left(\frac{ne}{10s}\right)^{10s}\left(\frac{11s}{n}\right)^{20s}\\
&= \sum_{s=1}^{n^{2/3}}\left(\frac{e^{11}11^{20}s^9}{10^{10}n^9}\right)^s\\
&= o(1). \qquad (5.21)
\end{aligned}
$$

Fix an arbitrary pair of vertices a,b. Define $S_i, i=0,1,\ldots$ to be the set of vertices v such that there exists an alternating path of length $2i$ in $\vec{\Gamma}_{20}$ from a to v. We let $S_0=\{a\}$ and given S_i we let $S_i'=N^+(S_i)\setminus\{b\}$ and $S_i''=\{w\neq b:\exists\{v,w\}\in M: v\in S_i'\}$. We now argue that (5.21) implies that w.h.p. $|S_i''|\geq 3|S_i|$, so long as $|S_i|=o(n^{7/12})$. Indeed, It follows from (5.21) that S_i has at least $9|S_i|$ out-neighbors T via a non-M edge. Now consider the edges of M incident with T. At most $|S_i|$ of these have one endpoint in S_i and one endpoint in T. At most $|T|/2$ have both endpoints in T. Thus at least $|T|-|T|/2-|S_i|>3|S_i|$ of these edges have one endpoint in T and one endpoint not in $S_i\cup T$.

We can therefore take S_{i+1} to be a subset of S_i'' of size $3|S_i|$. So w.h.p. there exists an $i_a\leq\log_3 n$ such that $|S_{i_a}|\in[n^{13/24},3n^{13/24}]$.

Repeat the procedure with vertex b, letting $T_j, j=0,1,\ldots$ be the set of vertices v such that there exists an alternating path of length $2i$ in $\vec{\Gamma}_{20}$ from v to b. Then let $T_{j+1}'=N^-(T_j)$ etc. By the same argument, there exists an $j_b\leq\log_3 n$ such that T_{j_b} is of size in $[n^{13/24},3n^{13/24}]$. Finally, the probability that there is no $S_{i_a}\to T_{j_b}$ edge is at most $(1-p)^{n^{13/12}}=o(n^{-2})$. This completes the proof of Lemma 5.13. □

The remainder of the proof of Lemma 5.12 is now exactly as in Section 5.2.3. This concludes the proof of Theorem 5.2.

5.4 Proof of Theorem 5.3

The proof of Lemma 5.8 allows us to claim that for any constant $K > 0$, with probability $1-O(n^{-K})$ the maximum length of an edge in the minimum cost perfect matching of $G = G_{n,n,p}$ or $G = G_{n,p}$ is at most $\xi = c_2 \frac{\log n}{np}$ for some constant $c_2 = c_2(K) > 0$. We now closely follow the ideas in Talagrand's proof [86] of concentration for the assignment problem. We let $\widehat{w}(e) = \min\{w(e), \xi\}$ and let $\widehat{C}(G)$ be the assignment cost using $\widehat{w}$ in place of w. We observe that

$$\mathbf{Pr}(\widehat{C}(G) \neq C(G)) = O(n^{-K}) \tag{5.22}$$

and so it is enough to prove concentration of $\widehat{C}(G)$.

For this we use the following result of Talagrand [86]: consider a family $\mathcal{F}$ of N-tuples $\boldsymbol{\alpha} = (\alpha_i)_{i \leq N}$ of non-negative real numbers. Let

$$Z = \min_{\boldsymbol{\alpha} \in \mathcal{F}} \sum_{i \leq N} \alpha_i X_i$$

where $X_1, X_2, \ldots, X_N$ are an independent sequence of random variables taking values in $[0,1]$.

Let $\sigma = \max_{\boldsymbol{\alpha} \in \mathcal{F}} ||\boldsymbol{\alpha}||_2$. Then if M is the median of Z and $u > 0$, we have

$$\mathbf{Pr}(|Z - M| \geq u) \leq 4 \exp\left\{-\frac{u^2}{4\sigma^2}\right\}. \tag{5.23}$$

We apply (5.23) with $N = n^2$ and $X_e = \widehat{w}(e)/\xi$. For $\mathcal{F}$ we take the $n!$ $\{0,1\}$ vectors corresponding to perfect matchings and scale them by ξ. In this way, $\sum_e \alpha_e X_e$ will be the weight of a perfect matching. In this case we have $\sigma^2 \leq n\xi^2$. Applying (5.23) we obtain

$$\mathbf{Pr}\left(|\widehat{C}(G) - \widehat{M}| \geq \frac{\varepsilon}{p}\right) \leq 4 \exp\left\{-\frac{\varepsilon^2}{4p^2} \cdot \frac{1}{n\xi^2}\right\} = \exp\left\{-\frac{\varepsilon^2 n}{(c_2 \log n)^2}\right\},$$

where $\widehat{M}$ is the median of $\widehat{C}(G)$.

It then follows from (5.22) that

$$\mathbf{Pr}\left(|C(G) - \widehat{M}| \geq \frac{\varepsilon}{p}\right) = O(n^{-K}). \tag{5.24}$$

Now

$$\mu(G) \leq \widehat{M} + \frac{\varepsilon}{p} + O(n^{-(K-2)}) \tag{5.25}$$

after using (5.24) and $\mathbf{E}\left[\sum_{v_i, v_j \in E(G)} w(v_i, v_j)\right] \leq n^2$.

In addition

$$\mu(G) \geq \left(\widehat{M} - \frac{\varepsilon}{p}\right)\left(1 - O(n^{-K})\right). \tag{5.26}$$

Theorem 5.3 follows easily from (5.22), (5.25) and (5.26).

5.5 Final remarks

We have generalised the result of [4] to the random bipartite graph $G_{n,n,p}$ and the result of [91] to the random graph $G_{n,p}$. It would be of some interest to extend the result in some way to random regular graphs. In the absence of proving Conjecture 5.1 we could maybe extend the results of [4], [91] to some special class of special graphs e.g. to the hypercube.

Chapter 6

Minimum cost of two spanning trees

This chapter corresponds to [39].

Abstract

Assume that the edges of the complete graph K_n are given independent uniform $[0,1]$ edges weights. We consider the expected minimum total weight μ_k of $k \geq 2$ edge disjoint spanning trees. When k is large we show that $\mu_k \approx k^2$. Most of the paper is concerned with the case $k = 2$. We show that μ_2 tends to an explicitly defined constant and that $\mu_2 \approx 4.1704288\ldots$.

6.1 Introduction

This paper can be considered to be a contribution to the following general problem. We are given a combinatorial optimization problem where the weights of variables are random. What can be said about the random variable equal to the minimum objective value in this model. The most studied examples of this problem are those of (i) Minimum Spanning Trees e.g. Frieze [42], (ii) Shortest Paths e.g. Janson [59], (iii) Minimum Cost Assignment e.g. Aldous [3], [4], Linusson and Wästlund [68] and Nair, Prabhakar and Sharma [74], Wästlund [92] and (iv) the Travelling Salesperson Problem e.g. Karp [61], Frieze [44] and Wästlund [93].

The minimum spanning tree problem is a special case of the problem of finding a minimum weight basis in an element weighted matroid. Extending the result of [42] has proved to be difficult for other matroids. We are aware of a general result due to Kordecki and Lyczkowska-Hanćkowiak [64] that expresses the expected minimum value of an integral using the Tutte Polynomial. The formulae obtained, although exact, are somewhat difficult to penetrate. In this paper we consider the union of k cycle matroids. We have a fairly simple analysis for $k \to \infty$ and a rather difficult analysis for $k = 2$.

Given a connected simple graph $G = (V, E)$ with edge lengths $\mathbf{x} = (x_e : e \in E)$ and a positive integer k, let $\mathrm{mst}_k(G, \mathbf{x})$ denote the minimum length of k edge disjoint spanning trees of G. ($\mathrm{mst}_k(G) = \infty$ if such trees do not exist.) When $\mathbf{X} = (X_e : e \in E)$ is a family of independent random variables, each uniformly distributed on the interval $[0,1]$, denote the expected value $\mathbf{E}[\mathrm{mst}_k(G, \mathbf{X})]$ by $\mathrm{mst}_k(G)$.

As previously mentioned, the case $k = 1$ has been the subject of some attention. When G is the complete graph K_n, Frieze [42] proved that

$$\lim_{n\to\infty} \mathrm{mst}_1(K_n) = \zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3}.$$

Generalisations and refinements of this result were subsequently given in Steele [83], Frieze and McDiarmid [48], Janson [58], Penrose [79], Beveridge, Frieze and McDiarmid [9], Frieze, Ruszinko and Thoma [51] and most recently in Cooper, Frieze, Ince, Janson and Spencer [19].

In this paper we discuss the case $k \geq 2$ when $G = K_n$ and define

$$\mu_k^* = \liminf_{n\to\infty} \mathrm{mst}_k(K_n) \text{ and } \mu_k^{**} = \limsup_{n\to\infty} \mathrm{mst}_k(K_n).$$

Conjecture: $\mu_k^* = \mu_k^{**}$ i.e. $\lim_{n\to\infty} \mathrm{mst}_k(K_n)$ exists.

Theorem 6.1.

(a)

$$\lim_{k\to\infty} \frac{\mu_k^*}{k^2} = \lim_{k\to\infty} \frac{\mu_k^{**}}{k^2} = 1.$$

(b) *With f_k and $c_2' \approx 3.59$ and $\lambda_2' \approx 2.688$ as defined in* (6.1), (6.5), (6.9),

$$\begin{aligned}\mu_2 &= 2c_2' - \frac{(c_2')^2}{4} + \int_{\lambda=\lambda_2'}^{\infty} \left(2 - \frac{\lambda e^\lambda}{2f_2(\lambda)} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right)\left(\frac{e^\lambda}{f_2(\lambda)} + \frac{\lambda e^\lambda}{f_2(\lambda)} - \frac{\lambda e^\lambda f_1(\lambda)}{f_2(\lambda)^2}\right) d\lambda \\ &= 4.17042881\dots\end{aligned}$$

There appears to be no clear connetion between μ_2 and the ζ function.

Note also, in connection with Theorem 6.1(a), that if n is even and $k = (n-1)/2$ and we take a partition of the edge set of K_n into spanning trees then w.h.p. $\mu_k \approx \frac{n^2}{4} \approx k^2$.

Before proceeding to the proof of Theorems 6.1 we note some properties of the κ-core of a random graph.

6.2 The κ-core

The functions

$$f_i(\lambda) = \sum_{j=i}^{\infty} \frac{\lambda^j}{j!}, \quad i = 0, 1, 2, \dots, \tag{6.1}$$

figure prominently in our calculations. We let

$$g_i(\lambda) = \frac{\lambda f_{2-i}(\lambda)}{f_{3-i}(\lambda)}, \quad i = 0, 1, 2.$$

Properties of these functions are derived in Section 6.8.

The κ-*core* $C_\kappa(G)$ of a graph G is the largest set of vertices that induces a graph H_κ such that the minimum degree $\delta(H_\kappa) \geq \kappa$. Pittel, Spencer and Wormald [80] proved that there exist constants, $c_\kappa, \kappa \geq 3$ such that if $p = c/n$ and $c < c_\kappa$ then w.h.p. $G_{n,p}$ has no κ-core and that if $c > c_\kappa$ then w.h.p. $G_{n,p}$ has a κ-core of linear size. We list some facts about these cores that we will need in what follows.

Given λ let $\mathrm{Po}(\lambda)$ be the Poisson random variable with mean λ and let

$$\pi_r(\lambda) = \Pr\{\mathrm{Po}(\lambda) \geq r\} = e^{-\lambda} f_r(\lambda).$$

Then

$$c_\kappa = \inf\left(\frac{\lambda}{\pi_{\kappa-1}(\lambda)} : \lambda > 0\right).$$

When $c > c_\kappa$ define $\lambda_\kappa(c)$ by

$$\lambda_\kappa(c) \text{ is the larger of the two roots to the equation } c = \frac{\lambda}{\pi_{\kappa-1}(\lambda)} = \frac{\lambda e^\lambda}{f_{\kappa-1}(\lambda)}. \tag{6.2}$$

Then **whp**[1] with $\lambda = \lambda_\kappa(c)$ we have that

$$C_\kappa(G_{n,p}) \text{ has } \approx \pi_\kappa(\lambda)n = \frac{f_\kappa(\lambda)}{e^\lambda} n \text{ vertices and } \approx \frac{\lambda^2}{2c} n = \frac{\lambda f_{\kappa-1}(\lambda)}{2e^\lambda} n \text{ edges.} \tag{6.3}$$

Furthermore, when κ is large,

$$c_\kappa = \kappa + (\kappa \log \kappa)^{1/2} + O(\log \kappa). \tag{6.4}$$

Łuczak [69] proved that C_κ is κ-connected **whp** when $\kappa \geq 3$.

Next let c'_κ be the threshold for the $(\kappa+1)$-core having average degree 2κ. Here, see (6.2) and (6.3),

$$c'_\kappa = \frac{\lambda e^\lambda}{f_\kappa(\lambda)} \text{ where } \frac{\lambda f_k(\lambda)}{f_{k+1}(\lambda)} = 2\kappa. \tag{6.5}$$

We have $c_2 \approx 3.35$ and $c'_2 \approx 3.59$.

6.3 Proof of Theorem 6.1(a): Large k.

We will prove Part (a) of Theorem 6.1 in this section. It is relatively straightforward. Part (b) is more involved and occupies Section 6.4.

In this section we assume that $k = O(1)$ and large. Let Z_k denote the sum of the $k(n-1)$ shortest edge lengths in K_n. We have that for $n \gg k$,

$$\mathrm{mst}_k(K_n) \geq \mathbf{E}[Z_k] = \sum_{\ell=1}^{k(n-1)} \frac{\ell}{\binom{n}{2}+1} = \frac{k(n-1)(k(n-1)+1)}{n(n-1)+2} \in [k^2(1-n^{-1}), k^2]. \tag{6.6}$$

This gives us the lower bound in Theorem 6.1(a).

[1] For the purposes of this paper, a sequence of events $\mathcal{E}_n$ will be said to occur *with high probability* **whp** if $\Pr\{\mathcal{E}_n\} = 1 - o(n^{-1})$

For the upper bound let $k_0 = k + k^{2/3}$ and consider the random graph H generated by the $k_0(n-1)$ cheapest edges of K_n. The expected total edge weight $\overline{E}_H$ of H is at most k_0^2, see (6.6).

H is distributed as G_{n,k_0n}. This is sufficiently close in distribution to $G_{n,p}, p = 2k_0/n$ that we can apply the results of Section 6.2 without further comment. It follows from (6.4) that $c_{2k} < 2k_0$. Putting $\lambda_0 = \lambda_{2k}(2k_0)$ we see from (6.3) that w.h.p. H has a $2k$-core C_{2k} with $\sim n\text{Pr}\left\{\text{Po}(\lambda_0) \geq 2k\right\}$ vertices. It follows from (6.2) that $\lambda_0 = 2k_0\pi_{2k-1}(2k_0) \leq 2k_0$ and since $\pi_{2k-1}(\lambda)$ increases with λ and $\pi_{2k-1}(2k + k^{2/3}) = \text{Pr}\left\{\text{Po}(2k + k^{2/3}) \geq 2k - 1\right\} \geq 1 - e^{-c_1k^{1/3}}$ for some constant $c_1 > 0$ we see that $\frac{2k+k^{2/3}}{\pi_{2k-1}(2k+k^{2/3})} \leq 2k_0$ and so $\lambda_0 \geq 2k + k^{2/3}$.

A theorem of Nash-Williams [75] states that a $2k$-edge connected graph contains k edge-disjoint spanning trees. Applying the result of Łuczak [69] we see that **whp** C_{2k} contains k edge disjoint spanning trees $T_1, T_2, \dots, T_k$. It remains to argue that we can cheaply augment these trees to spanning trees of K_n. Since $|C_{2k}| \sim n\text{Pr}\left\{\text{Po}(\lambda) \geq 2k\right\}$ **whp**, we see that **whp** $D_{2k} = [n] \setminus C_{2k}$ satisfies $|D_{2k}| \leq 2ne^{-c_1k^{1/3}}$.

For each $v \in D_{2k}$ we let S_v be the k shortest edges from v to C_{2k}. We can then add v as a leaf to each of the trees $T_1, T_2, \dots, T_k$ by using one of these edges. What is the total weight of the edges Y_v, $v \in D_{2k}$? We can bound this probabilistically by using the following lemma from Frieze and Grimmett [46]:

Lemma 6.1. *Suppose that $k_1 + k_2 + \cdots + k_M \leq a$, and $Y_1, Y_2, \dots, Y_M$ are independent random variables with Y_i distributed as the k_ith minimum of N independent uniform [0,1] random variables. If $\mu > 1$ then*

$$Pr\left\{Y_1 + \cdots + Y_M \geq \frac{\mu a}{N+1}\right\} \leq e^{a(1+\ln\mu-\mu)}.$$

Let $\varepsilon = 2e^{-c_1k^{1/3}}$ and $\mu = 10\ln 1/\varepsilon$ and let $M = k\varepsilon n$, $N = (1-\varepsilon)n$, $a = \frac{k(k+1)}{2}\varepsilon n$. Let $\mathcal{B}$ be the event that there exists a set S of size εn such that the sum of the k shortest edges from each $v \in S$ to $[n] \setminus S$ exceeds $\mu a/(N+1)$. Applying Lemma 6.1 we see that

$$\text{Pr}\left\{\mathcal{B}\right\} \leq \binom{n}{\varepsilon n} \exp\left\{k(k+1)\varepsilon n(1 + \ln\mu - \mu)/2\right\} \leq \left(\frac{e}{\varepsilon} \cdot e^{-\mu k^2/3}\right)^{\varepsilon n} = o(n^{-1}).$$

It follows that

$$\text{mst}_k(K_n) \leq o(1) + k_0^2 + \frac{\mu a}{N+1} \leq k^2 + 3k^{5/3}.$$

The $o(1)$ term is a bound $kn \times o(n^{-1})$, to account for the cases that occur with probability $o(n^{-1})$.

Combining this with (6.6) we see that

$$k^2 \leq \mu_k \leq k^2 + 3k^{5/3}$$

which proves Theorem 6.1(a).

6.4 Proof of Theorem 6.1(b): $k = 2$.

For this case we use the fact that for any graph $G = (V, E)$, the collection of subsets $I \subseteq E$ that can be partitioned into two edge disjoint forests form the independent sets in a matroid. This being the matroid which is the union of two copies of the cycle matroid of G. See for example Oxley [76] or Welsh [94]. Let r_2 denote the rank function of this matroid, when $G = K_n$. If G is a sub-graph of K_n then $r_2(G)$ is the rank of its edge-set.

We will follow the proof method in [6], [9] and [58]. Let F denote the random set of edges in the minimum weight pair of edge disjoint spanning trees. For any $0 \le p \le 1$ let G_p denote the graph induced by the edges e of K_n which satisfy $X_e \le p$. Note that G_p is distributed as $G_{n,p}$.

For any $0 \le p \le 1$, $\sum_{e \in F} 1_{(X_e > p)}$ is the number of edges of F which are not in G_p, which equals $2n - 2 - r_2(G_p)$. So,

$$\mathrm{mst}_2(K_n, \mathbf{X}) = \sum_{e \in F} X_e = \sum_{e \in F} \int_{p=0}^{1} 1_{(X_e > p)} dp = \int_{p=0}^{1} \sum_{e \in F} 1_{(X_e > p)} dp.$$

Hence, on taking expectations we obtain

$$\mathrm{mst}_2(K_n) = \int_{p=0}^{1} (2n - 2 - \mathbf{E}\left[r_2(G_p)\right]) dp. \tag{6.7}$$

It remains to estimate $\mathbf{E}\left[r_2(G_p)\right]$. The main contribution to the integral in (6.7) comes from $p = c/n$ where c is constant. Estimating $\mathbf{E}\left[r_2(G_p)\right]$ is easy enough for sufficiently small c, but it becomes more difficult for $c > c_2'$, see (6.5). When $p = \frac{c}{n}$ for $c > c_k$ we will need to be able to estimate $\mathbf{E}\left[r_k(C_{k+1}(G_{n,p}))\right]$. We give partial results for $k \ge 3$ and complete results for $k = 2$. We begin with a simple observation.

Lemma 6.2. *Let $C_{k+1} = C_{k+1}(G)$ denote the graph induced by the $(k+1)$-core of graph G (it may be an empty sub-graph). Let $E_k(G)$ denote the set of edges that are* **not** *contained in C_{k+1}. Then*

$$r_k(G) = |E_k(G)| + r_k(C_{k+1}).$$

Proof. By induction on $|V(G)|$. Trivial if $|V(G)| = 1$ and so assume that $|V(G)| > 1$. If $\delta(G) \ge k+1$ then $G = C_{k+1}$ and there is nothing to prove. Otherwise, G contains a vertex v of degree $d_G(v) \le k$. Now $G - v$ has the same $(k+1)$-core as G. If $F_1, ..., F_k$ are edge disjoint forests such that $r_k(G) = |F_1| + ... + |F_k|$ then by removing v we see, inductively, that $|E_k(G - v)| + r_k(C_{k+1}) = r_k(G - v) \ge |F_1| + ... + |F_k| - d_G(v) = r_k(G) - d_G(v)$. On the other hand $G - v$ contains k forests $F_1', ..., F_k'$ such that $r_k(G - v) = |F_1'| + ... + |F_k'| = |E_k(G - v)| + r_k(C_{k+1})$. We can then add v as a vertex of degree one to $d_G(v)$ of the forests $F_1', ..., F_k'$, implying that $r_k(G) \ge d_G(v) + |E_k(G - v)| + r_k(C_{k+1})$. Thus, $r_k(G) = d_G(v) + |E_k(G - v)| + r_k(C_{k+1}) = |E_k(G)| + r_k(C_{k+1})$. □

Lemma 6.3. *Let $k \ge 2$. If $c_k < c < c_k'$, then w.h.p.*

$$|E(G_{n,c/n})| - o(n) \le r_k(G_{n,c/n}) = |E(G_{n,c/n})|.$$

Proof. We will show that when $c < c_k'$ we can find k disjoint forests $F_1, F_2, \ldots, F_k$ contained in C_{k+1} such that

$$|E(C_{k+1})| - \sum_{i=1}^{k} |E(F_i)| = o(n). \tag{6.8}$$

This implies that $r_k(C_{k+1}) \geq |E(C_{k+1})| - o(n)$ and because $r_k(C_{k+1}) \leq |E(C_{k+1})|$ the lemma follows from this and Lemma 6.2.

Gao, Pérez-Giménez and Sato [53] show that when $c < c'_k$, no subgraph of $G_{n,p}$ has average degree more than $2k$, w.h.p. Fix $\varepsilon > 0$. Cain, Sanders and Wormald [16] proved that if the average degree of the $(k+1)$-core is at most $2k-\varepsilon$, then w.h.p. the edges of $G_{n,p}$ can be oriented so that no vertex has indegree more than k. It is clear from (6.3) that the edge density of the $(k+1)$-core increases smoothly w.h.p. and so we can apply the result of [16] for some value of ε.

It then follows that the edges of $G_{n,p}$ can be partitioned into k sets $\Phi_1, \Phi_2, \ldots, \Phi_k$ where each subgraph $H_i = ([n], \Phi_i)$ can be oriented so that each vertex has indegree at most one. We call such a graph a *Partial Functional Digraph* or PFD. Each component of a PFD is either a tree or contains exactly one cycle. We obtain $F_1, F_2, \ldots, F_k$ by removing one edge from each such cycle. We must show that w.h.p. we remove $o(n)$ vertices in total. Observe that if Z denotes the number of edges of $G_{n,p}$ that are on cycles of length at most $\omega_0 = \frac{1}{3}\log_c n$ then

$$\mathbf{E}[Z] \leq \sum_{\ell=3}^{\omega_0} \ell! \binom{n}{\ell} \ell p^\ell \leq \omega_0 c^{\omega_0} \leq n^{1/2}.$$

The Markov inequality implies that $Z \leq n^{2/3}$ w.h.p. The number of edges removed from the larger cycles to create $F_1, F_2, \ldots, F_k$ can be bounded by $kn/\omega_0 = o(n)$ and this proves (6.8) and the lemma. □

Lemma 6.4. *If $c > c'_2$, then w.h.p. the 3-core of $G_{n,c/n}$ contains two edge-disjoint forests of total size $2|V(C_3)| - o(n)$. In particular, $r_2(C_3(G_{n,c/n})) = 2|V(C_3)| - o(n)$.*

The proof of Lemma 6.4 is postponed to Section 6.6. We can now prove Theorem 6.1 (b).

6.5 Proof of Theorem 6.1 (b).

As noted in (6.7),

$$\text{mst}_2(K_n) = \int_{p=0}^{1} (2n - 2 - \mathbf{E}[r_2(G_p)])dp.$$

After changing variables to $x = pn$,

$$\text{mst}_2(K_n) = \int_{x=0}^{n} (2 - 2n^{-1} - n^{-1}\mathbf{E}\left[r_2(G_{x/n})\right])dx$$

By Lemmas 6.2 and 6.3, for $x < c'_2$ we have $\mathbf{E}\left[r_2(G_{x/n})\right] = |E(G_{x/n})| - o(n) = xn/2 - o(n)$. By Lemma 6.4, for $x > c'_2$ we have $\mathbf{E}\left[r_2(C_3(G_{x/n}))\right] = 2|V(C_3)| - o(n)$. So by Lemma 6.2 $r_2(G_{x/n}) = |E(G_{x/n})| - |E(C_3)| + 2|V(C_3)| - o(n)$, and

$$\mu_2 = \int_{x=0}^{c'_2} \left(2 - \frac{x}{2}\right) dx + \int_{x=c'_2}^{n} \left(2 - \frac{1}{n}\left(\frac{xn}{2} - |E(C_3(G_{x/n})) + 2|V(C_3(G_{x/n}))|\right)\right) dx + o(1)$$

We have from (6.3) that for $p = x/n$ we have

$$\begin{aligned} \frac{1}{n}|V(C_3)| &= \frac{f_3(\lambda)}{e^\lambda} + o(1) \\ \frac{1}{n}|E(C_3)| &= \frac{\lambda f_2(\lambda)}{2e^\lambda} + o(1) \end{aligned}$$

where λ is the largest solution to $\lambda e^\lambda / f_2(\lambda) = x$. So

$$\mu_2 = \lim_{n\to\infty} \text{mst}_2(K_n) = \int_{x=0}^{c_2'} \left(2 - \frac{x}{2}\right) dx + \int_{x=c_2'}^{\infty} \left(2 - \frac{x}{2} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right) dx$$

To calculate this, note that

$$\frac{dx}{d\lambda} = \frac{e^\lambda}{f_2(\lambda)} + \frac{\lambda e^\lambda}{f_2(\lambda)} - \frac{\lambda e^\lambda f_1(\lambda)}{f_2(\lambda)^2}$$

so

$$\int_{x=c_2'}^{\infty} \left(2 - \frac{x}{2} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right) dx$$
$$= \int_{\lambda_2'}^{\infty} \left(2 - \frac{\lambda e^\lambda}{2f_2(\lambda)} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right)\left(\frac{e^\lambda}{f_2(\lambda)} + \frac{\lambda e^\lambda}{f_2(\lambda)} - \frac{\lambda e^\lambda f_1(\lambda)}{f_2(\lambda)^2}\right) d\lambda$$

where, see (6.5),

$$\lambda_2' = g_0^{-1}(4) \approx 2.688 \tag{6.9}$$

is the unique solution to $\lambda f_2(\lambda)/f_3(\lambda) = 4$, see Section 6.8. Attempts to transform this into an explicit integral with explicit bounds have been unsuccesful. Numerical calculations give

$$\mu_2 \approx 4.1704288\ldots$$

The Inverse Symbolic Calculator (https://isc.carma.newcastle.edu.au/) has yielded no symbolic representation of this number. An apparent connection to the ζ function lies in its representation as

$$\zeta(x) = \frac{1}{\Gamma(x)} \int_{\lambda=0}^{\infty} \frac{\lambda^{x-1}}{e^\lambda - 1} d\lambda$$

which is somewhat similar to terms of the form

$$\int_{\lambda=\lambda_2'}^{\infty} \frac{\text{poly}(\lambda)}{e^\lambda - 1 - \lambda} d\lambda$$

appearing in μ_2, but no real connection has been found.

6.6 Proof of Lemma 6.4.

6.6.1 More on the 3-core.

Suppose now that $c > c_3'$ and that the 3-core C_3 of $G_{n,p}$ has $N = \Omega(n)$ vertices and M edges. It will be distributed as a random graph uniformly chosen from the set of graphs with vertex set $[N]$ and M edges and minimum degree at least three. This is an easy well known observation and follows from the fact that each such graph H can be extended in the same number of ways to a graph G with vertex set $[n]$ and m edges and such that H is the 3-core of G. We will for convenience now assume that $V(C_3) = [N]$.

The degree sequence $d(v), v \in [N]$ can be generated as follows: We independently choose for each $v \in V(C_3)$ a truncated Poisson random variable with parameter λ satisfying $g_0(\lambda) = 2M/N$, conditioned on $d(v) \geq 3$. So for $v \in [N]$,

$$\Pr\{d(v) = k\} = \frac{\lambda^k}{k! f_3(\lambda)}, \quad k = 3, 4, 5, \ldots, \quad \lambda = g_0^{-1}\left(\frac{2M}{N}\right)$$

Properties of the functions f_i, g_i are derived in Section 6.8. In particular, the g_i are strictly increasing by Lemma 6.7, so g_0^{-1} is well defined.

These independent variables are further conditioned so that the event

$$\mathcal{D} = \left\{ \sum_{v \in [N]} d(v) = 2M \right\} \tag{6.10}$$

occurs. Now λ has been chosen so that $\mathbf{E}[d(v)] = 2M/N$ and then the local central limit theorem implies that $\Pr\{\mathcal{D}\} = \Omega(1/N^{1/2})$, see for example Durrett [23]. It follows that

$$\Pr\{\mathcal{E} \mid \mathcal{D}\} \le O(n^{1/2}) \Pr\{\mathcal{E}\}, \tag{6.11}$$

for any event $\mathcal{E}$ that depends on the degree sequence of C_3.

In what follows we use the configuration model of Bollobás [11] to analyse C_3 after we have fixed its degree sequence. Thus, for each vertex v we define a set W_v of *points* such that $|W_v| = d(v)$, and write $W = \bigcup_v W_v$. A random configuration F is generated by selecting a random partition of W into M pairs. A pair $\{x, y\} \in F$ with $x \in W_u, y \in W_v$ yields an edge $\{u, v\}$ of the associated (multi-)graph Γ_F.

The key properties of F that we need are (i) conditional on F having no loops or multiple edges, it is equally likely to be any simple graph with the given degree sequence and (ii) for the degree sequences of interest, the probability that Γ_F is simple will be bounded away from zero. This is because the degree sequence in (6.11) has exponential tails. Thus we only need to show that Γ_F has certain properties w.h.p.

6.6.2 Setting up the main calculation.

Suppose now that $p = c/n$ where $c > c_2'$. We will show that w.h.p., for any fixed $\varepsilon > 0$,

$$i(S) = |\{e \in E(C_3) : e \cap S \ne \emptyset\}| \ge (2 - \varepsilon)|S| \text{ for all } S \subseteq [N]. \tag{6.12}$$

Proving this is the main computational task of the paper. In principle, it is just an application of the first moment method. We compute the expected number of S that violate (6.12) and show that tis expectation tends to zero. On the other hand, a moments glance at the expression $f(\mathbf{w})$ below shows that this is unlikely to be easy and it takes more than half of the paper to verify (6.12).

It follows from (6.12) that

$$E(C_3) \text{ can be oriented so that at least } (1-\varepsilon)N \text{ vertices have indegree at least two.} \tag{6.13}$$

To see this consider the following network flow problem. We have a source s and a sink t plus a vertex for each $v \in [N]$ and a vertex for each edge $e \in E(C_3)$. The directed edges are (i) $(s, v), v \in [N]$ of capacity two; (ii) (u, e), where $u \in e$ of infinite capacity; (iii) $(e, t), e \in E(C_3)$ of capacity one. A $s - t$ flow decomposes into paths s, u, e, t corresponding to orienting the edge e into u. A flow thus corresponds to an orientation of $E(C_3)$. The condition (6.12) implies that the minimum cut in the network has capacity at least $(2 - \varepsilon)N$. This implies that there is a flow of value at least $(2 - \varepsilon)N$ and then the orientation claimed in (6.13) exists.

Thus w.h.p. C_3 contains two edge-disjoint PFD's, each containing $(1 - \varepsilon)N$ edges. Arguing as in the proof of Lemma 6.3, we see that we can w.h.p. remove $o(N)$ edges from the cycles of these PFD's and obtain forests. Thus w.h.p. C_3 contains two edge-disjoint forests of total size at least $2(1-\varepsilon)N - o(N)$. This implies that $\mathbf{E}\left[r_2(C_3(G_{n,c/n}))\right] \geq 2(1-\varepsilon)N - o(N)$ and since $N = \Omega(n)$, we can have $\mathbf{E}\left[r_2(C_3(G_{n,c/n}))\right] = 2(1-\varepsilon)N - o(n)$. Because ε is arbitrary, this implies $r_2(C_3(G_{n,c/n})) = 2N - o(N)$ whenever $c > c_2'$.

6.6.3 Proof of (6.12): Small S.

It will be fairly easy to show that (6.13) holds w.h.p. for all $S| \leq s_0$ where

$$s_0 = \left(\frac{3(1+\varepsilon)}{e^{2+\varepsilon}c}\right)^{1/\varepsilon} n.$$

We claim that w.h.p.

$$|S| \leq s_0 \text{ implies } e(S) < (1+\varepsilon)|S| \text{ in } G_{n,p}. \tag{6.14}$$

Here $e(S) = |\{e \in E(G_{n,p}) : e \subseteq S\}|$.

Indeed,

$$\Pr\{\exists S \text{ violating (6.14)}\} \leq \sum_{s=4}^{s_0} \binom{n}{s} \binom{\binom{s}{2}}{(1+\varepsilon)s} p^{(1+\varepsilon)s} \leq$$
$$\sum_{s=4}^{s_0} \left(\frac{ne}{s}\right)^s \left(\frac{sec}{2(1+\varepsilon)n}\right)^{(1+\varepsilon)s} = \sum_{s=4}^{s_0} \left(\left(\frac{s}{n}\right)^{\varepsilon} \frac{e^{2+\varepsilon}c}{2(1+\varepsilon)}\right)^s = o(1).$$

For sets A, B of vertices and $v \in A$ we will let $d_B(v)$ denote the number of neighbors of v in B. We then let $d_B(A) = \sum_{v \in A} d_B(v)$. We will drop the subscript B when $B = [N]$.

Suppose then that (6.14) holds and that $|S| \leq s_0$ and $i(S) \leq (2-\varepsilon)|S|$. Then if $\bar{S} = [N] \setminus S$, we have

$$e(S) + d_{\bar{S}}(S) \leq (2-\varepsilon)|S| \text{ and } d(S) = 2e(S) + d_{\bar{S}}(S) \geq 3|S|$$

which implies that $e(S) \geq (1+\varepsilon)|S|$, contradiction.

6.6.4 Proof of (6.12): Large S.

Suppose now that C_3 contains an S such that $i(S) < (2-\varepsilon)|S|$. Let such sets be *bad*. Let S be a minimal bad set, and write $T = [N] \setminus S$. For any $v \in S$, we have $i(S \setminus v) \geq (2-\varepsilon)|S \setminus v|$ while $i(S) < (2-\varepsilon)|S|$. This implies $d_T(v) = i(S) - i(S \setminus v) < 2$.

We will start with a minimal bad set and then carefully add more vertices. Consider a set S such that $i(S) < 2|S|$ and $d_T(v) \leq 2$ for all $v \in S$. If there is a $w \in T$ such that $d_T(w) \leq 2$, let $S' = S \cup \{w\}$. We have $i(S') \leq i(S) + 2 < 2|S'|$. This means we may add vertices to S in this fashion to aquire a partition $[N] = S \cup T$ where $d_T(v) \leq 2$ for all $v \in S$ and $d_T(v) \geq 3$ for all $v \in T$. We further partition $S = S_0 \cup S_1 \cup S_2$ so that $d_T(v) = i$ if and only if $v \in S_i$. Denote the size of any set by its lower case equivalent, e.g. $|S_0| = s_0$.

We now start to use the configuration model. Partition each point set into $W_v = W_v^S \cup W_v^T$, where a point is in W_v^S if and only if it is matched to a point in $\cup_{u \in S} W_u$. The sizes of W_v^S, W_v^T uniquely determine $\mathbf{w} = (s_0, s_1, s_2, D_0, D_1, D_2, D_3, t, M)$. Here $D_i = d_S(S_i), i = 0, 1, 2$ and $D_3 = d_T(T)$.

Estimating the probability of w.

We have $D_i \geq (3-i)s_i$ for $i = 0,1,2$ and $D_3 \geq 3t$. Define degree sequences $(d_i^1, \ldots, d_i^{s_i})$ for $S_i, i = 0,1,2$ and $(d_3^1, \ldots, d_3^t)$ for T. Furthermore, let $\widehat{d}_1^j = d_1^j - 1$, $\widehat{d}_2^j = d_2^j - 2$ and $\widehat{d}_3^j \geq 0$ be the S-degrees of vertices in S_1, S_2, T, respectively.

Dealing with S_0:
Ignoring for the moment, that we must condition on the event $\mathcal{D}$ (see (6.10)), the probability that S_0 has degree sequence $(d_0^1, \ldots, d_0^{s_0})$, $d_0^i \geq 3$ for all i, is given by

$$\prod_{i=1}^{s_0} \frac{\lambda^{d_0^i}}{d_0^i! f_3(\lambda)}$$

where λ is the solution to

$$g_0(\lambda) = \frac{2M}{N}.$$

Hence, letting $[x^D]f(x)$ denote the coefficient of x^D in the power series $f(x)$, the probability $\pi_0(S_0, D_0)$ that $d(S_0) = D_0$ is bounded by

$$\begin{aligned}
\pi_0(S_0, D_0) \leq \sum_{\substack{d_0^1+\cdots+d_0^{s_0}=D_0 \\ d_0^i \geq 3}} \prod_{i=1}^{s_0} \frac{\lambda^{d_0^i}}{d_0^i! f_3(\lambda)} \quad &= \quad \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} \sum_{\substack{d_0^1+\cdots+d_0^{s_0}=D_0 \\ d_0^i \geq 3}} \prod_{i=1}^{s_0} \frac{1}{d_0^i!} \\
&= \quad \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} [x^{D_0}] \left(\sum_{d_0 \geq 3} \frac{x^{d_0}}{d_0!} \right)^{s_0} \\
&= \quad \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} [x^{D_0}] f_3(x)^{s_0} \\
&\leq \quad \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} \frac{f_3(\lambda_0)^{s_0}}{\lambda_0^{D_0}}
\end{aligned}$$

for all λ_0. Here we use the fact that for any function f and any $y > 0$, $[x^{D_0}]f(x) \leq f(y)/y^{D_0}$. To minimise (6.15) we choose λ_0 to be the unique solution to

$$g_0(\lambda_0) = \frac{D_0}{s_0}. \tag{6.16}$$

If $D_0 = 3s_0$ then $\lambda_0 = 0$ by Lemma 6.6, Section 6.8. In this case, since $f_3(\lambda_0) = \frac{\lambda_0^3(1+O(\lambda_0))}{6}$, we have

$$\pi_0(S_0, D_0) \leq \left(\frac{\lambda^3}{6 f_3(\lambda)} \right)^{s_0}, \quad \text{when } D_0 = 3s_0. \tag{6.17}$$

Dealing with S_1:
For each $v \in S_1$, we have $W_v = W_v^S \cup W_v^T$ where $|W_v^T| = 1$. Hence, the probability $\pi_1(S_1, D_1)$ that

$d(S_1) = D_1 + s_1$ is bounded by

$$\pi_1(S_1, D_1) \leq \sum_{\substack{\widehat{d}_1^1+\cdots+\widehat{d}_1^{s_1}=D_1 \\ \widehat{d}_1^i \geq 2}} \prod_{i=1}^{s_1} \binom{\widehat{d}_1^i + 1}{1} \frac{\lambda^{\widehat{d}_1^i+1}}{(\widehat{d}_1^i + 1)! f_3(\lambda)} = \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} \sum_{\substack{\widehat{d}_1^1+\cdots+\widehat{d}_1^{s_1}=D_1 \\ \widehat{d}_1^i \geq 2}} \prod_{i=1}^{s_1} \frac{1}{\widehat{d}_1^i!}$$
$$= \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} [x^{D_1}] f_2(x)^{s_1}$$
$$\leq \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} \frac{f_2(\lambda_1)^{s_1}}{\lambda_1^{D_1}}.$$

We choose λ_1 to satisfy the equation

$$g_1(\lambda_1) = \frac{D_1}{s_1}. \tag{6.19}$$

Similarly to what happens in (6.17) we have $\lambda_1 = 0$ when $D_1 = 2s_1$ and we have $f_2(\lambda_1) = \frac{\lambda_1^2(1+O(\lambda_1)}{2}$ and then we have

$$\pi_1(S_1, D_1) \leq \left(\frac{\lambda^3}{2f_3(\lambda)}\right)^{s_1}, \quad \text{when } D_1 = 2s_1. \tag{6.20}$$

Dealing with S_2:

For $v \in S_2$, we choose 2 points from W_v to be in W_v^T , so the probability $\pi_2(S_2, D_2)$ that $d(S_2) = D_2 + 2s_2$ is bounded by

$$\pi_2(S_2, D_2) \leq \sum_{\substack{\widehat{d}_2^1+\cdots+\widehat{d}_2^{s_2}=D_2 \\ \widehat{d}_2^i \geq 1}} \prod_{i=1}^{s_2} \binom{\widehat{d}_2^i + 2}{2} \frac{\lambda^{\widehat{d}_2^i+1}}{(\widehat{d}_2^i + 2)! f_3(\lambda)} \leq \frac{\lambda^{D_2+2s_2}}{f_3(\lambda)^{s_2}} \frac{f_1(\lambda_2)^{s_2}}{\lambda_2^{D_2}} 2^{-s_2} \tag{6.21}$$

where we choose λ_2 to satisfy the equation

$$g_2(\lambda_2) = \frac{D_2}{s_2}. \tag{6.22}$$

Similarly to what happens in (6.17) we have $\lambda_2 = 0$ when $D_2 = s_2$ and we have $f_1(\lambda_2) = \lambda_2(1 + O(\lambda_2))$ and then we have

$$\pi_2(S_2, D_2) \leq \left(\frac{\lambda^3}{2f_3(\lambda)}\right)^{s_2}, \quad \text{when } D_2 = s_2. \tag{6.23}$$

Dealing with T:

Finally, the degree of vertex i in T can be written as $d_3^i = \widehat{d}_3^i + \overline{d}_3^i$ where $\widehat{d}_3^i \geq 0$ is the S-degree and $\overline{d}_3^i \geq 3$ is the T-degree. Here, with $t = |T|$, we have

$$\sum_{i=1}^{t} \widehat{d}_3^i = d_S(T) = s_1 + 2s_2$$

by the definition of S_0, S_1, S_2. So the probability $\pi_3(T, D_3)$ that $d_T(T) = D_3$, given s_1, s_2 can be bounded by

$$\begin{aligned}
\pi_3(T, D_3) &\leq \sum_{\substack{\widehat{d}_3^1+\cdots+\widehat{d}_3^t=s_1+2s_2 \\ \widehat{d}_3^i\geq 0}} \sum_{\substack{\overline{d}_3^1+\cdots+\overline{d}_3^t=D_3 \\ \overline{d}_3^i\geq 3}} \prod_{i=1}^t \binom{\widehat{d}_3^i+\overline{d}_3^i}{\widehat{d}_3^i} \frac{\lambda^{\widehat{d}_3^i+\overline{d}_3^i}}{(\widehat{d}_3^i+\overline{d}_3^i)!f_3(\lambda)} \\
&= \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \sum_{\substack{\widehat{d}_3^1+\cdots+\widehat{d}_3^t=s_1+2s_2 \\ \widehat{d}_3^i\geq 0}} \sum_{\substack{\overline{d}_3^1+\cdots+\overline{d}_3^t=D_3 \\ \overline{d}_3^i\geq 3}} \prod_{i=1}^t \frac{1}{\widehat{d}_3^i!\overline{d}_3^i!} \\
&= \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \left([x^{D_3}]f_3(x)^t\right)\left([x^{s_1+2s_2}]e^x\right) \\
&\leq \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \frac{f_3(\lambda_3)^t}{\lambda_3^{D_3}} \frac{t^{s_1+2s_2}}{(s_1+2s_2)!},
\end{aligned} \tag{6.24}$$

where we choose λ_3 to satisfy the equation

$$g_0(\lambda_3) = \frac{D_3}{t}. \tag{6.25}$$

Similarly to what happens in (6.17) we have $\lambda_3 = 0$ when $D_3 = 3t$ and we have $f_3(\lambda_3) = \frac{\lambda_3^3(1+O(\lambda_1))}{6}$ and then we have

$$\pi_3(T, D_3) \leq \frac{\lambda^{D_3+s_1+2s_2}}{(6f_3(\lambda))^t} \frac{t^{s_1+2s_2}}{(s_1+2s_2)!}, \quad \text{when } D_3 = 3t.$$

Putting the bounds together.

For a fixed $\mathbf{w} = (s_0, s_1, s_2, D_0, D_1, D_2, D_3, t, M)$, there are $\binom{t+s}{s_0,s_1,s_2,t}$ choices for S_0, S_1, S_2, T. Having chosen these sets we partition the $W_v, v \in S$ into $W_v^S \cup W_v^T$. Note that our expressions (6.15), (6.18), (6.21), (6.24) account for these choices. Given the partitions of the W_v's, there are $(D_0 + D_1 + D_2)!!D_3!!(s_1 + 2s_2)!$ configurations, where $(2s)!! = (2s-1)\times(2s-3)\times\cdots\times 3\times 1$ is the number of ways of partitioning a set of size $2s$ into s pairs. Here $(D_0 + D_1 + D_2)!!$ is the number of ways of pairing up $\bigcup_{v\in S} W_v^S$, $D_3!!$ is the number of ways of pairing up $\bigcup_{v\in T} W_v^T$ and $(s_1 + 2s_2)!$ is the number of ways of pairing points associated with S to points associated with T. Each configuration has probability $1/(2M)!!$. So, the total probability of all configurations whose vertex partition and degrees are described by $\mathbf{w}$ can be bounded by

$$\begin{aligned}
&\binom{t+s}{s_0,s_1,s_2,t} \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} \frac{f_3(\lambda_0)^{s_0}}{\lambda_0^{D_0}} \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} \frac{f_2(\lambda_1)^{s_1}}{\lambda_1^{D_1}} \frac{\lambda^{D_2+2s_2}}{f_3(\lambda)^{s_2}} \frac{f_1(\lambda_2)^{s_2}}{\lambda_2^{D_2}} 2^{-s_2} \\
&\times \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \frac{f_3(\lambda_3)^t}{\lambda_3^{D_3}} \frac{t^{s_1+2s_2}}{(s_1+2s_2)!} \frac{(D_0+D_1+D_2)!!D_3!!(s_1+2s_2)!}{(2M)!!} \\
&= \binom{t+s}{s_0,s_1,s_2,t} \frac{\lambda^{2M}}{f_3(\lambda)^N} \frac{f_3(\lambda_0)^{s_0}}{\lambda_0^{D_0}} \frac{f_2(\lambda_1)^{s_1}}{\lambda_1^{D_1}} \frac{f_1(\lambda_2)^{s_2}}{\lambda_2^{D_2}} 2^{-s_2} \frac{f_3(\lambda_3)^t}{\lambda_3^{D_3}} \frac{t^{s_1+2s_2}}{(s_1+2s_2)!} \\
&\times \frac{(D_0+D_1+D_2)!!D_3!!(s_1+2s_2)!}{(2M)!!}
\end{aligned}$$

Write $D_i = \Delta_i s$, $|S_i| = \sigma_i s$, $t = \tau s$, $M = \mu s$ and $N = \nu s$. We have $k!! \sim \sqrt{2}(k/e)^{k/2}$ as $k \to \infty$ by Stirling's formula, so the expression above, modulo an $e^{o(s)}$ factor, can be written as

$$f(\mathbf{w})^s = \left(\frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau} \frac{\lambda^{2\mu}}{f_3(\lambda)^\nu} \frac{f_3(\lambda_0)^{\sigma_0}}{\lambda_0^{\Delta_0}} \frac{f_2(\lambda_1)^{\sigma_1}}{\lambda_1^{\Delta_1}} \frac{f_1(\lambda_2)^{\sigma_2}}{\lambda_2^{\Delta_2}} \frac{f_3(\lambda_3)^{\tau}}{\lambda_3^{\Delta_3}} \frac{(\tau e)^{\sigma_1+2\sigma_2}}{2^{\sigma_2}} \frac{(\Delta_0+\Delta_1+\Delta_2)^{(\Delta_0+\Delta_1+\Delta_2)/2}\Delta_3^{\Delta_3/2}}{(2\mu)^\mu} \right)^s$$

We note that

$$\begin{aligned} \sigma_2 &= 1-\sigma_0-\sigma_1, \\ \Delta_3 &= 2\mu - \Delta_0 - \Delta_1 - \Delta_2 - 2\sigma_1 - 4\sigma_2 \\ &= 2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1 \\ \nu &= 1+\tau. \end{aligned} \tag{6.26}$$

Hence σ_2, Δ_3, ν may be eliminated, and we can consider $\mathbf{w}$ to be $(\sigma_0, \sigma_1, \Delta_0, \Delta_1, \Delta_2, \tau, \mu)$. When convenient, Δ_3 may be used to denote $2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1$. Define the constraint set F to be all $\mathbf{w}$ satisfying

$$\Delta_0 \geq 3\sigma_0, \Delta_1 \geq 2\sigma_1, \Delta_2 \geq 1-\sigma_0-\sigma_1, \Delta_3 \geq 3\tau. \tag{6.27a}$$

$$\frac{\Delta_0+\Delta_1+\Delta_2}{2} + \sigma_1 + 2(1-\sigma_0-\sigma_1) < 2-\varepsilon \qquad \text{since } i(S) < (2-\varepsilon)|S|, \qquad \text{see (6.12)} \tag{6.27b}$$

$\sigma_0, \sigma_1 \geq 0, \sigma_0 + \sigma_1 \leq 1.$

$0 \leq \tau \leq (1-\varepsilon)/\varepsilon$ since $|S| \geq \varepsilon N$.

$\mu \geq (2+\varepsilon)(1+\tau)$ since $M \geq (2+\varepsilon)N$.

$$\sigma_0 < 1, \qquad \text{otherwise } C_3 \text{ is not connected.} \tag{6.27c}$$

Here ε is a sufficiently small positive constant such that (i) we can exclude the case of small S, (ii) satisfy condition (6.12) and (iii) have $M \geq (2+\varepsilon)N$ since $c > c_2'$.

For a given s, there are $O(\text{poly}(s))$ choices of $\mathbf{w} \in F$, and the probability that the randomly chosen configuration corresponds to a $\mathbf{w} \in F$ can be bounded by

$$\sum_{s \geq \varepsilon N} \sum_{\mathbf{w}} O(\text{poly}(s)) f(\mathbf{w})^s \leq \sum_s (e^{o(1)} \max_F f(\mathbf{w}))^s \leq N(e^{o(1)} \max_F f(\mathbf{w}))^{\varepsilon N}. \tag{6.28}$$

As $N \to \infty$, it remains to show that $f(\mathbf{w}) \leq 1-\delta$ for all $\mathbf{w} \in F$, for some $\delta = \delta(\varepsilon) > 0$. At this point we remind the reader that we have so far ignored conditioning on the event $\mathcal{D}$ defined in (6.10). Inequality (6.11) implies that it is sufficient to inflate the RHS of (6.28) by $O(n^{1/2})$ to obtain our result.

So, let

$$\begin{aligned} & f(\Delta_0, \Delta_1, \Delta_2, \sigma_0, \sigma_1, \tau, \mu) = \\ & \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau} \frac{\lambda^{2\mu}}{f_3(\lambda)^{\tau+1}} \frac{f_3(\lambda_0)^{\sigma_0}}{\lambda_0^{\Delta_0}} \frac{f_2(\lambda_1)^{\sigma_1}}{\lambda_1^{\Delta_1}} \frac{f_1(\lambda_2)^{1-\sigma_0-\sigma_1}}{\lambda_2^{\Delta_2}} \frac{f_3(\lambda_3)^{\tau}}{\lambda_3^{\Delta_3}} \\ \times \; & \frac{(e\tau)^{2-2\sigma_0-\sigma_1}}{2^{1-\sigma_0-\sigma_1}} \frac{(\Delta_0+\Delta_1+\Delta_2)^{(\Delta_0+\Delta_1+\Delta_2)/2}\Delta_3^{\Delta_3/2}}{(2\mu)^\mu} \end{aligned}$$

We complete the proof of Theorem 6.1(b) by showing that

$$f(\mathbf{w}) \le \exp\left\{-\frac{\varepsilon^2}{3}\right\} \text{ for all } \mathbf{w} \in F.$$

Eliminating μ

We begin by showing that it is enough to consider $\mu = (2+\varepsilon)(1+\tau)$. We collect all terms involving μ, including Δ_3, λ and λ_3 whose values are determined in part by μ. It is enough to consider the logarithm of f. We have

$$\frac{\partial \log f}{\partial \mu} = 2\log\lambda + \frac{\partial \lambda}{\partial \mu}\left(\frac{2\mu}{\lambda} - \nu\frac{f_2(\lambda)}{f_3(\lambda)}\right) + \frac{\partial \lambda_3}{\partial \mu}\left(\tau\frac{f_2(\lambda_3)}{f_3(\lambda_3)} - \frac{\Delta_3}{\lambda_3}\right) - 2\log\lambda_3 + \log\Delta_3 + 1 - \log 2\mu - 1$$

by definition of λ, λ_3, we have

$$\frac{2\mu}{\lambda} - \nu\frac{f_2(\lambda)}{f_3(\lambda)} = 0 \text{ and } \frac{\Delta_3}{\lambda_3} - \tau\frac{f_2(\lambda_3)}{f_3(\lambda_3)} = 0,$$

and so

$$\frac{\partial \log f}{\partial \mu} = 2\log\left(\frac{\lambda}{\lambda_3}\right) + \log\left(\frac{\Delta_3}{2\mu}\right)$$

We have $\Delta_3 \le 2\mu$ and furthermore, $\lambda \le \lambda_3$ since g_0 is an increasing function. Indeed, writing $\iota = i(S)/s \le 2$, we have $\Delta_3 + 2\iota = 2\mu \ge 4(\tau+1)$, so

$$g_0(\lambda_3) - g_0(\lambda) = \frac{\Delta_3}{\tau} - \frac{2\mu}{\nu} = \frac{2\mu - 2\iota}{\tau} - \frac{2\mu}{\tau+1} = \frac{2\mu - 2\iota(\tau+1)}{\tau(\tau+1)} \ge \frac{4-2\iota}{\tau} \ge 0.$$

This shows that $\log f$ is decreasing with respect to μ, and in discussing the maximum value of f for $\mu \ge (2+\varepsilon)(1+\tau)$ we may assume that $\mu = (2+\varepsilon)(1+\tau)$.

We now argue that to show that $f \le \exp\{-\varepsilon^2/3\}$ when $\mu = (2+\varepsilon)(1+\tau)$, it is enough to show that $f \le 1$ when $\mu = 2(1+\tau)$. Let $2(1+\tau) < \mu < (2+\varepsilon)(1+\tau)$. Then by (6.26) and (6.27a)

$$\begin{aligned} \Delta_3 &= 2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1 \\ &\le 2\mu - 4 - 3\sigma_0 - 2\sigma_1 - (1 - \sigma_0 - \sigma_1) + 4\sigma_0 + 2\sigma_1 \\ &= 2\mu - 5 + 2\sigma_0 + \sigma_1 \\ &\le 2\mu - 2 \end{aligned}$$

and since $\tau \le 1/\varepsilon - 1$, $\mu \le (2+\varepsilon)(1+\tau)$ implies $\mu \le 2/\varepsilon + 1 < 3/\varepsilon$. So,

$$\frac{\partial \log f}{\partial \mu} \le 2\log\left(\frac{\lambda}{\lambda_3}\right) + \log\left(\frac{2\mu-2}{2\mu}\right) \le \log\left(1 - \frac{\varepsilon}{3}\right)$$

So, fixing $\mathbf{w}' = (\sigma_0, \sigma_1, \Delta_0, \Delta_1, \Delta_2, \tau)$, let $\mu = 2(1+\tau)$ and $\mu' = (2+\varepsilon)(1+\tau)$. If $f(\mathbf{w}', \mu) \le 1$, then

$$\log f(\mathbf{w}', \mu') \le \log f(\mathbf{w}', \mu) + \varepsilon(1+\tau)\log\left(1 - \frac{\varepsilon}{3}\right) \le -\frac{\varepsilon^2}{3}.$$

This shows that it is enough to prove that $f(\mathbf{w}) \le 1$ for $\mathbf{w} \in F'$, defined by

$$\Delta_0 \ge 3\sigma_0, \Delta_1 \ge 2\sigma_1, \Delta_2 \ge 1-\sigma_0-\sigma_1, \Delta_3 \ge 3\tau \tag{6.29a}$$
$$\Delta_0+\Delta_1+\Delta_2 \le 4\sigma_0+2\sigma_1 \tag{6.29b}$$
$$\sigma_0, \sigma_1 \ge 0, \sigma_0+\sigma_1 \le 1$$
$$0 \le \tau < \infty$$
$$\mu = 2(1+\tau).$$

We have relaxed equation (6.27b) to give (6.29b) in order to simplify later calculations. In F', λ is defined by

$$g_0(\lambda) = \frac{2\mu}{\nu} = \frac{4(1+\tau)}{1+\tau} = 4,$$

so in the remainder of the proof

$$\lambda = g_0^{-1}(4) \approx 2.688 \text{ is fixed.}$$

It will be convenient at times to write $\Delta = \Delta_0+\Delta_1+\Delta_2$. We observe that $3\sigma_0+2\sigma_1+(1-\sigma_0-\sigma_1) = 2\sigma_0+\sigma_1+1$, so by (6.29a), (6.29b),

$$2\sigma_0+\sigma_1+1 \le \Delta \le 4\sigma_0+2\sigma_1. \tag{6.30}$$

Note also that $\mu = 2(1+\tau)$ implies

$$\Delta_3 = 2\mu-4-\Delta_0-\Delta_1-\Delta_2+4\sigma_0+2\sigma_1 = 4\tau+4\sigma_0+2\sigma_1-\Delta. \tag{6.31}$$

The quantity $2\sigma_0+\sigma_1$ will appear frequently. We note that (6.30) and $\sigma_0+\sigma_1 \le 1$ imply

$$1 \le 2\sigma_0+\sigma_1 \le 2.$$

Eliminating τ

We now turn to choosing the optimal τ. With $\mu = 2(1+\tau)$,

$$\begin{aligned} f(\sigma_0,\sigma_1,\Delta_0,\Delta_1,\Delta_2,\tau) &= \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau}\left(\frac{\lambda^4}{f_3(\lambda)}\right)^{\tau+1}\frac{f_3(\lambda_0)^{\sigma_0}}{\lambda_0^{\Delta_0}}\frac{f_2(\lambda_1)^{\sigma_1}}{\lambda_1^{\Delta_1}} \\ &\quad\times\frac{f_1(\lambda_2)^{1-\sigma_0-\sigma_1}}{\lambda_2^{\Delta_2}}\frac{f_3(\lambda_3)^\tau}{\lambda_3^{\Delta_3}}\frac{(e\tau)^{2-2\sigma_0-\sigma_1}}{2^{1-\sigma_0-\sigma_1}}\times\frac{\Delta^{\Delta/2}\Delta_3^{\Delta_3/2}}{(4+4\tau)^{2+2\tau}}. \end{aligned}$$

Here $\lambda_0 = \lambda_0(\Delta_0,\sigma_0)$, $\lambda_1 = \lambda_1(\Delta_1,\sigma_1)$, $\lambda_2 = \lambda_2(\Delta_2,\sigma_0,\sigma_1)$, $\lambda_3 = \lambda_3(\sigma_0,\sigma_1,\Delta_0,\Delta_1,\Delta_2,\tau)$ as defined in (6.16), (6.19), (6.22), (6.25). Since $\tau f_2(\lambda_3)/f_3(\lambda_3)-\Delta_3/\lambda_3 = 0$ by the definition of λ_3, the partial derivative of $\log f$ with respect to τ is given by

$$\begin{aligned} \frac{\partial}{\partial\tau}\log f(\sigma_0,\sigma_1,\Delta_0,\Delta_1,\Delta_2,\tau) &= \log(\tau+1)+1-\log\tau-1+\log\left(\frac{\lambda^4}{f_3(\lambda)}\right) \\ &\quad+\frac{\partial\lambda_3}{\partial\tau}\left(\tau\frac{f_2(\lambda_3)}{f_3(\lambda_3)}-\frac{\Delta_3}{\lambda_3}\right)+\log(f_3(\lambda_3))-4\log\lambda_3 \\ &\quad+\frac{2-2\sigma_0-\sigma_1}{\tau}+2\left(1+\log\Delta_3\right)-2\log(4+4\tau)-2 \\ &= \log(\tau+1)-\log\tau+\log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right)+\frac{2-2\sigma_0-\sigma_1}{\tau} \\ &\quad+2\log\Delta_3-2\log(4+4\tau) \end{aligned}$$

This is positive for τ close to zero. This is clear as long as $2\sigma_0+\sigma_1<2$. But if $2\sigma_0+\sigma_1=2$ then $\sigma_0+\sigma_1\le 1$ implies that $\sigma_0=1,\sigma_1=0$. But then if $\tau>0$ we have that C_3 is not connected and that if $\tau=0$, $S=[N]$ which violates (6.27c). On the other hand, $\frac{\partial}{\partial\tau}\log f$ vanishes if

$$2-2\sigma_0-\sigma_1-\tau\left[\log\left(1+\frac{1}{\tau}\right)-2\log\left(\frac{\Delta_3}{4\tau}\right)-\log\left(\frac{\lambda^4 f_3(\lambda_3)}{\lambda_3^4 f_3(\lambda)}\right)\right]=0. \tag{6.33}$$

So any local maximum of f must satisfy this equation. If no solution exists, then it is optimal to let $\tau\to\infty$. We will see below how to choose τ to guarantee maximality. For now, we only assume τ satisfies (6.33).

Eliminating $\Delta_0,\Delta_1,\Delta_2$.

We now eliminate $\Delta_0,\Delta_1,\Delta_2$. Fix σ_0,σ_1. For $\Delta_i>(3-i)\sigma_i$ such that $\Delta_0+\Delta_1+\Delta_2<4\sigma_0+2\sigma_1$,

$$\begin{aligned}\frac{\partial}{\partial\Delta_i}\log f &= \frac{\partial\lambda_i}{\partial\Delta_i}\left(\sigma_i\frac{f_{2-i}(\lambda_i)}{f_{3-i}(\lambda_i)}-\frac{\Delta_i}{\lambda_i}\right)-\log\lambda_i+\log\lambda_3\\ &\quad+\frac{\partial}{\partial\tau}\log f\ \ \frac{\partial\tau}{\partial\Delta_i}+\frac12\log\Delta+\frac12-\frac12\log\Delta_3-\frac12\\ &= -\log\lambda_i+\log\left(\lambda_3\sqrt{\frac{\Delta}{\Delta_3}}\right),\end{aligned}$$

since $g_i(\lambda_i)=\Delta_i/\sigma_i$ by definition of λ_i, and the term $\frac{\partial}{\partial\tau}\log f\ \ \partial\tau/\partial\Delta_i$ vanishes because (6.33) is assumed to hold. We note that $\lambda_i>0$ when $\Delta_i>(3-i)\sigma_i$ (Section 6.8), allowing division by λ_i.

As Δ_i tends to its lower bound $(3-i)\sigma_i$, we have $\log\lambda_i\to-\infty$ while the other terms remain bounded, so the derivative is positive at the lower bound of Δ_i. Any stationary point must satisfy $\lambda_0=\lambda_1=\lambda_2=\lambda_3\sqrt{\Delta/\Delta_3}=:\widehat{\lambda}$. This can only happen if

$$\sigma_0 g_0(\widehat{\lambda})+\sigma_1 g_1(\widehat{\lambda})+(1-\sigma_0-\sigma_1)g_2(\widehat{\lambda})=\sigma_0\frac{\Delta_0}{\sigma_0}+\sigma_1\frac{\Delta_1}{\sigma_1}+(1-\sigma_0-\sigma_1)\frac{\Delta_2}{1-\sigma_0-\sigma_1}=\Delta.$$

So we choose $\widehat{\lambda}$, Δ,τ to solve the system of equations

$$\begin{aligned}\widehat{\lambda} &= \lambda_3\sqrt{\frac{\Delta}{\Delta_3}}\\ \Delta &= \sigma_0 g_0(\widehat{\lambda})+\sigma_1 g_1(\widehat{\lambda})+(1-\sigma_0-\sigma_1)g_2(\widehat{\lambda})\\ 2-2\sigma_0-\sigma_1 &= \tau\left[\log\left(1+\frac{1}{\tau}\right)-2\log\left(\frac{\Delta_3}{4\tau}\right)-\log\left(\frac{\lambda^4 f_3(\lambda_3)}{\lambda_3^4 f_3(\lambda)}\right)\right]\end{aligned}$$

In Section 6.7 we show that this system has no solution such that $2\sigma_0+\sigma_1+1\le\Delta\le 4\sigma_0+2\sigma_1$ (see (6.30)). This means that no stationary point exists, and $\log f$ is increasing in each of $\Delta_0,\Delta_1,\Delta_2$. In particular, it is optimal to set

$$\Delta_0+\Delta_1+\Delta_2=4\sigma_0+2\sigma_1 \text{ which implies that } \Delta_3=4\tau, \text{ see (6.31).} \tag{6.36}$$

This eliminates one degree of freedom. We now set

$$\Delta_2=4\sigma_0+2\sigma_1-\Delta_0-\Delta_1.$$

Then for Δ_0, Δ_1, we have

$$\frac{\partial}{\partial \Delta_i} \log f = -\log \lambda_i + \log \lambda_2, \quad i = 0, 1.$$

To see this note that (6.34) has to be modified via the addition of $\frac{\partial}{\partial \Delta_2} \log f \times \frac{\partial \Delta_2}{\partial \Delta_i}$, for $i = 0, 1$.

So it is optimal to let $\lambda_0 = \lambda_1 = \lambda_2 = \overline{\lambda}$, defined by

$$\sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1 - \sigma_0 - \sigma_1) g_2(\overline{\lambda}) = 4\sigma_0 + 2\sigma_1 \tag{6.37}$$

This has a unique solution $\overline{\lambda} \geq 0$ whenever $2\sigma_0 + \sigma_1 \geq 1$, since for fixed σ_0, σ_1, the left-hand side is a convex combination of increasing functions, by Lemma 6.7, Section 6.8. This defines $\Delta_i = \Delta_i(\sigma_0, \sigma_1)$ by

$$\Delta_0 = g_0(\overline{\lambda})\sigma_0, \quad \Delta_1 = g_1(\overline{\lambda})\sigma_1, \quad \Delta_2 = g_2(\overline{\lambda})(1 - \sigma_0 - \sigma_1) \tag{6.38}$$

We note at this point that $\overline{\lambda} \leq \lambda$. Indeed, by (6.36) and (6.27a),

$$\Delta_0 = 4\sigma_0 + 2\sigma_1 - \Delta_1 - \Delta_2 \leq 4\sigma_0 + 2\sigma_1 - 2\sigma_1 - (1 - \sigma_0 - \sigma_1) \leq 4\sigma_0,$$

so

$$g_0(\overline{\lambda}) = \frac{\Delta_0}{\sigma_0} \leq 4 = g_0(\lambda) \tag{6.39}$$

implying that $\overline{\lambda} \leq \lambda$, since g_0 is increasing.

This choice (6.38) of $\Delta_0, \Delta_1, \Delta_2$ simplifies f significantly. With $\Delta = 4\sigma_0 + 2\sigma_1$ we have $\Delta_3 = 4\tau$, see (6.36), and so

$$\lambda_3 = g_0^{-1}\left(\frac{4\tau}{\tau}\right) = \lambda \tag{6.40}$$

is fixed. In particular, the relation (6.33) for τ simplifies to

$$2 - 2\sigma_0 - \sigma_1 = \tau \log\left(1 + \frac{1}{\tau}\right) \tag{6.41}$$

Let $\phi(\tau) = \tau \log(1 + 1/\tau)$. Then $\phi''(\tau) = -\tau^{-1}(\tau+1)^{-2}$, so ϕ is concave and then $\lim_{\tau \to 0} \phi(\tau) = 0$, $\lim_{\tau \to \infty} \phi(\tau) = 1$ implies that ϕ is strictly increasing and takes values in $[0, 1)$ for $\tau \geq 0$. This means that (6.41) has a unique solution if and only if $2\sigma_0 + \sigma_1 > 1$. When $2\sigma_0 + \sigma_1 = 1$, f is increasing with respect to τ, and we treat this case now.

If $2\sigma_0 + \sigma_1 = 1$, then (6.30) implies that $\Delta = 2$. Furthermore, $\Delta_3 = 4\tau$ (see (6.31)) and $\lambda_3 = \lambda$ (see (6.40)) and $g_i(0) = 3 - i$ implies that

$$\sigma_0 g_0(0) + \sigma_1 g_1(0) + (1 - \sigma_0 - \sigma_1) g_2(0) = 2\sigma_0 + \sigma_1 + 1 = 4\sigma_0 + 2\sigma_1,$$

so $\overline{\lambda} = 0$ is the unique solution to (6.37). Then since $\Delta_i/\sigma_i = g_i(0) = 3 - i$ (Lemma 6.6, Section 6.8), we have $\Delta_i = (3 - i)\sigma_i$, $i = 0, 1, 2$, and as in (6.17), (6.20), (6.23),

$$\frac{f_3(\overline{\lambda})^{\sigma_0} f_2(\overline{\lambda})^{\sigma_1} f_1(\overline{\lambda})^{1-\sigma_0-\sigma_1}}{\overline{\lambda}^{\Delta}} = \left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^3}\right)^{\sigma_0} \left(\frac{f_2(\overline{\lambda})}{\overline{\lambda}^2}\right)^{\sigma_1} \left(\frac{f_1(\overline{\lambda})}{\overline{\lambda}}\right)^{1-\sigma_0-\sigma_1} = \frac{1}{6^{\sigma_0}} \frac{1}{2^{\sigma_1}}$$

so when $2\sigma_0+\sigma_1=1$, (6.32) becomes

$$f(\sigma_0,\sigma_1,\tau)=\frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau}\frac{\lambda^4}{f_3(\lambda)}\frac{1}{6^{\sigma_0}}\frac{1}{2^{\sigma_1}}\frac{e\tau}{2^{1-\sigma_0-\sigma_1}}\frac{2^{2/2}(4\tau)^{2\tau}}{(4+4\tau)^{2+2\tau}}.$$

In this computation we also used the fact that $\lambda=\lambda_3$ (see (6.40)) and $\Delta_3=4\tau$ (see (6.31)) to find that

$$\left(\frac{\lambda^4}{f_3(\lambda)}\right)^{\tau+1}\frac{f_3(\lambda_3)^\tau}{\lambda_3^{\Delta_3}}=\frac{\lambda^4}{f_3(\lambda)}.$$

Here $\lambda^4/f_3(\lambda)\approx 7.05$ is fixed. We show in Section 6.7 that in this case, the partial derivative in τ is positive for all τ, so we let $\tau\to\infty$. Substituting $\sigma_1=1-2\sigma_0$ we are reduced to

$$\begin{aligned}f(\sigma_0) &= \lim_{\tau\to\infty}\frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}(1-2\sigma_0)^{(1-2\sigma_0)}\sigma_0^{\sigma_0}\tau^\tau}\frac{\lambda^4}{f_3(\lambda)}\frac{1}{6^{\sigma_0}}\frac{1}{2^{1-2\sigma_0}}\frac{e\tau}{2^{\sigma_0}}\frac{2(4\tau)^{2\tau}}{(4+4\tau)^{2+2\tau}}\\ &= \frac{\lambda^4}{16f_3(\lambda)}\frac{1}{\sigma_0^{2\sigma_0}(1-2\sigma_0)^{1-2\sigma_0}3^{\sigma_0}}\end{aligned}$$

This has the stationary point $\sigma_0=2-\sqrt{3}$, and $f(2-\sqrt{3})\approx 0.95$. We also have $f(0)\approx 0.44$ and $f(1/2)\approx 0.51$ at the lower and upper bounds for σ_0.

Dealing with σ_0,σ_1

With this, we have reduced our analysis to the variables σ_0,σ_1 in the domain

$$E=\{(\sigma_0,\sigma_1):\sigma_0,\sigma_1\ge 0,\sigma_0+\sigma_1\le 1,2\sigma_0+\sigma_1\ge 1\}.$$

We just showed that $f\le 1$ in

$$E_0=\{(\sigma_0,\sigma_1)\in E:2\sigma_0+\sigma_1=1\}.$$

Further define

$$\begin{aligned}E_1&=\{(\sigma_0,\sigma_1)\in E:0.01\le\sigma_1\le 0.99\},\\ E_2&=\{(\sigma_0,\sigma_1)\in E:0\le\sigma_1<0.01\},\\ E_3&=\{(\sigma_0,\sigma_1)\in E:0.99<\sigma_1\le 1\}.\end{aligned}$$

We will show that $f\le 1$ in each of these sets, whose union covers E.

From this point on, let $\partial_i=\frac{\partial}{\partial\sigma_i}, i=0,1$. As noted above, $\Delta=4\sigma_0+2\sigma_1$ simplifies f. Specifically, if $2\sigma_0+\sigma_1>1$ then (6.32) becomes, after using (6.36) and (6.40),

$$\begin{aligned}f(\sigma_0,\sigma_1) &= \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau}\frac{\lambda^4}{f_3(\lambda)}\frac{f_3(\overline{\lambda})^{\sigma_0}f_2(\overline{\lambda})^{\sigma_1}f_1(\overline{\lambda})^{1-\sigma_0-\sigma_1}}{\overline{\lambda}^{4\sigma_0+2\sigma_1}}\\ &\quad\times\frac{(e\tau)^{2-2\sigma_0-\sigma_1}}{2^{1-\sigma_0-\sigma_1}}\frac{(4\sigma_0+2\sigma_1)^{2\sigma_0+\sigma_1}(4\tau)^{2\tau}}{(4+4\tau)^{2+2\tau}}\end{aligned}$$

In (6.41), (6.37) respectively, τ and $\overline{\lambda}$ are given as functions of σ_0, σ_1. Recall that $\lambda = g_0^{-1}(4)$ is constant. So

$$
\begin{aligned}
&\partial_0 \log f(\sigma_0, \sigma_1) = \\
&- \log \sigma_0 - 1 + \log(1 - \sigma_0 - \sigma_1) + 1 + \log f_3(\overline{\lambda}) - \log f_1(\overline{\lambda}) \\
&- 4 \log \overline{\lambda} - 2 \log(e\tau) + \log 2 + 2 \log(4\sigma_0 + 2\sigma_1) + 2 \\
&+ \frac{\partial \overline{\lambda}}{\partial \sigma_0} \left(\sigma_0 \frac{f_2(\overline{\lambda})}{f_3(\overline{\lambda})} + \sigma_1 \frac{f_1(\overline{\lambda})}{f_2(\overline{\lambda})} + (1 - \sigma_0 - \sigma_1) \frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{4\sigma_0 + 2\sigma_1}{\overline{\lambda}} \right) \\
&+ \frac{\partial \tau}{\partial \sigma_0} \left(\log(\tau + 1) + 1 - \log \tau - 1 + \frac{2 - 2\sigma_0 - \sigma_1}{\tau} + 2 \log 4\tau + 2 - 2 \log(4 + 4\tau) - 2 \right) \\
&= \log \left(\frac{1 - \sigma_0 - \sigma_1}{\sigma_0} \right) + \log \left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^4 f_1(\overline{\lambda})} \right) - 2 \log \tau + \log 2 + 2 \log(4\sigma_0 + 2\sigma_1) \qquad (6.43)
\end{aligned}
$$

where, as expected, the terms involving $\partial_0 \tau$ and $\partial_0 \overline{\lambda}$ vanish since $\tau, \overline{\lambda}$ were chosen to maximize $\log f$. (See (6.41) and (6.37) respectively).

Similarly,

$$
\begin{aligned}
&\partial_1 \log f(\sigma_0, \sigma_1) = \\
&- \log \sigma_1 - 1 + \log(1 - \sigma_0 - \sigma_1) + 1 + \log f_2(\overline{\lambda}) - \log f_1(\overline{\lambda}) \\
&- 2 \log \overline{\lambda} - \log(e\tau) + \log 2 + \log(4\sigma_0 + 2\sigma_1) + 1 \\
&+ \frac{\partial \overline{\lambda}}{\partial \sigma_1} \left(\sigma_0 \frac{f_2(\overline{\lambda})}{f_3(\overline{\lambda})} + \sigma_1 \frac{f_1(\overline{\lambda})}{f_2(\overline{\lambda})} + (1 - \sigma_0 - \sigma_1) \frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{4\sigma_0 + 2\sigma_1}{\overline{\lambda}} \right) \\
&+ \frac{\partial \tau}{\partial \sigma_1} \left(\log(\tau + 1) + 1 - \log \tau - 1 + \frac{2 - 2\sigma_0 - \sigma_1}{\tau} + 2 \log 4\tau + 2 - 2 \log(4 + 4\tau) - 2 \right) \\
&= \log \left(\frac{1 - \sigma_0 - \sigma_1}{\sigma_1} \right) + \log \left(\frac{f_2(\overline{\lambda})}{\overline{\lambda}^2 f_1(\overline{\lambda})} \right) - \log \tau + \log 2 + \log(4\sigma_0 + 2\sigma_1). \qquad (6.44)
\end{aligned}
$$

Any stationary point must satisfy

$$
(\partial_0 - 2\partial_1) \log f = \log \left(\frac{\sigma_1^2}{\sigma_0(1 - \sigma_0 - \sigma_1)} \right) + \log \left(\frac{f_1(\overline{\lambda}) f_3(\overline{\lambda})}{f_2(\overline{\lambda})^2} \right) - \log 2 = 0. \qquad (6.45)
$$

Now we show in Lemma 6.8, Section 6.8 that

$$
1 \leq \frac{f_2(\overline{\lambda})^2}{f_1(\overline{\lambda}) f_3(\overline{\lambda})} \leq 2.
$$

This means from (6.45) that if $(\partial_0 - 2\partial_1) \log f = 0$ then

$$
2 \leq \frac{\sigma_1^2}{\sigma_0(1 - \sigma_0 - \sigma_1)} \leq 4.
$$

In particular, the lower bound implies $\sigma_0 \geq (1 - \sigma_1)/2 + \sqrt{1 - 2\sigma_1 - \sigma_1^2}/2$ and the upper bound implies $\sigma_1 \leq -2\sigma_0 + \sqrt{4\sigma_0 - 4\sigma_0^2}$. The latter bound is used only to conclude that $\sigma_1 < 1/2$, by noting that $-2\sigma_0 + \sqrt{4\sigma_0 - 4\sigma_0^2} \leq (5^{1/2} - 1)/3 < 1/2$ for $0 \leq \sigma_0 \leq 1$. In conclusion,

$$
(\partial_0 - 2\partial_1) \log f = 0 \Longrightarrow \begin{cases} \sigma_0 \geq (1 - \sigma_1)/2 + \sqrt{1 - 2\sigma_1 - \sigma_1^2}/2. \\ \sigma_1 < 1/2. \end{cases} \qquad (6.46)
$$

Case One. $E_1 = \{(\sigma_0, \sigma_1) \in E : 0.01 \leq \sigma_1 \leq 0.99\}$

When $\sigma_0 < 0.99$, we need a lower bound for $\overline{\lambda}\tau$. We first note that $g_i(\overline{\lambda}) \leq 3 - i + \overline{\lambda}$ (Lemma 6.6, Section 6.8) implies

$$4\sigma_0 + 2\sigma_1 = \sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1 - \sigma_0 - \sigma_1) g_2(\overline{\lambda}) \leq 2\sigma_0 + \sigma_1 + 1 + \overline{\lambda}$$

so

$$\overline{\lambda} \geq 2\sigma_0 + \sigma_1 - 1 = 1 - \tau \log(1 + 1/\tau).$$

Here we have used (6.41).

For τ, note that $\sigma_0 < 0.99$ and $\sigma_0 + \sigma_1 \leq 1$ implies $\tau \log(1 + 1/\tau) = 2 - 2\sigma_0 - \sigma_1 \geq 1 - \sigma_0 > 0.01$. The function $\tau \log(1 + 1/\tau)$ is increasing in τ by the discussion after (6.41). This implies

$$\tau > 10^{-3}, \tag{6.47}$$

since $0.001 \log(1001) < 0.01$.

If $\tau \leq 1.1$,

$$\overline{\lambda} \geq 1 - 1.1 \log 2 > 0.1.$$

So, if $\tau \leq 1.1$,

$$\overline{\lambda}\tau \geq 10^{-4}.$$

If $1.1 < \tau$ then we use $\log(1 + x) \leq x - x^2/2 + x^3/3$ for $|x| \leq 1$ to write

$$\overline{\lambda}\tau \geq \tau - \tau^2 \log(1 + 1/\tau) \geq \frac{1}{2} - \frac{1}{3\tau} \geq \frac{1}{6}.$$

So, in E_1, we have

$$\overline{\lambda}\tau \geq 10^{-4}. \tag{6.48}$$

By definition of E_1, $\sigma_0 \geq 0.01$ and $\sigma_1 \geq 0.01$. By (6.39), $0 \leq \overline{\lambda} \leq \lambda$. This implies $f_3(\overline{\lambda})/\overline{\lambda}^2 f_1(\overline{\lambda}) \leq 1/6$ and $f_2(\overline{\lambda})/\overline{\lambda} f_1(\overline{\lambda}) \leq 1/3$ (Lemma 6.8, Section 6.8). So after rewriting (6.43) slightly,

$$\begin{aligned}
\partial_0 \log f(\sigma_0, \sigma_1) &= \log\left(\frac{1 - \sigma_0 - \sigma_1}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_1(\overline{\lambda})}\right) - 2\log \overline{\lambda}\tau + \log 2 + 2\log(4\sigma_0 + 2\sigma_1) \\
&\leq \log\frac{1}{0.01} + \log\frac{1}{6} - 2\log 10^{-4} + \log 2 + 2\log 4 \\
&\leq 25.
\end{aligned}$$

Similarly, (6.44) is bounded by

$$\partial_1 \log f(\sigma_0, \sigma_1) \leq \log\frac{1}{0.01} + \log\frac{1}{3} - \log 10^{-4} + \log 2 + \log 4 \leq 15.$$

We now show numerically that $\log f \leq 0$ in E_1.

Numerics of Case One:
Since $\partial_i \log f$ is only bounded from above, $i = 0, 1$, this requires some care at the lower bounds of σ_0, σ_1, given by $\sigma_0 \geq (1 - \sigma_1)/2$ and $\sigma_1 \geq 0.01$. Note that if $\sigma_0 = (1 - \sigma_1)/2$, then $(\sigma_0, \sigma_1) \in E_0$ and it was shown above that $\log f(\sigma_0, \sigma_1) \leq \log 0.95 \leq -0.01$. Define a finite grid $P \subseteq E_1$ such that for any $(\sigma_0, \sigma_1) \in E_1$, there exists $(\overline{\sigma}_0, \overline{\sigma}_1) \in P \cup E_0$ where $0 \leq \sigma_0 - \overline{\sigma}_0 \leq \delta$ and $0 \leq \sigma_1 - \overline{\sigma}_1 \leq \delta$. Here

$\delta = 1/4000$. Numerical calculations will show that $\log f(\overline{\sigma}_0, \overline{\sigma}_1) \le -0.01$ for all $(\overline{\sigma}_0, \overline{\sigma}_1) \in P$. This implies that for all $\sigma_0, \sigma_1 \in E_1$,

$$\log f(\sigma_0, \sigma_1) \le \max_{\overline{\sigma}_0, \overline{\sigma}_1 \in P \cup E_0} \log f(\overline{\sigma}_0, \overline{\sigma}_1) + 25\delta + 15\delta \le -0.01 + 40\delta \le 0.$$

When calculating $\log f(\overline{\sigma}_0, \overline{\sigma}_1)$, approximations $\overline{\lambda}_{num}, \tau_{num}$ of $\overline{\lambda}(\overline{\sigma}_0, \overline{\sigma}_1), \tau(\overline{\sigma}_0, \overline{\sigma}_1)$ must be calculated with sufficient precision. By definition of $\overline{\lambda}$, $\partial \log f / \partial \overline{\lambda} = 0$, while

$$\begin{aligned}
&\left| \frac{\partial^2 \log f}{\partial \overline{\lambda}^2} \right| \\
&= \left| \sigma_0 \left(\frac{f_1(\overline{\lambda})}{f_3(\overline{\lambda})} - \frac{f_2(\overline{\lambda})^2}{f_3(\overline{\lambda})^2} \right) + \sigma_1 \left(\frac{f_0(\overline{\lambda})}{f_2(\overline{\lambda})} - \frac{f_1(\overline{\lambda})^2}{f_2(\overline{\lambda})^2} \right) \right. \\
&\qquad\qquad \left. + (1 - \sigma_0 - \sigma_1) \left(\frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{f_0(\overline{\lambda})^2}{f_1(\overline{\lambda})^2} \right) + \frac{4\sigma_0 + 2\sigma_1}{\overline{\lambda}^2} \right| \\
&= \frac{1}{\overline{\lambda}^2} \left| \sigma_0 \left(\overline{\lambda}^2 \frac{f_1(\overline{\lambda})}{f_3(\overline{\lambda})} - \frac{\overline{\lambda}^2 f_2(\overline{\lambda})^2}{f_3(\overline{\lambda})^2} \right) + \sigma_1 \left(\overline{\lambda}^2 \frac{f_0(\overline{\lambda})}{f_2(\overline{\lambda})} - \frac{\overline{\lambda}^2 f_1(\overline{\lambda})^2}{f_2(\overline{\lambda})^2} \right) \right. \\
&\left. + (1 - \sigma_0 - \sigma_1) \left(\overline{\lambda}^2 \frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{\overline{\lambda}^2 f_0(\overline{\lambda})^2}{f_1(\overline{\lambda})^2} \right) + \sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1 - \sigma_0 - \sigma_1) g_2(\overline{\lambda}) \right| \\
&= \frac{1}{\overline{\lambda}^2} | \sigma_0 g_0(\overline{\lambda})(g_1(\overline{\lambda}) - g_0(\overline{\lambda}) + 1) + \sigma_1 g_1(\overline{\lambda})(g_2(\overline{\lambda}) - g_1(\overline{\lambda}) + 1) \\
&\qquad\qquad\qquad\qquad + (1 - \sigma_0 - \sigma_1) g_2(\overline{\lambda})(\overline{\lambda} - g_2(\overline{\lambda}) + 1)| \\
&\le \frac{9}{\overline{\lambda}^2} |\sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1 - \sigma_0 - \sigma_1) g_2(\overline{\lambda})| \\
&= \frac{9}{\overline{\lambda}^2} |4\sigma_0 + 2\sigma - 1|, \qquad \text{by (6.37)} \\
&\le \frac{36}{\overline{\lambda}^2}.
\end{aligned}$$

Here we use the fact that $g_i(\overline{\lambda}) \le 4$ for $0 \le \overline{\lambda} \le \lambda$, $i = 0, 1, 2$ to conclude that $|g_1 - g_0 + 1|, |g_2 - g_1 + 1|, |\overline{\lambda} - g_2 + 1| \le 9$, and the final step uses $4\sigma_0 + 2\sigma_1 \le 4$. So the error contributed by $\overline{\lambda}_{num}$ is

$$|\log f(\overline{\sigma}_0, \overline{\sigma}_1; \overline{\lambda}_{num}) - \log f(\overline{\sigma}_0, \overline{\sigma}_1; \overline{\lambda})| \le (\overline{\lambda}_{num} - \overline{\lambda})^2 \frac{36}{\overline{\lambda}^2}$$

and to achieve a numerical error of at most 10^{-4}, we require that $|\overline{\lambda}_{num}/\overline{\lambda} - 1| \le 10^{-2}/6$.

Similarly by definition of τ, $\partial \log f / \partial \tau = 0$, while

$$\left| \frac{\partial^2 \log f}{\partial \tau^2} \right| = \left| \frac{1}{\tau(\tau + 1)} \right| \le 10^3, \text{ by (6.47).}$$

Thus to achieve a numerical error of at most 10^{-4}, it suffices to have $|\tau_{num}/\tau - 1| \le 10^{-2}$.

With the above precision, it is found that over all $(\overline{\sigma}_0, \overline{\sigma}_1) \in P \cup E_0$, $\log f(\overline{\sigma}_0, \overline{\sigma}_1) \le -0.0105$ numerically. With an error tolerance of 10^{-4}, this shows that $\log f(\overline{\sigma}_0, \overline{\sigma}_1) \le -0.01$.

Case Two. $E_2 = \{(\sigma_0, \sigma_1) \in E : 0 \le \sigma_1 < 0.01\}$

We divide E_2 into three subregions,

$$E_{2,1} = \{(\sigma_0, \sigma_1) \in E_2 : \sigma_1 = 0\},$$
$$E_{2,2} = \{(\sigma_0, \sigma_1) \in E_2 : \sigma_0 + \sigma_1 = 1\},$$
$$E_{2,3} = E_2 \setminus (E_{2,1} \cup E_{2,2}).$$

We begin by considering the point $(\sigma_0, \sigma_1) = (1, 0)$. Here $4\sigma_0 + 2\sigma_1 = 4$, and from (6.37) $\overline{\lambda}$ is defined by $g_0(\overline{\lambda}) = 4$. So $\overline{\lambda} = g_0^{-1}(4) = \lambda$. We also have $2 - 2\sigma_0 - \sigma_1 = 0$, and from the definition (6.41) of τ we have $\tau = 0$. Plugging this into the definition of f (6.42) gives $f(1, 0) = 1$.

Sub-Case 2.1a:
Now consider $E_{2,1}$, where $\sigma_1 = 0$. Here $\sigma_0 \geq 1/2$, from the definition of E and

$$\partial_0 \log f(\sigma_0, 0) = \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_1(\overline{\lambda})}\right) - 2\log\overline{\lambda}\tau + \log 2 + 2\log(4\sigma_0)$$

Within $E_{2,1}$, we consider two cases. First suppose $\sigma_0 \leq 0.99$. As noted in (6.48), $\sigma_0 \leq 0.99$ implies $\overline{\lambda}\tau \geq 10^{-4}$. Applying the same bounds as in (6.49),

$$\partial_0 \log f(\sigma_0, 0) \leq \log\frac{1}{6} - 2\log 10^{-4} + \log 2 + 2\log 4 \leq 21$$

and we show numerically that $f \leq 1$. The umerical calculations for this case now follow the same outline as above. The precision requirements given there will suffice in this case.

Sub-Case 2.1b:
Now suppose $\sigma_0 \geq 0.99$, still assuming $\sigma_1 = 0$. Here $\overline{\lambda} \leq \lambda$ (see (6.39)) implies $f_3(\overline{\lambda})/\overline{\lambda}^4 f_1(\overline{\lambda}) \geq 0.01$ by Lemma 6.8, Section 6.8. We have $\tau\log(1 + 1/\tau) = 2 - 2\sigma_0 - \sigma_1 = 2 - 2\sigma_0 \leq 0.02$ and since $\tau\log(1+1/\tau)$ is increasing (see (6.41)), it follows from a numerical calculation that $\tau \leq 0.004$. This implies

$$\frac{1-\sigma_0}{\tau^2} = \frac{\log\left(1 + \frac{1}{\tau}\right)}{2\tau} \geq 125\log 250$$

and

$$\begin{aligned}
\partial_0 \log f(\sigma_0, 0) &= \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^4 f_1(\overline{\lambda})}\right) - 2\log\tau + \log 2 + 4\log(4\sigma_0) \\
&= \log\left(\frac{1-\sigma_0}{\tau^2}\right) - \log\sigma_0 + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^4 f_1(\overline{\lambda})}\right) + \log 2 + 4\log(4\sigma_0) \\
&\geq \log(125\log 250) + \log 0.01 + \log 2 + 2\log 3.96 > 0
\end{aligned}$$

which implies $f(\sigma_0, 0) < f(1, 0) = 1$ for $\sigma_0 \geq 0.99$.

Sub-Case 2.2:
Now consider $E_{2,2}$, i.e. suppose $\sigma_0 + \sigma_1 = 1$ and $\sigma_1 < 0.01$. Then

$$\partial_0 \log f(\sigma_0, 1-\sigma_0) = \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})}\right) - \log\tau + \log(2 + 2\sigma_0) \tag{6.50}$$

By Lemma 6.8, Section 6.8, $\overline{\lambda} \leq \lambda$ implies

$$\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})} > 0.09.$$

As $\sigma_1 = 1 - \sigma_0$, τ is defined by $\tau \log(1 + 1/\tau) = 2 - 2\sigma_0 - \sigma_1 = \sigma_1$. So $\tau \log(1 + 1/\tau) \le 0.01$, implying $\tau \le 0.003$ since $\tau \log(1 + 1/\tau)$ is increasing, and

$$\frac{1-\sigma_0}{\tau} = \frac{\sigma_1}{\tau} = \log\left(1 + \frac{1}{\tau}\right) > \log 333.$$

So,

$$\begin{aligned} \partial_0 \log f(\sigma_0, 1-\sigma_0) &= \log\left(\frac{1-\sigma_0}{\tau}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})}\right) - \log\sigma_0 + \log(2 + 2\sigma_0) \\ &\ge \log\log 333 + \log 0.09 + \log 3.98 \\ &> 0 \end{aligned}$$

and for all $0.99 \le \sigma_0 < 1$, $f(\sigma_0, 1-\sigma_0) < f(1,0) = 1$.

Sub-Case 2.3:
Now consider $E_{2,3}$, i.e. suppose $0 < \sigma_1 < 1-\sigma_0$ and $\sigma_1 < 0.01$. We show that the gradient $\nabla \log f \ne 0$. Assume $(\partial_0 - 2\partial_1)\log f = 0$. By (6.46) we must have $\sigma_0 \ge (1-\sigma_1)/2 + \sqrt{1 - 2\sigma_1 - \sigma_1^2}/2$. Since $\sigma_1 \le 0.01$, we can replace this by the weaker bound $\sigma_0 \ge 1 - 1.1\sigma_1$. We trivially have $1 - \sigma_0 \ge (2 - 2\sigma_0 - \sigma_1)/2$, so

$$\frac{\sigma_1}{\tau} \ge \frac{1}{1.1}\frac{1-\sigma_0}{\tau} \ge \frac{1}{2.2}\frac{2-2\sigma_0-\sigma_1}{\tau} = \frac{1}{2.2}\log\left(1+\frac{1}{\tau}\right)$$

Since $\tau\log(1+1/\tau) = 2 - 2\sigma_0 - \sigma_1 \le 1.2\sigma_1 \le 0.012$, we have $\tau < 0.002$. So $\sigma_1/\tau \ge \log(500)/2.2$.

This allows us to show that if $(\partial_0 - 2\partial_1)\log f = 0$ and $\sigma_1 \le 0.01$, then $(\partial_0 - \partial_1)\log f \ne 0$. Noting that $4\sigma_0 + 2\sigma_1 \ge 4(1 - 1.1\sigma_1) + 2\sigma_1 \ge 3.976$,

$$\begin{aligned} (\partial_0 - \partial_1)\log f &= \log\left(\frac{\sigma_1}{\tau}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})}\right) - \log\sigma_0 + \log(4\sigma_0 + 2\sigma_1) \\ &\ge \log(\log(500)/2.2) + \log 0.09 + \log 3.976 \\ &= 1.038445... - 2.407945... + 1.380276... \\ &> 0 \end{aligned}$$

This shows that $\nabla \log f \ne 0$ in $E_{2,3}$. The boundary of $E_{2,3}$ is contained in $E_0 \cup E_{2,1} \cup E_{2,2} \cup E_1$. Since $f \le 1$ on the boundary of $E_{2,3}$ and $\nabla \log f \ne 0$ in $E_{2,3}$, it follows that $f \le 1$ in $E_{2,3}$.

Case Three: $E_3 = \{(\sigma_0, \sigma_1) \in E : 0.99 < \sigma_1 \le 1\}$.

Further divide E_3 into

$$\begin{aligned} E_{3,1} &= \{(\sigma_0,\sigma_1) \in E_3 : \sigma_0 + \sigma_1 = 1\}, \\ E_{3,2} &= E_3 \setminus E_{3,1}. \end{aligned}$$

Sub-Case 3.1:
Consider $E_{3,1}$, i.e. suppose $\sigma_0 + \sigma_1 = 1$ and $\sigma_0 < 0.01$. Then we write, see (6.50),

$$\partial_0 \log f(\sigma_0, 1-\sigma_0) = \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{1}{g_0(\overline{\lambda})}\right) - \log\overline{\lambda}\tau + \log(2 + 2\sigma_0)$$

To show that this is positive, we bound $\overline{\lambda}\tau$ from above. From (6.53) (Section 6.7) with $\Delta = 4\sigma_0+2\sigma_1$ we have $\tau \leq 1/(4\sigma_0+2\sigma_1-2)$. For $\overline{\lambda}$, we use the bound derived in Section 6.7 (6.54). Note that if $\Delta = 4\sigma_0+2\sigma_1$ then $L_2 = \overline{\lambda}$ in (6.54). So,

$$\overline{\lambda} \leq \frac{12(4\sigma_0+2\sigma_1-2\sigma_0-\sigma_1-1)}{6-3\sigma_0-2\sigma_1} \leq 12(2\sigma_0+\sigma_1-1) \leq 12.$$

These two bounds together imply $\overline{\lambda}\tau \leq 6$. For all $0 \leq \overline{\lambda} \leq \lambda$ we have $3 \leq g_0(\overline{\lambda}) \leq 4$ since $3 \leq \Delta_0/\sigma_0 \leq 4$ (see the discussion before (6.39)).

We conclude that

$$\partial_0 \log f(\sigma_0, 1-\sigma_0) \geq \log\frac{0.99}{0.01} + \log\frac{1}{4} - \log 6 + \log 2 > 0$$

This implies that for all $(\sigma_0,\sigma_1) \in E_{3,1}$, $f(\sigma_0,\sigma_1) \leq f(0.01, 0.99) \leq 1$, since $(0.01, 0.99) \in E_1$.

Sub-Case 3.2:
Now consider $E_{3,2}$. As noted in (6.46), any stationary point of $\log f$ must satisfy $\sigma_1 < 1/2$, so $E_{3,2}$ contains no stationary point. The boundary of $E_{3,2}$ is contained in $E_0 \cup E_1 \cup E_{3,1}$, and it has been shown that $f \leq 1$ in each of $E_0, E_1, E_{3,1}$. It follows that $f \leq 1$ in $E_{3,2}$.

This completes the proof of Lemma 6.4 and Theorem 6.1.

6.7 Appendix A

This section is concerned with showing that the system of equations (6.35) under certain conditions has no solution. Throughout the section, assume τ satisfies (6.33): Recall that $\Delta_3 = 4\tau+4\sigma_0+2\sigma_1-\Delta$,

$$\tau\left(\log\left(1+\frac{1}{\tau}\right) - 2\log\left(\frac{4\tau+4\sigma_0+2\sigma_1-\Delta}{4\tau}\right) - \log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right)\right) = 2-2\sigma_0-\sigma_1. \tag{6.51}$$

Here $\lambda = g_0^{-1}(4) \approx 2.688$ is fixed.

Define for $2\sigma_0+\sigma_1+1 \leq \Delta \leq 4\sigma_0+2\sigma_1$

$$L_1(\sigma_0,\sigma_1,\Delta,\tau) = \lambda_3\sqrt{\frac{\Delta}{4\tau+4\sigma_0+2\sigma_1-\Delta}}$$

and define $L_2(\sigma_0,\sigma_1,\Delta)$ as the unique solution to $G(\sigma_0,\sigma_1,L_2(\sigma_0,\sigma_1,\Delta)) = \Delta$, where G is defined by

$$G(\sigma_0,\sigma_1,x) = \sigma_0 g_0(x) + \sigma_1 g_1(x) + (1-\sigma_0-\sigma_1)g_2(x). \tag{6.52}$$

This is well defined because each g_i is strictly increasing, and for fixed σ_0,σ_1 we have $G(\sigma_0,\sigma_1,0) = 2\sigma_0+\sigma_1+1 \leq \Delta$ and $\lim_{x\to\infty} G(\sigma_0,\sigma_1,x) = \infty$ (see Section 6.8). Define

$$R = \left\{(\sigma_0,\sigma_1,\Delta,\tau) \in \mathbb{R}_+^4 : \sigma_0+\sigma_1 \leq 1; 2\sigma_0+\sigma_1 \geq 1;\ 2\sigma_0+\sigma_1+1 \leq \Delta \leq 4\sigma_0+2\sigma_1;\ (6.51) \text{ holds.}\right\}$$

We prove that the system (6.35) is inconsistent by proving

Lemma 6.5. *Let $(\sigma_0,\sigma_1,\Delta,\tau) \in R$. Then $L_1(\sigma_0,\sigma_1,\Delta,\tau) > L_2(\sigma_0,\sigma_1,\Delta)$*

Proof. Define $L(\sigma_0, \sigma_1, \Delta, \tau) = L_1(\sigma_0, \sigma_1, \Delta, \tau) - L_2(\sigma_0, \sigma_1, \Delta)$. We will bound $|\nabla L|$ in R in order to show numerically that $L \geq 0$. However, ∇L is unbounded for Δ close to 4 and $2\sigma_0 + \sigma_1$ close to 1. For this reason, define

$$\begin{aligned} R_1 &= \{(\sigma_0, \sigma_1, \Delta, \tau) \in R : \Delta \geq 3.6\}, \\ R_2 &= \{(\sigma_0, \sigma_1, \Delta, \tau) \in R : 2\sigma_0 + \sigma_1 \leq 1.1\}, \\ R_3 &= R \setminus (R_1 \cup R_2). \end{aligned}$$

Analytical proofs will be provided for R_1, R_2, and a numerical calculation will have to suffice for R_3.

First note that for any σ_0, σ_1 we have $L_2(\sigma_0, \sigma_1, 2\sigma_0 + \sigma_1 + 1) = 0$, since $G(\sigma_0, \sigma_1, 0) = 2\sigma_0 + \sigma_1 + 1$, see (6.52). Here we use the fact that $g_i(0) = 3 - i$, $i = 0, 1, 2$ by Lemma 6.6, Section 6.8. This implies that $L_1(\sigma_0, \sigma_1, 2\sigma_0 + \sigma_1 + 1, \tau) \geq 0 = L_2(\sigma_0, \sigma_1, 2\sigma_0 + \sigma_1 + 1)$, and we may therefore assume $\Delta > 2\sigma_0 + \sigma_1 + 1$.

We proceed by finding an upper bound for τ, given that it satisfies (6.51). Fix $\sigma_0, \sigma_1, \Delta$ and define

$$r(\zeta) = \zeta \left(\log \left(1 + \frac{1}{\zeta} \right) - 2 \log \left(\frac{4\tau + 4\sigma_0 + 2\sigma_1 - \Delta}{4\tau} \right) - \log \left(\frac{\lambda^4}{\lambda_3^4} \frac{f_3(\lambda_3)}{f_3(\lambda)} \right) \right)$$

We first derive a lower bound $r_1(\zeta) \leq r(\zeta)$.

For $x \geq 0$ we have $x - x^2/2 \leq \log(1 + x) \leq x$. This implies, that for all ζ,

$$2\zeta \log \left(1 + \frac{4\sigma_0 + 2\sigma_1 - \Delta}{4\zeta} \right) \leq 2\zeta \frac{4\sigma_0 + 2\sigma_1 - \Delta}{4\zeta} = \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2}$$

Let $h(x) = \log f_3(x) - 4 \log x$. Then $h'(x) = f_2(x)/f_3(x) - 4/x$, and we note that $h'(\lambda) = 0$, by definition of λ. The second derivative is $h''(x) = f_1(x)/f_3(x) - f_2(x)^2/f_3(x)^2 + 4/x^2$. Substituting $f_1(x) = f_3(x) + x + x^2/2$ and $f_2(x) = f_3(x) + x^2/2$, for all $x \geq \lambda$

$$\begin{aligned} h''(x) &= \frac{4}{x^2} + 1 + \frac{x + x^2/2}{f_3(x)} - 1 - \frac{x^2}{f_3(x)} - \frac{x^4}{4 f_3(x)^2} \\ &= \frac{4}{x^2} - \frac{x^2 - 2x}{2 f_3(x)} - \frac{x^4}{4 f_3(x)^2} \end{aligned}$$

Since $x \geq \lambda > 2$ we have $x^2 - 2x > 0$, and $f_3(x) < e^x$ implies

$$\begin{aligned} h''(x) &= \frac{4}{x^2} - \frac{x^2 - 2x}{2 f_3(x)} - \frac{x^4}{4 f_3(x)^2} \\ &\leq \frac{4}{x^2} - \frac{x^2 - 2x}{2e^x} \\ &\leq \frac{4}{x^2} + \frac{2x}{2e^x} \\ &\leq \frac{4}{x^2} + x^{1-\lambda} \end{aligned}$$

Here we use the fact that $e^x \geq x^\lambda$ for $x \geq \lambda$, since $\lambda < e$. Since $4x^{-2} + x^{1-\lambda}$ is decreasing, we have $h''(x) \leq 4\lambda^{-2} + \lambda^{1-\lambda} < 3/4$ for all $x \geq \lambda$.

By Taylor's theorem, for some $x \in [\lambda, \lambda_3]$

$$\begin{aligned}
\log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right) &= h(\lambda_3) - h(\lambda) \\
&= h(\lambda) + h'(\lambda)(\lambda_3 - \lambda) + \frac{1}{2}h''(x)(\lambda_3 - \lambda)^2 - h(\lambda) \\
&\leq \frac{3}{8}(\lambda_3 - \lambda)^2
\end{aligned}$$

Another application of Taylor's theorem lets us bound

$$\lambda_3 - \lambda = g_0^{-1}\left(4 + \frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right) - g_0^{-1}(4).$$

By Lemma 6.7, Section 6.8, we have $g_0'(x) \geq g_0'(\lambda) \geq 1/2$ for $x \geq \lambda$, so $dg_0^{-1}(y)/dy \leq 2$ for $y \geq 4$, and for some $y \geq 4$

$$\lambda_3 = \lambda + \frac{dg_0^{-1}(y)}{dy}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right) \leq \lambda + 2\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}$$

and so

$$\log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right) \leq \frac{3}{8}(\lambda_3 - \lambda)^2 \leq \frac{3}{2}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right)^2$$

Define τ_1 as the unique solution ζ to

$$2 - 2\sigma_0 - \sigma_1 = r_1(\zeta)$$

where

$$r_1(\zeta) = \zeta\left(\log\left(1 + \frac{1}{\zeta}\right) - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2\zeta} - \frac{3}{2}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\zeta}\right)^2\right).$$

Then $r_1(\zeta) \leq r(\zeta)$, and $r_1(\zeta)$ is strictly increasing. So, since $r_1(\tau_1) = r(\tau) = 2 - 2\sigma_0 - \sigma_1$, it follows that $\tau \leq \tau_1$.

Case of R_1:
Now fix $(\sigma_0, \sigma_1, \Delta, \tau) \in R_1$, i.e. suppose $\Delta \geq 3.6$. Then

$$\begin{aligned}
r_1\left(\frac{3}{4}\right) &= \frac{3}{4}\log\left(1 + \frac{4}{3}\right) - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2} - 2(4\sigma_0 + 2\sigma_1 - \Delta)^2 \\
&= \frac{3}{4}\log\frac{7}{3} - 2\sigma_0 - \sigma_1 + \frac{\Delta}{2} - 2(4\sigma_0 + 2\sigma_1 - \Delta)^2 \\
&\geq \frac{3}{4}\log\frac{7}{3} - 2\sigma_0 - \sigma_1 + \frac{3.6}{2} - 2(4 - 3.6)^2 \\
&> 2 - 2\sigma_0 - \sigma_1
\end{aligned}$$

We have $\lim_{\zeta\to 0} r_1(\zeta) \leq 0$, and r_1 is continous and increasing, so $\tau \leq \tau_1 < 3/4$. Since $\Delta \geq 3.6$ and $2\sigma_0 + \sigma_1 \leq 2$,

$$\Delta - (4\tau + 4\sigma_0 + 2\sigma_1 - \Delta) \geq 2\Delta - 3 - 4\sigma_0 - 2\sigma_1 \geq 7.2 - 7 > 0$$

This implies that

$$L_1(\sigma_0, \sigma_1, \Delta) = \lambda_3\sqrt{\frac{\Delta}{4\tau + 4\sigma_0 + 2\sigma_1 - \Delta}} > \lambda_3$$

Note that

$$G(\sigma_0, \sigma_1, \lambda) \geq G(\sigma_0, \sigma_1, \overline{\lambda}) = 4\sigma_0 + 2\sigma_1 \geq \Delta$$

implies that

$$L_2(\sigma_0, \sigma_1, \Delta) \leq \lambda = g_0^{-1}(4).$$

Also note that by (6.25) and (6.31) we have

$$\lambda_3 = g_0^{-1}\left(\frac{\Delta_3}{\tau}\right) = g_0^{-1}\left(4 + \frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right) \geq g_0^{-1}(4) = \lambda,$$

since g_0^{-1} is increasing (Lemma 6.7, Section 6.8). So

$$L_1(\sigma_0, \sigma_1, \Delta, \tau) > \lambda_3 \geq \lambda \geq L_2(\sigma_0, \sigma_1, \Delta)$$

for $(\sigma_0, \sigma_1, \Delta, \tau) \in R_1$.

Case of R_2, R_3:
For R_2, R_3 we will need a new bound on τ. Since $x - x^2/2 \leq \log(1+x)$ for all $x \geq 0$,

$$r_1(\zeta) \geq r_2(\zeta) = \zeta\left(\frac{1}{\zeta} - \frac{1}{2\zeta^2} - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2\zeta} - \frac{3}{2}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\zeta}\right)^2\right).$$

Let τ_2 be defined by $r_2(\tau_2) = 2 - 2\sigma_0 - \sigma_1$, which can be solved for τ_2;

$$\tau_2 = \frac{1 + 3(4\sigma_0 + 2\sigma_1 - \Delta)^2}{\Delta - 2}.$$

It follows from $r(\tau) \geq r_2(\tau)$ and the fact that r_2 is increasing that

$$\tau \leq \frac{1 + 3(4\sigma_0 + 2\sigma_1 - \Delta)^2}{\Delta - 2}. \tag{6.53}$$

An upper bound for $L_2(\sigma_0, \sigma_1, \Delta)$ will follow from bounding the partial derivative of $G(\sigma_0, \sigma_1, x)$ with respect to x. We have $g_0' \geq 1/4$, $g_1' \geq 1/3$ and $g_2' \geq 1/2$ by Lemma 6.7 (Section 6.8), so

$$\begin{aligned}
\frac{\partial}{\partial x} G(\sigma_0, \sigma_1, x) &= \sigma_0 g_0'(x) + \sigma_1 g_1'(x) + (1 - \sigma_0 - \sigma_1) g_2'(x) \\
&\geq \frac{\sigma_0}{4} + \frac{\sigma_1}{3} + \frac{1 - \sigma_0 - \sigma_1}{2} \\
&= \frac{6 - 3\sigma_0 - 2\sigma_1}{12}
\end{aligned}$$

and $G(\sigma_0, \sigma_1, 0) = 2\sigma_0 + \sigma_1 + 1$ implies

$$\begin{aligned}
\Delta &= G(\sigma_0, \sigma_1, L_2(\Delta)) \\
&\geq G(\sigma_0, \sigma_1, 0) + \min_x \frac{\partial}{\partial x} G(\sigma_0, \sigma_1, x) L_2(\Delta) \\
&\geq 2\sigma_0 + \sigma_1 + 1 + \frac{6 - 3\sigma_0 - 2\sigma_1}{12} L_2(\Delta)
\end{aligned}$$

So

$$L_2(\Delta) \leq \frac{12(\Delta - 2\sigma_0 - \sigma_1 - 1)}{6 - 3\sigma_0 - 2\sigma_1}. \tag{6.54}$$

So, to show $L_1(\sigma_0,\sigma_1,\Delta,\tau) \geq L_2(\sigma_0,\sigma_1,\Delta)$, it is enough to show that

$$\lambda_3\sqrt{\frac{\Delta}{4\tau+4\sigma_0+2\sigma_1-\Delta}} > \frac{12(\Delta-2\sigma_0-\sigma_1-1)}{6-3\sigma_0-2\sigma_1}$$

Solving for τ, this is equivalent to showing

$$\tau < \Delta\left[\frac{\lambda_3(6-3\sigma_0-2\sigma_1)}{24(\Delta-2\sigma_0-\sigma_1-1)}\right]^2 - \frac{4\sigma_0+2\sigma_1-\Delta}{4}$$

and by (6.53), and $\lambda_3 \geq \lambda$, it is enough to show

$$\frac{1+3(4\sigma_0+2\sigma_1-\Delta)^2}{\Delta-2} < \Delta\left[\frac{\lambda(6-3\sigma_0-2\sigma_1)}{24(\Delta-2\sigma_0-\sigma_1-1)}\right]^2 - \frac{(4\sigma_0+2\sigma_1-\Delta)}{4} \tag{6.55}$$

for $(\sigma_0,\sigma_1,\Delta,\tau) \in R_2 \cup R_3$.

Case of R_2:
Consider R_2, i.e. suppose $2\sigma_0+\sigma_1 \leq 1.1$. Then $4\sigma_0+2\sigma_1-\Delta \leq 2\sigma_0+\sigma_1-1 \leq 0.1$ since $\Delta \geq 2\sigma_0+\sigma_1+1$. This implies

$$\frac{1+3(4\sigma_0+2\sigma_1-\Delta)^2}{\Delta-2} \leq \frac{1.03}{\Delta-2}$$

Furthermore, $6-3\sigma_0-2\sigma_1 \geq 4.9-\sigma_0-\sigma_1 \geq 3.9$, while $2\sigma_0+\sigma_1 \geq 1$ implies $\Delta-2\sigma_0-\sigma_1-1 \leq \Delta-2$. We have $\lambda_3 \geq \lambda = g_0^{-1}(4) > 2.5$. So it holds that

$$\Delta\left[\frac{\lambda_3(6-3\sigma_0-2\sigma_1)}{24(\Delta-2\sigma_0-\sigma_1-1)}\right]^2 - \frac{(4\sigma_0+2\sigma_1-\Delta)}{4} > \Delta\left[\frac{2.5\times 3.9}{24(\Delta-2)}\right]^2 - 0.025$$

and it is enough to show that

$$\frac{1.03}{\Delta-2} \leq \Delta\left[\frac{2.5\times 3.9}{24(\Delta-2)}\right]^2 - 0.025$$

We have $\Delta \geq 2\sigma_0+\sigma_1+1 > 2$, so multipling both sides by $\Delta-2>0$, this amounts to solving a second-degree polynomial inequality. Numerically, the zeros of the resulting second-degree polynomial are $\Delta \approx -33$ and $\Delta \approx 2.37$. The inequality holds at $\Delta = 2.3$, and so it holds for all $2 < \Delta \leq 2.37$. In particular, it holds for $2\sigma_0+\sigma_1+1 < \Delta \leq 4\sigma_0+2\sigma_1$ when $1 \leq 2\sigma_0+\sigma_1 \leq 1.1$.

Case of R_3:
Lastly, consider R_3. Here more extensive numerical methods will be used, and we begin by reducing the analysis from three variables to two. Divide R_3 into four subregions,

$$\begin{aligned}
R_{3,1} &= \{(\sigma_0,\sigma_1,\Delta) \in R_3 : 1/2 \leq \sigma_1 \leq 1\},\\
R_{3,2} &= \{(\sigma_0,\sigma_1,\Delta) \in R_3 : 1/4 \leq \sigma_1 < 1/2\},\\
R_{3,3} &= \{(\sigma_0,\sigma_1,\Delta) \in R_3 : 1/8 \leq \sigma_1 < 1/4\},\\
R_{3,4} &= \{(\sigma_0,\sigma_1,\Delta) \in R_3 : 0 \leq \sigma_1 < 1/8\}.
\end{aligned}$$

Define

$$u_1 = 5.5, \quad u_2 = 5.75, \quad u_3 = 5.875, \quad u_4 = 5.9375.$$

Then

$$6-3\sigma_0-2\sigma_1=\left(6-\frac{\sigma_1}{2}\right)-3\sigma_0-\frac{3\sigma_1}{2}\geq u_i-\frac{3(2\sigma_0+\sigma_1)}{2}$$

in $R_{3,i}$, $i=1,2,3,4$.

Fixing i, (6.55) will hold in $R_{3,i}$ if we can show that

$$\frac{1+3(4\sigma_0+2\sigma_1-\Delta)^2}{\Delta-2}\leq\Delta\left[\frac{\lambda(u_i-3(2\sigma_0+\sigma_1)/2)}{24(\Delta-2\sigma_0-\sigma_1-1)}\right]^2-\frac{(4\sigma_0+2\sigma_1-\Delta)}{4} \tag{6.56}$$

Note that σ_0,σ_1 only appear as $\Sigma=2\sigma_0+\sigma_1$ in (6.56). For this reason we clear denominators in (6.56) and define for $i=1,2,3,4$,

$$\begin{aligned}\varphi_i(\Sigma,\Delta) &= \lambda^2\Delta(\Delta-2)(u_i-3\Sigma/2)^2-144(\Delta-2)(\Delta-\Sigma-1)^2(2\Sigma-\Delta)\\ &\quad -576(\Delta-\Sigma-1)^2-1728(\Delta-\Sigma-1)^2(2\Sigma-\Delta)^2.\end{aligned}$$

In which case, (6.56) is equivalent to $\varphi(\Sigma,\Delta)\geq 0$.

In $R_{3,1}$ we have $1.1\leq\Sigma\leq 1.5$ since $2\sigma_0+\sigma_1\geq 1.1$ is assumed, and $\sigma_1\geq 1/2$ and $\sigma_0+\sigma_1\leq 1$ imply $2\sigma_0+\sigma_1\leq 2-\sigma_1\leq 1.5$. For this reason define

$$\begin{aligned}\widetilde{R}_{3,1}&=\{(\Sigma,\Delta):1.1\leq\Sigma\leq 1.5,\Sigma+1\leq\Delta\leq 2\Sigma\}\\ \widetilde{R}_{3,2}&=\{(\Sigma,\Delta):1.5\leq\Sigma\leq 1.75,\Sigma+1\leq\Delta\leq 2\Sigma\},\\ \widetilde{R}_{3,3}&=\{(\Sigma,\Delta):1.75\leq\Sigma\leq 1.875,\Sigma+1\leq\Delta\leq\min\{2\Sigma,3.6\}\},\\ \widetilde{R}_{3,4}&=\{(\Sigma,\Delta):1.875\leq\Sigma\leq 2,\Sigma+1\leq\Delta\leq\min\{2\Sigma,3.6\}\}.\end{aligned}$$

Here $\Sigma+1\leq\Delta\leq 2\Sigma$ is (6.30).

Equation (6.56) will follow from showing that $\varphi_i(\Sigma,\Delta)\geq 0$ in $\widetilde{R}_{3,i}$, $i=1,2,3,4$.

The φ_i are degree four polynomials, and bounds on $|\nabla\varphi_i|$ are found by applying the triangle inequality to the partial derivatives of φ_i. The same bound will be applied to $\nabla\varphi_i$ for all i. using,

$$2\leq\Sigma+1\leq\Delta\leq 2\Sigma\leq 4,\quad u_i\leq 6,\quad \lambda<3$$

from which we obtain

$$\begin{aligned}&u_i-\frac{3\Sigma}{2}\leq\frac{9}{2},\ -1\leq 3\Sigma-2\Delta+1\leq 1, -2\leq 4\Sigma-3\Delta+2\leq 1,\\ &(\Delta-\Sigma-1)(2\Sigma-\Delta)\leq\frac{(\Sigma-1)^2}{4}\leq\frac{1}{4}.\end{aligned}$$

we have

$$\begin{aligned}\left|\frac{\partial\varphi_i}{\partial\Sigma}\right| &= |-3\lambda^2\Delta(\Delta-2)(u_i-3\Sigma/2)+288(\Delta-2)(\Delta-\Sigma-1)(2\Sigma-\Delta)\\ &\quad -288(\Delta-2)(\Delta-\Sigma-1)^2+1152(\Delta-\Sigma-1)\\ &\quad +3456(\Delta-\Sigma-1)(2\Sigma-\Delta)^2-6912(\Delta-\Sigma-1)^2(2\Sigma-\Delta)|\\ &\leq 3\lambda^2\Delta(\Delta-2)(u_i-3\Sigma/2)+288(\Delta-2)(\Delta-\Sigma-1)|3\Sigma-2\Delta+1|\\ &\quad +1152(\Delta-\Sigma-1)+3456(\Delta-\Sigma-1)(2\Sigma-\Delta)|4\Sigma-3\Delta+2|\\ &\leq 27\cdot 4\cdot 2\cdot 9/2+288\cdot 2\cdot 2\cdot 1+1152\cdot 2+3456\cdot 3/4\cdot 2\\ &= 9612\end{aligned}$$

For Δ,

$$\begin{aligned}
\left|\frac{\partial \varphi_i}{\partial \Delta}\right| &= |\lambda^2\Delta(u_i-3\Sigma/2)^2+\lambda^2(\Delta-2)(u_i-3\Sigma/2)^2-144(\Delta-\Sigma-1)^2(2\Sigma-\Delta) \\
&\quad -288(\Delta-2)(\Delta-\Sigma-1)(2\Sigma-\Delta)+144(\Delta-2)(\Delta-\Sigma-1)^2 \\
&\quad -1152(\Delta-\Sigma-1)-3456(\Delta-\Sigma-1)(2\Sigma-\Delta)^2+3456(\Delta-\Sigma-1)^2(2\Sigma-\Delta)| \\
&\le \lambda^2(2\Delta-2)(u_i-3\Sigma/2)^2+144(\Delta-\Sigma-1)^2(2\Sigma-\Delta)+288(\Delta-2)(\Delta-\Sigma-1)(2\Sigma-\Delta) \\
&\quad +144(\Delta-2)(\Delta-\Sigma-1)^2+1152(\Delta-\Sigma-1)+3456(\Delta-\Sigma-1)(2\Sigma-\Delta)|2\Delta-3\Sigma-1| \\
&\le 9\cdot 6\cdot(9/2)^2+144\cdot 2^2\cdot 1+288\cdot 2\cdot 2\cdot 1+144\cdot 2\cdot 2^2+1152\cdot 2+3456\cdot 3/4\cdot 1 \\
&= 8383.5
\end{aligned}$$

so $|\nabla\varphi_i| \le 12755$ for $i = 1, 2, 3, 4$.

For each i, a grid $\mathcal{P}_i \subseteq \widetilde{R}_{3,i}$ of $4\cdot 10^6$ points is generated such that for each $x \in \widetilde{R}_{3,i}$, there exists an $x_0 \in \mathcal{P}_i$ for which $|x - x_0| \le 0.001$. On this grid, φ_i is calculated numerically, and it is found that

$$\min_{x_0\in\mathcal{P}_i} \varphi_i(x_0) = \begin{cases} 22.49, & i=1 \\ 25.50, & i=2 \\ 27.08, & i=3 \\ 19.04, & i=4 \end{cases}$$

So for any i and any $x \in \widetilde{R}_{3,i}$, there exists an x_0 such that $|\varphi_i(x) - \varphi_i(x_0)| \le |\nabla\varphi_i||x - x_0| \le 12755 \cdot 0.001 < 13$, which implies $\varphi_i(x) > \varphi_i(x_0) - 13 > 0$. This proves (6.55) for $\sigma_0, \sigma_1, \Delta \in R_3$.

□

6.8 Appendix B

This section is concerned with the functions

$$f_0(x) = e^x \text{ and } f_k(x) = e^x - \sum_{j=0}^{k-1} \frac{x^j}{j!},\ x \ge 0,\ k = 1, 2, 3,$$

and the related functions

$$g_0(x) = \frac{xf_2(x)}{f_3(x)}, \quad g_1(x) = \frac{xf_1(x)}{f_2(x)}, \quad g_2(x) = \frac{xf_0(x)}{f_1(x)}.$$

Since $f_k(0) = 0$ for $k \ge 1$, we define $g_i(0) = \lim_{x\to 0} g_i(x) = 3 - i$. Note that

$$\frac{d}{dx} f_k(x) = f_{k-1}(x), \quad k \ge 1$$

Lemma 6.6. *For all $x \ge 0$ and $i = 0, 1, 2$,*

$$x < g_i(x) \le 3 - i + x$$

with equality in the upper bound if and only if $x = 0$.

Proof. Fix i. By definition, $g_i(0) = 3 - i$. For $x > 0$ consider

$$g_i(x) - x = \frac{xf_{2-i}(x)}{f_{3-i}(x)} - x = \frac{x(f_{2-i}(x) - f_{3-i}(x))}{f_{3-i}(x)} = \frac{x^{3-i}}{(2-i)!f_{3-i}(x)}.$$

Since $f_{3-i}(x) > 0$ we have $g_i(x) - x > 0$. Now

$$(3-i)(2-i)!f_{3-i}(x) - x^{3-i} = (3-i)! \sum_{k \geq 3-i} \frac{x^k}{k!} - x^{3-i} = (3-i)! \sum_{k \geq 4-i} \frac{x^k}{k!} > 0$$

for $x > 0$, implying $g_i(x) - x < 3 - i$. □

Lemma 6.7. *The functions g_0, g_1, g_2 are convex, and $g_i'(x) \geq 1/(4-i)$ for $x \geq 0$, $i = 0, 1, 2$.*

Proof. Consider g_0. Since $f_2(x) = f_3(x) + x^2/2$, g_0 can be written as

$$g_0(x) = \frac{xf_2(x)}{f_3(x)} = x + \frac{x^3}{2f_3(x)}$$

Let $q(x) = f_3(x)/x^3 = \sum_{j \geq 0} x^j/(j+3)!$. Then $g_0(x) = x + 1/2q(x)$, and

$$g_0'(x) = 1 - \frac{q'(x)}{2q(x)^2}, \quad g_0''(x) = \frac{2q'(x)^2 - q(x)q''(x)}{2q(x)^3}$$

and we show that $2q'(x)^2 - q(x)q''(x) \geq 0$. We have $q'(x) = \sum_{j \geq 0}(j+1)x^j/(j+4)!$ and $q''(x) = \sum_{j \geq 0}(j+1)(j+2)x^j/(j+5)!$, so the jth Taylor coefficient of $2q'(x)^2 - q(x)q''(x)$ is given by

$$\begin{aligned}
[x^j][2q'(x)^2 - q(x)q''(x)] &= \sum_{\substack{j_1, j_2 \geq 0 \\ j_1 + j_2 = j}} 2\frac{(j_1+1)}{(j_1+4)!}\frac{(j_2+1)}{(j_2+4)!} - \frac{1}{(j_1+3)!}\frac{(j_2+1)(j_2+2)}{(j_2+5)!} \\
&= \sum_{j_1, j_2} \frac{2(j_1+1)(j_2+1)(j_2+5) - (j_1+4)(j_2+1)(j_2+2)}{(j_1+4)!(j_2+5)!} \\
&= \sum_{j_1, j_2} \frac{(j_2+1)(2(j_1+1)(j_2+5) - (j_1+4)(j_2+2))}{(j_1+4)!(j_2+5)!} \\
&= \sum_{j_1, j_2} \frac{(j_2+1)(j_1 j_2 + 8j_1 - 2j_2 + 2)}{(j_1+4)!(j_2+5)!}
\end{aligned}$$

It is seen that this is positive for $j = 0, 1, 2$. Let $Q(j_1, j_2)$ denote the summand. If $j \geq 3$ then since $Q(j_1, j_2) \geq 0$ whenever $j_1 \geq 2$.

$$\begin{aligned}
\sum_{\substack{j_1, j_2 \geq 0 \\ j_1 + j_2 = j}} Q(j_1, j_2) &\geq Q\left(\left\lfloor \frac{j}{2} \right\rfloor, \left\lceil \frac{j}{2} \right\rceil\right) + Q(0, j) + Q(1, j-1) \\
&= \frac{(\lfloor j/2 \rfloor + 1)\left(\lceil j/2 \rceil \lfloor j/2 \rfloor + 8\lceil j/2 \rceil - 2\lfloor j/2 \rfloor + 2\right)}{(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{2(j^2-1)}{24(j+5)!} - \frac{j^2 - 11j}{120(j+4)!} \\
&\geq \frac{j^3}{8(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{j^2}{12(j+5)!} - \frac{j^2 - 11j}{120(j+4)!} \\
&= \frac{j^3}{8(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{10j^2 + (j^2 - 11j)(j+5)}{120(j+5)!} \\
&\geq \frac{j^3}{8}\left(\frac{1}{(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{1}{15(j+5)!}\right).
\end{aligned}$$

(To get the final inequality, consider $j \leq 11$ and $j > 11$ seperately).

It remains to show that $a_j = (\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!$ is smaller than $b_j = 15(j+5)!$ for $j \geq 3$. For $j = 3$, $a_3 = 6! \cdot 6! < 15 \cdot 8! = b_3$. For the induction step, $a_{j+1}/a_j \leq j/2 + 6$ while $b_{j+1}/b_j = j + 6$, so $a_3 < b_3$ implies $a_j < b_j$ for all $j \geq 3$. So $2q'(x)^2 - q(x)q''(x) \geq 0$, and it follows that g_0 is convex. Similar arguments show that g_1, g_2 are convex.

For $i = 0, 1$,

$$\begin{aligned} g_i'(x) &= \frac{f_{2-i}(x)}{f_{3-i}(x)} + \frac{xf_{1-i}(x)}{f_{3-i}(x)} - \frac{xf_{2-i}(x)^2}{f_{3-i}(x)^2} \\ &= \frac{f_{2-i}(x)f_{3-i}(x) + xf_{1-i}(x)f_{3-i}(x) - xf_{3-i}(x)^2}{f_{3-i}(x)^2}. \end{aligned}$$

Now

$$\begin{aligned} &f_{2-i}(x)f_{3-i}(x) + xf_{1-i}(x)f_{3-i}(x) - xf_{3-i}(x)^2 = \\ &x^{6-2i}\left(\frac{1}{(2-i)!(4-i)!} + \frac{1}{(3-i)!^2} + \frac{1}{(1-i)!(4-i)!} + \frac{1}{(2-i)!(3-i)!} - \frac{2}{(2-i)!(3-i)!} + O(x)\right) \\ &= x^{6-2i}\left(\frac{1}{(3-i)!(4-i)!} + O(x)\right). \end{aligned}$$

And

$$f_{3-i}(x)^2 = x^{6-2i}\left(\frac{1}{(3-i)!^2} + O(x)\right).$$

So, for $i = 0, 1$ we have

$$g_i'(x) = \frac{1}{4-i} + O(x).$$

For $i = 2$ we have

$$g_2'(x) = \frac{e^x}{f_1(x)} + \frac{xe^x}{f_1(x)} - \frac{xe^{2x}}{f_1(x)^2} = e^x\left(\frac{f_1(x)(1+x) - xe^x}{f_1(x)^2}\right) = e^x\left(\frac{\frac{x^2}{2} + O(x^3)}{x^2 + O(x^3)}\right) = \frac{1}{2} + O(x).$$

And by the convexity of g_i we have $g_i'(x) \geq 1/(4-i)$ for all $x \geq 0$. □

Lemma 6.7 allows us to define inverses g_i^{-1}, $i = 0, 1, 2$.

Lemma 6.8. *For $0 \leq x \leq \lambda = g_0^{-1}(4)$, the following inequalities hold.*

$$\begin{aligned} &(i) \quad 1 \leq \frac{f_2(x)^2}{f_1(x)f_3(x)} \leq 2 \\ &(ii) \quad 0.09 < \frac{f_3(x)}{x^2 f_1(x)} \leq \frac{1}{6} \\ &(iii) \quad \frac{f_2(x)}{xf_1(x)} \leq \frac{1}{3} \\ &(iv) \quad 0.01 < \frac{f_3(x)}{x^4 f_1(x)} \\ &(v) \quad 0.09 < \frac{f_3(x)}{x^2 f_2(x)} \end{aligned}$$

Proof. For the lower bound, let $x > 0$ and consider the equation $f_2(x)^2 = f_1(x)f_3(x)$. By definition of f_i, this equation can be written as

$$(e^x - 1 - x)^2 = (e^x - 1)\left(e^x - 1 - x - \frac{x^2}{2}\right)$$

Expanding and reordering terms, we have

$$e^x\left(x + \frac{x^2}{2}\right) = x + \frac{x^2}{2}$$

which clearly has no positive solution. Since $f_2(0)^2/f_1(0)f_3(0) = 3/2 > 1$, this implies that $f_2(x)^2/f_1(x)f_3(x) > 1$ for all $x \geq 0$.

For the upper bound we consider the equation $f_2(x)^2 = 2f_1(x)f_3(x)$. This simplifies to

$$(e^x - 1)^2 = x^2e^x \text{ or } e^x = 1 + xe^{x/2}$$

which has no positive solution.

Since g_0, g_1 are increasing by Lemma 6.7 and positive, the expressions in (ii) – (v) are all decreasing;

$$\frac{f_3(x)}{x^2f_1(x)} = \frac{1}{g_0(x)g_1(x)}, \quad \frac{f_2(x)}{xf_1(x)} = \frac{1}{g_1(x)}, \quad \frac{f_3(x)}{x^4f_1(x)} = \frac{1}{x^2g_0(x)g_1(x)}, \quad \frac{f_3(x)}{x^2f_2(x)} = \frac{1}{xg_0(x)}$$

The upper bounds are obtained by noting that $g_i(0) = 3 - i$ by Lemma 6.6, while the lower bounds are obtained numerically by letting $x = 2.688 > \lambda$. $\square$

Bibliography

[1] L.A. Adamic and B.A. Huberman. Power-law distribution of the World Wide Web. *Science*, 287:2115, 2000.

[2] R. Albert and A-L. Barabási. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[3] D. Aldous. Asymptotics in the random assignment problem. *Probability Theory and Related Fields*, 93:507–534, 1992.

[4] D. Aldous. The $\zeta(2)$ limit in the random assignment problem. *Random Structures and Algorithms*, 4:381–418, 2001.

[5] N. Alon and J. Spencer. *The probabilistic Method*. John Wiley and Sons Inc., New Jersey, third edition, 2008.

[6] F. Avram and D. Bertsimas. The minimum spanning tree constant in geometrical probability and under the independent model: a unified approach. *Annals of Applied Probability*, 2:113–130, 1992.

[7] A.L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A*, 311:590–614, 2002.

[8] J. Beardwood, J.H. Halton, and J.M. Hammersley. The shortest path through many points. *Mathematical Proceedings of the Cambridge Philosophical Society*, 55:299–327, 1959.

[9] A. Beveridge, A.M. Frieze, and C. McDiarmid. Random minimum length spanning trees in regular graphs. *Combinatorica*, 18:311–333, 1998.

[10] T. Bohman and A.M. Frieze. Hamilton cycles in 3-out. *Random Structures and Algorithms*, 35:393–417, 2010.

[11] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled graphs. *European Journal on Combinatorics*, 1:311–316, 1980.

[12] B. Bollobás. *Random Graphs*. Academic Press, London, 1985.

[13] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Math.*, 1(1):1–35, 2003.

[14] B. Bollobás, O. Riordan, J. Spence, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random structures and algorithms*, 18:279–290, 2001.

[15] B. Bosek, D. Leniowski, P. Sankowski, and A. Zych. Online bipartite matching in offline time. In *55th Annual IEEE Symposium on Foundations of Computer Science*, 2014.

[16] J.A. Cain, P. Sanders, and Nick Wormald. The random graph threshold for k-orientability and a fast algorithm for optimal multiple-choice allocation. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 469–476, 2007.

[17] K. Chaudhuri, C. Daskalakis, R. Kleinberg, and H. Lin. Online bipartite matching with augmentations. In *28th IEEE INFOCOM*, pages 1044–1052, 2009.

[18] F. Chung and L. Lu. Coupling online and offline analyses for random power law graphs. *Internet Math.*, 1(4):409–461, 2004.

[19] C. Cooper, A.M. Frieze, N. Ince, S. Janson, and J. Spencer. On the length of a random minimum spanning tree. arXiv:1208.5170.

[20] C. Cooper, A.M. Frieze, and J. Vera. Random deletion in a scale free random graph process. *Internet math.*, 1(4):463–483, 2003.

[21] K.S. Crump and C.J. Mode. A general age-dependent branching process. I. *Journal of mathematical analysis and applications*, 24:494–508, 1968.

[22] K.S. Crump and C.J. Mode. A general age-dependent branching process. II. *Journal of mathematical analysis and applications*, 25:8–17, 1969.

[23] R. Durrett. *Probability: Theory and examples.* Wadsworth and Brooks/Cole, 1991.

[24] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

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

[26] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[27] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta. Math. Acad. Sci. Hungar.*, 8:261–267, 1961.

[28] P. Erdős and A. Rényi. On random matrices. *Publ. Math. Inst. Hungar. Acad. Sci.*, 8:455–461, 1964.

[29] 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.

[30] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. In *ACM SIGCOMM, Boston*, 1999.

[31] T. Fenner and A.M. Frieze. On the connectivity of random m-orientable graphs and digraphs. *Combinatorica*, 2:347–359, 1982.

[32] A.D. Flaxman, A.M. Frieze, and J. Vera. A geometric preferential attachment model of networks. *Internet math.*, 3(2):187–205, 2006.

[33] A.D. Flaxman, A.M. Frieze, and J. Vera. Adversarial deletion in a scale free random graph process. *Combinatorics, Probability and Computing*, 16:261–270, 2007.

[34] A.D. Flaxman, A.M. Frieze, and J. Vera. A geometric preferential attachment model of networks II. *Internet math.*, 4(1):87–112, 2007.

[35] D. Fotakis, R. Pagh, P. Sanders, and P. Spirakis. Space efficient hash tables with worst case constant access time. *Theory of Computing Systems*, 8:229–248, 2005.

[36] N. Fountoulakis and K. Panagiotou. Sharp load thresholds for cuckoo hashing. *Random Structures and Algorithms*, 41:306–333, 2012.

[37] N. Fountoulakis, K. Panagiotou, and A. Steger. On the insertion time of cuckoo hashing. *SIAM Journal on Computing*, 42:2156–2181, 2013.

[38] A. Frieze and T. Johansson. Minimum-cost matching in a random graph with random costs. To appear in *SIAM Journal on Discrete Mathematics.*

[39] A. Frieze and T. Johansson. On edge disjoint spanning trees in a randomly weighted complete graph. Pre-print arXiv:1505.03429.

[40] A. Frieze and T. Johansson. On random k-out sub-graphs of large graphs. To appear in *Random Structures and Algorithms.*

[41] A. Frieze and T. Johansson. On the insertion time of random walk cuckoo hashing. To appear at *ACM-SIAM Symposium on Discrete Algorithms (SODA17).*

[42] A.M. Frieze. On the value of a random minimum spanning tree problem. *Discrete Applied Mathematics*, 10:47–56, 1985.

[43] A.M Frieze. Maximum matchings in a class of random graphs. *Journal of Combinatorial Theory B*, 40:196–212, 1986.

[44] A.M. Frieze. On random symmetric travelling salesman problems. *Mathematics of Operations Research*, 29(4):878–890, 2004.

[45] A.M. Frieze, N. Goyal, L. Rademacher, and S. Vempala. Expanders via random spanning trees. *SIAM Journal of Computing*, 43(2):497–513, 2014.

[46] A.M. Frieze and G.R. Grimmett. The shortest path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10:57–77, 1985.

[47] A.M. Frieze and M. Krivelevich. On the non-planarity of a random subgraph. *Combinatorics, Probability and Computing*, 22:722–732, 2013.

[48] A.M. Frieze and C.J.H. McDiarmid. On random minimum length spanning trees. *Combinatorica*, 9:363–374, 1989.

[49] A.M. Frieze and P. Melsted. Maximum matchings in random bipartite graphs and the space utilitization of cuckoo hashtables. *Random Structures and Algorithms*, 41:334–364, 2012.

[50] A.M. Frieze, P. Melsted, and M. Mitzenmacher. An analysis of random walk cuckoo hashing. *SIAM Journal on Computing*, 40:291–308, 2011.

[51] A.M. Frieze, M. Ruszinko, and L. Thoma. A note on random minimum length spanning trees. *The electronic journal of combinatorics*, 7, 2000.

[52] A.M. Frieze, M. Ruszinko, and L. Thoma. A note on random minimum length spanning trees. *Electronic Journal of Combinatorics*, 7:R41, 2000.

[53] P. Gao, X. Pérez-Giménez, and C.M. Sato. Arboricity and spanning-tree packing in random graphs with an application to load balancing. arXiv:1303.3881, extended abstract published in Proc. SODA 2014, Portland, Oregon, 317–326 (2014).

[54] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30(4):1141–1144, 1959.

[55] G. Grimmett and D. Stirzaker. *Probability and random processes.* Oxford University Press, Oxford, third edition, 2001.

[56] A. Gupta, A. Kumar, and C. Stein. Maintaining assignments online: matching, scheduling and flows. In *Proceedings of the 25th annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014.

[57] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.

[58] S. Janson. The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. *Random Structures and Algorithms*, 7:337–355, 1995.

[59] S. Janson. One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8:347–361, 1999.

[60] T. Johansson. Deletion of oldest edges in a preferential attachment graph. Pre-print arXiv:1610.04588.

[61] R.M. Karp. A patching algorithm for the non-symmetric traveling salesman problem. *SIAM Journal on Computing*, 8:561–573, 1979.

[62] R.M. Karp. *An upper bound on the expected cost of an optimal assignment.* Academic Press, New York, 1987.

[63] B.J. Kim and S.M. Park. Distribution of Korean family names. *Physica A*, 347:683–694, 2005.

[64] W. Kordecki and A. Lyczkowska-Hanćkowiak. Exact expectation and variance of minimal basis of random matroids. *Discussiones Mathematicae Graph Theory*, 33:277–288, 2013.

[65] M. Krivelevich, C. Lee, and B. Sudakov. Robust Hamiltonicity of Dirac graphs. *Transactions of the American Mathematical Society*, 366:3095–3130, 2014.

[66] M. Krivelevich, C. Lee, and B. Sudakov. Long paths and cycles in random subgraphs of graphs with large minimum degree. *Random Structures and Algorithms*, 46(2):320–345, 2015.

[67] M. Krivelevich and W. Samotij. Optimal packings of Hamilton cycles in sparse random graphs. *SIAM Journal on Discrete Mathematics*, 26:964–982, 2012.

[68] S. Linusson and J. Wästlund. A proof of Parisi's conjecture on the random assignment problem. *Probability Theory and Related Fields*, 128:419–440, 2004.

[69] T. Łuczak. Size and connectivity of the k-core of a random graph. *Discrete Mathematics*, 91:61–68, 1991.

[70] M. Mitzenmacher. Some open questions related to cuckoo hashing. In *Proceedings of the 17th Annual Europeans Symposium on Algorithms (ESA)*, pages 1–10, 2009.

[71] M. Mitzenmacher and S. Vadhan. Why simple hash functions work: exploiting the entropy in a data stream. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 746–755, 2008.

[72] S. Miyazima, Y. Lee, T. Nagamine, and H. Miyajima. Power-law distribution of family names in Japanese societies. *Physica A*, 278:282–288, 2000.

[73] S.A. Myers, A. Sharma, P. Gupta, and J. Lin. Information network or social network? the structure of the Twitter follow graph. In *WWW*, pages 493–498, 2014.

[74] C. Nair, B. Prabhakar, and M. Sharma. Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures. *Random Structures and Algorithms*, 27:413–444, 2005.

[75] C. St. J. A. Nash-Williams. Edge-disjoint spanning trees of finite graphs. *Journal of the London Mathematical Society*, 36:445–450, 1961.

[76] J. Oxley. *Matroid Theory*. Oxford University Press, Oxford, 1992.

[77] R. Pagh and F. Rodler. Cuckoo hashing. *Journal of Algorithms*, 51:122–144, 2004.

[78] G. Parisi. A conjecture on random bipartite matching. Physics e-Print archive, 1998.

[79] M. Penrose. Random minimum spanning tree and percolation on the n-cube. *Random Structures and Algorithms*, 12:63–82, 1998.

[80] B. Pittel, J. Spencer, and N. Wormald. Sudden emergence of a giant k-core in a random graph. *Journal of Combinatorial Theory (B)*, 67:111–151, 1996.

[81] O. Riordan. Long cycles in random subgraphs of graphs with large minimum degree. *Random Structures and Algorithms*, 45(4):764–767, 2014.

[82] E. Shamir and E. Upfal. One-factor in random graphs based on vertex choice. *Discrete Mathematics*, 41:281–286, 1982.

[83] J.M. Steele. On Frieze's $\zeta(3)$ limit for lengths of minimal spanning trees. *Discrete Applied Mathematics*, 18:99–103, 1987.

[84] J.M. Steele. *Probability theory and combinatorial optimization*. SIAM CBMS series, 1996.

[85] T. Szele. Combinatorial investigations concerning complete directed graphs. *Mat. es Fiz. Lapok*, 50:223–236, 1943. (in Hungarian).

[86] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathmatiques de L'IHÉS*, 81:73–205, 1995.

[87] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. The anatomy of the Facebook social graph. arXiv:1111.4503.

[88] D.W. Walkup. On the expected value of a random assignment problem. *SIAM Journal on Computing*, 8:440–442, 1979.

[89] D.W. Walkup. Matchings in random regular bipartite graphs. *Discrete Mathematics*, 31:59–64, 1980.

[90] J. Wästlund. A simple proof of the Parisi and Coppersmith-Sorkin formulas for the random assignment problem. *Linköping Studies in Mathematics*, 6, 2005.

[91] J. Wästlund. Random matching problems on the complete graph. *Electronic Communications in Probability*, pages 258–265, 2008.

[92] J. Wästlund. An easy proof of the $\zeta(2)$ limit in the random assignment problem. *Electronic Communications in Probability*, 14:261–269, 2009.

[93] J. Wästlund. The mean field traveling salesman and related problems. *Acta Mathematica*, 204:91–150, 2010.

[94] D.J.A. Welsh. *Matroid Theory*. Academic Press, London, 1976.

[95] R.M. Young. Euler's constant. *Mathematical Gazette*, 75:187–190, 1991.